Rationality, competition and evolution : entry (deterrence) in dynamic barrier market theory

Citation for published version (APA):

van Witteloostuijn, A. (1990). *Rationality, competition and evolution : entry (deterrence) in dynamic barrier market theory*. [Doctoral Thesis, Maastricht University]. Faculty of Economics and Business Administration, University of Limburg. <https://doi.org/10.26481/dis.19900913aw>

Document status and date:

Published: 01/01/1990

DOI:

[10.26481/dis.19900913aw](https://doi.org/10.26481/dis.19900913aw)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

RATIONALITY, COMPETITION AND EVOLUTION

Entry (Deterrence) in Dynamic Barrier Market Theory

RATIONALITY, COMPETITION AND EVOLUTION

Entry (Deterrence) in Dynamic Barrier Market Theory

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de Rijksuniversiteit Limburg te Maastricht, op gezag van de Rector Magnificus, Prof. Dr. F.I.M. Bonke, volgens het besluit van het College van Dekanen, in het openbaar te verdedigen op donderdag, 13 september 1990 om 14.00 uur

door

Adriaan van Witteloostuijn
geboren te Haarlem in 1960

Promotores:

Prof. Dr. J.A.H. Maks
Prof. Dr. S.K. Kuipers

Beoordelingscommissie:

Prof. Dr. J. Muysken (voorzitter)
Prof. Dr. M.J. Holler
Prof. Dr. H. Schreuder

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Witteloostuijn, Adriaan van

Rationality, competition and evolution: entry
(deterrence) in dynamic barrier market theory / Adriaan
van Witteloostuijn. - Maastricht : Datawyse. - Ill., fig., tab.

Proefschrift Maastricht. - Met lit. opg. - Met
samenvatting in het Nederlands.

ISBN 90-5291-028-6

SISO 340 UDC 330(043.3)

Trefw.: industriële economie / economische psychologie
/ micro-economie.

© 1990 A. van Witteloostuijn, Banholt

Layout: Aad van Mourik

Druk: Datawyse Maastricht / Krips Repro Meppel

Behoudens uitzonderingen door de Wet gesteld, mag zonder schriftelijke toestemming van de rechthebbende op de auteursrechten c.q. de rechthebbende gemachtigd namens deze op te treden, niets uit deze uitgave worden vermenigvuldigd en/of openbaar gemaakt door middel van druk, fotokopie, micro-film of anderszins. No part of this book may be reproduced in any form, by print, photoprint, micro-film or any other means without a written permission from the publisher.

A dissertation is to be written, not to be read

Dedicated to my parents and Hettie

FOREWORD

This book collects a nonrandom sample of the words I wrote in the past four years. This has much (not to say everything) to do with the status of this book as a dissertation. It is my "maiden speech" in the economic "parliament". In framing a selection procedure so as to construct a dissertation in social sciences one may follow two (common) extremes. First, one may opt for a *restrictive*, predominant theme. Each statement between the beginning and the end of the book fits in. Asides are avoided and all topics which show no close relation to the running thread are skipped. Concentrated focus prevails to the detriment of broadness of scope. Second, one may present a statement which covers *all* the work carried out during the (dissertation) years involved. Nothing is left out. Concentrated focus is sacrificed for the sake of completeness. It is, however, likely that the clarity of the exposition and the cohesion of the argument are reduced.

Although the first approach enjoys increasing popularity in recent (post-war) decades (it is supposed to be more scientific), the selection procedure employed here is closer to the second extreme. So, for example, extensive review material and a lengthy list of references are included. This is done, because I think that the nonspecialized reader often considers surveys to be the most interesting parts of a thesis. In line with the purport of the book's content, though, an attempt is made to trace an (expected) optimum in the tradeoff between two economies involved: concentrated focus and broadness of scope. Of course, my perception of the (proofs of the) printed result illustrates that decisions *ex ante* are more often than not associated with disappointed expectations *ex post*.

The following should be noted in advance with respect to the *general* approach employed in this book. Here it suffices to list a few well-known adjectives and nouns. First, the thesis mainly deals with *theory* rather than empirical research or practice. Second, the major arguments are pursued along *deductive* lines. Third, the way in which the material is expressed is eclectic in the use of *techniques*. Literary reasoning, graphical illustrations, mathematics and simulation experiments are all used in varying mixtures in each chapter.

Fourth, the economics in this book is mainstream. It describes essentially a *neoclassical, maximizing, equilibrium* framework. Here I do not deviate from a century-old tradition (which is described so excellently by J.A. Schumpeter in his *The History of Economic Analysis*, 1954). If the reader perceives an ignorance of a tradeoff between economic efficiency (the question of allocation) and social justice (the question of distribution), then (s)he must be aware of the fact that this book *only* deals with the former aspect. Abstraction is an inevitable and necessary tool of (scientific) economists. Of course, "[i]t is practically impossible to achieve the elimination of extra-scientific values from scientific activity. The situation is similar with respect to objectivity: we cannot rob the scientist of his partisanship without also robbing him of his humanity. ... the purity of science is an idea which is presumably unattainable It is an ideal for which we constantly fight - and should fight - by means of criticism" (K. Popper in: G. Adey and D. Frisby - *The Positivist Dispute in German Sociology*, 1976, p. 97). Hereby criticism is invited.

ACKNOWLEDGEMENTS

This book would, of course, be nonexistent without the help of many. Here only a few are mentioned. First of all, the knowledge account of the balance of payments with Professor W. J. Baumol (Princeton and New York Universities) shows a clear deficit on my part. With this book I hope to induce a first adjustment towards equilibrium. Second, I wish to thank the co-authors of the papers which underlie this thesis: H.B.A. Bierings (Centraal Bureau voor Statistiek), H. van Ees (University of Groningen), J. de Haan (University of Groningen), L.H. Hoogduin (De Nederlandsche Bank), Professor J.A.H. Maks (University of Limburg) and M. van Wegberg (University of Limburg). Third, I have been able to improve the argument considerably on the basis of the comments of Professor J.T. Addison (University of South Carolina), P. van Cayseele (University of Leuven), A. de Grip (University of Limburg), G. Hendrikse (University of Brabant), Professor M.J. Holler (University of Aarhus), P. Kooreman (University of Brabant), Professor J. Muysken (University of Limburg), Professor J. Pen (University of Groningen), Professor J.H.R. van de Poel (University of Limburg), E. de Regt (University of Limburg), Professor H. Schreuder (University of Limburg), M. Vendrik (University of Limburg) and the inevitable anonymous referees. Of course, the usual disclaimer applies. The editorial comments of Bob Wilkinson have contributed significantly to the readability of this thesis. Ad Klancnik and Aad van Mourik polished up the final layout. Last but not least, I am very grateful for the (emotional) support of Hettie and my parents.

Arjen van Witteloostuijn
Maastricht, March 1990

CONTENTS

FOREWORD	i
----------	---

ACKNOWLEDGEMENTS	ii
------------------	----

LIST OF TABLES AND FIGURES	xix
----------------------------	-----

LIST OF SYMBOLS	xxi
-----------------	-----

Part I	INTRODUCTION: History, Scope, Results and Guide	1
--------	---	---

1.	HISTORY AND SCOPE	3
----	-------------------	---

2.	RESULTS AND GUIDE	7
----	-------------------	---

Part II	ECONOMIC PSYCHOLOGY: Rationality and Decision Rules	11
---------	---	----

3.	RATIONALITY IN ECONOMICS: A Confusing Debate	13
----	--	----

3.1	Rationality and Division of Opinion	13
-----	-------------------------------------	----

3.2	Rationality and Perfect Foresight	13
-----	-----------------------------------	----

3.3	Rationality and Selfishness	16
-----	-----------------------------	----

3.4	Rationality and Empirical Evidence	17
-----	------------------------------------	----

3.5	Rationality and Tautology	19
-----	---------------------------	----

3.6	Rationality and Determinism	20
-----	-----------------------------	----

3.7	Conclusion	22
-----	------------	----

4.	DECISION RULES: Maximizing or Satisficing?	23
----	--	----

4.1	Introduction: Caricatures	23
-----	---------------------------	----

4.2	Behaviorism: Critique of the Maximizing Postulate	24
-----	---	----

4.3	The Maximizing Postulate: Goals, Time Variability and Groups	26
-----	--	----

4.4	The Maximizing Postulate: Uncertainty and Routines	28
-----	--	----

4.4.1	Introduction	28
-------	--------------	----

4.4.2	A Taxonomy of Uncertainty Concepts	29
-------	------------------------------------	----

4.4.3	Choice Theory and Perfect Foresight	31
-------	-------------------------------------	----

4.4.4	Choice Theory and Complete Knowledge	32
-------	--------------------------------------	----

4.4.5	Choice Theory and Routines and Computational Disabilities	33
-------	---	----

4.5	The Flexibility of the Maximizing Framework	34
-----	---	----

4.6	The "As If" and Natural Selection Arguments	35
-----	---	----

4.7	Asking the Right Question	37
-----	---------------------------	----

Part III	INDUSTRIAL ORGANIZATION: Static and Dynamic Competition	41
5.	THE DILEMMA: The Tradeoff between Static and Dynamic Efficiency	43
6.	STATIC EFFICIENCY: Price Competition	47
6.1	Introduction: (Minimum) Average Cost Pricing	47
6.2	Perfect and Bertrand Competition: Actual Rivalry	48
6.3	Potential Competition	51
6.4	(Perfect) Contestability: Potential Rivalry	52
6.4.1	Introduction: Intuition	52
6.4.2	Essential Assumptions	53
6.4.3	Crucial Definitions	56
6.4.4	Implications for Market Performance and Market Structure	59
6.5	Critique of Static Benchmark Cases of Competition	64
6.5.1	Two Lines of Critique	64
6.5.2	Restrictive Assumptions	64
6.5.3	Static Theory	69
6.6	Sunk Investment and Identification of Potential Entrants	71
7.	DYNAMIC EFFICIENCY: Entry Barriers and Sunk Investment	75
7.1	Actual and Potential Competition	75
7.2	Bain's Predecessors and Contemporaries	76
7.3	Bain	76
7.4	Definitions, Categories and Sources of Entry Barriers	78
7.5	Sylos' Postulate and Excess Capacity Hypothesis	80
7.6	Strategic Interactions	82
7.7	Market Performance	85
7.8	Entry Barriers, Imperfect Competition and the Source of Potential entry	87
8.	FIRMS' R&D OUTLAYS: Process and Product Innovations	91
8.1	Introduction: A Crucial Dynamic Economy and Entry Barrier	91
8.2	A Hicksian and a Schumpeterian Perspective	93
8.3	Hicksian Models of Induced Invention	96
8.4	Schumpeterian Equilibrium Models of R&D Decision Making	98
8.5	Schumpeterian Disequilibrium Models of Evolutionary Selection	103
8.6	Innovation and Market Performance	105
8.7	Conclusion: Complexity and Creativity	107
Part IV	WORKABLE COMPETITION AND THE BARRIER MARKET: A Benchmark Case of Intermediate Competition	109
9.	STATIC AND DYNAMIC EFFICIENCY: Workable Competition and the Threat of Entry	111
9.1	Introduction: A Way Out of the Efficiency Dilemma	111

Contents

9.2	Workable Competition	113
9.3	Recent Related Literature	116
9.4	Workably Competitive Scenarios	120
9.5	Essentials, Scenarios and Amortization	121
10.	THE ESSENTIALS: The Barrier Market as an Intermediate Case of Competition	125
10.1	Definition	125
10.2	Assumptions and Conditions of Entry and Exit	126
10.3	A Benchmark Model of the Sequence of Events, Rules of the Game and Equilibrium Concept	128
10.4	Implications for Market Behavior	131
10.5	A Benchmark Case of Competition	133
Part V	BARRIER MARKET SCENARIOS: Sources of Potential Entry and Profit Incentives	135
11.	SOURCES OF POTENTIAL ENTRY: Investment Disincentives, Excess Productive Capacity and (Un)protected Home Market	137
11.1	Introduction: Sources of Potential Entry	137
11.2	Investment Disincentives and Asymmetry of Sunk Cost	139
11.2.1	Two Necessary Assumptions for Perfect Barrier Market Outcomes	139
11.2.2	Dropping Assumption (A): Potential Entrants Sink Cost upon Entry	140
11.2.3	Dropping Assumption (B): Potential Entrants Can Leapfrog after a Positive Entry Lag	141
11.2.3.1	Leapfrogging	141
11.2.3.2	Incumbent Firms' Payoff	142
11.2.3.3	Potential Entrants' Payoff	144
11.2.3.4	Blockaded Entry and Imperfect Barrier Market	145
11.3	Excess Productive Capacity and Dumping	147
11.3.1	Entry with Excess Productive Capacity	147
11.3.2	Sustainability of Excess Capacity	148
11.3.3	Reciprocal Dumping	148
11.4	(Un)protected Home Market and Lag Structure	149
11.4.1	Home Market Disequilibrium	151
11.4.2	Protected Home Market	151
11.4.3	Unprotected Home Market	152
11.5	The Credibility of the Entry Threat	154
12.	PROFIT INCENTIVES: Escalating Commitment and Strategic Entry-Deterring Investment	159
12.1	Profit, Sunk Investment and the Opportunity Cost of Entry	159
12.1.1	Profit in a Barrier Market	159
12.1.2	It Can Pay to Be Superior	160
12.1.3	Zero Profit Investment	162
12.1.4	Zero or Normal Profit	164

Contents

12.1.5	Perfect and Imperfect Barrier Market	166
12.2	Nonstrategic and Strategic Investment	167
12.3	Types of Barrier Markets	168
Part VI	THE BARRIER MARKET AND SUSTAINABILITY: Unique Cost Minimizing Production and Amortization Procedures	171
13.	SUNK COST, UNIQUE COST MINIMIZING PRODUCTION AND SUSTAINABILITY	173
13.1	Introduction: Temporary and Intertemporal Unsustainability	173
13.2	Sunk Costs, Average Cost Functions and Sustainability	173
13.3	Imperfect Barrier Markets, Entry Barriers and Sustainability	176
13.4	Uncertainty, Unsustainability and Adjustment	176
13.5	Conclusion	179
14.	SUNK COST, AMORTIZATION AND SUSTAINABILITY	181
14.1	Introduction: Amortization and Unsustainability	181
14.2	Intertemporal Sustainability and Amortization Rules	182
14.2.1	Amortization Rules	182
14.2.2	Conventional Rules	183
14.2.3	Zero-Profit Amortization	185
14.2.4	Intertemporal Unsustainability	187
14.3	Unique Entry-Deterring Amortization Rule	188
14.4	Conventions, Routines and Focal Points	190
14.5	Conclusion	192
Part VII	COMMODITY MARKET: Cost Reduction and Product Innovation	193
15.	PROCESS INNOVATION: Minimization of Price	195
15.1	Introduction: Cost Reductions	195
15.2	Process Innovation: The R&D Technology	195
15.3	Minimization of the Future Price	197
15.4	Optimal R&D Outlays	199
15.5	Conclusion	201
16.	PRODUCT INNOVATION: Maximization of Buyers' Utility	203
16.1	Introduction: Increased Quality	203
16.2	Maximization of Buyers' Utility	203
16.2.1	Two-Stage Game with Product Innovation	203
16.2.2	Demand for Quality	204
16.2.3	Product Innovation	205
16.3	Product Innovation and (Un)sustainability	206
16.3.1	Efficient Scale	206
16.3.2	Demand Change	207
16.3.3	Exit	208

16.3.4	Entry	208
16.3.5	Increased Probability of Unsustainability	209
16.4	Conclusion	209
Part VIII	LABOR MARKET: Static-Dynamic Efficiency Tradeoff, Wage Rate and Human Capital Improvement	211
17.	CONTESTABLE MARKETS, ENTRY BARRIERS AND LABOR SUPPLY	213
17.1	Introduction: Industrial Organization and Labor Supply	213
17.2	Wage Rate and Human Capital Investment	214
17.2.1	The Economics of Human Capital	214
17.2.2	An Illustrative Two-Period Model	215
17.3	Contestability and Minimum Wage Rates	218
17.4	Barriers to Entry in the Labor Market	221
17.5	Imperfect Competition, Entry Barriers and Positive Returns	222
17.6	Conclusion: The Tradeoff	225
18.	HUMAN CAPITALIMPROVEMENTS: Maximization of Attractiveness	227
18.1	Introduction: Barrier Labor Markets	227
18.2	Zero Returns and Maximum Attractiveness	228
18.2.1	Barrier Labor Market	228
18.2.2	Zero Returns	228
18.2.3	Maximum Attractiveness	230
18.3	A Model	231
18.4	Conclusion	234
Part IX	MULTIMARKET COMPETITION: Microfoundation of Macroeconomics, Industrial Organization and Strategic Management	235
19.	MULTIMARKET COMPETITION: Vertical and Horizontal Spillovers and Strategies	237
19.1	Introduction: Set of Related Markets	237
19.2	Framework of Multimarket Competition	237
19.3	Focus of Rivalry	239
19.4	Ease or Difficulty of Entry	240
19.5	Intermarket Spillover Effects	240
20.	MICROFOUNDATION OF MACROECONOMICS: Rationing, Entry and Exit	245
20.1	Introduction: General Equilibrium Theory and Dynamic Competition	245
20.2	Microfoundation of Macroeconomics	246
20.2.1	Temporary and Intertemporal Equilibrium	246

Contents

20.2.2	Walrasian Equilibrium Models	247
20.2.3	Non-Walrasian Equilibrium Models	248
20.2.4	Dynamic Competition	251
20.3	Association Technology and Human Capital	253
20.4	Homogeneous Firms and Laborers	254
20.5	Heterogeneity of Firms and Price Equalization	259
20.6	Heterogeneity of Laborers, Rigid Wage Rates and Regimes	262
20.7	Tendency to Increased Wage Rate Dispersion	264
20.8	Conclusion	266
21.	INDUSTRIAL ORGANIZATION AND STRATEGIC MANAGEMENT: Entry(-Deterring) Strategies and Types of Competition	269
21.1	Introduction: Strategic Management's Perspective	269
21.2	Credibility of the Entry Threat	270
21.2.1	Entry Conditions	270
21.2.2	Internal Competition	272
21.2.3	Pool of Potential Entrants	272
21.2.4	Retaliation Lag	273
21.2.5	Simultaneous Entry Threat	273
21.2.6	Scale of Entry	274
21.2.7	Profit Motivation	274
21.3	Sources of Potential Entry and (Price) Strategies	276
21.3.1	Multimarket Competition and Strategy	276
21.3.2	(Non)contestability	276
21.3.3	Leapfrogging	277
21.3.4	Dumping	278
21.3.5	Hit-and-Run	280
21.3.6	Hit-and-Stay	281
21.4	An Example: (Excess) Productive Capacity	282
21.4.1	Bertrand Competition, Capacity Constraints Identified Entry and Multimarket Modeling	282
21.4.2	Demand, Capacity Choice, Price Setting and Entry Deterrence	284
21.4.3	Scenarios of Market Interactions	288
21.5	Types of Competition	291
Part X.	EVOLUTION: Competitive Dynamics and Intertemporal Welfare	295
22.	CHANGES OVER TIME: Dynamics and Welfare	297
22.1	Evolution	297
22.2	Competitive Dynamics	297
22.3	Intertemporal Frameworks	298
22.4	Simulation Techniques	304
22.5	A Complication: Intertemporal Welfare	306
23.	COMPETITIVE DYNAMICS IN A BARRIER MARKET:	

Contents

	Diverging Expectations, Entry and Exit Movements and Barriers	307
23.1	Introduction: A Simulation Experiment	307
23.2	Price, Quantity and Investment	308
23.3	Increasing and Decreasing Demand in a Contestable Market	310
23.4	A Contestable Market with Costless Capacity Expansion and Reduction	313
23.5	A Barrier Market with Costly Process Innovation	317
23.6	Competitive Dynamics and Market Performance	321
24.	WELFARE IMPLICATIONS AND POLICY INTERVENTIONS	323
24.1	Introduction: Welfare Evaluations in Economics	323
24.2	Intervention and Information	324
24.3	Welfare Criteria	327
24.3.1	Pre-Paretian and Paretian Welfare Norms	327
24.3.2	Compensation Principles, Economic Surplus Conditions, Second-Best Rules and Superfairness Criteria	329
24.4	Welfare Features of the Barrier Market: First Remarks	333
24.5	Future Research	335
25.	MULTIMARKET COMPETITIVE DYNAMICS	337
AFTERWORD		339
APPENDICES		343
A.	Minimum Average Total Cost	343
B.	Decreasing Returns R&D Technology	343
C.	Existence Optimal R&D Outlays	344
D.	Uniqueness Optimal R&D Outlays	344
E.	Existence and Uniqueness Optimal Human Capital Investment	345
REFERENCES		347
SAMENVATTING		389
CURRICULUM VITAE		395

LIST OF TABLES AND FIGURES

Figure 2.1	Readers' Guide	396
Figure 4.1	Focus of Uncertainty	30
Figure 6.1	Perfect Competition	48
Figure 6.2	Extensive Form Tree of the Game in a Contestable Market	61
Figure 6.3	(Minimum) Average Cost Pricing	62
Figure 6.4	Flat Bottom and Sustainability ($k > 1$)	63
Table 9.1	Assumptions Underlying Workably Competitive Results	122
Figure 10.1	Structure of the Barrier Market Game	130
Table 10.1	Assumptions of Contestable, Barrier and Profit Markets	133
Table 10.2	Implications of Contestable, Barrier and Profit Markets	134
Table 11.1	Payoff Matrix Perfect Barrier Market with Dumping	150
Table 12.1	Payoff Matrix Perfect Barrier Market with Exit Cost	163
Table 12.2	Payoff Matrix Perfect Barrier Market with Investment Alternative	164
Table 12.3	Payoff Matrix Perfect Barrier Market with Normal Profits	165
Figure 13.1	Average Total Cost Functions	175
Figure 13.2	Flat-Bottomed Average Total Cost Curve	177
Figure 14.1	The Game on Convention	185
Figure 15.1	The Game on Process Innovation	199
Figure 15.2	Optimal Innovative Activity	200
Figure 17.1	Game in a Contestable Labor Market	220
Figure 17.2	Game in a Labor Market with Entry Barriers	224
Figure 18.1	Second-Stage Game on Wage Rates in a Barrier Labor Market	229
Figure 18.2	Optimal Investment in Human Capital	233
Figure 19.1	Multimarket Dimensions	238
Figure 19.2	Games of Multimarket Rivalry	239
Figure 19.3	Intermarket Spillover Effects	242
Figure 20.1	Malinvaud's (1977) Regimes	250
Figure 20.2	Regimes in the Labor Market with Homogeneity	255
Figure 20.3	Regimes in the Commodity Market with Homogeneity	257
Figure 20.4	Regimes in the Commodity Market with Heterogeneity	261
Figure 20.5	Heterogeneity of Laborers and Regimes	262
Figure 20.6	Regimes in the Labor Market with Heterogeneity	263
Figure 21.1	Credibility of Entry Threat	271
Figure 21.2	Sources of Potential Entry and (Price) Strategies	275
Table 23.1	Increasing Demand and Entry in a Contestable Market	311
Table 23.2	Increasing Demand and Prices in a Contestable Market	311
Table 23.3	Decreasing Demand and Exit in a Contestable Market	312
Table 23.4	Increasing Demand and Sustainability of a Contestable Monopoly	314
Table 23.5	Increasing Demand and Unsustainability of a Contestable Market	315
Table 23.6	Exited Firms	316
Table 23.7	Process Innovation, Prices and Entry	318
Table 23.8	Innovative Opportunities	319
Table 23.9	Decisions and Expectations of Firm 3 and 8	319
Table 23.10	Results in Period $t = 4$	320
Table 23.11	Results in Period $t = 5$	321

LIST OF SYMBOLS

α	parameter/amortization procedure/labor elasticity of production
a	productivity
A	capital-augmenting coefficient/average (variable or nonsunk) cost function/adjustment lag/market index (superscript)
ac	average cost (function)
aec	actual exit cost
aic	average incremental sunk cost
asc	average sunk cost/actual sunk cost
avc	average variable or nonsunk cost
atc	average total cost
β	parameter/price of selling a unit of capital/material elasticity of production
b	new capital index (superscript)/budget constraint
B	labor-augmenting coefficient/starting date of investment/market index (superscript)
BE	blockaded entry index (superscript)
bm	price-cost margin
BR	blockaded re-entry index (superscript)
c	cost
c_K	capital cost component
c_L	labor cost component
C	(average total) cost function/commodity market index (superscript) /constrained (superscript)
δ	discount factor
d	quantity demanded from a firm
D	market or aggregate demand (function)
ϵ	parameter/state of technology/growth rate demand
e	(potential) entrant index (superscript)
E	entry date/exit payoff/technology change function/expectational operator
ec	exit cost/entry cost
ED	entry-detering index (superscript)/excess demand
EE	enduring exit index (superscript)
EP	entry period index (superscript)
f	probability that the rival decides not to invest/supplier index (superscript)/probability of entry/number of actual entrants
F	probability/fixed cost/marginal product change function/financial restriction index (superscript)
f^1	probability that at least one potential entrant sinks a cost
f^2	probability that at least one potential entrant is able to leapfrog over the incumbent firm's innovation
f^3	probability that the potential entrant is able to introduce a better technology than the incumbent
g	growth rate/supplier index (superscript)
G	probability/gross index (superscript)/time index (subscript)
h	supplier index (superscript)/R&D parameter
H	home market index (superscript)/time index (subscript)
hc	stock of human capital
HC	human capital stock change function
i	incumbent index (superscript)/discount rate/supplier index

	(superscript)/incumbent's market index (subscript)
I	inferior firm index (subscript)
ic	(alternative) investment cost
IK	added productive capacity
j	supplier index (superscript)/market index (subscript)
k	ratio parameter/productive capacity
K	capital (scale)/productive capacity change function
L	labor (supply)/investment lag/Lagrange function/labor market index (superscript)/number of working hours
LR	labor efficiency function
μ	R&D parameter
m	research line/market/number of commodities
ma	result of imitation efforts
mac	minimum average (nonsunk) cost
matc	minimum average total cost
mc	marginal cost
mm	agent's welfare
mp	value of marginal product
MP	marginal product function
mr	routine expenditures per unit of capital if imitation is undertaken
mu	markup
M	monopoly profit/markup function/material
F_n	total number of firms
L_n	total number of workers
n	research line
N	normal profit/net index (superscript)
na	results of innovation efforts
nc	number of community members
ne	number of potential entrants
NE	net entry index (superscript)
ni	number of incumbent firms
nr	routine expenditures per unit of capital if innovation is undertaken
NT	net total index (superscript)
π	payoff or profit
M_p	average cost price
MC_p	(minimum) average total cost price
N_p	minimum average cost price
$\underline{p}$	price/profit flow from innovation
$\bar{p}$	margin above minimum average total cost
P	price function
p^E	market equilibrium price without entry (threat)
PE	pre-entry index (superscript)
pec	pure exit cost
p^m	unit price research line m
p^n	unit price research line n
pwg	marginal product - wage rate gap
L_q	left hand side of flat bottom average nonsunk cost function

List of Symbols

m_q	average cost supply
n_q	minimum average cost supply
R_q	right hand side of flat bottom average nonsunk cost function
q	firm's supply
Q	market supply
r	reaction function/interest rate/profit flow from noninnovation/ rate of entry barriers
R	response date/average nonsunk cost reduction function/regime
RD	residual demand/re-entry-deterring index (superscript)
σ	rate of physical deterioration capital
s	game stage
S	superior firm index (subscript)/last entrant index (superscript)/ aggregate supply/scale average cost reduction function
sb	strategic barrier investment
sc	sunk cost
SK	scrapped productive capacity
τ	parameter/price of buying a unit of capital
t	time index (subscript)
T	end of period under consideration/date when innovation is outdated
u	introduction date of innovation/utility
U	utility function/unconstrained (superscript)
ui	utility of unemployment
uic	up-front investment cost
v	imitation date/product quality
V	R&D payoff/product quality change function
max_w	reservation wage rate of firms
min_w	reservation wage rate of workers
R_w	zero-returns wage rate
w	date when innovation is outdated/wage rate
W	expected flow of earnings/payoff from alternative investment/wage rate change function
wc	community's welfare
WC	community's welfare function
w_M	price per unit of material
l_{w_R}	zero-returns wage rate of superior workers
x	introduction date of new product/components welfare function
X	payoff from one-sided dumping
Y	payoff from one-sided nondumping
z	advance production parameter
Z	payoff from reciprocal dumping
$*$	optimal or conventional index (superscript)

PART

I

INTRODUCTION

History, Scope, Results and Guide

KEY WORDS. *Part I introduces the thesis' argument by relating the key concepts (rationality, competition and dynamics) to the history of economic thought, outlining the scope (Chapter 1), providing the intuition of the results and offering a readers' guide (Chapter 2). So, Part I marks the contours of the content of this book.*

This book deals with three key concepts in economic science: rationality, competition and dynamics. These notions are as old as economics itself. Joseph Schumpeter (1954) traces economic thinking back to ancient times. Long before Adam Smith published his *The Wealth of Nations* (1776) philosophers, clergymen, businessmen, public servants and professors went into describing, explaining and prescribing economic events. Of course, they often failed to move beyond common sense knowledge. However, traces of analytic economics can be found in the works of Greek philosophers, Scholastic monks and Mercantilist businessmen. Paying attention to such essentials as individual (ir)rationality, competitive forces and dynamic processes appears to be a rule rather than an exception, even in pre-Smithian times. Five exemplary quotations of Schumpeter (1954) may illustrate this point.

(i) Aristotle based his economic analysis squarely upon wants and their satisfactions. Starting from the economy of self-sufficient households, he then introduced division of labor, barter, and, as a means of overcoming the difficulties of direct barter, money (p. 60).

(ii) [W]e must nevertheless credit him [Duns Scotus] with having discovered the condition of competitive equilibrium which came to be known in the nineteenth century as the Law of Cost (p. 93).

(iii) If the idea [Utilitarianism] was of ancient origin and grew so slowly as to defy dating, the slogan itself may be dated more precisely; so far as I know, it occurs first in Hutcheson ..., then in Beccaria ...; after that in Priestley Hume does not have the slogan, but should be included in the series all the same. The word Utilitarianism is Bentham's (p. 132).

(iv) But he [Quesnay] also is the most important of all the founding fathers of the doctrine that will henceforth be referred to as the Maximum Doctrine of Perfect Competition That is to say, he held that maximum satisfaction of wants for all members of society, taken together, will result if, conditions of perfect competition prevailing, everyone be allowed to act freely upon his own self-interest (p. 233).

(v) Some of the 'mercantilist' writers went to surprising, in fact to Keynesian, lengths. There is nothing startling in Sir William Petty's saying that it is better to produce useless things than not to produce at all (p. 350).

The classical period in economics has produced highlighted formulations of theories of rationality, competition and dynamics, associated with the great names of, for example, Smith (1776), Malthus (1798), Say (1803), Ricardo (1817), Von Thünen (1826), Senior

(1836), Mill (1848) and Marx (1863). The pillar of classical economics is Mill's *Principles of Political Economy* (1848). As Schumpeter points out, "Mill's *Principles* was ... the most successful treatise of the period" (Schumpeter 1954, p. 527). It is worthwhile noting that Mill was the theorist who introduced the terms statics and dynamics into economic theory.

The marginalist "revolution" in the 1870s (Jevons 1871, Walras 1874 and Menger 1874) induced a significant change in the economists' way of expression (Hutchison 1978). Mathematics rather than literary reasoning became the standard language. The current pillar of "scientific" economics is general equilibrium theory. In particular Léon Walras (1874) shaped modern economics by introducing maximizing calculus and a system of simultaneous equations so as to frame individual decision making (ruled by a maximizing procedure) and equilibrium (through perfect competition) respectively (Jaffé 1983). Since then the maximizing postulate and (both perfect and imperfect) competition became the frontrunners of economic modeling. However, although Léon Walras recognized (but not modeled) the crucial importance of *dynamic* processes, mathematical constraints induced a predominance of the analytic focus on (comparative) statics.

Of course, the history of economic analysis did not proceed as simply as that. Not only did the marginal revolution not appear like a penny from heaven [consider, for example, Isnard's (1781) equilibrium equations (Schumpeter 1954 and Jaffé 1983) and Dupuit's (1844) and Gossen's (1854) notions of the marginal utility principle], but additional developments are also to be traced. First, the concept of (maximizing) rationality has been modified and elaborated in many ways (Hirshleifer and Riley 1979, Schoemaker 1982, Van Witteloostuijn 1988a and Fishburn 1989). Although the maximizing postulate has been and still is under continual attack, it still dominates economic modeling. In particular, the critique from behaviorist "satisficing" theorists (for instance, Simon 1945, Cyert and March 1963 and Nelson and Winter 1982) has not (yet) met with massive support from economists. Second, by taking up Augustin Cournot's (1838) thread, economists have introduced all kinds of models of imperfect competition, including the well-known contributions of Chamberlin (1933), Robinson (1933) and Bain (1956). Third, general equilibrium theory has been rendered dynamic by taking explicit notice of the role of expectations in a temporary equilibrium framework (Hicks 1939 and Grandmont 1977). In the chapters to follow the (recent) history and current status of economic theorizing with regard to each of the three key concepts are outlined. Of course, since space is restricted, these examinations can be but brief.

In this book essentially three propositions are put forward. First, it is argued that the maximizing postulate still does not face a credible alternative. Although the (use of) the maximizing postulate shows some deficiencies (by taking, in particular, insufficient notice of uncertainty and learning), it makes, as it stands, no sense to throw away the baby with the bath water. Therefore, here the maximizing principle is the point of departure. Second, it is argued that there is an intermediate case of competition between perfect (or contestable) and imperfect (or entry barrier) competition. In doing so, a model is provided in which the tradeoff between static and dynamic efficiency of market performance is weakened. This barrier market framework seeks to provide a benchmark case of competition. The barrier market theory can induce instability of market configurations. Barrier market assumptions are applied to competition for customers (commodity market) and jobs (labor market). Third, on the basis of a review of the literature on general equilibrium theory and economic dynamics it appears that so far insufficient attention has been paid toward modeling intermarket spillovers, multimarket competition and competitive, evolutionary processes over time. Here simulation

experiments may offer a first step (Nelson and Winter 1982). By way of illustration a simple sequential barrier market game is simulated.

The scope of the thesis implies that many areas of research are touched upon in the pages to follow. In particular, economic psychology, industrial organization, labor economics, economic dynamics and the microfoundation of macroeconomics receive attention. This implies that surveys of the literature show all the features of a (non)random sample. However, hopefully they succeed in communicating the import of the message of the literature concerned. In the next chapter an outline of the main arguments is presented so as to announce which topics are treated where.

CHAPTER

2

RESULTS AND GUIDE

The two chapters of Part II deal with *economic psychology*. Although the subject matter of economic psychology is, at least, as old as economics and psychology themselves, the specialization of economic psychology is a recent invention. For example, the *Journal of Economic Psychology* was only founded in the 1980s. Not surprisingly, economic psychology explores the interface between the two sciences involved. It deals with the psychology of economic behavior and the economic mind. In doing so, it examines tax evasion, children's learning of economic concepts, perception of inflation, properties of entrepreneurship, and so on and so forth. Here the examination is restricted to the hard core of economic psychology: decision making. Study is made of whether the economic concept of *rationality*, particularly the popular "maximizing" representation, offers a reasonable guiding principle in describing and explaining economic decision making (Chapter 3). In particular, it is investigated whether or not the behaviorist "*satisficing*" notion constitutes a superior alternative (Chapter 4).

Parts III, IV and V are devoted to *industrial organization*. The main argument centers on presenting and elaborating an intermediate case of competition: the *barrier market*. Before rendering concrete the behavior that occurs in barrier markets (Parts IV and V), a *survey* of the literature seeks to locate the barrier market approach in the existing body of theory (Part III), particularly from the welfare theoretical perspective of benchmark cases of competition.

The four chapters of Part III indicate that a central question in industrial organization is concerned with the tradeoff between static [*i.e.*, (minimum) average cost prices] and dynamic (*i.e.*, optimal innovations) economies of market behavior (Chapter 5). The theories of *perfect competition*, *Bertrand competition* and *perfect contestability* introduce static optimal price setting to the detriment of incentives to introduce innovations (Chapter 6), whereas the theories of *imperfect competition* and *entry barriers* provide contrary results (Chapter 7). In commodity markets dynamic economies enter the picture through, in particular, process and product innovations. The literature on *innovative activity* is impressive (Chapter 8).

Part IV introduces barrier market theory. The barrier market framework offers a formal elaboration of a concept of *workable competition*. The concept of workable competition (and *barrier market theory*) as well as a few recent models seek to trace circumstances that generate an optimal tradeoff between static and dynamic efficiency of market performance (Chapter 9). Under particular conditions, *careful* price setting or wage demands as well as *incentives* to introduce innovations or human capital improvements are induced, which defines the barrier market (Chapter 10).

The two chapters of Part V play around with assumptions and conditions so as to reach barrier market scenarios. The barrier market concept is only valid under particular conditions (Chapter 11). The basic assumption is that a predominance of a threat of *hit-and-run* entry prevails (based upon, for example, home market arguments), which forces firms and laborers to adopt careful price setting respectively wage demands and, at the same time, to undertake (expected) optimal innovative investments. This outcome can even be obtained if the equilibrium strategy only gives a *zero (or normal) profit* (Chapter 12).

Part VI examines the stability (sustainability) features of barrier market configurations. The barrier market approach moves beyond the contestability concept by taking explicit notice of (nonstrategic and strategic) *investment* behavior which is, by definition, associated with the sinking of costs. Sunk costs, however, render particular difficulties in an environment of free entry. First, the introduction of sunk costs is associated with a reduced likelihood that market configurations appear to be *sustainable*, as any positive sunk cost gives a *unique* cost minimizing output level (Chapter 13). Second, if the cost and benefit of an innovation are spread over a sequence of periods, the existence of a (unique or conventional) *zero-profit amortization rule* is necessary to preserve stability (sustainability) of market behavior (Chapter 14).

Part VII applies the barrier market theory to competition in the commodity market. That is, the incumbent firms' equilibrium strategy in a barrier commodity market is modeled. It appears that firms in a commodity market are inclined to minimize the average total *cost* price and to maximize the *utility* of buyers embodied in the commodities offered by introducing process (Chapter 15) and product (Chapter 16) innovations respectively. Both strategies are necessary to reach a position of exit avoidance (and, probably, entry deterrence).

Part VIII comprises two chapters focusing on the labor market. In labor markets dynamic economies of market behavior take the form of, for example, *improvements in human capital*. The *static - dynamic efficiency tradeoff* is also apparent in the labor market (Chapter 17). The barrier market theory can offer an intermediate benchmark case of labor competition. Workers in a barrier labor market are induced to maximize their *attractiveness* to firms by implementing improvements in human capital (Chapter 18). Moreover, this innovative investment behavior shows up side by side with *careful* (zero-returns) wage demands.

However, competition is reflected not only in partial equilibrium conditions, but in general frameworks as well. The four chapters of Part IX examine the role *multimarket models* play in economics in general and industrial organization in particular. The literature on the microfoundation of macroeconomics and entry-(detering) strategies in a multimarket setting suggests a general framework of multimarket competition in which vertical and horizontal spillovers between related markets play a critical role (Chapter 19). The literature on general equilibrium theory pays considerable attention to *vertical* multimarket (commodity-labor) spillovers that follow from *rationing*. An examination of the implications of partial rationing points out that in a multi barrier market model *tdtonnement* processes are partly replaced by entry and exit movements. A tendency to price and cost equalization and a possibility of the (co-)emergence of differentiated unemployment and vacancies is indicated (Chapter 20). Industrial organization explores the *horizontal* multimarket (commodity-commodity) spillovers which can show up in competition among firms in *related* markets. The implications of multimarket competition are illustrated in a two-market model with investment in the scale of productive capacity

(Chapter 21).

The four chapters of Part X conclude the thesis by focusing on dynamic modeling and welfare theory. The argument starts off with a review of the methods and techniques of dynamic (or intertemporal) modeling in economics (Chapter 22). The analysis of the dynamics of *competitive processes* in industrial organization requires the introduction of three distinguishing features: expectational errors, heterogeneous agents and nonsimultaneous decision making. For illustrative purposes, a first attempt to carry out such an analysis is presented by *simulating* a simple sequential barrier market game (Chapter 23). The introduction of a benchmark case of competition raises the question whether or not policy *interventions* are needed so as to reach as near as possible to the preferred ideal. The barrier market framework focuses on allocative efficiency only. Hence, insufficient notice is taken of the implications to *distributive justice* (Chapter 24). The thesis concludes with a brief argument on the major *interrelations* between the fields of research that were discussed in the preceding 24 chapters (Chapter 25).

Three remarks can conclude the introduction. First, Figure 2.1 (see the last page of this study) offers a readers' guide by summarizing the major links, conclusions and questions that indicate the thread running through the argument of this thesis. Of course, other cross-references can be made and other results and puzzles can be mentioned. However, one hopes that Figure 2.1's road map is able to spare readers the feeling of getting lost while making their way through the ten parts and twenty-five chapters. Second, survey chapters and sections may be redundant for those insider readers already familiar with the literature under review. However, for interested outsider readers these review chapters and sections hopefully communicate the import of the arguments in the literature. Third, the many quotations permit the superb formulations of experts to be a natural part of the argument in this book: honor where it is due, to whom it is due.

 

PART

II

ECONOMIC PSYCHOLOGY

Rationality and Decision Rules

KEY WORDS. *Part II reviews two essential postulates of (micro)economics. First, the rationality debate is considered (Chapter 3). The common attitude among the opponents of the rationality concept results too often in throwing out the baby with the bath water. Second, a survey of the boundary of economic modeling points out that the maximizing postulate is able to assimilate many of the pronounced anomalies (Chapter 4). The reflection on the rationality debate and maximizing postulate ends in the choice of the maximizing principle as this thesis' mode of modeling.*

CHAPTER

3

RATIONALITY IN ECONOMICS: A Confusing Debate

3.1 RATIONALITY AND DIVISION OF OPINION

The rationality concept in economics bears a certain resemblance to a magnet. Economists either take irresistible fancy to it or incessantly challenge the postulate. The result is not at all elevating: Paradoxes are shaped; contradictions are kept alive; and sketches of strawmen are not exceptional. The catchword rationality is used in and out of season, often as a matter of course. As Furnham and Lewis (1986) note:

'Rationality' becomes confusing as it can mean so many things: it can mean dispassionate, reasonable, logical, explainable, purposeful; it is not always made plain quite what is intended (Furnham and Lewis 1986, p. 12).

The main reason for this somewhat disappointing state of art lies in the fact that the rationality concept (and the associated maximizing postulate¹) are not at all unequivocally interpretable. The fact that Schoemaker (1982), for example, discerns nine varieties of and four approaches to the "Expected Utility Model" is telling.

Nevertheless, theorists often adopt "the" rationality concept as if it were a clearcut label. However, because the opposite holds, the result is a confusion of tongues. This chapter presents a description of the literature which reflects the confusion of tongues. Rationality is loosely defined as subjective maximization. The examination is ordered by means of a confrontation of "the" rationality concept and perfect foresight (Section 3.2), the notion of the selfishness of human beings (Section 3.3), the results of empirical research (Section 3.4), the properties of a tautology (Section 3.5) and the philosophy of determinism (Section 3.6). The examination strengthens the conclusion that the debate about the "rationality of behavior" has resulted in, on the one hand, many pronounced contradictions which rest, in fact, upon paradoxes and, on the other hand, the transformation of "the" rationality concept into an unproductive misnomer (Section 3.7).

3.2 RATIONALITY AND PERFECT FORESIGHT

Simon (1945) - the founding father of the satisficing notion - regards perfect rationality, perfect knowledge and perfect foresight as different facets of the same cube, because

¹ The maximizing postulate is by far the most popular device to model rational decision making (Chapter 4).

the subject, in order to perform with perfect rationality in this scheme, would have to have a complete description of the consequences following from each alternative strategy and would have to compare these consequences. He would have to know in every single respect how the world would be changed by his behaving one way instead of another, and he would have to follow the consequences of behaviour through unlimited stretches of time, unlimited reaches of space, and unlimited sets of values (Simon 1945, p. 69).

This interpretation of mainstream economics of decision making - being based upon the marginalist maximizing postulate - has been common until the present days. Since the mid-1970s, Nelson and Winter, for instance, have consolidated the identification of marginalism on the one hand and perfect knowledge and perfect foresight on the other time and again by stressing that

[i]n contemporary formal theory, profit *seeking* behavior is typically represented as profit *maximizing* behavior with choice sets precisely known and given (Nelson and Winter 1980, p. 179; Nelson and Winter's italics).¹

In order to enrich the economics of decision making Simon (1945) and many successors (for instance, Cyert and March 1963, Nelson and Winter 1982 and Selten 1987) propose to replace perfect rationality by bounded or limited rationality, which amounts to

the behaviour of human beings to *satisfy* because they have not the wits to *maximize* (Simon 1945, p. xxviii; Simon's italics).

Satisficing differs from maximizing, because

[i]t is obviously impossible for the individual to know *all* the ... consequences [of the list of *all* alternative strategies], and this impossibility is a very important departure of actual behaviour from the model of objective rationality (Simon 1945, p. 67; Simon's italics).²

However, Simon and his successors oppose a self-created strawman of choice theory, because maximizing on the one hand and information scarcity and uncertainty on the other are not at all incompatible phenomena.

The founding fathers of marginalism, a specific variant of maximizing, already had an open eye for the role of uncertainty in a maximizing framework. Van Witteloostuijn and Maks (1989 and 1990) point to Walras' (1874) insights into uncertainty. Besides, Robbins (1932, pp. 77-79) and Hicks (for instance, 1939, p. 58) treated in detail the compatibility of uncertainty and marginalism. In fact, the founding fathers of marginalism may be regarded as forerunners of the well-known "Expected Utility Hypothesis" of Morgenstern and Von Neumann (1944), which introduces *risk* into a maximizing framework.³ Moreover, Arrow and Hurwicz (1972) even have incorporated *ignorance* in a marginalist theory. Recently, Fuhrer (1987) captured *model uncertainty* in a maximizing analysis of learning behavior and expectations formation (Van Witteloostuijn 1990a). In addition,

¹ Recently, Nelson (1986) admitted that the maximizing postulate allows for the introduction of probabilistic uncertainty (pp. 450-451).

² In addition, advocates of the satisficing notion wrongly claim to move beyond the scope of the maximizing postulate through offering theories of goal setting, expectations and group choice (Chapter 4).

³ Morgenstern (1935) already contributed to the issue of perfect foresight (and perfect knowledge) and equilibrium.

Stigler's (1961) seminal contribution set into motion the development of *search theory*, which reflects the incorporation of incomplete knowledge of the set of alternative courses of action in a maximizing context.

Although the application of marginalist principles does not require a list of *all* alternative strategies, its use is conditional upon a *continuity* property of this list. However, decision analysis (for instance, Raiffa 1968 and Keeney and Raiffa 1976) reflects a *discrete* version of the maximizing postulate by means of introducing the premise that decision makers choose just the *best* strategy from a number of action alternatives, an important one being to postpone the ultimate decision in order to learn along the way through time by exploiting

the gradual, time-dependent unfolding of uncertainties (Keeney and Raiffa 1976, p. 11).¹

In this way, the maximizing postulate too includes routinized behavior. The relevance of routinized behavior has been stressed by "satisficing" theorists:

The organization uses standard operating procedures and rules of thumb to make and implement choices. In the short-run these procedures dominate the decisions made (Cyert and March 1963, p. 113).

However, besides discrete decision analysis allowing for the introduction of routinized behavior, continuous marginalism also enables the incorporation of "habit formation". In this respect, Houthakker and Taylor's (1966) and Akerlof's (1980) application to demand theory and labor supply behavior respectively are worth mentioning. Besides, Heiner (1983 and 1986) offers a formal model in which routinized behavior is the result of maximizing decision making in an uncertain environment.

The fact of the matter is that routinized behavior can be rational. As Von Mises (1949) puts it:

Most of a man's daily behavior is simple routine. He performs certain acts without paying special attention to them. He does many things because he was trained in his childhood to do them, because other people behave in the same way, and because it is customary in his environment. He acquires habits, he develops automatic actions. But he indulges in these habits only because he welcomes their effects. As soon as he discovers that the pursuit of the habitual way may hinder the attainment of ends considered as more desirable, he changes his attitude (Von Mises 1949, pp. 46-47).

In fact, one may wonder whether the satisficing postulate is a specialization of the maximizing principle, rather than an alternative and opposing approach (see, for instance, Schoemaker 1982, p. 539). If the former thesis holds, the satisficing notion amounts to an *ad hoc* specification of the maximizing principle in the case of, for example, high cost of information gathering and processing (Chapter 4). This line of reasoning leads to arguing that results of satisficing decision making converge with those of maximizing decision making. Day (1967), for instance, has shown that under particular conditions maximizing and satisficing decision rules on price setting lead to the same results.

¹ In this way, a point of contact is offered to pay attention towards endogenizing the *moments* of decision making (see, for example, Hoogduin and Van Witteloostuijn 1987a and Van Witteloostuijn 1989a).

3.3 RATIONALITY AND SELFISHNESS

Economic rationality is often associated with self-interest. In this respect, Hirshleifer (1985) traces the dominance of self-interested rationality back to Smith (for instance, 1776, p. 14)¹ and points to Edgeworth (1881) as a clear example of this attitude in the neo-classical tradition²:

The first principle of Economics is that every agent is actuated only by self-interest (Edgeworth, 1881, p. 16).

A recent advocate of the premise of self-interest is Posner (1977):

Economics ... explores and tests the implications of assuming that man is a rational maximiser of his ends in life, his satisfactions - what we shall call his 'self-interest' (Posner 1977, p. 3).

Akerlof (1984) condemns the assumed tradition of self-interest in economics by arguing that

[t]he unwritten rules that only *economic phenomena* be considered in economic models, with agents as individualistic, selfish maximisers, restrict the range of economic theory and in some cases even cause the economic profession to appear peculiarly absurd (Akerlof 1984, p. 2; Akerlof's italics).

Of course, economists may defend their position by arguing that all goals are, in essence, self-interested. However, as Hirshleifer (1985) properly observes, the distinctive opportunities of the difference between selfishness and nonselfishness would then be lost in a verbal trick. Therefore, a native judgement of the assumed inseparable pair of twins - economic rationality and selfishness - should hinge on other arguments. In this context three remarks are in order.

First, the exclusive link between economics and selfishness has already been questioned in the first half of the century. For example, Robbins (1932) has argued that the association of economic maximization and selfishness does an injustice to the flexibility of the maximizing framework. The modeling of economic behavior is by no means associated with specific (selfish) goals on *a priori* grounds.

Second, the ensemble acting of self-interested maximizers may, obviously, lead to nonselfish results. In this respect, the belief of many contributors to economic theory, from ancient to modern times, in the beneficial working of the free market process³, although individual participants are self-interested, tells its own story. Becker's (1976)

¹ However, in his *Theory of Moral Sentiments* (1759) Smith goes beyond pure self-interest. Holler (1986) points out that "[i]n fact, from his 'Theory of Moral Sentiments' (1795) follows that human action in a society was checked and balanced by moral sentiments. This also applies to the pursuit of self-interest: 'How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it, except the pleasure of seeing it' (Smith, 1966, p. 1)" (Holler 1986, p. 224).

² The belief in the dominance of self-interest is prominent in neo-Austrian theory as well. Neo-Austrian economists (for example, Hayek 1949 and Kirzner 1973) argue that self-interested behavior leads to profitable results for society as a whole (provided that state intervention is limited).

³ For instance, Smith's (1776) invisible hand, Walras' (1874) *tâtonnement* and the neo-Austrian (Hayek 1949 and Kirzner 1985) arguments of the discovery process.

"Rotten-Kid Theorem" reflects another example, aggravated to family behavior.

Third, in relation with the first remark, it should be stipulated that the maximizing postulate is not incompatible with nonselfishness. The utility framework can capture selfish as well as nonselfish goals. In this respect, game theory has a long tradition of including altruist and cooperative behavior in a utility framework (see, for instance, McClintock 1972, Griesinger and Livingston 1973 and Liebrand 1981).¹ Moreover, recently economists introduced nonself-interested motives of behavior in maximizing frameworks. Worth mentioning are Collard (1978), Margolis (1982), Akerlof (1983) and Hirshleifer (1985). Akerlof, for example, models the educational behavior of parents. The parents are motivated by the utility of their children. This points to a useful definition of nonselfishness. Hirshleifer (1985), in line with Collard (1978, p. 7), argues that

someone is non-self-interested to the extent that he or she attaches utility to the impact of events upon the bodies or psyches of other parties (Hirshleifer 1985, p. 55).

This definition makes clear that an association of rationality - in the sense of the maximizing postulate - with selfishness does an injustice to the flexibility of this framework.

3.4 RATIONALITY AND EMPIRICAL EVIDENCE

On the basis of an extensive review of the empirical research bearing on the "Expected Utility Model", Schoemaker (1982) concludes that

[t]he research reviewed in this article suggests that at the individual level EU maximization is more the exception than the rule, at least for the type of decision tasks examined (Schoemaker 1982, p. 552).

In fact, Schoemaker undermines the position of the Expected Utility Models by means of two arguments. First, empirical evidence is in conflict with the axioms of the theory. Second, the Expected Utility Model fails to assimilate many real life aspects of decision making, in particular the limited information processing abilities of human beings, context effects and psychological aspects of decision making behavior. However, the lines of defence of the maximizing postulate are not easily demolished.

The opposition against the maximizing postulate through undermining the empirical relevance of its axioms can be moderated in six ways. First, violations of the continuity property do not undermine the maximizing postulate. As argued in Section 3.3, decision analysis reflects a *discrete* version of the maximizing postulate. Second, Machina (1982) shows that Expected Utility Analysis is *robust* to failures of the independence axiom.² Machina's "generalized expected utility analysis" is able to generate predictions consistent

¹ Of course, one may argue that cooperative behavior raises the individual's level of utility so that it can be traced to self-interest. However, as pointed out above, this verbal trick leads to reduced distinctive opportunities. Whether or not cooperative behaviour should be regarded as being guided by self-interested or nonself-interested motives depends on the role played in the decision maker's utility function by the utility level of others.

² "One of several equivalent versions of this axiom reads 'a risky prospect A is weakly preferred (i.e. preferred or indifferent) to a risky prospect B if and only if a $p:(1-p)$ chance of A or C respectively is weakly preferred to a $p:(p-1)$ chance of B or C, for arbitrary positive probability p and risky prospects A, B, and C'" (Machina 1982, p. 278).

with several well-known paradoxes, in particular the Allais (1953) paradox, the Friedman-Savage (1948) hypothesis, the St. Petersburg Paradox (for instance, Samuelson 1977), the Markowitz (1952) hypothesis and decision makers' systematic oversensitivity to the chance of low probability events (for example, Kahneman and Tversky 1979). Third, the maximizing postulate and the Expected Utility Hypothesis as a *decision rule* are not just two sides of the same coin. In particular, risk attitudes may provoke decision makers to adopt other types of maximizing decision rules like, for instance, the minimax and maximax rules in order to maximize dominant goals (risk aversion and risk seeking respectively)¹, in this way introducing varying decision *weights*. Fourth, in relation to the third line of defence, economists have introduced *reformulations* of the utility functions involved. Chew and McCrimmon (1979) and Hagen (1979) respecify the mathematical forms of utility functions such that they accommodate empirical violations of expected utility theory. Regret theory (Bell 1982 and 1985; Loomes and Sugden 1982 and Sugden 1986) is based upon the introduction of an additional goal - the avoidance of regret or disappointment - in an utility framework, which leads to the assimilation of violations of the independence axiom like, for instance, preference reversals (Grether and Plott 1979). Fifth, the *nonparametric* approach to utility maximizing avoids the necessity to introduce a particular specification of utility functions (Samuelson 1948, Houthakker 1950 and Afriat 1963 and 1967). The nonparametric method has been employed to test whether data sets are consistent with the hypothesis of utility maximization (for example, Koo 1963, Maks 1978, Varian 1983, and Houtman and Maks 1985). Sixth, the last line of defence is pronounced through the famous "*as if*" and "*natural selection*" arguments (Alchian 1950, Friedman 1953 and Machlup 1967), which boil down to the assertion that those decision makers who behave "*as if*" they apply a maximizing decision rule survive the economic struggle and, hence, dominate in reality.²

Moreover, the Expected Utility Model is accused of putting important real life phenomena outside the scope of the framework. However, this proposition is contradicted by the arguments in the preceding sections and the preceding paragraph. The fertile soil for the birth of the satisficing notion has been the "limited information processing abilities" argument. However, limited abilities can be represented in an utility framework through additional constraints (for example, *via* the introduction of a discrete formulation). A recent example is Sah and Stiglitz' (1985, 1986 and 1988) analysis of the relationship between organizational structures and limited computational abilities of maximizing decision makers.

The implications of the limitations of human information processing, through empirical evidence concerning the relevance of context effects³, provoked the development of prospect theory. Prospect theory (Kahneman and Tversky 1979 and 1986) seeks to offer an explanation of preference reversals by pointing to the relevance of the way of presenting and the structure of information. In addition, Schoemaker (1982) argues that Expected Utility Theory fails to absorb many psychological principles of judgement and choice. In particular, he emphasizes the importance of the decomposition of decision problems, the complexity of the task, the principle of isolation, the role of reference

¹ The influence of risk attitude on decision making is examined in portfolio theory (for example, Markowitz 1959). Huber (1980) discusses several decision rules in the context of managerial decision making.

² The "*as if*" argument is in itself sufficient to plead in favor of the lack of relevance of the empirical validity of the axioms underlying the maximizing postulate. Chapter 4 offers further details.

³ In relation with context effects Schoemaker admits the success of regret theory as regards to the accommodation of particular empirical violations of the axioms (1982, p. 548).

points and the bias to overconfidence. However, the points made so far stipulate the fact that the maximizing postulate is not inconsistent with such psychological principles. In effect, the introduction of psychological principles into a maximizing framework reflects a popular and fruitful hunting field of research. Akerlof and Dickens (1982), for instance, include Festinger's (1957) theory of cognitive dissonance in a utility model.

3.5 RATIONALITY AND TAUTOLOGY

The flexibility of the maximizing postulate with respect to the assimilation of points of criticism and apparent anomalies in its general framework has brought to the surface the argument that utility theory is, in fact, an effete tautology. As Schoemaker argues:

The essential premise of the postdictive EU view is that all *observed* human behavior is optimal (in the EU sense), *provided* it is modeled in the appropriate manner. Seeming suboptimality is explained, *ex post facto*, by introducing new considerations ... that account for the anomalies, so as to make them optimal. ... It is the latter degrees of freedom, however, that may make the postdictive approach tautological: i.e., non-empirical and non-falsifiable (Schoemaker 1982, p. 539; Schoemaker's italics).

Opponents of the maximizing postulate are often surprised by the flexibility of the utility framework, because they misinterpret the framework as maximizing a *particular* thing, rather than *something*. However, the maximizing postulate is not linked *a priori* to particular goals, restrictions or specifications. Therefore, the maximizing premise

has the form which is called an "All-and-some statement". All-and-some statements are neither verifiable nor refutable! ... Any alleged counterexample is unverifiable *even if the counterexample is true*. ... In other words, the verification of the counterexample requires the refutation of a strictly existential statement; and ... we all agree one cannot refute existential statements (Boland 1981, p. 1034; Boland's italics).

However, to say that a premise is untestable is one thing, but to call it a tautology is another. The tautological property of the maximizing postulate depends on the way it is employed.

Applied *ex post* utility theory, as many theories (including the behaviorist framework), represents a tautology, because human behavior can always be explained in retrospect with the help of formulating the appropriate goals and restrictions. However, applied *ex ante* the maximizing postulate no longer reflects a tautological argument, because the goals, utility functions and constraints have, in advance, to be adapted to concrete modeling. However, Schoemaker moves a step further by accusing a *series* of respecifications of an utility maximizing model to be tautological *ex post*:

The possible respecifications of a particular model are of course not arbitrary and must gain the approval of fellow practitioners. Furthermore, most model respecifications imply testable predictions. However, if these testable predictions were not to hold, additional respecifications would be sought until all relevant past (i.e., observed) behavior would be accounted for as indeed being optimal. It is in this sense that the perspective is *ex post* (Schoemaker 1982, p. 540).

Therefore,

[a] better alternative is to examine closely the type of anomalies reported, and the cognitive reasons underlying them (Schoemaker 1982, p. 554).

However, in connection with Schoemaker's line of reasoning three remarks are in order.

First, the maximizing postulate and the optimality concept are by no means equivalent. Theorists who suggest that the rationality and optimality concepts are one and the same (compare with Section 3.2), confuse expectations *ex ante* and evaluations *ex post* with one another.¹ In fact, the maximizing postulate implies the proposition that a decision maker decides in such a way as to reach his/her most preferred *expected* position, subject to the opportunities open to him/her (Hicks 1939, p. 58). Von Mises (1949) already pointed out that

[i]t is a fact that human reason is not infallible and that man very often errs in selecting and applying means. An action unsuited to the end sought falls short of expectation. It is contrary to purpose, but it is rational, i.e., the outcome of a reasonable - although faulty - deliberation and an attempt - although an ineffectual attempt - to attain a definite goal (Von Mises 1949, p. 20).

However, expected optimality at the moment of decision making will not necessarily prove to be optimal in retrospect, except in situations of perfect foresight. Second, Schoemaker's description of the *ex post* nature of the utility framework reflects the *ex post* nature of scientific progress in general, whether in the form of respecifying an existing framework or developing a new one. Therefore, his learning by testing argument does not necessarily lead to his conclusion of the inferiority of the respecification procedure.² Third, in relation to the second remark it is not clear why the examination of anomalies and the cognitive foundations concerned should be incompatible with the maximizing postulate.

3.6 RATIONALITY AND DETERMINISM

In particular, neo-Austrian and post-Keynesian economists argue in favor of the *voluntarist* component of human action (Van Witteloostuijn 1990b). A determinist stance amounts to assuming that environmental input determines the decision making output, whereas a voluntarist position is based upon the reversed relation so that

the environment is viewed as an output rather than an input (Weick 1979, p. 228).

Neo-Austrian and post-Keynesian authors trace the indeterminacy of decision making to the properties of uncertainty and expectations (Van Witteloostuijn 1989a).

The neo-Austrian economist Lachmann (1943), for instance, argues that

economic action, concerned with the future, so far from strictly determined by the set of objective "data", is often decided upon in a penumbra of doubt and uncertainty, vague hopes

¹ The impression that rationality and optimality are an inseparable pair of twins is fortified by general equilibrium theory, because the general equilibrium tradition is based upon assuming the maximizing postulate on the one hand and the equilibrium concept - defined as a situation without the disappointment of expectations (see, for example, Hicks 1939, Debreu 1959 and Grandmont 1977) - on the other (Chapter 20).

² In fact, Schoemaker bases the latter conclusion upon his impression of the empirical evidence violating the maximizing postulate (Section 3.4).

and inarticulate fears, in which ultimate decisions may well depend on mental alertness, ability to face the unknown. ... experience, before being transformed into expectations, has, so to speak, to pass through a 'filter' in the human mind, and the undefinable character of this process makes the outcome unpredictable. ... We now realise that ultimately it is the *subjective* nature of these beliefs which imparts indeterminateness to expectations as it is their *mental* nature which renders them capable of explanation (Lachmann 1943, pp. 65, 67 and 73; Lachmann's italics).

Post-Keynesian theorists, referring to Keynes' chapter 12 in *The General Theory* (1936) and the 1937 article¹, adopt a similar stance (for instance, Minsky 1975 and Kregel 1976 and 1982). Shackle (1972) asserts that

[e]xpectations and determinacy are incompatible and mutually exclusive. The rational ideal must therefore exclude expectation, and expectation, since it is real and insistently present and accompanies all activity, must destroy the rational ideal (Shackle 1972, p. 156)

Similar arguments have been expressed by the neo-Austrian economists Kirzner (1985, pp. 46-47) and O'Driscoll and Rizzo (1985, pp. 24-25). However, the advocates of the voluntarist position partly put the finger on the weak spot of the maximizing framework.

The major weakness of Shackle's argument is reflected in his interpretation of the maximizing postulate. In line with the authors mentioned in Section 3.2 Shackle associates rationality with perfect knowledge and perfect foresight (see, for instance, Boland 1981, p. 1032). Therefore, it is not surprising to notice that Shackle points to the incompatibility of indeterminate expectations and rationality. However, a maximizing postulate that captures incomplete knowledge and uncertainty is by no means in conflict with indeterminacy of expectations. Nevertheless, Shackle and Kirzner point to the major lacuna in the economics of decision making:

The body of orthodox economic doctrine in the main has not dreamt of regarding economic policy making ... as an *originitive and imaginative art* (Shackle 1972, pp. 364-365; Shackle's italics);

and

human history ... contains a strong element of the unexplained and even the spontaneous (Kirzner 1985, p. 47).

However, recognizing the indeterminate creative component of human decision making is one thing, but explaining it is another.

The explanation of creativity, not to mention the prediction of creative acts, is not only lacking in economics, but in psychology as well (Van Witteloostuijn 1990a):

Perhaps because of the same problem of complexity, professional psychology also neglected the question of creative thinking Psychology was challenged by the problem of creativity and its relation to thinking, as, for example, in Guilford's work (1964) on creativity and in his distinction between 'convergent' and 'divergent' thinking (1959), but it had not as yet produced a massive response. ... Creativity is often a process of 'rearrangement' and there are many remaining questions as to how, when, why, and by whom the new and old raw materials

¹ See, for example, Keynes (1936, p. 149 and 1937, p. 214).

... and under what conditions the original, the unpredictable, the psychologically new, and socially important appear (Murphy and Kovach 1972, p. 345).

Until the present day, the explanation of creativity reflects the boundary of the economics (and psychology) of human decision making, whether or not being based upon the maximizing postulate. The prediction of the *content* of a creative act is even logically impossible (Loasby 1976). Predicting the content of the creative acts of the future means transferring this creativity to the present.

3.7 CONCLUSION

The debate about the rationality of behavior in economics shows all the characteristics of a confusion of tongues. Opponents of the maximizing postulate often attack a self-created strawman of utility theory, leading to a useless verbal combat. Neither the association of rationality with perfect foresight, selfishness, and tautological arguments nor undermining the position of the maximizing postulate with the help of empirical evidence has been conducive to the reduction of the confusion of tongues. The list of confusing subject matters can be lengthened by posing the question whether rationality is incompatible with group decision making, dynamic processes, goal setting and heterogeneity (Chapter 4). The debate has transformed the rationality concept into an unproductive misnomer.

Of course, the combat has not only resulted in confusion, but has also stimulated the emergence of new insights, such as regret theory, the introduction of social and psychological phenomena in economics, and prospect theory. However, the common attitude among the opponents of utility theory results too often in throwing out the baby with the bath water. As long as opponents fail to formulate a workable alternative to the rationality concept, this attitude seems to result in an arid discussion.

CHAPTER

4

DECISION RULES: Maximizing or Satisficing?

4.1 INTRODUCTION: Caricatures

In Chapter 3 (in particular, Section 3.2) it is argued that the maximizing and satisficing principles amount to *one and the same* decision rule. Here this argument is fleshed out. An often adopted strategy to stress the relevance and renewal of an approach is to attack a caricature of an existing - usually "mainstream" - view (see, for example, Yunker 1986). In this respect, choice theory is time and again exposed to similar critique, an early example being Veblen (1898, p. 73). Advocates of behaviorism, being based upon the "satisficing" principle, also correspond to this attitude by presenting a slashing critique of a strawman of choice theory. Simon (1945, p. 69) - the founding father of economic behaviorism¹ - associates choice theory (in his words, perfect rationality) with complete knowledge, perfect foresight and computational ingenuity. A similar line of reasoning still flourishes, as, for instance, the way Earl (1984, p. xii) introduces his book shows by stipulating the superiority of behaviorism (with respect to the allowance for uncertainty and informational disabilities). Another clear, recent example of this standpoint is reflected in the work of Nelson and Winter (for instance, 1982).

Advocates of choice theory seem to adopt a different attitude. In general, they do not pay any attention at all to the attacks and achievements of behaviorist theorists. Their strategy of keeping dark the behaviorist tradition seems to reflect a belief in the superiority of maximizing models. In as far as they respond to criticism, it is in the context of the rationality debate (see, for example, Boland 1981). In all, the current state of affairs seems to suggest the existence of a contradiction between choice theory and behaviorism. The maximizing postulate and the satisficing principle are generally regarded as conflicting frameworks.

This chapter advances an argument that the contradiction between maximizing and satisficing disappears like snow under a hot sun as soon as behaviorist frameworks are compared with well-designed choice-theoretic models. Although the contradiction stands up if basic textbook versions of behaviorism and choice theory are compared, pioneering research shows that satisficing and maximizing are equivalent rather than opposite concepts. This proposition is based upon a review of the literature.

The behaviorist critique of choice theory is adopted as a starting point, because

¹ Economic and psychological behaviorism should be clearly distinguished. Behaviorism in psychology is based upon the stimulus-response learning theory (for example, Skinner 1938).

"satisficing" theorists are explicitly opposed to the maximizing postulate. In Section 4.2 the satisficing principle is briefly outlined. Moreover, the major criticisms of the maximizing postulate are presented. In Section 4.3 and 4.4 it is argued on the basis of old and recent elaborations of choice theory that the behaviorist critique overshoots the mark. In Section 4.3 this argument is preceded by a short introduction of the maximizing postulate. In general, it seems that the critique of the maximizing postulate originates from an underestimation of the flexibility of the framework. This proposition is illustrated in Section 4.5. A well-known debate in the context of the maximizing postulate centers on selection and "as if" arguments. In Section 4.6 some attention is paid to this controversy. Section 4.7 offers final remarks.

4.2 BEHAVIORISM: Critique of the Maximizing Postulate

Behaviorism is put forward by Simon (1945) as an alternative explanation of (organizational) decision making behavior that seeks to replace the traditional theory of perfect or objective rationality (choice theory)¹:

The task of decision involves three steps: (1) the listing of all alternative strategies (2) the determination of all the consequences that follow upon each of these strategies; (3) the comparative evaluation of these sets of consequences. The word 'all' is used advisedly. It is obviously impossible for the individual to know all their consequences, and this impossibility is a very important departure of actual behaviour from the model of objective rationality (Simon 1945, p. 67; Simon's italics)

In order to come close to a description of real world decision making in a complex and uncertain environment, Simon proposes to replace the maximizing postulate by the satisficing principle. The behaviorist theory of decision making has been supplemented and elaborated by March and Simon (1958) and Cyert and March (1963) in order to shape a general behaviorist theory of organizational behavior. In this context the latter argue that

[i]n order to develop an alternative theory, we need more satisfactory theories of organizational goals, organizational expectations, organizational choice, and organizational control (Cyert and March 1963, p. 21; Cyert and March's italics).²

In the context of a discussion of the relation between the maximizing postulate and the satisficing principle the alternative theory of choice is of particular interest. Besides, the theories of (organizational) goals and expectations show their relevance, because both reflect additional criticisms of choice theory.

In short, the three relevant building stones of behaviorism amount to the following:

¹ Objective rationality is opposed to bounded rationality. The adjective "bounded" indicates the computational limitations human beings are facing. The notion of bounded rationality is now in common use in frameworks separate from the behavioral theory of organizations. For example, the concept is a cornerstone of transaction cost economics (Williamson 1975 and 1979). However, transaction cost economics retains the maximizing principle at its basic level of analysis (through the postulate of the minimization of transaction cost).

² Although the satisficing principle applies to individual decision makers in general, much behaviorist research is engaged with organizational behavior. Therefore, many of the arguments proceed at the level of organizations.

(i) *Group decision making.* Rather than analyzing decision behavior of an individual in isolation, behaviorist theorists focus on decision making of a group of individuals, which implies that

[t]he goals of a business are a series of more or less dependent constraints imposed on the organization through a process of bargaining among potential coalition members and elaborated over time in response to short-term pressures. Goals arise in such a form because the firm is, in fact, a coalition of participants with disparate demands, changing foci of attention, and limited ability to attend to all organizational problems simultaneously.

In the long run, studies of goals of a business firm must reflect the adoption of goals to changes in the coalition structure (Cyert and March 1963, p. 43).

(ii) *Uncertainty and expectations.* With respect to expectations behaviorist theorists tackle certainty and uncertainty by discerning three situations:

(a) *Certainty:* theories that assume the decision maker has complete and accurate knowledge of the consequences that will follow on each alternative. (b) *Risk:* theories that assume accurate knowledge of a probability distribution of the consequences of each alternative. (c) *Uncertainty:* theories that assume that the consequences of each alternative belong to some subset of all possible consequences, but that the decision maker cannot assign definite probabilities to the occurrence of particular consequences (March and Simon 1958, p. 137; March and Simon's italics).

(iii) *Satisficing decision rule.* The essence of behaviorism is that it reflects

the theory of intended and bounded rationality - of the behavior of human beings who *satisfy* because they have not the wits to *maximize* (Simon 1945, p. xxviii; Simon's italics).

Firms' decision making leads to the adoption of routines or search behavior. As Cyert and March (1963) argue with respect to the former:

The organization uses standard operating procedures and rules of thumb to make and implement choices. In the short-run these procedures dominate the decisions made (Cyert and March 1963, p. 113).

Search behavior, in contrast, takes the form of

[a]n approximate sequential consideration of alternatives. The first satisfactory alternative evoked is accepted. Where an existing policy satisfies the goals there is little search for alternatives. When failure occurs, search is intensified (Cyert and March 1963, p. 113).

To recognize whether or not goals are satisfied, decision makers compare realizations of behavior with *aspiration levels*, the latter being the concrete manifestation of goals. If the realized results satisfy the aspiration levels, routinized behavior is continued; otherwise, a problem oriented search for alternative courses of action is initiated, the first satisfactory alternative evoked being accepted. The acceptance of an alternative compels the adoption of new or the adjustment of old routines.

The critique of behaviorist theorists of the maximizing postulate is closely linked to the three building stones of their alternative framework. In the first place, behaviorist theorists accuse choice theory of a (relative) ignorance of goal setting procedures, variability of goals over time and group decision making. These criticisms are not so much

related to the satisficing principle in the strict sense, but rather to the behaviorist theory of organizations. In the second place, behaviorist theorists argue that the maximizing postulate reflects an ideal picture of human decision making behavior. This critique arises from the satisficing decision rule itself and leads to the condemnation of the maximizing principle.

Simon (1945) sketches perfect rationality (his words for choice theory) in the following way:

It has already been remarked that the subject, in order to perform with perfect rationality in this scheme, would have to have a complete description of the consequences following from each alternative strategy and would have to compare these consequences. He would have to know in every single respect how the world would be changed by his behaving one way instead of another, and he would have to follow the consequences of behaviour through unlimited stretches of time, unlimited reaches of space, and unlimited sets of values (Simon 1945, p. 69).

Ever since the mid-1970s Nelson and Winter have consolidated Simon's interpretation and critique of choice theory by stressing, for example, time and again that

[i]n contemporary formal theory, profit *seeking* behavior is typically represented as profit *maximizing* behavior with choice sets precisely known and given (Nelson and Winter 1980, p. 179).

Hence, the second line of critique amounts to associating choice theory with complete knowledge and perfect foresight in combination with computational ingenuity. In this way, maximizing frameworks fail to introduce uncertainty and routinized behavior into the scope of human decision making.

However, is the behaviorist critique of choice theory justified? In Section 4.3 this question is examined in the light of the first line of critique (ignorance of goal setting procedures, group decision making and variability of goals). Section 4.4 is devoted to the assumed incompatibility of choice theory and uncertainty and routinized behavior.

4.3 THE MAXIMIZING POSTULATE: Goals, Time Variability and Groups

Choice theory shows itself in the form of marginalism and decision theory. The former reflects a continuous and the latter a discrete variant of the maximizing postulate. Marginalism enjoys a dominant popularity in economics, whereas decision theory is mainly elaborated in psychology (and business administration). Choice theory became firmly embedded in economic science through the introduction of *marginalism* by neo-classical authors. Marginalism originates from the seminal contributions to economic theory of Jevons (1871), Menger (1871) and, in particular, Walras (1874). Marginalism is based upon the maximizing postulate in the sense that (individuals') decision making is guided by maximizing an objective (profit, utility) function subject to (technological, budget) constraints. Marginal decision making leads to a choice of the levels of the decision variables such that the benefit and cost of behavior are equated at the margin. The application of marginalist principles requires the existence of a continuous set of action alternatives.¹ *Decision analysis* (see, for instance, Raiffa 1968, Keeney and Raiffa

¹ This does not mean that continuity always implies marginalism.

1976 and Huber 1980) constitutes a discrete variant of choice theory. Decision analysis is based upon the premise that (individual) decision makers choose the best alternative course of action out of a number of action alternatives.

The question is whether choice theory can handle goal setting procedures, time variability of goals and group decision making. With respect to goal setting three remarks are in order. First, Cyert and March (1963) approach the misty arena of the questioning of means and ends by arguing that a firm's goals change over time. To be precise, the firm's restrictions change over time as the goals of organizational members are transformed into organizational constraints through bargaining. However, does this imply that firms may drop the ultimate goal of (maximizing the probability of) survival in favor of other ends? The objectives (or organizational constraints) Cyert and March are concerned about, are at least secondary goals or even means. However, the choice of neither goals nor means lies outside the scope of choice theory (see Section 4.4, in particular search theory). In as far as the endogenization of the driving ends of human beings are concerned, neither behaviorism nor choice theory offers a solution. It has been argued that the exogeneity of the basic ends in economic frameworks is inevitable (Robbins 1932, Hennipman 1945, Shackle 1972, Kirzner 1985, Langlois 1986 and Van Witteloostuijn 1990a).

Second, a theory of goal setting is complementary rather than contradictory to both choice theory and behaviorism, if these frameworks are regarded as strict rules of decision making. The application of either maximizing or satisficing decision rules is a sequel to goal setting (or restriction formulation). Goal setting considerations cannot plead in favor of either maximizing or satisficing. After goals having been defined, behaviorist as well as choice-theoretic procedures can be applied to make decisions.

Third, Cyert and March argue that goals are the result of "a process of bargaining among potential coalition members" (Cyert and March 1963, p. 43), so that goals vary over time due to "changes in the coalition structure" (Cyert and March 1963, p. 43). However, again this argument is unable to plead in favor of the plausibility of either behaviorism or choice theory. Bargaining game theory, being based upon the maximizing postulate, offers advanced tools to tackle the (time dependency of) goal setting of groups (see, for instance, Peters 1986). The (bargaining) interactions among individual maximizers determine the choice of the ends. A changed composition of participants leads to a renewed process of goal (or means) setting or learning. In this way, the opportunity arises to introduce time variability of goals into a choice-theoretic framework.

Hence, the *time variability* of ends is not incompatible with choice theory. Related examples concern the introduction of preference changes in a maximizing framework (for instance, Houthakker and Taylor 1966, Stigler and Becker 1977, Pollak 1978, Kahneman and Tversky 1979, Akerlof 1980, Kapteijn 1985 and Frank 1987). For example, Akerlof introduces reputation and social norms in a choice-theoretic model to explain (part of) unemployment phenomena. Reputation and social norms are arguments in the utility function as well as the "social" constraint. The utility function varies over time, because social norms (and thereby social restrictions) change in the long run.

The change of preferences over time may be caused not only by changing utility functions, but by changing information as well, as "prospect" theory shows. Prospect theory (Kahneman and Tversky 1979) seeks to offer an explanation of "preference drifts". A preference drift shows itself in preference reversals (see, for instance, Grether and Plott 1979). Depending on the structure of information individual decision makers prefer one alternative course of action to another. However, a change in the structure of the information received may lead to reversed preferences. As Schoemaker (1982) and Van

de Poel (1986) argue, prospect theory is not a departure from choice theory, but rather a particular variant which includes decision weights.

A third question concerns to the (in)compatibility of *group decision making* and choice theory. Again, choice-theoretic models have something to say on this phenomenon. Psychologists, for example, have produced choice-theoretic notions of decision making behavior of groups (in particular, Edwards 1976 and 1977), being based upon an aggregation procedure of the wishes of individual participants containing deliberation rounds and the composition of an aggregate utility function. The latter is the subject of research in social choice theory (for example, Arrow 1951 and Sen 1982)¹, being engaged in analyzing group decision making of society as a whole to find "a function of individual preferences to a social preference order" (Elster and Hylland 1986, p. 2).

Another angle is to regard a group decision as result of joint acting of individual participants' decision making behavior, leaving room for the introduction of maximizing participants. In this respect, an important decision is whether to stay in or leave a group (for example, Thibaut and Kelley 1959). The results of group decision making are constraints to individual participants. As noted above, game theory (for instance, Morgenstern and Von Neumann 1944 and Shubik 1984) offers tools to study the outcome of decision making of interacting maximizing individuals.

The argument in this section leads to the conclusion that the alternative elements of behaviorism - attention to goal setting procedures, time variability of ends and group decision making - are not incompatible with choice-theoretic models. Hence, those aspects of decision making do not reflect arguments that can plead in favor of either behaviorism or choice theory. However, do the strict decision rules in the form of the maximizing postulate and the satisficing principle conflict with one another?

4.4 THE MAXIMIZING POSTULATE: Uncertainty and Routines

4.4.1 Introduction

The second line of criticism of choice theory is the association of the maximizing postulate with complete knowledge and perfect foresight in combination with computational ingenuity. It is argued that maximizing frameworks, therefore, fail to introduce uncertainty and routinized behavior into the scope of human decision making. This critique touches upon the assumed contradiction of maximizing and satisficing as strict decision rules.

The fertile soil for the association of maximizing with complete knowledge and perfect information is often assumed to be the application of maximizing principles in the context of equilibrium theory (Chapter 20) in line with Walras (1874). Prominent contributions in the first half of the twentieth century worth mentioning, which engendered the emergence of general equilibrium theory as a dominant framework in economics, are Hayek (1928), Lindahl (1929), Wald (1936), Von Neumann (1937), Hicks (1939) and Samuelson (1941). In the mid-1950s general equilibrium theory cumulated in the well-known ADM model of Arrow, Debreu and McKenzie (Arrow and Debreu 1954, Debreu 1959 and McKenzie 1959). However, in the ADM model individuals feel

no need to revise strategies, because the choice of a strategy has already taken account of the

¹ This issue is part of welfare theory (Chapter 24).

structure of information in the future, i.e., what information will be available at each date. ... the present value of each producer and each consumer is *known with certainty at the beginning* (Radner 1968, p. 34; italics added).

In fact, general equilibrium theory rests upon the adoption of Walras' assumption of, at least, short-term perfect information (Walras 1874, p. 317).¹ Probably, this link between equilibrium, based upon the perfect information assumption, and marginalism provoked the persistence of the association of choice theory with perfect foresight (see, for instance, Nelson 1986, pp. 450-451).

However, Robbins (1932), formulating the essentials of economic theory and marginalism, argued that the perfect foresight assumption

is not to foster the belief that the world of reality corresponds to the constructions in which they figure, but rather enable, in isolation, tendencies which, in the world of reality, operate only in conjunction with many others, and then, by contrast as much as by comparison, to turn back to apply the knowledge thus gained to the explanations of more complicated situations (Robbins 1932, p. 94).

Therefore, it is not surprising that the marginalist theorists transformed the extreme variant of choice theory in the ADM model - with its assumptions of complete knowledge and perfect foresight - into general forms of choice theory by introducing uncertainty into the focus of the individual decision maker.²

In this context, clarification may be brought by examining the way in which choice theorists incorporate different kinds of uncertainty in their models. The introduction of uncertainty is associated with incomplete knowledge and computational disabilities. In Subsection 4.4.2 a taxonomy of uncertainty concepts is examined. Subsection 4.4.3 is devoted to the association of choice theory with perfect foresight. In Subsection 4.4.4 the link between the maximizing postulate and complete knowledge is considered. In Subsection 4.4.5 the attempts to introduce routines and computational abilities into a maximizing framework are presented.

4.4.2 A Taxonomy of Uncertainty Concepts³

The components of the decision process point to the possibility of differentiating uncertainty by focus (Figure 4.1). To make a decision an individual decision maker evaluates a choice set of alternative courses of action with respect to the cost and benefit of the alternatives concerned. In order to derive a clear picture of the decision problem the individual requires information about the choice set as well as the cost and benefit of alternatives. However, information scarcity may result in uncertainty about the composition of the choice set of alternative courses of action (*action uncertainty*), the cost and benefit of the alternatives concerned (*yield uncertainty*) and/or the availability of additional information and the cost and benefit of processing and gathering it (*information uncertainty*).

¹ Walras includes long-term uncertainty in his general equilibrium theory (for example, Walras 1874, p. 310), as argued by Van Witteloostuijn and Maks (1989 and 1990).

² In fact, pre-war contributions like Hayek's (1928), Lindahl's (1929) include a role for uncertainty to play. Moreover, an alternative formulation of the ADM model runs in terms of markets for contingent claims (Debreu 1959), which allows economic agents to have different foresights (Chapter 20).

³ This section is based upon Van Witteloostuijn (1989a), which offers an examination of a differentiated approach to uncertainty in the context of psychology.

Figure 4.1
Focus of Uncertainty

Moreover, uncertainty is not only differentiated by focus, but also by type. With respect to the latter it is important to indicate the extent to which (potentially) available information permits the reduction of uncertainty. The extent to which uncertainty prevails even when all available information is processed is called fundamental uncertainty. In this respect, Keynes' (1921) and Knight's (1921) contribution to the distinction of types of uncertainty is worth mentioning. Knight (1921) argues that

[i]t will appear that a *measurable* uncertainty, or 'risk' proper, as we shall use the term, is so far from different from an *unmeasurable* one that it is not in effect an uncertainty at all. We shall accordingly restrict the term 'uncertainty' to cases of the non-quantitative type (Knight 1921, p. 20; Knight's italics).

In the case of measurable uncertainty, or *risk*, individuals are able to attach probability estimates to particular variables or conditions, whereas in the case of unmeasurable

uncertainty, or *fundamental uncertainty*, they feel unable to do so.¹ Hence, the behaviorist distinction between certainty, risk and uncertainty (March and Simon 1958, p. 137) is preceded by far by Keynes' and Knight's!²

4.4.3 Choice Theory and Perfect Foresight

A major impulse to the introduction of uncertainty into choice theory is reflected in Hicks' pre-war publications, in particular *Value and Capital* (1939). A first attempt to incorporate uncertainty into economic choice theory in a formal way is put into equations by Morgenstern and Von Neumann (1944). As Hey (1979) describes the basic assumptions of their analysis:

We are concerned with the decisions of an individual when confronted with sets of alternative *choices*. Although the eventual outcome of any particular choice will be one of a set of *basic outcomes* (whose worth to the individual is certain), *which* basic outcome will occur is not known in advance. However, it is assumed that the individual can attach probabilities (either objective or subjective) to the various possible basic outcomes of each alternative choice (Hey 1979, p. 27; Hey's italics).

Hence, Morgenstern and Von Neumann just introduce yield risk concerning given and known alternatives. Morgenstern and Von Neumann's "Expected Utility Hypothesis" is extensively used by many theorists, and has received its place in "mainstream" textbooks of economics (for example, Koutsoyiannis 1979 and Varian 1984). The probability estimates attached to "basic outcomes" are divided into cost and benefit probability measures.

In psychology the same kind of probability estimates are employed by decision analysts (see, for instance, Keeney and Raiffa 1976). Moreover, Maital (1982, pp. 266-268) argues that many psychological theories are based on a variant of the "Expected Utility Hypothesis". In this context, Lewin's field theory (1951), Rotter's social learning framework (1954) and Edwards' multi-attribute rating technique (1977) are worth

¹ Keynes (1921 and 1936) discerns situations of fundamental uncertainty as well, as has recently been stipulated again by Lawson (1985) and Hoogduin (1987). Post-Keynesian economists like Davidson (1972) and Kregel (1976 and 1982) support Keynes' emphasis on fundamental uncertainty. In fact, Keynes' insights concerning the implications of fundamental uncertainty, individuals' trust in conventions (Keynes 1936, p. 152), looks remarkably similar to Simon's routines (Van Witteloostuijn 1989a). Recently, theorists traced Keynes' thinking back to his *Treatise on Probability* (1921) to recover an additional type of uncertainty (Stohs 1980 and Hoogduin 1987). This type of uncertainty refers to the individuals' perception of the balance between the amount of available and nonavailable information upon which decisions and expectations have to be based. As argued by Hoogduin (1987) and Hoogduin and Van Witteloostuijn (1987a), the influence of this type of uncertainty extends to the moments of decision making on long-term commitments. This points to a possible rationale for the adoption of routinized or conventional behavior. The relevance of confidence is also stipulated by Savage (1972). Savage distinguishes three types of probability theory. *Personalistic* theories include the confidence an individual has in the truth of a particular proposition (1972, p. 3). Besides, the notion of confidence bears a family resemblance to ambiguity (for example, Einhorn and Hogarth 1986).

² The distinctions of Koopmans and Davidson show a family resemblance to the difference between risk and fundamental uncertainty. Koopmans (1957) argues that *primary* (in contrast to *secondary*) uncertainty is concerned with random acts of nature and unpredictable changes in consumers' preferences. Davidson (1983) distinguishes *ergodic* and *nonergodic* stochastic processes. Only ergodic processes can be associated with risk measures, whereas nonergodic processes are by definition unpredictable.

mentioning. In addition, Lea *et al.* (1987, p. 130) point to Skinner's reinforcement notions (1938).

An example of the formal introduction of fundamental yield uncertainty - traced to the benefit side of alternative courses of action - into a choice-theoretic framework is the contribution of Arrow and Hurwicz (1972). Their 1972 paper is based upon their work in the early 1950s and intends

to offer a possible characterization of the concept of complete ignorance. Like other formulations, the problem is taken to be that of choice of an action from a given set when the consequences of any action are functions of an unknown state of nature. However, the properties regarded as defining an optimal choice are designed to reflect completely the idea that there is no *a priori* information available which gives any state of nature a distinguished position. Most importantly, the optimality criterion differs from those in the now more standard subjective probability framework by not presupposing a fixed list of states of nature (Arrow and Hurwicz 1972, p. 1).

Moreover, recently choice theorists have introduced ambiguity into a maximizing framework. Einhorn and Hogarth (1986) relate ambiguity to risk and ignorance by saying that

[w]e can ... distinguish between ignorance, ambiguity, and risk, according to the degree to which one can rule out alternative distributions; that is, ambiguity is an intermediate state between ignorance (no distributions are ruled out) and risk (all distributions but one are ruled out) (Einhorn and Hogarth 1986, p. S229).

A significant contribution to the prevalence of, in particular, yield uncertainty may be provided by model uncertainty. Individual decision makers may doubt the validity of the model of the world they employ, so as to estimate the cost and benefit of alternative courses of action. Recently, Fuhrer (1987) and Van Witteloostuijn (1990a) introduced model uncertainty in the context of a maximizing framework of learning behavior and expectations formation.

In all, the association of the maximizing postulate with perfect knowledge is unjustified. Choice theorists have been able to incorporate yield risk and even fundamental yield uncertainty into the scope of maximizing decision making behavior. In particular, the popularity of (variants of) the "Expected Utility Hypothesis" in the economic and psychological literature speaks for itself.¹

4.4.4 Choice Theory and Complete Knowledge

The economic theory of information in connection with information scarcity and cost and benefit of gathering and processing information offers a key to introduce action uncertainty (concerning the composition of the set of action alternatives) into choice theory. The economic theory of information originates from Stigler (1961). Because individuals are not well-informed, in advance, about the attainable action alternatives, they possess the possibility to gather and process additional information, if potentially

¹ In fact, the introduction of informational deficiencies in maximizing models can be traced back to at least the nineteenth century. For example, Cournot's (1834) and Bertrand's (1883) early models of imperfect competition already include an element of informational imperfection.

obtainable, and to search for alternative courses of action. Action uncertainty is the predecessor of search behavior. Because neither gathering and processing information nor searching behavior are costless activities, individuals have to weigh the cost and benefit of those activities.

Stigler (1961) elaborates these insights in the context of household behavior by means of answering the following question: What is the optimal number of shops the consumer should visit before purchasing a good, if he or she is uncertain about the prevailing distribution of prices? Stigler offers a choice-theoretic evaluation of the cost and benefit of information gathering and processing. In the course of the 1960s and 1970s search theory became embedded in economics. Rothschild (1973), Lippman and McCall (1976) and Stiglitz (1989) offer surveys of microeconomic search models. Popular applications of search theory are, on the one hand, the introduction of search behavior in labor market models in order to explain macroeconomic unemployment (for example, Alchian 1969 and Phelps 1970), and, on the other hand, the incorporation of search behavior in commodity market models in order to explain the persistence of price distributions (for instance, Diamond 1971).

Moreover, in the mid-1970s Darby (1976) and Feige and Pearce (1976) introduced Stigler's notions in the context of models of expectations formation. They incorporate cost of gathering and processing information in a rational expectations framework.¹ Feige and Pearce, for example, model expectations which

are formed by balancing the predictive benefits derivable from alternative information sets against the costs of acquiring and utilizing alternative information sets (Feige and Pearce 1976, p. 519).²

Search theory is a sequel to information and action uncertainty. Therefore, it reflects a clear example of the introduction of incomplete knowledge as regards to the composition of the set of action alternatives in a choice-theoretic framework.³

4.4.5 Choice Theory and Routines and Computational Disabilities

A major result of the behaviorist framework is the introduction of routinized behavior. As long as aspiration levels are satisfied, routinized behavior is continued. However, routines

¹ The rational expectations debate centers upon the relevance of Muth's (1961) "rational expectations hypothesis" that "expectations, since they are informed predictions of future events, are essentially the same as the prediction of the relevant theory" (Muth 1961, p. 316). Important advocates of this hypothesis are Lucas (for instance, 1972 and 1975) and Sargent and Wallace (1975 and 1976).

² However, Feige and Pearce (1976) beg the question by introducing incomplete information in a rational expectations framework by assuming complete information with respect to the cost and benefit of gathering and processing information. Although search theory enfeebles the critique of, for instance, Nelson *et al.* (1976) that "[t]he prototypical model in orthodoxy is one of full equilibrium under conditions of perfect and costless information" (Nelson *et al.* 1976, p. 90), now the problem of infinite regress occurs. In this context, Winter (1975) argues that "[t]he latter alternative - the optimization whose scope covers all consideration including its own costs - sounds like it may involve the logical difficulties of self-reference" (Winter 1975, p. 83). The infinite regress problem also shows its relevance in the context of endogenizing the ends in a maximizing framework (for example, Winter 1964 and Kirzner 1982). However, behaviorism faces the same difficulties. Conlisk (1988) argues that decision makers avoid the infinite regress problem either through quick convergence of the process or by making a shortcut judgement at some rather low level.

³ Many other examples can be found in the literature on industrial organization (Part III).

need not to be in conflict with maximizing decision making (Chapter 3). In effect, in the case of, for example, considerable uncertainty or high cost of search behavior, routines reflect the maximizing strategy. Four examples of the introduction of routinized behavior in a maximizing framework are worth mentioning.

First, decision analysts argue that an important alternative course of action consists in postponing the ultimate decision to learn along the way through time by exploiting

the gradual, time-dependent unfolding of uncertainties (Keeney and Raiffa 1976, p. 11).

In this way, routinized behavior is introduced in a choice-theoretic framework. Second, Heiner (1983 and 1986) offers a formal model in which routinized behavior is the result of maximizing decision making in an uncertain environment. Third, Houthakker and Taylor (1966) show, in the context of demand theory, that "habit formation" may arise from the decision making behavior of maximizing consumers. Fourth, Akerlof (1980) argues that routinized behavior in a labor market may reflect a maximizing strategy as a consequence of the existence of social norms and reputation considerations.

An important rationale underlying behaviorist theories is the prevalence of striking computational disabilities of human beings. It is argued that the application of the maximizing postulate requires computational ingenuity, because the decision maker has to possess a "complete description of the consequences following each alternative strategy" (Simon 1945, p. 69) and "choice sets precisely known and given" (Nelson and Winter 1980, p. 179). However, in the preceding sections it has been shown that the association of choice theory with complete knowledge and perfect foresight is unjustified. Moreover, choice theorists have analyzed the implications of computational limitations in a maximizing framework.

A recent example is Sah and Stiglitz' (1985, 1986 and 1988) analysis of the relationship between organizational structures and limited computational abilities of decision makers. Sah and Stiglitz' models show that individual participants - taking into account their limited communicational and informational capabilities - are able to maximize the payoff of organizational decision making by choosing the appropriate aspiration ("reservation") levels and organizational structure ("architecture").

4.5 THE FLEXIBILITY OF THE MAXIMIZING FRAMEWORK

Many modifications of the maximizing models took place both before and after Simon's (1945) critique. In general, the foregoing seems, therefore, to suggest that opponents of choice theory underestimate and misinterpret the range and flexibility of choice theory. The maximizing framework can absorb (the implications of) goal setting procedures, time variability of ends, group decision making, incomplete knowledge, uncertainty and computational disabilities. In this section the flexibility of choice-theoretic frameworks is illustrated with the help of a brief examination of two additional examples by repeating some of the arguments of Chapter 3.

First, maximizing models leave room for the introduction of all kinds of "noneconomic" phenomena by elaborating on the objective functions and restrictions. In this context, Becker is a controversial advocate of the imperialism of economics, an example being his maximizing model of marriage (1975). Akerlof (1980) introduces sociological determinants (i.e., social norms and reputation) of behavior in a maximizing framework. In cooperation with Dickens, Akerlof (1982) incorporates the psychological theory of cognitive dissonance (Festinger 1957) into a choice-theoretic model. However,

probably the most appealing example is the debate about altruism in economics. Several theorists have introduced altruism into the utility functions of individual maximizers (for instance, Collard 1978, Margolis 1982, Akerlof 1984 and Hirshleifer 1985).

Second, empirical research has produced a critique of choice theory by asserting that maximizing behavior is in conflict with the observations of actual decision making (see, for example, Schoemaker 1982). However, choice theorists have been able to elaborate maximizing frameworks such that many of the empirical violations are assimilated (Fishburn 1989). In this context, the "generalized expected utility analysis" (Machina 1982) and "regret theory" (Sugden 1986) are examples worth mentioning.

The flexibility of the maximizing postulate gave birth to the thesis that choice theory is, in fact, a powerless tautology. However, this thesis belies an error of reasoning. Applied *ex post* choice theory represents a tautology, because human behavior can always be explained in retrospect with the help of formulating the appropriate goals and restrictions. In fact, the same accusation can be directed at behaviorism. However, applied *ex ante* choice theory (and behaviorism) no longer reflect(s) a tautological argument, because then goals, utility functions and constraints have, in advance, to be adapted to concrete modeling (compare with Boland 1981 and Schoemaker 1982). The point is that *general* (unspecified) theories, in contrast to *specific* (specified) models, are untestable (for example, Klant 1972). However, this argument applies to both the maximizing and satisficing theory.

4.6 THE "AS IF" AND NATURAL SELECTION ARGUMENTS

The "as if" and natural selection arguments occupy a special place in the controversy that surrounds the maximizing postulate. Both arguments reflect a last line of defence of choice theorists. In particular, the criticism that maximizing requires an unrealistic assumption of human computational abilities can be enfeebled, on the one hand, by arguing that individuals behave "as if" they are guided by maximizing principles and, on the other hand, by putting forward a natural selection argument (Alchian 1950, Friedman 1953 and Machlup 1967). The natural selection argument is employed in the context of firms' behavior in particular. It says that

unless the behavior of businessmen in some way or other approximated behavior consistent with the maximization of returns, it seems unlikely that they would remain in business for long. Let the apparent immediate determinant of business behavior be anything at all - habitual reaction, random change, or whatnot. Whenever this determinant happens to lead to behavior consistent with the rational and informed maximization of returns, the business will prosper and acquire resources with which to expand; whenever it does not, the business will tend to lose resources and can be kept in existence only by the addition of resources from outside. The process of 'natural selection' thus helps to validate the (maximization of returns) hypothesis - or rather, given natural selection, acceptance of the hypothesis can be based largely on the judgement that it summarizes appropriately the condition for survival (Friedman 1953, p. 35).

The "as if" and natural selection arguments together lead to the proposition that individuals who behave more "as if" they apply a maximization rule have a greater

chance to survive the economic struggle and, hence, dominate in reality.¹

The "as if" and natural selection arguments win cogency by the results of the analyses of Nelson and Winter. In a number of articles and books Nelson and Winter (in particular, 1982) severely criticize neo-classical theory as being based upon complete knowledge, perfect foresight and equilibrium. Therefore, they suggest an alternative framework which captures behaviorist decision rules on the one hand and evolutionary theory on the other.² In an evolutionary environment, natural selection mechanisms operate in an economy consisting of satisficing firms. Firms are engaged in routinized behavior, unless trouble emerges. In the case of trouble, problem solving search activity is induced. These behavior patterns are adapted to concrete modeling with the help of *ad hoc* aspiration levels and stochastics.

Nelson and Winter's behaviorist model of firms' decision making in the context of an innovative struggle leads to natural selection:

Firm growth is linked to profitability ... The fact that growth is linked to profitability brings 'economic selection' into play as an influence on average productivity and on average policies toward innovation and imitation (Winter 1984, p. 288).

In combination with the introduction of firms' exit and entry movements, and in light of the "as if" and natural selection arguments, Nelson and Winter, in fact, *support* the hypothesis that firms which behave more "as if" they are guided by maximizing decision rules have a greater chance to survive the economic struggle, and dominate in reality. Firms with high aspiration levels appear to be superior searchers for profitable alternative courses of actions (in particular, innovations), and are, therefore, relatively flexible adaptors to the changing environment. Rigid firms with low aspiration levels, in contrast, adapt slowly to new circumstances and are, eventually, expelled from the market by better-equipped firms.³

Of course, it may be argued that Nelson and Winter's models bring to the surface the *a priori* postulated "as if" and natural selection arguments by showing a competitive process. However, Nelson and Winter's behaviorist model of a competitive struggle among heterogeneous decision makers can be represented in choice-theoretic terms as well. In the latter case, those agents equipped with superior knowledge or capabilities or

¹ Recently, the natural selection argument has been modified by arguing that the proposition that ("as if") *absolute* profit maximizers survive economic selection can be incorrect outside perfect competition, since strategic spillover effects (positive externalities) imply that ("as if") *relative* profit maximizers use the fittest strategy (Friedman and Rosenthal 1986 and Schaffer 1989). That is, "[w]hen firms have market power, the possibility of 'spiteful' behaviour exists: a firm may forego profit-maximisation and lower its profits and even its survival chances, but if the profits of its competitors are lowered still further, the 'spiteful' firm will be the more likely survivor" (Schaffer 1989, p. 44). So, it is maximization of *fitness* that matters, which indicates that firms in imperfectly competitive markets have to take strategic interdependencies into consideration. The key point is that a firm's (relative) profit maximization must proceed conditional upon (expectations of) the rivals' (re)actions, which resembles the familiar oligopoly argument (Part III).

² A detailed discussion of Nelson and Winter's Schumpeterian evolutionary theory is offered in Chapter 8.

³ A similar type of evolutionary argument is employed by Witt (1985 and 1986) to describe the natural selection of preferences (in particular, in connection with the debate about altruism and self-interest) and decision rules respectively. Strictly speaking, the specifics of the maximizing search strategy depend on environmental characteristics. For example, organizational ecology (or, to be precise, niche width theory) points to the role of the frequency and variance of environmental changes (Hannan and Freeman 1989). Here this complication is ignored. The key point is that the ("as if") maximizing strategy is defined relative to environmental restrictions.

less uncertainty survive the competitive struggle (Chapter 23). The assumption of nonmaximizing decision procedures is by no means a necessary condition for reaching a natural selection process. Instead, this position is reserved for the adoption of *heterogeneity* of whatever kind. The importance of heterogeneity is also stressed by choice theorists like Arrow (1978, p. 164) and Radner (1974). For example, the latter argues that

[w]e can extend the theory of the Arrow-Debreu economy to allow for differences in information among the [maximizing] economic agents (Radner 1974, p. 49).

Moreover, by allowing a shift of individuals' knowledge and information over time, and hence by introducing learning processes (Van Witteloostuijn 1990a), transitions over time of the individuals' decision behavior can be incorporated in a choice-theoretic framework.

Recent work on near rationally behavior (for example, Akerlof and Yellen 1985a and 1985b) may throw doubt on the argument that natural selection results in a dominance of those agents who behave "as if" they employ maximizing decision rules. Near rational behavior means

that agents have relatively wide latitude for deviating from full optimization without incurring significant losses (Akerlof and Yellen 1987, p. 138).

Hence, then near rationally behaving agents are able to survive the natural selection process. However, in that case near rational behavior reflects, in fact, a maximizing strategy! As McCallum (1987) puts it:

The point is that if the model used in the implicit definition of 'rational behavior' neglects some small computational or adjustment costs, then the agents' choices hypothesized by Akerlof and Yellen may in fact be *entirely* rational (McCallum 1987, p. 128; McCallum's *italics*).

In the context of the controversy of behaviorism and choice theory, it can, in addition, be argued that behaviorism is in need of an "as if" argument as well. Individual decision makers may, for example, rarely specify the values of their aspiration levels explicitly. Moreover, as in the case of maximizing, satisficing requires advanced calculations in order to determine the expected yields of the alternative courses of action under investigation.

4.7 ASKING THE RIGHT QUESTION

In the preceding sections it has been argued that the behaviorist critique of the maximizing postulate overshoots the mark. Contrary to the behaviorist suggestions, well-designed choice-theoretic models succeed in introducing goal setting procedures, time variability of ends, group decision making, uncertainty, routines and computational disabilities. In general, the foregoing seems to suggest that opponents of choice theory underestimate and misinterpret the range and flexibility of choice theory. The key point is that both the rationality and maximizing principle as descriptions of human decision making are *subjective* in nature. Rationality describes an economic *principle* rather than an economic motive (Hennipman 1945).

In effect, one may wonder whether behaviorism and choice theory do not reflect two

languages saying similar things; both enabling the incorporation of the same phenomena and both leading to identical results. However, it may still be argued that behaviorism deviates from choice theory because the diverging decision rules - satisficing and maximizing - lead to diverging compositions of the choice set of alternative courses of action. Given expectations and goals, maximizing decision making implies evaluating a choice set in order to locate the *best* alternative. Given expectations and aspiration levels, satisficing decision making leads to the adoption of the *first satisfactory* alternative evoked, an important one being to continue routinized behavior. However, given similar goals and expectations¹, maximizing and satisficing may lead to comparable results. In this respect, three remarks are in order.

First, at the level of the individual decision maker chance and the height of the aspiration levels influence the comparability of the results of maximizing and satisficing decision making. Irrespective of the height of the aspiration levels search for the first satisfactory alternative may by chance result in finding the maximizing course of action. Moreover, if the height of the aspiration levels is such that it exceeds or equals the - given information and expectations - maximally attainable results, then searching for the first satisfactory alternative leads to locating the best course of action (perhaps associated with a downward adjustment of the aspiration levels).

Second, the results of maximizing and satisficing decision making may converge after well-designed modifications of the models involved (see the preceding sections). In this case, it is a question of adopting the appropriate goals, restrictions, aspiration levels, information, etc.. This argument has already been put into words by Baumol and Quandt (1964), who argue that in view of costly computational resources satisficing and maximizing results can be identical.² Other examples are presented by Day (1967) and Van Helden (1972). For example, Day shows that under particular conditions choice-theoretic and behaviorist decision rules lead to the similar results. It can even be that in particular situations the restrictions dominate, so that neither satisficing nor maximizing plays a role in the decision making process (Frey and Foppa 1986 and Langlois 1986).

Third, at the level of a group of individuals, Alchian's, Friedman's and Machlup's "as if" and natural selection arguments point to the fact that individuals who decide more "as if" they are maximizing, or "as if" they possess higher aspiration levels, have a greater chance to survive the economic struggle and, therefore, dominate in reality. In this respect, the results of the analyses of Nelson and Winter are illustrative.

In the light of the natural selection argument, the following critique of the behaviorist approach occurs: why should decision makers only adjust behavior when trouble emerges and when it is perhaps too late; and why should decision makers stop searching for action alternatives as soon as a satisfactory alternative is found? The dominance of this kind of

¹ In particular, the adaptive (Cagan 1956) and rational (Muth 1961) expectations enjoy popularity in the context of choice-theoretic modeling.

² A recent elaboration on this argument is offered by Winston (1989). Winston points out that an "optimal" degree of rationality" (p. 69) can be calculated with the help of Becker's (1965) household production framework, which implies the modeling of a constrained utility maximizing decision rule. What matters, is the "allocation of rationality" (p. 70) as fully rational decision making is a costly and time-consuming exercise. This means that agents can decide to be imperfectly rational: "On the one hand, one never wants to put more energy or goods and services into being rational than is justified (on the intensity margin) by their effect on present and future satisfactions: too much rationality carries a cost in alternative use of its resources. On the other hand, one never wants to continue working at being rational about a choice - paying attention to it (on the duration margin) - when other activities give one's time a higher value: rationality carries a cost in terms of other things one could be doing" (Winston 1989, p. 73).

"crisis management" requires a sound rationale, which is difficult to provide in the context of a competitive struggle (see, for example, Weick 1979, pp. 246-247 and Koutsoyiannis 1979, p. 401). Moreover, the behaviorist satisficing principle introduces two additional *ad hoceries* which ask for an explanation: the height of the aspiration levels (and the extent of adaptive learning) and the sequence in which alternative courses of action are evaluated.¹

It seems justified to conclude that the assumed contradiction between the maximizing postulate and the satisficing principle rests on shaky grounds. It is a wrong question to ask whether human beings maximize or satisfice. Behaviorism and choice theory are two ways to model and analyze similar phenomena. Lea *et al.* argue on the basis of an extensive survey of economic psychology that

[f]or every kind of behavior we have considered, we have been able to find a description in terms of rationality, even in the face of apparently gross irrationality. For example, observed gambling, giving and saving behavior all involve substantial deviations from a rationality account, at least a superficial one, yet for each of them an 'economic' theory exists that identifies sources of utility which the observed behavior would maximize (Lea *et al.* 1987, p. 480).

The right question to ask is *what* (individual or group) decision makers maximize or satisfice (see Rachlin 1980 and Lea *et al.* 1987).

¹ An example of a theory of search sequence is Lindblom's (1965) "muddling through".

PART

III

INDUSTRIAL ORGANIZATION

Static and Dynamic Competition

KEY WORDS. *Industrial organization focuses on the study of competition. Part III reviews the literature on industrial organization that is relevant in the context of the static - dynamic efficiency tradeoff (Chapter 5). First, the static benchmark theories of competition (perfect competition, Bertrand models and perfect contestability) are evaluated (Chapter 6). Second, the theory of entry barriers is examined as barriers to entry introduce dynamic competition to the detriment of static efficiency (Chapter 7). Third, R&D models are reviewed, since innovation is an important aspect of dynamic competition (Chapter 8).*

1944

1944

1944

1944

1944

1944

1944

CHAPTER

5

THE DILEMMA: The Tradeoff Between Static and Dynamic Efficiency

In the centuries before the publication of Smith's *The Wealth of Nations* (1776), theorists had already philosophized about the concept of competition in a normative tradition (McNulty 1968). McNulty (1968) points out that classical economists emphasized that the price level is the result of the scale of competition as a regulatory force. The cost ("natural") prices are reached in an environment of free competition. Then, prices of commodities are such that the cost of offering is precisely covered. Classical economists do not associate *free* competition with a large number of independent, price-taking firms on *a priori* theoretical arguments. Neoclassical economists framed the reference point of *perfect* competition (Walras 1874). Ever since (at least) Schumpeter's (1943) seminal contribution there has been a debate going on about the optimality features of perfect competition (Hennipman 1954). The literature deals with the assumed *tradeoff* between *static* and *dynamic* efficiency of market behavior.¹ Static efficiency of market behavior refers to (minimum) average cost pricing, whereas dynamic efficiency is concerned with the fact that, for instance,

new products may be introduced, new qualities of existing products may be developed, new methods of production may be ventured, new forms of industrial organization, financing, or tackling risk may be developed (Kirzner 1985, p. 30).

Schumpeter (1943) argues that perfect competition undermines the firms' incentives to introduce dynamic economies of market behavior by pointing out that

the fundamental impulse that sets and keeps the capitalist engine in motion comes from the new consumers' goods, the new methods of production or transportation, the new markets, the new forms of industrial organization that capitalist enterprise creates. ... A system - any system, economic or other - that at every point of time fully utilizes its possibilities to the best advantage may yet in the long-run be inferior to a system that does so at no given point of time, because the latter's failure to do so may be a condition for the level or speed of long-run performance. ... But in capitalist reality, as distinguished from its textbook picture, it is not that kind of [price] competition which counts but the competition among firms from the new

¹ The efficiency terminology is often used in welfare theoretic arguments in the industrial organization field. It is, for example, predominant in Kamien and Schwartz' (1982) excellent survey of the economics of innovation. However, the reader must be aware of the fact that efficiency in industrial organization has a meaning which is narrower than the one in welfare economics.

technology, the new source of supply, the new type of organization (the largest-scale unit of control for instance) - competition which commands a decisive cost or quality advantage and which strikes not at the margins of the profits but at their foundations and their very lives. ... It is hardly necessary to point out that competition of the kind we now have in mind acts not only when in being but also when it is merely an ever-present threat. It disciplines before it attacks (Schumpeter 1943, pp. 83-85).

So, Schumpeter points to the essential distinctions of price *versus* nonprice (Stigler 1968) and actual *versus* potential (Bain 1956) competition.

Arguments like Schumpeter's imply that

[f]ew, if any, economists maintain that perfect competition efficiently allocates resources for technical advance (Kamien and Schwartz 1975, p. 2).

In effect, the literature suggests that there is a tradeoff between average cost pricing (static efficiency as being associated with perfect competition) and innovative activity (dynamic economies as being associated with imperfect competition). That is,

under monopoly, innovation occurs but at a slower pace than is socially optimal, whereas under perfect competition there is none at all. This of course leads to the consideration of the trade-off between perfect competition and its static efficiency properties and monopoly, which lacks static efficiency but allows for innovation (Kamien and Schwartz 1982, p. 191).

The main purpose of the chapters of Part III is to review the literature on industrial organization in light of the (assumed) tradeoff between static and dynamic efficiency of market behavior. Before presenting a brief outline of Part III, four remarks are in order.

First, the study of competition is nowadays the focus of a specialization in economics that is called *industrial organization* (or, alternatively, industrial economics).¹ The dominant school in industrial organization is the *structure-conduct-performance* (henceforth SCP) paradigm (for review expositions see Scherer 1980, Reid 1987 and Geroski 1988), which is

associated with the pioneering work of Edward S. Mason (1939, 1949). ... Put simply, the content of this approach is that exogenous basic conditions determine market structure and that there is a unidirectional flow of causality from market structure, through conduct, to performance (Reid 1987, p. 11).

Except for the fact that industrial economists no longer adhere to the belief that there is a unidirectional causality, the SCP-paradigm still constitutes a unifying framework. However, although much of the literature referred to below fits into the SCP paradigm, only a sidelong glance is taken at this framework.

Ironically, one of the greatest industrial economists, Stigler, argues that "there is no such subject as industrial organization. The courses taught under this heading have for their purpose the understanding of the structure and behavior of the industries (goods and service producers) of an economy. These courses deal with the size structure of firms (one or many, 'concentrated' or not), the causes (above all the economies of scale) of this size structure, the effects of concentration on competition, the effects of competition upon prices, investment, innovation, and so on. But this is precisely the content of economic theory - price or resource allocation theory, now often given the infelicitous name of microeconomics" (Stigler 1968, p. 1).

Second, a judgement of the performance of a market requires a welfare theoretic criterion. As Reid (1987, p. 11) points out, the benchmark that is usually employed in industrial organization is Pareto optimality. However, a welfare analysis abounds with pitfalls. Broadly speaking, the literature on industrial organization assumes that static efficiency is associated with average cost pricing [and so zero profit (Chapter 12)], whereas dynamic efficiency may follow from the maximization of an intertemporal welfare function. It is important to stress the observation that industrial economic models (by and large) view the static - dynamic efficiency tradeoff from a *partial* perspective, where the partial welfare function generally takes the form of the sum of the consumers' and producers' surplus (Tirole 1988, pp. 6-12).¹ Without further reference opinions of welfare performance in the chapters of Part III fit into this partial tradition. Particularly in Parts IX and X further remarks are made on this topic.

Third, there are many essential distinctions in use inside the field of industrial organization that must be kept in mind (among which the ones mentioned above; *i.e.*, price *versus* nonprice competition, static *versus* dynamic theory and actual *versus* potential rivalry). At this point two of these twins of concepts deserve mentioning: normative *versus* positive theory and theoretical *versus* empirical research. Although reference is made to some positive arguments and empirical research in passing², the main focus below is on *normative theory* (in the sense of a welfare theoretic reference point). The survey of the literature is concerned with the following question: *are there theories which describe a normative benchmark for competition in the sense that both static and dynamic efficiency are induced?*

Fourth, although below a considerable part of the literature is reviewed, space limitations imply that many contributions to industrial organization (and labor economics) are not covered. However, hopefully the sample of references is such that a reliable picture is presented of the current state of the art of industrial organization with respect to the question of the tradeoff between static and dynamic efficiency of market behavior. Here it is to be noted that in addition Part VIII pays some attention to static and dynamic efficiency in the labor market on the basis of some theories of *labor economics*.

The plan of Part III is the following. Chapter 6 describes the concepts of competition which yield static efficiency. (Long-run) perfect competition and contestable markets introduce (minimum) average (marginal) cost pricing to the detriment of incentives to introduce innovations. Chapter 7 examines the theories of competition that provide contrary results. The theories of imperfect competition and entry barriers are associated with prices in excess of (minimum) average cost on the one hand and incentives to carry out (sunk) investments on the other. Chapter 8 focuses on the literature on an important manifestation of dynamic economies of market behavior: innovative (R&D) activity in commodity markets.³

¹ For example, Schmalensee (1978) notes that "[t]he usual index of social welfare in such studies is the sum of consumers' surplus and producers' excess profits" (p. 319).

² This illustrates the fact that the analytical boundaries between positive and normative theory and theoretical and empirical research are difficult to trace in practice. It is all a question of emphasis.

³ It is clear that the focus of this part is on the *supply* side of the (commodity) market. The demand side is only introduced if it supports or complements the arguments that underlie the analysis of the decision making of suppliers and the associated market performance.

 

CHAPTER

6

STATIC EFFICIENCY: Price Competition

6.1 INTRODUCTION: (Minimum) Average Cost Pricing

Static efficiency is manifest in (minimum) average cost pricing. There are three *static* theories of *price* competition that are based upon assumptions of predominating rivalry such that the competitive pressure of rivals forces firms to adopt careful price taking or setting: (long-run) perfect competition, Bertrand competition (with product homogeneity and nonbinding capacities) and contestable market theory. The first two frameworks stipulate the predominance of *actual* competition, whereas the latter theory is based upon conditions of pervasive *potential* competition.

The well-known benchmark case for competition that guarantees static efficiency is *perfect competition*. The concept of perfect competition is embedded in economics not only through industrial organization, but also (perhaps even mainly) *via* general equilibrium theory.¹ Perfect competition is associated with forceful actual competition. In the long-run equilibrium state the force of actual competition is such that equilibrium prices prove to be equal to the minimum average and marginal cost of offering the commodities involved. The framework of perfect competition stands or falls with strict assumptions, particularly with respect to price taking and market structure.

The concepts of Bertrand competition and contestable markets seek to relax the strict assumptions of price taking and market structure that are associated with perfect competition. On the one hand, Bertrand competition still fits into the tradition which assumes that an ever-present force of actual competition operates as a disciplinary device. On the other hand, the framework of perfect contestability is based upon an assumption of the predominance of the threat of potential competition.

This chapter explains why the assumptions associated with the static theories of perfect competition, Bertrand competition and perfect contestability induce static efficiency in the sense that firms are forced to accept or set (minimum) average cost prices. The core of the chapter is organized as follows: Section 6.2 examines perfect and Bertrand competition; Section 6.3 introduces potential competition; Section 6.4 focuses on contestable markets; and Section 6.5 is concerned with the basic flaws of the theories involved. Section 6.6 contains concluding remarks. Since the theory of contestable markets is (i) a recent development and (ii) the point of departure of the barrier market

¹ Chapter 20 examines the interface of general equilibrium theory and models of (perfect and imperfect) competition.

concept, the examination of this framework takes more space than the discussion of either of the other theories. The fact of the matter is that the theories of perfect and Bertrand competition are only briefly outlined, because both are well-established in modern textbooks on microeconomics or industrial organization. A brief introduction to perfect competition and Bertrand rivalry is enclosed as both theories are a basis to contestability.

6.2 PERFECT AND BERTRAND COMPETITION: Actual Rivalry

Following the formal theory of, for example, Walras (1874) and the subsequent refinements of Clark (1915) and Knight (1921) the concept of perfect competition as a benchmark state of competition is founded in its final form, which is presented in modern (micro)economics textbooks.¹ Perfect competition is defined as a market structure with a large number of independent producers of some uniform product which take rather than set price. For example, in Varian's (1984) microeconomics textbook, the implications of perfect competition for price determination are described in the following way:

Figure 6.1
Perfect Competition

¹ Perfect competition can be viewed from both a general and partial perspective. This chapter (and part) focuses on partial theories, whereas general equilibrium theory (with or without perfect competition) is discussed in Chapter 20.

As long as there is perfect information about the prices being charged, and as long as the product really is identical across firms, it is clear that each firm that sells the product must sell it for the same price; for if any firm attempted to set its price at a level greater than the market price, it would immediately lose all of its customers. ... In a world with many such competing firms ... there can be only one price for output: the market price. Any firm that wishes to sell any output at all must do so at this price. Thus each firm must take the market price as a given, exogenous variable when it determines its supply decision (Varian 1984, p. 82).

So, the concept of perfect competition follows from assumptions of perfect price information¹, a large number of firms and a homogeneous product. That is, information is not perfect in every respect. For example, in the short run firms are not well-informed of technological progress.

The concept of perfect competition can be differentiated by distinguishing the short from the long-run case. In the short run (cost) heterogeneity of firms can be sustainable. In the long run exit and entry movements induce firms' homogeneity: only efficient (lowest cost) suppliers survive the competitive process. Without further reference this thesis deals with the long-run equilibrium position, since long-run perfect competition is associated with static efficiency (*i.e.*, partial maximization of total surplus as the deadweight loss is zero in a static context with constant U-shaped cost function): the market price is equal to the minimum average and marginal cost of offering the product involved. The result for an individual firm is depicted in Figure 6.1, where p denotes the market price, q the individual firm's output, ac the average cost and mc the marginal cost.

The efficiency feature of perfect competition follows from the predominance of actual rivalry among incumbent firms in a market that is implied by the assumed market structure and perfect information about the prices being charged. Bertrand competition circumvents the assumption of forceful rivalry in a different way in an environment in which firms set rather than take prices.

The elementary Bertrand (1883) model of competition is a reaction to Cournot's assumption that quantity is the decisive competitive variable.² The elementary Bertrand model is based upon *price* as the strategic weapon of rivals. Apart from cost and product homogeneity, it is assumed that each rival can supply the entire market at the minimum average (marginal) cost price.³ A firm i chooses a price level so as to maximize profits (π), conditional upon the price set by the rival(s) j ,

¹ Some equivalent assumptions will do the job as well. Particularly an assumption of a complete set of future markets enjoys popularity in general equilibrium theory (Arrow and Hahn 1971).

² Cournot's (1838) model is generally not associated with static efficiency.

³ Waterson (1984) points out that "Edgeworth (1925) developed an alternative to this model where the firms faced capacity constraints. In his model, as a consequence, no equilibrium need result" (Waterson 1984, p. 25). Here it suffices to describe the elementary Bertrand theory. Other types of Bertrand models are those with binding capacity constraints (Kreps and Scheinkman 1983, Davidson and Deneckere 1986 and Osborne and Pitchik 1986) and product heterogeneity (Dixon 1988). It is important to note that these alternative Bertrand models are not necessarily associated with static efficient market performance. Broadly speaking, Bertrand frameworks are part of the class of oligopoly models (Dixon 1988 and Shapiro 1989a), which are generally framed in terms of game theory (by modeling reaction functions).

$$(6.1) \quad \text{Max}_{i,p} \pi = p \cdot d(p | p) - i \text{ac}(q),$$

where, provided that the products are uniform, it is assumed that

$$(6.2) \quad \begin{cases} d(p) = 0 & \text{if } p > p, \\ d(p) = \tau \cdot D(p) & \text{if } p = p, \\ d(p) = D(p) & \text{if } p < p, \text{ and} \end{cases}$$

p represents the price set by a firm, d denotes the quantity demanded from an individual firm, $\text{ac}(\cdot)$ indicates the average cost function, D is market demand and τ describes the share of market demand that flows to firm i if all rivals j set prices at the same level. For example, in the symmetric duopoly case: [$\text{ac}(\cdot) = \text{ac}(\cdot)$, $i, j = 1, 2$, $i \neq j$] gives $\tau = 0.5$.

The essential feature of the Bertrand model is that a firm's individual demand function is *discontinuous*. It is assumed that buyers switch to the firm(s) which offer(s) the lowest price before rivals are able to respond. So, provided that no collusion (is allowed to) occur(s), in order to avoid the negative profits that follow from zero demand, a firm is inclined to respond to a rival's price setting by slightly undercutting the rival's price level, subject to a nonnegative profit constraint. Hence, a firm's reaction curve is

$$(6.3) \quad p = r(p) = p - \mu,$$

subject to $\pi \geq 0$ and $\mu > 0$ but very (infinitely) small. It is now easily shown that equilibrium is reached when $p = p = \text{ac}(q) = \text{ac}(q)$ if it is assumed that $\text{ac}(\cdot) = \text{ac}(\cdot)$.¹ Hence, the equilibrium situation is characterized by average cost pricing and zero profits.

So, provided that it is assumed that the individual demand functions are discontinuous, products and firms are homogeneous, collusion does not occur, capacity constraints are nonbinding, buyers switch to the lowest priced firm before rivals can respond, and price is the competitive variable, Bertrand competition is associated with static efficiency. The forceful rivalry of actual competitors inclines firms to adopt average cost pricing. That is to say, average cost pricing follows *irrespective* of the prevailing market structure, since the fear of undercutting by *one* rival is sufficient to enforce zero profits. This result is well-established in the (textbook) literature. Waterson (1984), for example, introduces the Bertrand model in the following way:

Bertrand (1883) asserted that even in duopoly, price would fall to the competitive level. The reason is that he assumes each firm chooses a price rather than an output level. Each also, crucially, has enough capacity to fulfil the entire market demand at a price equal to marginal cost (assumed the same for both, for simplicity). Because products are identical, the firm with the lowest price makes all the sales. Now if firm 1, say, charges a price p_1 greater than marginal cost, firm 2 would undercut him slightly to make all sales. In turn, firm 1 would

¹ If $\text{ac}(\cdot) \neq \text{ac}(\cdot)$, then the firm offering the lowest price survives at the expense of the higher cost rival (Chapter 23), provided that both firms face nonbinding capacity restrictions. It has, in addition, to be assumed that firms deploy Cournot conjectures: $dp/dp = 0$. That is, firms expect that rivals will not react to its price cut by lowering their prices.

undercut firm 2 slightly to regain the market. The only equilibrium would be at price equal to marginal cost, where no firm has an incentive either to lower or to raise prices, but where no profits are made (Waterson 1984, p. 25).¹

In Section 6.4 the arguments underlying this result are examined in detail.

6.3 POTENTIAL COMPETITION

The importance of forces of potential competition is well-known in the economic literature (Gilbert 1989a). Hause and Du Rietz (1984) point out that the recognition of the existence and implications of (the threat of) entry is at least as old as 1776. The following quotation of Smith (1776) illustrates this point:

When by an increase in effectual demand, the market price of some particular commodity happens to rise a good deal above the natural price, those who employ their stocks in supplying that market are generally careful to conceal this change. If it was commonly known, their great profit would tempt so many new rivals to employ their stocks in the same way, that, the effectual demand being fully supplied, the market price would soon be reduced to the natural price, and perhaps for some time even below it. If the market is at great distance from the residence of those who supply it, it may sometimes be able to keep the secret for several years together, and may so long enjoy their extraordinary profits without any new rivals. Secrets of this kind, however, it must be acknowledged, can seldom be long kept; and the extraordinary profit can last very little longer than they are kept (Smith 1776, p. 60).

Moreover, it is interesting to note that Clark (1915), one of the founding fathers of the theory of perfect competition, and Robinson (1933), a pioneer of the theory of imperfect competition, point to the existence of potential competition and barriers to entry:

And the thing that holds it [the monopoly firm] firmly in check is *potential competition*. The fact that a rival *can* appear and *will* appear if the price goes above the reasonable level at which it stands, induces the corporation to produce goods enough to keep that price level. Under such a nearly ideal condition the public good would get the full benefit of the economy which very large production gives, notwithstanding that no actual competition would go on. Prices would still hover near the low level of cost (Clark 1915, pp. 380-381; Clark's italics);

and

[a] demand curve of this form might be found where a monopolist is subject to potential competition. A monopolist may have some advantage over his rivals, whose costs are higher than his, but may be aware that if he raises his price beyond a certain critical level, his rivals will find it profitable to produce and will begin to invade his market (Robinson 1933, p. 81).

So, it is not surprising that Schumpeter (1943) argues that

[i]t is hardly necessary to point out that competition of the kind we now have in mind acts not only when in being but also when it is merely an ever-present threat. It disciplines before it

¹ Note that a Bertrand equilibrium exists only in exceptional cases (where the minimum average cost price clears the market) if average cost curves are U-shaped. This is not the case with linear cost functions (or a U-shaped curve with a flat bottom).

attacks (Schumpeter 1943, p. 85).

Modern advocates of the importance of potential competition are Chicago economists. The thesis of the Chicago school is that two firms may be sufficient to guarantee optimality of market behavior, as long as potential competition rules the market (for example, Demsetz 1968 and Brozen 1969). Brozen points out that

Professor Harold Demsetz has argued that a large number of producers is not a necessary condition for competition, although more than one bidder may be required. ... Open entry is, it seems to me, a necessary condition if a competitive result is to prevail in a market. With the abundant supply of entrepreneurs in the economy, I am willing to say that it is also a sufficient condition (Brozen 1969, p. 8).¹

The theory of (perfectly) contestable markets seeks to offer a formal microfoundation of the thesis of the Chicago school, for particular competitive circumstances.² Since the contestable market theory is one of the points of departure of the barrier market framework, it is justified to spend considerable space on examining the assumptions and implications of Baumol *et al.*'s theory.

6.4 (PERFECT) CONTESTABILITY: Potential Rivalry

6.4.1 Introduction: Intuition

The current pillar of the assumption of Bertrand competition, with *potential* rather than actual rivalry, is perfect contestability.³ In 1982 Baumol heralded the framework of contestable markets in his presidential address as an uprising in the theory of industry structure. The contestable market notion, as elaborated in the seminal contribution to industrial organization of Baumol *et al.* (1982a), has received considerable attention since then; it even showed up in the court room (Baumol and Willig 1986).

Baumol (1982) introduces the concept of perfect contestability in the following way:

The address of the departing president is no place for modesty. Nevertheless, I must resist the temptation to describe the analysis I will report here as anything like a revolution. Perhaps such terms as 'rebellion' or 'uprising' are rather more apt. But, nevertheless, I shall seek to convince that the work my colleagues, John Panzar and Robert Willig, and I have carried out

¹ The importance of potential competition as such as a determinant of market behavior has only rarely been the subject of empirical research, some exceptions being Hause and Du Rietz (1984), Van Herck (1984) and Gilbert (1989a). For example, Hause and Rietz conclude that the "results certainly suggest responsiveness of new firm entrants to changes in costs and returns, which provides some evidence supporting the belief in the importance of potential competition. There is strong statistical evidence that new firms entry is positively related to industry growth" (Hause and Du Rietz 1984, p. 752). In subsequent sections and chapters more will be said about the credibility of the entry threat.

² Other examples of formal free entry models are Fama and Laffer (1972) and Novshek and Sonnenschein (1987). Moreover, the assumption of free entry (intersectoral competition) is essential to modern classical economics (Semmler 1984).

³ Baumol *et al.* (1982a, Chapter 1) admit that the contestability framework fits into an established tradition. In particular, as Schwartz (1986) points out, the "basic idea that threat of entry may constrain pricing in concentrated industries has long been recognized. The voluminous 'limit pricing' literature dating at least to Bain (1949) makes precisely this point" (Schwartz 1986, p. 38). However, in contrast to the limit pricing literature, perfect contestability is associated with zero profits. See also Chapter 7.

and encapsulated in our new book enables us to look at industry structure and behavior in a way that is novel in a number of respects, that it provides a unifying analytical structure to the subject area, and that it offers useful insights for empirical work and for the formulation of policy (Baumol 1982, p. 1).¹

Baumol defines a

contestable market [as] one into which entry is absolutely free, and exit is *absolutely costless* (Baumol 1982, p. 3; Baumol's italics).

So, the essential feature of a perfectly contestable market is the *absence of entry and exit barriers*.² A contestable market is vulnerable to "hit-and-run" entry. That is,

[e]ven a very transient opportunity need not be neglected by a potential entrant, for he can go in, and, before prices change, collect his gains and then depart without cost, should the climate grow hostile (Baumol 1982, p. 4).

Schwartz (1986) nicely summarizes the implications of a pervasive entry threat. He argues that

the threat of new entry may be sufficient to discipline incumbent firms. In the extreme, benchmark case of perfect contestability, threat of entry ensures satisfactory performance regardless of the size distribution of incumbent firms and regardless of any oligopolistic interactions among them. ... To prevent costless hit-and-run entry incumbents must set price where average cost intersects market demand, which maximizes welfare subject to a breakeven constraint (Schwartz 1986, p. 37-38).

The next subsections examine where these results stem from. The focus is on the single product case. Subsection 6.4.2 describes the essential assumptions underlying perfect contestability; Subsection 6.4.3 briefly indicates crucial definitions; and Subsection 6.4.4 presents a formal proof of the major implications of perfect contestability for market performance and market structure. A review of the critique of the contestable market theory is postponed to Section 6.5.

6.4.2 Essential Assumptions

To clarify matters, the assumptions (some of which are tacit) of contestability theory are summarized. The contestable market theory is based upon (at least) eight assumptions. The first one runs as follows.

ASSUMPTION 6.1 (Free entry). *Entry barriers are absent so that entry is free. The*

¹ In effect, the contestability framework offers two, largely independent, contributions to industrial organization: a normative theory and a analytical apparatus for the study of multiproduct firms. Here the focus is mainly on the former aspect. Hence, the perspective in this thesis resembles Shepherd's (1984) attitude: "I try here only to appraise the analysis of ultra-free entry, as a normative contribution to industrial organization. It is largely separable from the multiproduct and sustainability analysis, which I leave for other writers and occasions" (Shepherd 1984, p. 572).

² An examination of entry barriers follows in Chapter 7. Here it suffices to note that free entry implies the absence of entry barriers.

(potential) entrant "suffers no disadvantage in terms of production technique or perceived product quality relative to the incumbent (Baumol 1982, p. 4).

Not only is entry free, but besides exit is costless. In effect, the literature suggests that Assumption 6.2 is necessary in order to sustain Assumption 6.1.

ASSUMPTION 6.2 (*costless exit*). Exit barriers are absent. Exit is costless at any exit date.

The fact of the matter is that the literature suggests that sunk costs are generally incompatible with free entry (contestability). This proposition is defended by, for example, Baumol and Willig (1981), Grossman (1981), Baumol *et al.* (1982a), Stiglitz (1987), Dasgupta and Stiglitz (1988a) and Martin (1989). To clarify matters, it is worthwhile to quoting a lengthy passage of Baumol and Willig (1981).

Sunk costs to some degree share with barriers to entry the ability to impede the establishment of new firms. The need to sink money into a new enterprise, whether into physical capital, advertising, or anything else imposes a difference between the incremental cost and the *incremental risk* that are faced by an entrant and an incumbent. The latter's funds are already committed and are already exposed to whatever perils participation in the industry entails. On the other hand, a new firm must take the corresponding amount of liquid capital and turn it into a frozen asset if it enters the business. Thus, the incremental cost, as seen by entrants, includes the full amount of the sunk costs, which is a bygone to the incumbent. ... The risk of losing unrecoverable entry costs, as perceived by a potential entrant, can be increased by the threat (or the imagined threat) of retaliatory strategic or tactical responses of the incumbent (Baumol and Willig 1981, p. 418; Baumol and Willig's italics).

The assumption underlying this result is nicely reflected in Grossman's (1981) argument:

If K_0 is the *ex ante* cost of the plant and K_1 is the *ex post* value of the plant (*i.e.*, what it can be sold for), then $K_0 - K_1$ represents the amount of the sunk cost (p. 1165).

Therefore,

our model suggests that it is the sunk costs rather than fixed costs which may deter entry. This is because if one firm 'sinks' some costs, then its (*ex post*) average cost of production is lower than firms which *have not yet sunk costs* (Grossman 1981, pp. 1170-1171; italics added).

The essential assumption now is that potential entrants have *not yet sunk the costs* necessary to eliminate (capacity, technology, goodwill or whatever) disadvantages relative to the incumbent firms, which indicates noncontestability. A further discussion of the role of sunk and exit cost follows in Subsection 6.4.3 and Chapter 10.

Two further assumptions are necessary so as to make the fear of "hit-and-run" entry effective.

ASSUMPTION 6.3 (*Bertrand-Nash expectations*). (Some or all) potential entrants have Bertrand-Nash expectations. They "find it appropriate to evaluate the profitability of entry in terms of incumbent firms' pre-entry prices (Baumol 1982, p. 4).

Assumption 6.4 resembles one of the conditions underlying Bertrand competition.

ASSUMPTION 6.4 (*Demand contestability*). *Buyers switch to supplier(s) which offer(s) the lowest price (or a preferred product) before the higher priced firm is able to respond or retaliate.*

Baumol *et al.* (1982a) defend Assumption 6.3 (and 6.4) by arguing that

[p]otential entrants may not fear such retaliation for one of several reasons: First, because entry and exit are so inexpensive that a profit opportunity, even one expected to be very short in duration, will attract new competitors; second, because incumbents are restrained by law or other impediments from undertaking retaliatory moves; or, third, because some potential entrants have 'Bertrand-Nash' expectations, believing that entry will not lead to price responses by incumbents (Baumol *et al.* 1982a, p. 350).

Fourth, as Schwartz (1986) adds, because long term contracts can protect the entrant from strategic responses by an incumbent. Referring to this fourth argument, Brock (1983) clarifies matters by asserting that

Dixit (1982), for example, has argued that the economic conditions that must be present for perfect contestability to exist ... are extremely stringent: (i) all producers must have access to the same technology, (ii) this technology may have scale economies such as fixed costs, but must not involve any sunk cost, (iii) incumbents can change prices only with a nonzero time lag, and (iv) consumers must respond to price differences with a shorter lag. Baumol *et al.* (1982b) argue that iii and iv are not needed if entrants can write firm contracts with consumers for delivery over some fixed period length t (Brock 1983, p. 1057).

Section 6.5 and Chapter 11 offer further discussion of the importance of the assumptions of entry and exit lags.

Four further assumptions are worthy of mention. First, it is supposed

that there is always an entrant around (Grossman 1981, p. 1159).

In the extreme it is even assumed

that there is no bound to the number of firms with access to a particular cost function (Novshek and Sonnenschein 1987, p. 1290).

ASSUMPTION 6.5 (*potential entrants*). *Incumbent firms reckon that there are potential entrants around which constitute a credible entry threat.*

Second, in the literature it is of course assumed that potential entrants capture zero profits in the incumbents' market if entry is abandoned.

ASSUMPTION 6.6 (*zero profits*). *Incumbent firms fear that there are credible potential entrants which capture a zero profit in the incumbents' market if they refrain from entry.*

In general, this assumption is rendered concrete by arguing that potential entrants are not engaged in production if they refrain from entry (see, for instance, Grossman 1981, Stiglitz 1987, and Maskin and Tirole 1988). As Novshek and Sonnenschein (1987) argue,

[i]n equilibrium no firm makes negative profit, because exit is possible and yields zero profit

(Novshek and Sonnenschein 1987, p. 1291).

Third,

if a firm is indifferent between producing and not producing, then it produces if this doesn't cause excess supply (Grossman 1981, p. 1159).

ASSUMPTION 6.7 (supply preference). *If producing and selling are associated with zero profits, then firms decide to produce rather than become or remain passive.*

Finally, it is supposed that incumbent firms have a first-mover advantage.¹

ASSUMPTION 6.8 (incumbents' first-mover advantage). *Incumbent firms announce prices before entrants. Successful entry is only possible if entrants offer prices below the incumbents' level.*

Assumption 6.8 implies that entry with $e^p = i^p$ is not worthwhile [where superscript e and i denote (potential) entrant and incumbent firm respectively]. This may be the case for two reasons. First, since buyers have no preference for either equally equipped supplier, it is likely that they will follow their routine buying behavior rather than switch to the entrant. Second, even if some buyers decide to go to the entrant, it is unlikely that all the incumbent's customers will do so. Hence, the quantity demanded from the entrant is insufficient to permit the recovery of cost against the original price ($e^d < e^q$). From assumption 6.6 it follows that an entrant can always do better elsewhere.

6.4.3 Crucial Definitions

A major innovation associated with the contestability theory is *sustainability* analysis. In the context of a sustainability analysis of market configurations a number of definitions are worth clarifying (Baumol *et al.* 1982a, pp. 24-25).

DEFINITION 6.1 (feasibility). *An industry configuration of n_i (= number of incumbent firms) firms is feasible if $\sum_{i=1}^{n_i} i^q = D(p)$ and $p \cdot i^q - C(i^q) \geq 0$ for $i = 1, \dots, n_i$.*

A market configuration is characterized by the number (n_i) and size (i^q) of incumbent firms. A market configuration is feasible where market demand is precisely met without any firm facing losses.

DEFINITION 6.2 (sustainability). *A feasible industry configuration is sustainable if $e^p \cdot e^q \leq ac(e^q) \cdot e^q$ for all $e^p \leq p$ and $e^q \leq D(e^p)$.*

Hence, a feasible industry configuration is sustainable if no entrant is able to enter profitably. It easily follows that

¹ Perry (1984) points out that (single product) firms quote a uniform price. Otherwise, sustainable positive profits can arise.

[i]n a perfectly contestable market with profit-seeking firms, only a sustainable configuration can constitute an equilibrium (Baumol *et al.* 1982a, p. 25).

The reason for this result is that unsustainability induces entry and/or exit movements (Subsection 6.4.4).

A further clarifying distinction is the one between *temporary* sustainability at a point in time and *intertemporal* sustainability over time (Baumol 1982 and Baumol *et al.* 1982a).

DEFINITION 6.3 (temporary sustainability). A market configuration shows temporary sustainability if market demand at a point in time permits a sustainable configuration.

DEFINITION 6.4 (intertemporal sustainability). A market configuration shows intertemporal sustainability if neither entry nor exit occurs as a result of changes over time.

The questions of temporary and intertemporal (un)sustainability are examined in Subsection 6.4.4 and Part VI. Of course, intertemporal sustainability is a sufficient (but not a necessary) condition for temporary sustainability.

There is a further concept left that must be introduced: *natural* monopoly. The implications of an assumption of perfect contestability are different for (natural) monopoly.

DEFINITION 6.5 (natural monopoly). An industry is said to be a natural monopoly if, over the entire range of outputs, the firm's cost function is associated with nondecreasing returns.¹

Definition 6.5 implies that $\partial ac/\partial q \leq 0$ for all q so that $n_i = 1$ irrespective of the scale of demand. A demand monopoly occurs if an incumbent firm with nondecreasing returns up to q^R can produce $q = D(p) \leq q^R$. A demand monopoly means that $\partial ac/\partial q \leq 0$ for $q \leq q^R$ but $\partial ac/\partial q > 0$ for $q > q^R$.

One concept is at the heart of the contestability theory: sunk cost. The difference between fixed and sunk cost is crucial for understanding the essentials of perfect contestability. Therefore, it is worthwhile quoting Baumol *et al.* (1982a) extensively:

As usual, taking 'the long run' to refer to a period of time sufficient for all current commitments to be liquidated, it follows that in the long run all sunk costs are zero. Long-run fixed costs are those costs that are not reduced, even in the long run, by decreases in output so long as production is not discontinued altogether. But they can be eliminated in the long run by total cessation of production. ... Sunk costs, on the other hand, are costs that (in some short or intermediate run) cannot be eliminated, even by total cessation of production. As such, once committed, sunk costs are no longer a portion of the opportunity cost of production. ... It should be emphasized that here fixed costs mean costs that are fixed in the long run as well as the short run. Thus, investments in large-scale plant and equipment do not generally qualify. For, as the textbook aphorism says, such costs do indeed become variable

¹ Strictly speaking, this definition departs from Baumol *et al.*'s (1982a), since the reference to the subadditivity of cost is missing. In this study Definition 6.5 can, however, do the job, since the focus is on the single product firm.

eventually. They are *sunk costs*, as distinguished from fixed costs that can only be diminished, even in the long run, by closing the enterprise down altogether (Baumol *et al.* 1982a, pp. 280-281; Baumol *et al.* 's italics).

Broadly speaking, sunk cost follows from *irreversible investment in specific capital* (Caves and Porter 1977 and Martin 1989).

The key assumption of perfect contestability is that exit cost is zero irrespective of the date of exit (Assumption 6.2). The literature usually identifies this assumption with a condition of zero sunk cost. That is,

[e]xit is perfectly free, at no sacrifice of any cost. Sunk cost is zero (Shepherd 1984, p. 572).

This means that Baumol *et al.*'s (1982a, pp. 280-281) intuition of the meaning of sunk cost has to be highlighted (Martin 1989 and Gilbert 1989b). To be precise,

suppose that a unit of capital purchased at a price of β per unit could be sold or utilized elsewhere ... for a unit salvage value of $\alpha \leq \beta$. Thus it is possible to parametrize continuously the degree of sunkeness of capital from zero ($\alpha = \beta$) to absolute sunkeness ($\alpha = 0$) (Baumol *et al.* 1983, p. 494).

DEFINITION 6.6 (*sunk costs*). *Sunk costs are costs that (in some short or intermediate run) cannot be eliminated, even by closing down the enterprise altogether: sunk cost follows from the part of capital which is unrecoverable (that is, which has zero selling value) outside the market in which the products that are produced with the capital involved are sold.*

So, perfect contestability can be assumed to be compatible with investment. If the not yet amortized investment cost can be recouped at any exit date, exit cost is zero. Perfect contestability assumes that firms are always able to benefit from cost effective selling or alternative use of the capital outside the market: exit cost is zero at any exit date.

However, the identification of sunk cost as exit cost can miss the point. Sunk cost only introduces a positive exit cost if exit occurs before the sunk investment is fully amortized¹, since the sunk investment is not recoupable outside the market.

DEFINITION 6.7 (*exit costs*). *Exit costs are the sunk costs which have not yet been amortized (that is, recovered through selling the products that are produced with the sunk capital involved) in the market at the date of exit.*

At this point of the argument it suffices, first, to associate perfect contestability with free entry and costless exit and, second, to recognize the literature's usual practice of regarding zero sunk cost, implying zero exit cost, as an essential contestability assumption. Here it is telling to observe that Baumol *et al.* (1982a) argue that

we may conclude that where the productive techniques and market demands available to incumbents are also freely available to potential entrants, markets must be perfectly contestable if entry is costlessly reversible. ... Clearly, when entry requires the sinking of substantial costs, it will not be reversible because, by definition, the sunk costs are not recoverable [i.e.,

¹ Chapter 14 deals with the assumptions of amortization.

recoupable outside the market]. However, if efficient operation requires no sunk outlays, then entry can, by and large, be presumed to be reversible, and the market can be presumed to be contestable (Baumol *et al.* 1982a, p. 7).

So, for the moment the case where a *positive* sunk cost is associated with a *zero* exit cost is ignored. Chapter 10 goes on to explain the role of sunk and exit cost from the perspective of the barrier market.

6.4.4 Implications for Market Performance and Market Structure

The implications of perfect contestability for market performance and market structure follow directly from the assumptions underlying the theory. First, entry deterrence requires that incumbent firms make *static optimal* price and quantity decisions; second, provided that auxiliary assumptions are introduced, the (set of) contestable *market structure(s)* is *endogenous* rather than *exogenous* to the theory. Subsequently, both results are briefly examined.

Brock (1983) and Schwartz and Reynolds (1983) point out that Baumol *et al.* (1982a) do not precisely define the game played in a contestable market. Below a typical price game in a perfectly contestable market is described. In a contestable market incumbents and entrants play a Bertrand-Nash game on price (Knieps and Vogelsang 1982, Schwartz and Reynolds 1983 and Brock 1983; Assumption 6.3). Besides, incumbents have a first-mover advantage in the sense that entrants are only able to outperform incumbents if $e_p < i_p$ (Assumption 6.8). In order to illustrate the incumbents' first-mover advantage, a three-staged structure is described. In t^1 incumbents announce prices. The potential entrants' responses follow in t^2 . The game is closed in t^3 , where transactions take place. (So, formally the contestable market theory is associated with a two-staged game.) Contestability is based upon a Bertrand-Nash assumption, which implies that incumbent firms are unable to respond in t^3 . The game is one which has complete and perfect information.

Definition 6.1 and 6.2 together imply that a feasible and sustainable market configuration is associated with a zero-profit condition. A zero-profit condition implies that

$$(6.4) \quad i_p = ac(i_q), \text{ for } i = 1, \dots, n_i.$$

An incumbent firm has two sets of relevant strategies (i_s): $i_p > ac(i_q)$ (i_s^1) and $i_p = ac(i_q)$ (i_s^2). Actions showing $i_p < ac(i_q)$ are dominated, because they yield a negative payoff. The two relevant regions of the set of reply options of a potential entrant (e_s) are $e_p < i_p$ (e_s^1) and non-entry (e_s^2). The entrants' strategy to offer $e_p \geq i_p$ is not worthwhile, because then they are unable to enter successfully.

A firm's payoff (π) is determined by

$$(6.5) \quad \pi = p \cdot d - ac(q) \cdot q,$$

When entry deterrence is successful: $i_d = i_q$ (Assumption 6.4). If a potential entrant succeeds in outperforming an incumbent firm, then $e_d = e_q$ and $i_d = 0$. Non-entry is associated with a zero profit in the incumbents' market: $e_\pi = 0$ (Assumption 6.6). Hence,

$$(6.6) \quad \left\{ \begin{array}{l} i, e \pi > 0 \text{ if } i, e p > ac(i, e q) \text{ and } i, e d = i, e q, \\ i, e \pi = 0 \text{ if } i, e p = ac(i, e q) \text{ and } i, e d = i, e q, \\ i, e \pi = -ac(i, e q) \cdot i, e q < 0 \text{ if } i, e d = 0, \text{ and } \\ e \pi = 0 \text{ (non-entry).} \end{array} \right\} \begin{array}{l} \text{(entry deterrence)} \\ \text{or successful entry)} \\ \text{(exit or unsuccessful entry)} \end{array}$$

The payoff matrix (6.6) is associated with the assumption that firms produce products *before* they announce price. This means that the negative payoff involving exit is $-z \cdot ac(q) \cdot q$, with $z = 1$. However, a negative payoff of exit holds for any $z > 0$. Hence, the critical assumption is that firms produce *some* supply in advance. Alternatively, Assumption 6.7 can be adopted.

The payoff matrix (6.6) points to an inconsistency in contestability theory. Strictly speaking, advance production introduces a positive sunk exit cost, which contradicts contestability. Advance production enforces the assumption (Assumption 6.7) that firms prefer zero-profit production to (forced) exit. Assumption 6.7 is necessary to permit entry deterrence and non-entry to be the *unique* Nash equilibrium. Otherwise, incumbent firms are indifferent between entry deterrence and exit, as both give a zero payoff. This means that for entry deterrence (*i.e.*, average cost pricing) to be the unique Nash equilibrium strategy in a contestable market, a positive exit cost must be assumed, which contradicts the very definition of contestability! However, it suffices to have an infinitely small exit cost: $i, e \pi = -\epsilon$ if $i, e d = 0$ for $\epsilon \rightarrow 0$. Porter (1976) points out that this exit cost may be psychological (Chapter 19).¹

It is now easily proven that compliance with a zero-profit condition by incumbent firms constitutes a Nash equilibrium, where a Nash equilibrium (Nash 1951) is a set of strategies, one for each player, such that the set of strategies maximizes the payoff of each player, given the strategies of the others.²

PROPOSITION 6.1 (zero-profit condition). *A necessary condition for sustainability of a contestable market configuration is that incumbent firms yield zero-profits: $[p - ac(i, q)] \cdot i, q = 0$ for $i = 1, \dots, ni$, conditional upon*

- (i) *incumbents fear that there are some potential entrants which receive $e \pi = 0$ in the incumbents' market if they refrain from entry (credibility; Assumption 6.5 and 6.6)³, and*
- (ii) *expelled firms incur a negative payoff: *i.e.*, they produce some supply in advance: z*

¹ An alternative argument can be that, since potential entrants only decide to enter *after* observing a profit opportunity, unsuccessful entry (and thus a positive exit cost) does not occur. As Gilbert (1989b) argues, "[t]he direct cost of exit is irrelevant if ... potential entrants correctly forecast post-entry profits and avoid markets where earnings would fall below the opportunity cost of their capital" (Gilbert 1989b, p. 521). However, this argument is not very convincing as the existence of a positive exit cost (which opposes the very definition of contestability) is not eliminated.

² This means that starting from a Nash equilibrium position none of the players can increase his/her payoff by deviating from the equilibrium strategy.

³ Equivalent results emerge if it is assumed that incumbents take into account a nonnegative probability of the existence of some zero-profit potential entrants. This means that Selten's (1975) "trembling hand" perfect equilibrium is imposed.

> 0 (uniqueness; Assumption 6.7).¹

Compliance with the zero-profit condition by incumbent firms constitutes then a unique Nash equilibrium.

Proof Proposition 6.1. If an incumbent firm deploys $i^p > ac(q)$, then a potential entrant can enter profitably by offering $e^p < i^p$, bearing in mind that $i^{ac} = e^{ac}$, because incumbent firms and potential entrants face identical (technological) opportunities (Assumption 6.1 and 6.2). From Assumption 6.4 it immediately follows that $i^d = 0$ and $e^d = q$. So, an entrant which receives $e^\pi = 0$ if it refrains from entry, will outperform the incumbent. Whenever the incumbent firm adopts $i^p = ac(q)$, a potential entrant is unable to enter successfully without incurring losses. Now, an extensive form tree of the game can be constructed, showing that the strategy pair $i^p = ac(q)$ and non-entry constitutes a Nash equilibrium (Figure 6.2).

Figure 6.2. Extensive Form Tree of the Game in a Contestable Market

If firms produce some supply in advance (or, alternatively, Assumption 6.7 holds), then *only* the incumbent firm's strategy $i^p = ac(q)$ is associated with a nonnegative payoff, given the set of optimal replies of the potential entrant. Hence, in order to prevent a potential entrant being able to benefit from a second mover advantage, the incumbent firm adopts average cost pricing. *Q.E.D.*

¹ The assumption that expelled firms incur a negative payoff is plausible for many additional reasons. For example, exit is probably associated with a reduced probability of long-run survival and the need to fire employees.

Proposition 6.1 shows that for $n_i > 1$ sustainability is associated with *minimum* average and marginal cost pricing and thus static Pareto optimality.¹ A price in excess of the minimum average cost level provokes profitable entry. For $n_i = 1$ the result is different. A (natural) monopolist can face increasing returns. Provided that there is a potential (natural) monopolist behind the scene, perfect contestability now induces average cost pricing. Price exceeds marginal cost. However, average cost prices satisfy a second-best condition.² The results are illustrated with the help of Figure 6.3. A nonmonopoly firm is forced to produce the quantity that is associated with minimum average cost (nq) so that minimum average cost pricing (np) is feasible. A (demand) monopoly, however, faces a demand constraint (D). On the one hand, a price below average cost (mp) yields negative profit, while, on the other, a price in excess of average cost, $^mp > ac(D)$, induces entry, since a potential (natural) monopolist is then able to enter profitably by slightly undercutting the incumbent's price.

Figure 6.3
(Minimum) Average Cost Pricing

The shape of the average cost curve is not only important for distinguishing the (natural) monopoly from the nonmonopoly case, but is also crucial for (i) the determination of the

¹ Strictly speaking, the contestable market theory deals with a long run equilibrium position. For example, Cairns and Mahabir (1988) argue that Baumol *et al.* assume "a timeless model, in which there is no short run and an instantaneous long run" (Cairns and Mahabir 1988, p. 270).

² Ramsey optimality (Ramsey 1927) implies that buyers' utility is maximized, subject to the financial viability of the firms (Baumol *et al.* 1982a, p. 6). The single product analogue of Ramsey optimality can read as average cost pricing.

likelihood of sustainability and (ii) the endogenization of (the set of) market structure(s) for $n_i > 1$. It appears that sustainability of a market configuration is facilitated by a *flat* bottom of the average (variable and fixed) cost curve. The argument behind this result runs as follows. Suppose that the usual U-shaped average cost curve applies. Let q^* denote the unique quantity that is associated with minimum average cost. Sustainability requires that $D(\text{mac})/q^*$ gives an integer, since otherwise $RD = D(\text{mac}) - \sum_{i=1}^n q^* < q^*$ (where RD denotes residual demand), which violates the very definition of sustainability. The introduction of a flat-bottomed U-shaped average cost curve¹ increases the probability that there is an incumbents' output level q^* associated with the minimum average cost level such that $D(\text{mac})/q^*$ gives an integer number of firms in equilibrium.

Let the flat bottom range from L_q to R_q (Figure 6.4), where $L_q > 0$ and $R_q = k \cdot L_q$.

Figure 6.4.
Flat Bottom and Sustainability ($k > 1$)

Source: Baumol *et al.* (1982a)

¹ The assumption that average cost is minimal over a range of output levels is not in conflict with results of empirical research (see, for example, Shepherd 1979, p. 89 and Scherer 1980, p. 87). As Koutsoyiannis, however, notes: "The empirical evidence ... supports the view that there are no diseconomies of scale at large scales of output. The empirical evidence has not, however, established conclusively whether costs remain constant beyond a certain minimum optimal scale, or fall continuously with scale" (Koutsoyiannis 1983, p. 137). This means that the empirical results do not support either assumption on the shape of cost curves.

Then, for $k \geq 2$ the probability of temporary unsustainability is even zero (Ten Raa 1980, Baumol 1982 and Baumol *et al.* 1982a). Moreover, it is now easily proven that even a (natural or demand) monopoly is forced to deploy minimum average cost pricing for $L_q \leq D(\text{mac}) \leq R_q$. Figure 6.4 illustrates the implications of the flat bottom for $k < 2$. The larger the flat bottom of the average cost curve: ($L_q - R_q$) and the larger the equilibrium number of firms (n_i), the higher the probability of temporary sustainability. Moreover, Figure 6.4 shows that the set of feasible and sustainable market structures (*i.e.*, number and size of the incumbent firms in equilibrium) follows *endogenously* from demand and cost properties. Note that the flat bottom implies that a multitude of equilibrium market structures can be derived (Holler 1985).

6.5 CRITIQUE OF STATIC BENCHMARK CASES OF COMPETITION

6.5.1 Two Lines of Critique

Immediately following Baumol's (1982) address a barrage of questions arose regarding the merits and robustness of the concept of (perfect) contestability (Brock 1983, Spence 1983, Schwartz and Reynolds 1983, Weitzman 1983 and Shepherd 1984). The purport of the critique is that perfect contestability only represents an "odd special case" (Shepherd 1984, p. 577) for two reasons. First, it is argued that the first-best results are generated on the basis of *very special (restrictive)* assumptions, in particular with regard to entry and exit lags (Schwartz and Reynolds 1983, p. 488). Second, it is said that the framework reflects a *static* theory.

This section reviews the critique. The focus is, again, on the theory of contestable markets. However, some of the essential points made below also apply to perfect and Bertrand competition. Subsection 6.5.2 describes the critique that is based upon indicating the special and restrictive assumptions that underlie perfect contestability. Subsection 6.5.3 examines the flaws following from the static features of the frameworks discussed in this chapter.

6.5.2 Restrictive Assumptions

The first line of critique follows from the argument that perfect contestability is based upon very special assumptions. These restrictive assumptions are (i) absence of (price) retaliation, (ii) zero sunk and exit cost, (iii) Bertrand competition, (iv) (indeterminacy of) market structure and (v) pervasive potential competition. Below these five points of critique are examined subsequently.

(i) According to Schwartz and Reynolds (1983) the assumed *absence of price retaliation* is associated with specific assumptions of entry and exit lags (see also Brock 1983, p. 1057). To be precise,

perfect contestability requires two implausible conditions: (i) in response to high prices, an entrant can enter instantaneously at any scale, that is, there is no entry lag; and (ii) an entrant can undercut an incumbent's price and exit with no loss of fixed costs before the incumbent can adjust price, that is, the incumbent's price adjustment lag exceeds the exit lag. If these conditions are relaxed even slightly, the results can differ dramatically from those obtained under perfect contestability. Specifically, if (ii) holds but (i) is violated, the incumbent may

charge a high price pre-entry even though entry will eventually bring price down. If (ii) is violated, entry may not occur; whether it occurs depends upon the nature of the nondegenerate oligopolistic game post-entry (Schwartz and Reynolds 1983, p. 488).¹

Baumol *et al.* (1983) respond by arguing that

contrary to the perceptions of Schwartz and Reynolds, the static equilibrium theory of perfectly contestable markets contains *no assumptions whatever* about the possibility of instantaneous entry or the ability of new firms to exit more quickly than an incumbent can change its prices. Such dynamic properties are best thought of as attempts to examine the plausibility and range of applicability of the static equilibrium concepts. Conditions (i) and (ii) of Schwartz and Reynolds are *not* required by perfect contestability, nor, as we have seen, are they a logical consequence of the absence of *economically* sunk costs, the attribute of a market which, we believe, makes the perfect contestability paradigm applicable. In short, while the conditions which Schwartz and Reynolds cite may be *sufficient* to ensure perfect contestability, they are certainly not *necessary* (Baumol *et al.* 1983, p. 496; Baumol *et al.*'s italics).²

However, Schwartz and Reynolds' argument indicates that an investigation of scenarios that illustrate the applicability and plausibility of perfect contestability is needed (Baumol *et al.* 1983, p. 496). In this respect, it is important to note Brock's (1983) observation that

[s]ome game forms will yield perfect contestability as an equilibrium outcome, others will not. It seems that we cannot expect costlessly reversible entry, by itself, to lead us automatically to the nirvana of social optimality (Brock 1983, p. 1062).

For example, Knieps and Vogelsang (1982), Brock (1983), Schwartz and Reynolds (1983), Schwartz (1986), Stiglitz (1987), Van Wegberg and Van Witteloostuijn (1989) and Chapter 11 pursue this line by examining games and assumptions that generate (non)contestability.

(ii) The hard core of the contestability theory is the assumption of *zero sunk and exit cost* (and so costless exit).³ The condition of zero sunk and exit cost raises the question of the

¹ Shepherd (1984) even states that an assumption of total entry is inconsistent with Bertrand-Nash expectations: "The two assumptions, of trivial entry and total entry, are opposites. If entry is trivial, it has no force. If it is total (or even merely significant), then the no-response assumption is not tenable. ... Total entry, which would entirely duplicate and replace even a monopolist, would be particularly absurd in a Bertrand-Nash model" (Shepherd 1984, p. 576). See Subsection 6.4.2 for Baumol *et al.*'s defence of the assumption of the absence of price retaliation.

² Schwartz and Reynolds (1983) point to a further argument in favor of contestability: "In some markets there may exist outside institutions that prevent price reactions by incumbents or negate their effect on entrants. For example, long-term contracts may enable entrants to contract for a fixed price before entering" (Schwartz and Reynolds 1983, p. 490). The plausibility of this argument is debated (Baumol *et al.* 1983, Schwartz 1986 and Stiglitz 1987).

³ Entry and exit lags [critique (i)] and sunk costs [critique (ii)] are closely related phenomena. Schwartz (1986) argues that the "degree of sunk cost generally increases with exit speed" (Schwartz 1986, p. 41). Farrell (1986) indicates that sunk costs and response lags must be viewed together in order to determine the degree of contestability of a market (Chapter 11). McLeod (1987) clearly distinguishes the exit lag and unrecoverability components of sunk cost. This again points to the fallacy of the identification of sunk cost as exit cost (Subsection 6.4.3). The key point is that sunk cost only gives an exit cost if amortization is not yet completed at the date of exit. So, a positive sunk cost can be associated with costless exit if the exit lag

plausibility, generality and robustness of the theory of perfectly contestable markets.¹ Shepherd (1984), for example, argues that

[t]o establish ultra-free entry² as a general case, one will probably need to show that its normative results are robust even to large departures from the assumptions Fixed and sunk costs commonly overlap and are sizable Zero sunk cost is therefore a doubtful, counterfactual assumption for a general theory (Shepherd 1984, p. 577).

Stiglitz (1987) points out that (endogenous) sunk costs (technology) are crucially important in industrial organization:

The general point remains: the incumbent firm's choice of technology will be affected by the threat of entry. It may choose a technology that will deter entry. It may choose a technology that will make another firm's entry unprofitable if the two firms compete vigorously. The incumbent firm may choose a technology that will make the entrant believe that it will not exit. Or it may choose a technology that will make the entrant believe it will not cooperate. The incumbent will choose to sink costs: the contestability doctrine requires that no technology with sunk costs will be available, a clearly inadmissible assumption. And the technology chosen in response to the threat of competition will not, in general, be the efficient technology (Stiglitz 1987, p. 910).

It is in particular argued that even small sunk (exit) costs undermine the favorable implications of contestability (Brock 1983, Schwartz and Reynolds 1983, Schwartz 1986, Stiglitz 1987, Dasgupta and Stiglitz 1988a and Martin 1989³). Two quotations illustrate this point:

To summarize, Baumol's theory is not robust; once we deviate even slightly from the strict assumptions of perfect contestability, pricing and entry decisions depend upon the nature of firm interactions (Schwartz and Reynolds 1983, p. 489); and

There is, however, a widespread belief that even if actual competition is absent, potential competition may be sufficient to discipline firms in an industry and to hold prices down. This argument demands that scale economies be due to costs that are not sunk. But it has not been recognized precisely how sensitive the conclusions are to the assumptions. I have shown that the presence of arbitrarily small sunk costs can serve as an absolute barrier to entry and make potential competition completely ineffective as a discipline device (Stiglitz 1987, p. 932).

is sufficiently long (as a long exit lag permits the recovery of the sunk cost in the market, which gives a zero exit cost).

¹ Moreover, there are doubts about Baumol *et al.*'s use of cost concepts. First, Weitzman (1983) argues that "[a]s a matter of formal theory, you cannot have a range of decreasing average cost without sunk costs" (Weitzman 1983, p. 486). Second, Shepherd (1984) asserts that "fixed costs are claimed (or defined) to be irrelevant to entry barriers. But that is true only if the entrant can match the incumbent firm's output completely, by means of total entry" (Shepherd 1984, p. 577). For a response see Baumol *et al.* (1983 and 1986).

² Shepherd argues that "the term contestability has two defects. It misplaces attention from entry conditions to a nonexistent post-entry struggle. And it implies degrees of variation in entry that are not logically permissible. The phrase *ultra-free entry* more precisely reflects the crucial conditions" (Shepherd 1984, p. 573; Shepherd's italics).

³ Martin (1989) even "demonstrates that financial markets can place entrants at a cost disadvantage, even if capital costs are not sunk. Financial markets increase the range within which incumbent firms may engage in strategic behavior" (Martin 1989, p. 1096).

Baumol *et al.* (1983) admit the crucial role of sunk and exit cost. They argue that the degree of contestability depends *continuously* on the scale of sunk cost (p. 495). In effect,

[i]t should be emphasized that models which support the robustness of contestability analysis follow a relatively long tradition going back at least to the work of J.S. Bain. This tradition holds that increased ease of entry and exit improves the welfare performance of firms and industries. On this subject, the theory of contestability has only sought to contribute insights on the underpinning of that judgement (Baumol *et al.* 1983, p. 494).¹

(iii) The results of the contestability theory depend critically upon the assumption of *Bertrand competition*. Perfect contestability is not necessarily generated if other assumptions are introduced. Two examples are Cournot competition and collusion. For illustrative purposes, here the results of some models with Cournot competition in a free entry framework are briefly indicated.²

Knieps and Vogelsang (1982), Brock (1983), Calem (1988), and Maskin and Tirole (1988) employ the assumption that quantity rather than price is the competitive weapon of firms in a free entry setting. On the one hand, Knieps and Vogelsang (1982) and Brock (1983) show that Cournot competition is generally *not* associated with average cost pricing. Brock indicates that

as in the quantity-setting model criticized by Baumol *et al.*, the incumbent can protect some monopoly profit, no matter how fine the partition, even though entry is costlessly reversible (Brock 1983, p. 1061).

On the other hand, Calem (1988) and Maskin and Tirole (1988) present models which illustrate that, under particular conditions, Cournot competition and contestability are compatible (Chapter 9 and 11). Again, it appears that the result of the models depends upon the specifics of the rules of the game.

(iv) The contestability framework is liable to the critique that the equilibrium market configuration is *indeterminate* (Subsection 6.4.4). From a static point of view this means, as Holler (1985) argues, that

there is no mechanism indicated which will ensure that any equilibrium will be realized as a consequence of the actions of the individual firms. This result challenges a theory that proposes to analyse the determination of the industry structure endogenously and simultaneously with the pricing, output, and other decisions of the industry's firms. The possibility of indeterminateness, which is inherent to the theory of contestable markets, is but a second version of the well-known fact, that the number of firms as well as their individual output is indeterminate in competitive markets, if technology is linear (Holler 1985, p. 67).³

¹ Chapter 12 argues that in particular circumstances zero profits and positive sunk costs can even go together. This is the perfect barrier market. Hence, the perfect barrier market notion supports the proposition that the contestability results can be robust to deviations from the zero sunk cost assumption.

² Stiglitz (1987) examines models in which firms are allowed to collude. These models generate noncontestable results (*i.e.*, prices above average cost).

³ Shepherd (1984) even accuses Baumol *et al.* of being unfamiliar with the literature: "Pure (or reasonably effective) competition has also yielded this result [endogenous market structure], despite Baumol *et al.*'s denial. Competition enforces minimum-cost production. Indeed the claim that in 'older' research, structure was 'determined exogenously in a manner totally unspecified' (Baumol, 1982, p. 7) suggests an

So, the flat bottom of the cost curve (*i.e.*, the range of constant returns) increases the probability of temporary sustainability on the one hand, but introduces a multitude of equilibrium market structures on the other. Only the *set* of sustainable market configurations is endogenized.

(v) The key element of the theory of perfect contestability is the assumed pervasive force of potential competition. The extreme dominance of the threat of entry raises empirical issues. Although the contestable market approach mainly seeks to offer a normative benchmark case of competition, Baumol *et al.*'s own positive arguments have provoked the criticism of the *empirical plausibility* of the theory and its assumptions. In this respect, Baumol's (1982) reference to the airline industry as an example of a contestable industry and Baumol *et al.*'s (1982a) agenda for empirical research (Chapter 16, p. 467) are worth noting. In response to Baumol *et al.*'s suggestions, Shepherd (1984) makes a vicious attack on the empirical validity of contestability:

There seems to be no general theory here, only an odd special case. ... Theorizing before testing is often an acceptable approach, but Baumol *et al.* have taken it beyond reasonable bounds of valid scientific method. Implausible assumptions have been applied on an abstract plane to reach not only 'insights', but also emphatic conclusions and wide policy lessons. The system hangs in the air, lacking a foundation or even plausibility. ... A large inventory of research findings, accrued over many decades, has shown internal conditions to be the primary force (Shepherd 1984, pp. 577 and 579).

Spence (1983), for example, takes a more moderate stance:

My personal view is that the range of applicability of the theory as a descriptive tool for analyzing short and medium run industry configurations is limited by the widespread existence of partial irreversibility of investment. ... entry barriers lack permanence or durability. If this is correct, it follows that for certain analytical purposes relating to the long-run evolution of industries, the theory's range of descriptive applicability may be greater than many of us whose eyes are trained mainly upon competitive strategy and entry barriers might imagine (Spence 1983, pp. 986 and 988).

The results of actual empirical research are ambivalent.¹ Supportive findings on the basis of tests of the contestability of industries have been reported by Cassidy (1982; ocean transport industry), Froeb and Geweke (1984; aluminum industry), Davies (1986; liner shipping industry) and Nathan and Neave (1989; financial market). However, studies of Shaw and Shaw (1977; polyester fiber industry) and Call and Keeler (1984; airline industry) invalidate the contestability theory. Experimental studies of contestability (Coursey *et al.* 1984a and 1984b, Harrison and McKee 1985, Harrison 1986 and Harrison *et al.* 1987) generally support the theory of contestable markets. However, Schwartz (1986) argues that

unfamiliarity with the literature" (Shepherd 1984, p. 578). There is indeed literature on the determination of market structure in a free entry setting. Some examples are Stackelberg (1932) and Prescott and Visscher (1977) and more recently Bernheim (1984), McLean and Riordan (1985) and Eaton and Ware (1987). However, Baumol *et al.* (1982a) do refer to some of the literature that is related to their approach.

¹ In this respect, the difference in tone of Baumol and Willig's (1986) and Schwartz' (1986) review of the empirical literature is striking.

[f]irst, the evidence offered cannot be interpreted as favorable or unfavorable to contestability theory. Second, and more important, it is doubtful whether the entire experimental approach can contribute anything to appraising the empirical importance of contestability theory (Schwartz 1986, p. 50).

On the basis of a survey of empirical evidence on potential competition Gilbert (1989a) concludes that

[t]he observations presented here suggest that potential competition is important, but not as powerful as the theory of contestable markets implies (Gilbert 1989a, p. 123).

6.5.3 Static Theory

The second line of critique stems from the static features of perfect contestability. The major critique of the static benchmark theories of competition in general focuses precisely on the static features of these frameworks. Baumol *et al.* admit that perfect contestability is a static concept. They style their framework as

the static equilibrium theory of perfect contestable markets (Baumol *et al.* 1983, p. 496).

Kirzner (1973), for example, argues that

the emphasis on equilibrium hampered any possible appreciation of the notion of competition which we have seen to be the outstanding characteristic of the market process. By definition, a state of equilibrium does not permit activity designed to *outstrip* the efforts of others in catering to the wishes of the market (Kirzner 1973, p. 27; Kirzner's italics).

Static theories shed insufficient light upon the dynamic aspects of competition.

The literature suggests that perfect competition implies that there are no incentives to be engaged in innovative activity (Kamien and Schwartz 1982, p. 191). The reason is simple: perfect competition is associated with zero profits; it does not pay to invest. The same argument applies to Bertrand competition and contestability. Existing contestable market models even exclude meaningful investment by definition, since sunk exit cost is zero. Existing contestable market models deal with the case where investment cannot occur if it is associated with sunk exit costs (Schwartz and Reynolds 1983 and Shepherd 1984).

The abstraction of sunk exit costs implies that perfect contestability (i) fails to endogenize innovative activity, (ii) is likely to be intertemporally unsustainable and (iii) ignores *ex post* (oligopolistic) interactions. Below the three points are examined in turn.

(i) Shepherd (1984) notes that sunk (exit)

costs include many categories besides physical capital, such as R&D, advertising to establish brand loyalty, and training to create special workers' skills. These intangible forms are often more fully 'sunk' than physical capital, which can be leased or resold (Shepherd 1984, note 19, p. 580).

For example, R&D outlays are an important type of sunk investment. Stiglitz (1987) argues that

"[m]ost expenditures on R&D are, by their very nature, sunk costs. The resources spent on a scientist to do research cannot be recovered (Stiglitz 1987, p. 928).

This is not to say that the contestable market framework does not leave room for the introduction of innovations. To the contrary,

we can prove that even if it is run by a monopoly, a contestable market will ... yield zero profits ... and in addition enforce efficiency of production, the adoption of new improved techniques as they become available (Baumol and Willig 1981, p. 420).

In effect,

[t]he efficiency and innovation properties follow from the same argument as the zero-profit condition. For if the incumbent does not use least-cost techniques for production, he obviously invites profitable entry (Baumol and Willig 1981, note 13, p. 420).¹

However, in order to uphold the contestability features of the market, the literature suggests that the adoption of new techniques is *not* allowed to be associated with sunk exit costs. As Stiglitz (1987) argues,

much of the literature on contestability seems to depend on the existence of technologies with no sunk costs (Stiglitz 1987, p. 907).

In effect, the adoption of new techniques takes the form of buying new (and selling old) capital which comes *exogenously* available (Baumol *et al.* 1982a, Chapter 13). The *endogenization* of innovations (*i.e.*, R&D activity) requires, however, the introduction of sunk exit costs (Stiglitz 1987).

(ii) From a dynamic angle it appears that there are many sources of the intertemporal unsustainability of contestable market configurations. Baumol *et al.* are the first to recognize this difficulty. This is particularly clear from Chapter 14 in their 1982 book:

Unsustainability can come closer to being the rule rather than the exception. ... We begin our discussion with the source of the difficulties that appear to be most general and most serious - economies of scale in sunk construction costs (Baumol *et al.* 1982a, pp. 405-406).²

So, the argument is that any sunk investment is particularly detrimental to the intertemporal sustainability of contestable market configurations. Parts V and VI explore the conditions under which this proposition is true.

(iii) The argument that perfect contestability ignores *ex post* oligopolistic interaction is closely related to the criticism that Baumol *et al.* fail to model dynamic games [which is indirectly related to critique (i) and (ii) in Subsection 6.5.2]. The combination of the assumptions of the absence of (price) retaliation and Bertrand competition implies that the

¹ Spulber (1988) models a monopolist in a contestable market which introduces second-best product liability rules. Here too sunk exit costs are absent.

² Besides, Baumol *et al.* (1982a) elaborate on three other sources of dynamic unsustainability: learning by doing (pp. 429-434), postponability of fixed outlays (pp. 434-437) and pure public goods with postponable purchase date (pp. 437-443).

ex post game (i.e., the game after entry) is degenerate.¹ As Spence (1983) and Brock (1983) put it:

What is set aside in the theory are elements of strategic interaction that we have come to associate with entry deterrence; matters such as preempting market positions by making irreversible moves first (Spence 1983, p. 982); and

Oligopolistic interactions matter. Baumol *et al.* (1982a) argue that such oligopolistic indeterminacy (and hence the necessity of making unpalatable assumptions about entrants' conjectures that cannot be supported by economic 'common sense') is avoided in the presence of perfectly reversible and frictionless entry. Baumol *et al.* need some argument for this conclusion because they impose the dual (equally unpalatable) assumption that entrants enter in the expectation that incumbents' prices remain fixed. ... But nowhere in the book is a game precisely defined as modern theory expects ... That is not easy, but until it is done contestability theory must be used with caution. Unpalatable hidden assumptions may be necessary to obtain Baumol *et al.*'s conclusions (Brock 1983, p. 1059; Brock's italics).

Knieps and Vogelsang (1982) and Subsection 6.4.4 describe games that give contestability outcomes, whereas Brock (1983), Schwartz (1986), Stiglitz (1987) and Maskin and Tirole (1988), among others, present dynamic games which generally fail to generate contestability. However, according to Baumol *et al.* the critique that the theory of contestability ignores dynamic games is besides the mark, because the opponents

seek to criticize what is essentially a static equilibrium theory by *imputing to it* a dynamic mechanism of *their own creation*, and attacking the reasonableness of the assumptions *they* have made about the dynamic process governing the behavior of agents in contestable markets outside equilibrium. This is not to say that dynamic considerations are unimportant, rather that they should not be interferred casually from an informal statement of a model's equilibrium properties (Baumol *et al.* 1983, p. 495; Baumol *et al.*'s italics).

6.6 SUNK INVESTMENT AND IDENTIFICATION OF POTENTIAL ENTRANTS

Much of the critique of the contestability theory ensues from misunderstandings.² Two crucial features of the theory must always be kept in mind: (i) it is a *static* equilibrium theory; (ii) it seeks to offer a *normative* standard.³ As Spence (1983) nicely puts it (see also Gilbert 1989b, p. 527):

remember that we are talking for the moment about a theory that provides a normative standard against which market outcomes can be judged. There may also be conditions under

¹ Appelbaum and Lim (1985) present a free entry model with an *ex ante* and *ex post* game in order to endogenize the degree of contestability. Mankiw and Whinston (1986) model a free entry framework that yields inefficient (i.e., excessive) entry *ex post*.

² Compare with Davies and Lee (1988), who offer a critique of the contestability theory from a post-Keynesian perspective. In particular, they indicate three basic flaws: (i) time has no real meaning (short and long run collapse in instantaneity); (ii) the definition of the market is too broad (actually, there is an infinite number of competitors); and (iii) long-run strategic behavior is missing (full-cost pricing is ignored).

³ The confusion is clearly expressed by Shepherd (1984): "Because the whole set of writings is so audacious and complex, mingling pure theory with claims about reality, a full evaluation is difficult" (Shepherd 1984, p. 579).

which the theory is a reasonable descriptive approximation to reality as well. Though sunk costs and related aspects of structure are important influences on the dynamic aspects of competition, I think it perfectly appropriate to develop a theory that focusses on some other important structural aspect of many markets. It is neither fair nor reasonable to require microeconomic models to capture all relevant aspects of reality. No model that is tractable and informative does. The only caution is that because models do leave things out, one needs to exercise some discretion and judgment in using them (Spence 1983, p. 982).

However, two assumptions of contestability theory deserve careful consideration: zero sunk and (almost) zero exit cost and unidentified potential entrants. The critique of the assumption of zero sunk (and exit) cost is outlined in Subsection 6.5.2 and 6.5.3: without sunk investment the static - dynamic efficiency tradeoff cannot be considered. With respect to the source of potential entry Cairns and Mahabir (1988) point out that

Baumol *et al.* (1982) give a rather confused picture of entry into an industry characterized by production by single-product firms. The contest between entrants and incumbents is not so much degenerate (as claimed by Schwartz and Reynolds, 1983) as ephemeral. One is asked to imagine a world populated by a reserve army of phantom entrants, described variously as 'latent', 'unidentified' or 'in the wings'. But in discussing the key concept of the difference between fixed and sunk costs, Baumol *et al.* refer to moving *existing* locomotives, aeroplanes, etc., from one market to another. ... The main questions to be asked are: *Who* are the entrants? Are they indeed idle, thus not earning their opportunity cost, as the terms 'latent' and 'in the wings' suggest? ... If they are not idle, from what other activity do they, their labour and their equipment come? ... Caves and Porter (1977, p. 243) have already dubbed such a reserve army as 'an inconceivable list of unconceived corporate children' (Cairns and Mahabir 1988, p. 270; Cairns and Mahabir's italics).

This means that the *credibility* of the entry threat, which is the key foundation of contestability (outcomes), hangs in the air.

Moreover, the fact that the introduction of sunk investment (and so dynamic economies) goes beyond the perfect contestability framework invalidates the theory as a benchmark case for *dynamic* competition. For example, if perfect contestability is incompatible with the introduction of sunk R&D outlays which produce *process* and product innovations (Stiglitz 1987), an essential element of the efficiency features of competition is missing. The barrier market theory seeks to introduce exit cost and sunk (R&D) investment in a free entry framework. Furthermore, additional remarks can be made about the identification of the potential entrants (Chapter 9, 10, 11 and 21 in particular).¹ In this way, a *dynamic multimarket* theory of contestability is offered which gives in to some of the main objections listed above. The key question is: can dynamic

¹ Van Wegberg and Van Witteloostuijn (1989) analyze the contestability assumption from a multimarket perspective, which permits the *identification* of potential entrants. The literature's ignorance of the strategic dimension in the potential entrants' decision making is noted by Schwartz and Reynolds (1986): "If there are multiple *entrants*, the problem becomes more complicated since an entrant must know the number of such entrants, and whether they would enter sequentially or simultaneously. With a large number of entrants moving simultaneously, it is possible that none would enter because of the losses that would be incurred if all entered simultaneously (see, for example, Roger Sherman and Thomas Willett, 1967). The implications of different numbers of entrants and different interactions among them have not been thoroughly analyzed to our knowledge" (Schwartz and Reynolds 1983, p. 489, note 6; Schwartz and Reynolds' italics). The problems associated with multiple entrants are also indicated by, for example, Goldberg and Moirao (1973), Kalish *et al.* (1978) and Nti (1989).

market economies be introduced in a contestability (or, to be precise, free entry) framework without damage to the static efficiency features of pricing? This question appears to be closely related to the credibility issue and the topic of the sources of potential entry. Before answering this question, a selection of literature on dynamic economies is reviewed. Here the issue of entry barriers (Chapter 7) and innovation (Chapter 8) are particularly relevant.

CHAPTER

7

DYNAMIC EFFICIENCY: Entry Barriers, Sunk Investment and Entry Deterrence

7.1 ACTUAL AND POTENTIAL COMPETITION

In Chapter 6 time and again the term entry barrier is used. The essential feature of contestable markets [and, for that matter, barrier markets (Part V)] is precisely the absence of entry barriers. The literature suggests that barriers to entry and sunk (exit) costs are but two sides of the same coin (for example, Baumol and Willig 1981, Grossman 1981, Farrell 1986, Stiglitz 1987, Dasgupta and Stiglitz 1988a and Martin 1989). Entry barriers and sunk investment are closely associated with dynamic competition. Therefore, an introduction to the literature on entry barriers cannot be missed in a discussion of the static - dynamic efficiency tradeoff. This chapter offers in a nutshell an examination of the basic concepts and a review of the literature. Of course, the survey can only be brief, since the literature on entry barriers is immense (Lyons 1988, Schmalensee 1988, Tirole 1988 and Gilbert 1989b). Ever since Bain's (1949, 1951 and 1956) seminal analysis of barriers to entry, the body of literature on this topic has been growing exponentially (Van Cayseele and Schreuder 1988 and Van Witteloostuijn 1989b).

In the centuries before the publication of Smith's *The Wealth of Nations* (1776), theorists had already philosophized about the concept of competition. Competition in the classical framework implies a dynamic process of actual competition (McNulty 1968). Although early neoclassical economics shifted the focus from dynamic processes to equilibrium states, the first neoclassical theories still took actual competition as the point of departure. Here the well-known models of perfect (Clark 1915 and Knight 1921) and imperfect (Chamberlin 1933 and Robinson 1933) competition are worth noting. The modern literature on industrial organization, however, gives plenty of room to the study of the implications of *potential* competition.

Although the classical and early neoclassical authors had an understanding of the role of entry threats, the formal modeling of the effects of potential competition did not take shape until the 1950s. The pillar of the literature on potential competition is the theory of barriers to entry. The recognition of entry barriers as an important determinant of market performance follows logically from the development of economic thought on the influence of potential competition on the behavior of incumbent firms.

Section 7.2 briefly deals with Bain's predecessors and contemporaries. Section 7.3 examines the outlines of Bain's contribution. Section 7.4 points out the definitions, categories and sources of entry barriers indicated in the literature. Section 7.5 discusses two well-established conjectural assumptions: Sylos' postulate and excess capacity

hypothesis. Section 7.6 outlines the literature on the strategic component of entry deterrence. Section 7.7 puts the entry barrier literature into a normative perspective by examining the impact of barriers to entry on market performance. Section 7.8 offers final observations.

7.2 BAIN'S PREDECESSORS AND CONTEMPORARIES

The recognition of entry barriers as an important determinant of market performance follows logically from the development of the economic insights of the influence of potential competition on the behavior of incumbent firms. Although Smith (1776, p. 60), Clark (1915, pp. 380-381) and Robinson (1933, p. 81), among many others, already recognized the force of potential competition and its relation to entry barriers, the advanced analysis of barriers to entry only began in the 1950s. Here the contributions of Harrod (1953), Hicks (1954), Hahn (1955) and of course Bain (1956) are worth mentioning.

Harrod (1953) explicitly assumes that incumbent firms take into account the threat of entry while deciding on their short-run policy. Therefore, incumbent firms capture only normal profits in order to deter entry by potential entrants, because they seek to maximize long-term profit. Hicks (1954) points out that the behavior of incumbent firms is not determined *a priori* by long-term profit maximization. Incumbent firms decide on the basis of the maximization of a profit function that consists of a weighted average of short-term and long-term profit. In this respect, incumbent firms may well be heterogeneous. Hicks focuses on two extreme types: the "snatcher" and the "sticker". A snatcher is mainly interested in seizing quick profit, whereas the sticker seeks to build up a steady business. That is, a sticker is much more inclined to carry out entry-detering strategies than a snatcher. Moreover, Hicks stipulates that particularly expectations of post-entry behavior determine the decisions (and actions) of incumbent firms and potential entrants. Hicks introduces two contrasting types of conjectures: oligopolistic and polypolistic expectations. Polypolistic expectations imply the firm's belief that its short-run policy does not influence the probability of future entry, while oligopolistic expectations predict the opposite. Hahn (1955) modifies Hick's two-period setting by introducing a multi-period profit function.

7.3 BAIN

The founding father of the growing literature on entry barriers is no doubt Bain (1949, 1951 and 1956), in particular through his seminal *Barriers to New Competition* (1956). In his 1956 book Bain seeks to explain why entry may be more or less free. He offers a detailed examination of the influence of (the height of) barriers to entry on the behavior of incumbent firms and potential entrants.¹ Bain defines the crucial concept of entry barriers as

the advantage of established sellers in an industry over potential entrant sellers, their advantage being reflected in the extent to which established sellers can *persistently* raise their prices above a competitive level without attracting new firms to enter the industry (Bain 1956, p. 3; italics added),

¹ Moreover, Bain's work gave the impetus to much empirical research in the tradition of the structure-conduct-performance paradigm (Schmalensee 1988 and 1989).

in which the

'competitive level of prices' is defined ... as the minimum attainable average cost of production, distribution, and selling for the good in question, such cost being measured to include a normal interest return on investment in the enterprise. In effect, this is equivalent to the level of price hypothetically attributed to long-run equilibrium in pure competition (Bain 1956, p. 6).

Bain (1951) offers a precise definition of the zero profit that is associated with the competitive level of prices:

A firm 'breaks even' for a period if it recovers in current revenues all currently incurred, currently allocable costs, plus the original cost of all previously acquired assets legitimately amortized, adjusted for price level change since acquisition, plus a current market interest return on original cost value of owners' equity, similarly adjusted for price level change. This identifies the theoretical zero-profit level with capital maintenance in the sense of maintenance of 'assets of constant purchasing power' in the hands of owners. True profits are found in earnings above this level (Bain 1951, p. 307).

So, the competitive level of prices is equal to *total* average cost, where total average cost includes variable, fixed and sunk components. The price which firms can persistently set above a competitive level without attracting new firms is called the limit price. A limit price just deters entry, whereas any price slightly above the limit level induces entry.

Bain argues that barriers to entry must be considered as long-run and structural conditions of a market.¹ He distinguishes three types of entry barriers that follow from incumbents' advantages with respect to (i) absolute cost, (ii) product differentiation and (iii) economies of scale.² Heterogeneity of (incumbent and potential) firms implies that barriers to entry cannot be represented by a unique absolute value. Therefore, Bain introduces the general condition of entry:

The *general condition of entry* then refers to the succession of values of immediate condition of entry as entry to industry occurs - to the distribution of price-minimal cost gaps just necessary to induce successively less favored firms or groups of firms to enter an industry consecutively, beginning with the most favored firm. At any stage in its development, each industry has a general condition of entry in prospect (Bain 1956, p. 9).

¹ There is a problem of the time interval considered here. As Waterson (1984) notes: "It is fairly obviously not the theoretical short run, for in that case there would not be any possibility of entry. It cannot easily be the theoretical long run either, for that would mean no firm having, for example, any product differentiation advantage over another, and all firms both potential and actual would have access to capital on the same terms. Nevertheless, Bain feels that such structural features as entry barriers are not ephemeral short-lived advantages, which leads Williamson (1963) to consider that we are talking about an 'intermediate-run' situation (see, for example, his p. 113, n. 6), some time period long enough for society to be concerned about" (Waterson 1984, p. 57).

² It may be argued that Bain considers four sources of entry, the fourth one being absolute capital costs (for example, Schmalensee 1988, p. 663). Opponents point out that "Bain's argument that an entrant's need to invest absolutely large sums of money might serve as a barrier to entry has been widely criticized because it seems to rest on capital market imperfections; incumbents also had to invest large sums" (Schmalensee 1988, p. 665). This critique, however, misses the point as daily practice proves that capital markets are imperfect indeed (Stiglitz and Weiss 1981 and 1987). Models in which potential entrants face an internal finance barrier are Judd and Petersen (1986) and Poitevin (1989).

Furthermore, Bain recognizes that there are many intervening factors that influence the limit price. He particularly points to (i) the time lag between the decision to enter and actual entry (*i.e.*, the entry lag), (ii) the elasticity of demand, (iii) the degree of collusion among incumbent firms and (iv) the conjectures of incumbent firms and potential entrants as regards to post-entry behavior (in particular, price and quantity replies of incumbents upon entry). With respect to the crucial fourth factor Bain assumes the following:

Initially, therefore, we may suggest that going price policies of established firms are likely to be read to the effect that post-entry price is unlikely to exceed pre-entry price. ... we will also assume that he [the potential entrant] will see a non-negligible and perhaps dominant probability that post-entry prices will be lower than pre-entry prices if he enters at a significant scale (Bain 1956, p. 28).

Bain's notions of entry barriers and his four intervening factors have been subject to subsequent research.¹

7.4 DEFINITIONS, CATEGORIES AND SOURCES OF ENTRY BARRIERS

A natural sequel to Bain's (1956) contribution to the literature is to further differentiate the definitions, categories and sources of entry barriers. This section is an intermezzo on these topics.

(i) *Definitions.* As Demsetz (1982) points out, "[t]he discussion of barriers in the economic literature hardly reflects consensus" (Demsetz 1982, p. 47). Consider, for example, the following two well-known definitions.

DEFINITION 7.1 (*Bain's entry barriers*). *Entry barriers follow from "the extent to which, in the long run, established firms can evaluate their selling prices above the minimal average costs of producing and distribution ... without inducing potential entrants to enter the industry" (Bain 1956, p. 252), where average costs include variable, fixed and sunk components.*

DEFINITION 7.2 (*Stigler's entry barriers*). *Entry barriers are "a cost of producing (at some or every rate of output) which must be borne by a firm which seeks to enter an industry but is not borne by firms already in the industry" (Stigler 1968, p. 67).*

As Demsetz (1982) argues, the definitions imply different opinions about the sources of entry barriers (see also Baumol and Willig 1981). In particular, only Bain's definition includes economies of scale *per se* as an entry barrier, whereas Stigler's definition implies that economies of scale are not an entry barrier so long as (potential) entrants have access to the same production technology. Contestability theory suggests that Stigler's argument is valid (Chapter 6) as scale economies are compatible with an assumption of free entry.

¹ For illustrative purposes, De Bondt (1976 and 1978) can be referred to as an example of the literature that focuses on other intervening factors of influence. De Bondt examines the role of entry lags. It appears that "a pre-entry price may also increase with a lowering of the barriers to entry when the established firm realizes that there is a positive finite entry lag between a rival's decision to enter and his appearance as an effective competitor" (De Bondt 1978, p. 267).

Since this thesis takes contestability as the starting point of the argument, entry barriers are used in the Stiglerian meaning unless otherwise indicated.

This means that free entry implies that Stiglerian *entry* barriers are absent. Any positive entry cost indicates an entry barrier. However, even if entry barriers are absent, firms may face *exit* barriers. On the one hand, this is the case whenever incumbents leave the market before sunk capital is fully amortized (Definition 6.7). On the other hand, exit cost can be an entry barrier for potential entrants (Section 6.4.2). However, a positive exit cost does not *necessarily* confront potential entrants with an entry barrier. Positive exit costs are only an entry barrier if potential entrants foresee a risk of not being able to amortize fully the sunk capital in the entry market (Section 11.2).

(ii) *Categories*. Since Caves and Porter's (1977) contribution to industrial organization, it has been common practice to distinguish *categories* of entry barriers. Caves and Porter point out that there are *exogenous* (structural) and *endogenous* barriers to entry. Endogenous, as opposed to exogenous, entry barriers derive from the (profit-maximizing) behavior of incumbent firms. Exogenous barriers to entry ensue from structural market conditions (for example, geographical features). Following Salop (1979), endogenous entry barriers can be further divided into innocent and strategic categories:

An *innocent* entry barrier is unintentionally erected as a side effect of innocent profit maximization. In contrast, a *strategic* entry barrier is purposely erected to reduce the probability of entry (Salop 1979, p. 335; Salop's italics).

It should be noted, though, that raising a strategic entry barrier is also meant to be a profit-maximizing activity. The point is that strategic entry barriers are *purposely* erected in the face of potential entry.

Strategic, unlike nonstrategic, investments are either (i) carried out *only* with the purpose of reducing the probability of entry or (ii) associated with a larger magnitude of investment. On the one hand, the word "only" is crucial, since there are many investments which serve an innocent profit-maximizing and strategic entry-detering goal at the same time. For example, efficient process innovations are likely to be introduced irrespective of the absence or presence of an entry threat, although they clearly reduce the likelihood of entry as well. On the other hand, the magnitude of the investment may make it strategic. Suppose that there is an (expected) optimal innocent investment I_1 if entry deterrence is irrelevant, but actual investment $I_2 > I_1$ if entry effects are considered. Then, $I_2 - I_1$ is strictly strategic.

Caves and Porter (1977) offer a further differentiation by defining *mobility* barriers. Mobility barriers *within* an industry can indicate the existence of distinctive strategic groups:

A strategic group is a grouping of organizations which pursue similar strategies with similar resources (Hatten and Hatten 1987, p. 329).

Mobility barriers are entry barriers that impede movements from one market segment to another within a particular industry (McGee and Thomas 1986, Cool and Schendel 1987, Hatten and Hatten 1987 and Van Cayseele and Schreuder 1989).

A last category of entry barriers can be labelled institutional. *Institutional* barriers to entry follow from the (intended or unintended) actions of the public sector. A clear example is the public policy to impose a license system on particular economic activities

(Quirk and McDougall 1981). In effect, many of the public regulations generate institutional entry barriers (Kahn 1970), as is particularly clear in the context of international trade (Mayes 1978). Here environmental, safety and quality requirements are appealing examples. Institutional entry barriers can be both exogenous and endogenous. Pressure groups can affect bureaucrats' and politicians' decision making on institutional barriers (Brock 1983, Salop *et al.* 1984, Ordover and Saloner 1989, and Van Mierlo and Van Witteloostuijn 1990).

(iii) *Sources.* As regards to the literature on types of entry barriers in the commodity market it suffices to note that Bain's successors pursue along the lines of his scale economies (Schmalensee 1981 and Perrakis and Warskett 1986), product differentiation (Spence 1980 and Schmalensee 1982)¹ and absolute cost advantages (Bhattacharya 1984 and 1985, and Krattenmaker and Salop 1986).² A typical example of an absolute cost advantage is incumbents which control lower cost production techniques. An example of barriers to entry that are caused by product differentiation is the preference of buyers for established brand names and reputations ("goodwill"). Economies of scale advantages follow from cost benefits as regards large-scale functional activities like production, distribution and/or advertising (Porter 1980).

Moreover, other sources of barriers of entry have been identified: for example, patenting (Fudenberg *et al.* 1983), switching cost (Klemperer (1985), network economies (Katz and Shapiro 1985), inventory holding (Ware 1985), divisionalization (Schwartz and Thompson 1986), trade unions (Dewatripont 1987), contracts (Aghion and Bolton 1987) and unlimited liability (Carr and Mathewson 1988).³ The key contribution of the modern literature is, however, the advanced study of the strategic dimension of entry (deterrence). Raising and demolishing entry barriers are strategic options to incumbent firms and potential entrants respectively. This implies that expectations of the rivals' reactions to strategic actions play a crucial role in the modern theory of potential competition, which means that Hicks' (1954) and Bain's (1956) emphasis on the expectational aspects of entry (deterrence) appear to be the recurrent theme in the recent literature.

7.5 THE SYLOS' POSTULATE AND EXCESS CAPACITY HYPOTHESIS

The immediate reaction to Bain's *Barriers to New Competition* (1956) was that the intervening role in models of entry deterrence of the expectations about rivals' post-entry (re)actions was brought into the center of attention. With regard to assumptions on (expectations of) post-entry reactions of rivals the relative size of the (potential) entrant is

¹ A topic related to product differentiation is advertizing or sales promotion (Nelson 1975, and Comanor and Wilson 1979). For example, Schmalensee's (1974 and 1982), Cubbin's (1981) and Waagstein's (1982 and 1983) models describe a setting with sales promotion as a barrier to entry.

² For example, Waterson (1984), Encaoua *et al.* (1986), Schmalensee (1988) and Schmalensee and Willig (1989) offer reviews of the literature. In the text primarily references are to recent examples. Particularly the literature on absolute cost advantages (process innovations) and product differentiation (product innovations) is expanding very rapidly. Innovative activity is an important source of barriers to entry (DeBondt 1977, Stiglitz 1987 and Van Witteloostuijn and Maks 1988). Here space limitations rule out a review of the literature on innovation. For example, Kamien and Schwartz (1982), Scherer (1984), Stiglitz and Mathewson (1986), Reinganum (1989) and Chapter 8 offer surveys of models of innovative activity.

³ This series of (references to) entry barrier sources is of course only a sample. Other examples are presented in Schmalensee and Willig (1989). In particular, Gilbert's (1989b) contribution to these books is worth noting.

of crucial importance. Dixit (1979) describes the models with small (potential) entrants as follows (see also Encaoua *et al.* 1986):

In the case where the prospective entrants are a fringe of small price-taking firms, the established firms are assumed to calculate their residual demand, given the entrants' supply curves, and then act as to maximize their own profit. The fringe supply is thus a reaction function, and the established firms are acting as von Stackelberg leaders, allowing entry to the extent that suits their own best interests. This seems a proper solution concept, and can be extended to sophisticated dynamic settings (Dixit 1979, p. 20).

A well-known example of a dynamic model with small entry is Gaskins (1971).¹

The core of the literature on entry barriers is, however, concerned with the implications of (the threat of) large scale entry. Large scale entry, contrary to small scale entry, does not leave prices and/or quantities unchanged. Therefore, assumptions as regards (expectations of) post-entry price and/or quantity reactions of rivals are of crucial importance. The assumptions that were first imposed are the "Sylos' postulate" and the "excess capacity hypothesis".

The *Sylos' postulate* originates from the Italian economist Sylos-Labini (1957). Modigliani (1958) is responsible for its introduction (and modification) into the English literature. The Sylos' postulate implies that

potential entrants behave as though they expected existing firms to adopt the policy most unfavorable to them, namely, the policy of maintaining output while reducing price (or accepting reductions) to the extent required to enforce such an output policy (Modigliani 1958, p. 217).²

This assumption closely resembles Bain's (1956, p. 28). Both Sylos-Labini and Modigliani assume that incumbents set limit prices so as to preclude entry altogether. The result of the model is that

[u]nder Sylos' postulate there is a well-defined maximum premium that the oligopolists can command over the competitive price, and this premium tends to increase with the importance of economies of scale and to decrease with the size of the market and the elasticity of demand (Modigliani 1958, p. 220).³

Kamien and Schwartz (1971), among others, introduce dynamics in Sylos-Labini's and

¹ The tools employed in almost the entire literature are derived from game theory. Therefore, the adjective dynamic must be seen from a game-theoretic perspective. This means that if the time horizon of the players exceeds one period (or one point in time) and if lags are taken into account, a model is considered dynamic. For example, Spulber (1981) presents "a two-period model of entry in which a monopolist chooses output and capacity in the first period with the knowledge that another firm will enter in the second period" (Spulber 1981, p. 505). Fudenberg and Maskin (1987), Schmalensee (1988) and Tirole (1988) offer brief introductions to the role of game theory in industrial organization.

² Note that the Sylos' postulate does not necessarily imply the strategy most unfavorable to the incumbent firm. The key point is that in many cases other post-entry strategies are more profitable (Dixit 1980 and Spulber 1981).

³ The Sylos' postulate is mainly used in models with scale economies.

Modigliani's static theory.¹ However, the Sylos' postulate suffers from a basic flaw, because just one of the possible post-entry reactions is assumed *a priori*, despite the fact that such behavior is not necessarily profitable to incumbents in all circumstances. Schmalensee (1988) points out that

this argument has a serious game-theoretic problem: the implicit threat to maintain output in response to entry is not credible, since the incumbent (quantity-setting) firm would generally do better to reduce production (Schmalensee 1988, p. 663).

Strategic interactions between incumbent firms and potential entrants are ignored. Moreover, a significant extent of (tacit) collusion is necessary to guarantee that all (price-setting) incumbents set the limit prices.²

Pashigan (1968) introduces the *excess capacity hypothesis* as an alternative to the Sylos' postulate. This hypothesis has been worked out in detail by Wenders (1971), whose

purpose ... is to show [that] excess capacity might be profitably used by established firms to invalidate Sylos' postulate by threatening to increase output if an entrant appeared (Wenders 1971, p. 15).

An incumbent firm with excess capacity disseminates the threat of increasing output (and hence reducing prices) after the appearance of an entrant. Spence (1977) incorporates the excess capacity hypothesis in a static model. Incumbents are assumed to threaten potential entrants to reduce eventually the pre-entry monopoly price to the post-entry competitive price, irrespective of the profitability of such a strategy, by increasing supply with the help of the excess capacity that they possess for that purpose. Spence (1979) elaborates on the dynamics of a model with excess capacity by introducing growing demand. However, the excess capacity hypothesis too may not imply credible post-entry actions: this assumption also imposes just one of the possible post-entry replies upon entry of the incumbent firms on *a priori* rather than (profit) maximizing grounds [although there are scenarios where the installation of excess capacity is the profit maximizing strategy (Gilbert 1989b)].

7.6 STRATEGIC INTERACTIONS

The foundation of (the expectations of) post-entry (re)actions in maximizing decision making introduces strategic interactions into the models. In this context, Eaton and Lipsey (1980) emphasize that Schelling's (1956) distinction between threats and commitments is particularly relevant to the study of entry (deterrence):

¹ Moreover, Kamien and Schwartz respond to an important critique of Stigler (1968) by introducing demand growth, expectations of future rates of return and incumbents' strategies which retard (rather than preclude altogether) the rate of entry. Here it should be noted that the two-polar case, either preventing or freely allowing entry, is not Bain's. Bain (1956) distinguishes four scenarios: blockaded entry, easy entry, ineffectively impeded entry and effectively impeded entry (Dixit 1979).

² Probably (depending on further assumptions as regards demand), the entry threat disciplines the incumbents' price setting such that the result resembles (tacit) collusion. If quantity is the strategic variable, the result may be different, because excess capacity (or supply) is a public good (Caves and Porter 1977 and Gilbert and Vives 1986).

In his seminal essay on bargaining, Thomas Schelling (1956) distinguishes threats and commitments. Both are designed to influence a competitor by impressing him with the consequences of his actions. Both take the same form: 'If you take action X, I shall take action Y, which will make you regret X.' The distinguishing characteristic of a threat is that the actor has no incentive to carry out action Y either before or after action X. The distinguishing characteristic of a commitment is that, X having occurred, it is in the actor's self-interest to take action Y. A fanatic may carry out a threat, and may thus make 'credible threats'. A maximizer would not carry out his threats and thus cannot make his threats credible. There would seem, therefore, to be little place for threats in maximizing models (Eaton and Lipsey 1980, p. 721).¹

That is, only post-entry actions that are profitable to incumbent firms are taken seriously by potential entrants.

Dixit (1979) drops the assumptions of *a priori* given post-entry reactions in favor of profit-maximizing decision making. In a static setting incumbent firms can choose to either maintain (Sylos' postulate) or increase (excess capacity hypothesis) output after entry, depending on the profitability of both strategies. Dixit (1980) extends the set of possible post-entry strategies that are available to incumbent firms. Spulber (1981) shows in a dynamic framework that the Sylos' postulate and the excess capacity hypothesis constitute just two extreme examples of entry-detering strategies.² A particular strategy is only optimal in special circumstances. Perrakis and Warskett (1983) introduce uncertainty of demand and rivals' utility functions in a dynamic model with strategic interactions.³ That is, threats of post-entry behavior have to be credible. Fudenberg and Tirole (1983a) and Thepot (1989) are two further examples of sophisticated models along these lines. The key point is Schelling's (1960) argument that

a threat can be made credible by entering into a prior commitment that made its fulfillment optimal ex post (Dixit 1982, p. 13).⁴

The introduction of incomplete information and signaling devices can further extend the set of feasible entry-detering strategies. Salop's (1979) argument that firms may

¹ Selten's (1975) concept of subgame perfectness takes account of the credibility of threats in game-theoretic (equilibrium) terms.

² Moreover, Fudenberg and Tirole (1984), Bulow *et al.* (1985a) and DeBont and Veugelers (1989), among others, have shown that in particular circumstances the incumbent's optimal strategy is to *underinvest* rather than to *overinvest* [relative to the nonstrategic (or open loop) monopoly solution].

³ Moreover, there is some related literature on *actual* capacity - price rivalry. These models generally describe a two-stage game. The first stage examines capacity choice and the second Bertrand price setting (for example, Kreps and Scheinkman 1983, Davidson and Deneckere 1986 and Osborne and Pitchik 1986). Depending on the assumptions made [in particular, as regards to demand (rationing)], the Nash (perfect) equilibrium can be shown to approach (in the extreme) either the Bertrand or Cournot result.

⁴ An uncertain market environment may undermine the strategic value of commitments. Vives (1989) points out that "[r]eceived common wisdom provides us with two presumptions" (p. 386): first, "[m]ore uncertainty will induce firms to seek more flexible technological positions" (p. 386) and, second, "[u]ncertainty will dilute the strategic commitment value of the technological choice" (p. 386). For example, Appelbaum and Lim (1985) offer arguments in favor of both presumptions. However, Vives (1989) presents scenarios which reveal opposite results. In particular, "we have concluded that an increase in uncertainty always raises the value of strategic commitment for risk-neutral firms. Therefore the suspicion that the importance of strategic commitment may fade under uncertainty is shown not to be well founded. This, we think, adds to the robustness of the results obtained in the certainty literature on capacity commitment" (Vives 1989, p. 407).

purposely manipulate the (price) signals upon which (potential) rivals (partly) base their decisions, has induced

[a] number of recent studies [that] have argued that potential entrants might well attach some positive probability (assumed of course to be common knowledge) to the possibility that an established monopoly is irrational - that it will always prey on entrants regardless of the costs. Then a rational established firm facing a finite set of potential entrants will often find it optimal to prey on the first few entrants in order to build (or, more precisely, to avoid destroying) a useful *reputation* for irrationality (Kreps and Wilson, 1982). With incomplete information, predation may also serve to lower the cost to the predator of acquiring the prey (Saloner, 1987) (Schmalensee 1988, p. 665; Schmalensee's italics).¹

In particular, Kreps and Wilson (1982) emphasize the crucial role which reputation considerations may play in games with imperfect information.²

The existence of reputation and signaling effects points to the crucial question of the *credibility* of entry-detering strategies (Gilbert 1986). As Encaoua *et al.* (1986) put it:

It is clear that when we allow pre-play communication, we introduce the possibility of strategic action through pre-commitment, and so must face the problem of distinguishing between empty and non-empty threats (Encaoua *et al.* 1986, p. 58).

Milgrom and Roberts (1982) argue that a threat may imply a commitment, even when a post-entry price war is unprofitable in the short run for the incumbent firm, as a result of the influence of reputation:

Economists often argue that predatory practices are irrational, since there exist cheaper or more certain means to gain or maintain a monopoly. Our game-theoretic, equilibrium analysis suggests that if a firm is threatened by several potential entrants, then predation may be rational against early entrants, even if it is costly when viewed in isolation, because it yields a reputation which deters other entrants. Asymmetric information plays a crucial role in our analysis, since it provides the rationale for entrants to base their expectations of firm's future behavior on its past actions (Milgrom and Roberts 1982, p. 280).

Harrington (1986) responds to Milgrom and Roberts' argument by asserting that the result, under particular circumstances, can be reversed. Harrington's argument is based upon the alternative assumption that the potential entrant is uncertain of its post-entry cost of production and selling. In combination with the belief that firms' cost functions are positively correlated, incumbent firms' high prices signal a high cost level, which reduces the entry incentives. If the correlation is (expected to be) sufficiently strong, high prices signal low post-entry profitability for an entrant. So, Milgrom and Roberts' postulate of asymmetric information³ may yield low as well as high limit prices, depending on the

¹ Note that this literature uses a specific (strict) interpretation of the rationality concept. In this literature rationality means profit maximization. A reputation for *short-run* irrationality (i.e., deviation from short-run profit maximization) can facilitate *long-run* rationality (i.e., long-run profit maximization).

² In this way, Kreps and Wilson (1982) solve Selten's (1978) "Chain Store Paradox" (although not in Selten's setting with perfect information).

³ However, "[u]nfortunately, since unobservable beliefs play a critical role in reputation models, these models place relatively weak restrictions on observed behavior; they imply the potential rationality of predation under almost any observable conditions" (Schmalensee 1988, p. 665).

informational assumptions made (compare with Riordan 1985, Arvan 1986, Roberts 1986 and 1987 and Milgrom and Roberts 1987).

A key point is that the manipulation of information can be used to reduce the attractiveness of entry. The strategic feature of entry deterrence through signaling can take many shapes. Stigler (1966), Smith (1981) and Smiley (1988) point to three illustrative examples. Stigler (1966) observes that incumbent firms with divisions can avoid revealing the profitability of operations in specific areas to make the attractiveness of entry less apparent. Smith (1981) argues that the announcement of strategic investment projects can be employed so as to discourage potential entrants. Smiley (1988) indicates that incumbent firms can keep new products or processes secret as long as possible.

Credible threats of post-entry competition - being signaled through the manipulation of information (in the form of commitments, prices, announcements, *etc.*) - are important entry-detering devices. The success of entry deterrence critically depends on (the manipulation of) the potential entrants' expectations of post-entry competition.¹ However, threats of post-entry toughness (Fudenberg and Tirole 1984) are not Stiglerian entry barriers *per se*. In effect, even if Stiglerian entry barriers (*i.e.*, potential entrants' cost disadvantage) are absent, credible threats of unprofitable post-entry competitions can keep potential entrants out.² The study of entry barriers and entry deterrence cannot do without the analysis of the impact of the incumbent firms' threatening strategies.

7.7 MARKET PERFORMANCE

The implications of entry barriers for market performance cannot be established in a straightforward way. For example, Schmalensee (1988) points out that, (see also Gilbert 1989b, pp. 528-530),

formal models of imperfect competition rarely generate unambiguous welfare conclusions. In such models, feasible policy options usually involve movements *toward* but not *to* perfect competition, so that welfare analysis involves second-best comparisons among distorted equilibria. In particular, there is no guarantee that making markets 'more competitive' will generally enhance welfare, particularly if non-price rivalry is intensified (Schmalensee 1988, p. 677; Schmalensee's *italics*).

This section offers a brief illustration of Schmalensee's observation in the context of a welfare evaluation of entry barriers. More is said about the welfare implications of market forms in subsequent chapters (in particular, Chapter 24).

The easy part is the effect on price, which follows directly from the definition of barriers to entry. In terms of Stigler (1968) entry barriers imply that potential entrants face a cost disadvantage: $atc <^* atc$. The margin between price and average total cost (bm), $bm = p - atc$, indicates the height of barriers to entry. For $bm = 0$ there is no entry barrier. For $bm > 0$ (*i.e.*, if an entry barrier is present) the profit of an incumbent

¹ Moreover, strategic interactions *among* potential entrants may matter (Sherman and Willett 1967, Goldberg and Moirao 1973, Kalish *et al.* 1978 and Nti 1989). That is, the potential entry threat may correlate *negatively* with the number of potential entrants.

² Threats can be entry barriers in the sense of Bain. This is particularly clear from Gilbert's (1989b) definition of entry barriers, which proceeds along the lines of Bain's. Gilbert "takes the view that a *barrier to entry is a rent that is derived from incumbency*. It is the additional profit that a firm can earn as a sole consequence of being established in an industry" (Gilbert 1989b, p. 478; Gilbert's *italics*).

firm, $\pi = q \cdot p - mc$, is positive (ignoring time and discounting). Price exceeds average and marginal cost: $p > mc$. Hence, the essential condition of static efficiency is violated. The literature suggests that in a static and partial context entry barriers imply static inefficiency.

However, this is not to say that entry barriers necessarily introduce a welfare loss. The fact that barriers to entry can be closely related to dynamic efficiency reflects the snake in the grass. The welfare implications depend critically on the source of the entry barrier. In the literature the normative analysis of entry barriers is generally carried out with the help of a welfare function that consists of a producers' and consumers' surplus (Reid 1987). Profits are usually regarded as the producers' surplus, whereas the consumers' surplus follows from a utility function.¹ Here it suffices to note that (i) the producers' surplus increases with price and quantity and decreases with average (total) cost and (ii) the consumers' surplus generally increases with product quality, variety of products and quality of information and decreases with price and transaction cost.

The producers' surplus follows from the profit of incumbents and entrants. On the one hand, from the incumbent firms' point of view entry barriers clearly, *ceteris paribus*², raise welfare. Zero entry barriers are associated with free entry and so with a tendency to profit dissipation, whereas barriers to entry are so defined that they yield a positive profit. The higher the barrier to entry, the higher the profit rate that the incumbent firms are facing.³ On the other hand, entry barriers generally imply a reverse welfare effect on potential entrants, since barriers to entry restrict entry (and so profit) opportunities.

From the consumers' perspective the source of the entry barriers is crucial.⁴ Take one typical example: advertizing and goodwill. Advertizing has a good and a bad side. Nelson (1975) argues that

there is much, indeed, in advertising that is informational: price advertisements, help-wanted advertisements, and pictures of dresses and furniture available in stores. ... But surely ..., this is not the whole story. What about the frequent endorsement of a brand by announcers, actors, or celebrities - all of whom are paid for their efforts? It is advertisements of this character that have produced much of the skepticism about advertising's information role (Nelson 1975, p. 213).

So, advertising may introduce countervailing effects: if advertising raises an effective entry barrier, the limit price is increased, but transaction cost is lowered and information is increased.⁵ If advertising has a positive influence on price, the increased price may raise the incumbent firms' profits, but lowers the consumers' surplus. The net effect of

¹ Consumers' surplus is often derived from the area under the average revenue curve above a given price (Reid 1987).

² If actual rivalry among incumbent firms is such that zero profits are earned (perfect or Bertrand competition), then entry barriers make no difference. However, then there is no incentive to enter anyway. Here this case is ignored.

³ Conditional of course upon the assumption that potential competition dominates actual rivalry (Shepherd 1984).

⁴ This is, for example, immediately clear from the extreme case where the incumbent firms' entry-detering strategies raise the rivals' costs (compare with Salop and Scheffman 1983).

⁵ Another example of a normative evaluation of entry barriers that received and still receives much attention is the literature on innovation. This literature, which has been inspired by Arrow's (1962) seminal contribution, also shows the complexity of this kind of study (for example, Dasgupta and Stiglitz 1980a and Spence 1986). Chapter 8 offers a survey.

advertising on the consumers' surplus depends upon the quality of the information. So, the debate centers on the quality of the information that is embodied in advertisements (Comanor and Wilson 1979, pp. 472-473). Informative and noninformative advertisements have a different impact (Schmalensee 1978 and 1986).

In effect, as with the case with entry barriers through informative advertising, the tradeoff between static and dynamic efficiency is generally at stake. Entry barriers facilitate the profit rate of incumbent firms, but restrict the entry opportunities of potential entrants. The impact on consumers' utility critically depends on the source of the barrier to entry.¹ Von Weizsäcker (1980a and 1980b) seeks a way out of this dilemma by redefining entry barriers:

We then define barriers to entry in the following way, building on Stigler's definition: a barrier to entry is a cost of producing which must be borne by a firm which seeks to enter an industry but is not borne by firms already in the industry and which implies a distortion in the allocation of resources from the social point of view (Von Weizsäcker 1980a, p. 400).

However, the definition that entry barriers are socially undesirable is empty if it is not made clear what is socially desirable and what is not (Maks 1986, p. 119). So, a critical analysis of the aspects of dynamic competition involved is still needed. Here Von Weizsäcker (1980a and 1980b) provides some nice examples of Schmalensee's (1988, p. 677) observation by showing that entry may either be insufficient (*i.e.*, entry barriers too high) or excessive (*i.e.*, entry barriers too low), depending on the specific circumstances of the case at hand. For example, Von Weizäcker shows

that there are plausible parameter configurations for both economies of scale and goodwill (which is a variant of product differentiation) under which welfare would be improved by increasing, rather than decreasing, the protection of incumbents from the competition of entrants. Greater attention to detail in the analysis of industry circumstances and greater caution in reaching 'obvious' welfare conclusions are needed (Von Weizäcker 1980a, p. 399).

So, the fact of the matter is that there is no simple answer to the question of the welfare implications of barriers to entry.²

7.8 ENTRY BARRIERS, IMPERFECT COMPETITION AND THE SOURCE OF POTENTIAL ENTRY

Before concluding the argument, two final remarks are in order.³ First, this chapter has

¹ With respect to the welfare implications of entry barriers for consumers, an extreme argument can be pursued: "in a less-than-perfect world, what is called an entry barrier may actually reflect consumers' interests, when broadly considered (Demsetz, 1982 ...). Rather than allowing resources to move into high-profit industries, it might be better from the point of view of consumer welfare more generally to take into account the role of externalities, information and transaction costs. One could then consider entry barriers to be a valuable second-best answer to real world frictions" (Encaoua *et al.* 1986, p. 74).

² The literature on standardization also illustrates this point (Dybvig and Spatt 1983, Katz and Shapiro 1985, Farrell and Saloner 1986 and Besen and Johnson 1986). Standardization is associated with welfare costs as well as benefits.

³ Moreover, the empirical literature on barriers to entry cannot be ignored. The number of empirical studies on entry barriers is very large (Lyons 1988). Three methods of research dominate the literature (Reid 1987). First, following Bain (1951 and 1956), cross section data are used to compare (on the basis of

dealt with the theory of entry barriers, but by doing so the theory of imperfect competition has been touched upon as well. Imperfect competition in a market can only exist by the grace of barriers to entry in the sense that *imperfect competition* is incompatible with free entry. As Stigler (1968) puts it:

The *sufficient* conditions for competition [i.e., static efficiency] are unlimited numbers and perfect knowledge and divisibility, for market competition; and absence of all barriers to entry for industrial competition (Stigler 1968, p. 16; Stigler's italics).¹

Along the same lines Waterson (1984) argues that

we may say that a necessary (but not sufficient) condition for the number of firms in an industry to affect the level of profits in that industry is that entry is not completely free, costless and quick (Waterson 1984, p. 56).

statistical analyses) the entry barriers in two or more industries at one point in time. Second, Comanor and Wilson (1967) gave the impetus to time series studies of the evolution over time of barriers to entry in one particular industry with the help of econometric techniques. Third, case studies, as already propagated by Mason (1939), can be used to focus on a specific example of an entry barrier. Broadly speaking, empirical research offers diverse results. Take the following five examples. (i) Hannan (1983) concludes that "charging prices less attractive to potential entrants or investing in capacity both appear to be ways in which established firms *may* deter entry" (Hannan 1983, p. 550; Hannan's italics). (ii) Sengupta *et al.* (1983) present empirical results which "suggest the plausibility of the limit price model for the computer industry for the period analysed here" (Sengupta *et al.* 1983, p. 298). (iii) Masson and Shaanan (1986) argue that a "small cross-section sample of national industries was available. From this sample it appears that entry is in fact deterred by lower prices, higher excess capacity and/or greater entry barriers. But a simultaneous test of excess capacity and pricing leads to little evidence that firms strategically add capacity to deter entry" (Masson and Shaanan 1986, p. 373). (iv) Lieberman (1987) points out that the "empirical evidence in this study suggests that the excess capacity entry barriers identified in the theory are not very common in practice. While significant excess capacity was held by firms in the sample, most was maintained to accommodate demand variability and investment lumpiness. The statistical tests fail to show that incumbent firms expanded pre-emptively in an effort to deter entry" (Lieberman 1987, p. 624). (v) Lyons (1988) draws some conclusions out of a survey of the empirical literature, one being that "[l]ittle evidence is available to either support or deny the excess capacity hypothesis" (Lyons 1988, p. 63). A basic flaw of most empirical studies is that they fail to identify the structural or strategic dimension of barriers of entry (Smiley 1988, p. 167 and Lyons 1988, pp. 63-64). Four studies illustrate the fact that the identification of *strategic* entry deterrence requires alternative research methods: Biggadike (1976), Smith (1981), Yip (1982a) and Smiley (1988). For example, on the basis of a *questionnaire* Smiley (1988) recently reports that "[f]or new products, firms attempt to limit entry through patent preemption and the creation of product loyalty through heavy advertising expenditures. Limit pricing is used less often. For existing products, firms attempt to limit entry through filling all available product niches, through masking product specific data on profitability, and again through the creation of product loyalty through advertising expenditures. The practice of entry deterrence is surprisingly important to the firms involved. More than half of the respondents reported that attempts to deter entry were comparable in importance to other strategic marketing and production decisions" (Smiley 1988, pp. 177-178).

¹ From the contestable market theory it is clear that these conditions are sufficient, not necessary. Bertrand competition and contestable markets achieve competition even with a limited number of rivals. As a matter of fact, Stigler continues his argument in the following way: "this note will argue (but not rigorously prove) that a large number of rivals is sufficient to achieve competition" (Stigler 1968, p. 17). Perfect contestability is the current pillar of the literature on free entry, since its basic assumption is that entry and exit barriers are absent (Baumol 1982 and Baumol *et al.* 1982). Baumol's (1982) introduction of the contestability theory gave rise to an interesting and illuminating discussion on the underlying assumptions and empirical relevance of free entry conditions (for example, Shepherd 1984, Schwartz 1986 and Chapter 6), which is also interesting from the perspective of the literature on entry barriers.

Without entry barriers, imperfect competition and high profit in a market generally induce entry (industrial competition) that generates profit dissipation. Behind protecting (Stiglerian) entry barriers monopolistic and oligopolistic rivalry may yield a positive profit.

Second, entry barrier models generally ignore the issue of the *source of potential entry*. As in contestability theory (Subsection 6.5.4), by and large industrial organization's models of entry deterrence leave potential entrants unidentified (Caves and Porter 1977). However, the source of potential entry may be crucial in the evaluation of the credibility of the entry threat (Part V). The usual approach is to assume that a pool of potential entrants (or at least one potential entrant) exists, where a potential entrant is inclined to enter if entry is perceived to give a positive (however small) profit. This implies an assumption of a zero opportunity cost of entry. Particularly Chapter 21 critically evaluates (the implications of) this assumption.

AMERICAN MEDICAL ASSOCIATION
PUBLISHED WEEKLY
CHICAGO, ILL., U.S.A.

Volume 10, No. 1, January 1917
Subscription price, \$5.00 per annum in advance
Single copies, 15 cents
Entered as Second-Class Matter, May 2, 1902
Postpaid
Acceptance for mailing at special rate of postage provided for in Act of October 3, 1917
Authorized by Act of October 3, 1917
Copyright, 1917, by American Medical Association
Printed at the American Medical Association, 535 North Dearborn Street, Chicago, Ill.

CONTENTS
Original Articles
The Medical Profession and the Public
The Medical Profession and the Public
The Medical Profession and the Public

Editorial
The Medical Profession and the Public
The Medical Profession and the Public
The Medical Profession and the Public

Correspondence
The Medical Profession and the Public
The Medical Profession and the Public
The Medical Profession and the Public

Obituary
The Medical Profession and the Public
The Medical Profession and the Public
The Medical Profession and the Public

Index
The Medical Profession and the Public
The Medical Profession and the Public
The Medical Profession and the Public

Advertisements
The Medical Profession and the Public
The Medical Profession and the Public
The Medical Profession and the Public

Subscription Service
The Medical Profession and the Public
The Medical Profession and the Public
The Medical Profession and the Public

Index
The Medical Profession and the Public
The Medical Profession and the Public
The Medical Profession and the Public

Index
The Medical Profession and the Public
The Medical Profession and the Public
The Medical Profession and the Public

CHAPTER

8

FIRMS' R&D OUTLAYS: Process and Product Innovations

8.1 INTRODUCTION: A Crucial Dynamic Economy and Entry Barrier

A survey on innovative activity is included for three reasons. First, perhaps the most striking manifestation of dynamic efficiency of market performance is technical change. As Kamien and Schwartz (1982) put it:

It hardly seems necessary these days to point out the importance of technical advance. We look to it to rescue us from the consequences of exhausting essential natural resources; abate inflation through productivity increases; improve our balance of payments deficit; eliminate famine; and cure cancer, heart disease, and a variety of other ailments. Our faith in technical advance is bolstered by achievements such as the atomic bomb, electronic computers, the landing of a man on the moon, heart transplants, and test-tube babies. We no longer ask if something is possible, but how soon it can be done and at what price (Kamien and Schwartz 1982, p. 1).

Second, this chapter concurs with Chapter 7, because innovative activity is also an important source of entry barriers (DeBontdt 1977, Dasgupta and Stiglitz 1980a and Dasgupta 1986). Third, specific barrier market models (in particular, Part VII) focus on innovation.

The study of technical advance gained enormous popularity in the decades of fast technical change that followed upon World War II.¹ Therefore, it is not surprising that this chapter cannot offer more than a caricature of an all-embracing survey by referring to a (random) sample of the (micro)economic models of innovative activity. This implies that only minor attention is paid toward such important phenomena as the nature of invention and innovation, the cyclical nature of technical development and empirical research.² Moreover, the contributions of the pioneer economists are bypassed, since they generally assumed technological progress to be determined exogenously. It should,

¹ Here the literature that puts forward technical change as an important (*ad hoc*) explanation of the high post-war growth rates is illustrative. Solow (1957) and Denison (1962 and 1967) are of course the classic references.

² This means, among other things, that the stages of invention and innovation are not distinguished. For a survey of (some of) these and other topics see, for example, Nelson (1981), Kennedy and Thirlwall (1981), Freeman (1982), Kamien and Schwartz (1982), Nelson and Winter (1982), Stoneman (1983), Scherer (1984), Stiglitz and Mathewson (1986), Dosi (1988), Dosi *et al.* (1988), Davies (1988) and Beath *et al.* (1989).

however, be stressed that they recognize the role of technical change; one need only think of Smith's (1776) argument on the relation between inventions and the division of labor, Ricardo's (1817) famous chapter on machinery and Marx's (1867) well-known emphasis on the destructive role of technological progress. However, the general attitude was to take technology and technical change as given. Walras (1874), for instance, argues that

[i]t should, therefore, be clearly understood that every time the production function itself undergoes a change, we have a case of technical progress brought about by science and that every time the coefficients of production made up of land-services increases without any change in the production function, we have a case of economic progress resulting from saving. In reality, both kinds of progress may take place simultaneously ... But in this discussion we shall abstract from technical progress and consider economic progress only (Walras 1874, p. 386).¹

The founding fathers of the modern theory of innovation are Hicks (1932) and Schumpeter (1943). Hicks introduced the notion of *induced invention* and Schumpeter described the process of *creative destruction*. Schumpeter's theory in particular is the dominant source of inspiration of many outstanding scholars of innovative activity, as expressed in, for example, Kamien and Schwartz (1982), Nelson and Winter (1982) and Scherer (1984). Scherer, for instance, points out that

[t]he innovation theme has been central to much of my work as an economist. I was started on this course at the age of twenty-one, when I was captivated by the writings of Joseph A. Schumpeter, whose centennial we celebrated in 1983. Three decades later I continue to be a Schumpeterian in the most fundamental sense (Scherer 1984, p. viii).²

The survey of the literature is structured along the following lines. Section 8.2 offers a brief introduction to economic thinking on innovative activity from an overview perspective. Section 8.3 pays attention to (microeconomic) models of induced innovations in the Hicksian tradition. Kamien and Schwartz (1982) distinguish two equilibrium approaches to the study of innovation that start from a Schumpeterian perspective. Section 8.4 focuses on decision-theoretic models and the game-theoretic approach. Section 8.5 examines the disequilibrium models of a Schumpeterian evolutionary process. Section 8.6

¹ Van Witteloostuijn and Maks (1989 and 1990) present an interpretation of Walras' contribution to economics by stipulating the fact that Walras had an understanding of the dynamics of economies.

² Dasgupta (1986) points to "the sheer volume of effort that has been spent over the past quarter of a century trying to demonstrate that economic forces are a prime architect of technological change. If there is a single driving force behind these empirical investigations, it is the writings of Joseph Schumpeter (most especially perhaps in his *Capitalism, Socialism and Democracy* ...). Since recent developments in the theory of technological competition have addressed a few of the empirical findings, they reflect this heritage as well. Economists are prone to having hero-figures. In the economics of technological change, Joseph Schumpeter continues to reign as the undisputed godfather" (Dasgupta 1986, p. 519). However, Dasgupta continues his argument by asserting that "[a]s with all such personalities, there is much discussion about what he actually *means*. ... The point is that, like other system-builders, Schumpeter did not write with the utmost clarity" (Dasgupta 1986, pp. 519-520; Dasgupta's italics). This observation is illustrated below, in particular through the contrast between the equilibrium and disequilibrium approach to Schumpeterian economics.

contains remarks on market structure, R&D outlays and market performance.¹ Section 8.7 offers a conclusion.

In advance, two remarks are in order. First, for illustrative purposes, the review of each of the three broad categories of innovation models consists of an indication of the mode of modeling with the help of an example set of equations on the one hand and important references on the other. Second, although there are important differences in emphasis and key elements, the microeconomic models of technical advance have much in common.

8.2 A HICKSIAN AND A SCHUMPETERIAN PERSPECTIVE

A first approach to the study of innovations progresses along Hicksian lines. Hicks' (1932) introduction of purposeful innovative activity by firms settled a research agenda that focuses on the explanation of *induced inventions*. According to Hicks, induced inventions are caused by changes in the relative prices of production factors:

We must put on one side those inventions which are the result of a change in the relative prices of the factors, let us call these "induced" inventions. The rest we may call "autonomous" inventions (Hicks 1932, p. 125).

Moreover, Hicks proposes a classification of inventions in labor-rising, labor-saving and neutral technical changes.² These terms indicate the factor requirement that is reduced in greatest proportion after the introduction of an innovation (for example, Binswanger 1978, p. 20). In the first instance, the follow-up of Hicks' insights mainly consisted of the application of the microeconomic model of induced innovation to macroeconomic theories.³

First, there was a discussion about the validity of Hicks' foundation of the induced invention hypothesis upon changes in the factor prices. Salter (1960) argues that the

entrepreneur is interested in reducing costs in total, not particular costs such as labor costs or capital costs. When labor's costs rise, any advance that reduces total costs is welcome, and whether this achieved by saving labor or capital is irrelevant (Salter 1960, p. 43).

Fellner (1961) points out that it is the anticipation of a continuous rise in the price of the progressively scarce production factor that brings about a labor-saving or labor-rising bias in technical advance.

Second, there was the introduction of macroeconomic models of induced inventions. Kennedy (1964) presents a macroeconomic model which is popular in growth theory. Two other well-known growth models with macro-induced innovations are Samuelson (1965) and Drandakis and Phelps (1966). An important feature of these models is the assumption of an innovation possibility frontier (IPF). The IPF describes all efficient

¹ For example, Kamien and Schwartz (1982), Scherer (1984) and the papers in Stiglitz and Mathewson (1986) offer impressive accounts of the current state of the art in the economics of technical advance and technological competition.

² Although Harrod and Solow have introduced a classification as well, Hicks' is the one most often employed.

³ Vintage models (for example, Van de Klundert and De Groof 1977, Kuipers and Van Zon 1982, Gelauff *et al.* 1985 and Muysken and Van Zon 1987) only partially fit into the macroeconomic Hicksian tradition as these models focus on the diffusion, rather than innovation, question.

potential innovations that an economy is facing at a particular point in (or period of) time. Wan (1971) notes that it is implicitly assumed that the *ex ante* IPF is identical to the frontier of actual innovation possibilities. So, unless the firms know the IPF, the approach breaks down. The IPF-models are associated with a (tacit) assumption of complete and perfect knowledge of the innovation possibilities. In response to a change in relative factor prices, firms choose a new optimal production technique. This may yield a labor-saving or labor-rising bias in aggregate technical advance. Ahmad's (1966) historic innovation possibility curve is an IPF like concept. Ahmad permits the curve to shift over time as a result of progress in basic research.

Third, there was a response to the macroeconomic modeling from a microeconomic perspective. A basic flaw of the macroeconomic models of induced invention is that they lack a microfoundation, which induced microeconomic modeling (for example, Lucas 1967, Binswanger 1974 and 1978, Radner 1975 and Evenson and Kislev 1975). The microeconomic models try to assimilate the critique of the Kennedy approach that, for example, (i) technical change comes about without cost and (ii) expectations are ignored (Binswanger 1978, pp. 35-38). For illustrative purposes, Binswanger's (1978) Hicksian microeconomic model is discussed in Section 8.3.

The major part of the research on innovative activity has been (and still is) done from a Schumpeterian perspective. Following Marx (1867), Schumpeter (1943) presented an evolutionary theory of the working of the capitalist system. He described an evolutionary disequilibrium process that is driven by the innovation-inducing forces within capitalist society. In his short (five-pages) chapter "The Process of Creative Destruction" (1943, pp. 81-86; for important quotations see Chapter 5) Schumpeter stipulated the importance of evolution, market structure, long-run performance, potential competition and nonprice competition as the essential features of an accurate description and analysis of the economic process in general and technical advance in particular. These key elements of the Schumpeterian view of the economy have generated an impressive body of literature on innovative activity. The Schumpeterian research program can be divided into three parts.

First, much empirical research¹ has been done (and still is done) so as to test two important so-called "Schumpeterian hypotheses":

Two broad hypotheses are associated with Schumpeter: (1) There is a positive relationship between innovation and monopoly power with the concomitant above normal profits. (2) Large firms are more than proportionately more innovative than small firms (Kamien and Schwartz 1982, p. 22).

As Kamien and Schwartz note, Galbraith (1952) should receive substantial credit for the second hypothesis. The Schumpeterian hypotheses have been refined in two ways by the introduction of the "technology-push hypothesis" (Nelson 1959 and Phillips 1966) and the "demand-pull hypothesis" (Schmookler 1966).² Kamien and Schwartz (1975 and 1982)

¹ Nice summaries of the results of empirical studies are provided by Kamien and Schwartz (1982, Chapter 5), Dasgupta (1986, pp. 522-527) and Cohen and Levin (1989).

² Of course, the technology-push and demand-pull explanations need not to be contradictory (for example, Gort and Wall 1986).

offer a summary of the results of the empirical investigation (see also Dosi 1988).¹ Three more or less general tendencies appear²: (i) the relationship between profit and innovation seems to be significant and bidirectional; (ii) innovative activity does not typically increase faster than firm size; and (iii) little support has been found for the hypothesis that R&D outlays increase with monopoly power. More evidence has arisen that an intermediate market structure promotes the highest rate of innovative activity.

Second, a significant body of literature has been (and still is) shaped so as to offer Schumpeterian microeconomic models of firms' decision making as regards to R&D expenditures. Kamien and Schwartz (1982) characterize these models as follows:

In the modern theories of market structure and innovation competition among potential innovators is thought of as a race to be the first. ... the main question addressed in the modern theories are how profits, costs and the 'intensity of rivalry' determine the speed of development of an innovation. The 'intensity of rivalry' is related to market structure and to the number of rivals but is identical to neither. That is, a potential innovator may feel threatened, as Schumpeter claimed, even if he appears to be alone (Kamien and Schwartz 1982, pp. 105-106).

The equilibrium-type of models can be classified as to the mathematical tools employed. Starting with Scherer (1966 and 1976) and Barzel (1968), models were (are) built with the help of decision theory. However, these models fail to take account of the interactive effect between rivals. The game-theoretic approach (for example, Loury 1979 and Gilbert 1986) seeks to assimilate this critique.³ Section 8.4 reviews some of the Schumpeterian

¹ Empirical research on Schumpeterian topics goes on continuously. Post-1982 references are, for example, Kaplinsky (1983), Loeb (1983), Bosworth and Westaway (1984) and Cohen *et al.* (1987). Cohen and Levin (1989) summarize the disappointing state of the art of the results of empirical research on the Schumpeterian hypotheses: "The empirical results concerning how firm size and market structure relate to innovation are perhaps most accurately described as fragile. The failure to obtain robust results seems to arise, at least in part, from the literature's inadequate attention to the dependence of these relationships on more fundamental conditions" (Cohen and Levin 1989, p. 1078). For example, the "most notable feature of this considerable body of empirical research on the relationship between firm size and innovation is its inconclusiveness" (Cohen and Levin 1989, p. 1069).

² However, some snakes in the grass prevent the happy state of consensus. For illustrative purposes, two snakes are mentioned. Following, for example, Scherer (1979 and 1984), Angelmar (1985) modifies the direct causality that flows from concentration to rate of innovation. On the basis of an empirical study he concludes: "Where the cost and uncertainty of R&D are high and conditions favor speedy imitation by competitors, concentration has a statistically significant and substantial positive impact on research intensity. But its impact is clearly negative when R&D and uncertainty are low, and when high barriers to imitation exist" (Angelmar 1985, p. 69). Kaplinsky (1983) argues that empirical research on the influence of firm size on inventive activity generally takes place with the help of a (largely) static framework. This line of research yields a set of contrasting views on the sign of the correlation, varying from clearly positive (for example, Soete 1977 and 1979) to clearly negative (for example, Rothwell and Zegveld 1982). According to Kaplinsky the lack of consensus can be explained by the dynamic character of technical change. Kaplinsky's data show that "it is not possible to determine whether large or small firms are inherently more innovative since it depends largely on which period of industry life-cycle one is measuring" (Kaplinsky 1983, p. 57).

³ Dasgupta (1986) argues that Kamien and Schwartz (1982) "catalogue modern approaches to theories of market structure and innovation according to whether they are *decision-theoretic* or *game-theoretic*. It transpires that by the former they mean investigations that study a firm's response to an exogenously given (or evolving) market environment and by the latter those which allow firms to interact strategically. But these are not different approaches at all. The former simply asks a more restricted set of questions" (Dasgupta 1986, p. 526; Dasgupta's italics).

equilibrium models.

Third, in the mid-1970s an important line of research was initiated by Nelson and Winter. The focus of this approach is on the disequilibrium aspects of (innovation) competition that are based upon Schumpeter's evolutionary notions. Nelson and Winter (1974) introduce the essentials of the evolutionary theory by arguing that

[f]or Schumpeter the most important firms are those that serve as the vehicles for action of the real drivers of the system - the innovating entrepreneurs. Firms (and entrepreneurs) may seek profit, and may innovate or imitate to achieve higher profit. However, the emphasis on careful calculation over well-defined choice sets is absent. The competitive environment within which firms operate is one of struggle and motion. It is a dynamic selection environment, not an equilibrium one. The essential forces of growth are innovation and selection (Nelson and Winter 1974, p. 890).

During the 1970s and 1980s Nelson and Winter continued to carry out research in the evolutionary tradition (Nelson and Winter 1982). Their work and related contributions (for example, in the classical tradition) are briefly examined in Section 8.5.

8.3 HICKSIAN MODELS OF INDUCED INVENTION

The essential feature of the Hicksian models of induced invention is that the direction of innovative efforts depends on (expected) changes in the relative factor prices (see, for example, Clarkson and Miller 1983, pp. 389-391). What matters is the *relative* (not the absolute) use of the production factors. Let K_t and L_t denote the scale of capital and labor respectively in use before ($t = 0$) and after ($t = 1$) the innovation. If $K_0/L_0 = K_1/L_1$, the innovation is Hicksian neutral; $K_0/L_0 > K_1/L_1$ indicates a capital-saving innovation; and $K_0/L_0 < K_1/L_1$ points to a labor-saving change in technology (provided that w/k is constant, where w denotes wage rate and k the user cost per unit capital). The question is what influence a change in the relative factor prices has on innovation.

The answer to this question was first provided at an aggregate level (Section 8.2). A basic flaw of the macroeconomic models of induced invention is that they lack a microfoundation. As Binswanger (1978) notes:

The failure of the Kennedy approach clearly is a consequence of the theory's lack of a microeconomic foundation (Binswanger 1978, p. 38),

which, according to Nordhaus (1967), implies that the Kennedy-like

model is then just a disguised version of the neoclassical model with exogenous technical change (Nordhaus 1967, p. 64).

Technological change is just imposed through the assumption of an aggregate innovation possibility curve. Therefore, Lucas (1967), Radner (1975), Binswanger (1974 and 1978) and Sato and Suzawa (1983), among others, present models that include a microfoundation of the Hicksian factor price-biased induced invention hypothesis.¹

¹ Moreover, as Binswanger (1978) points out, the "evolutionary models of economic behavior and technical change developed by Richard Nelson and Sidney Winter were not initially intended to be induced innovation models. Nevertheless, even their earliest models contain a simple induced innovation mechanism"

The microeconomic modeling of induced inventions tries to accommodate the critique of the macroeconomic theory of Kennedy (1964) and his successors. The microeconomic models therefore introduce the (maximizing) decision rules of firms into the framework. The firms' decision making on technology in response to changes in the (relative) factor prices is explicitly modeled. So, by modeling the firm's decision making on R&D a microfoundation is framed in the sense that the incentives (that is, stimulating antecedents) of induced invention are incorporated. Since the models in Parts VI to X of this thesis fit into the Schumpeterian tradition (that is, the role of factor prices is ignored), here it suffices to outline briefly, for illustrative purposes, just one example of a microeconomic model with induced invention. The exemplary model is Binswanger (1978).

Binswanger follows Evenson and Kislev (1975) in presenting a stochastic model of applied research so as to endogenize the introduction of innovations in a microeconomic setting. It is assumed that basic research determines (exogenously) the distribution of the potential payoffs to applied research. Applied research is assumed to be subject to diminishing returns.¹ Moreover, it is supposed that firms are risk neutral and maximize expected profit (or, to be precise, minimize expected cost). The Hicksian hypothesis is introduced by assuming that

research administrators or firms have a choice between several research activities, each of which affects the factor-intensity characteristics of the production process in a different way [so that] we can set up the model as a portfolio-choice model of research projects (Binswanger 1978, p. 98).

To simplify matters, the set of possible research activities is restricted to contain only two alternatives:

In the formal model we will consider only one relatively capital-saving research process m and one labor-saving process n , with the choice of bias achieved by linear combinations of the two processes (Binswanger 1978, p. 98),

where the focus is on process innovations.²

The impact of R&D on the production technology is framed with the help of a Markov process model. The outcome of R&D input in period t redefines the change in the production parameters in period $t+1$. Let ΔA and ΔB denote the change in the capital-augmenting and labor-augmenting coefficients of the production technology respectively. Then, the invention possibilities are assumed to be

(Binswanger 1978, p. 29). See Section 8.5.

¹ It must be stipulated that this is a common assumption (Kamien and Schwartz 1982, pp. 64-70 and 194, and Dasgupta 1986, p. 523).

² Binswanger *et al.* defend this assumption by arguing that, "since all development of new intermediate products can be viewed as process innovation from the standpoint of final production, the emphasis on process innovation is clearly appropriate" (Binswanger *et al.*, p. 22). This is the common view. As Spence (1986) puts it, for example: "In many markets, firms compete over time by expending resources with the purpose of reducing their costs. Sometimes the cost-reducing investments operate directly on costs. In many instances, they take the form of developing new products that deliver what customers need more cheaply. Therefore product development can have the same ultimate effect as direct cost reduction. In fact, if one thinks of the product as the services it delivers to the customer (in the way that Lancaster pioneered), then product development often is just a cost reduction" (Spence 1986, p. 475).

$$(8.1) \quad \begin{cases} \Delta A = \mu(m)\alpha^m + \mu(n)\alpha^n, \text{ and} \\ \Delta B = \mu(m)\beta^m + \mu(n)\beta^n, \end{cases}$$

where μ is the common scale function and α and β are the research productivity parameters that determine the factor-saving features of research line m and n .¹ (That is, for $\alpha^m = \alpha^n$ and $\beta^m = \beta^n$ invention possibilities are neutral.)

Suppose that the savings of production cost (ΔC) follow from

$$(8.2) \quad \Delta C = C_t - C_{t+1} = c_K \cdot \Delta A + c_L \cdot \Delta B,$$

where c_K denotes the capital cost component and c_L represents the labor cost component. So, the firm seeks to maximize

$$(8.3) \quad V = c_K \cdot \Delta A(m,n) + c_L \cdot \Delta B(m,n) - m \cdot P^m - n \cdot P^n,$$

where P^m and P^n are the prices that are associated with a unit of research of line m and n respectively. Here it suffices to note that a change in the relative factor prices (reflected in a change in the ratio of c_K and c_L) induces, *ceteris paribus*, an invention bias (reflected in the ratio of m and n).² The result is a labor-saving or capital saving innovation.

8.4 SCHUMPETERIAN EQUILIBRIUM MODELS OF R&D DECISION MAKING

Schumpeterian equilibrium models seek to describe the cost-benefit analysis that underlies the R&D decision making of firms in particular environments. The models in the Schumpeterian equilibrium tradition can be distinguished according to the assumptions employed. Crucial are the assumptions on expectations, rivalry and payoff structure (that is, the rules of the game). Dasgupta (1986) points out that an essential distinction is the one between continuous and discontinuous games.³ This distinction is particularly important in the context of the barrier market theory, since this approach is clearly of the discontinuous type (Parts IV and V). A continuous game implies that

the pay-off functions of firms are all continuous. Thus the model does not distinguish winners and losers sharply: if a firm alters his action slightly, his net reward changes slightly. Put another way, the reward schedules faced by the competitors are not of the form of a *tournament*, where performance is rewarded on the basis of its *rank* within the set of all realised performances (Dasgupta 1986, p. 535; Dasgupta's italics).

Kamien and Schwartz (1982) offer an excellent survey of *continuous* games. They argue that models are characterized by either decision-theoretic or game-theoretic features.

¹ An illustrative case is the one with $\alpha^m > 0 > \beta^m$ and $\beta^n > 0 > \alpha^n$. That is, "the m process reduces capital requirement but increases labor requirement, whereas the n process reduces labor but increases capital requirements" (Binswanger 1978, p. 135).

² Binswanger elaborates on the model, for example, by introducing the impact of budget constraints and market structure.

³ Beath *et al.* (1989) and Reinganum (1989) offer surveys of the literature on continuous and discontinuous (or, alternatively, nontournament and tournament) innovation games that has developed since 1980.

It is typical of a *decision-theoretic* setting that

certain elements of the firm's environment are taken as fixed or exogenous, including the rewards for innovation and imitation, the development cost function, and the intensity of rivalry. Of particular importance in the decision-theoretic framework is the assumption that the firm believes its choice of the level of R&D spending does not influence the R&D spending levels of its rivals (Kamien and Schwartz 1982, pp. 107-108).¹

So, a decision-theoretic model implies that (i) Cournot conjectures are employed and (ii) the intensity of rivalry (and so market structure) is given. Following Scherer (1967) and Barzel (1968), Kamien and Schwartz have published a series of articles in this tradition (1972, 1974a, 1974b, 1976, 1978, 1980 and 1982).

Kamien and Schwartz' models focus on the time-cost tradeoff that dominates much R&D. That is, they determine how rivalry affects the optimal development period of a given innovation (Waterson 1984, p. 158). This is done by analyzing the decision model of a representative firm with the help of dynamic programming. The basic argument runs as follows. A firm is assumed to maximize the expected present value of future profit flows, taking into account the intensity of rivalry. So, a firm seeks to balance

the relative importance of the opportunity for new extraordinary profits, the 'carrot', versus the defense of existing profits, the 'stick', as motivators of inventive activity. The decision to accelerate or decelerate development upon rival precedence rests importantly on which of these motives is dominant (Kamien and Schwartz 1982, p. 110).

In order to illustrate the mode of modeling in the Schumpeterian equilibrium approach, the essentials of the decision(-theoretic) model of a firm are described. Let r 's and p 's denote receipts on the noninnovated and innovated good respectively. The given values of the r 's and p 's reflect the assumption of the intensity of rivalry. In the decision period $t = 0$ the profit flow is r_0 . If the market grows with rate g and an innovation is introduced at date $t = u$, then the firm has captured a gross profit flow (ignoring discounting)

$$(8.4) \quad \pi_{0-u} = \int_0^u e^{gt} \cdot r_0 dt$$

at u . If the firm is the first to introduce the innovation, whereas rivals imitate at time v , then the firm receives

$$(8.5) \quad \pi_{u-v} = \int_u^v e^{gt} \cdot p_0 dt,$$

over the interval of time in which it is a monopolist, where p_0 is the monopolist's gross profit at u . Assume

$$(8.6) \quad \pi_{v-w} = \int_v^w e^{gt} \cdot p_1 dt$$

to be the value of the innovator's gross profit flow after imitation, where of course $p_1 <$

¹ Note that decision theory in the literature on innovation is different from the discrete variant of the maximizing postulate in Part II.

p_0 and w denotes the date when the innovation ceases to be useful.¹ If a rival innovates at u before the firm under consideration does, then the gross profit from the noninnovated good until the imitation date v is

$$(8.7) \quad \pi_{u-v} = \int_u^v e^{gt} \cdot r_1 dt,$$

where of course $r_1 < r_0$. After imitation the firm captures a gross profit flow of the value

$$(8.8) \quad \pi_{v-w} = \int_v^w e^{gt} \cdot p_2 dt.$$

Let $F(u)$ be the firm's assessment at date t of the probability that a rival will have introduced the innovation by time $t = u$. Hence, $G(u) = 1 - F(u)$ indicates the expected probability that no rival has brought out the innovation at u . The derivative of $F(u)$, $F'(u)$, represents the probability density of being the first but an imitator appearing at u , and of being a follower if a rival preceded at u . Discounting future profits at rate i , the firm's expected flow of earnings (W) if it introduces the innovation at date u is

$$(8.9) \quad W(u) = \int_0^u G(u) \cdot e^{kt} \cdot r_0 dt + \int_u^v G(u) \cdot e^{kt} \cdot p_0 dt + \int_v^w F'(u) \cdot e^{kt} \cdot p_1 dt + \int_u^v F(u) \cdot e^{kt} \cdot r_1 dt + \int_v^w F'(u) \cdot e^{kt} \cdot p_2 dt,$$

where $k = -(i - g)$.

The R&D cost is assumed to decrease at an increasing rate as the completion time u is extended, which implies a time-cost tradeoff. If R&D are undertaken under contract², a firm chooses an introduction date u which maximizes the net profits

$$(8.10) \quad V(u) = W(u) - C(u),$$

where $C(u)$ is the R&D cost function with $\partial C / \partial Ku < 0$ and $\partial^2 C / \partial u^2 < 0$. Hence,

[t]he marginal condition for choice of introduction time involves balancing the marginal gain from a moment's delay against the marginal cost from doing so (Kamien and Schwartz 1982, p. 117).

Kamien and Schwartz (1982) elaborate on the basic structure of the decision-theoretic model by introducing, for example, noncontractual R&D and technical uncertainty. The main result of the comparative statics as regards to the intensity of rivalry is that intermediate competition generates the shortest development period.³

¹ After $t = w$ the innovated product is superceded by a subsequent innovation, the firm exits or the market ceases to operate.

² This implies that the R&D program cannot be modified after a rival's introduction of the innovation before date u .

³ Waterson (1984) points out that Kamien and Schwartz "find that the shortest development period may either occur when there is no rivalry at all or (particularly when the project is an important one) when there is an intermediate intensity of rivalry. The intuitive explanation they give for this is as follows: If, from a situation free of rivalry, the threat of entry increases, there is a probability of later profits being pre-empted. This will encourage the firm to go for a shorter development time. However as rivalry increases still further, pre-emption becomes very likely, the prospect of winning the race recedes, and the cost of shortening development time rises ever faster, so that the firm may well come to the point of deciding to cut

The contribution of the *game-theoretic* approach is that it makes the intensity of rivalry an *endogenous* variable. That is, the number of (innovating) firms in equilibrium depends on the characteristics of the model. Game-theoretic models describe a decision procedure of firms that closely resembles the one above. A firm faces a tradeoff between time compression and R&D cost. In particular, the uncertainties surrounding the tradeoff determine the equilibrium number of innovating firms. Important models in this tradition are Loury (1979), Lee and Wilde (1980), Dasgupta and Stiglitz (1980a and 1980b) and Reinganum (1979 and 1981).¹

Dasgupta and Stiglitz (1980b) note that three elements of a model are particularly important in determining the structure of the equilibrium:

- (1) The *actions* that are available to the firms. For example, Dasgupta and Stiglitz (1980a and 1980b) develop only one strategic variable, the once and for all choice of the size of the R&D budget, whereas Reinganum (1981) introduces dynamic strategies, the choice of an entire path of R&D expenditures through time.
- (2) The *timing* of the actions and the *information* available to firms about the actions taken by rivals.
- (3) The *beliefs* about the actions of rivals. Here Cournot or Stackelberg expectations are popular. Cournot conjectures imply the belief that rivals will not react, whereas Stackelberg expectations predict an active response by competitors.²

It appears that in the determination of the equilibrium number of (innovating) firms the uncertainty in outcomes of R&D plays a critical role (for example, Dasgupta and Stiglitz 1980b, p. 26 and Reinganum 1981, p. 22). This uncertainty is introduced by assuming that the *probability* of reaching a given innovation by any date can be increased (generally under decreasing returns) by devoting more resources to R&D. Kamien and Schwartz (1982) note an important result of a selection of game-theoretic models that the

essential linkage is through technical uncertainty in the innovation process, which generates market uncertainty, [since] without technical uncertainty, the game-theoretic formulation suggests either no innovation or at most one innovator. This somewhat surprising result was independently recognized by Dasgupta and Stiglitz (1980b) and by Reinganum (1979) (Kamien and Schwartz 1982, p. 176).³

Discontinuous games depart from the continuous models by assuming a framework in which the winner takes all (Dasgupta 1986). A firm's payoff function is discontinuous. That is, the winner (the firm which innovates first) applies the innovation, whereas losers (firms which innovate later) remain empty handed. The natural setting that reflects this assumption is characterized by patent races in combination with Bertrand competition.

its losses and go for a longer development period" (Waterson 1984, p. 158).

¹ A further feature of these models is that market performance is compared with social optimal outcomes. This topic is treated in Section 8.6.

² For example, Nakao (1989) presents a model of a cost-reducing oligopoly with conjectural variations.

³ However, the "presence of technical uncertainty alone ... is not enough to assure that a game-theoretic formulation can explain several firms simultaneously seeking to innovate. As Dasgupta and Stiglitz have observed, it is also necessary that the development activities of each of the firms are uncorrelated. That is, there are several paths to successful development of the innovation. If this is not true, then each firm, by spending enough, can virtually assure itself a tie for being first and we are back to the situation in the absence of technical uncertainty" (Kamien and Schwartz 1982, p. 177).

The 1980s show a growing literature on patent races (Gilbert and Newberry 1982, Fudenberg *et al.* 1983, Reinganum 1983 and 1989, Fudenberg and Tirole 1985, Harris and Vickers 1985, Gilbert 1986, Sah and Stiglitz 1987 and Connor 1988). The discontinuous games sometimes do and sometimes do not take into account strategic interactions (Harris and Vickers 1985, p. 195 and Beath *et al.* 1989, p. 171).¹

The essential feature is that the

analysis is based on a model in which the gains to firms are determined in a market characterized by Bertrand competition. This assumption gives our model a "winner-takes-all" feature similar to that in the patent race literature. There is a private return to innovation only when a single firm innovates; when two (or more) firms innovate, Bertrand competition means that they earn no profits (Sah and Stiglitz 1987, p. 98).

The race to be the first is generally associated with preemptive innovation, which can provide an explanation for either the persistence of monopoly or profit equalization and/or dissipation. Take, for example, the following two arguments. First, the persistence of monopoly can arise because

[o]nce a player is 'far ahead enough' in the race, it ceases to be a real contest. The other player gives up and saves his efforts in the knowledge that any effort that he could afford to make would immediately be outdone by its rival. Being far ahead in this sense does not depend literally upon respective distances from the finishing line; rather, it also has to do with how much the competitors value the price (Harris and Vickers 1985, p. 206).²

The key point is that a monopolist has a greater incentive to win the next patent, because it has more to lose (for example, Dasgupta 1986, p. 536; for an opposite view see Reinganum 1983). Second, strong competition (with in the limit profit dissipation) is, however, facilitated if an entrant is able to leapfrog over the incumbent.³ Fudenberg *et al.* (1983) present two examples:

The first is a model in which the R&D process must pass through a number of discrete steps, so that even though one firm may be favored in the current stage a follower may still jump ahead. The second reason for leapfrogging is that each firm may be unable to monitor the current R&D efforts of its rivals, observing their progress only with a lag. The information lag allows a follower to catch up before the leader can respond. ... the possibility of leapfrogging can give rise to vigorous competition (Fudenberg *et al.* 1983, pp. 4-5; Fudenberg *et al.*'s italics).

¹ One may question the game-theoretic nature of discontinuous games without strategic interactions.

² A monopolist may even deter entry by adopting sleeping patents. As Gilbert (1986) puts it: "Notice at once that the patent may be for a technology which is *strictly inferior* to the one which the monopolist holds. The pre-emption argument goes through, but the monopolist will not actually use the new patent. His incentive for winning it is simply to prevent his rival from using it. In this case, the monopolist will take pre-emptive action and hold a sleeping patent" (Gilbert 1986, p. 537; Gilbert's italics).

³ Broadly speaking, leapfrogging models can be distinguished from catchup frameworks. The former describe the case "where a follower is allowed at any given time to compete directly with the leader for the new best-practice technology" (Beath *et al.* 1989, p. 179). The latter assume that "at any given time a follower is not able to compete for the same position as the leader. They must first move, in a step-by-step fashion, to the leader's position" (Beath *et al.* 1989, p. 179). Beath *et al.* (1989) review and compare both types of models.

8.5 SCHUMPETERIAN DISEQUILIBRIUM MODELS OF EVOLUTIONARY SELECTION

The Schumpeterian equilibrium approach to innovation can be criticized for describing essentially

a one-shot game [that] fail[s] to situate it in a long-run evolutionary process of industrial development (Iwai 1984a, p. 160).¹

In a series of publications (1974, 1975, 1976, 1977a, 1977b, 1978, 1982, 1984 and 1986) Nelson and Winter present a well thought-out response to this critique.

In the first appearance of the evolutionary framework in the literature, Nelson and Winter (1974) clearly express the basics of their alternative approach.² Starting from a scathing (but unjustified; Chapter 4) critique of the neoclassical theory, they develop a Schumpeterian evolutionary theory by replacing both cornerstones, maximizing decision rules and equilibrium, by satisficing decision routines and disequilibrium selection respectively.³ The first key feature means that firms follow routinized behavior as long as the aspiration levels of the goal variables are satisfied. Nelson and Winter (1974) note that the

first major commitment of the evolutionary theory is to a "behavioural" approach to individual firms. ... While neoclassical theory would attempt to deduce these decision rules from maximisation on the part of the firm, the behavioural theory simply takes them as given and observable (Nelson and Winter 1974, p. 891).

If a firm fails to reach an aspiration level, a search for a better action alternative is started. The second key element of an evolutionary model implies that the inevitable uncertainty and heterogeneity of firms induces disequilibrium⁴ and external selection.⁵

¹ Here it is illuminating to note Kamien and Schwartz' (1982) lament that "[i]deally, we would like to construct a model of the interaction between the process of technical advance and market structure that incorporates all these features and others as well. Firms should be able to revise their development plans in response to the successes or failures of rivals. Rewards might accrue to imitation as well as to innovation. There should be possibilities for strategic behavior both with respect to identifiable existing rivals and unidentifiable potential entrants. Firms should be allowed to learn from their past innovation experience. Exogenous shocks representing technological breakthroughs complementary to the particular market under study might also be incorporated. *The main question to be addressed in this framework is how market structure evolves through time*" (Kamien and Schwartz 1982, p. 219; italics added).

² The immediate reason for Nelson and Winter's (1974) introduction of the evolutionary theory is to close the analytical gap between the macroeconomic neoclassical growth theory (with its *ad hoc* explanation of growth on the basis of technical change) and the microeconomic theory of technical advance (with its neglect of macroeconomic implications). In particular, Nelson and Winter link their simulation model to Solow's (1957) data so as to offer an explanation of economic growth on the basis of a microeconomic foundation.

³ Nelson and Winter's critique (as, in particular, put into words in 1975, 1980 and 1982) is largely wide of the mark. The (potential) similarities between maximizing and satisficing decision procedures is discussed in Chapter 4. The compatibility of neoclassical theory and disequilibrium is touched upon in Chapter 23. In effect, Nelson and Winter model temporary equilibrium settings. Here these discussions are ignored.

⁴ Nelson and Winter's disequilibrium argument bears a family resemblance to the neo-Austrian and classical theory of the competitive process (see Chapter 23).

That is,

[e]volutionary theory involves, finally, explicit analysis of the economic selection mechanism - the change in the weighing of different decision rules that comes about through the expansion of firms using profitable rules and the contraction of firms using unprofitable ones (Nelson and Winter 1974, p. 893).

Since the two key characteristics of an evolutionary model introduce a scale of complexity that is difficult to handle analytically, a major part of Nelson and Winter's analysis is based upon simulation techniques (Nelson and Winter 1977).

A typical evolutionary model is presented in Winter (1984). Winter nicely introduces the essentials of the framework by arguing that the

model employed is much the same as in earlier work. It is a Markov model of a single industry in which firms produce a homogeneous product and in which cost reduction through productivity improvement is the major competitive weapon. ... Exogenously changing technological opportunities provide the setting for the struggle to increase productivity, but taking advantage of those opportunities requires costly and uncertain innovative efforts. Firms may choose to try to imitate the successful methods of other firms instead of trying to innovate themselves, but this strategy too is costly and uncertain. Firm growth is linked to profitability, but responds negatively to market share for firms that are large relative to the market. The fact that growth is linked to profitability brings economic "natural selection" into play as an influence on industry average productivity and on industry average policies toward innovation and imitation. ... the emphasis on the role of entrants demands that the basic model be augmented by a model of entry (Winter 1984, p. 288).

Assume that single product firms face downward sloping demand curves. Techniques vary with respect to capital productivity (a , which is the number of products produced by a unit of capital). Firm i 's profit per unit of capital (π) in a period t follows from

$$(8.11) \quad i\pi = (p_t - ac) \cdot i a_t - i m r_t - i n r_t,$$

where ac denotes the average unit cost of production per unit of capital and mr and nr are the routine expenditures per unit of capital that are spent on imitation respectively innovation. Productivity is determined by a two-stage random procedure. First, a yes or no draw indicates whether a firm carries out imitative and/or innovative activities in period t . The second stage determines the outcome of the imitative and/or innovative efforts. The probability that a firm is successful is proportional to its spending on imitation and innovation activities. The result of imitation efforts (ma_t) follows from a draw from the population of (capital-weighted) technologies employed by rivals, whereas the outcome of innovative activity (na_t) is determined by a draw from the population of

⁵ A similar type of evolutionary argument is employed by Witt (1985 and 1986) and Hannan and Freeman (1977 and 1989). Witt describes the natural selection of preferences (in particular, in connection with the debate about altruism and self-interest; Chapter 3) and decision rules (Chapter 4) respectively, whereas Hannan and Freeman focus on the ecology of organizational populations. Here a resemblance to the literature on the interface between biology and economics (for example, Becker 1976 and Hirschleifer 1985) is touched upon. Nelson and Winter recognize the parallel between their framework and biological theories of selection (for example, 1974, note 1, p. 891).

technological opportunities.¹ The productivity in period $t+1$ is given by

$$(8.12) \quad a_{t+1} = \max [a_t, ma_t, na_t],$$

which implies a one-period investment lag.

A firm adjusts its routine R&D policy (mr_t, nr_t) if a profit aspiration level is not reached. This means that internal selection occurs. In effect,

a firm that has an unsatisfactory performance makes incremental adjustments of its policy in the direction of industry average policy, modified by random disturbances (Winter 1984, p. 301).

External selection follows from the fact that a firm's (negative or positive) growth is linked to its profitability. If a firm's capital stock shrinks below a given minimum or if its rate of return is persistently below a given critical level, (permanent) exit is inevitable. Potential entrants, which are imitating and innovating behind the scene, enter the market to seize perceived profit opportunities.² Simulation experiments are employed to analyze the implications of the model under different conditions (for example, the nature of technological opportunities and the relative cost of imitation and innovation). The result can be that either imitation or innovation is the most profitable strategy (Winter 1984, p. 469). Moreover, market structure is determined endogenously.

Nelson and Winter's evolutionary approach to technical change has appeared to be an example that is worth following. For instance, Futia (1980), Iwai (1984a and 1984b) and Grabowski and Vernon (1987) pursue an argument along evolutionary lines³, since the aim of these models is to analyze a process of creative destruction by describing the interplay between imitation, innovation and growth: (i) Futia's (1980) setting departs from Nelson and Winter's framework by modeling maximizing rather than satisficing decision making; (ii) Iwai (1984a and 1984b) bypasses the modeling of firms' decision rules by introducing simple, *ad hoc* efficiency principles; and (iii) Grabowski and Vernon (1987) focus on product rather than process imitation and innovation.

8.6 INNOVATION AND MARKET PERFORMANCE

As in the entry barrier literature, particularly in the Schumpeterian equilibrium tradition the scale and speed of innovation are evaluated with regard to market performance by studying the impact of particular conditions on producers' and consumers' surplus. The socially optimal scale or speed of innovation is calculated by maximizing the total surplus. The first attempt to carry out a welfare analysis of R&D and market structure is Arrow (1962). Subsequent studies have elaborated on Arrow's analysis by introducing additional arguments (for example, the role of uncertainty). Some recent examples worth mentioning are Dasgupta and Stiglitz (1980a and 1980b), Von Weizsäcker (1980a and

¹ Technological opportunities depend on the firm's current productivity and the results of (basic) research carried out outside the firm. Changes in the latter are determined exogenously.

² The potential entrants' "background" R&D activity is funded outside the industry under consideration. The size of an entrant's initial capital stock follows from a draw. The scale of entry is supposed to be relatively small. Potential entrants face entry barriers. They decide to enter if $p_t \cdot a_t - c > r + \mu_t$, where r is the rate of entry barriers and μ a random error.

³ Other examples are reviewed and presented in Dosi (1988) and Dosi *et al.* (1988).

1980b), Reinganum (1981), Spence (1984 and 1986), Gallini and Kotowitz (1985), Stiglitz (1986 and 1987) and Bond and Samuelson (1987). The complexity of the welfare evaluation of innovation and market performance is immense. Below only a few remarks are made to illustrate this point.

Three problems are associated with R&D and market performance (Spence 1986, pp. 475-476): (i) if R&D is associated with sunk cost, the market structure may show imperfect price competition (entry barriers); (ii) R&D outlays generally introduce fixed and sunk costs, which implies that the private revenues of innovation may understate or overstate the social benefit; and (iii) externalities (*i.e.*, imitation) may bias the private benefit of innovation.

The first problem is well-known from the entry barrier literature. Therefore, there is no need for a further exploration here (Chapter 7). The second problem may be associated with a suboptimal speed of innovation. For example, duplication of R&D (that is, many firms carrying out equivalent innovation projects) may result in excessive expenditures from a social point of view.¹ Only the third problem is typical for R&D, since the first two follow directly from the sunk nature of much innovations. The third problem (externalities or appropriability problems) introduces

really two versions of the same problem. If the R&D for the single firm is not appropriable, *i.e.*, the benefits flow without payment to other firms, the initial incentives to do R&D are reduced. On the other hand, ... restoring appropriability is sometimes regarded as a second-best solution to the problem of providing incentives, because it creates monopoly or monopoly power. It may do that. But it is important to note that it also prices incorrectly the good that the R&D has created. And that by itself has its social costs (Spence 1986, p. 476).²

It is clear that the three problems indicated above do not allow unequivocal answers to the question of innovation and market performance. So it is not surprising to note that

[i]n spite of the long-standing concern about the relation between competition and innovation, there is no consensus. One view holds that monopolies have too little incentive to engage in R&D; without the spur of competition, monopolies will simply enjoy their monopoly rents. The other view holds that, without some degree of monopoly power, firms will be unable either to support R&D programmes or to appropriate the returns from research. As is so often the case, there is undoubtedly a grain of truth in both views: the question is; under what conditions, under what circumstances, is each view appropriate? (Stiglitz 1986, p. 401).

To make things worse, Stiglitz (1986, pp. 439-440) points to three further complications that are ignored in the literature. For example, current models reflect a

partial-equilibrium rather than general-equilibrium analysis. The analysis might be correct for

¹ Dasgupta (1986) moves one step further by arguing that a firm's internal incentive structure needs careful consideration. This introduces a quest for optimal incentive structures that induce managers to make correct decisions (Sah and Stiglitz 1985, 1986 and 1988).

² Several ways out have been suggested. To mention two examples: (i) Spence (1986) argues in favor of corrective subsidies; and (ii) Gallini (1984), Salant (1984), Gallini and Winter (1985) and Katz and Shapiro (1985 and 1987) point to the instrument of licensing. Moreover, Cohen and Levinthal (1989) point out that "contrary to the traditional result, intra-industry spillovers may encourage equilibrium industry R&D investment" (p. 569) as "firms invest in R&D not only to pursue directly new process and product innovation, but also to develop and maintain their broader capabilities to assimilate and exploit externally available information" (p. 593).

an isolated sector; that is, it might provide the correct answer to this question. If all sectors of the economy but one were competitive, and that one was the only one in which R&D could occur, would there be too much or too little research? (Stiglitz 1986, pp. 439-440).¹

However, although the question of innovation and market performance is complicated, the theoretical literature suggests that *intermediate* competition provides the strongest incentives to innovate (Kamien and Schwartz 1982). Part IV goes on to examine conditions of (potential) competition that facilitate both average total cost pricing and efficient innovation by mixing the insights of contestability and entry barrier theory.

8.7 CONCLUSION: Complexity and Creativity

The literature on innovation is immense. It leaves the reader with the impression that the complexity of the R&D process itself is reflected in the economics of innovation. Here it is telling that the above survey ignores (or mentions only in passing) many of the topics that are related to the economics of innovation. To name four examples: (i) diffusion trajectories (Salter 1960, Mansfield 1968 and Davies 1979); (ii) spillovers of multiproject R&D (Van Cayseele 1986 and 1987 and Dasgupta and Maskin 1987); (iii) product innovation and differentiation (Gabsewicz and Thisse 1980 and 1986, Shaked and Sutton 1982 and 1983 and Sutton 1986); and standardization of technologies (Farrell and Saloner 1985 and 1986, Katz and Shapiro 1985 and Besen and Johnson 1986).

A last remark is in order as regards the extent to which the economics of innovation has succeeded in endogenizing technical change. The boundary of the theory of endogenous innovations is located at the point where *creativity* comes in.² It is only the (scale and speed of the) *use* of *existing* technological opportunities that can be endogenized. In all models there is some assumption of an exogenously determined (and probably changing) fund of general technological possibilities. As Binswanger (1978) puts it:

Thus we arrive at a pessimistic conclusion. It is impossible to build a very long-term theory of economic growth that treats technical change in a truly endogenous way. This means that it is impossible to have a long-term theory of growth at all, because technological change is an essential part of growth. The fundamental problem is to know *ex ante* how basic and supporting sciences will develop in the future. Growth models cannot bridge that gap in our knowledge; they can only trace the consequences of various assumptions about the long-term development of basic knowledge (Binswanger 1978, p. 159).

Hence, the

usefulness of induced innovation models, therefore, should not be sought in their contribution to formal growth theory, but rather in what they imply for research policy (Binswanger 1978, p. 159).

For example, theory can be useful in making explicit the market conditions that facilitate innovation or the exploitation of existing technological opportunities.

¹ The two other point mentioned by Stiglitz are the distortive effects of corrective measures and the effect of competition on the risk of undertaking R&D. At this point, these and other complications (for example, the existence of winners and losers) are ignored. Some further remarks are offered in Chapter 24.

² For an underlying argument, see Chapter 3.

PART

IV

WORKABLE COMPETITION AND THE BARRIER MARKET

A Benchmark Case of Intermediate Competition

KEY WORDS. Part IV serves a twofold purpose. First, the theory of workable competition and related literature is discussed (Chapter 9). Workable competition describes intermediate competition which gives favorable market performance in terms of static as well as dynamic efficiency. Recent related literature presents formal models of firms' behavior which resembles workably competitive outcomes. Second, the barrier market theory is introduced (Chapter 10). The barrier market theory offers a microfoundation of a workability concept by introducing sunk investment and exit cost in a free entry framework. This means that the barrier market theory generalizes the recent models of workably competitive scenarios. As workable competition, the barrier market is defined with regard to its implications for market performance.

1944

1945

1946

1947

1948

1949

1950

1951

1952

1953

1954

1955

1956

1957

1958

1959

1960

1961

1962

1963

1964

1965

CHAPTER

9

STATIC AND DYNAMIC EFFICIENCY: Workable Competition and the Threat of Entry

9.1 INTRODUCTION: A Way Out of the Efficiency Dilemma

On the one hand, Baumol *et al.*'s (1982a) contestable market theory offers a *static* benchmark case of competition. Its advantage over perfect competition is that static efficient outcomes are guaranteed, irrespective of the prevailing market configuration (that is, number and size distribution of firms), provided that the configuration is sustainable. Bertrand competition in a free entry single product framework implies that incumbent firms are forced to adopt *average cost pricing* in order to avoid exit and deter entry.¹ However, the contestable market theory is still a *static* equilibrium notion (Baumol *et al.* 1983). Hence, it takes insufficient notice of the tradeoff between static and dynamic efficiency of market behavior (Chapter 6). On the other hand, entry barriers introduce dynamic aspects of competition to the detriment of static efficiency, since barriers to entry permit incumbents to exploit a cost advantage, which implies that prices are (persistently) set above the average cost level (Chapter 7). This efficiency dilemma induced an impressive literature on the optimal tradeoff between statically efficient prices and dynamically efficient innovations (Chapter 8).

In this chapter it is argued that there are market conditions which induce incentives to adopt careful price setting and to introduce innovation at one and the same time.² This result is reached by introducing sunk investment and exit cost in a multimarket free entry setting. In effect, the introduction of firms' *endogenous* (sunk) investment strategies constitutes a way to introduce *dynamics* into the contestability framework. This result resembles the intentions of the concept of workable competition. Moreover, in order to endogenize the credibility of the entry threat, competition is viewed from a *multimarket* perspective. This permits the identification of potential entrants. More is said about this in Chapters 10, 11 and 21.

It should be emphasized that the following discussion is restricted to the single product case. Investment behavior is associated with sunk costs so that incumbent firms can face exit barriers. However, a major issue to address is whether an assumption of free entry is compatible with the presence of sunk costs. So far, the literature suggests that sunk costs

¹ Parts IV to VII of the thesis are phrased in terms of the commodity market. Of course, the argument can be extended to the labor market (Part VIII).

² Much of the argument is focused on innovation as a specific manifestation of a sunk cost investment.

introduce entry barriers.¹ To elucidate the place of the argument presented below in the literature on industrial organization, reference must also be made to recent articles by Stiglitz (1987), Dasgupta and Stiglitz (1988a) and Cairns and Mahabir (1988).

As Dasgupta and Stiglitz (1987) argue,

if those arguments [underlying the contestable market theory] could be extended to the particular kinds of non-convexities associated with technical change, then we would at last have an intellectual foundation for the belief in the efficiency of capitalist industrial economies (Dasgupta and Stiglitz 1987, p. 570).

Dasgupta and Stiglitz are, however, skeptical about the possibility of providing such an extension. The reason for their skepticism is straightforward:

Baumol, Panzar and Willig recognized that for potential competition to ensure zero profits sunk costs could not be too large. It turns out, however, that *if there are even small sunk costs, potential competition may not be effective in ensuring either that profits go to zero or that efficient outcomes obtain* (Dasgupta and Stiglitz 1987, p. 571; Dasgupta and Stiglitz' italics).

Stiglitz (1987) and Dasgupta and Stiglitz (1988a) support this proposition by presenting some illustrative examples where even small sunk (exit) costs undermine the threat of entry. Moreover, Dasgupta and Stiglitz (1987) note that

[a] general principle emerges: *the more competitive ex post competition... the less effective ... the market discipline provided by potential competition* (Dasgupta and Stiglitz 1987, p. 572; Dasgupta and Stiglitz' italics).

This result implies an argument in favor of the proposition that there is a tradeoff between static and dynamic efficiency.

There is no doubt that Stiglitz' (1987) and Dasgupta and Stiglitz' (1988a) examples are valid. However, here it is argued that there are *counterexamples* which show that under particular circumstances the *opposite* results apply (Chapter 11). That is to say, it is pointed out that under particular conditions the first proposition of Dasgupta and Stiglitz underlined above is *reversed*: to paraphrase Dasgupta and Stiglitz, *even if there are sunk (exit) costs, potential competition may be effective in ensuring that profits go to zero and that (efficient) innovations obtain*. Cairns and Mahabir (1988) even argue that the validity of an assumption of contestability *increases* if sunk investment is permitted. Here they point to

an already existing firm in a market which is a prototype of the market to be contested (Cairns and Mahabir 1988, p. 270).

So, they

propose that the idea of potential entry as a disciplinary force requires latent excess capacity, latent fungibility, and the ability of firms to alter endogenously the industry product set through their own strategies. In such a situation the existence of sunk costs may actually

¹ For example, Baumol and Willig (1981), Grossman (1981), Baumol *et al.* (1982a), Farrell (1986), Stiglitz (1987) and Dasgupta and Stiglitz (1988a). See also Chapter 6.

improve the chances of a firm to contest a market (Cairns and Mahabir 1988, p. 272; italics added).¹

Particularly Part V argues that the key point is that, although Stiglitz' (1987) and Dasgupta and Stiglitz' (1988a) examples as well as the cases presented below

depend on exit and entry (sunk) costs [which] are *endogenous* chosen by the firms as part of their strategies (Dasgupta and Stiglitz 1988a, p. 573; Dasgupta and Stiglitz' italics),

the results depend critically upon the *opportunity cost* of alternative strategies (*i.e.*, investment, pricing, entry, exit and stay).

The chapter is organized as follows. Section 9.2 briefly reviews the literature on workable competition, while Section 9.3 describes recent (and rare) examples of models that support the results presented below (*i.e.*, the compatibility of sunk and exit cost and free entry). Section 9.4 summarizes the argument by indicating workably competitive scenarios. Section 9.5 marks the contours of the argument in the other three chapters of Parts IV and V.

9.2 WORKABLE COMPETITION

Clark (1940) introduced the concept of workable competition² in order to replace the unrealistic norm of perfect competition (Clark 1940, pp. 241-242) by a benchmark that constitutes second-best performance³:

One central point may be put abstractly. If there are, for example, five conditions, all of which are essential to perfect competition, and the first is lacking in a given case, then it no longer follows that we are necessarily better off for the presence of any one of the other four. In the absence of the first, it is *a priori* quite possible that the second and third may become positive detriments; and a workably satisfactory result may depend on achieving some degree of 'imperfection' in these other two factors. ... imperfect competition may be too strong as well as too weak; and ... workable competition needs to avoid both extremes (Clark 1940, pp. 242-243).

Clark presented a list of the ten most important criteria that facilitate workable competition (Clark 1940, pp. 243-244).⁴ Moreover, he pointed to

[m]odified, intermediate or hybrid competition. ... The most important cases involve formally

¹ In this respect, Cairns and Mahabir stress the role of *multiproduct* firms. Particularly Part V argues that their proposition may also hold for the single product case.

² The concept of workable competition is closely linked to the structure - conduct - performance tradition (Sosnick 1958, p. 380). Other notable analyses of workability are Bain (1950), Clark (1955 and 1961), Ferguson (1964), Liebhafsky (1971) and Reid (1987).

³ Reid (1987) takes the stance that "Clark's basic insight will be regarded as a precursor of the theory of the second best" (Reid 1987, p. 115). Reid offers a treatment of the workability concept in second-best theory.

⁴ These ten criteria specify conditions concerning (1) the standardized or unstandardized character of the product, (2) the number and size-distribution of producers, (3) the general method of price making, (4) the general method of selling, (5) the character and means of market information, (6) the geographical distribution of production and consumption, (7) the degree of current control of output, (8) variation of cost with varying size of plant or enterprise, (9) variation of cost with short-run fluctuations of output and (10) flexibility of productive capacity.

free entry, but no exit without loss (Clark 1940, p. 245; Clark's italics).

This intermediate competition is precisely what the barrier market theory seeks to describe¹!

In 1958 Sosnick² pointed out that eighteen authors have listed criteria of workability that can be grouped into norms for performance, conduct and structure.³ The selection of norms follows from an inquiry that

would be concerned in the case of performance norms to decide what state of the various performance dimensions would imply maximum service to buyers' desires, and in the case of structure and conduct norms, to decide what state of their dimensions would imply maximum effects on performance (Sosnick 1958, p. 395).⁴

The dynamics of competition are very important:

Indeed, market incentives are able to designate and evoke aspects of desirable behavior not merely when enterprises are constrained by conditions like substitutes' availability and opposites' knowledge, but when producers have freedom to innovate and latitude when successful (Sosnick 1958, p. 397).

To give an idea of the nature of the workability criteria listed in the literature, Reid's (1987, p. 125) list of Sosnick's norms is presented.

Structure norms

- (1) No dominance, and traders as large as economies of scale will permit
- (2) Quality differentials which are moderate and sensitive to prices
- (3) No impediments to mobility
- (4) Reasonable availability of market information
- (5) Some uncertainty about responses to price cutting
- (6) Freedom from legal restraint
- (7) Development of new markets and trade contracts

¹ It is interesting to note that Clark indicates some other conditions of competition that are crucial in the current debate. To name two examples: (i) "the more attention centers on the imperfections of active competition, the more important become the forces of potential competition" (Clark 1940, p. 246); and (ii) "there is typically no definite 'optimum size,' but rather a wide optimum range of size, within which most plants fall; and ... economies due to size are far from being such vitally important factor as is suggested by the type of theoretical cost curve now in general use" (Clark 1940, p. 249).

² Sosnick strengthened the condemnation of perfect competition as a norm by concluding "that the perfectly competitive structure and conduct are unattainable in any real market; that closer approximations to them may entail worse performance than more distant; and that the closest possible approximation would entail actual and even equilibrium performance of dubious desirability" (Sosnick 1958, p. 384).

³ Sosnick recognized the well-known fallacy of causality: "It is also important to notice that this separation into structure, conduct, and performance corresponds only imperfectly to the direction of causation" (Sosnick 1958, p. 387). Moreover, he pointed out that obstacles of deriving (performance) norms are distribution effects and the occurrence of winners and losers (Sosnick 1958, pp. 393-394) and that the criteria listed are intertwined and related (Sosnick 1958, p. 399).

⁴ The requirement that the service to buyers' desires is maximized resembles the consumers' surplus maximization criterion (subject to the condition that production is feasible) that is used in the modern literature (Chapter 8 and 24).

Conduct norms

- (1) Independent rivalry, in pursuit of profit
- (2) No shielding of inefficient rivals, suppliers, or customers
- (3) No unfair, exclusionary, predatory, or coercive tactics
- (4) No unreasonable discrimination
- (5) No misleading sales promotion
- (6) Rapid response by buyers to differentials in attributes of products

Performance norms

- (1) Efficient production and distribution
- (2) No excessive promotional expenses
- (3) Profits sufficient to reward investment, efficiency, and innovation
- (4) Output consistent with efficient resource allocation
- (5) Prices that do not intensify cyclical problems
- (6) Quality consistent with consumers' interest
- (7) Appropriate exploitation of improved products and techniques
- (8) Conservation requirements respected
- (9) Sellers responsive to buyers' needs
- (10) Entry as free as the industry sensibly permits
- (11) Regard for national security requirements
- (12) Avoidance of excessive political and economic power in few hands
- (13) Regard for employees' welfare

The fact of the matter is that the workability concept can be improved by providing a formal microfoundation to the informal intuition.¹ This is useful, since the

difficulty with this approach [workability] lies in its sheer eclecticism. The criteria developed seem frequently arbitrary and vague, and occasionally inconsistent (Reid 1987, p. 115).

The lack of analytical rigor goes hand in hand with unanswered questions: what is the relative importance of the criteria listed; in what ways may interdependencies between criteria interfere; what critical levels are associated with the criteria listed? Attempts to elaborate formally on the workable competition concept seem therefore to be justified. To focus attention, two essential criteria of workable competition, which capture many of Sosnick's (particularly performance) norms, have to be kept in mind. First,

[p]rofits should be at levels just sufficient to reward investment, efficiency and innovation (Scherer 1980, p. 42).

¹ Reid (1987) points out that the "final approach to workability ... is the 'contestability markets' analysis of Baumol *et al.* (1982). They recognize the problems that arise from vagueness in the criteria laid down for workability but applaud the attempt to work with a competitive norm other than that of perfect competition. What they attempt to produce is an analysis which is both precise and consistent, as well as being concerned with a feasible workably competitive norm. However, it is not free of weaknesses, the most significant of which is its dependence on an assumption of ultra-free entry. The search for rigor and policy relevance is however commendable, and indicative of an attitude which may lead to a new and potentially fruitful approach to the doctrine of workable competition" (Reid 1987, p. 116).

Second,

[o]pportunities for introducing technically superior new products and processes should be exploited (Scherer 1980, p. 42).

Section 9.3 evaluates recent literature which suggests scenarios that give both workably competitive outcomes.

9.3 RECENT RELATED LITERATURE

Maks (1986), Kim (1987), Cairns and Mahabir (1988), Calem (1988), Maskin and Tirole (1988) and Mills (1988) offer (rare and recent) examples of arguments that support the result that there are scenarios which constitute an environment in which careful pricing goes hand in hand with investment and which indicate that the presence of sunk and exit costs is compatible with a free entry condition. Below these six examples are briefly discussed.

(i) Maks (1986) introduces the first contours of the barrier market theory by distinguishing two cases. First,

[a] first simplified analysis considers an equi-barrier market in a multi-period setting. This concept admits sunk costs but at the same time assumes equal exit costs for both incumbents and potential entrants. This situation prevents the hit and run activities of the contestable market. However, if one assumes that the potential entrant perceives the observed prices as reliable indicators of the minimum average cost of the incumbents, it is very dangerous to set prices higher than these costs. The potential entrant has no reason to hesitate about entry if he thinks that the expelled competitor is unlikely to react because of his higher costs. So, again, the price setting behaviour may remain close to the minimum cost under the threat of potential entry (Maks 1986, p. 119).

Second, Maks (1986) points to asymmetric sunk exit cost in favor of incumbents by arguing that

[t]he more general case of the barrier market acknowledges the possibility of higher exit costs for potential entrants (Maks 1986, p. 119),

whereas the entry lag is indicated by pointing out that

higher prices may stimulate entry within a short number of periods (Maks 1986, p. 119),

where the incumbent firm can set new prices at the beginning of each new period. The result is that

price setting behaviour in the neighborhood of the minimal average cost may be generated if the incumbents are aware of alert potential entrants, say existing capable enterprises who are able to quickly convert a barrier market with above normal profit opportunities to their advantage. ... It may also strengthen the attempts to create new barriers, leading to cheaper and better products (Maks 1986, p. 119).

So, Maks argues that a credible entry threat can derive from alert (new or existing)

potential entrants which will prepare (future) entry if profitable entry opportunities are observed. This preparation can imply (in the general case) that potential entrants (i) have to sink a cost upon entry which incumbents can spare and (ii) anticipate a probability that they are able to outperform the incumbent firms' innovation after an entry lag which exceeds the incumbent firms' adjustment delay. New firms start from scratch, whereas existing firms come from a home market.

(ii) Kim (1987) introduces product heterogeneity in a natural monopoly model. He analyzes

a firm that can provide a continuum of different price and quality pairs for certain ranges of qualities. We assume that the cost of production is related to the range of qualities produced as well as the quality level of each output. We assume an increasing returns to scale technology by assuming that a firm must incur a fixed setup cost when it begins operation and that the fixed setup cost is an increasing function of the measure of quality sets produced by the firm. ... we assume that the firm knows the distribution of consumer tastes, but cannot distinguish individual consumers (Kim 1987, pp. 231-232).

The fixed setup cost incurred before operation begins resembles a sunk entry cost. Provided that there is an equally equipped (but unidentified) potential entrant around, Kim proves that the natural monopolist's entry-deterring decision complies with Ramsey optimality for a particular specification of the demand side (*i.e.*, utility function and distribution of consumers). Hence, the natural monopolist

maximizes consumer surplus subject to a zero profit constraint (Kim 1987, p. 236).

The conditions that potential entrants are facing remain in the dark. However, the argument that unsustainability arises if

a potential rival could profitably take away the consumers who are bunched together by offering a finite number of discrete qualities only for those consumers (Kim 1987, p. 238),

drives the model.

(iii) The key argument of Cairns and Mahabir's (1988) essay is that

contestability theory leaves unanswered questions about the identity of entrants and the source of their resources. ... Rather than as a hindrance to potential entry, sunk costs are viewed as being of key significance for the disciplining of oligopolists through potential entry (Cairns and Mahabir 1988, p. 269).

The source of the entry threat are multiproduct firms. In fact,

contestability is inseparable from the multi-product firm analysis presented by Baumol *et al.* (Cairns and Mahabir 1988, p. 269).

Latent fungibility and excess productive capacity provide the basis for expansion into other markets, because

then one may argue that the conundrum of potential entrants not receiving their opportunity

costs does not arise (Cairns and Mahabir 1988, p. 271).

This means that only *existing* firms can induce a credible entry threat:

Established firm entry is easier than new firm entry, an important factor being reserve or excess capacity in the short run which may be shifted to producing other types of product. ... Altering the product set requires latent fungibility as well as latent capacity. The latter may be provided by previously sunk expenditures for capital: the former requires discovering the latent fungibility, through R and D and other expenditures that are necessarily sunk, as well as the use of the goodwill of the firm (Cairns and Mahabir 1988, p. 273).

(iv) Calem (1988) offers a model of penetrable markets:

The potential entrant may be a firm producing an identical product but operating in a geographically distinct market, in which case ease of entry would derive from low transport costs; or it may be a firm producing a distinct but technologically related product, in which case ease of entry would derive from substitutability in production. A market faced with entry as such may be deemed a *penetrable market*. It is plausible that a firm in a penetrable market would be unable to adjust its total output as rapidly as a potential entrant could transfer some of its product into the market. ... the rivalry between a firm in a penetrable market and a firm threatening to enter that market may be reciprocal. As these firms operate in related markets, each may be a potential entrant into the other's market (Calem 1988, p. 171; Calem's italics).

In line with Shepherd (1984) and Green (1987), Calem (1988, pp. 172-173 and 180) argues that the penetrable market model applies particularly to (the threat of) foreign competition.

Calem's model describes Cournot rather than Bertrand competition: quantity is the strategic variable. It is essential that incumbent firms and potential entrants decide on output (productive capacity) *simultaneously*. That is,

[a] novel feature of our model is that the monopolist in a market, and the firm threatening entry into the market, both make strategic choices (choose their total outputs) during the pre-entry stage (Calem 1988, p. 172; Calem's italics).

Entry is anticipated by installing excess capacity. Further assumptions are that two monopolists in two distinct markets face a joint production technology, whereas the entry lag and transfer (entry) cost are positive. Moreover, a firm is committed to productive capacity for one period (with a pre-entry and post-entry stage).

The result of the model is that

[i]n a penetrable market, as in a contestable market, a threat of entry can limit the exercise of monopoly power (Calem 1988, p. 173).

This result depends critically upon the features of the setting involved. To be precise,

[a] firm facing a threat of entry will employ an entry-detering strategy in its home market only if its transfer costs are large enough to rule out its being a potential entrant into its rival's market. Moreover, in the asymmetric case, the firm facing a threat of entry into its market will expand its output to an entry-detering level only if its rival's transfer costs are in a medium range (not if those costs are small) (Calem 1988, p. 181).

This means that market performance resembles workable competition if the entry threat is one sided. The entrant's home market is safe against entry by expelled incumbents.

(v) Maskin and Tirole (1988) examine a model of a monopolist facing one potential entrant. They are concerned with the impact of short term commitments on profit levels. A commitment is defined as follows:

When we say that firm 1 is *committed* to a particular action in the short-run - whether a quantity or a price - we mean that it cannot change that action for a finite (although possibly brief) period, during which time other firms might act. By firm 2's *reaction* to 1 we mean the response it makes, possibly after some lag, to 1's chosen action. Short-run commitment ensures that, by the time firm 2 reacts, firm 1 will not already have changed its action (Maskin and Tirole 1988, p. 549; Maskin and Tirole's italics).

That is to say, a commitment resembles the adjustment lag a first mover faces after announcing a quantity or price. Hence, a commitment amounts to the absence of a fear for retaliation for a (short) period of time.

In effect, Maskin and Tirole adopt a *one-period adjustment lag*. They assume that the two firms move alternately. A defence of this assumption is provided by arguing that

[w]e have in mind primarily exogenous or technological reasons for commitment, e.g., installed capital that has little scrap value or lags in producing and disseminating price lists. Alternatively, short-term contracts might serve to bind the firm temporarily (Maskin and Tirole 1988, p. 550).

Maskin and Tirole restrict their analysis to a *monopoly* market in which *quantity (productive capacity)* is the strategic investment variable. The monopolist faces the entry threat by an identical nonproducing firm which can decide on entry in reaction to the incumbent's quantity setting. The entry lag is one period. The model is rendered concrete by arguing that firms

maximize the discounted sum of instantaneous profits, with discount factor δ . If q is chosen to be strictly positive, we shall assume that the firm incurs a fixed cost F . We shall suppose that this cost is incurred up-front. But, since the firm is committed to the capacity q for two periods, we can think of $f = F/(1 + \delta)$ as the per-period or flow equivalent of F . Viewed this way, the fixed cost can be thought of as a 'short-term' *sunk cost* (Maskin and Tirole 1988, p. 555; italics added).

The key point is that the commitment is associated with an investment of sunk cost up-front. Maskin and Tirole show that the equilibrium solution depends critically upon the value of δ . To quote their major result:

The deterrence level q monotonically increases in the discount factor δ (and decreases in the fixed cost f). When δ is comparatively high ..., q is above the monopoly quantity q^m ... That is, to drive out its rival or deter it from entering, a firm must operate above the monopoly level. If the firm actually uses all the capacity it has installed, it, therefore, charges less than the monopoly price. Given these restrictions, the firm will produce exactly q . This is a result reminiscent of the limit pricing literature ...: an incumbent firm sells at a price sufficiently low that the immediate short-run losses outweigh the longer run gains. ... notice that, as δ tends to 1, $\pi^1(q, 0)$ tends to zero. That is, instantaneous profit is driven down to the competitive level. Hence, our model confirms the heuristic stories of Grossman (1981) and Baumol, Panzar and

Willig (1982) if firms place sufficient weight on future profits (Maskin and Tirole 1988, pp. 561-562).

(vi) Mills (1988) models a framework with two firms which bears close resemblance to the patent race literature (Chapter 8). The essential feature of the model is that firms decide on the *timing* of a sunk investment. Mills describes two scenarios: one with costless threats and one with costly threats. The former illustrates workable competition, whereas the latter shows that sunk cost can erect entry barriers. The essential assumptions of the setting with costless threats are

that there are two firms equally capable of making the investment and capturing the payoff. Assume that they have full information about each other and behave in a noncooperative fashion. Assume that both firms would receive negative payoffs if both invest, because the second investment would be redundant. Unless both firms invest simultaneously, the prospect of certain loss is enough to prevent a second investment. Neither firm invests once it learns that its rival has invested. ... assume that firms must act sequentially (Mills 1988, p. 116).

Sequential investment implies a positive entry and adjustment lag, where entry always occurs before the incumbent is able to adjust. Moreover, it is assumed that, first, t_m is the unique optimal introduction date of the sunk investment. That is, if a firm invests a sunk cost at date t_m before a rival has done so, then the investment yields a maximum payoff. Second, $\pi_{t-1} < \pi_t$ for $t \leq t_m$. Third, a zero profit accrues to the firm which refrain from investment (or entry). The result is that

[t]he investment described above can be viewed as a natural monopoly where potential monopoly rents are dissipated by rivalrous behavior. While this outcome results from the potential for 'hit-and-run-entry' in contestable markets, in this model it results from preemptive timing. In a contestable market rent dissipation occurs because entry involves no unrecoverable costs. Here it occurs because entry can be timed strategically (Mills 1988, p. 117).

By way of conclusion it is argued that the above literature describes scenarios that fit into the workable competition tradition. Notwithstanding the fact that incumbent firms and potential entrants incur a sunk cost up-front (and so face a positive exit cost), the condition of free entry is sustained so that an entry threat disciplines the behavior of incumbent firms.

9.4 WORKABLY COMPETITIVE SCENARIOS

The arguments of Maks (1986), Kim (1987), Cairns and Mahabir (1988), Calem (1988), Maskin and Tirole (1988) and Mills (1988) indicate scenarios that can give workably competitive results. As in contestability theory, the common assumption is that incumbent firms and potential entrants have access to the same (production and investment) technologies. The scenarios raise, however, (at least) six questions which are associated with underlying assumptions and conditions.

- (1) *(A)symmetry of sunk (exit) cost*: Is the (potential) entrants' scale of sunk and exit cost higher than, equal to or lower than the incumbents'?
- (2) *Order of investment*: Do (potential) entrants invest after the incumbent firms do so or do they sink cost simultaneously?

- (3) *Entry and adjustment lag*: Does the entry lag exceed the incumbents' adjustment delay?
- (4) *Response lag*: Is the incumbents' response lag zero or positive?
- (5) *Opportunity cost of entry*: Is the cost or profit foregone with entry zero or positive?
- (6) *Credibility of the (source of the) entry threat*: What is the identity of (potential) entrants?

Assume that at the beginning of period $t = 0$ incumbent firms make a decision which can induce entry. Suppose that entry takes place at the beginning of period $t = E$, whereas the incumbent firms are able to respond at the beginning of period $t = R$. That is, $0-E$ denotes the entrants' entry lag and $E-R$ represents the incumbents' response lag after entry has occurred. Incumbent firms can adjust plans (*i.e.*, price, quantity and investment) after A periods. So, A indicates the incumbents' adjustment delay. That is, the incumbent firm can lower price (below the higher level that signalled profit opportunities to the potential entrants) before actual entry occurs if $E > A$, whereas the incumbent is unable to change behavior before entry takes place if $E < A$.

The literature's assumptions and conditions are summarized in Table 9.1. Table 9.1 suggests the assumptions and conditions that underlie workably competitive market performance. Chapter 10 and 11 examine whether the assumptions and conditions listed in Table 9.1 stand up against a critical evaluation.

9.5 ESSENTIALS, SCENARIOS AND AMORTIZATION

The barrier market framework offers a formal elaboration of a concept of workability. The argument in Parts IV and V seeks to show that most of Sosnick's workability norms are associated with the assumptions and implications of the barrier market theory.¹ Here it is particularly noteworthy to quote a conclusion of Sosnick that the

most useful criterion for the original delineation appears to be that of symmetrical high arc price-quantity cross-elasticities of demand.² The idea is perhaps best expressed by saying that the workability criterion should formulate standards that apply to the *related* activities of all those who supply or demand all outputs *almost* any two of which are so akin that a *moderate ceteris paribus reduction* from the *prevailing* price of *either* would *soon* transfer to it a *substantial* proportion of the quantity demanded of the other (Sosnick 1958, p. 401; Sosnick's italics).

The threat of entry by producers operating in different but related markets can be a cornerstone assumption of the barrier market theory. An implication of this crucial assumption is, for example, that the aspects of static and dynamic efficiency that are associated with the performance norms (1), (3), (4), (6), (7) and (9) follow from the entry-detering strategies of incumbent firms in a barrier market (Chapters 10 and 11).

¹ Some crucial norms are associated with perfect contestability. For example, structure norm (3) and conduct norm (6) point to assumptions of supply respectively demand contestability.

² "High cross-elasticities of supply are more conveniently acknowledged under the condition of entry" (Sosnick 1958, note 9, p. 401).

Table 9.1
Assumptions Underlying Workably Competitive Results

	Sunk and exit cost	Investment order	Entry-Adjustment lag	Response lag (R)	Opportunity cost	Credibility of entry threat
Maks 1986	Symmetric or asymmetric in favor of incumbent	Simultaneous or sequential	$E > A$	?	Zero ?	New or capable existing firms
Kim 1987	?	?	?	?	?	Unidentified firm
C&M 1988	Symmetric	Simultaneous	?	?	Zero	Multiproduct firms with excess capacity and latent fungibility
Calem 1988	Asymmetric in favor of potential entrant	Simultaneous	$E < A$	Positive	Zero	Firms with excess capacity from safe home market
M&T 1988	Symmetric	Sequential	$E < A$	Positive	Zero	Identical nonproducing firm
Mills 1988	Symmetric	Sequential	$E < A$	Positive	Zero	Costless commitment (threat) to timing

C&M = Cairns and Mahabir

M&T = Maskin and Tirole

A = adjustment lag

E = date of entry

R = date of response

? = not available

Chapter 10 introduces the essentials of the barrier market notion: the barrier market is defined, the underlying assumptions and conditions are outlined and the implications for market performance are indicated. Chapters 11 and 12 deal with the subtle interplay of assumptions and some of the associated difficulties. Chapter 11 critically evaluates the assumptions and conditions of the workably competitive scenarios as to their necessity and validity. Chapter 12 goes on to examine the role of a zero-profit condition in a free entry framework with long-run investment. Hopefully, the chapters of Parts IV and V communicate an intuition of the sensitivity of market outcomes for the precise composition of the set of assumptions and conditions that is used to frame the argument.

AMERICAN MEDICAL ASSOCIATION
PUBLISHED WEEKLY
CHICAGO, ILL., U.S.A.
Subscription price, \$5.00 per annum in advance.
Single copies, 15 cents.
Entered as Second-Class Matter, May 2, 1902.
Postage paid at Chicago, Ill., and at additional mailing offices.
Acceptance for mailing at special rate of postage provided for in
Section 1103, Act of October 3, 1917, authorized on July 1, 1962.
Copyright, 1962, by American Medical Association
All rights reserved. Reproduction by any means without permission is prohibited.

CONTENTS

Original Articles
The Role of the General Practitioner in the Management of the Patient with
Hypertension
The Role of the General Practitioner in the Management of the Patient with
Hypertension
The Role of the General Practitioner in the Management of the Patient with
Hypertension

Editorial
The Role of the General Practitioner in the Management of the Patient with
Hypertension

Book Reviews
The Role of the General Practitioner in the Management of the Patient with
Hypertension

Index
The Role of the General Practitioner in the Management of the Patient with
Hypertension

Advertisements
The Role of the General Practitioner in the Management of the Patient with
Hypertension

Corrections
The Role of the General Practitioner in the Management of the Patient with
Hypertension

Index
The Role of the General Practitioner in the Management of the Patient with
Hypertension

CHAPTER

10

THE ESSENTIALS: The Barrier Market as an Intermediate Case of Competition

10.1 DEFINITION

As workable competition, the barrier market is defined with regard to its performance. The barrier market offers a formal microfoundation to a workability concept by introducing sunk investment in a contestability framework. This means, first, that a case of intermediate competition is framed and, second, that contestable market theory is transformed into a dynamic benchmark.

DEFINITION 10.1 (barrier market). *A barrier market is a market in which incumbent firms follow a strategy of careful [(nearly) average total cost] pricing as well as investment in innovation.*

This means that a barrier market is associated with market conditions that induce incumbent firms' behavior which gives in to both static and dynamic efficiency considerations (Maks 1986 and Van Witteloostuijn and Maks 1987). The barrier market (just as workable competition) is an *umbrella* notion that covers all scenarios which generate the indicated performance.

Moreover, for the sake of the argument in Chapter 11, a perfect barrier market is defined as a market in which average total cost pricing is the incumbents' equilibrium strategy, whereas an imperfect barrier market is associated with incumbents' careful but above average total cost prices. That is, a perfect barrier market gives a zero profit, while in an imperfect barrier market the incumbents' equilibrium strategy indicates a price below Bain's (1956) limit price but above average total cost, so that a positive profit is captured (Van Witteloostuijn 1988b). This distinction is explained further in Section 11.5. The argument in this chapter focuses on the *perfect* barrier market. The key point is that in a barrier market the incumbent firms' performance is disciplined by a threat of entry, notwithstanding the sunk nature of (innovation) investment.

This short chapter introduces the contours of the barrier market theory so as to offer an intuition that can be used as a stepping stone for the arguments in the following chapters by focusing on the *perfect* barrier market; that is, the barrier market which gives average total cost pricing and so zero profit. Section 10.2 examines the assumptions and conditions of entry and exit. Section 10.3 describes a benchmark model of the sequence of events in a barrier market which is used in subsequent chapters. Section 10.4 briefly indicates the implications for market behavior, while Section 10.5 relates the barrier market to other theories of competition.

10.2 ASSUMPTIONS AND CONDITIONS OF ENTRY AND EXIT

Existing contestable market models deal with the case where sunk investment and exit cost cannot occur (Schwartz and Reynolds 1983 and Shepherd 1984). The *endogenization* of innovations (*i.e.*, R&D activity) requires the introduction of sunk costs, since R&D cost can generally not be (fully) recouped outside the market (Stiglitz 1987). Hence, in order to introduce firms' sunk investment in a free entry framework, Assumption 6.2 must be dropped. This is precisely what the barrier market theory does (compare with Cairns and Mahabir 1988).

ASSUMPTION 10.1 (*positive sunk cost*). *Firms are allowed to undertake investment that is associated with a positive sunk cost (Definition 6.6). The sunk cost can give a positive exit cost (Definition 6.7) if exit occurs before the sunk cost is fully amortized.*

The pervasive fear of entry is retained by modifying Assumption 6.1.

ASSUMPTION 10.2 (*entry threat*). *Incumbent firms anticipate full-scale (*i.e.*, replacing) entry through the response of (a) alert potential entrant(s) which will take advantage of any profit opportunity by entering the incumbents' market, notwithstanding the sunk nature of investment.*

A remark is in order as regards the term barrier market. Strictly speaking, this terminology can be misleading, since in a barrier market incumbent firms anticipate responses of potential entrants which face *free entry*. However, barriers enter the picture through the exit barrier which firms are facing before they have fully amortized the sunk cost incurred and the entry barrier which inferior potential entrants have to overcome in the case of actual entry. So, the barrier market theory deviates from perfect contestability, since the latter concept assumes the *total* absence of entry and exit barriers.

Subtle arguments underlie the validity of the combination of Assumptions 10.1 and 10.2, some of which are indicated in Chapters 9 and 11. Here a brief remark is made on the implications for the conditions of entry and exit. The incumbents' fear of entry after the introduction of the innovation is only rational if there is at least one potential entrant which does face free entry after the incumbents' introduction of sunk R&D, since otherwise incumbent firms are able to capture positive profits by setting a limit price. Here it is worth recalling that potential entrants' entry or exit costs (barriers), however small, are a source of barriers to entry (Baumol *et al.* 1982a, Chapter 11). This means that sunk costs cannot be asymmetric in favor of incumbent firms. Section 11.2 goes into this proposition in detail.

Baumol *et al.* (1982a) argue that

it should be emphasized that the need to sink substantial cost upon entry is not an entry barrier in itself (Baumol *et al.* 1982a, p. 291).

Shepherd (1984) notes that Baumol *et al.*'s

writings [mostly] stress exit as a major new element, which is separable from entry barriers and can be important even if entry barriers are not (Shepherd 1984, note 11, p. 576).

Entry barriers only emerge when sunk costs introduce a *disadvantage* for (potential)

entrants relative to incumbent firms. The barrier market framework applies to those cases where incumbent firms anticipate that the sunk cost incurred *ex ante* (in a two-period model this means in period $t = 0$) does not raise entry barriers *ex post* (period $t = 1$ in a two-period model) to all potential entrants (Assumption 10.2). Henceforth *ex ante* indicates decision making *before* an investment (innovation) is carried out, whereas *ex post* points to the circumstances occurring *after* an investment (innovation) has been introduced.

The investment costs incurred *ex ante* introduce *costly* exit to the *incumbent firms ex post* if they leave the market before the sunk investment is fully amortized. If an incumbent firm is underpriced by an entrant ($^ep < ^ip$), then it fails to recover fully the sunk costs invested. Forced exit can occur as a result of careless price setting. Let *p_t be the price associated with zero profits in a period t . The (perfect) barrier market theory applies to those cases where incumbent firms fear that there is at least one superior potential entrant which evaluates entry to be profitable for any $^*p_t < ^ep_t < ^ip_t$ (Assumption 10.2). Then, the optimal strategy of incumbent firms boils down to $^ip_t = ^*p_t$. Incumbents face forced exit whenever they offer a price $^ip_t > ^*p_t$. The exit costs introduce an additional incentive to avoid (forced) exit by complying with the zero profit condition.

The key point is that sunk capital can give an exit cost. This contradicts contestability. Recall that contestability is compatible with investment, if the investment cost not yet amortized can be recouped at any exit date. That is, contestability assumes that firms are always able to benefit from cost effective selling or alternative use of the capital outside the market, so that exit cost is zero by assumption irrespective of the date of exit. This is not the case in the barrier market. If exit occurs before the sunk investment is fully amortized, selling or alternative use of the sunk capital is not cost effective. That is, (part of) the investment cost is not fully recoupable outside the market. This means that (part of) the investment cost has to be recovered in the barrier market. For the sake of notational simplicity, unless indicated otherwise, investment cost is assumed to be fully unrecoverable¹: that is, the sunk investment is to be fully amortized in the market so as to avoid a negative payoff. Exit cost follows from the part of the sunk investment that is not yet amortized at the exit date.

This thesis focuses on innovation. In an innovative market environment superior potential entrants can be distinguished from inferior ones.² *Superior potential entrants* are defined as firms which have already introduced their sunk costs *ex ante* (either in a home market or in anticipation of future entry opportunities) to attain an innovation equivalent to the incumbents'. This means that incumbent firms and superior potential entrants have to bear a *zero* incremental sunk cost upon staying in (incumbents) or entering (superior entrants) the market *ex post*. The sunk costs incurred *ex ante* to introduce innovations are ancient history *ex post*. Thus, superior potential entrants face free entry after the introduction of the innovation.³

¹ Of course, pure exit cost can be introduced as firms can discount the net present value of pure exit cost (*i.e.*, cost incurred at the date of exit) in current calculations.

² Strictly speaking, the introduction of inferior potential entrants is not necessary at this point of the argument. However, their introduction clarifies the intuition. Moreover, inferior potential entrants play a role in subsequent chapters (particularly Chapters 11 and 23).

³ This scenario resembles Maks' (1986) equi-barrier market. Entrants' unsuccessful entry and incumbents' premature exit are associated with equal exit costs, since the equal sunk investment costs have not yet been recovered at the date of exit. However, ignoring the possibility of mistakes (and provided that potential

It immediately follows that *inferior potential entrants*, which have failed to mirror the incumbents' innovation, face *entry barriers*, since they only have access to the new technology or product after incurring *costly* (sunk) outlays. Inferior potential entrants have to bear a *positive* incremental sunk cost upon entering *ex post*. It is now intuitively clear that two incentives underlie the incumbent firms' eagerness to implement sunk innovative investments *ex ante*.¹ First, innovative investments are necessary so as to avoid exit *ex post* as a result of entry by superior potential entrants. Second, innovations limit the entry opportunities to inferior potential entrants, which then face entry barriers *ex post*.

10.3 A BENCHMARK MODEL OF THE SEQUENCE OF EVENTS, RULES OF THE GAME AND EQUILIBRIUM CONCEPT

Chapter 9 indicated that the sequence of events in a market is an important element of competition. The order of investment, entry and adjustment lag and retaliation delay are part of scenarios that give workably competitive outcomes. For the sake of illustration, this section gives the outline of an elementary discrete time model that describes competition over one particular investment. This model offers the time structure that is used in subsequent chapters. Moreover, the rules of the barrier market game and the equilibrium concept are clarified.

Transactions take place only once in a period t . Assume that $t = 0, \dots, T$: that is, the time interval from $t = 0$ to $t = T$ is the total period under consideration. In the total time period $0-T$ the investment technology (R&D opportunities) is (are) assumed to be constant. In period $t = 0$ the investment (R&D) technology becomes feasible, whereas in period $t = T+1$ the investment (innovation) under consideration is outdated [*i.e.*, superceded by a subsequent innovation so that the old (process or product) technology is uneconomic]. This means that before $t = T+1$ the invested sunk costs must be recovered.

B^i and B^e denote the period in which incumbent firms and potential entrants respectively start undertaking the investment. L represents the investment lag: that is, the introduction of the investment takes place L periods² after $t = B$. A potential entrant is able to enter E periods after the entry decision, whereas the incumbent firm can retaliate R periods after the occurrence of entry. The incumbents' adjustment delay is A periods, where $E < A$ so that A can be ignored: incumbent firms are not able to lower price just before actual entry occurs. Note that the time index t can indicate two meanings: (i) the date t at the beginning of the period; and (ii) the period's time interval $[t, t+1]$. The following time schedule can illustrate the model. Suppose that an incumbent signals (*via* price) an entry opportunity immediately after the introduction of the investment. Moreover, take the case where incumbent firms and potential entrants invest simultaneously ($B^i = B^e = B$) and $E < R$ (recall that the potential entrants' investment does not necessarily prepare entry). The following sequence of events then applies

entrants either can recover the sunk cost in a home market or only undertake investment in anticipation to profitable entry opportunities), a potential entrant will never enter unsuccessfully, since this strategy is not worthwhile. As in contestability theory, a potential entrant is able to monitor the incumbent firms' prices before entry (Section 10.3). Hence, entry only occurs when the evaluation of post-entry profitability on the basis of pre-entry prices (Assumption 6.3) indicates profit opportunities.

¹ These points are clarified in Chapter 14.

² Since incumbent firms and (superior) potential entrants are assumed to have access to the same investment technology, both face the same investment lag L .

(provided that $B+L+E+R < T+1$):

Without loss of content, this thesis assumes $B^i = 0$ and $L = 1$ throughout.

Moreover, the argument generally (particularly in Parts VI to X) proceeds along the lines of a special case: a simple two-period - two-market model with period $t = 0, 1$ and market $m = i, e$. Incumbent firms anticipate the entry threat by innovative potential entrants which capture a zero profit if they refrain from entry. In a two-market framework this means that incumbent firms in market $m = i$ take notice of the strategies of potential entrants in market $m = e$ by assuming that $^i s_{c0} = ^e s_{c0}$ and $^e \pi_1 = 0$ if the potential entrant refrains from entry (where $^m s_{c0}$ is the sunk R&D outlay in market $m = i, e$ at period $t = 0$ and π denotes the payoff). The two-market setting reflects three usual free entry assumptions (Chapter 6). First, it is supposed "that there is always an entrant around" (Grossman 1981, p. 1159). Second, incumbent firms and potential entrants have access to the same pool of technologies. Third, it is assumed that the opportunity cost of entry is zero (Chapters 11 and 12).

The products traded in market $m = i$ and $m = e$ are perfect substitutes. Both markets are characterized by localized demand. That is, only firms are mobile between markets. Here it can be plausible to think of a two-country setting, where potential entrants reflect the fear of imports (Calem 1988). There can be a *reciprocal* entry threat: incumbent firms in market $m = i$ are potential entrants into market $m = e$ and *vice versa*. The development of a (process or product) innovation takes a fixed period of time: that is, R&D outlays have to be incurred up-front in period $t = 0$. In period $t = 1$ the firm introduces the innovation. It sets a price p_1 and serves demand up to an output level q_1 . Following the introduction of an innovation a firm's profit $\pi_1 = p_1 \cdot q_1 - avc_1 \cdot q_1 - s_{c0}$ in period $t = 1$ can be calculated (where avc denotes nonsunk unit cost and discounting is ignored). For the innovation project $\{s_{c0}, p_1, q_1\}$ to be profitable, it has to meet two criteria: (i) it permits the recovery in period $t = 1$ of the sunk cost incurred in period $t = 0$ ($\pi_1 \geq 0$); and (ii) it does not invite entry in period $t = 1$ by an entrant offering a lower price or better product to consumers. A potential entrant has the opportunity to enter profitably in period $t = 1$ if the entrant's innovation project introduces a larger cost reduction or better product than the new technology or product of an incumbent firm. Such entry forces the incumbent firm to exit in period $t = 1$. To summarize: $B^i = B^e = 0$, $L = 1$, $E = 0$, $A = R = 1$ and $T = 1$.

In a barrier market incumbent firms and potential entrants play a two-stage game on innovation and price. The second-stage pricing game (in period $t = 1$) exerts influence on the first-stage R&D game (in period $t = 0$). Incumbents show a first-mover advantage in the sense that entrants can only outperform incumbents if they offer a strictly lower price or better product. In order to illustrate the incumbent firms' first-mover advantage, a four-stage structure is adopted for the second-stage R&D game. To reduce complexity, reciprocal entry is ignored except in Chapters 11, 12 and 21. The argument is described from the perspective of an incumbent firm in market $m = i$ which does not consider entry into market $m = e$ (potential entrants in market $m = e$ being aware of this attitude), because, for example, incumbent firms, contrary to the potential entrants, face an entry

and/or exit barrier.¹

In t^1 incumbent firms announce prices in their home market. The potential entrants' price responses follow in t^2 : that is, the entrants' entry prices are declared *after* the incumbents' announcement of home market prices. In t^3 both the incumbents and the entrants produce/supply output for/in market $m = i$. Supply is equal to the quantity demanded from the incumbent or entrant (d). For the game on process innovation this means, for instance, that firms compete over price only. So, as in contestability theory, for incumbent firms this implies that $^i q = ^i d > 0$ if $^i p \leq ^e p$ and $^i q = ^i d = 0$ if $^i p > ^e p$, whereas (potential) entrants face $^e q = ^e d > 0$ if $^i p > ^e p$ and $^e q = ^e d = 0$ if $^i p \leq ^e p$ (for the case without excess demand). The game is closed in t^4 , where transactions take place. The game is one which has complete and imperfect information. The structure of the game is summarized in Figure 10.1.

Figure 10.1
Structure of the Barrier Market Game

	Incumbents firms	(Potential) entrants
$t = 0$	Sunk investment: $^i sc_0$	Sunk investment: $^e sc_0$
$t = 1$ {	t^1 Introduction innovation Home market prices: $^i p_1$	Introduction innovation
	t^2	Entry market prices: $^e p_1$
	t^3 Home market output: $^i q_1$ ($= ^i d_1$)	Entry market output: $^e q_1$ ($= ^e d_1$)
	t^4 Market transactions	

The barrier market models make use of the (subgame) perfect equilibrium concept. As Tirole (1988) points out,

[i]n a Nash equilibrium the players take their opponents' strategies as given and therefore do not consider the possibility of influencing them. In games in which a player chooses some actions after observing some of his opponents' actions (what we call *dynamic* games), this conjecture is naive and leads to some absurd Nash equilibria [Perfect equilibrium] is a refinement of Nash equilibrium for dynamic games that mitigates the deficiencies of Nash

¹ For example, Calem (1988) introduces the assumption of one-sided entry (threat) by arguing that the incumbent firms' entry cost is sufficiently large to keep it from entering the potential entrants' market (p. 175) or, alternatively, by supposing that there exist legal or regulatory barriers which prevent incumbent firms from being potential entrants into market $m = e$ (note 5, p. 182). Reciprocal entry (threats) introduce(s) complicated strategic interactions (Bulow *et al.* 1985a, Pinto 1986, Cairns and Mahabir 1988, Calem 1988, Casas 1989 and Chapter 11 and 21). These complications can be ignored by assuming, first, zero opportunity cost of entry and, second, one-sided entry threat.

equilibrium (Tirole 1988, pp. 428-429; Tirole's italics).

The perfect equilibrium concept (Selten 1965 and 1975) gives in to this critique as it selects Nash equilibria that do not involve noncredible threats (Tirole 1988, p. 429). To be precise,

[a] (subgame) *perfect equilibrium* (Selten 1965) is a set of strategies for each player such that in any subgame the strategies (truncated to this subgame) form a Nash equilibrium. Thus, *perfection requires that strategies be in equilibrium whatever the location (understand subgame) in the game tree, and not only along the equilibrium path* (Tirole 1988, p. 429; Tirole's italics).¹

perfect equilibrium can be obtained through backward induction by "folding back the extensive game tree" (Tirole 1988, p. 429).

Strictly speaking, the barrier market benchmark game is one of complete and imperfect information.² *Complete* information implies that

each player knows (a) who the set of players is, (b) all actions available to all players, and (c) all potential outcomes to all players. ... This is like saying that each player knows the whole game tree, including the payoffs listed at each terminal node (Friedman 1986, p. 9).

The players' information is *imperfect*: on the one hand, they are not informed about the rivals' investment scale at the time of their own decision making on investment, since both incumbent firms and rivals decide simultaneously on investment at $t = 0$; on the other hand, price information is perfect in the sense that the potential entrants' observe the incumbents' prices at $t = 1$ ¹ before they decide on entry (prices) at $t = 1$ ².

10.4 IMPLICATIONS FOR MARKET BEHAVIOR

The barrier market framework can be regarded as a benchmark case of competition, because the underlying assumptions imply favorable market outcomes.³ This section briefly outlines the major implications for two reasons: the argument, first, illustrates the barrier market's definition and intuition and, second, puts subsequent chapters into perspective. Assume that the two-period - two-market model applies (Section 10.3): This means that *ex ante*, in a period $t = 0$, incumbent firms face the decision whether to invest sc_0 sunk costs so as to introduce an innovation in period $t = 1$. It pays the incumbent to invest in period $t = 0$, and, consequently, to undertake innovative activity that is associated with nonzero sunk costs (Assumption 10.1), if the incumbents anticipate that a

¹ The game tree and Nash equilibrium concept were introduced in Chapter 6 (Section 6.4.4). A subgame is defined "as a subset of the initial game tree that (1) begins with an information set that contains only one node, (2) is closed under succession (if a node is in the subgame, so are all of its successors), and (3) is such that all the information sets of the subgame are information sets of the initial game" (Tirole 1988, p. 429). The barrier market benchmark game (Figure 10.1) has two subgames: a first-stage game on investment (played in period $t = 0$) and a second-stage game on price (played during period $t = 1$).

² The repeated dumping game in Chapter 11 is an example of a game of almost complete information: "[t]he simplest example of such a game is the 'repeated game' in which a simple one-period simultaneous move game ... is repeated T times and, at date t , players know all the moves before t " (Tirole 1988, p. 431).

³ Here the conclusions only hold in a partial context. This is the common angle in the literature on industrial organization (Part III). For further discussion, see Chapter 24.

situation of free entry will prevail *ex post*. Let *sc_0 denote the sunk costs associated with the (expected) optimal innovation, while *p_1 represents the price level such that the sunk costs *sc_0 are precisely recovered in period $t = 1$. Under the free entry Assumption 10.2, the incumbent firm is forced to (i) invest *sc_0 and (ii) set *p_1 . The reason is straightforward. If an incumbent firm fails to introduce the optimal innovation or offers too high a price in period $t = 1$, then an entrant which does adopt the optimal innovation is able to push aside the incumbent. As in contestability theory (Assumption 6.7), it is assumed that incumbents prefer to stay in the market, rather than (to be forced) to exit, conditional upon capturing nonnegative profits. Two examples may illustrate the implications: process and product innovations.

First, Chapter 15 shows that under particular conditions there exists a unique entry-detering sc_0 that minimizes average total cost in period $t = 1$ (Van Witteloostuijn 1989c). Let atc_1 represent the minimum average total cost in period $t = 1$, and let *sc_0 denote the associated investment of sunk costs in period $t = 0$. Suppose that the incumbent invests a $^i sc_0 \neq ^*sc_0$. Then, $^i atc_1 > ^*atc_1$. Hence, a potential entrant which does invest $^e sc_0 = ^*sc_0$ is able to enter profitably in period $t = 1$, since $^e p_1 = ^*atc_1 < ^i p_1 = ^i atc_1$. Hence, only *sc_0 is an entry-detering investment, since only $^i p_1 = ^*atc_1$ forestalls entry. Second, Chapter 16 argues that under particular conditions there exists a unique *sc_0 such that a profit-maximizing and optimal (from the buyers' perspective) product innovation is offered in period $t = 1$ (Van Witteloostuijn 1990c). The equilibrium product innovation is associated with offering a product which embodies optimal utility u_1 to the consumers in period $t = 1$. Suppose that the incumbent invests a $^i sc_0 \neq ^*sc_0$. Then, $^i u_1 < ^*u_1$. Hence, a potential entrant introducing $^e sc_0 = ^*sc_0$ is able to enter profitably in period $t = 1$, since $^e u_1 = ^*u_1 > ^i u_1$. Therefore, again only *sc_0 is the entry-detering investment, since only *u_1 forestalls entry.¹

The implication that $^i p_1 = ^*p_1$ simply follows from the contestable market theory (Chapter 6, Subsection 6.4.4). In the case of process innovations, for instance, it means that $^i p_1 = ^*p_1 = ^*atc_1$. A crucial assumption is that entry is free *ex post* (period $t = 1$ in the illustrations); this follows from Assumption 10.2, which asserts that the incumbents expect that $^e sc_0 = ^*sc_0$ so that $^e atc_1 = ^*atc_1$ or $^e u_1 = ^*u_1$.²

In the two-period framework it is easily shown that, for a given sunk cost invested in a period $t = 0$ that must be recovered in a period $t = 1$, there is indeed a unique price - quantity combination ($^*p_1 - ^*q_1$) such that the associated revenues (provided that *q_1 can be sold) are just sufficient to recover the sunk costs invested at a minimum price level *p_1 (Chapter 13). In a multi-period framework, if the sunk costs invested must be recovered in a period from $t = 1$ to $t = T$ ($1 < T < \infty$) because after period $t = T$ the innovation concerned is outdated, then the time path of prices will matter. Let *avc* denote the average variable and fixed cost level; so *avc* does not cover the sunk costs invested. Chapter 14 argues that a necessary, though not sufficient, condition for entry deterrence is that

$$(10.1) \quad \sum_{t=1}^T \delta^t \cdot (p_t - avc_t) \cdot q_t \leq sc_0,$$

¹ Kim (1987) and Rashid (1988) present similar arguments.

² Of course, these results can also be derived in a setting with actual Bertrand competition (with firms' homogeneity). However, the (static) efficiency of actual Bertrand competition requires at least two incumbent firms. Moreover, the observation that potential competition can discipline incumbents' behavior is ignored.

where q denotes sales and δ the discount factor which takes into account the firm's time discount rate r [$\delta^t = \exp(-rt)$]. Equation (10.1) implies a zero-profit condition. Moreover, if a conventional (or unique cost-minimizing) zero-profit amortization rule prevails, then a necessary condition for entry deterrence is compliance with the rule, which gives a unique sequence of price levels p_t such that the invested sunk costs are precisely recovered. For the special case of process innovations Baumol *et al.* (1982a, Chapter 13) show that a determinate solution exists giving the cost minimizing sequence of price levels (p_t , where $t = 1, \dots, T$).

10.5 A BENCHMARK CASE OF COMPETITION

As noted in Section 10.2, sunk exit costs generate entry barriers to inferior potential entrants. However, so long as there is at least one potential entrant which faces no disadvantage relative to incumbent firms, the entry threat is credible. In effect, in order to retain the free entry condition (Definition 10.1), it is necessary that at least one potential entrant does face free entry (Assumption 10.1). If all potential entrants do face entry barriers (or exit costs), then Bain's (1956) entry barrier theory applies (Chapter 7 and Section 9.2). Suppose that potential entrants face a sunk entry cost e_c . Incumbent firms are now able to capture positive profits by offering a limit price such that $\pi = e_c$. Potential entrants' entry cost indicates an entry barrier (Baumol *et al.* 1982a). Of course, if the limit price exceeds the profit-maximizing price, then incumbent firms stick to the latter.

Table 10.1

Assumptions of Contestable, Barrier and Profit Markets

Market	Assumptions		
	Free entry	Incumbent firms	Costly exit Potential entrants
Contestable markets	Yes	No	No
Barrier markets	Yes	Yes	Yes
Profit markets	No	Yes	Yes

It may be illuminating to compare the assumptions of the theories of (perfect) contestable markets (and Bertrand and perfect competition, Chapter 6), (perfect) barrier markets (and workable competition, Chapter 9) and limit pricing (and imperfect competition; henceforth called profit markets, Chapter 7). The theories of contestable and barrier markets, unlike that of profit markets, are based upon an assumption of free entry. The contestability concept, unlike the frameworks of barrier and profit markets, is associated with costless exit for incumbent firms *and* potential entrants. In a profit market potential entrants have to overcome entry barriers, whereas *both* incumbent firms and potential entrants may face exit barriers. In a barrier market (superior) potential entrants face free

entrants have to overcome entry barriers, whereas *both* incumbent firms and potential entrants may face exit barriers. In a barrier market (superior) potential entrants face free entry, while both incumbent firms and potential entrants may have to incur exit costs. The point is that in a profit market, unlike in a barrier market, *all* potential entrants face barriers to entry that are higher than the *exit* barriers (which may even be absent). As Shepherd (1984) argues:

If entry barriers are higher, then exit barriers do not matter (Shepherd 1984, p. 578).

In Table 10.1 the assumptions are summarized.

The distinction of the three types of markets is worthwhile, since they are associated with *different* implications for market behavior. In contestable and barrier markets, unlike profit markets, a credible entry threat forces incumbent firms to satisfy a zero profit condition. Existing contestable market models, however, unlike the barrier and profit market frameworks, fail to take account of the case where firms' (innovative or strategic entry-deterring) sunk investment behavior can occur. Without going into further detail here, the implications are summed up in Table 10.2.

Table 10.2
Implications of Contestable, Barrier and Profit Markets

Market	Behavior	
	Zero profit condition	Entry-deterring investment
Contestable market	Yes	No
Barrier markets	Yes	Yes
Profit markets	No	Yes

A last remark is in order with respect to the approach that is followed as regards to the formal elaboration on a workability concept. The barrier market theory is a quest for market conditions that generate incentives to (incumbent and potential) suppliers to adopt careful pricing on the one hand and to innovate and satisfy buyers' desires on the other. So, at this point there is not yet an explicit analysis of a welfare function. Much of the literature on the tradeoff between static and dynamic efficiency focuses on a benchmark welfare function in the form of the sum of the producers' and consumers' surplus. Reid (1987) adopts a Bergsonian approach to workability. However, to avoid the immense difficulties that surround the application of welfare function in evaluating market performance (Ng 1979), in subsequent chapters the focus is on market conditions that imply incentives to firms to behave in the consumers' interest.¹

¹ More is said about this in Chapter 24.

BARRIER MARKET SCENARIOS

Sources of Potential Entry and Profit Incentives

KEY WORDS. *The barrier market is defined with regard to its implications for market performance. Part V examines scenarios which can give barrier market outcomes. That is, by playing around with assumptions Part V provides an insight into market conditions that can facilitate favorable performance. The literature on workably competitive scenarios suggests that the source of potential entry is of particular importance. Barrier market conditions have to tune in to the identity of potential rivals (Chapter 11). Does the entry threat ensue from new or unrelated firms or suppliers which are incumbent in a related market? Have the latter disposal of excess productive capacity or are they facing binding capacity restrictions? Moreover, profit incentives cannot be ignored (Chapter 12). Does an established firm escalate its commitment to incumbency if it invests a positive sunk cost in prospect of a zero profit (as the barrier market theory predicts)?*

 

[illegible]

10

BARRIER MARKET SCENARIOS: Investment Disincentives, Excess Productive Capacity and (Un)protected Home Market

11.1 INTRODUCTION: Sources of Potential Entry

Contestability theory *a priori* assumes a credible entry threat. That is, the entry threat is exogenous to the theory. Among the many critiques of the contestable market theory, one line stands out that is relevant in the current context. Shepherd (1984), Calem (1988), Cairns and Mahabir (1988) and Chapter 6 argue that the existence and credibility of the entry threat in a perfectly contestable market needs careful examination. What is the identity of these entrants that hit-and-run just to benefit from short-lived entry opportunities? If potential entrants are new firms that still need to build up capacity, their speed of response is unlikely to be as fast as hit-and-run requests. Besides, entry by new firms is unlikely to be associated with a zero sunk entry cost.¹ So, they probably are existing firms which contest, for example, not by investing in the entry market, but rather by exporting goods (Shepherd 1984, p. 584 and Green 1987, p. 485). Established firms are engaged in potential entry into each other's market (Calem 1988). For example, import competition does indeed constitute a major threat against market shares of dominant firms (Scherer 1980, p. 241). In the barrier market theory the credibility of the entry threat is even more crucial, because firms face sunk costs which have to be recovered.

Following Cairns and Mahabir (1988) and Calem (1988) home market arguments can be used to identify credible potential entrants. The key point is that the foundation of the credibility of the entry threat is facilitated by focusing on a multimarket framework (Van Wegberg and Van Witteloostuijn 1990a and 1990b). A multimarket approach is implicit in many accounts of the contestable market (for example, Baumol *et al.* 1982a, p. xxi). Making it explicit, however, breaks with one dominant assumption in the literature on entry and exit in industrial organization: the exogeneity of the *opportunity cost of entry*. The common assumption is that the opportunity cost of entry is zero: potential entrants earn a zero profit if they do not enter. Hence, they enter only if entry profits are strictly positive. This can be the case because the alternative to entry is either zero production and investment or stable home market profit. Home market profit does not influence the

¹ Contestability gets around this dilemma by assuming that capital is not sunk. That is, capital can be bought and sold in an outside capital market such that the selling firms only lose a user cost. However, this only begs the question: what rationale underlies the existence of such a perfect capital market?

entry decision because of the usual key assumption that entry leaves the entrant's home market profit unchanged. However, in a multimarket context the assumption of zero opportunity cost of entry is not so obvious.

Home market arguments rely on the assumption that both incumbent firms and potential entrants undertake sunk investment in their home markets: to be precise, sunk cost can be assumed *not* to be *asymmetric* in favor of the incumbent. This points to the possibility of reciprocal entry threats (Bulow *et al.* 1985a, Pinto 1986, Calem 1988 and Krugman 1989). In a setting with parallel sunk costs home market arguments, however, give rise to complicated intermarket spillovers. At this point, two examples are worth noting. First, if firms face nonbinding productive capacity restrictions (excess capacity), dumping can occur (Brander and Krugman 1983 and Pinto 1986). If a potential entrant has excess capacity, then this firm is, *ceteris paribus*, inclined to use this excess capacity to supply in the incumbents' market for any price above marginal entry cost, where the marginal entry cost of dumping a unit of production is calculated by taking account of intermarket spillover effects. For example, joint scale and/or scope economies imply that entry is associated with an increase of the profit margin in the home market (an explanation of this point follows in Chapter 21). Any additional revenues in excess of marginal entry cost facilitate profit. Moreover, an expelled incumbent firm may retaliate by reciprocal entry into the entrant's home market. Second, if firms have to take account of binding capacity constraints (fungible capital), leaving the home market can induce entry (Cairns and Mahabir 1988), because a

shift reduces capacity in the original market, thereby creating profits for remaining firms. That market will then be invaded by yet another entrant. But that leaves no room for the first invader to return to the original market (Cairns and Mahabir 1988, p. 271).

Intermarket spillovers can be avoided if the entry threat is assumed to ensue from new or unrelated existing firms which sink cost upon entry: that is, if sunk (and so exit) cost is supposed to be *asymmetric* in favor of incumbent firms (Maks 1986).¹ Maks (1986) argues that a credible entry threat can derive from alert potential entrants which will prepare for future entry if profitable entry opportunities are observed.

The opportunity cost of entry affects the incumbent firm's pricing rule. If the opportunity cost is zero, any perceived positive entry profit, however small, invites entry. Hence, to deter entry the incumbent firm sets the average total cost price. If the opportunity cost is strictly positive, a potential competitor enters only if price exceeds cost sufficiently so as to warrant entry profits that are at least equal to the opportunity cost of entry. This allows for an entry-detering price with a positive profit. The present chapter goes through this argument by examining the credibility of the entry threat for different sets of assumptions. This means that the literature on workably competitive

¹ Of course, other scenarios can be assumed as well. Two examples are worth noting. First, the inverse asymmetry - incumbent firms facing higher sunk and exit costs than potential entrants - can be assumed. It is then likely that potential entrants can benefit from a competitive advantage so that incumbents firms may not be able to avoid entry (and probably their own exit). Second, potential entrants may face lower sunk cost (for example, because imitation is less costly than innovation) but higher exit cost (for example, because incumbent firms can amortize rapidly during a significant "safe against entry" period as a result of a long entry lag) than incumbent firms. This case resembles that in Section 11.2, since exit costs reflect an entry barrier. These and other cases are ignored as the examples in this chapter suffice to illustrate the existence of barrier market scenarios.

scenarios (Chapter 9) is critically evaluated. Chapter 21 discusses the (implications of the) different sources of potential entry in a broader context.

The plan of the chapter can be indicated by three questions:

- (1) Is an entry threat by new (or unrelated existing) firms, which implies the assumption of asymmetric sunk cost in favor of incumbent firms, credible (Section 11.2)?
- (2) Can dumping strategies be avoided if firms face excess productive capacities (Section 11.3)?
- (3) Can a credible entry threat exist if potential entrants have to take into account the implications of binding productive capacity restrictions (Section 11.4)?

Section 11.5 summarizes the argument by stipulating the importance of the credibility of the entry threat.

11.2 INVESTMENT DISINCENTIVES AND ASYMMETRY OF SUNK COST

11.2.1 Two Necessary Assumptions for Perfect Barrier Market Outcomes

In a setting with parallel sunk costs home market arguments can give rise to complicated intermarket spillovers through use of excess productive capacity and reciprocal entry threats (Section 11.3 and 11.4). Intermarket spillovers can be avoided if the entry threat is assumed to ensue from new or existing firms which sink cost upon entry: that is, if sunk cost is supposed to be *asymmetric* in favor of incumbent firms (which, with an imperfect second-hand capital market, implies higher exit cost for potential entrants relative to incumbent firms). This section focuses on the plausibility of the alternative credibility argument that is proposed by Maks (1986). Maks' (1986) general setting avoids the assumption of parallel sunk cost by arguing that a credible entry threat can derive from alert (new or existing) potential entrants which will prepare for future entry if profitable entry opportunities are observed (Section 9.3). This preparation implies that potential entrants (i) have to sink a cost upon entry which incumbent firms can spare and (ii) anticipate a probability that they are able to outperform the incumbent firms' innovation after an entry lag. Intermarket spillovers are avoided. Excess entry capacity is built up against a sunk cost so that anticipative dumping is not feasible and home market losses are absent. New firms start from scratch, whereas existing firms keep serving the home market.

The key question is whether assumptions of asymmetric sunk and exit cost in favor of incumbent firms and outperformance after a positive entry lag, can give sustainable perfect barrier market outcomes. This section suggests a negative answer to this question. It appears that the outperformance condition implies the assumption that potential entrants can leapfrog over the incumbent firms' innovation. In fact, the intuition is that two very opposite assumptions are likely to be *necessary* to reach perfect barrier market results:

- (A) Credible potential entrants are not facing a sunk cost disadvantage (absence of asymmetry in favor of incumbent firms); and
- (B) Credible potential entrants cannot leapfrog over the incumbent firms' innovation if the latter adopt an efficient R&D policy (zero probability of leapfrogging).

If one (or both) assumption(s) is (are) dropped, perfect barrier market outcomes cannot be obtained. However, imperfect barrier market results can be reached, which supports

Maks' (1986) argument as regards investment alertness as a credible source of potential entry. Although potential entry does not discipline the incumbent firms up to the point where minimum average total cost pricing is adopted throughout, the entry threat does keep incumbents from capturing the blockaded entry profit. This proposition is illustrated below. Following Bain (1956), blockaded entry is defined as the case where barriers are such that established firms could price even at the monopoly level yet still not incur entry (Waterson 1984, p. 57). That is, blockaded entry occurs if the monopoly price falls below the limit price. Internal competition dictates a price which is so low that profitable entry is blockaded.

In advance, three remarks can be made. First, it should be noted that the argument presented in Section 11.2.2 and 11.2.3 shows a family resemblance to Schwartz' (1986) critical treatment of the robustness features of contestability assumptions and Fudenberg *et al.*'s (1983) patent race model. The result that the assumptions of absence of asymmetric sunk costs in favor of incumbent firms is necessary to reach perfect barrier market outcomes, has its mirror in the contestability literature (Chapter 6), whereas the patent race models point to the role of leapfrogging opportunities (Chapter 8).¹ Second, the model in this subsection focuses on the special case of process innovation. Third, the model is preliminary: only the payoff structure of the game is described without explicit calculation of equilibrium features. However, the preliminary model indicates, first, critical determinants and, second, intuitive propositions of barrier market scenarios. Future efforts can be directed at the game-theoretic analysis of the complicated potential rivalry described below.

11.2.2 Dropping Assumption (A): Potential Entrants Sink Cost upon Entry

The key implication of dropping Assumption (A) is that the potential competitors sink cost in reaction to the incumbent firms' price signals, while the incumbent firms have already sunk this cost: sunk cost is asymmetric in favor of incumbent firms. For the sake of convenience, assume that the incumbents' sunk cost is zero ($sc = 0$), whereas the potential rivals' sunk cost upon entry is positive ($sc > 0$).

The decisions of incumbent and potential competitors at time $t = 0$ can illustrate the argument. Suppose that a potential entrant faces a zero entry lag ($E = 0$; this assumption is dropped in Section 11.2.3), whereas an expelled incumbent firm is able to respond (that is, to re-enter) after R periods, where $R > 0$. The market ceases to operate after T periods (where T is allowed to approach infinity). The key point is that profitable entry requires amortization of the sunk entry cost (given an assumption of zero scrap value). The amortization has to take place during the period of the incumbent's exit: R periods after exit an expelled incumbent firm is inclined to re-enter for any $p_t > avc_t$ ($t \geq R$, recall that p indicates price and avc denotes the average nonsunk cost), since $p_t > avc_t$ is associated with a positive contribution to profit. This means that the entrant's (gross) payoff stream from period $t = R$ to $t = T$ is zero ($\pi_{R-T} = 0$): at $t = R$ the entrant either decides to adopt the entry-detering price or prefers to exit.

A potential rival decides to enter (at $t = 0$ with $E = 0$) if (provided that $\pi_{R-T} = 0$ and ignoring discounting)

¹ Leapfrogging can also occur if a change in technological opportunities over time permits a potential entrant to benefit from a second-mover advantage as the entrant, contrary to the incumbent, is not yet committed to an outdated R&D path. Section 11.2 ignores this case.

$$(11.1) \quad {}^e\pi_{0-R-1} = \left[\sum_{t=0}^{R-1} ({}^e p_t - {}^e avc_t) \cdot {}^e q_t \right] - {}^e sc_0 > 0,$$

where q is the quantity sold by the firm. On the basis of a common amortization rule incumbent firms can calculate the first period's entry-detering price (${}^{ED}p_0$) by setting profit ${}^e\pi_{0-R-1} = 0$, which gives a positive profit provided that both incumbent firms and potential entrants have access to the same nonsunk cost function (${}^i avc_t = {}^e avc_t$) or incumbents face a lower-cost technique than entrants (${}^i avc_t < {}^e avc_t$).

The incumbents' positive profit that follows from limit pricing can easily be illustrated for the two limit cases case where $R = 1$ respectively $R \rightarrow \infty$. $R = 1$ gives

$$(11.2) \quad {}^{ED}p_0 = avc_0 + {}^e sc_0 / {}^e q_0 > {}^{MC}p_0 = atc_0,$$

where ${}^{MC}p$ denotes the incumbents' (minimum) average (total) cost price. Farrell (1986) describes the case with $R \rightarrow \infty$. Suppose that ec denotes sunk entry cost per unit output (that is, amortization markup). The result is that incumbent firms set the limit price ${}^{ED}p_t = avc_t + {}^e sc_t$ throughout (i.e., for $t = 0, \dots, T$). So,

[i]n other words, if the (self-enforcing) etiquette is for incumbents to leave if entered against, then we get average-cost limit-pricing, and no entry in equilibrium. This result is not hard to understand: since there is no entry against incumbents in equilibrium, a potential entrant contemplating entry will calculate on the assumption that he would have the market forever, and thus amortize on that basis (Farrell 1986, p. 69).

For $R > 0$ the amortization rule is crucial. Farrell (1986) tacitly assumes flat amortization. The key point is that both incumbent firms and potential entrants have to calculate limit prices and entry opportunities respectively on the basis of a *common* amortization rule, since otherwise the market is unstable (Chapter 14). Provided that a common amortization rule is used for $R > 0$, any asymmetric sunk cost in favor of incumbents is associated with limit pricing by incumbent firms that yields a positive profit if ${}^i avc_t \leq {}^e avc_t$ (provided of course that actual competition in the incumbents' market permits a $p > {}^{MC}p$).

11.2.3 Dropping Assumption (B): Potential Entrants Can Leapfrog after a Positive Entry Lag

11.2.3.1 Leapfrogging

The assumption that ${}^i avc_t \leq {}^e avc_t$ is of course crucial. Contestability theory assumes ${}^i avc_t = {}^e avc_t$ at any exit date t : that is, barriers to entry (and exit) are absent. Farrell (1986) points out that

[i]f the entrant's and incumbent's costs differ, but are common knowledge, then entry occurs precisely when it saves costs overall ...; and the price will be the entrant's long-run average cost if the incumbent chooses to prevent entry (Farrell 1986, p. 70).

That is, profitable entry deterrence is feasible if ${}^i avc_t < {}^e avc_t + {}^e sc_t$ (where ${}^i avc_t > {}^e avc_t$).

The condition that ${}^i avc_t > {}^e avc_t$ points to leapfrogging of entrants over the

incumbents' innovation (Fudenberg *et al.* 1983). This subsection explores the implications for the incumbents' strategy if leapfrogging can occur. Incumbent firms sink a cost $^i sc_0$ in period $t = 0$ upon which potential entrants can react by incurring a sunk entry cost $^e sc_0$. Suppose that incumbent firms anticipate that a price above average total cost will increase the probability that potential entrants start to prepare for future entry by increasing R&D outlays. This means that the probability of at least one potential entrant sinking a cost (f^1) is an increasing function (F^1) of the pre-entry price in period $t = 0$ ($^{PE}p_0$). That is,

$$(11.3) \quad f^1 = F^1(^{PE}p_0), \text{ with } 0 \leq f^1 \leq 1,$$

where $dF^1/d^{PE}p_0 > 0$ for $^{PE}p_0 > ^i avc_0$. Moreover, the probability that at least one potential rival is able to leapfrog over the incumbent firm's innovation in period $t = E$ (f^2) follows from

$$(11.4) \quad f^2 = F^2(^i sc_0, ^e sc_0),$$

where $\partial F^2/\partial ^i sc_0 < 0$ and $\partial F^2/\partial ^e sc_0 > 0$.¹ Suppose that potential rivals face an entry lag (E) which exceeds the incumbents' adjustment delay (A) but which permits entry before the life time of the innovation is over: $0 < A = 1 < E < T$. Moreover, $R > E$: that is, an expelled firm faces a response lag after actual entry has occurred.² Assume that at the beginning of the entry period $t = E$ incumbent firms are well-informed about the R&D results of potential rivals.³ This means that, first, the incumbent's profit consists of two parts and, second, the incumbent firm immediately adopts the equilibrium (*i.e.*, profit-maximizing) strategy.

11.2.3.2 Incumbent Firms' Payoff

From period $t = 0$ to $t = E-1$ the incumbent can capture a temporary pre-entry gross profit ($^{PE}\pi$). So,

$$(11.5) \quad ^{PE}\pi = \sum_{t=0}^{E-1} \delta^t \cdot (^{PE}p - ^i avc) \cdot i q_t,$$

¹ A complication is that $^e sc_0$ in equation (11.4) is not easily defined if the number of potential entrants exceeds one. In this case both the number and size distribution of R&D budgets are relevant.

² So, not only the original incumbent firm but also the entrant faces a response lag R after actual entry has occurred (because the entrant is now incumbent).

³ Of course, the entry period can be modeled in many ways, depending on the informational assumptions made. Two other examples are worth indicating. First, suppose that the information on the potential entrants' innovation becomes available at the beginning of period $t = E+1$. Incumbents use period $t = E$ as a try out by adopting $^p p_E = ^i avc_E$. If a potential rival is able to leapfrog over $^i avc_E$, entry occurs with $^e avc_E < ^p p_E < ^i avc_E$. Otherwise, the potential entrant stays out, which indicates that the incumbent faces the opportunity to increase price to $^e avc_{E+1}$ ($\geq ^i avc_{E+1}$) in period $t = E+1$. Second, assume that only the (yes or no) entry decision of potential competitors is a source of information. Incumbents have to experience a sequence of price strategies from $t = E$ to some $t > E$ in order to find out the potential rivals' R&D result. The assumption that incumbents receive information on the potential entrants' innovation at the beginning of the entry period, simplifies matters considerably, since in this case incumbent firms immediately choose the equilibrium strategy (exit, limit price or blockaded entry price).

where $PE_{\pi} \geq 0$ for $PE_p \geq i_{avc}$ (assuming PE_{p_t} and i_{avc_t} constant for $t = 0, \dots, E-1$).¹ Gross profit indicates the payoff before the sunk cost is amortized. After E periods the incumbent firm adopts the equilibrium strategy (which is feasible, because $A < E$ and $E \leq 2 \cdot T$)² on the basis of the knowledge of the potential rivals' R&D result. In equilibrium three strategies (and associated payoffs) can occur.

First, if none of the potential entrants has sunk a cost up-front so as to prepare for potential (future) entry in period $t = E$, the incumbent firm can maximize profit by setting the blockaded entry price (BE_p). Potential rivals are impotent, because they have not obtained access to the incumbent firms' technology. The probability of this event happening, is indicated by $(1-f^1)$. The blockaded entry gross profit (BE_{π}) is

$$(11.6) \quad BE_{\pi} = \sum_{t=E}^T \delta^t \cdot (BE_p - i_{avc}) \cdot i_{q_t}.$$

Of course, $BE_{\pi} > 0$ (assuming BE_{p_t} and i_{avc_t} to be constant for $t = E, \dots, T$ and profitable internal market conditions).

Second, a potential entrant which has incurred a sunk cost up-front in anticipation to future entry opportunities in period $t = E$ but which has failed to obtain a better technology ($e_{avc_E} \geq i_{avc_E}$), is inclined to enter for any $e_{p_t} > e_{avc_t}$ (for $t \geq E$) so as to recover (part of) the sunk cost. This condition occurs with probability $[f^1 \cdot (1-f^2)]$. This means that the incumbent firm can set an entry-detering price (ED_p): $ED_{p_t} = e_{avc_t}$ for $t = E, \dots, T$. The entry-detering gross profit (ED_{π}) is

$$(11.7) \quad ED_{\pi} = \sum_{t=E}^T \delta^t \cdot (ED_{p_t} - i_{avc_t}) \cdot i_{q_t},$$

where $ED_{\pi} \geq 0$ for $i_{avc} \leq e_{avc}$ (assuming e_{avc_t} and so ED_{p_t} and i_{avc_t} to be constant for $t = E, \dots, T$).

Third, if at least one potential entrant is able to leapfrog over the incumbent firm's innovation ($e_{avc_E} < i_{avc_E}$), irreversible exit is inevitable. The incumbent firm is willing to stay in (or re-enter) the market for any price $p_t \geq i_{avc_t}$ for $t = R, \dots, T$. However, the potential entrants can profitably block this strategy by setting $e_{avc_t} < e_{p_t} < i_{avc_t}$. The probability of leapfrogging is $(f^1 \cdot f^2)$. The gross profit associated with enduring exit (EE_{π}) is zero as exit implies the absence of both production and sales. (So, EE_{π} implies gross profit if profitable re-entry is blocked).

The incumbent firm's net total profit from period $t = 0$ to $t = T$ (NT_{π}) follows from

$$(11.8) \quad NT_{\pi} = PE_{\pi} + (1-f^1) \cdot BE_{\pi} + [f^1 \cdot (1-f^2)] \cdot ED_{\pi} + (f^1 \cdot f^2) \cdot (EE_{\pi}) - i_{sc_0}.$$

Net profit follows from gross profit minus sunk cost. Since $EE_{\pi} = 0$, equation (11.8) reduces to

¹ This simplifying assumption (and subsequent similar ones) indicate(s) a condition of an average nonsunk cost function which shows constant returns in the relevant output interval. This can be Baumol *et al.*'s (1982a) flat bottom.

² If $A > E$, the pre-entry period cannot show an incumbents' strategy different from the post-entry period, which is a trivial case in the current discussion. If $E > 2 \cdot T$, incumbent firms are induced to sustain a entry-detering strategy up to period $t = T - E$ (including a number of periods in which entry can be effectuated), which complicates the analysis without changing its essence.

$$(11.9) \quad \begin{matrix} NT\pi = PE\pi + (1-f^1) \cdot BE\pi + [f^1 \cdot (1-f^2)] \cdot ED\pi - sc_0, \\ (\geq 0) \quad \quad \quad (> 0) \quad \quad \quad (\geq 0) \quad \quad (< 0) \end{matrix}$$

where the contribution of the profit or cost components to net total profit is indicated between brackets. Incumbent firms maximize equation (11.9) over the R&D outlay (sc_0) and pre-entry price ($^{PE}p_0$). The precise content of the incumbent's strategy depends particularly on the entry lag E , the life time of the innovation T , the discount rate δ and the probability functions F^1 and F^2 .

11.2.3.3 Potential Entrants' Payoff

A potential entrant decides to sink an entry cost if entry profits promise to be positive on the basis of the incumbent firms' pre-entry price and the probability of leapfrogging. The probability that the potential competitor is able to introduce a better technology than the incumbent's is denoted by f^3 , which depends on $^{e}sc_0$ through¹

$$(11.10) \quad f^3 = F^3(^{e}sc_0, ^i sc_0),$$

where $\partial F^3 / \partial ^{e}sc_0 > 0$, $\partial F^3 / \partial ^i sc_0 < 0$ and $F^3(0) = 0$. After entry the potential rival's gross payoff consists of three parts: the gross profit in the entry period (in period $t = E$), the gross profit before (from period $t = E+1$ to period $t = R-1$) and after (from period $t = R$ to period $t = T$) the incumbent firms' retaliation lag is passed.

If a potential entrant has incurred a sunk cost up-front in anticipation of period $t = E$'s entry opportunities, the entrant is prepared to enter for any $^e p_t > ^e avc_t$ for $t = E, \dots, T$ so as to recover (part of) the sunk cost. If the potential entrant is able to leapfrog over the incumbent's innovation ($^e avc_E < ^i avc_E$), entry occurs in period $t = E$ with a price slightly below the incumbent firm's average nonsunk cost (which indicates the point where the incumbent accepts exit). The gross profit in the entry period ($^{EP}\pi$) is

$$(11.11) \quad ^{EP}\pi = \delta^E \cdot (^{EP}p_E - ^e avc_E) \cdot ^e q_E,$$

where $^{EP}\pi > 0$, since $^e avc_E < ^{EP}p_E < ^i avc_E$.

The entrant expects to be able to increase price in period $t = E+1$ to the incumbents' pre-entry price in period $t = 0$, $^{PE}p_0$. This price level can be sustained up to period $t = R$ when the expelled incumbent firm is able to re-enter. It is here where signaling comes in (Chapter 7). *Ex ante* the potential entrant formulates its expectation of post-entry pricing opportunities on the basis the incumbents' pre-entry price signal in period $t = 0$, $^{PE}p_0$, rather than on the actual blockaded re-entry price, $^{BE}p_t$ (for $t = E, \dots, R-1$). So, the expected blockaded re-entry gross profit ($^{BR}\pi$) is

$$(11.12) \quad ^{BR}\pi = \sum_{t=E+1}^{R-1} \delta^t \cdot (^{PE}p_0 - ^e avc) \cdot ^e q_t,$$

where $^{BR}\pi$ critically depends on $^{PE}p_0$ and $^e avc$ (assuming that $^e avc_t$ is constant for $t = E+1, \dots, R-1$): $^{BR}\pi > 0$, since $^{PE}p_0 \geq ^i avc > ^e avc$.

¹ If the number of potential entrants exceeds one, F^3 differs from F^2 , since F^3 is defined for *one* potential entrant, whereas F^2 follows from the R&D strategies of *all* potential rivals.

From period $t = R$ to $t = T$ the expelled incumbent firm can and is willing to re-enter for any $i_{p_t} > i_{avc}$, so as to recover (part of) the sunk cost. Hence, the entrants' re-entry-detering price (RD_p) is equal to the incumbent's average nonsunk cost: $RD_p = i_{avc}$ (assuming i_{avc} to be constant for $t = R, \dots, T$). This gives the re-entry-detering gross profit (RD_π)

$$(11.13) \quad RD_\pi = \sum_{t=R}^T \delta^t (RD_p - e_{avc}) \cdot e_{q_t},$$

where $RD_\pi > 0$, since leapfrogging implies that $RD_p = i_{avc} > e_{avc}$.

If the potential entrant fails to introduce a lower-cost technology than the incumbent ($e_{avc_E} \geq i_{avc_E}$), entry is abandoned provided that the incumbent firm sets the entry-detering price $ED_{p_t} = e_{avc_t}$ for $t = E, \dots, T$. The gross profit associated with non-entry (LF_π) is of course zero, since production and sales are absent. (Net profit may be negative as a result of unrecovered sunk cost.) Bearing in mind that the probability of leapfrogging is f^3 , the net entry profit (NE_π) is

$$(11.14) \quad NE_\pi = f^3 \cdot (EP_\pi + BR_\pi + RD_\pi) + (1-f^3) \cdot LF_\pi - e_{sc_0}.$$

Since $LF_\pi = 0$, equation (11.14) reduces to

$$(11.15) \quad NE_\pi = f^3 \cdot (EP_\pi + BR_\pi + RD_\pi) - e_{sc_0},$$

(>0) (<0)

where the contribution of the profit or cost components to net total profit is indicated between brackets. Potential entrants maximize equation (11.15) over the R&D outlay (e_{sc_0}). The precise expected value of net entry profit particularly depends on the incumbents' response lag R , the life time of the innovation T , the discount rate δ , the incumbents' pre-entry price signal PE_{p_0} and the probability function F^3 .

11.2.3.4 Blockaded Entry and Imperfect Barrier Market

The discussion of three illustrative cases can indicate the implications of the assumption of a (non)zero probability of leapfrogging for the incumbent firms' performance. The first two cases are trivial. First, if $f^1 = 0$, none of the potential entrants prepares for future entry by sinking a cost up-front. Incumbent firms are able to capture the blockaded entry gross profit BE_π [equation (11.6)]. This case resembles Bain's (1956) blockaded entry scenario and the well-known theories of imperfect competition without entry (threats). Incumbent firms' profit is fully determined by internal market conditions. Second, if $f^2 = f^3 = 0$, the probability of leapfrogging is zero. Hence, potential entrants which incur a sunk R&D cost up-front are only able to reach a production technology such that $e_{avc_E} \geq i_{avc_E}$. Provided that incumbent firms are prepared to set the entry-detering price $ED_{p_t} = e_{avc_t}$ for $t = E, \dots, T$, the net entry profit NE_π is negative [equation (11.15)]. So, potential entrants' equilibrium strategy is to remain passive: $e_{sc_0} = 0$. The market shows blockaded entry.

The third case is interesting in the context of the barrier market theory. Assume that incumbent firms set the pre-entry price in period $t = 0$ equal to average total cost: $PE_{p_0} = i_{atc_0}$. This price follows from the common amortization rule (Chapter 14) which

predicts a zero net total profit if entry does not occur: $NT_{\pi} = 0$. Moreover, if the incumbent firm adopts the R&D strategy which it expects to give the lowest-cost technology (i^*sc_0), the incumbent undertakes entry deterrence as much as is feasible. Suppose that entry deterrence is successful: none of the potential entrants anticipates period $t = E$'s entry opportunities by sinking R&D cost up-front. One proposition suggests that sustained minimum average (total) cost pricing, and so a zero profit, is the incumbents' optimal choice, because it reduces the exit probability to zero. This proposition, however, does not hold, because $f^1 = 0$ permits incumbents to increase price to ^{BE}p in period $t = A$, since the market is associated with blockaded entry. Hence, a positive profit is captured.

It goes without saying that the strategy $^{PE}p_0 = iatc_0$ and i^*sc_0 minimizes f^1 respectively f^2 and f^3 [equation (11.3), (11.4) and (11.10) respectively].¹ However, does the incumbent firms' entry-detering strategy bring f^1 , f^2 and f^3 to zero so that blockaded entry is obtained? That is, the key question is whether potential entrants are inclined to remain passive if incumbent firms adopt $^{PE}p_0 = iatc_0$ and i^*sc_0 . The potential entrants' net entry profit function (11.15) indicates that this need not be the case if f^3 is positive, notwithstanding the incumbents' entry-detering strategy. The assumption that $F^1(iatc_0), F^2(i^*sc_0), F^3(i^*sc_0) > 0$ can be grounded by introducing heterogeneity of firms. The argument can run two ways.

First, if potential entrants face a superior R&D technology, e^*sc_0 gives a lower-cost production technique than i^*sc_0 . This argument implies that an essential contestability assumption is dropped, since incumbent firms and potential entrants do not have access to the same set of technologies. Moreover, if incumbent firms are aware of the existence of superior potential entrants, their equilibrium strategy can be to accept exit in period $t = E$ but to cream off temporary pre-entry profits on the one hand and to abandon sunk R&D on the other. If incumbents fail to anticipate entry by lower-cost entrants, forced exit in period $t = E$ goes hand in hand with exit cost. In both cases in the post-entry equilibrium higher-cost incumbent firms are expelled from the market, so that the heterogeneity of incumbent and potential competitors in favor of the latter is not sustainable if superior potential entrants actually sink cost up-front.

Second, technical uncertainty in combination with diverging expectations (that is, potential entrants having expectations different from incumbent firms) can also introduce $e^*sc_0 \neq i^*sc_0$. If either e^*sc_0 or i^*sc_0 gives the lowest-cost technique (that is, if one of the expectations models is superior), blockaded entry (if incumbents introduce the lowest-cost technology) or equilibrium-restoring entry (if potential entrants have the lowest-cost production technique) occurs. As with potential entrants with better innovation technologies, heterogeneity of incumbent and potential competitors in favor of the latter is unsustainable if superior potential rivals actually undertake sunk R&D.²

¹ A key point is that the manipulation of information can be used to reduce the attractiveness of entry. The strategic feature of entry deterrence through signaling can take many shapes. Stigler (1966) and Smith (1981) point to two illustrative examples worth noting in the current context. Stigler (1966) observes that incumbent firms can avoid revealing the profitability of operations to make the attractiveness of entry less apparent. Smith (1981) argues that the announcement of strategic investment projects can be employed so as to discourage potential entrants.

² The case with diverging expectations is illustrated in Chapter 23. Note that other equilibria may occur. For example, with appropriate lag conditions (in particular, A, E and R) an incumbent and entrant can decide to take turns on entering in combination with above average cost pricing. The entering firm prices the rival out of the market at the margin. This implies that the dynamic adjustment path toward the zero

By way of conclusion it can be argued, on the one hand, that Maks' (1986) scenario cannot be retained if the assumption is adopted that $f^1 = F^1(atc_0) = 0$. On the other hand, the assumption that $f^1 = F^1(atc_0) > 0$ can only be sustained if potential entrants expect that they can outperform incumbent firms even when the latter adopt i^1, sc_0 : $f^2 = F^2(i^1, sc_0) > 0$ and $f^3 = F^3(i^1, sc_0) > 0$. This implies the assumption that incumbent firms differ from potential entrants (either with respect to R&D capabilities or technical expectations), which is likely to give unsustainability. The argument's intuitive result is the observation that a positive probability of *leapfrogging* can induce imperfect barrier market outcomes if the parameter set $\{E, R, T, \delta\}$ and functional specifications of F^1 , F^2 and F^3 indicate the incumbents' equilibrium strategy to lower price below the blockaded entry price (but above average total cost) during an interval within the life time of the innovation. The argument in this section suggests (but does not prove) the intuition that if the entry threat ensues from (new or unrelated) firms which have to sink a cost upon entry (and which can face a positive exit cost), imperfect barrier market outcomes can be obtained under the assumption of a positive probability of leapfrogging.

11.3 EXCESS PRODUCTIVE CAPACITY AND DUMPING

11.3.1 Entry with Excess Productive Capacity

Cairns and Mahabir (1988) suggest that the credibility of the entry threat ensues from firms with excess capacity in related markets. They

argue that firms in related industries have an advantage *because* of their own sunk costs. An advantage (over completely new firms or firms in unrelated industries) may arise if costs are sunk and (1) there is short-run excess capacity because of unpredictable demand fluctuations; (2) there are regular peak and off-peak periods and the firm is actively seeking a use for its off-peak excess capacity; (3) it is normal to have some capacity reserve that can be utilized if need be, perhaps at higher operating costs; (4) the firm has excess capacity, created in order to deter entry to its *own* market ...; or (5) the firm is a member of a monopolistically competitive industry and, for that reason, has excess capacity (Cairns and Mahabir 1988, p. 273; Cairns and Mahabir's italics).

However, the proposal of the use of excess productive capacity so as to ground the credibility of the entry threat is subject to two critical remarks.

First, excess capacities are not always sustainable in a barrier market (Subsection 11.3.2). Second, firms with idle productive capacity have an incentive to dump (Subsection 11.3.3). If a firm has excess capacity, then this firm is, *ceteris paribus*, inclined to use this excess capacity so as to supply commodities in the rivals' market for any price above marginal entry cost. Here this behavior is called *dumping*. The barrier market theory is incompatible with dumping, because either the incumbents' retaliatory dumping threat invalidates the credibility of potential entry or reciprocal dumping strategies give a negative payoff to both parties. Subsections 11.3.2 and 11.3.3 explain these results.

profit state gives a positive profit to both rivals. The key point is that the lag conditions are crucial determinants of the equilibrium strategies.

11.3.2 Sustainability of Excess Capacity

Excess capacity can exist for several reasons. In addition to Cairns and Mahabir's (1988) list, five further sources are worth mentioning. First, it can be that a (natural) monopolist or firm which serves a residual demand, produces short of the cost minimizing output level *q . This yields an excess capacity that follows from $^*q - i_q$, where i_q is the actual quantity supplied in the home market. The firm is a potential entrant if it endeavors to expand output to *q and selling the additional output in the rival's market (which also decreases unit cost of production). Second, a firm which faces a flat-bottomed average cost curve can expand output without increasing unit cost up to an upper limit of output. Suppose that the constant returns interval of the cost function ranges from L_q to R_q , where $0 < L_q < R_q$. Beyond R_q decreasing returns set in. If $i_q < R_q$, then $R_q - i_q$ can be regarded as excess capacity. Third, footloose productive capacity that is installed so as to underprice incumbent firms constitutes excess capacity. Fourth, firms which are expelled from the home market face idle productive capacity. Fifth, unsold inventories are a source of excess supply. This also brings in an intertemporal aspect. Provided that the goods involved are nonperishable, a firm that fails to sell its products in the current period, because it is underpriced, can supply these commodities in a future period.¹

Barrier market theory indicates that excess productive capacity can be unsustainable, however. A positive sunk cost is likely to introduce an average total cost function with a *unique* cost minimizing output level *q , since sunk cost is by definition subject to global increasing returns (Chapter 13). This means that contestability theory predicts that a nonmonopoly market configuration is unsustainable (Baumol *et al.* 1982a). Operation at the left hand side of *q implies that too many firms are incumbent (inducing exit of some incumbent firms), whereas production at the right hand side of *q indicates entry opportunities (actual entry forcing incumbent firms to contract output). Cairns and Mahabir (1988) point out that

[o]ne possible source of entry by existing firms is excess capacity. It is not unlikely that Baumol *et al.* had in mind the possibility of such a mechanism, as they cite Clemens (1951, pp. 1-2) ... The existence of such idle capacity with which firm 2 could move into firm 1's market would be impossible in a perfectly contestable market, however. If capacity were truly idle, a putative firm 3 could have earlier built the optimal capacity in market 2 and replaced firm 2. One possible way out is that a firm could use short-run reserve capacity to move into new lines, as suggested by Andrews (1949). Andrews cites the use of contingency reserves and overtime. But contestability theory precludes reserve capacity, since it is presented in a

¹ The first two sources of excess capacity are worth considering in little more detail in the context of contestability theory. The first case offers a critique of the treatment of (demand or natural) monopoly in the theory of contestable markets. Take, for the sake of convenience, a demand monopoly. The literature (Baumol *et al.* 1982a, p. 31) often refers to a contestable monopoly with $m_q < L_q$ (or *q). In a multimarket framework it is, however, unlikely that this monopoly configuration is sustainable, since the market is contestable only if other markets exist in which the firms are potential entrants into the monopolist's market; otherwise, the existence of potential entrants which can outperform the monopolist before the latter is able to respond, is not very likely. However, if such potential entrants are around, the monopolist prefers to produce *q rather than m_q so as to sell excess supply $^*q - m_q$ in other markets. If the firm fails to produce *q units of goods, unit cost exceeds the level of minimum average cost so that potential entrants can underbid the monopolist. The second case strengthens the significance of the flat-bottomed cost curve. The flat bottom is more than a condition that facilitates the likelihood of static sustainability (Baumol *et al.* 1982a, pp. 31-40). It allows an entrant to expand output so as to take over an incumbent's market, while continuing to serve the home market. This supports the credibility of the entry threat.

context of certainty, contingency reserves would be redundant and wasteful. Overtime would require higher rewards to factors than in the original equilibrium; therefore, it would not be competitive with efficient production. Thus, latent idle capacity is impossible in a contestable market (Cairns and Mahabir 1988, p. 271).

This suggests that only particular sources of excess capacity can be sustainable in a barrier market. First, internal market conditions can indicate the efficiency of excess capacity: for example, excess capacity is maintained so as to satisfy regular peak periods or unpredictable demand fluctuations. Second, potential entrants with excess capacity can have home markets which are noncontestable: that is, entry into (and exit from) the potential entrants' home market is (can be) costly. Third, excess capacity can be costless. The buildup cost of productive capacity is already amortized and excess capacity is associated with a zero holding cost. Fourth, if firms face the need to sink cost so as to install an efficient productive capacity which provides scale economies in terms of average nonsunk cost, the equilibrium strategy in a multimarket setting with reciprocal entry can be to build up excess capacity (Chapter 21).¹

11.3.3 Reciprocal Dumping

Actual entry is likely to diminish the incumbent firm's profit - if it does not drive the incumbent firm off the market altogether. The incumbent firm may retaliate by reciprocal entry into the entrant's home market (Bulow *et al.* 1985a and 1985b and Calem 1988). Bulow *et al.* (1985a and 1985b) point out strategic interactions that can occur in a multimarket framework. In an example of two monopolists A and B, which are potential entrants into each other's markets, they argue that

B's entry will change A's equilibrium output in the market where it is incumbent and therefore possibly alter its decision of whether to enter B's market. ... If B enters A's market then it may be profitable for A to retaliate. So the threat that deters B's expansion is a credible one (Bulow *et al.* 1985a, p. 505).

The point is that exit gives idle productive capacity. The expelled incumbent firm can increase profits (or decrease losses) by selling output in the entrant's market for any price above average marginal entry cost. As exit is associated with zero production, an expelled firm can profitably enter the entrant's market for any price above average variable cost. Dumping by expelled incumbent firms does not undermine the credibility of the entry threat either if the potential entrants' home market is safe against entry by expelled incumbent firms (Subsection 11.4.2) or if the incumbents' response lag is large enough to invalidate the retaliatory dumping threat (Subsection 11.4.3). The case is different, however, if both incumbent firms and potential entrants have sustainable excess capacities (Subsection 11.3.2). For illustrative purposes, an argument can be put forward so as to defend the assumption that dumping does not occur.² Assume that potential entrants take

¹ A further complication arises if the use of excess capacity gives economies of scale (single product monopoly) or scope (multiproduct diversifying firm). This intermarket spillover is recognized by Bulow *et al.* (1985a). Chapter 21 discusses this issue further.

² Of course, if potential firms face a *binding capacity constraint* the case is different. First, it can be that potential entrants cannot feasibly produce commodities in excess of H_q . If the feasible quantity demanded from a firm in the home market is $H_d \geq H_q$, then the capacity constraint is binding. The potential entrant will, *ceteris paribus*, only transfer supply to the incumbents' market if $p > p^*$ (Section 11.4). Second,

into account the threat of *reciprocal* dumping. That is, incumbent firms have the opportunity to dump into the potential entrants' home market. Therefore, incumbent firms and potential entrants play a game on excess capacity (dumping) and price, which may or may not yield a Nash equilibrium that is associated with (im)perfect barrier market outcomes and the absence of (anticipative) dumping.

Table 11.1
Payoff Matrix Perfect Barrier Market with Dumping

		Firm j	
		No Dumping	Dumping
Firm i	No Dumping	(0,0)	(Y,X)
	Dumping	(X,Y)	(Z,Z)

Take two symmetric representative firms i and j in market m_1 respectively m_2 . Both firms play a (Bertrand-)Nash game on dumping. Ignoring lag conditions, the payoff matrix is indicated in Table 11.1 [depicting (π, π) per strategy pair].

The static (one-period) dumping game clearly reflects a noncooperative Prisoners' Dilemma, since $Y < Z < (0 <) X$, which yields a Nash equilibrium in which both firms decide to dump (Z,Z), although the solution where neither firm dumps (0,0) is preferable. The dynamic (multiperiod) repeated dumping game is more complicated (Friedman 1986). At least three (related) arguments can be used to indicate that the strategy pair that neither firm decides to dump, is a (Nash) equilibrium.

First, *reputation* considerations can call into question the equilibrium features of the case where both firms are dumping. Firm i can communicate to firm j that firm i 's dumping will be followed by reciprocal dumping, which leaves firm j worse off.¹ This argument bears a family resemblance to the predation literature (Milgrom and Roberts 1982 and Roberts 1986; see also Chapter 7). Second, the reputation effect wins cogency if firms have or behave as if they have an *infinite* planning horizon. In effect, the Folk

potential entrants can deliver from inventories. Inventories are built up so as to meet demand in the home market only. Anticipative use of excess capacity, in order to build up inventories that permit dumping in the future, does not occur. The role of the *limited inventories* is equivalent to the one of binding capacity constraints.

¹ Moreover, firm j has to take into account the observation that, if firm j repeatedly decides to dump, it is likely that firm i will start to *expect* that firm j will dump. This induces firm i 's best *anticipative* reply to dump as well, which leaves firm j worse off.

Theorem¹ shows that the preferred payoff combination can be a Nash equilibrium in an infinitely repeated Prisoners' Dilemma (Pinto 1986).² The assumption of planning as if the horizon is infinite is plausible, since firms are generally uncertain of the date at which they will cease to operate. Fisher (1989) points out that

real-life incumbents do not face a well-defined finite set of potential entrants. Corporations in most contexts are assumed to have an infinite horizon and surely cannot believe that any particular fight will be the last (Fisher 1989, p. 123).

Third, firms can (tacitly) agree upon not dumping. The dumping game is noncooperative but gives (tacitly) *cooperative* outcomes. The case where neither firm dumps is clearly the cooperative outcome.³

11.4 (UN)PROTECTED HOME MARKET AND LAG STRUCTURE

11.4.1 Home Market Disequilibrium

Binding productive capacity constraints can introduce a positive opportunity cost of entry. Binding capacity restrictions force the firm to give up sales in its home market in order to be able to serve an entry market. Cairns and Mahabir (1988) point out that

[t]he entrant shifts capacity from an original [market] ..., expecting to return after a hit-and-run entry. But that shift reduces capacity in the original market, as compared with the original equilibrium there, forcing a disequilibrium in which price must rise to clear the market, thereby creating profits for remaining firms. That market will then be invaded by yet another entrant. But that leaves no room for the first invader to return to the original market (Cairns

¹ To be precise, the Folk Theorem "clarifies the role of trigger strategies in supporting cooperative outcomes by means of noncooperative equilibrium strategy combinations" (Friedman 1986, p. 104). The theorem originates from Aumann (1981) in a setting without discounting. Elaborations (with discounting) are offered by Fudenberg and Maskin (1983), Abreu (1983), Benoit and Krishna (1985) and Friedman (1985). High discount factors facilitate cooperative outcomes. Friedman even shows that "it is obvious that trigger strategies can support in finite horizon repeated games the same sort of outcomes that they can support in infinite horizon games: those that give to each player a payoff no lower than the worst payoff he can receive under a single-period Nash equilibrium" (Friedman 1986, p. 104). Moreover, Radner's (1980) epsilon equilibrium concept yields the result that "if players are content to use strategies that are within ϵ of being best replies, then trigger strategy ϵ -equilibria are possible in games having sufficiently long, but finite, horizons" (Friedman 1986, p. 94). Shapiro (1989a) offers a survey of the literature on (both Bertrand and Cournot) trigger strategies.

² (Reciprocal) dumping models are explored in the theory of international trade in particular (Brander and Krugman 1983, Pinto 1986 and Krugman 1989). The natural frame of reference in international economics is of course multimarket (or country) models. The key question focuses on a quest for the conditions that are associated with the existence or absence of (reciprocal) trade or dumping: in terms of industrial organization, under what conditions is (reciprocal) entry the equilibrium strategy (pair) and when is entry-deterrence successful? Pinto (1986) points out that dumping does not occur in models with infinite repetitions and sufficiently little discounting.

³ Pinto (1986) presents an argument that bears a family resemblance to the one presented above. Pinto's model shows in an international trade model with a Cournot duopoly that the static Prisoners' Dilemma yields reciprocal trade, whereas the repeated Prisoners' Dilemma can constitute a Nash equilibrium that is associated with the absence of trade. Although both firms face transportation costs (an entry barrier), the sunk cost that follows from investment (in innovation) is ignored. Strictly speaking, Pinto's model does therefore not describe dumping.

and Mahabir 1988, p. 271).

Hence, entry has repercussions to the home market. Since entry affects home market profit, a potential entrant takes the implications for home market demand into account. Revenues foregone by undersupplying the home market constitute the opportunity cost of entry.

In Subsections 11.4.2 and 11.4.3 six scenarios are described where existing firms operating elsewhere imply a credible entry threat, because they (i) have fungible capital or (ii) face a negligible transportation cost disadvantage. That is, there are potential entrants which can easily switch productive capacity or transfer (part of) the commodities produced to the incumbents' market. Recall that demand is localized so that only firms are mobile between markets (Section 10.3). Moreover, it is assumed that neither leapfrogging (Section 11.2) nor anticipative (reciprocal) dumping (Section 11.3) does occur.

Recall that incumbent firms announce prices and potential entrants decide on entry at the beginning of period $t = 1$ on the basis of the result of R&D in period $t = 0$ (investment lag $L = 1$). Prices remain fixed during a period¹ (adjustment delay $A = 1$). Here this means that, *ceteris paribus*, the response of incumbent firms in the case of exit takes one period: that is, the response lag is one period ($R = 1$). Discounting is ignored. Recall that before $t = T+1$ the invested sunk costs must be recovered (life time of the innovation expires after period $t = T$). Moreover, the home market argument is based upon the assumption that $^i\text{avc}_t = ^e\text{avc}_t = \text{avc}_t$ and $^i\text{sc}_0 = ^e\text{sc}_0 = \text{sc}_0$ (incumbent firms and potential entrants make use of the same production and investment technology), while potential entrants enter the market if they perceive any profit opportunity (Assumption 10.2). This is Cairns and Mahabir's (1988) and Calém's (1988) prototype market argument (Section 11.3). Note that the outcome of the investment sc_0 can be applied to both markets. However, with binding capacity restrictions a firm is not able to serve both markets simultaneously.

This section explores barrier market scenarios by playing around with assumptions as regards to the lag structure and potential entrants' home market conditions. The key argument is that sets of assumptions can be traced which give perfect barrier market outcomes. The scenarios can be divided into two types on the basis of potential entrants' home market conditions. Section 11.4.2 examines two examples in which potential entrants come from a protected home market: that is, the entrants' home market is safe against entry. Section 11.4.3 presents four cases that start from the assumption that the entrants' home market is unprotected. In this case Cairns and Mahabir's (1988) argument on the repercussions of home market disequilibria is relevant.

11.4.2 Protected Home Market

Home market arguments can be facilitated by a restriction to one-sided entry (threats). That is, the potential entrants' home market is safe against entry. This means that entrants can anticipate an easy return to the home market after hit-and-run entry. For example, Calém (1988) introduces the assumption of one-sided entry (threats) by arguing that the

¹ This assumption is in accordance with Hicks (1939, p. 122). The identification of a period with the interval in time during which prices are constant is nowadays adopted in the popular "temporary equilibrium" tradition (Grandmont 1977). See also Chapter 20.

incumbent firms' entry cost is sufficiently large to keep it from entering the potential entrants' market (p. 175) or, alternatively, by supposing that there exist legal or regulatory barriers which prevent incumbent firms from being potential entrants into market $m = e$ (note 5, p. 182).

It is important to distinguish the case where $T = 1$ from the one where $T > 1$. For $T = 1$, unlike $T > 1$, the assumption of (potential) entrants' Bertrand-Nash expectations (Assumption 6.3) is easily sustained.

EXAMPLE 11.1 (short-lived innovations). For $T = 1$ it is obvious that a potential entrant is able to adopt "hit-and-run" entry. Sunk cost must be recovered in period $t = 1$. If an incumbent firm is expelled from the market as a result of careless price setting in period $t = 1$ ($p_1 > p_1^*$), then the incumbent firm fails to recover the invested sunk costs, because the incumbent can only re-enter after one period ($R = 1$ so that re-entry cannot take place before period $t = 2$ when the innovation is outdated):

$$(11.16) \quad i\pi = -sc_0.$$

Hence, the expelled incumbent firm faces an exit cost. The entrant can capture positive profits so that no exit costs are incurred, provided that re-entry only occurs for a price $avc_1 < p_1 < p_1^*$:

$$(11.17) \quad e\pi = [(p_1 - avc_1) \cdot q_1] - sc_0 > 0.$$

Condition (11.17) gives entry if the potential entrant captures a zero profit in the home market. $T = 1$ corresponds, for instance, to the case where firms introduce incremental, short-lived innovations so that R&D expenditures must be amortized in a short (one) period of time.

For $T > 1$ it is likely that potential entrants anticipate the response of an expelled incumbent firm at the beginning of period $t = 2$. Hence, they only temporarily employ Bertrand-Nash expectations. Assumption 6.3 is now in need of modification.

ASSUMPTION 11.1 (re-entry). Potential entrants only employ Bertrand-Nash expectations for the first period after entry. They anticipate correctly the response of an expelled firm at the beginning of $t = 2$ (or $R+1$; see Example 11.4).

The likely response of an expelled incumbent firm follows from the argument that an incumbent is inclined to re-enter after exit at the beginning of period $t = 2$ for any price in excess of its average variable cost (that is, in excess of the shut-down-price). A price $p_t > avc_t$ (where $t > 1$) enables the recovery of (part of) the sunk cost invested.¹ It is the sunk nature of investment expenditures that induces the incentive to an expelled incumbent firm to re-enter for any $p_t > avc_t$, because the opportunity cost of refraining from entry ($i\pi_A$) exceeds the cost of re-entry ($i\pi_E$):

¹ This behavior does not imply dumping (as defined above), because dumping is associated with supply in excess of the quantity already supplied elsewhere so as to recover the sunk cost. That is, dumping yields a positive profit.

$$(11.18) \quad i\pi_A = -sc_0 < i\pi_E = \left[\sum_{t=2}^T (i p_t - avc_t) \cdot i q_t \right] - sc_0 < 0,$$

for any $i p_t > avc_t$. There are now five examples of cases where incumbent firms face a credible threat of entry, despite of their ability to respond with low-priced re-entry at the beginning of period $t = 2$. One example fits into the set of scenarios which assumes a protected home market condition.

EXAMPLE 11.2 (temporary cross-entry). Potential entrants deploy temporary cross-entry. Before the incumbent firm is able to respond, the entrant returns to its home market. Suppose that the entrant is able to sell a quantity H_{q_t} at price H_{p_t} in the period $2 - T$ after returning to its home market at the beginning of period $t = 2$. If an incumbent firm sets $i p_1 > p_1$, then the potential entrant enters even for $T > 1$ if the latter captures positive profits

$$(11.19) \quad e\pi = [(e p_1 - avc_1) \cdot e q_1] + \left[\sum_{t=2}^T (H_{p_t} - avc_t) \cdot H_{q_t} \right] - sc_0 > 0,$$

for any $e p_1 > p_1$. Hence, if condition (11.19) holds, the entrant does not face any exit cost. Re-entry of the expelled firm at the beginning of period $t = 2$ is associated with $i p_t = p_t$ (where $t = 2, \dots, T$) so as to avoid further exit. Hence, the incumbent firm, again, faces exit costs:

$$(11.20) \quad i\pi = \left[\sum_{t=2}^T (p_t - avc_t) \cdot i q_t \right] - sc_0 < 0.$$

Condition (11.19) corresponds to the case where (internal) conditions in the potential entrant's home market dictate price $H_{p_1} = p_1$ in the period $t = 1$, so that the opportunity cost of refraining from entry for any $e p_1 > p_1$ is positive, conditional upon the assumption that the potential entrant is able to return to its home market at the beginning of period $t = 2$ without suffering negative consequences (in the sense of losing the home market). So, $H_{p_1} = p_1$ gives the threshold entry-detering price level $i p_t = p_t$ where the potential entrant is indifferent between entry and non-entry.

11.4.3 Unprotected Home Market

Further examples follow from dropping the assumption of an easy return to a home market. Then, although an entrant is able to "hit" at the beginning of period $t = 1$, it is unable to "run" at the beginning of period $t = 2$. It may be that a return to the home market is impossible altogether or that a profitable return is only feasible after the sunk costs have been fully recovered at period $t = T$. Section 11.4.2 assumes that the incumbent firm faces *idle* productive capacity during the period of exit. This need not be the case. For example, it can be that the expelled incumbent firm enters *another* (perhaps even the entrant's) market. By shipping goods to an entry market, entrants without excess capacity withdraw supply from their home market. This may invite further entry into the entrant's market. Hence, when the entrant returns from its hit-and-run, it is likely to discover that it has lost its home market (Cairns and Mahabir 1988 and Van Wegberg and Van Witteloostuijn 1989). The next examples assimilate this argument by focusing on

"hit-and-stay" entry.

EXAMPLE 11.3 (small R&D budgets). For $T > 1$ the potential entrant responds with entry to a ${}^e p_1 > {}^* p_1$ (provided that home market price ${}^H p_1 = {}^* p_1$, which implies a zero opportunity cost of entry) if it is profitable to deter re-entry by the expelled incumbent firm by reducing ${}^e p_1$ to avc_2 at the beginning of period $t = 2$. Hence, if

$$(11.21) \quad [({}^e p_1 - avc_1) \cdot {}^e q_1] - sc_0 > 0,$$

then the potential entrant is able to capture positive profits without inducing re-entry. Condition (11.21) implies that a potential entrant is inclined to enter if the entry price permits the recovery of the sunk cost in one period. If condition (11.21) holds, the entrant faces no exit costs. Then, successful entry by the expelled firm requires ${}^e p_2 < avc_2$, which is not worthwhile, since condition (11.18) is violated. The expelled incumbent firm faces exit costs [equation (11.16)]. Obviously, this case is extremely sensitive to the precise value of sc_0 (i.e., the scale of investment). Low values of sc_0 (that is, small R&D budgets) facilitate the likelihood that this case is valid. However, in general condition (11.21) is unlikely to hold. Consider, therefore, three further scenarios.

EXAMPLE 11.4 (large response lag). Suppose that Baumol *et al.*'s (1982a) argument that

incumbents are restrained by law and other impediments from undertaking retaliatory moves (Baumol *et al.* 1982a, p. 350),

i.e., post-entry price reductions or low-priced re-entry, holds such that expelled incumbent firms suffer from a response lag $1 < R < T$. Then, condition (11.21) transforms into

$$(11.22) \quad [({}^e p_1 - avc_1) \cdot {}^e q_1] + \left[\sum_{t=2}^R ({}^e p_t - avc_t) \cdot {}^e q_t \right] - sc_0 > 0,$$

since, for $t = 2, \dots, R$, the entrant only fears "hit-and-run" entry by other potential entrants, so that ${}^e p_t = {}^* p_t > avc_t$ deters entry for $t = 2, \dots, R$. Again, the expelled firm, unlike the entrant, faces exit costs [equation (11.16)].

From condition (11.22) it follows that high values of R (that is, large response lags of expelled incumbent firms) make it more likely that this case is valid. In effect, this scenario shows close resemblance to Example 11.2 (for ${}^H p_t = {}^* p_t$), although the entrant is unable to return to its home market as, for example, other entrants have invaded its home market (before period $t = T$). Moreover, for $R \rightarrow T$ this case converges to Example 11.1.

EXAMPLE 11.5 (alternative use). Assume, for the sake of simplicity, that $T = 2$. Suppose that an expelled firm is able to benefit from an alternative use of its R&D output. Then, of course, the R&D expenditures (ic_0) are not fully unrecoverable. If sunk exit cost is denoted by sc_0 , then $sc_0 < ic_0$, so far, it has been tacitly assumed that $sc_0 = ic_0$. Suppose that

$$(11.23) \quad \Lambda \pi_T = \Lambda \pi_1 + \Lambda \pi_2 < ic_0,$$

where Λ_{π_T} denotes the (negative) profits from alternative use, composed of profits during period $t = 1$ (Λ_{π_1}) and $t = 2$ (Λ_{π_2}). Then, for $t = 2$ there is an $e p_2 > avc_2$ such that re-entry is deterred. This entry-detering price ($^{ED}p_2$), and associated quantity ($^{ED}q_2$), follows from

$$(11.24) \quad (e p_2 - avc_2) \cdot e q_2 = \Lambda_{\pi_2}.$$

Hence, the potential entrant now decides to enter if the following condition holds:

$$(11.25) \quad [(e p_1 - avc_1) \cdot e q_1] + \Lambda_{\pi_2} - ic_0 > 0.$$

If condition (11.25) holds, the entrant does not face an exit barrier. The expelled firm is only able to enter successfully by offering a $i p_2 < ^{ED}p_2$, which is not worthwhile, since then

$$(11.26) \quad (i p_2 - avc_2) \cdot i q_2 < \Lambda_{\pi_2}.$$

The incumbent's exit costs follow from assumption (11.23):

$$(11.27) \quad i \pi = -i sc_0 = \Lambda_{\pi_T} - ic_0 < 0.$$

Here it follows that the likelihood of a credible fear of "hit-and-stay" entry is facilitated by high benefits from the alternative use of R&D output.

EXAMPLE 11.6 (reciprocal entry). Suppose that an expelled incumbent firm can take the place of the entrant in the latter's home market.¹ If competitive conditions in the potential entrants' home market dictate price $H_{p_t} = {}^*p_t$ for $t = 1, \dots, T$, the entrant can deter re-entry by the expelled incumbent after period $t = 2$ ($R = 1$) by setting the average total cost price from period $t = 2$ to period $t = T$. Hence, entry is profitable if

$$(11.28) \quad [(e p_1 - avc_1) \cdot e q_1] + \left[\sum_{t=2}^T ({}^*p_t - avc_t) \cdot e q_t \right] - sc_0 > 0,$$

which holds for any $H_{p_1} = {}^*p_1 < e p_1 < i p_1$. Hence, if expelled incumbent firms are able to undertake reciprocal entry into the potential entrants' home market against H_{p_t} , only $i p_1 = {}^*p_1$ deters entry when $H_{p_t} = {}^*p_t$ (for $t = 2, \dots, T$). The expelled incumbent firm faces a zero exit cost as capital can be used profitably in another market. This example points out that incumbent firms' low exit cost (as a result of the opportunities of reciprocal entry), increases the credibility of the "hit-and-stay" entry strategy of potential entrants.²

¹ Alternatively, it can be assumed that the expelled incumbent firm is able to enter a third market. Porter (1980, pp. 84-85) calls competition over a third market "cross-parry".

² The argument in Examples 11.5 and 11.6 resembles Eaton and Lipsey's (1980) proposition that "[t]o make an entry-detering threat ... the sitting monopolist must threaten that in the event of entry, he will stay in the market *long enough* that the entrant's present value at the time of entry will be nonpositive. 'Long enough' is Δ^* periods. ... Δ^* can be interpreted as the monopolist's minimum commitment to the market or as the minimum *barrier to his exit*. It is in this sense that barriers to exit are barriers to entry" (Eaton and Lipsey 1980, p. 728; Eaton and Lipsey's italics). Examples 11.5 and 11.6 mirror this result by pointing out that the credibility of the incumbent's re-entry threat decreases if the incumbent's exit barrier diminishes.

11.5 THE CREDIBILITY OF THE ENTRY THREAT

An important critique of contestability theory is that the identity of potential entrants remains in the dark (Chapter 6). This means that the credibility of the entry threat is assumed *a priori*. This flaw is even more serious if sunk investment is permitted. This chapter has provided an insight into assumptions and market conditions which indicate a credible entry threat, notwithstanding the introduction of a positive sunk cost, such that (im)perfect barrier market outcomes can be obtained. That is, this chapter has described examples of workably competitive scenarios.

- (i) It appears to be difficult to set up scenarios which offer a foundation of the case where *new* (or existing) firms which have to sink a cost upon entry constitute a credible source of potential entry. If potential entrants are firms that still need to build up capacity, their speed of response is unlikely to be as fast as hit-and-run requests. An assumption of a positive probability of leapfrogging (that is, heterogeneity of incumbent and potential suppliers) can save imperfect barrier market outcomes (Section 11.2).
- (ii) Firms from a *related* (*prototype*) market face a zero opportunity cost of entry if they can make use of *excess productive capacity*. However, only particular sources of excess capacity are compatible with an assumption of free entry. Moreover, excess productive capacity in combination with a positive sunk cost is likely to invite dumping. Barrier market outcomes can be obtained by imposing specific assumptions (with regard to, for example, parity, reputation, planning horizon and discounting) on the dynamic Prisoners' Dilemma game on dumping (Section 11.3).
- (iii) If potential entrants in a *related* market face *binding capacity restrictions*, the opportunity cost of entry is not necessarily zero. Actual entry can induce home market disequilibria which the potential entrant takes into account while deciding on entry. A careful consideration of assumptions regarding, for example, market conditions (protected or unprotected home market), lag structure (short-lived innovation or long response lag) and the incumbent firms' exit barrier (alternative use of investment or reciprocal entry) shows that barrier market outcomes can be obtained.

The key point is that the source of potential entry cannot be ignored while searching for sets of assumptions which give barrier market outcomes. The barrier market theory is associated with a subtle interplay of the conditions which characterize the environment in which incumbent firms and potential entrants operate.

CHAPTER

12

PROFIT INCENTIVES: Escalating Commitment and Strategic Investment

12.1 PROFIT, SUNK INVESTMENT AND THE OPPORTUNITY COST OF ENTRY

12.1.1 Profit in a Barrier Market

So far, as in contestability theory, the argument is based upon the proposition that perfect barrier markets give a zero profit. This raises the question why incumbent firms are inclined to remain active (and invest sunk cost) in the first place. If the reward to investment is zero, a firm can earn the same profit by being inactive, since zero production and investment yields a zero payoff as well (without any effort or risk)! This argument suggests that incumbent firms are assumed to succumb to the fallacy of escalating commitment. Tang (1988) argues that escalating commitment implies that

individuals tend to choose to commit more resources to a previously invested but failing course of action rather than change the original course of action. The EC [escalating commitment] phenomenon implies not only that a firm does not change its strategy when the environment changes, but that the firm also commits more resources to the strategy even though the strategy is not effective (Tang 1988, p. 79).

To see whether established firms in a barrier market have escalated their commitment to being incumbent, three questions are examined:

- (i) Incumbent firms in a barrier market undertake sunk R&D so as to keep ahead of potential entrants. However, does it pay to be superior (Subsection 12.1.2)?
- (ii) The reward to investment in a perfect barrier market is zero: incumbent firms set average total cost prices. Is it worthwhile to undertake investment which gives a zero payoff (Subsection 12.1.3)?
- (iii) Outside the market incumbent firms can capture a normal return to investment (for example, by devoting money resources to government bonds or putting out money on interest paying banking accounts). Under what conditions is the incumbents' strategy to stay in a barrier market an equilibrium (Subsection 12.1.4)?

It appears that an incumbent firm in a barrier market does *not* escalate its commitment.

The argument is twofold. First, being committed to the market implies that the incumbent's opportunity cost of exit is positive (as exit cost exceeds zero). Second, it is the practice in the literature to define zero profit such that a normal return to investment is included (incorporated in the profit function *via* the discount rate). The argument results in distinguishing two types of barrier markets: perfect and imperfect barrier markets (Subsection 12.1.5). The model used in this section is the elementary two-period framework: $B = 0$, $L = 1$, $A = 1$, $E = 0$, $R = 1$ and $T = 1$ (Section 10.3 and Example 11.1 in Subsection 11.4). For the sake of simplicity of notation, time indices are suppressed. Moreover, demand is localized and entry with excess capacity does not occur.

12.1.2 It Can Pay to Be Superior

Suppose that there are three markets: in markets m_1 and m_2 superior firms operate, whereas m_3 is the home market of inferior suppliers.¹ The purpose of the distinction between superior and inferior firms is twofold. First, it is argued that under particular circumstances price setting of superior incumbent firms is disciplined by a threat of entry by superior potential entrants so that superior firms are prevented from capturing the (full) benefit from the entry barriers that inferior potential entrants are facing. Second, it appears that it can clearly pay to be superior.

The superiority or inferiority of firms is illustrated in terms of production cost or product quality. The sunk investment in period $t = 0$ is directed to either process or product innovation. In the case of process innovations the average total cost level (atc) of superior firms is lower than the one of inferior suppliers: $atc_1 = atc_2 = {}^*atc < atc_3$. Suppose that the process innovation is associated with a cost differential $atc_3 - {}^*atc$ which is the same for all quantity levels. In the case of product innovations inferior entrants face disadvantages in terms of product quality and/or goodwill: that is, they offer a product with a (perceived) unfavorable price - quality ratio. The demand function is Bertrand: consumers prefer to buy the product with the favored price - quality ratio. Inferior potential entrants have to bear a sunk entry cost upon entry so as to provide access to the superior production technology (process innovation) or favored price - quality ratio (product innovation). The average additional sunk cost is denoted by $aic (= sc/q)$, so that $atc = avc + aic$, where q follows from minimizing the average total cost function. Assume that $q_1 = q_2 \geq q_3$.²

Let BE_{p_1} , BE_{p_2} and BE_{p_3} represent the blockaded entry price levels that would result in

¹ Firms can be inferior as a result of, for example, technical inferiority, historical leeway or worse (technical) expectations.

² From Stigler's (1966) definition (Chapter 7) it immediately follows that the cost differential constitutes a barrier to entry. Cost minimization (over quantity) is a necessary condition for sustainability (Baumol *et al.* 1982a). A firm which chooses an output level short of or in excess of the cost minimizing quantity, can be outperformed by a potential entrant which adopts a lower-cost level of production (Chapter 13). The assumption that $q_1 = q_2 \geq q_3$ is plausible. Note that efficient scale increases with sunk cost (Chapter 13). On the one hand, $q_1 = q_2$ follows from the observation that $sc_1 = sc_2$. The familiar contestability assumption implies that superior firms in market m_1 and m_2 have disposal of the same production and investment technology, which they both have used to obtain the same innovation: $sc_1 = sc_2$ (which indicates equal efficient scales in both markets). On the other hand, the assumption that $q_3 \leq q_1 = q_2$ can be grounded by arguing that the (sunk) outlays associated with developing a higher-quality product exceed the expenditures needed to introduce a lower-quality good: $sc_3 < sc_1 = sc_2$ (which gives an efficient scale in market m_3 short of the one in markets m_1 and m_2). The assumption that $q_1 = q_2 \geq q_3$ facilitates the results on profit differentials.

the markets m_1 , m_2 and m_3 respectively if (a threat of) entry and exit are (is) conspicuous by their (its) absence. Moreover, p^E is the price level that would occur in the markets m_1 and m_2 if a reciprocal entry threat rules markets m_1 and m_2 while the threat of entry by inferior potential entrants is not credible. Recall that the reciprocal entry threat implies that firms in market m_1 can enter into market m_2 and *vice versa*, while an expelled firm can only respond (by re-entering the home market or entering the entrants' market) after one period (when the innovation is outdated). Depending on whether inferior potential entrants constitute a credible entry threat, two classes of scenarios can be distinguished.

Firstly, suppose that inferior entrants do not face profitable entry opportunities: ${}^{BE}p_3 + aic \geq p^E$. That is, inferior potential entrants face an effective entry barrier. The relation between ${}^{BE}p_1$ and ${}^{BE}p_2$ indicates the existence of three plausible scenarios of price setting behavior of superior firms. (i) If ${}^{BE}p_1 > {}^{BE}p_2$, firms in market m_2 face profitable entry opportunities if firms in market m_1 set a price $p_1 > {}^{BE}p_2$. Therefore, it is likely that firms in market m_1 opt for the entry-detering price: $p_1 = {}^{BE}p_2 = p^E$. (ii) If ${}^{BE}p_1 = {}^{BE}p_2$, the result is straightforward. Firms in both markets can offer the blockaded entry price: $p_1 = p_2 = {}^{BE}p_1 = {}^{BE}p_2 = p^E$. (iii) If ${}^{BE}p_1 < {}^{BE}p_2$, then the result for ${}^{BE}p_1 > {}^{BE}p_2$ is mirrored: $p_2 = {}^{BE}p_1 = p^E$. The key point is that a firm in market m_i is only prepared to enter market m_j for a price $p_i < p_j$, since p_i and p_j reflect the opportunity cost of entry of firms in markets m_i and m_j respectively (where $i, j = 1, 2$ and $i \neq j$).

The implications of the three scenarios for the profit level depend on whether $p^E \geq {}^*atc$. Negative profits result if in one or both markets the blockaded entry price is associated with ${}^{BE}p < {}^*atc$. If ${}^{BE}p_i = {}^*p = {}^*atc$ and ${}^{BE}p_j \geq {}^*atc$, where $i, j = 1, 2$ and $i \neq j$, then superior firms receive zero profits *ex post*, notwithstanding the sunk investments *ex ante*. The entry barriers, which inferior potential entrants are facing, allow superior firms to capture positive profits as long as ${}^{BE}p_1$ and ${}^{BE}p_2$ both exceed *atc . The key point is that the free entry condition of both markets in which superior firms operate implies that the equilibrium price level converges to the lowest value of ${}^{BE}p$: $p^E = \min \{{}^{BE}p_1, {}^{BE}p_2\}$. This means that perfect (if ${}^{BE}p_1$ and/or ${}^{BE}p_2$ equals *atc) or imperfect (if both ${}^{BE}p_1$ and ${}^{BE}p_2$ exceed *atc) barrier market outcomes are obtained. The result can be extended to the multimarket case: $p^E = \min \{{}^{BE}p_1, \dots, {}^{BE}p_n\}$, where n is the number of markets in which superior firms operate.¹ An important question to address is now whether there were incentives to be engaged in sunk investment behavior *ex ante* if nonpositive profits are expected *ex post*. An answer to this question is provided below. A first indication lies in the examination of the equilibrium in the market in which inferior firms operate.

Events in market m_3 are straightforward. In order to deter entry by superior entrants, inferior firms are forced to adapt their price level downward to at least p^E ; otherwise, superior firms face profitable entry opportunities. On the one hand, in the case of process innovations, profits of the inferior firms fall below those of superior suppliers, since then $atc_3 > {}^*atc$. So, it pays to be superior. On the other hand, in the case of product innovations, inferior firms are forced to offer a price $p_3 < p^E$ such that the lower price offsets the superior firms' quality and/or goodwill advantage: that is, price is adjusted downward to obtain an entry-detering price - quality ratio. The result regarding profit

¹ Of course, if all superior firms manage to collude, then the vector of equilibrium prices $p^E = [{}^{BE}p_1, \dots, {}^{BE}p_n]$ is obtained. However, as soon as one or more superior firms undermine the collusion agreement, the preceding result is likely to apply. Here it is assumed that the superior firms do not (manage to) coordinate their price setting behavior.

price - cost margin is larger than the unfavored product's.¹

Secondly, suppose that inferior potential entrants constitute a credible entry threat: that is, ${}^{BE}p_3 + aic < p^E$. Here it should be noted that aic may be equal to zero. In the case of process innovation the inferior firm may decide to enter while producing with the help of the inferior technology, since ${}^{BE}p_3 < p^E$. In the case of product innovation it may be that there is a ${}^{BE}p_3 < p_3 < p^E$ such that the inferior firms' lower price offsets the superior suppliers' higher quality. In any event, the equilibrium price is likely to deflate to the overall entry-detering price level. Superior firms in the markets m_1 and m_2 adapt their price level downward to $p_1 = p_2 = {}^{BE}p_3 + aic$ in order to deter entry by inferior lower priced firms. The profit levels of inferior and superior firms depend on whether ${}^{BE}p_3 \geq atc_3$ and ${}^{BE}p_3 + aic \geq {}^*atc$ respectively. That is, it is likely that it pays to be superior, since ${}^{BE}p_3 + aic \geq {}^{BE}p_3$ and ${}^*atc \leq atc_3$. Profits may turn out to be negative, zero or positive, depending on p_1 , p_2 , ${}^{BE}p_3$ and aic .

Note that it is of course likely that none of the firms will decide to invest *ex ante* if it is foreseen that profits *ex post* will be negative. In effect, if the price falls below the level of average variable and fixed cost, then the firms' shut-down-point is passed so that it pays (in terms of opportunity cost) to exit. This result has of course little to do with the market form that is (assumed to be) prevailing. If neither incumbent firms nor potential entrants decide to invest *ex ante*, the market is perfectly contestable *ex post* (Subsection 12.1.3): in absence of positive sunk cost, entry is free and exit is costless.

12.1.3 Zero Profit Investment

The prospect of positive profits can induce sunk investment behavior. The case where the inferior potential entrants' entry threat is credible invokes limit pricing ($p_1 = p_2 = {}^{BE}p_3 + aic$) on the part of the superior firms. The situation in which superior incumbent firms' price setting in market m_i is disciplined by a threat of entry by superior potential entrants in m_j ($i, j = 1, 2$ and $i \neq j$) is called an *imperfect barrier market* if the associated profits exceed zero (${}^*atc < p_i = {}^{BE}p_j < {}^{BE}p_3 + aic$). The threat of entry by superior potential entrants prevents incumbent firms from capturing the full benefit from the entry barriers that inferior potential entrants are facing. A *perfect barrier market* indicates the equilibrium in which a zero profit is received ($p_i = {}^{BE}p_j = {}^*atc < {}^{BE}p_3 + aic$). The question whether sunk investment is worthwhile undertaking in a perfect barrier market deserves further attention.

Suppose that incumbent firms face superior potential entrants which will capture a zero profit if they refrain from entry (${}^{BE}p_j = {}^*p = {}^*atc$). Assume further that there are two investment options: a firm invests either zero ($sc = 0$) or at an (expected) optimal scale (${}^*sc > 0$). Take a representative firm in market m_i and entrant in m_j which both face a nonbinding productive capacity. Both face the decision whether to invest *ex ante* in view of the payoff matrix that indicates the following *ex post* profits (calculated over both markets; that is, assuming a reciprocal entry threat).

If both firms fail to invest *ex ante*, then the markets are perfectly contestable *ex post*,

¹ If the opposite holds, the investment game in period $t = 0$ yields a different result. Like firms in market m_3 , suppliers in market m_1 and m_2 can decide to invest in the unfavored product, which gives higher profitability. Investment in the favored product is a disequilibrium strategy, since offering the favored product gives a lower payoff. Therefore, the superiority - inferiority distinction remains under cover. All firms offer equal products with the same price-quality ratio, so that the period $t = 1$ game resembles the one described in the preceding paragraphs.

If both firms fail to invest *ex ante*, then the markets are perfectly contestable *ex post*, which yields a zero profit to either rival. If one firm invests while the other does not, then the innovating firm is able to push aside the noninnovating rival. The expelled firm can face an exit cost (E), whereas the entrant captures a positive profit (M). If both firms innovate *ex ante*, they face a perfect barrier market [or, using Maks' (1986) terminology, equi-barrier market] *ex post* so that zero profits are earned.

Table 12.1
Payoff Matrix Perfect Barrier Market with Exit Cost

		Firm j	
		$sc = 0$	$^*sc = 0$
Firm i	$sc = 0$	(0,0)	(-E,M)
	$^*sc = 0$	(M,-E)	(0,0)

Zero sunk cost gives a zero exit cost. If $E = 0$, both perfect contestability ($^i sc = ^j sc = 0$) and a perfect barrier market ($^i sc = ^j sc = ^* sc$) are an equilibrium outcome in pure strategies. Moreover, both monopoly outcomes are sustainable ($^i sc = 0$ and $^j sc = ^* sc$, where $i, j = 1, 2$ and $i \neq j$), since neither the expelled firm nor the monopolist has an incentive to change strategy. Two observations are noteworthy. First, the contestable and barrier market equilibria are weak and imply coordination difficulties. Second, the combination of positive sunk cost and zero exit cost implies that exit occurs after a period sufficiently long to permit the recovery of the up-front sunk cost: this is a peculiar case.

The perfect barrier market can be a *unique* Nash equilibrium if one of two assumptions is introduced. First, if the game is played in mixed strategies, even the smallest probability of being a monopolist (f , which is the likelihood that the rival firm decides not to invest) induces positive investment: $(1-f) \cdot E + f \cdot M = f \cdot M > 0$ for $E = 0$ and any $f > 0$. Second, if $E > 0$, the perfect barrier market ($sc_i = sc_j = ^* sc > 0$ and zero profit to both firms) constitutes a unique Nash equilibrium which follows from dominating strategies. That is, both firms are inclined to invest *ex ante*, notwithstanding the zero profit *ex post*. The reason is simple: the investment strategy yields at least as much payoff as when investment is abandoned, while negative profits (*i.e.*, exit) are avoided. The barrier market equilibrium is strong, since the associated strategies are best replies *no matter* what the rival is doing. Moreover, this equilibrium is also the maximin solution.

Three remarks are worth making. First, it should be stressed that neither firm is prepared to push aside the rival for a price below the average total cost level ($^* atc - aic$ in the limit; that is, $^* atc - aic$ indicates the shut-down-price). The reason is straightforward. Any $p < ^* atc$ yields a loss, whereas non-entry is associated with a zero profit in the entry

market. That is, the opportunity cost of entry is positive. Second, the assumption that incumbent firms face an exit cost if they are expelled from the market by an entrant, boils down to arguing that incumbent firms, *ceteris paribus*, prefer staying in the market to exiting (provided that nonnegative profits are captured). A similar assumption underlies perfect contestability (Chapter 6). Third, Section 12.1.2 argues that nonoptimal investments are not worth considering (the associated game shows a payoff matrix which resembles the one above).

12.1.4 Zero or Normal Profit

Daily business practice indicates that the net returns to investments (zero in a perfect barrier market) must compare favorably to *alternative* investment opportunities. Suppose, therefore, that there is an alternative investment opportunity ic which yields payoff W *ex post*. Then, payoff matrix 12.2 appears.

Table 12.2
Payoff Matrix Perfect Barrier Market with Investment Alternative

		Firm j	
		$ic = 0$	$^{*}sc > 0$
Firm i	$ic = 0$	$(-W, -W)$	$(-E + W, M)$
	$^{*}sc > 0$	$(M, -E + W)$	$(0, 0)$

The alternative investment ic only constitutes a unique Nash equilibrium if $W > E$ and $W > M$. If both conditions are violated, the unique Nash equilibrium is associated with a perfect barrier market *ex post*. If only one of the conditions fails to hold, then further assumptions are needed in order to reach a unique solution.

First, take the case where $W > M$ but $W < E$. Then, the unique Nash equilibrium is still associated with a perfect barrier market *ex post* if, for example, the firms employ either a maxmin decision rule (which gives an equilibrium in dominating strategies) or mixed strategies in combination with an appropriate "trembling hand" (Selten 1978) condition. An appropriate "trembling hand" condition amounts to assuming that (the firms expect that) the probability that a rival opts for innovative investment behavior exceeds (to a sufficient extent) the likelihood of a rival deciding to carry out the alternative investment. That is to say, any f such that $W - f \cdot E < (1 - f) \cdot M$ will do, where f is the probability that the rival opts for $^{*}sc > 0$.

Second, suppose that $W < M$ and $W > E$. This is a plausible case, since perfect contestability indicates that E is close to zero. It is now possible to reach a unique Nash equilibrium which is associated with a barrier market *ex post* that yields a normal return to sunk investment. This can be done by introducing the assumption that *normal* profits (N) are earned in a barrier market such that $N > -E + W$. The following payoff matrix then occurs.

Table 12.3
Payoff Matrix Perfect Barrier Market with Normal Profits

		Firm j	
		ic > 0	*sc > 0
Firm i	ic > 0	(W, W)	(-E + W, M)
	*sc > 0	(M, -E + W)	(N, N)

It is easily seen that $sc_i = sc_j = {}^*sc$ constitutes a unique Nash equilibrium. Moreover, imposing an appropriate "trembling hand" condition gives innovative investment behavior as the unique Nash equilibrium strategy even in the case where $W > E$ and $W > M$. That is, any f such that $W - f \cdot E < (1 - f) \cdot M + f \cdot N$ implies that a barrier market *ex post* associated with normal profits constitutes a unique Nash equilibrium.

A remark is in order with respect to normal profits. Suppose that there is a credible alternative investment which yields a rate of return r . The normal rate of return to investment is then simply introduced into the barrier market theory through the discount factor. Assume, in terms of equation (10.1), that the alternative investment yields r in each period in the interval $1 - T$. Recall that after $t = T$ the innovative investment is outdated. That is, the sunk costs incurred in $t = 0$ must be recovered in the period from $t = 1$ to $t = T$. In light of the normal returns requirement, the necessary condition for entry deterrence (10.1) is associated with $\delta^t = \exp(-rt)$.¹

Three further remarks are also in order. First, alternatively, condition (10.1) can be adjusted (in order to capture the normal profit requirement) by replacing the right hand side by $sc_0 - E + W$. It can easily be seen that the normal profit requirement is only effective if $W > E$. Second, the level of normal profits follows from the assumptions underlying the barrier market theory. Normal profits are determined by $-E + W$ (where, it will be recalled, E is the once and for all exit cost and W the net present value of the flow of discounted payoffs to an alternative investment). That is, normal profits are the sum of exit costs and the profits associated with the "best" alternative investment opportunity (*i.e.*, outside the market).

Third, it is important, however, to note that the normal profit requirement has *little* to do with the market form that is (assumed to) prevail(ing) (compare with, for instance, Pyatt 1971). That is to say, if incumbent firms in, for instance, a contestable market (*sc*

¹ Ignoring, for the moment, exit costs. Otherwise, r must be adjusted downward in order to take E into account: that is, the alternative investment gives $-E + W$ rather than W .

= 0) face an alternative opportunity to invest (money) resources that offers a normal rate of return, then they too are inclined to introduce a normal profit requirement (which does not attract entry). The reason is simple: the normal rate of return r indicates the opportunity payoff that can be captured outside the market. Hence, firms are only inclined to stay in or enter the market if at least a normal rate of return is earned. Von Weizsäcker (1980a), for instance, argues that

[t]he actual entrant ... can expect a return on his investment, which corresponds to the opportunity cost of capital, i.e., the normal rate of return (Von Weizsäcker 1980a, p. 407).

In the same spirit, Baumol *et al.* (for example, Panzar and Willig 1977, Baumol and Willig 1981 and Baumol *et al.* 1982a) indicate time and again that incumbent firms in a contestable market yield "normal economic rents". In preceding and subsequent chapters the normal profit condition enters in the models through the discount rate. As is usual in economic theory, it can be assumed that the discount rate is equal to the prevailing market interest rate. Moreover, the zero ($r = 0$) and normal ($r > 0$) profit are used as synonyms.

12.1.5 Perfect and Imperfect Barrier Market

By way of summary it is argued that there are cases where the assumption of a perfect (sunk investment *ex ante* and zero or normal profit *ex post*) or imperfect (sunk investment *ex ante* and positive or above-normal profit below the blockaded entry level *ex post*) barrier market is valid. If one starts from the position that sunk investment only occurs where there is the prospect of above-normal profits, then an imperfect barrier market framework is plausible for any arbitrarily small value of β ($\beta > 0$) such that $p_t = {}^{BE}p_t = {}^{atc} + \beta$ (provided that a superior investment alternative is not available).

By way of illustration the conditions that underlie a *perfect* barrier market are stated explicitly for the case where the entry threat ensues from existing potential entrants which either are inclined to refrain from dumping strategies (Section 11.3) or face binding capacity restrictions in a home market (Section 11.4). In what follows, for reasons of brevity, the adjectives "superior" and "perfect" are suppressed: the key assumption is that incumbent firms and potential entrants are at par (in terms of technology and sunk exit cost). Moreover, to simplify matters, the discount factor is ignored (that is, $r = 0$). The credible fear of "hit-and-run" entry *ex post* can only be sustained if the innovating potential entrants are able to recover *precisely* the investment costs elsewhere. The reason is straightforward. Suppose that potential entrants face the opportunity to capture a (gross) profit (π)

$$(12.1) \quad H_{\pi} = \sum_{t=1}^T ({}^e p_t - avc_t) \cdot {}^e q_t$$

in a home market. Hence, (gross) profits H_{π} represent the *opportunity costs* of entry.

Assume, first, that *all* potential entrants capture a $H_{\pi} > sc_0$. This implies that potential entrants will only decide to enter when post-entry (gross) profits ${}^e \pi$ exceed H_{π} . Hence, incumbent firms are able to set a entry-detering price (${}^{ED}p_t$) without provoking entry, where ${}^{ED}p_t$ and associated ${}^{ED}q_t$ follow from

$$(12.2) \quad \sum_{t=1}^T (\text{ED} p_t - \text{avc}_t) \cdot \text{ED} q_t = H_{\pi},$$

provided that firms face sunk cost parity. Incumbent firms then capture *positive* (net) profits, since $H_{\pi} > sc_0$ (recall the assumption that $sc_0 = {}^e sc_0$). Suppose, second, that there are *some* potential entrants facing $H_{\pi} < sc_0$. An analogous argument indicates that the incumbent firms' limit price is now associated with *negative* (net) profits, which undermines the very incentive to innovate *ex ante*. So it follows that the usual (tacit) assumption that potential entrants earn a zero net profit outside the market (Chapter 6, Assumption 6.6) is very important. However, this assumption only reflects a *special case*. In effect, it is the *opportunity cost of entry* that matters.

12.2 NONSTRATEGIC AND STRATEGIC ENTRY-DETERRING INVESTMENT

A distinction well-known in the literature on industrial organization is that by Salop (1979) between nonstrategic (innocent) and strategic entry-detering investment (Chapter 7; Section 7.4). Recall that

[s]ince some expenditures ... will be made in the absence of entry, it is conceptually worth separating out that proportion of such expenditures as would be incurred by profit maximising firms in the absence from entry from that part which is strategically chosen in order to obstruct entrants (Encaoua *et al.* 1986, p. 57).

The study of (the implications of) deterrence instruments enjoys considerable popularity in the literature. In particular, excess capacity, advertizing and pre-emptive innovation are worth noting (Chapter 7 and 8). The models of strategic entry-detering investment are similar to the literature on limit pricing in the sense that strategic entry-detering investments are carried out in order to raise entry barriers so that incumbent firms can capture positive (above-normal) profits.

The (expected) optimal innovations that incumbent firms in a barrier market are inclined to introduce are carried out in order to reduce the probability of entry. An incumbent firm which fails to undertake (expected) optimal R&D investments invites profitable entry. The same holds in a contestable market (Chapter 6). If such efficient innovations are also carried out by firms which do *not* face an effective entry threat (blockaded entry), the market is called a *nonstrategic* barrier market.¹ A barrier market in which *strategic* entry-detering investment occurs is called a *strategic* barrier market. In a strategic barrier market incumbent firms (are inclined to) carry out investments which are only undertaken because they deter entry.²

The following scenario illustrates a strategic barrier market. Take the three-market model of Section 12.1.2. Suppose that the superior firms in the markets m_1 and m_2 captured a blockaded entry price ${}^{BE}p_1$ and ${}^{BE}p_2$ respectively. Take, for instance, the case where ${}^{BE}p_1 > {}^{BE}p_2 = p^E$. Furthermore, assume that inferior firms in market m_3 face a blockaded entry price ${}^{BE}p_3 < p^E$. To simplify matters, let the inferiority of the firms

¹ Although, probably, there may be a difference in the date at which innovations are introduced (Chapter 8).

² Recall that strategic entry-detering investments are profit maximizing in the face of a credible entry threat. They would not be profit maximizing if the entry threat was ineffective. However, in a strategic barrier market the entry threat is credible.

which operate in market m_2 be the result of an average total cost differential ($atc_1 = atc_2 = {}^*atc < atc_3$). Since ${}^{BE}p_3 < p^E$, inferior firms face profitable entry opportunities. Therefore, *ceteris paribus*, superior firms are inclined to adjust their price downward to the entry-detering level: $p_1 = p_2 = {}^{BE}p_3$. In this way, the benefit which firms in the markets m_1 and m_3 yield from their superiority is restricted.

A strategic barrier market can occur if superior firms can carry out an investment (sb) such that $({}^{BE}p_3 - {}^*atc) \cdot {}^*q < [(p^E - {}^*atc) \cdot {}^*q - sb]$, where *q is the cost minimizing output level of the firms facing the superior production technology. The strategic investment sb deters entry if it implies that inferior firms have to bear an *additional* average (sunk) expenditure upon entry (aic) such that ${}^{BE}p_3 + aic \geq p^E$. That is, entry is no longer worthwhile. Under these circumstances it pays *all* superior firms to undertake strategic entry-detering investment in a strategic barrier market. Strategic investments may be enhanced further if superior firms manage (and are allowed) to collude.

To clarify matters, two examples of strategic entry-detering investments are briefly described. First, superior firms are engaged in strategic advertizing if inferior firms are forced to bear even higher advertizing budgets so as to make entry successful. This may be the case when the inferiority of the firms in market m_3 is due to a lower product quality so that advertizing confronts inferior firms with a goodwill disadvantage. Second, superior firms may introduce excessive product innovation. That is, the R&D outlays exceed the efficient blockaded entry level. Such excessive innovations are strategic, however, if inferior firms are only able to enter successfully after bearing an even higher R&D budget. The higher R&D budget of firms in market m_3 may follow from their very (innovative) inferiority.

Of course, whether a strategic barrier market is perfect or imperfect depends on whether profits $[(p^E - {}^*atc) \cdot {}^*q - sb]$ are equal to or exceed zero (or a normal level). In any event, it is clear that the above scenario illustrates that in a strategic barrier market it pays to introduce strategic investments in order to reduce an *unlimited* threat of entry (by superior *as well as* inferior potential entrants) to a *limited* one (by superior potential entrants only).¹

12.3 TYPES OF BARRIER MARKETS

In all, there are four types of barrier markets to be distinguished:

- (i) A *perfect nonstrategic* barrier market in which innocent investments are introduced, although only zero (or normal) profits are earned;
- (ii) A *imperfect nonstrategic* barrier market in which innocent investments are carried out that are associated with positive (or above-normal) profits;
- (iii) A *perfect strategic* barrier market in which entry-detering investments occur, although they yield only zero (or normal) profits; and
- (iv) An *imperfect strategic* barrier market in which entry-detering investments are undertaken that result in positive (or above-normal) profits.

¹ Special reference can be made to Eaton and Lipsey (1980). They offer another example of the introduction of nonstrategic and strategic market structures: "We define *type-T natural monopoly* ... to be one in which *cost minimizing* decisions with respect to durability, replacement, and maintenance of capital imply that there is no point in time at which entry is profitable, and a *type-T artificial monopoly* to be one in which *strategic decisions* with respect to capital prevent there being any point in time at which entry is profitable" (Eaton and Lipsey 1980, p. 722; Eaton and Lipsey's italics).

The common denominator of all types of barrier markets is that superior incumbent firms are unable to benefit (fully) from the entry barriers which inferior firms are facing, since their behavior is disciplined by a threat of entry by superior potential entrants.

The key argument in this chapter is that barrier market outcomes can only be obtained under specific conditions, which strongly depend on the source of potential entry that is assumed. By playing around with scenarios and assumptions an insight has been provided regarding the sensitivity of the results for conditions of, for example, opportunity cost of entry, productive capacity restrictions and lag structure. Moreover, games are explored which give barrier market outcomes without incumbent firms being engaged in escalating commitment. Keeping these complicated arguments in mind, Parts VI, VII, VIII, IX and X go on to analyze the sustainability of a barrier market (Part VI), incumbent suppliers' decision models (Parts VII and VIII) and the (dynamic) interaction between incumbent firms and potential entrants (Parts IX and X) in the illustrative two-period + two-market model (Section 10.3).

A final remark on the role of sunk and exit cost can further clarify matters. Up to this point investment is assumed to be fully unrecoverable outside the market. Exit cost is determined only by unamortized investment. A general argument can, however, be based upon distinguishing four cost components:

- (i) *up-front investment cost* (uic) follows from the investment (R&D) outlays which precede the effectuation of the investment (introduction of innovation);
- (ii) *pure exit cost* (pec) is determined by the expenditures which the firm has to incur so as to effectuate exit (for example, the cost of firing employees and dismantling productive capacity);
- (iii) *actual sunk cost* (asc) at the investment date is the unrecoverable part of uic, whereas pec is sunk at the date of exit (of course, pec can be discounted *ex ante*); and
- (iv) *actual exit cost* (aec) consists of the unrecoverable part of uic and pec which have not yet been amortized in the market. If ${}^G\pi_E$ is the firm's gross profit at the exit date $t = E$, being insufficient to recover both uic and pec, then $aec = {}^G\pi_E - uic - pec$.

The key point is that contestability assumes pure and actual exit cost always to be zero, while both can be positive in a barrier market. Distinguishing strict from broad contestability can clarify matters.

Strict contestability assumes the four cost components to be zero ($uic = pec = asc = aec = 0$), whereas *broad contestability* can be defined as a contestable market in which the up-front investment cost is always fully recoupable outside the market and pure exit cost is zero ($uic > 0$ and $pec = asc = aec = 0$). The discussion in the literature (Chapter 6) focuses on strict contestability, which implies that both up-front investment cost and pure exit cost are assumed to be zero. However, strictly speaking, Baumol *et al.* (Chapter 13 in their 1982 book) also capture broad contestability.

The *barrier market theory* is an alternative to both variants of contestability, as this framework is associated with the assumption that, first, the up-front investment cost is not (fully) recoupable outside the market(s) and, second, the incumbent firms' pure exit cost is allowed to be positive ($uic, asc, aec > 0$ and $pec \geq 0$). Sunk capital can, however, be used in both the incumbent firms' and potential entrants' markets. However, outside these restricted set of markets the sunk capital does not give value for money. To focus attention, subsequent chapters, unless otherwise indicated, deal with a perfect barrier

market in which pure exit cost is zero and the up-front investment cost is fully unrecoverable outside two (the incumbent firms' and potential entrants') markets.

The new element in the argument in this section has been pure exit cost. A positive pec can influence the incumbent firms' and potential entrants' strategy in a barrier market. On the one hand, an incumbent firm's motivation to adopt careful pricing and entry-detering R&D is fortified so as to avoid the increased cost of exit. This extra motivation does not matter in a perfect barrier market, as incumbent firms are already engaged in maximum (within feasibility bounds) entry deterrence. The case is different in an imperfect barrier market. A positive probability of future exit (f^1 , f^2 in the model in Subsection 11.2.3) increases the cost of above average cost pricing, as the latter is positively correlated with the entry probability. On the other hand, as the literature indicates (Subsection 11.2.2), a potential entrant's pure exit cost can operate as a barrier to entry. If a potential entrant anticipates a positive probability of future exit, pec adds to the cost which has to be amortized after entry. So, a positive pure exit cost of incumbent firms induces further convergence to perfect barrier market outcomes, whereas a potential entrants' positive pure exit costs procures the opposite. In the latter case the pure exit cost gives an exit and entry barrier.

BARRIER MARKET SUSTAINABILITY**Unique Cost-Minimizing Production
and Amortization Procedures**

KEY WORDS. An important concept of contestability theory is sustainability. The barrier market theory imposes two new elements on sustainability analysis, because sunk cost is permitted to be positive. First, any positive sunk cost increases the probability that the average total cost function is associated with a unique cost-minimizing output level (Chapter 13). This fortifies the difficulty of (particularly temporary) sustainability. Second, if sunk cost has to be recovered over a number of periods ($T > 1$), the amortization procedure plays a critical role (Chapter 14). Even if incumbent firms and potential entrants face the same scale of sunk cost, the use of different amortization rules can induce (particularly intertemporal) unsustainability.

1944

1945

1946

1947

1948

1949

1950

1951

CHAPTER

13

SUNK COST, UNIQUE COST-MINIMIZING PRODUCTION AND SUSTAINABILITY

13.1 INTRODUCTION: Temporary and Intertemporal Unsustainability

Existing contestable market models deal with the case where firms' investment behavior cannot occur if it is associated with sunk exit cost (Schwartz and Reynolds 1983, Shepherd 1984 and Stiglitz 1987). However, in a dynamic framework investment behavior is of utmost importance. In the context of a dominant *potential* competition it is *pre-emptive* investment that matters (Gilbert and Newberry 1982, Fudenberg *et al.* 1983 and Gilbert 1986; see also Chapter 8). In order to deter entry incumbent firms need to introduce price reducing process innovations and favored product improvements. As stipulated time and again, innovative activity is associated with investment of *sunk* R&D expenditures (Kamien and Schwartz 1982, Shepherd 1984, note 19, p. 580 and Stiglitz 1987, p. 928). Dynamic nonprice competition introduces sunk exit costs. However, not only is the introduction of sunk exit costs in conflict with existing models of perfect contestability (Chapter 6), but it is also likely that it will harm the sustainability of market configurations (Baumol *et al.* 1982a).

This chapter examines the difficulties that sunk costs pose to the (particularly temporary) sustainability of industry configurations (Definitions 6.2 to 6.4)¹. To be precise, the focus is on the integer problem. So, unless indicated otherwise, this chapter deals with temporary sustainability. For the sake of the argument, Section 10.3's two-period benchmark model is used. The sunk cost incurred in period $t = 0$ can introduce temporary (un)sustainability in period $t = 1$. The chapter is organized as follows. Section 13.2 presents the core argument by showing the impact of sunk costs on the likelihood of temporary (and intertemporal) sustainability. Section 13.3 argues that imperfection of the barrier market and entry barriers facilitate the sustainability of a market configuration. Section 13.4 discusses two further complications: uncertainty and adjustment. Section 13.5 offers a conclusion.

13.2 SUNK COSTS, AVERAGE COST FUNCTIONS AND SUSTAINABILITY

This section examines the implications of the introduction of sunk costs in the illustrative

¹ So, this chapter present an argument along the lines of Baumol *et al.*'s (1982a) sustainability framework. The results with regard to perfectly contestable markets are summarized in Chapter 6 (Subsection 6.4.4).

framework with a one-period investment lag.¹ Suppose that firms undertake R&D programs in period $t = 0$ so as to reach an entry-deterring innovation in period $t = 1$. The sunk R&D expenditures (sc) in period $t = 0$ must be rewarded in period $t = 1$ with the help of a markup (μ) on average variable and fixed cost (avc). In period $t = 1$ the sunk costs introduce by definition increasing returns in production:

$$(13.1) \quad \mu_1 = M(q_1, sc_0), \text{ with}$$

$$(13.1a) \quad \partial M / \partial q_1 < 0, \text{ and}$$

$$(13.1b) \quad \partial M / \partial sc_0 > 0,$$

where M is a continuous, twice differentiable function with $M(q_1, 0) = 0$.

The pervasive fear of "hit-and-run" entry still forces incumbent firms to follow the zero-profit condition so as to avoid exit (Proposition 14.2). Hence, in the illustrative example with a one-period investment lag equation (13.1) boils down to

$$(13.2) \quad \mu_1 = sc_0 / q_1,$$

where, for the sake of simplicity of notation, discounting is ignored. In order to deter entry (and so maximize profit) incumbent firms seek to minimize the price offered in period $t = 1$.² An incumbent firm deploys minimum average total cost (atc) pricing (Chapter 15). So, the price-output decision then follows from

$$(13.3) \quad \min_{q_1} p_1 = C(q_1, \cdot) = A(q_1, \cdot) + M(q_1, \cdot)$$

for $i = 1, \dots, n_{i0}$, where C denotes the average total (nonsunk plus sunk) cost function and A the average variable and fixed (nonsunk only) cost function. A is a twice continuously differentiable function with $A(0, \cdot) = 0$ and $\partial^2 A / \partial q^2 > 0$ (suppressing indices). In what follows, for the sake of simplicity of notation, the incumbent's index is suppressed. The special case of the flat-bottomed U-shape of the average nonsunk cost function (Baumol *et al.* 1982a; see also Chapter 6, Subsection 6.4.4) means that

$$(13.4) \quad avc_1 = A(q_1, \cdot), \text{ where}$$

$$(13.4a) \quad \partial A / \partial q_1 < 0 \text{ for } q_1 < L_{q_1},$$

$$(13.4b) \quad \partial A / \partial q_1 = 0 \text{ for } L_{q_1} \leq q_1 \leq R_{q_1}, \text{ and}$$

$$(13.4c) \quad \partial A / \partial q_1 > 0 \text{ for } q_1 > R_{q_1}.$$

¹ Part VI points out that the effect of sunk cost is twofold. First, a positive sunk cost introduces increasing returns in the cost function. This is the subject matter of the present chapter. Second, in a multiperiod setting the time path of markups matters. Then, firms may be forced to follow some conventional zero-profit amortization rule. This implication is outlined in Chapter 14.

² The minimization of price as a necessary condition for sustainability is a general result of the free entry assumption, whether (Chapter 15) or not (Baumol *et al.* 1982a) sunk exit costs are present. Chapter 15 focuses on minimization over sc_0 , whereas this chapter examines minimization over q_1 .

A is a piecewise twice continuously differentiable cost function with $0 < L_{q1} \leq R_{q1}$.

In a contestable market $sc_0 = 0$ so that $C(q_1, \cdot) = A(q_1, \cdot)$ (or $atc_1 = avc_1$). If sunk costs are introduced this equality no longer holds, since then $atc_1 = A(q_1, \cdot) + M(q_1, \cdot)$. It is easily shown that minimization of the price gives a *unique* cost (and price, provided that firms adopt average total cost pricing) minimizing $^*p - ^*q$ combination for a given R&D input sc_0 .

PROPOSITION 13.1 (sunk cost and unique minimum average total cost). *Sunk costs introduce, ceteris paribus, a unique minimum in the average total cost function.*

The proof of Proposition 13.1 is offered in Appendix A. Proposition 13.1 implies that the unique minimum must *precisely* match individual demand at the minimum price in order to preserve sustainability of a market configuration from period $t = 0$ to $t = 1$: $^*q = ^*d$, where $^*d = D(^*p)/ni_0$. Or, equivalently, $D(^*p)/ni_0 = ^*q$. Otherwise, entry (if $^*q < ^*d$) or exit (if $^*q > ^*d$) occurs. Moreover, the uniqueness of the minimum of the average total cost function reduces the likelihood that temporary sustainability is reached. Temporary sustainability in period $t = 1$ requires that $D(^*p)/q^*$ gives an integer ni_1 . Hence, the introduction of sunk costs in a free entry framework seriously reduces the probability of sustainability of an industry configuration for $ni_0 > 1$.¹

Figure 13.1 (suppressing time indices) suggests an important implication (recall that $^*p = mac$).

Figure 13.1
Average Total Cost Functions

¹ Knieps and Vogelsang (1982) and Kim (1987) introduce sunk costs in a natural monopoly model. Then, sunk costs may even enhance the sustainability of the monopoly configuration.

If $sc_0 = 0$ (contestability), the minimum optimal scale is Lq . If $sc_0 > 0$ (barrier market), the minimum optimal scale (*q_1) is larger than Rq , and therefore $^*q_1 \gg ^Lq$. Hence, a small structural change (i.e., the appearance of sunk costs) changes the size of minimum optimal scale and, therefore, market structure.

The reduced likelihood of (temporary) sustainability also appears in Novshek and Sonnenschein's (1987) analysis of a general equilibrium model with free entry. They bypass the integer problem in two ways. First, they argue that there may be a slightly higher price $^*p + \alpha$

at which supply equals demand and excluded firms have *little* incentive to enter (Novshek and Sonnenschein 1987, p. 1286; italics added).

Hence, by introducing "little" rather than "no" incentive to enter they indicate that small variations in the price are allowed without attracting entry. In Section 13.3 a rationale for this assumption is presented. Second, Novshek and Sonnenschein focus on the limit case of perfect competition. That is to say, if

the efficient scale of each firm is an infinitesimal quantity and [if] there is an unbounded mass of available firms (Novshek and Sonnenschein 1987, p. 1287),

then

an appropriate mass of active firms achieves an exact equilibrium at price [*p] at which each produces at efficient scale (Novshek and Sonnenschein 1987, p. 1287).

This result follows from the observation that in the limit case of perfect competition the aggregate supply schedule is horizontal.

A last remark is in order with respect to the case where $ni_0 > ni_1$. Then, a firm which is predetermined to exit is willing to sacrifice the markup in order to stay in the market so that at least part of the investment cost can be recovered. In this case pricing below the average total cost markup is a credible threat. However, in line with the contestable market theory, the barrier market approach is based upon a Bertrand-Nash assumption. That is to say, potential entrants do *not* fear (immediate) retaliation by incumbent firms (Assumption 6.3). The prices announced remain fixed for a period of time sufficiently long to make the entry threat credible (Chapter 11). In the same fashion, the sacrifice of the markup, so as to avoid exit if $ni_0 > ni_1$ only matters if announced in advance. Then, it may be that the downward pressure on the market price establishes an equilibrium with $ni_0 = ni_1$. Two observations are noteworthy. First, the state with $ni_0 > ni_1$ gives, of course, coordination difficulties, since the underlying game is associated with multiple equilibria: some (but which?) firms invest, whereas some (but which?) rivals leave the market. Second, the associated negative profit implies a disincentive to invest in the first place (Chapter 12). Here it must be noted that the observation that $^*q_1 > ^Rq_0$ does not imply that $ni_0 < ni_1$ is the most likely case. The reason is straightforward: what matters is not Rq_0 , but rather *q_0 .

13.3 IMPERFECT BARRIER MARKETS, ENTRY BARRIERS AND SUSTAINABILITY

The introduction of sunk costs in a free entry framework implies that the assumption of a

flat-bottomed U-shaped average variable and fixed cost curve is not useful, since it fails to facilitate the likelihood of sustainability of a market configuration. An obvious way to promote the sustainability of industry configurations, if sunk costs emerge, is to re-introduce a flat bottom in the average total cost function. Then, it is necessary to postulate a (U-shaped) average nonsunk cost function such that the increasing returns of sunk costs are precisely outbalanced by the decreasing returns of average variable and fixed cost over a (small) interval of output levels, so that the average total cost function shows a flat bottom. This is illustrated in Figure 13.2.

Figure 13.2
Flat-Bottomed Average Total Cost Curve

Whether or not the average total cost function shows an (almost) flat bottom depends critically upon the numerical specification of the average variable and fixed cost function and the actual scale of sunk costs. Since both are subject to firms' decision making, the flat bottom may appear and disappear over time. The length of the (flat) bottom ($R_q - L_q \geq 0$) is partly endogenized through the firms' decisions on R&D outlays.

Besides, the sustainability of a market configuration is facilitated by permitting prices to fluctuate within a (small) range *without* provoking entry. If it is assumed that potential competitors will only enter for $p - p \geq \alpha$, then the likelihood of successful entry deterrence depends on α . A possible rationale underlying this assumption is, of course, that α reflects an (opportunity) cost of entry. On the one hand, this means that an imperfect barrier market is more robust (in terms of sustainability) than its perfect mirror image (Chapter 12). On the other hand, entry barriers facilitate the probability of sustainability of an industry configuration (Baumol *et al.* 1982a). However, (minimum) average cost pricing may still be evoked if uncertainty is introduced. Suppose that firms are unaware of the existence of α . The firms just act as if $\alpha = 0$, which implies a pervasive fear of entry and thus (minimum) average cost pricing. Alternatively, it may be that firms are uncertain of α 's precise value (Maks 1986, p. 119, note 20). Firms then realize that the likelihood of entry (f) depends on the extent of above average total cost

pricing ($\bar{p} = p - \alpha^*$) and the expected value of α (α^*). So,

$$(13.5) \quad f = F(\bar{p}, \alpha^*), \text{ where}$$

$$(13.5a) \quad \partial F / \partial \alpha^* < 0,$$

$$(13.5b) \quad f = 0 \text{ for } \bar{p} \leq 0, \text{ and}$$

$$(13.5c) \quad \partial F / \partial \bar{p} > 0 \text{ for } \bar{p} > 0.$$

The following profit function (π) now emerges (provided that $d = q$, as assumed in the model in Section 10.3):

$$(13.6) \quad \pi = (1 - f) \cdot [p - A(q)] \cdot q - sc.$$

If $\alpha = 0$ (as assumed in the preceding sections), then $f = 0$ for $p \leq \alpha^*$ ($\bar{p} \leq 0$) and $f = 1$ for $p > \alpha^*$ ($\bar{p} > 0$).¹

It appears that the occurrence of (minimum) average cost pricing is facilitated by the following phenomena:

- (i) incumbent firms expect α to be very small;
- (ii) incumbent firms are risk averse, which is reflected in an upgrading of f (to $f = 1$ for $p > \alpha^*$ in the limit);
- (iii) incumbent firms believe that f is very sensitive to increases in $\bar{p}$; and
- (iv) demand is considerable price elastic, so that a price increase is associated with falling q and so rising $T(q, \cdot)$.

However, if none of these moderating phenomena is present predominantly, then the likelihood of sustainability is increased to the detriment of the static efficiency of market performance, simply because $\alpha > 0$ allows $p > \alpha^*$ without provoking entry. This result is, of course, in line with the literature on entry barriers (for instance, Bain 1956 and Baumol *et al.* 1982a; see also Chapter 7).

13.4 UNCERTAINTY, UNSUSTAINABILITY AND ADJUSTMENT

An important source of intertemporal unsustainability is uncertainty. Take, for example, the case of technical uncertainty as regards process innovation (Chapters 15 and 23). If firms have to base decisions on innovative activity upon imperfect knowledge, then it is likely that the resulting price reductions are smaller than the maximally attainable ones. If firms employ *diverging* knowledge or different procedures of expectations formation, then entry and exit movements occur whenever there are potential entrants which have generated innovations that are superior to the ones of (some) incumbents (Chapter 23). A simple example may illustrate this point.

Assume that an incumbent firm and a potential entrant employ diverging mechanisms of expectations formation with regard to the uncertain innovative opportunities. Then, it is

¹ This scenario is different from the case in Section 11.2, because sunk cost is symmetric and the probability of leapfrogging is zero.

likely that they introduce different investment efforts: $^i sc \neq ^e sc$. This in turn yields different innovative and markup results, $^e \Delta A(q) \neq ^i \Delta A(q)$ and $^e \mu \neq ^i \mu$, which is likely to be associated with $^e p \neq ^i p$. If $^e p < ^i p$, then the potential competitor is able to enter profitably, so that the incumbent firm is expelled from the market. The result is intertemporal unsustainability. However, notwithstanding the uncertainty regarding innovative opportunities, firms are still inclined to implement the *expected* optimal R&D efforts so as to minimize the likelihood of future exit (Chapter 15).

A last remark is in order with respect to the *adjustment* mechanisms reflected in entry and/or exit movements. In the case of identical firms the theory, as it stands, fails to indicate which firms will be the victim or beneficiary of unsustainability (Chapter 6, Section 6.5). The multitude of equilibria implies a coordination problem. The introduction of heterogeneity (for instance, through diverging expectations of innovative opportunities) weakens, of course, the ambiguity. It is obvious that inferior firms are expelled from the market by superior entrants if the latter face a zero opportunity cost of entry. However, what happens if the number of inferior firms does not precisely match the number of superior entrants? Let subscript I and S denote inferior and superior firms respectively.¹ Then, ambiguity arises when $n_I < n_S$ or $n_I > n_S$, which re-introduces the coordination issue. Only if $n_I = n_S$, can it easily be indicated which firms enter and exit.² However, if $n_I \geq n_S$, then the threat of entry is no longer credible, since ultimately no *superior* potential entrant remains which is waiting for profitable entry opportunities (Chapter 20).

Moreover, if $D_0(\text{mac})/q$ fails to yield an integer number of superior firms in equilibrium, then residual demand may not be served at all, or it may be satisfied by a firm producing short (or in excess) of optimal scale. This niche induces behavior very different from the rest of the market. A niche firm may even decide to carry out another investment program. In a way, the niche firm is in the same position as a natural monopolist, since both face a demand constraint in their output decision.

13.5 CONCLUSION

The contestable market framework is a *static* theory of (efficient) competition. The introduction of dynamic aspects of competition in such a framework is associated with a reduced likelihood of temporary and intertemporal sustainability of an industry configuration. The introduction of (entry-detering) sunk investment in a free entry framework implies that the facilitating role of the flat bottom in the average variable and fixed cost curve loses power. Moreover, changes over time induce entry and/or exit movements. However, there are phenomena that moderate this result. To name two examples: (i) if the size of the firm is small relative to market demand, the integer problem is less serious; and (ii) a positive opportunity cost of entry and entry barriers facilitate the *sustainability* of a market configuration.

The difficulty of sunk cost and temporary and intertemporal unsustainability of *market configurations* is ignored in the following chapters, with the exception of Chapters 14, 16,

¹ For the sake of illustrative simplicity, just a dichotomy is employed. This assumption leaves the main argument unaffected. See Chapter 20 for further details.

² Chapter 23 offers a simple simulation experiment with inferior and superior firms as a result of diverging expectations of innovative possibilities. A similar kind of evolutionary natural selection process is simulated by Nelson and Winter (1982; see also Chapter 8). Of course, the contestable market theory describes a *static* equilibrium concept, implying that natural selection arguments are beyond the scope of the analysis.

20 and 23. The results of the subsequent chapters in Parts VII and VIII are not upset by unsustainability problems, since they follow from the analysis of firms' *decision rules* in a one or two-shot barrier market game: that is, the subject matter concerns strategies which facilitate the profitability (survival) of the individual supplier. It appears that an individual firm is always inclined to undertake entry-detering actions in order to avoid exit. However, first Chapter 14 goes on to examine a further possible source of (intertemporal) unsustainability: the amortization of sunk cost if the life time of the investment extends over a number of periods ($T > 1$).

CHAPTER

14

SUNK COST, AMORTIZATION AND SUSTAINABILITY

14.1 INTRODUCTION: Amortization and Unsustainability

The peculiarity of investment is that the cost precedes the benefit. The cost of investment today has to be rewarded by introducing a markup on average cost today and tomorrow. In practice, the amortization procedure is based upon conventional rules. In this chapter it is shown that in an environment of Bertrand competition the existence of a conventional amortization rule is of particular importance as soon as (1) sunk investment is permitted and (2) a unique entry-detering amortization rule is absent (Van Witteloostuijn 1988c). This means that the argument in this chapter is related to, at least, three branches of literature.

- (i) This chapter describes a market environment in which *conventional* (or routinized) behavior constitutes the optimal strategy. So, this chapter presents a setting which permits the derivation of optimal rules of thumb. This means that the argument is relevant in the context of the debate about the maximizing principle (Chapter 4) as the application of rules of thumb can be the *optimal* strategy.
- (ii) The intertemporal unsustainability of a contestable market derives from the *dynamic* aspects of competition. Sunk investments are particularly detrimental, since they require amortization. In effect, the literature suggests that the introduction of sunk costs implies a movement beyond the contestability framework (Chapter 6). This chapter argues that intertemporal sustainability can be retained if firms make use of conventional amortization rules.
- (iii) By making use of a conventional amortization rule a firm signals information with regard to investment efforts. So, the conventional amortization rule can operate as a *signaling device* (Chapter 7). Provided that competitors are informed of the rivals' production cost structure, a publicly known amortization procedure permits rivals to infer the firm's investment scale from quoted prices¹.

¹ A related argument can be found in the literature on information sharing in oligopolistic markets (Clarke 1983a and 1983b, Fried 1984 and Shapiro 1986), which investigates conditions which predict the profitability of sharing private information as regards demand or cost parameters with rivals in different (for example, Cournot or Bertrand) oligopoly settings. Broadly speaking, the result of this investigation is the observation that rivals only decide to share private information in the prospect of imperfect competitive [or (tacitly) cooperative] outcomes from the game on quantities or prices. The argument in the present chapter deviates from this literature by focusing on information about *multi-period* (amortization of investment)

The focus is on a perfect nonstrategic barrier market, although the results apply to any setting with Bertrand competition with cost and product homogeneity. Two cases can be distinguished.

First, if a unique entry-detering amortization rule *cannot* be calculated, the well-established coordination problem arises. Different amortization procedures give different prices and so a multitude of equilibria. In a free entry setting (or actual Bertrand competition) with long-lived investment ($T > 1$) this can induce market instability, since perceived entry opportunities depend on amortization differentials. Section 14.2 illustrates this proposition, and argues that in this case a necessary condition for intertemporal sustainability of a barrier market configuration is that incumbent firms follow a prevailing zero-profit *conventional* amortization rule (Proposition 14.1 and 14.2). In the absence of such a rule, a barrier market configuration is likely to be intertemporally unsustainable (Proposition 14.3). This proposition resembles the well-known argument that conventions are a means to overcome coordination problems in multi-period settings with multiple equilibria (for example, Luce and Raiffa 1957 and Schotter 1981). Second, if a unique entry-detering amortization procedure *does* exist, equilibrium is triggered by this rule. Section 14.3 describes two scenarios which predict a unique entry-detering amortization procedure.

The chapter is organized as follows. Section 14.2 shows that the prevalence of intertemporal sustainability of a barrier market configuration requires the existence of a (zero-profit) conventional rule if a unique entry-detering rule cannot be determined. Section 14.3 examines two illustrative cases that yield a unique entry-detering amortization rule. Section 14.4 reviews some literature that focuses on the role of routinized behavior in economics (as this theme is central to Chapter 4's argument). Section 14.5 offers a conclusion. The model used is based upon the home market benchmark with $T > 1$ discussed in Section 10.3. The focus is on the second-stage game on prices (in the periods $t = 1, \dots, T$) by assuming that both incumbent firms and potential entrants have introduced the same investment in the first stage (period $t = 0$).

14.2 INTERTEMPORAL SUSTAINABILITY AND AMORTIZATION RULES

14.2.1 Amortization Rules

Investment is associated with the sinking of costs. The investment costs have to be covered through the introduction of a markup on average variable and fixed cost (in line with, for example, Kalecki 1943, Andrews 1949 and Earl 1983). Hence,

$$(14.1) \quad p_t = avc_t + mu_t^1,$$

where mu_t denotes the markup in period t . Let sc_t be the sunk costs associated with an investment that is carried out in period t . Suppose that the lifespan of today's ($t = 0$) investment encloses T periods. After T the innovation involved is outdated. So, sc_0 must be recovered from period $t = 1$ to $t = T$. Then, a profitable amortization rule amounts to

rather than single period (demand or cost) parameters.

¹ Strictly speaking, avc_t is a function of q_t (and, in the case of process innovation, sunk R&D cost), whereas mu_t depends on both q_t and sunk cost (Chapter 13). Sunk R&D can shift the average nonsunk cost function if the firm invests in production technology (Chapter 15). For the sake of the argument, Section 14.2 assumes an average nonsunk cost function with constant returns and unchanging production technology.

$$(14.2) \quad G\pi_T = \sum_{t=1}^T \delta_t \cdot mu_t \cdot q_t \geq sc_0,$$

$$(14.3) \quad mu_t = \alpha_t \cdot mu_{t-1}, \text{ and}$$

the α -vector, where $G\pi_T$ denotes the net present value of the gross profit at the end of period $t = T$, $\delta_t = 1/(1+r)^t$, r represents the discount rate and q_t is the cost-minimizing production level.¹ The amortization procedure $(\alpha_2, \dots, \alpha_T)$ plus the desired discounted gross reward ($G\pi_T$) define the series of markups $(mu_1, \dots, mu_T)$ and prices $(p_1, \dots, p_T)$. So, amortization deals with the determination of the time path of markups. The time path of the parameter α_t determines the properties of the amortization rule deployed. If $\alpha_t = 1$ for $t = 2, \dots, T$, then simple flat amortizing results. Provided that $\alpha_t > 1$, for $t = 2, \dots, T$ the rule reflects accelerated amortization, whereas $\alpha_t < 1$ for $t = 2, \dots, T$ is associated with an decelerated amortization procedure. Of course, α_t is not necessarily constant over time. Section 14.2 assumes that a unique entry-detering α -vector cannot be calculated. In this case, a conventional rule can coordinate competition.

14.2.2 Conventional Rules

In practice, amortization procedures often follow a convention. A convention operates as a signaling device. If firms are not well-informed of the rivals' investments (and if the response lag exceeds the entry lag: $R > E$), they can judge rivals' prices on the basis of the prevailing conventional amortization rule. The convention offers a foundation upon which expectations of rivals' investment structure can be based. However, the signaling property of a convention opens the door to the manipulation of information. A firm may deviate from the convention in order to deceive rivals.² Suppose that internal market conditions imply that incumbent firms and potential entrants are able to set a series of markups $(mu_t \text{ for } t = 1, \dots, T)$ such that ${}^{i,G}\pi_T = {}^{e,G}\pi_T \geq {}^{i}sc_0 = {}^{e}sc_0$ [equation (14.2)]; so, ${}^{i,G}\pi_T$ is the blockaded entry profit. Moreover, assume that a conventional amortization rule does exist (superscript *): $mu^*_1, \alpha^*_2, \dots, \alpha^*_n$.³

The structure of the game on convention in a period t is the same as the one in a contestable market (Chapter 6, Subsection 6.4.4). Incumbent firms and potential entrants play a second-stage repeated Bertrand-Nash game (for $t = 1, \dots, T$). Assume that both incumbent firms and potential entrants use the convention up to some period $2 \leq G \leq T$. In period $t = G$ an incumbent firm has three regions of relevant strategies: ${}^i\alpha_G > \alpha^*_G$ (${}^i s^1$), ${}^i\alpha_G = \alpha^*_G$ (${}^i s^2$) and ${}^i\alpha_G < \alpha^*_G$ (${}^i s^3$). A potential entrant is able to reply with the help of two regions of relevant responses: ${}^e\alpha_G < {}^i\alpha_G$ (${}^e s^1$) and non-entry (${}^e s^2$). Actions showing ${}^e\alpha_G \geq {}^i\alpha_G$ are not worthwhile, because they are associated with ${}^e p_G \geq {}^i p_G$ and

¹ In practice, firms carry out a sequence of overlapping investments (Chapter 8) and, thereby, face a series of overlapping amortization charges. Then, the problem basically remains the same, although the required calculations become increasingly complicated. The competitors' signals now result from a composition of amortization charges which gives a summed markup.

² This type of phenomena is examined in, for example, the principal-agent literature (see, for example, Radner 1981 and 1986, Rogerson 1985 and Abreu 1986).

³ This chapter ignores the question of the origin of conventions. For example, conventions may come into existence through regulatory or evolutionary processes (Schotter 1981).

so unsuccessful entry.¹

For the sake of simplicity of notation, discounting is henceforth ignored. A firm's net payoff is determined by

$$(14.4) \quad {}^N\pi_T = \left[\sum_{t=1}^T p_t \cdot d_t \right] - \left[\sum_{t=1}^T \text{avc}_t \cdot q_t \right] - sc_0.$$

Recall that d denotes the quantity demanded from a firm. Whenever entry deterrence is successful: ${}^i d_t = {}^i q_t > 0$ and ${}^e d_t = {}^e q_t = 0$. If a potential entrant succeeds in outperforming an incumbent firm, then ${}^e d_t = {}^e q_t > 0$ and ${}^i d_t = {}^i q_t = 0$. The negative payoff of an expelled firm follows from the fraction μ_t of sc_0 which the firm fails to amortize ($\mu_t \cdot sc_0$), because its revenues have dropped to zero. The value of μ_t depends on the timing (t) of exit (Chapter 11). Hence, the payoff matrix is

$$(14.5) \quad \begin{cases} {}^{i,N}\pi_T \geq 0 \text{ if } {}^i d_t = {}^i q_t > 0 \text{ for } t = 1, \dots, T \text{ (entry deterrence),} \\ {}^{e,N}\pi_T \geq 0 \text{ if } {}^e d_t = {}^e q_t > 0 \text{ for } t = 1, \dots, T \text{ (successful entry),} \\ {}^{i,e,N}\pi_T = - {}^{i,e}\mu_t \cdot sc_0 < 0 \text{ if } {}^{i,e}d_t = 0 \text{ [exit (cost)], and} \\ {}^{e,N}\pi_T = 0 \text{ if no entry occurs (non-entry),} \end{cases}$$

where ${}^N\pi_T$ denotes the net present value of the net profit at the end of period $t = T$.

Assume an incumbent firm which deviates from the convention by deploying ${}^i\alpha_G > \alpha_G^*$ in the period $1 \leq G \leq T$. Then, given ${}^i\text{avc}_G = {}^e\text{avc}_G$, ${}^i sc_0 = {}^e sc_0 = sc_0$ (Assumption 10.2) and the assumption that both incumbent firms and potential entrants have used the amortization convention up to period $t = G$, the incumbent firm's price exceeds the one associated with α_G^* . Hence, a potential entrant can outperform the incumbent firm profitably by adopting $\alpha_G^* < {}^e\alpha_G < {}^i\alpha_G$. Moreover, if $G < T$ enduring and profitable re-entry in later periods is unlikely, because then profitability requires that (part of) sc_0 is recovered in fewer periods, leading to higher amortization charges and prices during the time left.

Suppose that an incumbent firm deviates from the convention by adopting ${}^i\alpha_G < \alpha_G^*$ in the period $1 \leq G \leq T$, which enables the firm to offer a lower price in period $t = G$. Then, none of the potential entrants (which use the conventional amortization rule) is able to enter profitably in period $t = G$. However, if $G = T$ or if the incumbent firm returns to the conventional amortization rule in later periods (for $G < T$), then ${}^{i,N}\pi_T = {}^{i,G}\pi_T - sc_0 < 0$, because the convention is associated with zero profits and ${}^i\alpha_G < \alpha_G^*$. (Above-)zero-profits require ${}^i\alpha_H > \alpha_H^*$ in some period $G < H \leq T$. Then, however, profitable entry is provoked. For the same reasons, a potential entrant cannot enter profitably whenever incumbent firms follow the convention, because then entry requires ${}^e\alpha_G < \alpha_G^*$ in the period $t = G$ of entry. In effect, potential entrants are able to capture ${}^{e,N}\pi_T = 0$ if they abandon entry. Hence, ${}^e\alpha_G < \alpha_G^*$ always reflects an inferior alternative. The extensive-form tree of the game can be constructed (Figure 14.1).

¹ Note that for period $t = 1$ this is in terms of μ_1 rather than α_G . For the sake of brevity, period $t = 1$ is henceforth ignored (bearing in mind that analogous reasoning can be applied).

Figure 14.1
The Game on Convention

Ignoring other conditions for sustainability (Baumol *et al.* 1982a and Chapter 13), the extensive-form tree shows that the strategy pair $i_G = \alpha_G^*$ and non-entry constitutes a unique Nash equilibrium provided that expelled firms capture a negative payoff.

PROPOSITION 14.1 (obedience). A necessary condition for intertemporal sustainability of a barrier market configuration is that incumbent firms comply with the prevailing conventional amortization rule: $i_\alpha^*, \dots, i_\alpha^*$ for $i = 1, \dots, ni$, conditional upon

- incumbents fearing that there is at least one equally-equipped potential entrant which faces a zero opportunity cost of entry (credibility), and
- expelled firms being confronted with a negative payoff: $\mu_t > 0$ for $t = 1, \dots, T$ (uniqueness).¹

Obedience to the convention constitutes a unique Nash equilibrium strategy.

14.2.3 Zero-Profit Amortization

Subsection 14.2.2 assumed that $i_G \pi_T = e_G \pi_T \geq i_{sc_0} = e_{sc_0} = sc_0$. However, the perfect barrier market defines $e_G \pi_T = e_{sc_0}$ if entry is abandoned (Chapters 11, 12 and 13). The incumbent firm can use the conventional amortization rule ($i_{\mu_1}, i_{\alpha_2}, \dots, i_{\alpha_T}$) so as to attain positive ($i_G \pi_T > sc_0$) or zero ($i_G \pi_T = sc_0$) profit (provided that internal market conditions permit a $i_N \pi_T > 0$). Is amortization that gives a positive profit sustainable? Incumbent firms and potential entrants use the amortization convention (Proposition 14.1). In a barrier market incumbent firms and potential entrants play a (second-stage repeated) Bertrand-Nash game on amortization (for $t = 1, \dots, T$). The structure of the game is the same as the one in Subsection 14.2.2.

Competitors are able to influence the revenues by manipulating prices. Incumbent

¹ This condition implies a movement beyond Assumption 6.7.

firms have two regions of relevant strategies: $i, G \pi_T > sc_0$ (s^1) and $i, G \pi_T = sc_0$ (s^2). Strategies reflecting $i, G \pi_T < sc_0$ are dominated, because they generate $i, N \pi_T < 0$. Potential entrants are able to reply with the help of two regions of relevant responses: $e, G \pi_T < i, G \pi_T$ (e, s^1) and non-entry (e, s^2). Actions showing $e, G \pi_T \geq i, G \pi_T$ are assumed not to be worthwhile, because they are associated with at least one $e, p_t \geq i, p_t$, which implies unsustainability (Subsection 14.2.2). The payoff matrix resembles matrix (14.5).

If a market is ruled by a threat of "hit-and-run" entry, incumbent firms are forced to adopt amortization rules which satisfy a zero-profit condition. In effect, zero-profit amortization is a Nash equilibrium strategy. The argument is equivalent to the one underlying Proposition 6.1. Suppose that an incumbent firm employs an amortization rule which implies the possibility to earn positive profits ($i, G \pi_T > sc_0$). However, then a potential entrant which captures a zero profit if entry is abandoned, is able to enter profitably by adopting an amortization rule which implies $sc_0 < e, G \pi_T < i, G \pi_T$.¹ That is, provided that both incumbent firms and potential entrants use the amortization convention (Proposition 14.1), this means that $e, p_t < i, p_t$ for $t = 1, \dots, T$. An alternative strategy is to adopt a zero-profit amortization rule. Then, $i, G \pi_T = sc_0$, so that a potential entrant cannot enter profitably. Successful entry requires $e, p_t < i, p_t$ and thus $e, G \pi_T < i, G \pi_T = 0$. The extensive-form tree of the game can be constructed, showing that the strategy pair $i, G \pi_T = sc_0$ and non-entry constitutes a unique Nash equilibrium if $\mu_t > 0$ (for $t = 1, \dots, T$). The extensive-form tree of the game is equivalent to the one of the game in a contestable market (Figure 6.2).

PROPOSITION 14.2 (zero-profit amortization). *A necessary condition for intertemporal sustainability of a barrier market configuration is that incumbent firms follow zero-profit conventional amortization rules: $i, G \pi_T = sc_0$ for $i = 1, \dots, n_i$, conditional upon*

- (i) *incumbents fearing that there is at least one equally-equipped potential entrant which faces a zero opportunity cost of entry (credibility), and*
- (ii) *expelled firms being confronted with a negative payoff: $\mu_t > 0$ for $t = 1, \dots, T$ (uniqueness).*

The use of a zero-profit amortization procedure is then a unique Nash equilibrium strategy.

Hence, two necessary conditions for sustainability of a barrier market configuration is that incumbent firms adopt the zero-profit amortization rule (Proposition 14.2) which complies with the prevailing convention (Proposition 14.1). Neither assumption is sufficient. On the one hand, the zero-profit condition is not sufficient, because a zero-profit markup rule is generally not associated with a unique amortization procedure ($\mu_1, \alpha_2, \dots, \alpha_n$). On the other hand, compliance with the prevailing market convention such that a positive profit will be captured if entry is abandoned, attracts profitable entry by potential entrants which calculate slightly lower markups. This insight is clarified in Subsection 14.2.4.

¹ Strictly speaking, it is moreover assumed that incumbent firms and potential entrants use the same discount rate. This can follow from the observation that the discount rate is equal to the prevailing interest rate. Here (and in other chapters) discounting is ignored.

14.2.4 Intertemporal Unsustainability

A convention breaks down if competitors deviate time and again from the prevailing rule. In the absence of a convention, competitors cannot rely on a fixed amortization rule while judging price setting behavior of rivals. The solid foundation of conjectures of rivals' amortization procedures is transformed into a shaky one. Heterogeneity of competitors as regards the amortization procedures employed is likely to prevail. Even if the condition of zero profits is sustained, the stability of the market configuration is no longer guaranteed. A firm can always undercut the rival's price by using a different amortization rule.

Assume that incumbent firms and potential entrants have to amortize the same investment outlays. The structure and payoff matrix of the game on nonconvention are equivalent to the one on convention. In the absence of a conventional amortization rule, *intertemporal unsustainability* of the market is likely to occur. The reason for this result is straightforward. Two cases can be distinguished as the specifics of the result depend on the assumption as regards to the (potential) entrants' anticipation of future exit.

Suppose, firstly, that entrants fail to foresee that it is likely that they are forced to exit one period after their entry. If an incumbent firm adopts some α_t in period t , a potential entrant is *always* able to undercut the incumbent firm's price by offering a $\alpha_t < \alpha_t$ (that is, a lower markup) and, therefore, to enter successfully. In fact, the heterogeneity of competitors regarding the amortization procedures employed induces successful entry, because it is likely that there is a potential entrant which follows (in a home market) an amortization procedure implying $\alpha_t < \alpha_t$. However, the potential entrant in period t is the incumbent firm in the next period $t+1$. The incumbent firm in period $t+1$ is vulnerable to "hit-and-run" entry in the same fashion as the one in period t . Hence, in each period t a second mover is able to outperform a first mover. So, the result is a process of recursive entry and exit.

Continually re-occurring entry and exit movements are of course associated with the opportunity to learn. Therefore, assume, secondly, that (potential) entrants are aware of the fact that staying in the market results in exit in the next period, and thereby losses: $\pi_{t+1} = \mu_{t+1} \cdot sc_0 < 0$. An expelled firm forfeits the opportunity to recover (part of) the invested cost sc_0 in period $t+1$.¹ Hence, the optimal strategy to follow is "hit-and-run" entry if an entrant is able to return to its home market after entry. A potential entrant employing an amortization procedure such that $\alpha_t < \alpha_t$ takes advantage from the temporary opportunity to collect higher revenues and departs without accruing a loss.

Nevertheless, whether or not (potential) entrants anticipate future exit, in the absence of a conventional amortization rule intertemporal sustainability of a barrier market configuration is unlikely to emerge.

PROPOSITION 14.3 (*intertemporal unsustainability*). *In the absence of a conventional amortization rule, intertemporal unsustainability of the market is likely to occur. If (potential) entrants fail to anticipate future exit, then firms enter and exit recursively. If (potential) entrants foresee their exit if they stay in the market, then "hit-and-run" entry results.*

¹ In line with the argument in Chapter 11 it can be that the expelled firm is either unable to sell any output (a positive response lag and an alternative entry market being absent) or can enter another market which is associated with lower profitability (reciprocal entry or entry into a third market).

14.3 UNIQUE ENTRY-DETERRING AMORTIZATION RULE

A convention is a coordinating device if selection among a multitude of equilibria is at stake. However, if a *unique* entry-detering amortization rule can be calculated, rivalry in a barrier market dictates the use of this rule. In this section two examples are discussed. On the one hand, a model is presented with a *product innovation* that is subject to time diseconomies.¹ On the other hand, Baumol *et al.*'s (1982a) example with process innovations is outlined. Both examples are described only by way of illustration.

First, take a firm which faces the need to introduce a *product innovation*.² Assume that there is an existing product 1 and a product innovation 2. Let product 1 be the dominant product in period $t = 0$. It is technologically feasible to introduce product 2. To do this, it is necessary to spend R&D outlay ic . The size of ic depends on the date of introduction of product 2 (x), where $ic(x)$ decreases in x . That is, $ic(x) > 0$ and $\partial ic/\partial x < 0$ for $x = 1, \dots, T$, where T is the date when the innovation is outdated (the life time of product 2 being $1-T$). Assume that the introduction date $x > 0$ is given. Let p_t be the sequence of prices for product 2 ($t = x, \dots, T$). Then, $ic(x)$ is recovered (provided that zero profits are earned; Proposition 14.2) if

$$(14.6) \quad ic(x) = \sum_{t=x}^T \mu_t \cdot q_t,$$

where, of course, $\mu_t = p_t - avc_t$ reflects the amortization charge necessary to recover the R&D outlays $ic(x)$. Moreover, discounting is ignored.

From condition (14.6) it follows that, *ceteris paribus*, the sequence of prices may be anything as long as the zero-profit condition is satisfied. If this is true, only a conventional amortization rule can guarantee intertemporal sustainability (Proposition 14.1). However, any firm (incumbent or entrant) which follows the innovator by introducing product 2 in the next period ($x+1$), faces the need to spend an R&D outlay of $ic(x+1)$. This does not introduce an asymmetry, since the first innovator too had the option to postpone the introduction of product 2 to period $x+1$. It may be that firms are indifferent to the introduction date. In period $x+1$ condition (14.6) holds in a modified way. That is,

$$(14.7) \quad ic(x+1) = \sum_{t=x+1}^T \mu_t \cdot q_t.$$

Combining condition (14.6) and (14.7) yields the simple result that the difference between $ic(x)$ and $ic(x+1)$ should be recovered in the first period. This provides the solution of p_x , which follows from $ic(x) - ic(x+1) = \mu_x \cdot q_x = [p_x - avc_x] \cdot q_x$. Or,

$$(14.8) \quad p_x = avc_x + [ic(x) - ic(x+1)]/q_x.$$

In the same manner, $p_t = \{p_{x+1}, \dots, p_T\}$ can be determined. This means that there is a *unique entry-detering* series of amortization charges and prices. Equation (14.8) reduces

¹ This example resembles Kamien and Schwartz' (1982) time-cost tradeoff dilemma and the patent race models (Chapter 8).

² This example is derived by Marc van Wegberg.

the multiperiod pricing decision for a long-lived innovation to a two-period one. In general terms this implies that $ic(t) - ic(t+1)$ is a sunk cost which is recovered by a markup implicit in the pricing rule reflected in equation (14.8). Here the sunk investment is defined as $sc_t = ic(t) - ic(t+1)$. Intertemporal *entry deterrence* is associated with employing the unique amortization schedule implicit in $p_t = \{p_x, \dots, p_T\}$.

Second, another example of the calculation of a unique entry-detering amortization schedule is Baumol *et al.*'s (1982a) model with *process innovation*. Baumol *et al.* (1982a, pp. 371-404) indicate that the intertemporal sustainability of market configurations is preserved in special cases. They discuss a model with intertemporally interdependent production. Then, sunk costs and the need for amortization are of crucial importance. In effect, this implies a movement beyond existing contestability models, because sunk cost may exceed zero. Baumol *et al.* restrict their analysis to those cases where the minimization of cost yields a unique solution. Then, in particular physical deterioration of capital, inflation or deflation, demand fluctuations and technological change enter the picture. The model focuses on investment in replacement, expansion or reduction of productive capacity and process innovations.

A firm in a market with free entry faces the need to minimize the intertemporal cost function. That is,

$$(14.9) \quad atc(q_1, \dots, q_T) = \min \sum_{t=1}^T avc_t(q_t, K_t) \cdot \delta_t \\ + \sum_{t=1}^T \tau_t \cdot IK_t \cdot \delta_t - \sum_{t=1}^T \beta_t \cdot SK_t \cdot \delta_t - \beta_{T+1} \cdot K_{T+1} \cdot \delta_{T+1},$$

subject to

$$(14.9a) \quad K_t = (1 - \sigma_{t-1}) \cdot K_{t-1} + IK_t - SK_t \geq 0,$$

$$(14.9b) \quad IK_t \geq 0, SK_t \geq 0 \text{ and } \tau_t > \beta_t,$$

where K denotes capital, τ the price of buying a unit of capital, IK the addition to productive capacity, β the price of selling a unit of capital, SK the scrapping of productive capacity and σ a parameter indicating the rate of physical deterioration. Capital is measured in physical units. Changes in capital emerge through physical deterioration and buying and selling of units of capital. The term $\beta_{T+1} \cdot K_{T+1} \cdot \delta_{T+1}$ captures the revenue from selling the remaining capital in period $t = T+1$, when the active life of the firm has ceased.

The minimization decision reflected in the equations (14.9) to (14.9b) introduces three reasons for investment. First, physical deterioration of capital demands for replacement investment. This leads, *ceteris paribus*, to $(1 - \sigma_{t-1}) \cdot K_{t-1} = IK_t$. Second, changing prices of capital, implying that β_t and τ_t vary over time, influence, *ceteris paribus*, the optimal timing of investment IK_t and disinvestment SK_t . Third, demand fluctuations force firms to adapt productive capacity. This result can be easily illustrated with the help of two extreme cases: continually increasing or decreasing demand. On the one hand, if demand increases in each period t , then, *ceteris paribus*, $SK_t = 0$ and $IK_t > (1 - \sigma_{t-1}) \cdot K_{t-1}$ such that the firm is able to meet increased demand. On the other hand, if demand is

continually decreasing, *ceteris paribus*, $IK_t = 0$ and $SK_t \geq 0$. SK_t exceeds zero if the deterioration of capital occurs insufficiently fast so as to keep pace with decreasing demand. Fourth, technological change induces the replacement of old by new capital. The firm decides on the time path of adaptations of the mixture of old and new capital. This mixture is associated with a variable cost function $avc_t(q_t, K_t, {}^bK_t)$, where superscript b denotes new capital. Baumol *et al.* (1982a, pp. 390-391) point out that technological progress in the form of a shift of the variable cost function is equivalent to a decrease of β_t over time.

Baumol *et al.* prove that the prevalence of intertemporal sustainability of a "contestable" market configuration requires zero-profit amortization (1982a, pp. 384-389 and 396-397). Moreover, their special case is associated with a determinate solution of the time path of investment, disinvestment and amortization charges (1982a, pp. 294-296). Given knowledge of the rate of physical deterioration of capital, price changes, demand fluctuations and technological change, minimization of equation (14.9) subject to conditions (14.9a) and (14.9b) enables the calculation of a unique solution. Incumbent firms are forced to minimize intertemporal cost (*atc) in order to deter entry. If an incumbent firm adopts some ${}^1atc < {}^*atc$, then a potential entrant can outperform the incumbent firm by implementing an investment schedule such that ${}^*atc \leq {}^1atc < {}^1atc$.

14.4 CONVENTIONS, ROUTINES AND FOCAL POINTS

Although conventional behavior is not located in the center of economic thinking, it has nevertheless received considerable attention (Chapters 3 and 4). In this section some literature on conventional behavior is mentioned in order to shed additional light upon the role conventional rules play in economic behavior. For the sake of brevity, the following is restricted to Keynes' conventions, Simon's routines and Schelling's focal points and Schotter's social conventions.

Keynes (1936) introduced conventions as a foundation of expectations formation in the context of investment decisions. He argued that

[i]n practice we have tacitly agreed, as a rule, to fall on what is, in truth, a *convention*.¹ The essence of this convention ... lies in assuming that the existing state of affairs will continue indefinitely, except in so far as we have specific reasons to expect a change (Keynes 1936, p. 152; Keynes' italics).

The advantage of a convention is that it

will be compatible with a considerable measure of continuity and stability in our analysis, so long as we can rely on the maintenance of the convention (Keynes 1936, p. 152; Keynes' italics).

Post-Keynesian economists (for example, Davidson 1972 and Kregel 1976) stipulate the role of conventions in reducing uncertainty. It is the facilitating role of conventions in maintaining continuity and stability which also emerges in the context of a market with free entry. In a barrier market compliance with a conventional amortization rule is necessary in order to deter entry. In the absence of a unique entry-detering convention,

¹ Van Witteloostuijn and Maks (1989 and 1990) argue that Walras (1874) introduces a procedure of expectations formation which shows a close resemblance to Keynes' convention.

instability occurs. The convention reduces uncertainty by coordinating rivalry.

An important concept of Simon's (for instance, 1945) behaviorist theory of organizations are *routines*. Cyert and March (1963) argue that

[t]he organization uses standard operating procedures and rules of thumb to make and implement choices. In the short-run these procedures dominate the decisions made (Cyert and March 1963, p. 113).

Routines are employed to shortcut decision making. In many circumstances routinized behavior occurs as an optimal strategy (Baumol and Quandt 1964 and Chapter 4). If a unique entry-detering amortization rule is absent, this is also the case in a barrier market in which firms are forced to follow the prevailing conventional amortization rule in order to deter entry.

An interesting notion of conventional behavior is reflected in Schelling's (1960) *focal points*. Schelling's focal points are substitutes of direct communication. If behavior has to be coordinated tacitly, in many occasions (with a multitude of equilibria) it makes sense to follow some focal point, irrespective of its origin. It is only required that the focal point is unique. Scherer (1980, pp. 190-193) applies Schelling's notion of focal points to pricing decisions in oligopolistic markets in which collusion is conspicuous by its absence. Scherer argues

that adherence to accepted pricing points facilitates coordination and discourages price warfare (Scherer 1980, p. 191).

In fact, this is precisely what a conventional amortization rule in an environment of free entry without a unique profit maximizing amortization rule brings about as well.

Schotter (1981) defines social conventions as

regularities in behavior which are agreed to by all members of a society and which specify behavior in specific recurrent situations (Schotter 1981, p. 9).

The key point is that

coordination can be achieved by establishing a social convention or regulatory in behavior to govern these situations [with otherwise multiple equilibria] ... Which specific convention will evolve is indeterminate, yet both [parties] are in equilibrium, because not only does neither party have an incentive to deviate from the convention, both desire that the other not deviate either. The convention determines equilibrium behavior (Schotter 1981, p. 10).

Again, this is the function attributed to a conventional (zero-profit) amortization rule in a barrier market without a unique entry-detering procedure (the two parties being incumbent firms and potential entrants). Price only is an insufficient information statistic to coordinate rivalry in imperfect competitive markets. That is,

prices are mechanisms that provide information about the societal scarcity of resources, whereas institutions [conventions] are mechanisms that supply information about the potential actions of other economic agents. Consequently, in markets where prices are not informationally sufficient to coordinate economic activity, some additional informational device will have to emerge to help this coordination, because a problem of strategic interdependence among agents exists and the need to coordinate them has arisen. The informational devices that emerge to settle such problems are institutions or conventions. ...

perfect competitive markets are markets in which the equilibrium is completely characterized by prices.

In what are traditionally called oligopolistic markets, this is not true. ... Each firm must have information about the intended actions of others, and this information is transmitted through the existing convention of behavior (Schotter 1981, p. 157).

14.5 CONCLUSION

The intertemporal unsustainability of (perfectly) contestable markets derives from *dynamic* aspects of competition, in particular insofar as it concerns the investment of sunk costs. The presence of any positive sunk cost introduces the need for amortization. If a unique entry-detracting amortization rule is absent, the analysis of intertemporal sustainability of a barrier market configuration shows that the firms' optimal strategy is to comply with some zero-profit conventional amortization rule. If a convention exists, a deviating incumbent firm is, sooner or later, outperformed by a potential entrant. In the absence of a convention, intertemporal unsustainability emerges in the sense that firms enter and exit recursively or deploy "hit-and-run" entry. In the next chapters the problem of amortization in a multiperiod setting is bypassed by assuming an innovation with a one-period life time which can be introduced after a one-period investment lag ($B = 0$, $L = T = 1$).

PART

VII

COMMODITY MARKET

Cost Reduction and Product Innovation

KEY WORDS. Part VII examines the strategies which incumbent firms in a (perfect) barrier commodity market can follow so as to reach a position of exit avoidance (provided that market unsustainability does not interfere). The results are illustrated with the help of the two-period - two-market benchmark model from Section 10.3. The focus is on process and product innovation. Incumbent firms have to undertake sunk R&D in order to keep ahead of (or, to be precise, to keep pace with) the potential entrants' innovation. Firms are forced to minimize future price (Chapter 15) and maximize future buyers' utility (Chapter 16) by carrying out efficient sunk R&D. Under particular conditions, a unique R&D scale does exist, which predicts the incumbent firms' equilibrium strategy.

 

 

 

CHAPTER

15

PROCESS INNOVATION: Minimization of Price

15.1 INTRODUCTION: Cost Reductions

The inventors of the contestable market also pay attention to the effects of sunk costs in a multiperiod framework. In particular, in Chapter 13 of their 1982 book Baumol *et al.* offer an analysis of intertemporally sustainable contestable markets in which firms are able to buy technologically improved capital (Chapter 14). That is, process innovation is introduced. However, in contestability theory technological progress (in the form of cheaper capital) occurs like pennies from heaven, because (i) it is associated with perfect foresight and (ii) it is not allowed to erect any sunk cost (Stiglitz 1987; see also Chapter 6). In the barrier market model technological progress is *endogenized* in an environment with technical uncertainty (Van Witteloostuijn 1989c).

In this chapter two important barrier market results are derived: (i) incumbent firms are forced to minimize the feasible future price by undertaking the (expected individually) optimal investment program in order to deter entry (Proposition 15.2); and (ii) under particular conditions there exists a unique optimal scale of R&D activity (Proposition 15.1, 15.3 and 15.4). In advance, three features of the model are worth noting: first, it deals with single product firms; second, process innovations are private goods¹; third, technical advance is associated with Hicksian neutrality. The two-period - two-market benchmark model from Section 10.3 is used as the point of departure.

The plan of the chapter is as follows. Section 15.2 introduces the R&D technology that governs cost reducing innovative activity. Section 15.3 points out that (incumbent) firms are inclined to minimize the (expected) future price in order to avoid exit. This is done by directing R&D efforts to reaching cost reductions. Section 15.4 shows that under particular conditions an (expected individually) optimal scale of R&D outlays exists. Section 15.5 offers a remark on welfare theory. Since it leaves the mathematics of the model unaffected, the results are presented as if there were no technical uncertainty.²

15.2 PROCESS INNOVATION: The R&D Technology

Firms' R&D is directed at reducing average cost. In accordance with Baumol *et al.* (1982a) the special case of a flat-bottomed U-shaped average variable and fixed (nonsunk)

¹ This means that imitation is excluded. The appropriability of the innovation is perfect. However, there are no exclusive patents. For different models see Chapter 8.

² The impact of technical uncertainty, however, re-occurs in Chapter 23.

cost curve is adopted. Hence (suppressing the firms' index),

$$(15.1) \quad \text{avc}_t = A(\epsilon_t, q_t)^1, \text{ with}$$

$$(15.1a) \quad \partial A / \partial \epsilon_t < 0,$$

$$(15.1b) \quad q_t < L_{q_t} \rightarrow \partial A / \partial q_t < 0 \text{ (increasing returns),}$$

$$(15.1c) \quad L_{q_t} \leq q_t \leq R_{q_t} \rightarrow \partial A / \partial q_t = 0 \text{ (constant returns), and}$$

$$(15.1e) \quad q_t > R_{q_t} \rightarrow \partial A / \partial q_t > 0 \text{ (decreasing returns),}$$

where ϵ denotes the state of the technology and $t = 0, 1$. A is a piecewise twice continuously differentiable cost function with $0 < L_{q_t} \leq R_{q_t}$ and $A(\epsilon_t, 0) = 0$. In accordance with Baumol *et al.*'s assumption the cost function (15.1) only contains *nonsunk* cost components.

The investment of sc_0 sunk costs in period $t = 0$ results in a benefit in period $t = 1$ in the form of reduced average cost (Δavc_1). The condition of zero profits (Proposition 14.2) determines the pricing rule [equation (14.1) and (13.2)]. Substitution of equation (13.2) into (14.1) shows that incumbent firms employ (minimum) average *total* cost pricing.

Innovative activity is assumed to be associated with decreasing returns², reflected in a function E :

$$(15.2) \quad \Delta \epsilon_1 = \epsilon_1 - \epsilon_0 = E(sc_0), \text{ with}$$

$$(15.2a) \quad \partial E / \partial sc_0 > 0 \text{ and,}$$

$$(15.2b) \quad \partial^2 E / \partial sc_0^2 < 0,$$

where E is a twice continuously differentiable, monotonic increasing function. The reduction of average cost simply follows from equation (15.1):

$$(15.3) \quad \Delta \text{avc}_1 = \text{avc}_0 - \text{avc}_1 = A(\epsilon_0, \cdot) - A(\epsilon_1, \cdot).$$

Bearing in mind that $\epsilon_1 = \epsilon_0 + \Delta \epsilon_1 = \epsilon_0 + E(sc_0)$, equation (15.3) gives Δavc_1 as a function (R) of sc_0 . In other words, the reduction of average cost depends on the investment outlays that were incurred in period $t = 0$:

$$(15.4) \quad \Delta \text{avc}_1 = A(\epsilon_0, \cdot) - A[\epsilon_0 + E(sc_0), \cdot] = R(sc_0, \cdot),$$

where R is twice continuously differentiable.

¹ Factor prices are not included in the cost function, because the assumption of Hicksian neutrality implies that these prices do not make any difference while deciding on process innovation. In the case of Hicksian nonneutrality it is necessary to take into account the prices of capital and labor in particular (Chapter 8).

² This is in accordance with theoretical and empirical research (for example, Kamien and Schwartz 1982, pp. 64-70 and 194; see also Chapter 8).

PROPOSITION 15.1 (decreasing returns average nonsunk cost reduction). *The decreasing returns property of the R&D technology re-occurs in a decreasing returns relationship between innovative activity and the reduction in average nonsunk cost if one of the following three conditions holds: (i) $\partial^2 A / \partial \epsilon_1^2 > 0$; (ii) $\partial^2 A / \partial \epsilon_1^2 = 0$; or (iii) $\partial^2 A / \partial \epsilon_1^2 < 0$ and $\partial^2 A / \partial \epsilon_1^2 \cdot \partial E / \partial sc_0 < \partial A / \partial \epsilon_1 \cdot \partial^2 E / \partial sc_0^2$.*

A proof is offered in Appendix B. Proposition 15.1 indicates the conditions which give the R&D technology predicting decreasing returns from sunk investment outlay in reducing average nonsunk cost: that is, $\partial R / \partial sc_0 > 0$ and $\partial^2 R / \partial sc_0^2 < 0$.

15.3 MINIMIZATION OF THE FUTURE PRICE

Equations (15.1) to (15.4) describe the implications of investment outlays for average nonsunk cost. However, in which way does a firm's decision making proceed? It appears that firms are forced to minimize the feasible future price in order to deter entry by adopting an (expected individually) optimal investment strategy. This is done by undertaking R&D such that (suppressing the firms' index)

$$(15.5) \quad \min_{\Delta avc_1, \mu_1} p_1 = avc_0 - \Delta avc_1 + \mu_1 = {}^*p_1.$$

$$(15.5a) \quad \text{subject to } \pi_1 \geq 0.$$

Substitution of equation (13.2) and (15.1) into (15.5) gives p_1 as a function (P) of sc_0 :

$$(15.6) \quad \min_{sc_0} p_1 = avc_0 - R(sc_0, \cdot) + M(sc_0, \cdot) = P(sc_0, \cdot),$$

subject to condition (15.5a). Equation (15.6) is associated with an optimal scale of R&D outlays, *sc_0 . P is twice continuously differentiable. Entry-detering strategy (15.6) implies that incumbent firms seek to minimize the future average total cost price.

Firms play a first-stage game on process innovation. Let *p_1 denote the minimum zero-profit (average total cost) price which a firm is able to reach with the help of the optimal scale of R&D outlays *sc_0 . For the moment, suppose that *sc_0 is unique (Proposition 15.4). The firms' actual average total cost prices ip_1 and op_1 are the result of the actual R&D expenditures isc_0 and osc_0 respectively (zero-profit pricing follows from the second-stage game on prices; Proposition 14.2). An incumbent firm has two regions of relevant strategies (is): ${}^ip_1 = {}^*p_1$ (${}^is^1$) and ${}^ip_1 > {}^*p_1$ (${}^is^2$). Actions yielding ${}^ip_1 < {}^*p_1$ are dominated, because they are associated with a negative payoff. The firm is then unable to recover fully the sunk costs invested. A potential entrant is able to reply with the help of two regions of relevant responses (es): ${}^op_1 < {}^ip_1$ (${}^es^1$) and non-entry (${}^es^2$). Strategies showing ${}^op_1 \geq {}^ip_1$ are not worthwhile, because then the entrant cannot outperform the incumbent.

Recall that a firm's payoff is determined by

$$(15.7) \quad \pi_1 = p_1 \cdot d_1 - A(\cdot, q_1) \cdot q_1 - sc_0.$$

Non-entry is associated with ${}^op_1 = 0$: that is, the opportunity cost of entry is zero. For the sake of simplicity of notation, the time index is suppressed. Bearing in mind that $d =$

$i_q > 0$ and $e_d = e_q = 0$ when entry deterrence is successful and that $i_d = i_q = 0$ and $e_d = e_q > 0$ if a potential entrant succeeds in outperforming an incumbent firm ("hit-and-run" entry), the following payoff matrix results:

$$(15.8) \quad \begin{cases} i_{\pi} = 0 \text{ if } i_d = i_q > 0 \text{ (entry deterrence),} \\ e_{\pi} > 0 \text{ if } e_d = e_q > 0 \text{ ("hit-and-run" entry),} \\ i_{\pi} = -sc \text{ if } i_d = i_q = 0 \text{ [exit (cost)],} \\ e_{\pi} = 0 \text{ if } e_d = i_q = 0 \text{ (non-entry).} \end{cases}$$

The immediate result is that pervasive fear of entry forces firms to introduce (expected individually) optimal investment programs.

PROPOSITION 15.2 (feasible price minimization). *In order to deter entry an incumbent firm in a barrier market is forced to minimize its feasible future price (p_1) by undertaking an optimal scale of R&D (sc_0). The introduction of the feasible minimum future price $p_1 = p_1$ by carrying out an optimal investment effort $sc_0 = sc_0$ constitutes a unique (subgame) perfect equilibrium strategy.*

Proof Proposition 15.2. The reason is straightforward (suppressing time indices). Perfect contestability (Proposition 6.1) and perfect barrier market theory (Proposition 14.2) predict the second-stage equilibrium to be average total cost pricing (provided that incumbent firms and potential entrants face parity). If an incumbent firm introduces investment outlays i_{sc} in the first stage such that the average total cost price i_p exceeds p in the second stage, then a potential entrant which receives zero profits if entry is abandoned, is able to enter profitably after undertaking an investment program e_{sc} yielding $p < e_p < i_p$. The expelled incumbent firm is then confronted with a negative payoff, while the entrant is able to collect temporary profits (Baumol 1982, p. 4).¹ Whenever an incumbent follows sc and so $i_p = p$, a potential entrant cannot enter successfully without incurring losses. Successful entry requires $e_p < i_p = p$, which is associated with the impossibility to recover fully the sunk costs invested, because then the entrant's price falls below the feasible minimal one. The extensive-form tree of the game can be constructed, showing that the strategy pair $p = p$ (associated with $i_{sc} = sc$) and non-entry constitutes a unique (subgame) perfect equilibrium (Figure 15.1).

Only the incumbent firm's strategy $p = p$ is associated with a nonnegative payoff, given the set of optimal replies of the potential entrant. Hence, in order to prevent a potential entrant from being able to profit from a second-mover advantage, the incumbent firm adopts price minimization by carrying out optimal innovative activities. *Q.E.D.*

¹ Aside from Proposition 14.2, the minimization of the future price facilitates of course a zero profit, for any $mu_1 > sc_0$ can be reduced to $mu_1 = sc_0$ without violation of condition (15.5a), which is associated with a further fall in the price.

Figure 15.1
The Game on Process Innovation

Here it should be stressed that it is likely that unsustainability arises whenever potential entrants can leapfrog over the incumbent firms' innovations (see, for instance, Fudenberg *et al.* 1983; see also Chapter 8 and 13). So, it is henceforth assumed that there is a series of incremental innovations [following from the R&D technology (15.2) and (15.4)] that makes it impossible for rival firms to leapfrog if incumbent firms do not deviate from optimality. Hence, a firm is forced to seek for a position of exit avoidance and entry deterrence through implementing (expected individually) optimal investment programs.

15.4 OPTIMAL R&D OUTLAYS

The conditions under which a unique optimal investment program exists remain, however, to be proven.

PROPOSITION 15.3 (*existence optimal investment outlays*). *There exists at least one optimal scale of R&D activity, which is determined by $\partial R/\partial sc_0 = 1/q_1$, if Proposition 15.1 applies.*

A proof is presented in Appendix C. The interpretation of Proposition 15.3 is straightforward. The future price is minimized if the marginal reduction of average nonsunk cost is equal to the marginal increase of the markup (*i.e.*, unit sunk cost), resulting from a marginal increase of the scale of sunk costs invested. In fact, the goal to minimize the feasible future price boils down to the minimization of average *total* cost with the help of innovative activity, where the cost consists of variable, fixed and sunk (investment) components.

PROPOSITION 15.4 (*unique optimal investment outlays*). *There exists a unique optimal scale of R&D outlays if four conditions hold:*

- (i) Proposition 15.1 applies;

that $\Delta \text{avc}_1 = \max \Delta \text{avc}_1$ for $sc_0 = 0$, so that again zero-investment is the optimal strategy (H in Figure 15.2). An extreme assumption is that $E(0), R(0,.) = 0$ [equation (15.2) and (15.4)], which implies the assumption that learning-by-doing is absent. It must be noted that the proof of Proposition 15.4 is facilitated by any moderate intertemporal externality: that is, $0 < E(0), R(0,.) < \max \Delta \text{avc}_1$. Then condition (iii) can be replaced by $R(0,.) > M(0,.) = 0$, which is satisfied by assumption.¹

Third, in practice a financial constraint often bounds the set of feasible R&D expenditures, so that the solution is, *ceteris paribus*, determined by the binding financial restriction (F_{sc_0}) if $F_{sc_0} < {}^*sc_0$.

As in the existing contestability models, (the set of) market structure(s) follows endogenously from demand, cost and innovation properties. That is, market demand $D(p_1)$ determines the number of incumbent firms (Baumol *et al.* 1982a): $[D(p_1)/q_1] \leq n_i \leq [D(p_1)/q_1] + 1$, where n_i denotes the number of incumbent firms in equilibrium and where D is a demand function with the usual properties. Once it is realized that in equilibrium all firms choose the same strategy sc_0 associated with the unique cost minimizing output level q_1 , the proof follows (provided that firms are assumed to be identical).² Note that the integer problem (Chapters 6 and 13) can arise.

In the case of a (demand or natural) monopoly ($n_i = 1$), equation (15.6) is not easily solved, because price enters recursively. A monopoly faces a demand constraint in the sense that equation (13.2) transforms into $mu_1 = sc_0/[D(p_1)]$. Price is not only the variable to decide on, but enters demand and so supply and markup as well. So, $q_1 = D(p_1)$ is also to be determined. If demand shows the usual downward slope ($\partial D/\partial p_1 < 0$), then it follows directly from equation (13.2) that $\partial M/\partial p_1 > 0$. This means that $M(sc_0, q_1)$ in Figure 15.2 is convex rather than linear in Δp_1 for the interval of investment outlays associated with a price reduction ($0 < sc_0 < {}^*sc_0$ in Figure 15.2).³ The larger the price reduction is, the lower the markup turns out to be. The monopolist is now able to determine the optimal (${}^*sc, {}^*p, {}^*q$) combination.

15.5 CONCLUSION

The barrier market model describes incumbent firms which adopt pre-emptive sunk innovation in order to reduce the likelihood of future exit. The basic departure from existing contestability models is that (i) entry-deterring *sunk R&D outlays* are *endogenized* and (ii) it is argued that contestability can be guaranteed even if sunk (exit) cost is *positive*. The threat of "hit-and-run" entry still forces firms to deploy careful price setting. On the one hand, firms are forced to adopt careful average total cost pricing, but,

¹ This is not to say that learning-by-doing is incompatible with the barrier market theory. If it is assumed that incumbent firms and potential entrants benefit symmetrically from these intertemporal externalities, the compatibility is clear. Here it is simply argued that excessive learning-by-doing reduces the incentive to spend sunk resources on R&D. It is worth mentioning that Dasgupta and Stiglitz (1988b) point out that learning-by-doing involves sunk costs (so that it may erect entry barriers) and is, therefore, relevant in the debate about the contestability theory.

² At this point it should be noted that unsustainability occurs when $q_1 < {}^Lq_1$ or $q_1 > {}^Rq_1$. Here these complications are ignored (for further details see Ten Raa 1980 and Baumol *et al.* 1982a). In effect, if sunk costs are allowed to occur, the flat-bottomed average total cost function transforms into a U-shaped one (with a unique minimum, i.e. ${}^Lq_1 = {}^Rq_1$), which reduces the likelihood of sustainability (Chapter 13). Proposition 15.1 shows that each level of sc_0 gives a unique cost-minimizing output level q_1 .

³ Of course, for $sc_0 > {}^*sc_0$ $M(sc_0, q_1)$ is concave in Δp_1 , simply because now $\Delta p_1 > 0$.

on the other, they are also inclined to introduce R&D outlays. In the special case of process innovations, firms' behavior in a barrier market brings about, *ceteris paribus*, (partial) Pareto improvements, because the entry threat entry induces minimization of feasible future prices.

PRODUCT INNOVATION: Maximization of Buyer's Utility

16.1 INTRODUCTION: Increased Quality

This chapter intends to develop a (perfect) equilibrium concept for quality innovation in a barrier market setting, which implies that the argument resembles that of Kim (1987).¹ Kim (1987) introduces endogenously chosen heterogeneous products in a contestable monopoly model by allowing for quality differences (compare with Rashid 1988). A natural monopolist supplies a quality scope. This chapter departs from Kim's model by focusing on firms' decisions on *single-product, incremental* quality improvements in a *ni-market configuration* ($n_i \geq 1$). In line with Kim, it will be shown that a strategy of maximizing consumers' utility, apart from zero-profit pricing (Proposition 14.2), is a necessary condition for sustainability of a market configuration. However, a *ni-market configuration* introduces additional sustainability requirements, since both of the preceding conditions fail to be sufficient (Van Witteloostuijn 1990c).

In effect, two arguments are outlined: (i) incumbent firms are forced to introduce product improvements that maximize buyers' utility (Proposition 16.1); and (ii) product innovation increases the likelihood of market unsustainability (Proposition 16.2). The argument is framed in terms of the two-period - two-market benchmark model of Section 10.3. The chapter is organized as follows. Section 16.2 shows that firms seek to undertake product innovations that maximize consumers' utility. Section 16.3 examines the impact of product innovation on the likelihood of (un)sustainability. Section 16.4 offers concluding remarks.

16.2 MAXIMIZATION OF BUYERS' UTILITY

16.2.1 Two-Stage Game with Product Innovation

Incumbent firms and potential entrants play a two-stage game on innovation and price. In the first stage ($t = 0$) firms decide on product innovation, which can be effectuated in the

¹ In this chapter only selected references are made to the literature on product innovation (and differentiation). Examples of models offered in the last two decades are Fisher and Shell (1971), Rosen (1974), Eaton and Lipsey (1975), Ohta (1975), Spence (1976), Dixit and Stiglitz (1977), Kihlstrom and Levhari (1977), Schmalensee (1978), Gabszewicz and Thisse (1980 and 1986), Hart (1980), Shaked and Sutton (1982 and 1983), Shapiro (1982 and 1983), Capozza and Van Order (1982), Brander and Eaton (1984), Sutton (1986), Hollander (1987), Bonanno (1987), Jones (1987), Raubitschek (1987) and Miller (1988). Eaton and Lipsey (1989) offer a survey of the product differentiation literature.

second stage ($t = 1$). In the second stage firms compete over prices. In accordance with theoretical and empirical research (Kamien and Schwartz 1982, pp. 64-70 and 194) innovative activity is assumed to yield decreasing returns. That is,

$$(16.1) \quad v_1 = V(sc_0),$$

where v denotes product quality, $\partial V/\partial sc_0 > 0$ and $\partial^2 V/\partial sc_0^2 < 0$. Sunk investment in period $t = 0$ raises price [equation (14.1)] and quality [equation (16.1)] in period $t = 1$. Following the introduction of an innovation, profit accrues to a firm in period $t = 1$ [payoff matrix (15.8)].

Incumbents show a first-mover advantage in the sense that entrants can only outperform incumbents if they offer a strictly lower price or better product. For the game on product innovation this means that firms compete over price-quality combinations. In order to illustrate the incumbent firms' first-mover advantage, the four-stage structure of Section 10.3 is adopted. So, provided that incumbent firms and potential entrants offer similar quality, for incumbent firms this implies that $i_q = i_d > 0$ if $i_p \leq e_p$ and $i_q = i_d = 0$ if $i_p > e_p$, whereas (potential) entrants face $e_q = e_d > 0$ if $i_p > e_p$ and $e_q = e_d = 0$ if $i_p \leq e_p$.

Let p_1 be the price that yields the revenues just sufficient to reward the investment $i_{sc_0} = e_{sc_0} = sc_0$ ($\pi_1 = 0$): that is, p_1 is the average total cost price. The perfect contestability assumption of firms' homogeneity is retained. Both incumbent firms and potential entrants have access to the same production and investment (product innovation) technology: $i_{avc_1} = e_{avc_1} = avc_1$ [or, to be precise, both groups of firms face the same production function (15.1)] and $i_{v_1} = e_{v_1} = v_1$. The second-stage game on prices proceeds along the lines of perfect contestability theory (although with zero sunk cost) predicts (Knieps and Vogelsang 1982 and Brock 1983). The pricing decision follows the markup rule that prices are determined by average nonsunk cost plus a zero-profit markup [Proposition 14.2, equation (14.1) and (13.2)]. The investment involves sunk R&D in order to increase the quality of the product.

16.2.2 Demand for Quality

In a contestable environment, the introduction of product variety can be adopted as a competitive weapon, either to deter entry or to enter (Kim 1987, Cairns and Mahabir 1988 and Rashid 1988). In the literature product differentiation is often simply restricted to a unidimensional quality continuum (Kihlstrom and Levhari 1977). The utility of the buyers depends on the quantities of m other commodities ($q^1, \dots, q^m$) and the quality (v) and quantity of the product involved (q): $u^j = U(q^1, \dots, q^m, q, v)$, where u^j denotes utility of buyer j and $\partial U/\partial v > 0$. For reasons of simplicity, buyers' homogeneity is assumed.¹ That is to say, all buyers are characterized by similar utility and demand functions and budget constraints. A representative buyer yields

$$(16.2) \quad u_t = U(q^1, \dots, q^m, q_t, v_t), \text{ and}$$

¹ This implies that this chapter describes a setting with vertical rather than horizontal product innovation (differentiation). Horizontal, contrary to vertical, product differentiation is based upon an assumption of heterogeneous preferences (Lancaster 1979 and Gabszewicz and Thisse 1986). That is, there is buyers' heterogeneity.

$$(16.3) \quad \left(\sum_{i=1}^m p_t^i \cdot q_t^i \right) + p_t \cdot q_t \leq b_t,$$

which gives

$$(16.4) \quad D_t = D(p_t^1, \dots, p_t^m, b_t, p_t, v_t),$$

where b represents budget constraint and D market demand for the product involved. The utility and demand function are subject to the usual assumptions of continuity, convexity and signs of the first-order derivatives. The composition of alternative spending opportunities $(q_t^1, \dots, q_t^m, p_t^1, \dots, p_t^m)$ and the budget constraint (b) are assumed to be given and constant for $t = 0, 1$.

Suppliers in a market with product innovation have two competitive weapons: price and quality. The free entry assumption is in line with many contributions to the product differentiation literature, from Hotelling (1929) and Chamberlin (1933) to Lancaster (1966 and 1979) and Gabszewicz and Thisse (1980). To reach a position of exit avoidance and entry deterrence an incumbent firm seeks to introduce a product which offers maximum utility to the buyers. Rashid (1988) suggests that

[i]n the long run the only way to stay is by pleasing customers; this requires providing them with the goods they really want, and this long-term dependence of producers upon customers is perhaps the most effective guarantee of quality (Rashid 1988, p. 248).

In the simple two-period context this result is easily illustrated.

16.2.3 Product Innovation

Incumbent firms and potential entrants play a first-stage subgame on product innovation. Let *u_1 represent the maximum utility of the representative buyer (through a p_1 - v_1 combination) which is feasible in period $t = 1$, where *u_1 follows from optimal sunk R&D outlays *sc_0 in period $t = 0$. That is,

$$(16.5) \quad ^*u_1 = \max_{q_1, v_1} u_1 = U(., q_1, v_1),$$

subject to the feasibility condition that $\pi_1 = 0$ (the binding nature of which follows from the second-stage subgame; Proposition 14.2). Using Lagrange multiplier transforms the maximization problem (16.5) into

$$(16.6) \quad ^*u_1 = \max_{q_1, v_1, \sigma} L_1 = U(., q_1, v_1) - \sigma \cdot (\pi_1 - 0)$$

Hence, $\partial L_1 / \partial q_1 = 0$, $\partial L_1 / \partial v_1 = 0$ and $\partial L_1 / \partial \sigma = 0$. Moreover, the matrix of second-order derivatives should be negative semidefinite.

Kim (1987) and Van Witteloostuijn and Van Wegberg (1988) solve equation (16.6) for R&D technology (16.1) and specified linear forms of the utility function, which gives an optimal sunk R&D outlay *sc_0 , average total cost price *p_1 and quality *v_1 . To be precise, Kim (1987) uses the utility function $u = v \cdot q - p$, whereas Van Witteloostuijn and Van Wegberg (1988) adopt the simple repackaging model (Deaton and Muellbauer

1980, pp. 260-263). The key point is that the assumption of consumers' homogeneity implies that all buyers select the same quality. For example, the simple repackaging model with buyers' homogeneity predicts that the representative consumer will prefer the product variety which offers the minimum p_1/v_1 ratio.¹

Incumbent firms have two regions of relevant strategies: $^i u_1 < ^* u_1$ and $^i u_1 = ^* u_1$ (where $^i u_1 < ^* u_1$ follows from $sc_0 \neq ^* sc_0$ and $^* u_1$ is associated with $^* sc_0$). A potential entrant's set of replies contains two relevant regions: $^e u_1 > ^i u_1$ and non-entry. Actions yielding $^e u_1 \leq ^i u_1$ are not worthwhile, because then the potential entrant is unable to outperform the incumbent. If an incumbent firm fails to offer a product which yields $^* u_1$ to buyers, then a potential entrant is able to push aside the incumbent firm by supplying a product which offers $^i u_1 < ^e u_1 < ^* u_1$ yielding $^e \pi_1 > 0$, which follows from the observation that $^i u_1 < ^e u_1$ permits the potential entrant to increase price above average total cost up to the point where the entrant's price-quality combination is just preferred to the incumbent's. If $^i u_1 < ^e u_1$, all buyers immediately switch from the incumbent firm to the potential entrant, given the assumptions of Bertrand competition, perfect barrier market and buyers' homogeneity. When the incumbent firm adopts $^i u_1 = ^* u_1$, a potential entrant's reply $^e u_1 > ^i u_1$ is not possible. The strategy combination $^i u_1 = ^* u_1$ and non-entry is a *unique* equilibrium as the second-stage equilibrium follows from offering a maximal utility product (associated with the first stage investment $^i sc = ^* sc$) against the average total cost price (Proposition 14.2).

In line with Kim's (1987) result it appears that entry deterrence (or exit avoidance) requires zero-profit pricing in combination with a strategy of investments in product quality so as to maximize consumers' utility. This means that behavior in the market yields *second-best optimality*. Second-best optimality implies that buyers' utility is maximized, subject to the financial viability of the firms (Baumol *et al.* 1982a, p. 6). By adopting $^i u_1 = ^* u_1$ the incumbent firm takes advantage of being the first mover. To uphold the pervasive threat of entry it is required that incumbent firms assume that there are potential entrants around which are able to leapfrog over an incumbent firm's innovation if the incumbent deviates from the optimal product innovation (Fudenberg *et al.* 1983; see also Chapter 8 and 11).

PROPOSITION 16.1 (*maximization buyers' utility*). *In a barrier market incumbent firms are forced to maximize buyers' utility in order to deter entry. This behavior constitutes a unique perfect equilibrium strategy.*

16.3 PRODUCT INNOVATION AND (UN)SUSTAINABILITY

16.3.1 Efficient Scale

Maximization of buyers' utility is, however, not sufficient for entry deterrence or exit avoidance. The point is that unsustainability of market configurations can interfere. Baumol *et al.* (1982a) illustrate the contestability framework with the help of the special case of a flat-bottomed U-shaped average nonsunk (*i.e.*, variable and fixed) cost curve. The constant returns interval is of crucial importance in providing (temporary) sustainabi-

¹ Only the accidental case where a number of varieties share the minimum p_1/v_1 ratio, gives more than one quality in equilibrium. For the current argument the (non)uniqueness feature of the optimal product innovation does not matter.

lity of contestable markets (Ten Raa 1980 and Baumol *et al.* 1982a; see also Chapters 6 and 13). A flat-bottomed average nonsunk cost curve is associated with switching of type of returns [equation (15.1)].

The switching returns property of the average nonsunk cost curve can result, *ceteris paribus*, in exit or entry as soon as product quality improvement is introduced. This proposition is related to the (un)sustainability of market configurations. This result is even enhanced if the average cost function is of a usual U-shape with a unique minimum, which is the case when sunk cost is positive (Chapter 13). By definition sunk cost introduces increasing returns in production. Unlike a natural monopolist (Kim 1987) a firm is now *not* facing nondecreasing returns over the whole range of quantities offered.

In combination with production technology (15.1) a positive sunk cost implies that product innovation introduces a unique cost-minimizing scale (*q) in period $t = 1$, where $^*q_1 > ^*q_0$: that is, period $t = 1$'s efficient scale exceeds period $t = 0$'s (Proposition 13.1). Contestability predicts that entry deterrence (or, to be precise, exit avoidance) requires the adoption of *q_1 and associated minimum average total cost price *p_1 . It is only by accident that $n_{i0} \cdot ^*q_1 = D(\cdot, ^*p_1, v_1)$, where n_{i0} is the number of incumbent firms in period $t = 0$. This follows not only from increased efficient scale (*scale effect*), but also from the demand change that follows upon the introduction of the product innovation (*demand effect*).

16.3.2 Demand Change

Recall that the perfect contestability assumption of firms' homogeneity is retained: both incumbent firms and potential entrants have access to the same production and R&D technology. That is, the average nonsunk cost curve is assumed to be equal for all incumbent firms and potential entrants. Moreover, all incumbent firms and potential entrants offer the same product quality. The key point is that product innovation exerts influence on demand through both price and quality [equation (16.4)]. Demand shifts can introduce intertemporal unsustainability. *Ceteris paribus*, two cases can be distinguished as to the direction of demand change.

First, the quality improvement can be associated with decreasing demand: $\partial D/\partial v < 0$. The reason can be that the same utility level can be reached with less quantity. For example, if the product innovation implies that washing can be done with less detergent, satisfaction of the washing need in period $t = 1$ can be reached by making use of a quantity of washing powder below period $t = 0$'s level. Second, the product innovation can induce increasing demand: $\partial D/\partial v > 0$. This is the argument that increased quality can bring about a re-allocation of money resources (within the budget bound) from other commodities to the new product variety. For example, the innovation of record to compact disc can spur demand in period $t = 1$ for sound recording mediums.

The scale and demand effect together indicate two potential sources of intertemporal unsustainability. The key argument is that the product innovation can increase efficient scale and change demand such that either $n_{i0} \cdot ^*q_1 > D(\cdot, ^*p_1, v_1)$ or $n_{i0} \cdot ^*q_1 < D(\cdot, ^*p_1, v_1)$, which indicates unsustainability. The first scenario emerges if product innovation goes hand in hand with either decreased demand or increased demand which cannot offset the scale effect. The second case occurs if demand is increased such that the scale effect is dominated.

16.3.3 Exit

Suppose that the product innovation introduces $ni_0 \cdot q_1 > D(\cdot, p_1, v_1)$. If $ni_0 = 1$, then a potential entrant is unable to avoid the incumbent firm's diseconomies of scale without suffering losses. This market configuration is a monopoly (Baumol *et al.* 1982a). For $ni_0 > 1$ the prevailing market configuration is unsustainable, because the quality improvement induces (probably negative) market growth insufficient for all incumbent firms to produce and sell at efficient scale. Then, either some incumbent firms exit or all incumbent firms are outperformed by a smaller number of potential entrants. As it stands, the theory is unable to predict which alternative adjustment mechanism occurs (Schwartz and Reynolds 1983). Moreover, if some incumbent firms exit, it is unclear which incumbent firms will be the victim of unsustainability (Chapter 6). None of the incumbent firms can be indicated as the candidate that has to exit on any *a priori* argument, since all firms are alike.

Nevertheless, equilibrium is reached when $D(\cdot, p_1, v_1) / ni_1 = q_1$, where $ni_1 < ni_0$. It is here where the implications of the flat bottom of the average cost curve come in. The existence of an equilibrium market configuration is promoted by an increasing length of the flat bottom of the average cost curve (Baumol *et al.* 1982a, pp. 36-40). If R&D sunk outlays remove the flat bottom from the scenery, the probability of temporary unsustainability (in the sense of a noninteger equilibrium number of firms) is increased. Market equilibrium can be retained if, for example, a niche incumbent or entrant is allowed to serve contingent demand at an average cost price (so as to deter further entry) in excess of p_1 .

16.3.4 Entry

Suppose that period $t = 0$'s investment induces $ni_0 \cdot q_1 < D(\cdot, p_1, v_1)$. The prevailing market configuration is unsustainable, because entry can be undertaken so as to satisfy contingent demand (DR). This holds both for monopoly and nonmonopoly configurations. For illustrative purposes, a specific entry mechanism is examined. The supply of period $t = 0$'s incumbent firms amounts to

$$(16.7) \quad iQ_1 = ni_0 \cdot q_1.$$

Incumbent firms now leave residual demand DR unfulfilled:

$$(16.8) \quad DR = D(\cdot, p_1, v_1) - iQ_1.$$

Potential entrants are able to enter profitably by absorbing residual demand.¹ The entry process can easily be illustrated when sequential entry is assumed. An entrant offers q_1 . Hence, post-entry residual demand is

$$(16.9) \quad DR = DR - q_1.$$

¹ This assumes a rationing scheme which permits profitable niche entry production (Chapter 21). This chapter ignores the specifics of the rationing scheme.

with,

$$(16.10) \quad {}^eQ_1 = \sum_{e=1}^f {}^e q_1,$$

where f denotes the number of entrants. A potential entrant's decision making proceeds as follows: if $DR \geq {}^*q_1$, then ${}^e q_1 = {}^*q_1$; otherwise ${}^e q_1 = DR$. After successful entry $f = f + 1$. If $DR = 0$, then the profitable entry opportunities are exhausted; otherwise history is repeated. It should be noted that the first entrant facing the entry decision is confronted with ${}^e Q_1 = 0$ [equation (14.9)]. A concluding remark is in order with regard to the supply of the last entrant (superscript S). It may be that ${}^S q_1 < {}^*q_1$. Then, the last entrant's price exceeds the one of the incumbent firms and preceding entrants.

16.3.5 Increased Probability of Unsustainability

The analysis of the entry or exit adjustment mechanism shows that market structure is endogenized. Given demand, cost and quality properties, an equilibrium in the number and size of incumbent firms results. However, sunk product innovation increases the likelihood of both temporary and intertemporal unsustainability. Three effects explain this result. First, positive sunk cost introduces an increased and unique efficient scale. Second, improved quality exerts influence on demand. This means that a further sustainability condition is that $ni_0 \cdot {}^*q_1 = D(\cdot, {}^*p_1, {}^*v_1)$, which implies that market growth (so $\partial D/\partial v$ has to be positive) *precisely* outbalances the effect of increased efficient scale. The scale and demand effect together can induce intertemporal unsustainability. Third, a positive sunk cost introduces a unique cost-minimizing scale of production, which increases the probability of temporary unsustainability. An integer number of firms in equilibrium can be supported by permitting a niche incumbent or entrant to satisfy contingent demand at an average cost price in excess of *p_1 .

PROPOSITION 16.2 (unsustainability). *Introduction of sunk product innovation in a barrier market implies that the likelihood of both temporary (scale effect) and intertemporal (scale and demand effect) unsustainability is increased.*

16.4 CONCLUSION

The concept of the barrier market seeks to offer a benchmark case of competition by introducing entry-detering sunk investment in a free entry framework. Although the models are based upon strict assumptions, they provide a framework to study the implications of deviating assumptions on the static and dynamic aspects of performance. It remains to be seen under what relaxed assumptions the results can be sustained. For example, the study of a framework with consumers' heterogeneity (so that competing product varieties can co-exist) deserves a high degree of priority.

Except for monopoly, perfectly contestable market configurations are associated with Pareto-optimal price setting, because incumbent firms are forced to adopt minimum average (and marginal) cost pricing in order to deter entry. In the case of a monopoly the second-best solution prevails, given monopolist's average cost pricing. In the single-product case, average cost pricing implies second-best optimality. However, existing contestable market models deal with the case where sunk investment (or, to be precise, sunk exit cost) cannot occur.

It is in this context that barrier market theory is relevant. This framework is based upon the introduction of sunk investment in a free entry framework. In this way, dynamic economies of market behavior are introduced. The investment cost has to be rewarded by means of a markup on average variable cost. Second-best optimality is guaranteed, because incumbent firms are forced to introduce a markup such that the cost of investment is precisely rewarded. The introduction of higher product quality reflects a Pareto-improvement, because incumbent firms are forced to maximize expected buyers' utility in order to avoid exit.

PART

VIII

LABOR MARKET

Static-Dynamic Efficiency Tradeoff, Wage Rate and Human Capital Improvement

KEY WORDS. *The static-dynamic efficiency tradeoff can also be applied to the labor market. Wage and nonwage (human capital) competition can be framed in terms of perfect contestability and entry barriers (Chapter 17). The analogy with the commodity market appears to be perfect. As in industrial organization, the barrier market theory can be a benchmark case of (intermediate) competition in labor economics (Chapter 18). The competition between incumbent and potential workers is explored in the two-period - two-market model of Section 10.3. Workers demand a wage rate and can undertake sunk human capital investment.*

CHAPTER

17

CONTESTABLE MARKETS, ENTRY BARRIERS AND LABOR SUPPLY

17.1 INTRODUCTION: Industrial Organization and Labor Supply

Up to this point the focus of the argument mainly rested on the commodity market. The labor market has only been mentioned in passing. However, many of the topics mentioned above can be applied to the supply side of the labor market as well (Van Witteloostuijn 1989d). In the labor market, the tradeoff between static and dynamic efficiency is associated with the distinction between wage and nonwage (rather than price and nonprice) competition. Then, static efficiency is expressed in wage rates that yield zero net returns to investment in human capital¹ (rather than prices that are equal to marginal cost), whereas dynamic efficiency can reveal itself in improvements of human capital (rather than innovations).

Thurow (1975) points to the tradeoff between static and dynamic efficiency in the labor market by arguing that

the types of wage and employment competition that are the essence of efficiency in simple, static neoclassical models may not be the essence of efficiency in a dynamic economy where the primary function of the labour market is to allocate individuals to on-the-job training ladders and where most learning occurs in work-related contexts (Thurow 1975, p. 84).²

The analogy with the commodity market seems to be perfect. The present chapter explores this parallel for the special case of human capital investment. The argument is restricted to the two concepts of competition that are most relevant in the context of the

¹ Two remarks are in order. First, the well-known marginal productivity theory of distribution predicts that in a perfectly competitive market labor is paid the value of the marginal product (for example, Koutsoyiannis 1979 and Varian 1984). However, here the demand side of the labor market (firms) plays a critical role. This chapter views the supply side of the labor market in isolation. Section 17.3 shows the limit tendency of the wage rate to fall down to the reservation level if there is strong competition among suppliers of labor. Second, distribution arguments may plead in favor of wage rates in excess of the reservation level. Here the aspect of distribution is ignored. (For further discussion see Chapter 24.) The analogy with the commodity market implies that the surplus of the demand side of the market is maximized subject to the constraint that the other side of the market is still willing to supply. In the commodity market this means that the consumers' surplus is maximized under the condition that the firms' surplus is zero.

² Creedy and Whitfield (1988) argue "that static Pareto efficiency is rather narrow and ... other features of organizations are more important for the development of a high productivity labour market" (Creedy and Whitfield 1988, p. 253).

barrier market framework: contestable markets and entry barrier competition.

Of course, this is not to say that the argument presented below is entirely new. Indeed, at many points the interface between industrial organization and labor economics has been analyzed extensively. For example, in the context of the theory of labor unions (Hirsch and Addison 1986), the economics of the internal labor market (Creedy and Whitfield 1988) and the insider - outsider theory (Lindbeck and Snower 1987) the focus is on the implications of imperfect competition in the labor market (*i.e.*, for the wage rate and employment).

The plan of the chapter is as follows. Section 17.2 briefly reviews the essentials of the economics of human capital and applies the two-period - two-market benchmark model of Section 10.3 to the labor market. Section 17.3 develops the theory of perfect contestability for the supply side of the labor market. Section 17.4 illustrates the relevance of barriers to entry in the labor market. Section 17.5 describes the implications of barriers to entry to the wage rate. Section 17.6 offers some final remarks. The argument shows that perfect contestability facilitates static efficiency (Proposition 17.1), whereas entry barrier competition is associated with incentives to undertake human capital investment to the detriment of static efficiency (Proposition 17.2).

17.2 WAGE RATE AND HUMAN CAPITAL INVESTMENT

17.2.1 The Economics of Human Capital

Although Smith (1776, Chapter 10) already made mention of the influence of training and education on labor earnings, the birth of the economics of human capital dates from the beginning of the 1960s with the seminal contributions of Schultz (1961 and 1963) and Becker (1962 and 1964). The theory of human capital is the result of drawing a parallel between the investment theory of the firm and an investment notion of labor suppliers. Becker (1964) introduces the essentials of the economics of human capital by arguing that

[t]his study is concerned with activities that influence future monetary and psychic income by increasing the resources in people. These activities are called investments in human capital (Becker 1964, p. 9).

In this subsection the main argument underlying the economics of human capital is briefly outlined. Useful surveys are offered by Blaug (1976) and Siebert (1985). Individual decisions on labor supply contain a quantitative and a qualitative aspect. The former is concerned with short-run supply (Killingsworth 1983) and search behavior (Lippman and McCall 1976), *given* the quality of the labor services offered, whereas the latter is devoted to the *manipulation* of the quality of the labor services. The qualitative aspect, unlike the quantitative one, mainly has to do with long-run decision making.

The economics of human capital precisely deals with suppliers' decisions on *investment* in the quality of labor services.¹ In this respect, Becker (1964) distinguishes specific and general investments. Specific investments, unlike general ones, are concerned with firm-specific skills and knowledge. A specific investment in human capital is only productive in *one* firm, whereas *several* firms are able to benefit from a general

¹ The notion of investment behavior in the human capital theory applies to many aspects of household decision making, for example allocation of time (Becker 1965), health care (Cooper and Cuyler 1973) and religion (Neuman 1986).

investment. A firm can come up with a purely specific investment, whereas a purely general investment must be paid by the (future) laborer. In general, however, a mixture of specific and general investments obtains (Becker 1964, p. 26).¹

Investments in human capital are manifest in many ways. Becker (1964) points, for instance, to schooling, re-education, on-the-job-training and learning-by-doing. Investments in human capital occur before and after entry into the market. Pre-entry schooling concerns a general investment in the form of a choice of occupation (for example, O'Connell 1972 and Rizzutto and Wachtel 1980). Post-entry schooling and training may take the form of an occupational switch or the updating of prevailing human capital. Pre-entry investments are of particular importance in explaining the life cycle pattern of wage earnings (for example, Becker 1964, Ben-Porath 1967 and Mincer 1974).

The economics of human capital applies the neoclassical maximizing postulate to laborers' investment by drawing an analogy with the investment theory of the firm (for example, Nickell 1978). The decision maker maximizes the expected net present value of the return to investment (for example, Addison and Siebert 1979). The investment in human capital is generally subject to (time, budget and ability) restrictions.² The benefits from an investment in human capital derive from increased wage and psychic income. The wage income from an investment in human capital is facilitated by an increased value of the marginal product. The costs of an investment follow from foregone earnings and direct investment costs. Foregone earnings reflect opportunity costs. The time involved with the investment defeats the opportunity to earn (wage and psychic) income. The direct costs of investment simply point to the expenditures that are necessary to reach a change in the stock of human capital.

A major challenge to human capital theory is the explanation of the distribution of earnings. This explanation is offered by introducing the additional assumption that the rates of return are equal at the margin (Becker 1964 and Thurow 1970). This is to say that perfect competition rules the markets in the economy such that allocation takes care of reaching the marginal equalities. However, what is the influence of competition in a labor market on the rates of returns to investment? Moreover, in what way does competition in a labor market generate incentives to introduce investments in human capital? It is here where industrial organization comes in.

17.2.2 An Illustrative Two-Period Model

For expository purposes, the argument proceeds along the lines of a simple two-stage game (Section 10.3). The model focuses on two essential variables, wage rate and human capital investment, in a two-period - two-market framework. The (schooling) cost incurred in period $t = 0$ yields an improvement of human capital in period $t = 1$. As a benchmark case assume that in period $t = 0$ workers either receive the reservation wage rate or are unemployed. Incumbent workers are employed in market $m = i$, whereas potential workers either have a job in market $m = e$ or are unemployed.³ The scale of the human capital investment in period $t = 0$ and the level of the wage rate in period $t =$

¹ The employee and employer then have to reach an agreement on sharing the costs and benefits of the investment. As a result, a number of additional arguments enter the picture, including the likelihood of turnover (Bartel and Borjas 1978) and bargaining (Hashimoto 1979).

² Ability constraints introduce a "nature-nurture" debate that questions the relative importance of the influence of the environment and inheritance (see, for example, Taubman 1976).

³ Section 17.3 indicates that benchmark competition in period $t = 0$ resembles contestability.

1 depend on the assumption of supply competition. The investment decision of a worker is described as if there were no uncertainty. Supply competition derives from the rivalry between incumbent and potential workers. The game is one which has complete and imperfect information.

It is likely that some minimum size of wage income exists at which workers are just willing to work. A further decrease of income (due to a further fall of the wage rate) leads to a predominance of the utility of leisure. Equivalently, there is presumably some maximal acceptable wage rate which firms are just willing to pay, simply because a higher wage rate is associated with a negative payoff. Here it is plausible to think that there can be a minimal or maximal acceptable or reservation wage rate (referred to as $\min w_t$ and $\max w_t$ respectively). A workers' reservation wage and firms' reservation marginal product (associated with a reservation wage rate) are well-known in search theory and job matching models (Lippman and McCall 1976 and Pissarides 1988). It is important to note that a worker prefers to be unemployed, rather than employed for $w_t < \min w_t$. The values of the reservation wage rates in period $t = 0$ are given: $\min w_0$ and $\max w_0$.

Investment in human capital is modeled along the lines of the argument in Subsection 17.2.1. The wage income from an investment in human capital is facilitated by an increased value of the marginal product. The prevailing value of the marginal product (mp) in a period depends on the stock of human capital (hc):

$$(17.1) \quad mp_t = MP(hc_t),$$

$$(17.1a) \quad \partial MP / \partial hc_t > 0,$$

where MP is twice continuously differentiable and $t = 0, 1$. The value of the marginal product is defined in terms of the product which the employers offer in the commodity market. For the sake of convenience, this chapter assumes single-product firms which supply a uniform product in the commodity market. The human capital stock in period $t = 0$ is the same for all workers. In the face of potential rivalry the wage rate demand (w) of a worker i depends on his/her marginal product (dis)advantage relative to rival j (Section 17.5): $\bar{w}_{impd} = \bar{w}_{mp} - \bar{w}_{jmp}$. So,

$$(17.2) \quad \bar{w}_t = W(\bar{w}_{impd}),$$

$$(17.2a) \quad \partial W / \partial \bar{w}_{impd} > 0,$$

where W is twice continuously differentiable with $t = 0, 1$ and $mp_0 = \bar{w}_{mp_0} = \bar{w}_{jmp_0}$ is given. (A further assumption is that $\max w_0 = mp_0 > \min w_0$.) Equations (17.1) and (17.2) show that the incremental wage rate in period $t = 1$ (Δw_1) depends, *ceteris paribus*, on the changes in the rivals' stock of human capital (Δhc_1).

The direct cost of investment in period $t = 0$ (ic_0) simply points to the expenditures that are necessary to reach a change in the stock of human capital in period $t = 1$. That is,

$$(17.3) \quad \Delta hc_1 = HC(ic_0), \text{ with}$$

$$(17.3a) \quad \partial HC / \partial ic_0 > 0, \text{ and}$$

$$(17.3b) \quad \partial^2 HC / \partial ic_0^2 < 0,$$

where HC is twice continuously differentiable with $HC(0) = 0$. Equation (17.3b) implies that investment is subject to decreasing returns. The assumption of decreasing returns is generally adopted in human capital models (Killingsworth 1983). Moreover, $\max w_1 = \max w_0 + \Delta mp_1$, where $\max w_0 = mp_0$.

Equations (17.1) to (17.3) define the restrictions that workers have to consider while deciding on wage rate and investment, which is central in the economics of human capital. The payoff (π_t) of employment per period is wage income, being the wage rate times the number of working hours (L_t) minus the (sunk) investment cost. Let u_i be a money measure of the utility of unemployment, where u_i is very close to $\min w_t$ ($u_i = \min w_t - \epsilon$ for $\epsilon \rightarrow 0$). The payoff matrix is normalized by assuming $L_t = 1$ (for $t = 0, 1$) and $\min w_0 = 0$. For the sake of simplicity, the following two assumptions are introduced: (i) discounting, foregone earnings and psychic income are ignored; and (ii) $\min w_t$ and u_i are constant over time (i.e., $\min w_0 = \min w_1 = 0 > u_0 = u_1$). Let sc_0 denote the sunk cost incurred in period $t = 0$ in order to stay or become employed in period $t = 1$. Recall that in period $t = 0$ workers either receive the reservation wage rate or are unemployed. That is,

$$(17.4) \quad \begin{cases} \pi_0 = \min w \text{ if } L_0 = 1, \\ \pi_0 = u_i \text{ if } L_0 = 0, \\ \pi_1 = \min w + \Delta w_1 - sc_0 \text{ if } L_1 = 1, \text{ and} \\ \pi_1 = u_i - sc_0 \text{ if } L_1 = 0. \end{cases}$$

It is important to point to the sunk nature of the human capital investment (Becker 1964). A firm that is closed down before all sunk costs are rewarded incurs a loss. Equivalently, a worker who stops working before the investment in human capital is rewarded faces a loss as well.¹ For the moment, this means that $sc_0 = ic_0$.

In the commodity market static efficiency is affiliated to a zero profit (that is, average total cost pricing). An important (perhaps the most important) manifestation of dynamic efficiency is innovation (Kamien and Schwartz 1982 and Stiglitz 1987). Reasoning by means of analogies implies that static efficiency in the labor market can be assumed to be associated with zero returns to the investment in human capital: $\pi_1 = 0$. Moreover, it is supposed that dynamic (in)efficiency is, *ceteris paribus*, reflected in the attractiveness of workers to employers (which critically depends on the scale of human capital investment).

A worker's attractiveness (pwg) to a firm is determined by the gap between the value of the marginal product and the wage rate:

¹ Conditional upon the assumption that the worker has paid for (part of) the investment. This implies that the human capital is not completely firm-specific, since then the employer has an incentive to pay fully for the investment.

$$(17.5) \quad \begin{cases} \text{pwg}_0 = \text{mp}_0 - \min_w, \text{ and} \\ \text{pwg}_1 = \text{mp}_1 - w_1, \end{cases}$$

where pwg_0 follows from the assumption that $w_0 = \min_w$ for employed workers. (Section 17.3 indicates that this can follow from a state of labor market contestability in the first period.) A worker's attractiveness in period $t = 1$ follows from her/his human capital investment in period $t = 0$ and wage rate in period $t = 1$. That is, a worker's attractiveness to employers depends on her/his own (wage rate and investment) strategies: attractiveness is driven by supply-side rivalry. Proposition 18.3 proves that under particular (plausible) conditions a unique (and positive) optimal investment scale gives a unique maximum attractiveness. For the moment it suffices to suppose that *ic_0 denotes the unique scale of human capital investment that yields a unique maximum attractiveness, $^*\text{pwg}_1$, in period $t = 1$, subject to the feasibility condition that $\Delta w_1 \geq ic_0$. That is, investment cost *ic_0 is assumed to be associated with a unique $^*\text{mp}_1$ and *w_1 that maximize $^*\text{pwg}_1$, where of course $^*w_1 = \min_w + ^*ic_0$ (or $\Delta w_1 = ^*ic_0$: that is, zero net returns: $\pi_1 = 0$). Moreover, $hc_0 < ^*hc_1$, $\text{mp}_0 < ^*\text{mp}_1$, $\min_w < ^*w_1$ and $\text{pwg}_0 < ^*\text{pwg}_1$. The maximum (feasible) attractiveness $^*\text{pwg}_1$ goes, *ceteris paribus*, hand in hand with dynamic efficiency.¹ Subsequent sections examine the implications of an assumption of perfect contestability (Section 17.3) and entry barriers (Section 17.5) for market performance in period $t = 1$.

17.3 CONTESTABILITY AND MINIMUM WAGE RATES

Perfect contestability of the labor market implies that incumbent suppliers of labor anticipate the reactions of credible potential rivals. The assumption of homogeneity means that incumbent and potential have access to the *same* stock of human capital such that all rivals offer the same value of the marginal product. Neither the potential entrants nor the incumbent workers suffer a disadvantage relative to rivals.

DEFINITION 17.1 (*perfect contestability*). A labor market is perfectly contestable if entry is free and exit is costless. All suppliers have access to identical human capital with equal productivity: $^i hc_1 = ^e hc_1$ and $^i mp_1 = ^e mp_1$ for $i = 1, \dots, n_i$ and $e = 1, \dots, n_e$, where n_i and n_e denote the number of incumbent and (equally equipped) potential workers respectively.²

In a contestable labor market incumbents and (potential) entrants play a Bertrand-Nash game on wage rates. Baumol *et al.* (1982a) argue that "entry is not impeded by fear of

¹ Chapter 18 argues that under particular conditions pwg_1 has a unique maximum. It appears that pwg_1 is a function (G) of ic_0 . The result that $\pi_1 = 0$ follows from the first-order condition: $\partial G / \partial ic_0 = 0$ holds if $\partial W / \partial ic_0 = \partial MP / \partial ic_0$. This implies that maximization of future attractiveness requires that the increase of the wage rate is equal to the one of the value of the marginal product. The maximum future attractiveness is associated with zero net returns, because pwg_1 can be increased for any $\Delta w_1 > ic_0$ by lowering the incremental wage rate to $\Delta w_1 = ic_0$ without violating the feasibility condition.

² The assumption of homogeneity is not facilitated when *specific* human capital enters the picture; because specific human capital is likely to be associated with heterogeneity (Section 17.4 and 17.5). The assumption of homogeneous labor supply is only valid in the case of identical stocks of specific human capital.

retaliatory price [wage] alterations" (Baumol *et al.* 1982a, p. 349), so potential entrants employ Bertrand-Nash expectations while judging post-entry profit opportunities. Moreover, incumbents have a first mover advantage (which, for that matter, does not permit capturing a positive payoff) in demanding a wage rate in the sense that entrants are only able to outperform incumbents if ${}^e w_1 < {}^i w_1$. When the entrant's wage rate is equal to the incumbent's, the firm will *not* fire the incumbent laborer so as to hire the entrant supplier.¹ However, if ${}^i w_1 > {}^e w_1$, then the firm replaces the incumbent laborer by the entrant, simply because the lower wage rate increases profit. In order to illustrate the incumbents' first-mover advantage, a three-stage structure is adopted (Chapter 6). In t^1 incumbent laborers announce the wage rates demanded. The potential entrants' [wage and (non-)entry] responses follow in t^2 . The game is closed in t^3 , where firms decide on dismissing incumbent laborers and hiring entrants. The game is one which has complete and perfect information.

As it stands, perfect contestability is a *static* equilibrium theory (Baumol *et al.* 1983). Existing contestability models assume sunk cost to be zero (Schwartz and Reynolds 1983, Shepherd 1984 and Stiglitz 1987): ${}^i sc_1 = {}^e sc_1 = 0$. Here this means that investment in human capital is zero: ${}^i hc_0 = {}^e hc_0 = 0$. Hence, ${}^i \Delta hc_1 = {}^e \Delta hc_1 = 0$ [equation (17.3)] so that ${}^i hc_0 = {}^e hc_1 = {}^e hc_0 = {}^e hc_1$ and ${}^i mp_0 = {}^i mp_1 = {}^e mp_0 = {}^e mp_1$ [equation (17.1)], which implies that, *ceteris paribus*, $\max w_0 = \max w_1$. In existing models of perfect contestability dynamic efficiency improvements are ignored. Together with the observation that ${}^i \Delta mp_1 = 0$ the equilibrium wage rate, ${}^i w_1 = \min w$ (Proposition 17.1), implies that $pwg_0 = pwg_1$ [equation 17.5)].

Incumbent laborers have two regions of relevant strategies (s): ${}^i w_1 > \min w$ (s^1) and ${}^i w_1 = \min w$ (s^2). If ${}^i w < \min w$, then the incumbent prefers not to supply any labor. The potential entrants' set of replies contains two regions of relevant responses: ${}^e w_1 < {}^i w_1$ (s^1) and non-entry (s^2). Actions reflecting ${}^e w \geq {}^i w$ are not worthwhile, because then the entrant cannot outperform incumbent laborers. It can easily be shown that perfect contestability is associated with incumbents demanding the reservation wage rate, provided that three conditions hold.

PROPOSITION 17.1 (demand for reservation wage rate). *In a perfectly contestable market incumbent suppliers are inclined to accept the reservation wage rate: ${}^i w_1 = \min w$ for $i = 1, \dots, n_i$, conditional upon*

- (i) ${}^e \pi_1 = u_i < 0$ or ${}^e \pi_1 = \min w = 0$ for some potential entrants if they refrain from entry (credibility),
- (ii) $u_i < \min w$ (uniqueness), and
- (iii) $\min w \leq \max w$ (feasibility).

Incumbents' demand for $\min w$ constitutes a unique Nash equilibrium strategy.

Proof Proposition 17.1. The intuition is straightforward. If an incumbent laborer adopts ${}^i w_1 > \min w$, then a potential entrant is able to outperform the incumbent by offering $\min w < {}^e w_1 < {}^i w_1$, so that the firm fires the incumbent so as to hire the entrant. Potential entrants are inclined to enter if ${}^e \pi_1 = u_i < 0$ or ${}^e \pi_1 = \min w = 0$ when they refrain from

¹ This assumption is likely to be valid in practice, because hiring and firing of laborers is associated with transaction (for example, contracting) costs.

entry [condition (i)], because entry yields ${}^e\pi_1 > 0$. The incumbent is then fired: ${}^i\pi_1 = u_i < \min w = 0$ [condition (ii)]. If the incumbent worker demands ${}^i w_1 = \min w$ (associated with ${}^i\pi_1 = 0$), then successful entry requires ${}^e w_1 < \min w$. However, a wage rate below $\min w$ is unacceptable, so that the optimal reply is not to enter (Figure 17.1).

Figure 17.1
Game in a Contestable Labor Market

The strategy ${}^i w_1 = \min w$ is associated with a unique Nash equilibrium, because then only the incumbent laborer's nonnegative payoff is guaranteed. Of course, this result is conditional upon $\min w \leq \max w_1$ [condition (iii)]. Q.E.D.

The interpretation of the three conditions is straightforward. Condition (i) assumes that incumbents anticipate the reply of potential entrants with ${}^e\pi_1 \leq 0$, which is necessary to preserve the *credibility* of the pervasive threat of "hit-and-run" entry. If incumbents conversely expect that all potential entrants capture ${}^e\pi_1 = {}^e w_1 > 0$, then demand for ${}^i w_1 = {}^e w_1$ also forestalls entry. Condition (i) means that ${}^e\pi_1 = {}^e w_1 = \min w$ or ${}^e\pi_1 = u_i$, which implies that credible potential entrants either receive the reservation wage rate or are unemployed. On the one hand, the fact that there are some potential entrants receiving $\min w$ indicates the co-existence of other perfectly contestable (or competitive) labor markets. On the other hand, the fear of unemployed potential entrants suggests that the degree of contestability of a labor market probably depends on the rate of unemployment. The key implication of condition (i) resembles the argument in Chapter 11 and 12: (perfect) labor barrier market outcomes can only be obtained if the opportunity cost of entry is zero.

The *uniqueness* of the Nash equilibrium follows from condition (ii): $u_i < \min w$. If incumbents are indifferent between unemployment and working for $\min w$, then two Nash equilibria prevail. Both strategies ${}^i w_1 > \min w$ and ${}^i w_1 = \min w$ are then associated with a zero payoff: ${}^i\pi_1 = u_i = \min w = 0$. The condition $u_i < \min w$ probably follows from the observation that a change of status (i.e., from employed to unemployed) is associated with adjustment costs. Moreover, it is likely that future considerations plead in favor of being employed, because the employed laborers benefit from a superior point of departure for

CURRICULUM VITAE

Arjen (Adriaan) van Witteloostuijn is op 12 oktober 1960 geboren in Haarlem. In 1979 behaalde hij zijn Atheneum-B diploma aan de RSG Epe (Gelderland). Gedurende de periode september 1979 - januari 1985 studeerde hij Economie en Psychologie (vanaf september 1981) aan de Rijksuniversiteit Groningen. Na het behalen van zijn doctoraaldiploma Algemene en Bedrijfseconomie (cum laude) is hij in februari 1985 als Wetenschappelijk Assistent in dienst getreden bij de Economische Faculteit van dezelfde universiteit (zijn studie Psychologie heeft hij medio 1986 voortijdig afgebroken). Het promotie-onderzoek bevond zich op het raakvlak van de micro- en macro-economie. Halverwege het jaar 1986 heeft hij het onderzoek voortgezet aan de Economische Faculteit van de Rijksuniversiteit Limburg (eveneens als Wetenschappelijk Assistent). Van maart tot juni 1988 is hij als gastonderzoeker werkzaam geweest aan de C.V. Starr Center for Applied Economics, Department of Economics van de New York University in de Verenigde Staten. Voor dit bezoek ontving hij een beurs van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek. Na een kortstondige periode als Universitair Docent "Micro-Economie" (van oktober 1988 tot maart 1989) is hij benoemd tot (beoogd) Universitair Hoofddocent "Externe Organisatie en Strategisch Management" bij de Sectie Organisatie van de Vakgroep Bedrijfseconomie binnen Faculteit der Economische Wetenschappen der Rijksuniversiteit Limburg.

achieving promotion (Doeringer and Piore 1971) and avoiding depreciation of human capital. Condition (ii) resembles contestability's Assumption 6.7.

The condition (iii) that $\min w \leq \max w_1$ of course guarantees the *feasibility* of the Nash equilibrium. If $\min w > \max w_1$, then supply and demand will not meet, so that the market equilibrium is empty. A laborer who demands $\min w > \max w_1$ is not hired by a firm, simply because this would add negatively to profits. Condition (iii) holds, since by assumption $\max w_1 = mp_1 = \max w_0 = mp_0 > \min w$.

17.4 BARRIERS TO ENTRY IN THE LABOR MARKET

The relevance of entry barriers to labor market performance is recognized in labor economics. For example, Siebert (1985) points out that Smith (1776) already stressed that

competition can be prevented as when entry to an occupation is restricted, or regional mobility is restricted (Siebert 1985, p. 6).

Smith argues, for instance, that laws hinder perfect mobility of labor, since

[t]he statue of apprenticeship obstructs the free circulation of labour from one employment to another, even in the same place. ... Corporation laws ... give less obstruction to the free circulation of stock from one place to another than to that of labour (Smith 1776, pp. 239-240).

Entry barriers and mobility imperfections are but two sides of the same coin. The segmentation theories of the labor market (Doeringer and Piore 1971 and Creedy Whitfield 1988) depend critically on an assumption of mobility imperfections and, thereby, entry barriers.¹ For example, Malveaux (1984) notes that

dual labor market theory develops the notion of one key barrier: a lack of mobility between two separate labor markets termed primary and secondary (Malveaux 1984, p. 107).

The implication of the entry barrier is straightforward:

The level of inter-firm mobility was deemed insufficient to promote the equalization of net advantages that is central to wage competition (Creedy and Whitfield 1988, pp. 250-251).

In the context of the economics of human capital (potential) entrants' goodwill and reputation *disadvantages* or *asymmetric heterogeneity* of incumbents and entrants as to human capital may operate as entry barriers.² Here two labor theories may illustrate this point: (i) the economics of internal labor markets and (ii) the insider - outsider theory.³ The common denominator in these two theories is the emphasis on the role of specific

¹ There are many segmentation theories (Friedman 1984, Creedy and Whitfield 1988 and Bierings and Van Witteloostuijn 1989). Here it suffices to note the common denominator: a segmentation of the labor market restricts the mobility of labor.

² The human capital approach has also been applied to questions of mobility (for instance, Sjaastad 1962 and Clark 1982).

³ Efficiency wage models (Yellen 1984 and Akerlof and Yellen 1986) show a family resemblance to the insider - outsider theory. There the wage premium follows from imperfect information (monitoring) arguments. For a comparison see Lindbeck and Snower (1987).

human capital in introducing asymmetric heterogeneity (and so monopoly power).

Doeringer and Piore (1971) are the founding fathers of the economics of the internal labor markets (Creedy and Whitfield 1988). They,

explained low mobility by the efforts of employers to reduce turnover because of the presence of skill-specificity; that is, the existence of skills which are only useful in a small range of jobs. The consequence of such skills is that firms draw a distinction between incumbents and otherwise similar workers outside the firm. ... The most important rules are those governing entry and promotion. Skill-specificity is seen to promote the restriction to entry into ILMs [internal labor markets] to lower job classifications (Creedy and Whitfield 1988, p. 251).¹

The skill-specificity permits incumbent laborers to benefit from a monopoly position by demanding a premium over their alternative wage (Creedy and Whitfield 1988, pp. 254-255).

The insider - outsider theory (Solow 1985 and Lindbeck and Snower 1987 and 1988) is based upon assuming

(a) an explicit labor turnover cost and (b) the ability of the full-fledged employees ('insiders') to exercise influence over their wages, without taking full account of the interests of the fledgling employees ('entrants') or the unemployed workers ('outsiders'). The insiders' market power arises from the turnover cost, and this power may also be devoted to augmenting that cost. Due to this cost, the insiders are able to raise their wage above the minimal level required to induce workers to become entrants, but firms nevertheless have no incentive to hire outsiders (Lindbeck and Snower 1987, p. 407).

So, incumbents can exploit their advantages over (potential) entrants. A likely source of the monopoly position is specific human capital (Lindbeck and Snower 1987, p. 411).

17.5 IMPERFECT COMPETITION, ENTRY BARRIERS AND POSITIVE RETURNS

Section 17.4 indicates that the apparatus devised by the literature on entry barriers in industrial organization is likely to be applicable to the labor market. The existence of entry barriers implies that potential entrants must incur an incremental entry cost in period $t = 1$ (ec_1) so as to enable successful entry. The entry costs reflect the fact that entry requires additional investment efforts. In the context of the economics of human capital it is plausible to think that there can be skill specificity of incumbent workers in period $t = 1$. This means that, *ceteris paribus*, $^1mp_1 > ^0mp_1$ in market $m = i$. Therefore, potential entrants can only overcome this skill disadvantage by incurring an incremental entry cost.²

DEFINITION 17.2 (*labor entry barriers*). *Barriers to entry derive from a cost disadvantage of potential entrants relative to incumbent laborers in the sense that*

¹ Kerr (1954) already indicated the importance of the distinction between inside (incumbent) and outside (potential) workers.

² The argument in this section resembles the model in Subsection 11.2.2. The line of reasoning can also be framed in terms of productivity differentials rather than incremental entry cost. That is, the entry-detering wage rate follows from the potential entrants' opportunity cost of entry (*i.e.*, the wage rate which potential entrants can capture outside the incumbents' market) plus the incumbency rent ($^1mp_1 - ^0mp_1$).

successful entry requires a burden of additional entry costs: ${}^e ec_1 > 0$ for $e = 1, \dots, ne$.

It is precisely the entrants' cost disadvantage that allows incumbent suppliers to receive positive net returns to investment.¹

Bain (1956, p. 3) defines entry barriers as market conditions which allow incumbent firms to set prices persistently above a competitive level without attracting entry. In terms of investments in human capital this means that ${}^i \pi_1 > 0$. Hence, the condition of static efficiency (i.e., zero net returns) is violated. Suppose that both incumbent workers and potential entrants have undertaken a similar scale of human capital investment in the first-stage period $t = 0$: ${}^i ic_0 = {}^e ic_0 = ic_0$. However, the very definition of entry barriers implies that ${}^i sc_1 = {}^i ic_0 < {}^e sc_1 = ic_0 + {}^e ec_1$, where ${}^e sc_1$ is the sunk cost that an entrant has to incur upon successful entry.

In a market with entry barriers incumbent laborers and potential entrants play a second-stage game on wage rate. Incumbents have a first mover advantage in the sense that entrants are only able to outperform incumbents if ${}^e w_1 < {}^i w_1$. In order to illustrate the incumbents' first-mover advantage, the three-stage structure is sustained. In t^1 laborers introduce the result of period $t = 0$'s investment and incumbents announce the wage rates demanded. The potential entrants' [wage and (non-)entry] responses follow in t^2 . The game is closed in t^3 , where firms decide on whether or not to fire incumbent laborers so as to hire entrants. The game is one with complete and perfect information.

Incumbent suppliers have two regions of relevant strategies (s): ${}^i \Delta w_1 > {}^e sc_1$ (s^1) and $ic_0 \leq {}^i \Delta w_1 \leq {}^e sc_1$ (s^2). Actions showing ${}^i \Delta w_1 < ic_0$ are of course rejected, since they yield a negative payoff. Potential entrants' set of replies contains two regions of relevant responses: ${}^e \Delta w_1 < {}^i \Delta w_1$ (s^1) and non-entry (s^2). Actions implying ${}^e \Delta w_1 \geq {}^i \Delta w_1$ are not worthwhile, because then ${}^e w_1 \geq {}^i w_1$ so that successful entry is not feasible. (Of course, ${}^i w_1 = \min_w + {}^i \Delta w_1$).

PROPOSITION 17.2 (above-normal net returns). *If entry barriers exist, incumbent suppliers are able to capture a net return to investment in human capital which is equal to the potential entrants' entry cost: ${}^i \Delta w_1 = ic_0 + {}^e ec_1 = {}^e sc_1$ so that ${}^i \pi_1 = {}^e ec_1$ for $i = 1, \dots, ni$, conditional upon*

- (i) ${}^e \pi_1 \geq 0$ for all potential rivals if they refrain from entry (credibility); and
- (ii) ${}^i w_1 \leq \max_w$ (feasibility).

Actions yielding ${}^i \pi_1 = {}^e ec_1$ constitute a second-stage Nash equilibrium strategy.

Proof Proposition 17.2. The payoff matrix follows from condition (i) and substitution of the incumbents' and entrants' sunk cost in equation (17.4). Whenever an incumbent supplier demands a wage rate that yields ${}^i \Delta w_1 > {}^e sc_1$, a potential entrant receiving zero net returns if entry is abandoned can enter profitably by demanding a wage rate such that ${}^e sc_1 < {}^e \Delta w_1 < {}^i \Delta w_1$. The incumbent supplier is expelled from the market, because ${}^i w_1 > {}^e w_1$. The incumbent's payoff is negative, because the incumbent forfeits the

¹ The production factor labor can also be used to raise entry barriers in the commodity market (Dewatripont 1987 and 1988). If incumbent firms employ workers with higher productivity (that is, higher stock of human capital) than potential competitors and if these employees are not inclined, *ceteris paribus* (i.e., constant wage rate), to switch employment to entrants, potential rivals face an entry barrier.

opportunity to recover the sunk cost invested (${}^i\pi_1 = -ic_0$). This follows from the fact that the incumbent's earnings drop to zero because of (temporary) unemployment. The entrant's payoff is positive, since ${}^e\Delta w_1 > {}^e\pi_1$. If the incumbent supplier adopts a wage rate that yields ${}^i\pi_1 \leq {}^e\pi_1$, none of the potential entrants is able to enter profitably, because successful entry requires ${}^e\Delta w_1 < {}^i\Delta w_1$ so that ${}^e\pi_1 < 0$. A non-entry strategy yields ${}^e\pi_1 = 0$ [condition (i)]. Backward induction points out that the region of strategies associated with a Nash equilibrium in pure strategies is $ic_0 \leq {}^i\Delta w_1 \leq {}^e\pi_1$ and non-entry, because only then the entrant's optimal reply leads to an incumbent's nonnegative payoff. An extensive-form tree of the game can be constructed (Figure 17.2).

Figure 17.2
Game in a Labor Market with Entry Barriers

The strategy ${}^i\Delta w_1 = {}^e\pi_1$ dominates, simply because it yields the highest ${}^i\pi_1$. However, the associated wage rate is only feasible if ${}^i w_1 \leq \max w_1$ [condition (ii)]; otherwise ${}^i w_1 = \max w_1$. An equilibrium is, probably, nonexistent if $\max w_1$ is so low that ${}^i w_1 = \max w_1$ generates ${}^i\pi_1 < 0$. *Q.E.D.*

Three remarks are in order. First, the assumption that there are some potential entrants waiting behind the scene who will receive ${}^e\pi_1 = 0$ if they refrain from entry, is facilitated by the co-existence of related labor markets. This resembles the home market argument of Chapter 11. For example, the entry threat can derive from employed outsiders who are waiting for an opportunity to enter an internal labor market. Second, if there are potential rivals who capture ${}^e\pi_1 < 0$ elsewhere, then these workers are inclined to enter for a ${}^e w_1 < \min w + ic_0 + {}^e\pi_1$. Third, the result that the incumbent suppliers demand a wage rate ${}^i w_1 = \min w + ic_0 + {}^e\pi_1$ that yields ${}^i\pi_1 = {}^e\pi_1$, is in accordance with the implications of the limit price theory of industrial organization (for example, Bain 1956, Sylos-Labini 1957, Modigliani 1958, Stigler 1968, Kamien and Schwartz 1971 and Stiglitz 1987; Chapter 7). Here this means entry barriers imply that $\text{pwg}_1 < {}^* \text{pwg}_1$, simply because the assumption of zero net returns is violated. Dynamic economies of market behavior (*i.e.*, investments in human capital) are introduced to the detriment of the static efficiency of market performance (*i.e.*, zero net returns).

17.6 CONCLUSION: The Tradeoff

In this chapter it has been shown that assumptions of supply competition in the labor market are of particular importance. Laborers' decisions on wage demands and investments in human capital depend critically on the competitive forces that are at work on the supply side of the labor market. Industrial organization shows that two extreme types of competition may characterize the market.

First, perfect competition and the contestable market framework focus on the implications of a pervasive threat of competition (actual and potential respectively) for the determination of the wage rate. In recent years the concept of perfect contestability has been growing in importance. Perfect contestability is based upon an assumption of a fear of "hit-and-run" entry. Entry barriers are conspicuous by their absence. This means that incumbents and (potential) entrants face an equivalent stock of human capital. From the predominance of the fear of entry it follows that the wage rate tends to a minimum reservation level, because incumbent suppliers are forced to adopt careful wage demands: $w_1 = w_{\min}$. The advantage of the theory of perfect contestability over perfect competition is that the results do *not* depend on the prevailing market structure. A basic flaw of the theories of perfect competition and contestable markets is that they include insufficient incentives for the introduction of investments in human capital. Therefore, both notions are only *static* benchmark cases of competition.

Second, the theories of imperfect competition and entry barriers introduce the *dynamic* aspects of market behavior. The economics of internal labor markets and the insider-outsider theory show that an important source of entry barriers is, at least partly, skill specificity: that is, specific human capital. Investments in specific human capital are associated with positive net returns, since they permit incumbents to benefit from a monopoly position. Therefore, a clear incentive exists to introduce improved specific human capital. Incumbent suppliers are able to capture $\pi_1 = \pi_{ec_1}$ without attracting new entry, because potential entrants face a cost disadvantage in the form of entry costs. Entry barriers and imperfect competition are merely two sides of the same coin. A tradeoff between static and dynamic efficiency may also arise in the labor market.

CHAPTER

18

HUMAN CAPITAL IMPROVEMENTS: Maximization of Attractiveness

18.1 INTRODUCTION: Barrier Labor Markets

As has been argued in Chapter 17, the tradeoff between static and dynamic efficiency of market performance is also succinctly present in the labor market. The concepts and theories of competition, as they are shaped in the literature on industrial organization (Part III), can therefore be usefully applied to the supply side of the labor market. For illustrative purposes, this thesis aggravates the efficiency dilemma in the labor market to the tradeoff between the wage level and investment in human capital. Chapter 17 shows that at first sight it seems that incentives to undertake human capital improvements are introduced to the detriment of static efficient wage rates (as these arise in perfectly contestable labor markets), since human capital differentials operate as entry barriers.

This chapter argues that the barrier market assumptions provide conditions in the labor market that simultaneously facilitate careful wage demands and human capital investments.¹ The fact of the matter is that market performance depends critically on the assumptions of (zero or positive) sunk cost and (a)symmetry of incumbent and potential laborers (Chapter 11). In effect, the argument presented below proceeds along the lines of Chapter 15. That is, wage rate and human capital improvement in the labor market setting take the place of price and cost reduction in the commodity market model. Therefore, bearing in mind the contents of Chapters 15 and 17, the argument in this chapter follows in a straightforward way. This implies that the presentation of the model in this chapter can be brief.

The plan of this chapter is as follows. Section 18.2 indicates the entry-detering strategies in a barrier labor market. Not surprisingly, it is shown that entry deterrence requires, on the one hand, that the wage rate is such that the human capital investment yields zero returns (Proposition 18.1) and, on the other, that (incumbent) laborers undertake human capital investment that seeks to maximize the laborers' attractiveness to firms (Proposition 18.2). Section 18.3 presents a simple model. It is proven that, provided that particular conditions hold, a unique entry-detering human capital investment can be identified (Proposition 18.3). Section 18.4 offers a conclusion. This chapter uses the two-period - two-market benchmark model (Chapter 17).

¹ Here it must be noted that this chapter focuses on the *supply* side of the labor market, which is in accordance with the intentions of Part VIII. It is only competition among (incumbent and potential) laborers that matters. This means that the argument is analogous to the one on the commodity market presented in Chapter 15. The demand side (*i.e.*, firms) merely enters as passive. Chapter 20 indicates some of the implications of an active demand side.

18.2 ZERO RETURNS AND MAXIMUM ATTRACTIVENESS

18.2.1 Barrier Labor Market

As has been noted time and again, existing contestable market models deal with the case where firms' investment cannot occur if it is associated with sunk cost (Schwartz and Reynolds 1983, Shepherd 1984 and Stiglitz 1987). Hence, perfect contestability is not associated with investment in human capital, because the latter requires foregone earnings and direct sunk costs of investment (Becker 1964; see also Chapter 17). A firm that is closed down before all sunk costs are rewarded incurs a loss. Equivalently, a laborer who stops working before the investment in human capital (provided that it is not fully firm-specific) is rewarded faces a loss as well.

The key barrier market assumptions (Assumption 10.1 and 10.2) apply to those cases where the incumbent suppliers anticipate that a positive sunk cost does *not* fully block free entry, which can be illustrated with the help of home market arguments (Chapter 11). The home market argument implies that incumbent suppliers anticipate responses of potential entrants who face similar sunk costs (sc_0) and carry out equivalent improvements of human capital (Δhc_1) in a home market (so that $mp_1 = mp_1$).

ASSUMPTION 18.1 (barrier labor market). *Incumbent laborers anticipate the entry threat of at least one (or some) potential entrant(s) who (i) has (have) incurred similar sunk costs, (ii) carries (carry) out equivalent improvements of human capital and (iii) faces (face) a zero opportunity cost of entry: $ic_0 = sc_0 = ic_0 = sc_0$, $\Delta hc_1 = \Delta hc_1$ and $\pi_1 = 0$ for $i = 1, \dots, ni$ and $e = 1, \dots, ne$.*

Assumption 18.1 cements the key barrier market assumption (Assumption 10.2) in a labor market setting.

In the context of the economics of human capital (Becker 1964, Blaug 1976 and Siebert 1985; see also Chapter 17) Assumption 18.1 implies that incumbents anticipate that there are some potential entrants around (*i.e.*, in a closely related profession or nearby region)¹ who independently and simultaneously introduce equivalent changes in human capital. If, for the sake of simplicity, foregone earnings and psychic income are ignored, then the net present value of returns to the investments in human capital (π) depends on incremental wage income (Δw) and the cost of investment ($ic = sc$).

18.2.2 Zero Returns

In a labor barrier market incumbents and entrants play a second-stage game on wage rates and a first-stage game on investments in human capital. The assumptions of a three-stage structure (but two-stage game) and complete and imperfect information are sustained (Chapter 17). Henceforth, generally discounting is ignored and time indices are suppressed. Assume that all laborers have undertaken the same investment in human capital in the first period: $hsc = esc$ for all incumbents and entrants h and g . Let R_w be the wage rate in period $t = 1$ such that $\pi_1 = 0$: that is $R_w = ic_0 = sc_0$. Incumbent laborers have two regions of relevant strategies (s): $iw > R_w$ (s^1) and $iw = R_w$ (s^2).

¹ For example, compare with Shepherd's (1984), Green's (1987), Cairns and Mahabir's (1988) and Calem's (1988) arguments in a commodity market setting (Chapters 9 and 11).

The strategies showing $i_w < R_w$ are dominated, because they yield $i_\pi < 0$. The potential entrants' set of replies contains two regions of relevant responses: $e_w < i_w$ (e_s^1) and non-entry (e_s^2). Actions showing $e_w \geq i_w$ are not worthwhile, because then the entrant cannot outperform incumbent laborers. The payoff matrix is similar to (17.4).

It is easily shown that the introduction of human capital investments in a free entry framework is associated with incumbents demanding R_w , implying $i_\pi = 0$, if the incumbent laborers only anticipate the entry threat of potential entrants with a zero opportunity cost of entry.

PROPOSITION 18.1 (zero returns). *Incumbent laborers demand a second-stage wage rate such that the investment in human capital is just rewarded: $i_w = R_w$ so that $i_\pi = 0$ for $i = 1, \dots, n_i$, conditional upon*

- (i) $e_\pi = 0$ for some and $e_\pi \geq 0$ for all equally equipped potential entrants when they refrain from entry (credibility), and
- (ii) $R_w \leq \max_w$ (feasibility).

The incumbent suppliers' compliance with the zero-returns condition constitutes a unique second-stage equilibrium strategy.

Proof Proposition 18.1. The proof bears close resemblance to that of Proposition 17.1. Given condition (i), incumbents' strategies associated with $i_w > R_w$ provoke entry, because then there are entrants who can capture $e_\pi > 0$ by entering against $R_w < e_w < i_w$. The incumbent is expelled from the market, so that $i_\pi < 0$. When the incumbent laborer adopts $i_w = R_w$, successful entry is associated with $e_w < i_w = R_w$, which yields $e_\pi < 0$. The optimal reply is not to enter (Figure 18.1).

Figure 18.1

•Second-Stage Game on Wage Rates in a Barrier Labor Market

Hence, backward induction indicates that the strategy pair $i_w = R_w$ and non-entry constitutes a unique second-stage equilibrium, conditional upon $R_w \leq \max_w$ [condition (ii)]. Incumbents are forced to demand a wage rate such that the net return to investment

is zero in order to deter entry. *Q.E.D.*

Condition (i) needs clarification (Chapter 11). The condition that incumbents anticipate that $e_{\pi} = 0$ ($e_w = R_w$) for some and $e_{\pi} \geq 0$ ($e_w \geq R_w$) for all equally equipped potential entrants whenever they refrain from entry, introduces a credible entry threat for any $i_w > R_w$. On the one hand, if all potential entrants capture $e_w > R_w$ elsewhere, then all potential entrants are only inclined to enter for $i_w > e_w > R_w$. Hence, incumbents can forestall entry by setting $i_w = e_w > R_w$, which implies that the incumbents capture *positive* net returns. On the other hand, potential entrants receiving $e_w < R_w$ face profitable entry opportunities for any $e_w < i_w < R_w$. Then, only $i_w = e_w < R_w$ deters entry, which means that incumbents receive *negative* net returns. In the extreme it may be that there are equally equipped potential entrants who either are unemployed (a payoff of u_i) or receive the minimum reservation wage in a home market ($e_w = \min_w$) so that the market is perfectly contestable in period $t = 1$. Each $\min_w < e_w < R_w$ constitutes an intermediate case between a perfectly contestable and a perfect barrier market.

It appears plausible, however, that equally equipped potential entrants (expect to) receive $e_{\pi} \geq 0$ elsewhere. The reason is straightforward. A potential entrant receiving $e_{\pi} < 0$ in period $t = 1$ (after the introduction of a human capital improvement) lacks the very incentive to be engaged in schooling in period $t = 0$ so as to uphold the possibility of future entry. Given the incumbents' strategy to demand $i_w = R_w$, successful entry is associated with $e_w < i_w < R_w$ and, therefore, $e_{\pi} < 0$. Hence, irrespective the strategy adopted (entry or non-entry), the potential entrant *always* faces a negative payoff. However, if the potential entrant decides not to carry out schooling investment in period $t = 0$, (s)he receives, at least, either e_{ui} (unemployment) or $\min_w$ (employment in a contestable market) in period $t = 1$, which is likely to be preferred to the negative payoff associated with schooling (since e_{ui} is close to $\min_w$, whereas $\min_w = 0$; see Chapter 17). In any event, the incumbents' zero-returns wage demands only occur if incumbent workers anticipate the entry threat of equally equipped potential entrants facing a zero opportunity cost of entry.

18.2.3 Maximum Attractiveness

In a barrier labor market incumbents and entrants play a game on attractiveness: that is, wage rates and investments in human capital. Proposition 18.1 predicts that in a (perfect) barrier labor market a zero return to investment is the incumbents' second-stage equilibrium strategy. This subsection focuses on the first-stage game on investment ($i_c = s_c$). For the moment, suppose that there is a unique s_c (associated with a unique s_{mp} and zero-returns s_w) that yields maximum attractiveness (Proposition 18.3).

In period $t = 0$ workers carry out investments in human capital that yield $mp - w$ differentials. Moreover, in t^1 incumbents announce the associated zero-returns wage rates. In t^2 the (potential) entrants' [wage and (non)-entry] replies follow. The game is closed in t^3 . Incumbent laborers have two regions of relevant strategies in period $t = 0$: $i_{sc} < (or >) s_c$ (s^1) and $i_{sc} = s_c$ (s^2). The potential entrants' set of replies contains two regions of relevant responses: $e_{sc} > i_{sc}$ (e^1) and non-entry (e^2). Actions showing $e_{sc} \leq (or \geq) i_{sc}$ are not worthwhile, because then the entrant cannot outperform incumbent laborers.

The pervasive fear of entry not only forces the adoption of zero-returns wage demands, but also facilitates the introduction of optimal investments in human capital in order to reduce the likelihood of future exit and entry. Investments in human capital reduce the probability of future exit and entry, because improved human capital facilitates

a laborer's attractiveness to the firms.

PROPOSITION 18.2 (maximization of future attractiveness). Incumbent workers are inclined to maximize future attractiveness by introducing optimal first-stage investments in human capital: ${}^i sc_0 = {}^* sc_0$ for $i = 1, \dots, ni$, conditional upon

- (i) ${}^e \pi_1 = 0$ for some and ${}^e \pi_1 \geq 0$ for all equally equipped potential entrants when they refrain from entry (credibility), and
- (ii) ${}^* w_1 \leq {}^{\max} w_1$ (feasibility).

The implementation of optimal investment efforts constitutes a unique first-stage equilibrium strategy.

Proof Proposition 18.2. Condition (i) and substitution of the sunk cost in equation (17.4) give the payoff matrix. If an incumbent laborer introduces a nonoptimal investment effort, a potential entrant is able to offer superior work, because ${}^i pwg = {}^i mp - {}^i w < {}^* pwg = {}^* mp - {}^* w$, by introducing ${}^e sc$ such that ${}^i mp - {}^i w < {}^e mp - {}^e w \leq {}^* mp - {}^* w$. Given condition (i), the entrant can outperform the incumbent profitably. The incumbent is confronted with the negative payoff that follows upon exit, whereas the entrant receives ${}^e \pi > 0$. If the incumbent adopts ${}^i sc = {}^* sc$, then the entrant cannot enter successfully, provided that the incumbent demands the zero-returns wage rate ${}^* w$ [Proposition (18.1)]. The strategy ${}^i sc = {}^* sc$ (in combination with non-entry in the second stage) constitutes a unique first-stage equilibrium if ${}^* w \leq {}^{\max} w$ [condition (ii)]. *Q.E.D.*

Propositions 18.1 and 18.2 describe the perfect equilibrium strategy of the two-stage game on investment (first stage: ${}^i sc_0 = {}^* sc_0$) and wage rate (second stage: ${}^i w_1 = {}^R w_1 = {}^* w_1$).

The condition (i) that incumbents anticipate that ${}^e \pi = 0$ for some and ${}^e \pi \geq 0$ for all (equally equipped) potential entrants whenever they refrain from entry, introduces a credible entry threat for any ${}^i w > {}^* w$ or ${}^i pwg < {}^* pwg$. That is, the source of the entry threat lies in workers employed in related barrier labor markets (Chapter 11). The result is that the wage rate is kept down and that the adoption of investments in human capital is given an impetus. The notion of the barrier market can also provide a formal elaboration of a concept of workable competition in a labor market setting (Chapter 9). Investments in human capital are employed as a pre-emptive strategy in an environment of a pervasive entry threat, resembles to the literature on strategic investment of firms (for instance, Gilbert and Newberry 1982 and Gilbert 1986). The anticipation that potential entrants are able to leapfrog if incumbents fail to invest optimally (Fudenberg *et al.* 1983), underlies the favorable outcomes in a barrier labor market.

18.3 A MODEL

Let the initial wage rate in period $t = 0$ be associated with

$$(18.1) \quad {}^i w_0 = \min w \text{ for } i = 1, \dots, ni.$$

All laborers face an equivalent stock of human capital ${}^f hc_0$ and, thereby, value of the marginal product ${}^f mp_0$ for $f = i = 1, \dots, ni$ and $f = e = 1, \dots, ne$. A laborer deploys maximization of future attractiveness (Proposition 18.2) in order to reach a position of

exit avoidance and entry deterrence. For that purpose incumbents increase the value of their marginal product by means of investments in human capital.

The prevailing value of the marginal product depends on the current stock of human capital [equation (17.1)]. Improved human capital is associated with an increased value of the marginal product, which in turn yields, *ceteris paribus*, the opportunity to demand a higher wage rate [equation (17.2)]. The cost of investment simply indicates the expenditures that are necessary to permit a change in human capital. When, for the sake of simplicity, a one-period investment lag is assumed, this means that the investment of sc_0 in period $t = 0$ determines the improvement of human capital hc_1 in period $t = 1$ [equation (17.3)]. Investment is assumed to be subject to decreasing returns [equation (17.3a)].¹

An investment effort in period $t = 0$ yields two countervailing effects in period $t = 1$. On the one hand, improved human capital raises the value of the marginal product, but, on the other, an increase of the wage rate is necessary to recover the investment cost. Noting that $hc_1 = hc_0 + \Delta hc_1$, substitution of equation (17.3) in (17.1) gives

$$(18.2) \quad mp_1 = MP(hc_1) = MP(hc_0 + \Delta hc_1) = MP[hc_0 + HC(sc_0)].$$

So, Δmp_1 appears as a monotonically increasing function (F) of sc_0 :

$$(18.3) \quad \Delta mp_1 = mp_1 - mp_0 = MP[hc_0 + hc(sc_0)] - mp_0 = F(sc_0).$$

It is assumed that (MP, HC and) F is (are) twice continuously differentiable. First,

$$(18.4) \quad \partial F / \partial sc_0 > 0.$$

Second, assume that the second-order derivative of MP [equation (17.1)] permits condition (17.3a) and (17.3b) to indicate that

$$(18.5) \quad \partial^2 F / \partial sc_0^2 < 0,$$

which implies that the decreasing returns property of investment in human capital enters the relation between sc_0 and Δmp_1 without hindrance.²

The wage rate increase must be just sufficient to reward the investment, for positive returns will induce entry (Proposition 18.1). So, the net return to investment must equal zero. Thus, Δw_1 appears as a function (G) of sc_0 :

$$(18.6) \quad \Delta w_1 = (1+r) \cdot sc_0 = G(sc_0).$$

An incumbent supplier seeks to maximize future attractiveness in order to reach a position of exit avoidance and entry deterrence (Proposition 18.2). Substitution of $mp_1 = mp_0 + \Delta mp_1$, $w_1 = w_0 + \Delta w_1$, equations (18.3) and (18.6) gives

¹ As argued in Chapter 17, the assumption of decreasing returns is generally adopted in human capital models (Killingsworth 1983) and the investment theory of the firm (Kamien and Schwartz 1982).

² Strictly speaking, an equivalent for Proposition 15.1 is necessary to ground condition (18.5). However, for the sake of brevity, Chapter 18 imposes an *a priori* assumption.

$$(18.7) \quad \max_{sc_0} \text{pwg}_1 = [mp_0 + F(sc_0)] - [\min_w + G(sc_0)] = H(sc_0).$$

By analogy to Proposition 15.3 and 15.4, it can be proven that the maximization rule (18.7) gives a unique solution, provided that three conditions hold:

PROPOSITION 18.3 (existence unique optimal investment effort). *There exists a unique optimal investment effort if three conditions hold:*

- (i) if $sc_0 = 0$, then $\partial F/\partial sc_0 > (1+r)$;
- (ii) $\lim_{sc_0 \rightarrow \infty} F(sc_0) = \max \Delta mp_1$, where $\max \Delta mp_1 \geq 0$; and
- (iii) if $sc_0 = 0$, then $\Delta mp_1 < \max \Delta mp_1$.

Proposition 18.3 is proven in Appendix E.

Proposition 18.3 may be illustrated with the help of Figure 18.2.

Figure 18.2
Optimal Investment in Human Capital

Superscript * denotes optimal values. The maximal increase of future attractiveness results when the gap between $F(sc_0)$ and $G(sc_0)$ (i.e., Δmp_1 and Δw_1) is maximized. The first-order condition requires that $\partial F/\partial sc_0 = \partial G/\partial sc_0$. This condition implies that $^*\Delta pwg_1 = ^*\Delta mp_1 - ^*\Delta w_1$. The optimal investment effort is only positive if $\partial F/\partial sc_0 > (1+r)$ for $sc_0 = 0$ [condition (i)], because otherwise the increase in the wage rate always exceeds the rise in the value of the marginal product so that a decrease of the future attractiveness occurs (1G). The increase of the wage rate is zero (2G) if learning-by-doing occurs such that $\Delta mp_1 = \max \Delta mp_1$ if $sc_0 = 0$ [condition (iii)]. A further requirement is that there is some $\max \Delta mp_1$ [condition (ii)], so that the change in human capital does not reach infinity when $sc_0 \rightarrow \infty$. This simply means that there is an upper bound to Δmp_1 , which is

economically plausible.

18.4 CONCLUSION

In a barrier market laborers are engaged in pre-emptive sunk investment in human capital in order to reach a position of exit avoidance and entry deterrence. The fear of potential entry induces incentives for zero-returns wage demands and maximization of expected future attractiveness to the firm. The notion of the barrier market can also provide a formal elaboration of a concept of workable competition in a labor market setting, because the free entry assumption yields (partial) second-best outcomes of market behavior. The barrier commodity market has a reflection in the labor market.

PART

IX

MULTIMARKET COMPETITION

Microfoundation of Macroeconomics, Industrial Organization and Strategic Management

KEY WORDS. Parts IV, V, VI, VII and VIII have presented models which can be characterized by indicating one crucial feature: the models view competition from a partial perspective: that is, the incumbent and potential suppliers' performance in one or two commodity or labor markets is studied. Part IX supplements the argument by putting the barrier market theory in a general perspective. Chapter 19 introduces the framework of multimarket competition. Chapter 20 focuses on the issue of the microfoundation of macroeconomics, which means that multimarket competition is considered in a commodity-labor market framework. Chapter 21 elaborates on the strategic dimension of multimarket competition in a commodity-commodity market setting with reciprocal entry (threats) by exploring the industrial organization - strategic management interface.

1944

1944

1944

1944

1944

1944

1944

1944

CHAPTER

19

MULTIMARKET COMPETITION: Vertical and Horizontal Intermarket Spillovers and Strategies

19.1 INTRODUCTION: Set of related Markets

The discussion on the sources of entry and barrier market scenarios (particularly Parts IV and V) can be summarized and clarified by introducing the concept of multimarket competition. Moreover, multimarket competition integrates the discussion in this part. Multimarket competition is defined as rivalry among competitors which operate in a set of related markets. Related markets are characterized by entry barriers to outside (new or unrelated) rivals, but ease of entry for competitors within the set of related markets. This argument naturally leads to emphasizing a general rather than partial perspective to competition. This points to interfaces with general (dis)equilibrium theory (Chapter 20) - particularly the debate about the microfoundation of macroeconomics - and strategic management (Chapter 21) - particularly the study of strategy analysis and choice. This chapter develops a stepping stone for the argument in Part IX by defining and classifying the key elements of multimarket competition.

Key elements of multimarket competition (intermarket spillovers and strategies: Section 19.2) are the focus of rivalry (types of games: Section 19.3), ease or difficulty of entry (excess resources, entry or exit barriers and one-sided or reciprocal entry: Section 19.4) and intermarket spillover effects (four categories: Section 19.5). Although, for the sake of clarity, the argument in this chapter proceeds along the lines of firms' rivalry in related commodity markets, analogous reasoning can be applied to labor-labor (Part VIII) or commodity-labor (Chapter 20) market interaction.

19.2 FRAMEWORK OF MULTIMARKET COMPETITION

Multimarket competition has both a horizontal and vertical dimension. Figure 19.1 illustrates the multimarket competition that a firm in an intermediate market (that is, a market which borders on a factor as well as customers' market) can face. Two crucial concepts are intermarket spillovers and intermarket strategies (Bulow *et al.* 1985a and Cooper and John 1988). Intermarket spillovers are defined as externalities between two or more markets in terms of payoffs: that is, the vector of equilibrium payoffs in market A has an impact on the equilibrium payoffs in market B and *vice versa*. Intermarket strategies are the rivals' actions that concern more than one market.

Broadly speaking, from the viewpoint of a firm in an intermediate market I2, five

categories of intermarket strategies/spillovers can be distinguished: (1) forward diversifying integration/spillover (I and III), (2) forward downstream integration/spillover (II), (3) horizontal diversification/spillover (IV and V), (4) backward diversifying integration/spillover (VI and VII) and backward upstream integration/spillover (VIII). Figure 19.1 assumes that both multimarket strategies (diversification and integration) and spillovers can be differentiated by distinguishing the location from product focus on the one hand and the vertical from horizontal dimension on the other. By and large, the argument in this chapter on strategy is concentrated on horizontal multimarket competition that ensues from geographical or product diversification. For example, I1, I2 and I3 indicate different product markets (IV and V depict horizontal product diversification and/or spillover), whereas I2, I4 and I5 are different geographical markets (IX and X denote horizontal geographical diversification and/or spillover).¹

Figure 19.1
Multimarket Dimensions

¹ Of course, Figure 19.1 simplifies matters. To mention two complications: first, diversification and integration strategies can overlap, whereas Figure 19.1 defines diversification and integration as strategies concerning horizontal and vertical competition respectively; and, second, an intermediate market is generally related to more than one factor and sales market, while Figure 19.1 depicts one-to-one relations. However, the five broad categories of strategies and spillovers can be used as a heuristic device so as to distinguish pure cases. Moreover, the discussion of vertical intermarket rationing spillovers ("dual decision hypothesis") that is specific for general non-Walrasian equilibrium theory, is postponed to Chapter 20. For the sake of brevity, Chapter 19 only focuses on commodity-commodity market interaction.

Figure 19.1 makes clear that multimarket competition is related to many different branches of literature. Worth mentioning is the literature on diversification (Lecraw 1984 and Mueller 1987), integration (Harrigan 1985 and Caves and Bradburd 1988), multiproduct firms (Teece 1982 and Panzar 1989), multinational enterprises (Caves 1982 and Dunning 1988), interbrand competition (Scherer 1980 and Judd 1985), transaction costs (Teece 1980 and Williamson 1989), international trade (Brander and Spencer 1984 and Krugman 1989) and non-Walrasian macroeconomics (Weintraub 1979 and Cooper and John 1988). This chapter can of course only refer to illustrative examples. To support the argument, a selection of exemplary pieces of literature is mentioned when convenient.

Multimarket competition introduces new elements into strategy choice. Figure 19.2 indicates three considerations which the incumbent firm has to take into account while evaluating the implications of multimarket competition.

Figure 19.2
Games of Multimarket Rivalry

First, the focus of rivalry can be characterized by distinguishing three types of games (Section 19.3). Second, the ease or difficulty of entry cannot be ignored (Section 19.4). Third, intermarket spillover effects can reflect important entry motivations (Section 19.5).

19.3 FOCUS OF RIVALRY

Multimarket competition is associated with three categories of games. First, incumbent firms play an *incumbents against incumbents* game. This type of game has been studied in the well-established theories of (im)perfect competition without (free) entry (Shapiro 1989a). Only internal market conditions determine competition. Second, incumbent firms and potential entrants are engaged in an *incumbents against entrants* game. The entry (deterrence) literature focuses on the features and implications of this category of games (Chapter 7). External conditions dominate over internal competition. Third, potential rivals are the players in an *entrants against entrants* game. This game has been explored only sporadically (Sherman and Willett 1967, Kalish *et al.* 1978, Nti and Shubik 1981 and Nti 1989). The entrants against entrants game includes the potential threat of simultaneous (or too much) entry (Holler 1987). The three games are interrelated: for

instance, the features of the entrants against entrants game influence the outcome of the incumbents against entrants game (Sherman and Willett 1967).

The characteristics of multimarket competition critically depend on the relative importance of the three types of games. Taking the extreme cases where one of the types of games dominates over both the others can clarify matters. First, if the incumbents against incumbents game is predominant, only internal market conditions matter. This means that the threat of entry is ineffective. Second, predominance of the entrants against entrants game negates the entry threat into the incumbents' market. Third, if the incumbents against entrants game cannot be ignored, the interesting case occurs where multimarket competition is in force.

19.4 EASE OR DIFFICULTY OF ENTRY

A key determinant of the incumbents against entrants game is the extent to which entry is easy or difficult. Brunner (1961) points out that new firms

seem somewhat insignificant alongside the much more obvious ease of entry by existing firms (Brunner 1961, p. 250).

So, a first distinction is the one between new and existing firm entry (Chapter 11). The proposition is that existing firm entry can be easy, as related firms, contrary to new (or unrelated) entrants, can benefit from home market assets, which permit related firms to avoid barriers (Chapter 7) that are unsurmountable for new entrants. If existing firms are potential entrants, the credibility of the entry threat is facilitated (hampered) by excess (binding) resources, one-sided (reciprocal) entry and positive (negative) intermarket spillovers. This section very briefly summarizes the arguments in Chapter 11 on the first two aspects, whereas the last element is discussed in Section 19.5.

Excess resources, if fungible, allow existing firms to economize on entry cost. An existing firm can divert excess resources from home to entry market. Entry is easy if the switching or transport costs (which both reflect an entry barrier) are low (Calem 1988, p. 171). An advantage is that the opportunity cost of using excess resources is zero (Section 11.3). *One-sided entry protects potential entrants against incumbents' retaliatory moves in their home market*. This reduces the opportunity cost of entry. If incumbent firms can undertake reciprocal entry, the entry threat is undermined (Section 11.4).

19.5 INTERMARKET SPILLOVER EFFECTS

The entry threat is particularly effective if potential entrants can exploit intermarket spillover opportunities. Bulow *et al.* (1985a) argue that a multimarket perspective introduces new elements in oligopoly theory. The key point is that

changes in a firm's opportunities in one market may affect its profits by influencing competitors' (or potential competitors') strategies in a second oligopoly market (Bulow *et al.* 1985a, pp. 488-489).

Aside from the reciprocal dependency of different firms in different markets, the strategy choice of multiproduct (or multimarket) firms points to the importance of the multimarket perspective. The key point is that intermarket spillovers can exert influence on strategy choice. To be precise, the opportunity to exploit intermarket spillovers can be an

important motive for (refraining from) entry. This observation is also recognized in the literature on diversification, multiproduct firms, integration and multinational enterprises. An early example of a list of sources of intermarket spillovers is Hines (1957).

Bulow *et al.* (1985a) distinguish strategic substitutes from strategic complements. That is,

[c]onventional substitutes and complements can be distinguished by whether a more 'aggressive' strategy by firm A (e.g., lower price in price competition, greater quantity in quantity competition, increased advertising, etc.) lowers or raises firm B's *total* profits. *Strategic* substitutes and complements are analogously defined by whether a more 'aggressive' strategy by A lowers or raises B's *marginal* profits (Bulow *et al.* 1985a, p. 489; Bulow *et al.* 's italics).

Strategic substitutes and complements can be explained in terms of reaction curves. Downward sloping reaction curves imply strategic substitutes, whereas upward sloping reaction functions indicate strategic complements.¹

The key point is that firms can improve upon their strategy choice by taking account of the nature of *intermarket strategic spillovers*. This argument extends over the literature on entry and entry deterrence (Chapter 7). Bulow *et al.* (1985a) illustrate this argument by considering, for example, the strategic effect of diversification in the case of Frontier Airlines. That is,

[i]n the early 1980s the firm expanded beyond its original Denver hub to capitalize on some apparently profitable opportunities. . . . After Frontier spread itself over several new markets, other airlines began to compete much more aggressively for shares of the Denver market. Some of this new competition may have been inevitable in a changing, deregulated environment, but some of it was probably due to a perceived weakness on Frontier's part that arose from its being 'spread thin' (Bulow *et al.* 1985a, p. 508).

Figure 19.3 summarizes eight examples of intermarket spillover effects.

Intermarket *demand* spillover effects can arise if buyers are mobile between markets. The strategy of firms in market 2 can influence the scale of demand in market 1 (and *vice versa*) if the cross-elasticity is nonzero. The intermarket demand spillover

is positive if a firms' demand in one market is complementary to its demand in the second . . . and would be negative if selling more in one market hurts prospects in the other (Bulow *et al.* 1985a, p. 509).

That is, complements are associated with a positive demand dependency (video recorders and discs), whereas the demands for substitutes are negatively correlated (public transport and private cars).

¹ Bulow *et al.* (1985b) construct a Cournot reaction curve which gives a change of the sign of the first-order derivative in the relevant production interval as a result of the market's strategic features. This means that the reaction function is hill-shaped. This provides an economic rationale for (and improves upon) Rand's (1978) and Dana and Montrucchio's (1986) abstract proposition that nonlinear reaction functions can yield chaotic quantity competition (Van Lier and Van Witteloostuijn 1990 and Van Witteloostuijn and Van Lier 1990).

Figure 19.3
Intermarket Spillover Effects

Spillovers	Positive	Negative
Demand	Goodwill	Badwill
	Complements	Substitutes
Supply	Scale economies	Scale diseconomies
	Scope economies	Scope diseconomies

Judd (1985) presents a multiproduct duopoly game which indicates that negative demand spillover effects may imply that an incumbent is inclined to withdraw a product from the market in the face of competition with an entrant in a substitute market. Champsaur and Rochet's (1989) multiproduct duopoly model indicates that in equilibrium both rivals avoid negative demand spillovers *via* the creation of a gap between both product lines. A multiproduct firm (supplying in a number of markets) can even face product cannibalism (Connor 1988). If product A and B are substitutes, upgrading product A's quality can (in the extreme) destroy product B's market. If both products are complements, change of the user technique of product A (compact disc player instead of pickup) can make the supply of product B (records) unprofitable. Goodwill (or badwill) transfer can introduce a further intermarket demand spillover to a multiproduct firm (Margolis 1989).

Joint economies of scale or scope can induce intermarket *supply* spillovers. Joint economies imply that a firm can decrease unit production cost of product A in market 1 by increasing supply of product A (scale economies) or B (scope economies) in market 2. Strategic use of the learning curve (Spence 1981 and Fudenberg and Tirole 1983b) provides a telling example. Take the case where a first mover (incumbent) can improve its position against a second mover (entrant) by expanding output in the pre-entry stage so as to benefit from the learning curve effect in the post-entry period. For example,

if a Japanese firm has decreasing costs over time (as in an industry with a "learning curve") ..., the firm may 'dump' output in the early stages of a market's development to encourage the competitor to either contract operations or withdraw from the market (Bulow *et al.* 1985a, p. 504).

The trade (dumping) models in international economics explore the implications of intermarket supply effects (Brander and Krugman 1983, Pinto 1986 and Casas 1989). Joint diseconomies of scale or scope can induce adverse effects. To be precise,

[n]atural resource models are the mirror image of learning curve models: They are games played over time (*i.e.*, sequentially) where greater output in period 1 will increase marginal cost in period 2. That is, firms' production costs will generally rise with cumulative output as 'digging deeper' is required for extraction. In this case ... a natural resource oligopolist will produce less in the first period and more in the second than if both periods' strategic variables were chosen simultaneously (Bulow *et al.* 1985a, p. 508).¹

The key point is that intermarket spillover effects can have important implications for multimarket competition. The list of examples in Figure 19.3 is of course not exhaustive. For instance, Caves (1982, Chapter 4) summarizes spillovers in the context of the multinational enterprise and Teece (1982) lists multimarket externalities which diversified firms can exploit. A key argument in this literature is the stipulation of exploitation of (excess) fungible but intangible assets as an explanation of (international) diversification (for example, R&D, marketing facilities, knowhow, managerial skills, *etcetera*). This indicates that intermarket spillover effects need not be restricted to *technical* factors. Two further examples are worth noting. First, *financial* benefits from diversification (in the form of risk spreading) can plead in favor of entry (Chang and Thomas 1989). Moreover, deep-pocket arguments suggest that existing firms can finance entry more easily (and cheaply) than new entrants by making use of financial resources earned elsewhere and/or easy access to imperfect capital markets (Poitevin 1989).

Second, Porter (1976) points to *managerial* exit barriers which can be associated with intermarket psychological spillover effects, whereas Teece (1982) argues that diversification can be efficient as a result of the exploitation of fungible but tacit organizational knowledge. The two following quotations from Porter illustrate this point:

Commitment and intense involvement with the business are, in fact, some of the important characteristics of a successful leader and effective manager. In the face of the commitment, however, it is difficult to 'give up' by making the divestment decision. Divestment is a blow to a manager's professional pride because it reflects his inability to make the business successful (Porter 1976, p. 25);

and

[i]f the managers are part of a larger company, incentive and reward systems often place no premium on exit decisions and may actually discourage them (Porter 1989, p. 25).

Porter's exit barrier argument can have a mirror image in the sense that managerial incentives may promote entry through, among others, pride and premium motives.

¹ This example points to the role of the sequence of moves. Sequential markets give results different from simultaneous choice. This is because sequential markets permit the communication of commitment.

 

MICROFOUNDATION OF MACROECONOMICS: Rationing, Entry and Exit

20.1 INTRODUCTION: General Equilibrium Theory and Dynamic Competition

The concept of multimarket competition offers an integrative framework which identifies common denominators in different literatures. Apart from rivalry in a related set of commodity markets (Section 19.3 to 19.5 and Chapter 21), multimarket competition can also capture the interaction between different labor markets (Part VIII) or commodity and labor markets (Chapter 20). This chapter explores the interface between literature on microfoundations of macroeconomics and industrial organization by introducing innovative behavior in, essentially, the two-market - two-period temporary equilibrium framework of Section 10.3. That is, the arguments of Chapters 13 and 18 are brought together so as to analyze commodity-labor market interactions in a potentially free entry framework. The natural frame of reference of multimarket modeling in economics is general equilibrium theory.¹ General equilibrium theory is general because it seeks to study the economy as a whole by analyzing the interactions between (many) agents (*i.e.*, firms and households) who operate in a (large) number of markets. The vertical intermarket spillovers from commodity to labor markets and *vice versa* are central to the microfoundation of macroeconomics debate.

Although the analysis of interactions between commodity and labor markets in the general equilibrium context of the microfoundations debate points to the importance of intermarket spillover effects, many important aspects of economic dynamics are missing. It is here where the insights of *industrial organization* are relevant. Ever since Schumpeter's (1943) seminal contribution, the literature on industrial organization emphasizes the crucial importance of dynamic competition (Part III, particularly Chapters 7 and 8), being associated with innovative activity in particular (Kamien and Schwartz 1982, Stiglitz and Mathewson 1986 and Jacquemin 1987). This chapter examines, in a preliminary way, which shape interactions among innovative firms and laborers take in particular regimes. Since the number of possible regimes is too large to offer a full scope of examination, attention is focused on some interesting cases.

The chapter is organized as follows. Section 20.2 briefly evaluates the literature on the microfoundation of macroeconomics. Section 20.3 introduces an association between

¹ A cursory look in prominent economic journals provides the reader with the impression that there is "the widespread belief that every economic theory must be fitted into the general equilibrium mold, if it is to qualify as rigorous science" (Blaug 1980, p. 192). Fitoussi (1983) argues that "[g]eneral equilibrium theory, which by virtue of this axiom is given the status of a meta theory, [is] a common structure within which all other theories must be expressed" (Fitoussi 1983, p. 2).

firms' process innovations and laborers' human capital improvements, in the sense that improved human capital imposes an incremental cost reduction on the firms' process innovation. Section 20.4 explores regimes for the cases where firms and laborers are homogeneous. Sections 20.5 and 20.6 successively examine what the implications are of introducing the heterogeneity of firms and laborers respectively. Section 20.7 shows that the free entry feature of the commodity market may introduce a tendency to increased wage rate dispersion. Section 20.8 summarizes the argument.

20.2 MICROFOUNDATION OF MACROECONOMICS

20.2.1 Temporary and Intertemporal Equilibrium

Obviously, the equilibrium concept is fundamental to general equilibrium theory. General equilibrium theory refers to a Hicksian notion of equilibrium. That is, an equilibrium is said to prevail if all economic agents can continue to execute their plans *without* any revision (Hicks 1939, p. 132). This means that in equilibrium expectations are confirmed. Following a Swedish tradition (Lindahl 1929), Hicks (1939) argues that time can be brought into a general equilibrium setting by describing the economy as a sequence of discrete periods.¹ A period is an interval in time in which prices and plans remain unchanged. That is, prices as well as the agents' plans are allowed to change only at the junction of two subsequent periods. Each period has its own demand and supply system. Within each period demand and supply meet and transactions occur only once. An agent's demand and supply depend on his or her expectations of current and future variables and conditions. The term *temporary* is reserved for within-period events and conditions, whereas *intertemporal* refers to events and conditions over periods. An equilibrium can be temporary and intertemporal in quality.

An intertemporal equilibrium is more than a sequence of temporary equilibria, because the expectations upon which each single period equilibrium is based, have to be consistent with one another (Hicks 1939, p. 24). A Hicksian temporary disequilibrium is associated with disappointed expectations of current variables and conditions (*temporary uncertainty*), whereas a Hicksian intertemporal disequilibrium follows from disappointed expectations of future variables and conditions (*intertemporal uncertainty*). Moreover, some authors refer to an equilibrium if markets clear. The situation where prices are fully flexible - and hence markets clear - is denoted a *Walrasian* equilibrium (Subsection 20.2.2). The opposite equilibrium is referred to as *non-Walrasian* (Subsection 20.2.3), which can be associated with the rationing of agents.

A further distinction defines temporary and intertemporal dynamics. *Temporary* dynamics refers to the adjustment mechanism that describes the process which brings about (dis)equilibrium *within* a period. That is, temporary dynamics points to the way in which agents coordinate their transactions in the current period of time (in which, recall, transactions take place only once). *Intertemporal* dynamics is concerned with the phenomena or behavior that relate(s) preceding to succeeding periods (Baumol 1970, p. 4). So, intertemporal dynamics indicates why and how economic agents change their plans and behavior from one period to another. Chapter 20 particularly examines general equilibrium theory's contribution to the study of *temporary* dynamics. Intertemporal

¹ Morishima (1977), Jaffé (1980) and Van Witteloostuijn and Maks (1989 and 1990) point out that Walras (1874) foreshadowed Hicks (1939) by using a temporary equilibrium construct.

dynamics is only framed in terms of a two-period model. A series of many periods is studied in Chapter 23. Therefore, the review of the literature in this section mainly focuses on temporary equilibrium models, whereas the discussion of intertemporal theories is postponed to Chapter 23.

20.2.2 Walrasian Equilibrium Models

Not surprisingly, Walrasian equilibrium models are a sequel to Walras' (1874) seminal contribution to economic theory. The characteristic feature of Walrasian equilibrium models is that prices are assumed to be *perfectly flexible*. That is, a Walrasian equilibrium model is associated with the assumption of *market clearing*. Supply and demand are coordinated such that either side of the market can sell or buy the quantity of commodities that is associated with the equilibrium price level. The Walrasian equilibrium method is based upon the assumption of price *tâtonnement*. That is, transactions only occur after the declaration of the equilibrium price vector. Although prices are exogenous to the individual agents (which implies that a microfoundation of price determination is missing), they are endogenous to the economy as a whole.

Walras' successors in the 1930s and 1940s assigned themselves the task of shaping Walras' static general nonmonetary equilibrium model in a way that answers the questions involving the existence, stability and uniqueness of an equilibrium. The contributions by Wald (1936), Von Neumann (1937), Hicks (1939), Samuelson (1941 and 1948) and Leontief (1941) are particularly worth mentioning. The pillar of general equilibrium theory is however the model that is shaped by Arrow, Debreu and McKenzie (henceforth called the ADM model). Arrow and Debreu (1954), Debreu (1959) and McKenzie (1959) gave particular impetus to the development of a sophisticated Walrasian general equilibrium model.

The ADM model is designed to cover the entire (or at least a large part of) the history of the economy. The economy appears as a sequence of states of the world (periods). Economic agents (consumers and producers) decide on their pattern of consumption and production of commodities at the beginning of the time period under consideration. A commodity is characterized by its physical properties, location, date of availability and the state of the environment. At the beginning of the history of the economy each commodity has its own market. Decision making proceeds decentralized. After a price *tâtonnement* all contracting occurs at the beginning of the economy's history. The contracts fix all current and future transactions (recontracting is excluded). Transactions are assumed to be costless. A complete set of conditional future markets permits the insurance of all risk. Agents are not necessarily well-informed about the actual sequence of future states of the world, which determines the course of the economy over time. In the future markets the economic agents can insure mischances through contingent contracting, which implies that recontracting is redundant. An alternative interpretation is that perfect foresight guards the agents against future disappointments (Grandmont 1977, p. 535). This implies that in the course of time incentives to adjust plans, and hence to recontract, are absent.

Not surprisingly, the focus of much subsequent research is on the dynamics of a

sequence of markets.¹ As Weintraub (1979) puts it:

The major lessons of macroeconomics suggest first that the ADM model be extended to the future; questions about uncertainty and intertemporal linkages begin to appear in a natural context (Weintraub 1979, p. 90).

The assumption that markets are open only once is to be dropped. The introduction of a *sequence* of periods (and hence a sequence of markets) implies that the ADM period length is reduced such that the period cannot enclose the entire history of the economy. Economic agents do not decide just once, but time and again at the junction of subsequent periods. Radner (1968, 1974 and 1982) indicates that uncertainty is the very reason for sequential trading (see also Green 1973). Hahn (1971) introduces transaction cost as an explanation of sequential trading in a perfect foresight environment. Hahn (1973), Ostroy (1973), Starrett (1973), Ostroy and Starr (1974) and Kurz (1974) offer transaction cost models of sequence economies (Ulph and Ulph 1975).

20.2.3 Non-Walrasian Equilibrium Models

The central message of Keynes' *The General Theory* (1936) is that the general Walrasian equilibrium is a special rather than a general case. To put it bluntly, market clearing prices only occur by accident.² Keynes' style of writing, however, induced an ongoing debate about the predominant forces that drive his system. On the one hand, Hicks' (1937), Modigliani's (1944) and Hansen's (1953) IS - LM framework integrated Keynes' argument in the existing body of knowledge of macroeconomics such that nonmarket clearing appeared as a special case of the general Walrasian market clearing setting.³ On the other hand, Post-Keynesian authors like Shackle (1972 and 1974), Minsky (1975) and Kregel (1976 and 1982) stressed time and again Keynes' departure from orthodox Walrasian theory by pointing to the overpowering role of fundamental uncertainty.

Patinkin (1956) contributed to the debate by arguing in favor of the foundation of

¹ Two other topics which received considerable attention are the introduction of money and production. Hahn (1965) introduced the well-known problem (that carries his name) of proving the existence of a *monetary* Walrasian equilibrium (for example, Clower 1967 and Grandmont 1983). Stigum (1969a, 1969b and 1972), Arrow and Hahn (1971), Radner (1972) and Sondermann (1974) are examples of general Walrasian equilibrium models with production.

² Keynes introduces his seminal book by asserting that "I shall argue that the postulates of the classical theory are applicable to a special case only and not to the general case, the situation which it assumes being a limiting point of the possible positions of equilibrium" (Keynes 1936, p. 3).

³ Here the argument is that price and wage rigidities underlie Keynesian results. Patinkin's (1956) argument of the real balance effect induced the belief that (in the long run) the convergence to a Walrasian equilibrium is retained (McCallum 1987). Well-known advocates of this line of reasoning in macroeconomics are monetarist economists (Friedman 1968 and Phelps 1970), who initiated the debate about the role of the Phillips curve (as an equilibrium rather than adjustment equation). The monetarism - Keynesianism debate in the 1960s and 1970s took largely place within the macroeconomic IS - LM framework. This means that the analytical structures that both camps employed are not very different (Spence 1974). This brings Friedman (1976) to the assertion that there actually is a nondebate. The difference of opinion follows from empirical rather than theoretical arguments. New classical economics, introducing the rational rather than the adaptive expectations hypothesis in general Walrasian equilibrium models, can be regarded as the microeconomic pendant of monetarism (Sheffrin 1983 and McCallum 1987). New classical economics even postulates short-run convergence to a Walrasian equilibrium (Lucas 1972, Sargent and Wallace 1975 and 1976, Barro 1976, Kydland and Prescott 1977 and McCallum 1980).

Keynesian macroeconomics in microeconomics. The research program that followed upon Patinkin's book argues that Keynesian macroeconomic implications should follow from the rational (i.e., maximizing) behavior of economic agents in a (general) equilibrium framework. Patinkin points out that if commodity markets are expected not to clear (that is, firms are rationed in the commodity market), firms take into account the associated quantity constraints in their decision on labor demand. Clower (1965 and 1967) applies Patinkin's rationing argument to households. Clower works out his proposition by introducing the "dual decision hypothesis" and the associated distinction between notional (Walrasian market clearing) and effective (non-Walrasian rationing) quantities. This hypothesis states that households take into account the quantity constraints they face in the labor market while deciding on demand for commodities.

Two further contributions to non-Walrasian equilibrium theory in the 1960s are worth mentioning: Hicks (1965) and Leijonhufvud (1968).¹ Hicks (1965) reached the methodology of the modeling of market nonclearance that characterizes many of the contributions in the 1970s: the fixprice method. Leijonhufvud (1968) elaborated on the non-Walrasian argument by pointing out that the essential feature of non-Walrasian equilibrium is that quantities move faster than prices. That is, in the short run quantity adjustments dominate price movements. This is essentially an assumption of fixed prices (Drazen 1980, p. 286).

Hicks (1965) elaborates on the temporary equilibrium method by introducing the assumption that diametrically opposes the Walrasian condition of flexible market clearing prices (Hicks 1939). The fixprice method assumes that sticky prices impede the establishment of a temporary Walrasian equilibrium. The essential assumption is that prices in a period are fixed at a predetermined (nonmarket clearing) level, implying that at least one economic agent cannot realize his or her notional intentions. The agent perceives binding quantity restrictions: The excess demand functions depend on prices as well as quantities. However, the fixprice method still implies a temporary equilibrium, since transactions only take place after the announcement of equilibrium quantities (rather than prices). That is, temporary dynamics follows from a *quantity tâtonnement*.

Barro and Grossman (1971, 1974 and 1976) brought Patinkin's firms' and Clower's households' rationing together in one model. Spillovers can run from the labor to the commodity market and *vice versa*. The general perspective of Barro and Grossman's model introduces the opportunity to focus on the macroeconomic implications of microeconomic responses to quantity constraints. Barro and Grossman's setup, in common with most of their immediate successors', is the Hicksian fixprice method. The follow up of Barro and Grossman's temporary non-Walrasian equilibrium model is particularly manifest in Benassy's (1975, 1977 and 1982) and Drèze's (1975) generalizations to a multimarket setting and Malinvaud's (1977) typology of regimes. Malinvaud (1977) elaborates on Barro and Grossman's two-market model by indicating that the features of the temporary equilibrium in the commodity and labor market determine specific macroeconomic regimes, depending on whether buyers or sellers face rationing in the commodity and labor market. Figure 20.1 illustrates Malinvaud's well-known regimes (Malinvaud 1977, p. 31).

¹ Surprisingly, Solow and Stiglitz' (1968) contribution is generally not mentioned by the standard surveys of non-Walrasian equilibrium theory, whereas Glustoff (1968) is only referred to in passing (Grandmont 1977, Weintraub 1979 and Drazen 1980). Nonetheless, both models clearly foreshadow the rigorous treatment of general non-Walrasian equilibria in the 1970s.

Malinvaud amplifies three regimes by arguing that

[w]hen supplies are rationed in both markets, there is unemployment and firms do not produce more because of lack of effective demand: this is the Keynesian case. When labour is not fully employed but firms sell all their supply, we may speak of classical employment. Finally when demand exceeds supply in all markets, there is an inflationary pressure, but the short-run price rigidities are such that inflation is somewhat repressed (Malinvaud 1977, pp. 31-32).

Figure 20.1
Malinvaud's (1977) Regimes

		Commodity Market Rationing	
		Supply	Demand
Labor Market	Supply	Keynesian Un- employment	Classical Un- employment
	Demand	Undercon- sumption*	Repressed Inflation

* Not treated by Malinvaud

Moreover, Muellbauer and Portes (1978) identify the underconsumption case,

where firms would like to trade more than households on both markets (Muellbauer and Portes 1978, reprinted in Branson 1979, p. 357).

A major drawback of the fixprice models is the reliance on the extreme assumption of predetermined sticky prices and wage rates. This means that the macroeconomics the models seek to derive (*i.e.*, Keynesian unemployment), is assumed rather than explained. Therefore, it is crucially important to study to what extent the assumption of rigid prices and wage rates can be rationalized. According to Hahn (1984, pp. 318-321), the microfoundation of price and wage rigidity will inevitably assume market imperfection. The literature suggests that a rationale for wage rigidity can be found in three theories in particular: (i) contract theory; (ii) efficiency wage models; and (iii) insider - outsider arguments. Contract theory (Baily 1974, Okun 1975 and 1981, Azariadis 1975, Kihlström and Laffont 1979 and 1983, Holmström 1983, Azariadis and Stiglitz 1983 and Chari *et al.* 1989) points out that (implicit or explicit) contracting in the labor market can fix wage rates for periods sufficiently long to justify an assumption of wage rigidity. The basic

underlying argument is that workers generally try to reduce the risks they face by signing (long-)lasting agreements (Azariadis and Stiglitz 1983, p. 17). The efficiency wage (Yellen 1984 and Akerlof and Yellen 1986) and insider - outsider theory (Solow 1985 and Lindbeck and Snower 1987 and 1988) are discussed in Chapter 17. The key point is that specific human capital and turnover cost can provide a monopolistic position to incumbent workers. This means that wage rates can be set above the market clearing levels. The fact of the matter is that the three theories introduce market imperfections so as to offer a rationale for nonmarket clearing wage rates; contract theory assumes informational deficiencies, whereas the efficiency wage and insider- outsider models imply imperfect competition with barriers.^{1,2}

The microfoundation of nonmarket clearing prices is generally obtained by introducing imperfect competition.³ The fixprice method is replaced by models in which firms undertake price setting. Arrow (1959) points out that imperfect competition naturally arises outside the Walrasian equilibrium. Following Negishi's (1960) and Arrow and Hahn's (1971) introduction of monopolistic elements in the ADM model, subsequent studies employed monopolistic competition. For example, Benassy (1976 and 1977), Grandmont and Laroque (1976), Hahn (1977 and 1978), Gale (1978) and Weitzman (1982) have succeeded in proving the existence of temporary non-Walrasian equilibria in a general equilibrium setting with price setting agents. The basic argument is that economic agents respond to (stochastic demand) disturbances by adjusting prices to an insufficient degree or insufficiently fast to reach the re-establishment of the market clearing equilibrium, which implies that quantities take part in the burden of adjustment.⁴

20.2.4 Dynamic Competition

To clarify Chapter 20's contribution to the debate, special reference is made to the

¹ Two other candidates that can offer a microfoundation of nonmarket clearing wage rates, are the internal labor market framework (Creedy and Whitfield 1988) and Akerlof's (1980) norm theory.

² In the literature on general equilibrium theory imperfect competition is generally regarded as a market imperfection (which simply means that economic agents act as price setters); that is, perfect competition is used as the benchmark case for perfection. In the literature on industrial organization this assumption is questioned (Part III). In many cases it is imperfect competition that is preferred to perfect competition.

³ Carlton (1989) reviews this literature (by presenting both empirical evidence and theoretical arguments) and evaluates the (possible) contribution of industrial organization to the explanation of (non)market clearing. He pleads in favor of the argument that *nonprice* allocation mechanisms must be taken into consideration (pp. 938-940).

⁴ Strictly speaking, an assumption of imperfect competition alone cannot guarantee that firms set nonmarket clearing prices (Chamberlin 1933). For illustrative purposes, three (related) factors that facilitate the establishment of nonmarket clearing prices can be mentioned: (i) conjectures (Hahn 1977 and 1978 and Neary and Stiglitz 1983); (ii) informational deficiencies (Iwai 1981, Negishi 1979, Ito and Honkapohja 1985 and Greenwald and Stiglitz 1989); and (iii) adjustment costs (Blinder 1982, Amihud and Mendelson 1983 and Ball *et al.* 1988). One important result of the recent literature on non-Walrasian economics cannot be ignored: *small microeconomic deviations from perfection can yield large macroeconomic departures from the Walrasian market clearing equilibrium* (Akerlof and Yellen 1985 and 1987, Mankiw 1985, Blanchard and Kiyotaki 1987 and Ball *et al.* 1988). Cooper and John (1988) offer a framework that seeks to unify the literature on non-Walrasian economics by shaping the argument in terms of game theory: It is the strategic complementarities in the economic agents' payoff functions (in combination with spillover and multiplier effects) that underlie the Keynesian features of non-Walrasian equilibrium models.

"Cournot-Marshallian-ADM synthesis" of Novshek and Sonnenschein (1987).¹ A lengthy quotation illustrates the features of this synthesis:

The Marshallian and ADM theories have many striking and essential differences, for example:

1. The ADM theory specifies a fixed finite number of firms. The Marshallian theory postulates a pool of firms, any number of which may be active in the market.
2. The ADM theory postulates that technology of each firm is convex, which rules out increasing returns to scale. The Marshallian theory postulates that the average cost function of each firm is U-shaped, based on the assumption of fixed costs of production or regions of increasing returns to scale.
3. The ADM theory assumes price-taking behavior regardless of the number of firms. The Marshallian theory assumes it only if the efficient scale is small relative to demand.
4. ADM theory, a general equilibrium theory, relates perfect competition and economic efficiency. Most of Marshallian analysis ignores intermarket effects.
5. Finally, the ADM theory is a static one, to which the adjoiment of dynamics via tâtonnement is not very satisfactory. The Marshallian analysis of equilibria is a dynamic one in the sense that entry and exit of firms cease in equilibria (Novshek and Sonnenschein 1987, pp. 1281-1282).

Novshek and Sonnenschein's synthesis seeks to capture the best of both the ADM and Marshallian world. That is to say, market structure is endogenized (1), technology is nonconvex (2), ("as if") price taking behavior is explained (3), intermarket (price and spillover) effects are taken into account (4), and dynamic entry and exit enhance coordination (5). As modern non-Walrasian equilibrium models take notice of price setting and intermarket spillovers, the key departure from (non-)Walrasian models lies in point (1), (2) and (5). The framework presented below shows the same features. A key departure from non-Walrasian equilibrium theory is that vertical intermarket spillovers through rationing are ignored. Instead, horizontal intermarket spillovers through entry and exit movements are taken into account. Although these vertical externalities can be incorporated in the framework, for simplicity reasons this is not yet done in the current analysis.

The framework of this chapter deviates from Novshek and Sonnenschein's synthesis in two essential ways. First, Novshek and Sonnenschein's results resemble the ADM model for the restrictive case of perfect competition. They

define perfectly competitive equilibrium as the limit of Cournot equilibrium with entry as firms become small relative to the market (Novshek and Sonnenschein 1987, pp. 1288-1289).

In this chapter, however, the results can be *independent* of market structure (that is,

¹ This work fits into the tradition that focuses on the relationship between Cournot-Nash equilibria and general Walrasian equilibrium theory. Further references are Gabszewicz and Vial (1972), Novshek and Sonnenschein (1978 and 1983), Mas-Colell (1980, 1983 and 1986), Novshek (1980), Ushio (1983) and Guesnerie and Hart (1985). Attention is not only directed to the case with free entry, but also to multimarket settings with entry barriers. The latter models introduce imperfect competition. The dominant conclusion of this literature is, however, that "[i]f firms face a fixed set-up cost and a constant marginal cost and there is free entry, then, if the economy is sufficiently large, a Cournot-Nash equilibrium exists and in such an equilibrium, price will be approximately equal to marginal cost" (Guesnerie and Hart 1985, p. 541). In the main text the focus is on Novshek and Sonnenschein (1987) as a representative example of this line of modeling.

number and size of firms) by exploring the implications of the *Bertrand* analogue of Novshek and Sonnenschein's Cournot assumption. This is done by exploring a commodity-labor barrier market setting as a benchmark case. Not surprisingly, it appears that cost and demand features and the number of suppliers are crucial determinants of the type of competition that rules the market. What matters is the credibility of the entry threat. Second, in addition to price or quantity competition in Baumol *et al.*'s contestable market theory and Novshek and Sonnenschein's Cournot-Marshallian-ADM synthesis, the barrier market theory implies that a major competitive weapon of firms (and laborers) is to undertake *sunk* (R&D) investment. Hence, investment in production technology by firms and in human capital by laborers are *endogenously* determined.¹

20.3 ASSOCIATION OF TECHNOLOGY AND HUMAN CAPITAL

Suppose that the workers' innovation induces an *incremental* reduction of the firms' average total cost. This means that the impact of laborers' attractiveness simply comes in through *efficiency* considerations. The improved human capital features allow production to proceed more efficiently without, however, reducing labor input per unit of output. Here it is particularly plausible that there can be reduced X-inefficiency and/or less waste of nonlabor inputs (e.g., materials and energy).² The wedge between the value of the marginal product and the wage rate determines the magnitude of the incremental labor induced efficiency cost reduction ($\Delta^L \text{atc}$):

$$(20.1) \quad \Delta^L \text{atc}_1 = \text{LR}(\text{pwg}_1), \text{ with}$$

$$(20.1a) \quad \partial \text{LR} / \partial \text{pwg}_1 > 0,$$

where LR is twice continuously differentiable. Of course, the rise of the wage rate associated with a human capital improvement (Δw_1) induces, *ceteris paribus*, an increase of the firms' cost level. However, equation (20.1) points to the assumption that the increased value of the marginal product (Δmp_1) more than offsets this effect via efficiency improvements.

Hence, given pwg_1 a firm's average total cost reduction ($\Delta \text{atc} \mid \text{pwg}_1$) proceeds through two steps. First, a firm reaches a cost reduction as a result of its own investment efforts in period $t = 0$ [equation (15.6)], given pwg_0 [$(\Delta^F \text{atc}_1 \mid \text{pwg}_0) = \Delta \text{avc}_1 + \mu u_1$]. Second, an additional cost reduction emerges through improved efficiency in period $t = 1$ as a result of the laborers' schooling [equation (18.7) and Propositions 18.1 and 18.2],

¹ Chatterjee and Cooper (1989) offer an overlapping generations model with entry and exit. However, they exclude "consideration of potentially interesting dynamic strategic interactions" (p. 354). It is therefore not surprising to note their observation that "amending the model to allow for some dynamic strategic considerations and learning-by-doing might generate some additional insights" (p. 357).

² In practice, effects of human capital improvements can enter the picture in, at least, two further ways: (i) human capital improvements can introduce *labor augmented* cost reductions (Hicks 1932; see also Chapter 8), so that a *quantitative* change in labor demand occurs; and (ii) process innovations are associated with a *binding* demand for a *particular* (range of) type(s) of human capital improvement(s). This implies that the situation may arise where unemployment and unfilled vacancies (U and V) exist *simultaneously*. Particularly in Europe, U-V analysis still enjoys considerable popularity to treat the concept of frictional unemployment (Hansen 1970 and Haveman 1978). However, the empirical establishment of U-V curves is in need of an appropriate theoretical underpinning. Here human capital considerations may play a significant role (Bierings and Van Witteloostuijn 1989). For the sake of simplicity, Chapter 20 ignores these complications.

associated with pwg_1 ($\Delta^L \text{atc}_1 \mid \text{pwg}_1$). Of course, the entry-deterring price level is $p_1 = \text{atc}_1$ (Proposition 15.2).¹

The association between firms' process innovations and laborers' improvements of human capital permits a first analysis of commodity-labor market interactions. It appears that the implications of rationing for the credibility of the entry threat and market outcomes depend on the answers to, at least, the following three questions:

- (i) Does rationing take place in the commodity and/or labor market?
- (ii) Which side of the market is being rationed?
- (iii) Is there heterogeneity of firms and/or workers?

The point of departure in Section 20.4 lies in answering the third question in the negative: both firms and laborers are assumed to be homogeneous. Sections 20.5 to 20.7 examine the implications of dropping this assumption.

20.4 HOMOGENEOUS FIRMS AND LABORERS

The literature on the microfoundations of macroeconomics points to intermarket spillover effects from the commodity to the labor market and *vice versa* (Patinkin 1956, Clower 1965 and Barro and Grossman 1976) that can be characterized with the help of particular regimes (Malinvaud 1977 and Muellbauer and Portes 1978). Although, of course, a microfoundation of macroeconomics requires the modeling of (profit and utility) maximizing frameworks in which rationing plays a role, here it suffices to introduce the overall regimes. (Figure 20.1 illustrates Malinvaud's and Muellbauer and Portes' well-known regimes.) Homogeneity of firms and laborers means that all firms yield the same average total cost level and that all workers offer the same wedge between the value of the marginal product and the wage rate.

If firms and laborers are assumed to be homogeneous, then Malinvaud's and Muellbauer and Portes' classification applies. Then, four logical possibilities of (isolated) rationing arise: (i) firms' rationing in the labor market; (ii) laborers' rationing in the labor market; (iii) firms' rationing in the commodity market; and (iv) laborers' rationing in the commodity market. In this section the four cases of rationing are examined under *ceteris paribus* conditions. It is assumed that no further rationing prevails, so that the other market is faced with an effective entry threat *ex post*. Henceforth, for the sake of brevity, *ex ante* refers to period $t = 0$ and *ex post* indicates period $t = 1$. This means that vertical intermarket rationing spillovers are ignored.

Suppose that firms and laborers have decided on innovative expenditures, wage rates, prices and output in period $t = 0$ in the way described in Chapters 15 and 18 respectively. Thus, in period $t = 0$ the following holds: $F^N > F^{ni}_0$ and $L^N > L^{ni}_0$, where F^N and L^N denote the total number of firms and laborers respectively: $F^N = F^{ni} + F^{ne}$ and $L^N = L^{ni} + L^{ne}$. The entry threat was effective *ex ante*, because there were

¹ If demand is price elastic, then a monopolist faces a demand constraint $\mu_1 = sc_0/d_1(p_1)$. Hence, in order to calculate the (expected) optimal investment efforts in period $t = 0$, a monopolist must take into account the implications in period $t = 1$ of (expected) laborers' schooling, because pwg_1 enters Δatc_1 and so p_1 . The recursive influence of p_1 is no obstacle for proving that a unique interior solution to the firms' investment-price-output decision exists (Chapter 13).

equally equipped potential competitors waiting for an opportunity to enter.¹ In period $t = 1$, however, rationing may occur (unless indicated otherwise; assume that F^N and L^N remain constant over time).

Broadly speaking, in the labor market eight regimes can occur. Figure 20.2 depicts the eight cases for the assumption that the standard working week $L = 1$.

Figure 20.2
Regimes in the Labor Market with Homogeneity

This assumption simplifies matters by excluding mixed scenarios with partial entry and exit. S and D denote aggregate supply and demand and superscript C and L indicate commodity and labor respectively. Rationing of labor in the labor market means that $L^N > L^D_1$ (L^D_I , L^D_{II} , L^D_{III} and L^D_{IV}). Note the metaphorical nature of L^D_I , L^D_{IV} and L^D_V as opposed to the precise nature of L^D_{II} and L^D_{III} . Unemployment of labor is associated with an effective entry threat *ex post*, because $L^N > L^{ni}_1$. Without additional information it is not possible to determine whether the labor market configuration is intertemporally sustainable. If $L^D_1 = L^D_0$, then sustainability prevails; otherwise, entry (if $L^D_1 > L^D_0$) or exit (if $L^D_1 < L^D_0$) movements introduce unsustainability. Rationing of firms in the labor market is associated with $L^N < L^D$ (L^D_V). Firms are unable to hire all the laborers they want. Shortage of labor makes the entry threat *ex post* ineffective in the labor market, simply because there are no potential entrants to fear. Entry has occurred such that $L^N - L^{ni}_1 = 0$ (or $L^{ni}_1 = L^N$), so that the labor market shows intertemporal

¹ In effect, $L^N > L^{ni}_0$ means that in period $t = 0$ there were laborers ($L^N - L^{ni}_0$) rationed in the labor market. If emigration is ignored, this means that in period $t = 0$ unemployment flourished. Equivalently, it can be argued that $F^N > F^{ni}_0$ points to firms' rationing in the commodity market, unless $F^N - F^{ni}_0$ contains firms in (nearby or related) industries which are waiting for an opportunity to cross-entry (probably by way of expansion).

unsustainability.¹ The labor market is characterized by blockaded entry *ex post*.

Subtle arguments underlie the implications of a credible entry threat *ex post* for the equilibrium wage rate. The key point is that the opportunity cost of entry matters, which points to the importance of the source of potential entry (Chapter 11). For the sake of brevity, only two cases are considered: potential entry either ensues from unemployed workers who have mirrored the incumbents' schooling investment in anticipation of future entry opportunities² (supply curve L^S_A) or comes from potential entrants who are employed in (nearby or related) barrier labor markets (supply curve L^S_B). These two cases already permit the characterization of eight regimes (R_i , where $i = 1, \dots, 8$). A regime can be defined according to assumptions of the source of potential entry, exit opportunities, location of demand and/or wage rate (in)flexibility.

Exit into barrier markets (R1). Demand has declined to L^{D^I} or remained at $L^{D^{II}}$, while the entry threat comes from barrier labor markets. The location of demand predicts exit as $L^{D^I}_1 < L^{ni}_0$ for $w_1 = w_1^*$. If the expelled incumbent workers can enter nearby perfect barrier markets, the entry-detering wage rate follows from $w_1 = w_1^*$ (as $L^{ni}_1 < L^{ni}_0 < L^N$). Hence, the equilibrium outcome is associated with a zero-returns wage rate.

Exit into unemployment (R2). The demand location L^{D^I} is associated with excess supply ($L^{D^I}_1 < L^{ni}_0$) if $w_1 = w_{\min}$. This is the relevant benchmark if expelled incumbent workers are not able to enter other markets: the inevitable consequence of exit is unemployment. The entry threat is fortified ($L^{ni}_1 < L^{ni}_0 < L^N$). In this case contestability predicts convergence to the reservation wage rate.

Contestable sustainability (R3). Expelled incumbent workers cannot enter into other markets. This is the case if, for example, potential entrants are unemployed. The demand location $L^{D^{II}}$ permits sustainability ($L^{D^I}_1 = L^{ni}_0$) if $w_1 = w_{\min}$. Contestability theory predicts convergence to the reservation wage rate [as the entry threat is still effective ($L^{ni}_0 = L^{ni}_1 < L^N$)].

Barrier market sustainability (R4). The increase in demand to $L^{D^{III}}$ precisely allows incumbent workers to be employed against $w_1 = w_1^*$, while the entry threat comes from workers in perfect barrier markets ($L^{ni}_0 = L^{ni}_1 < L^N$). This means that the entry-detering wage rate is w_1^* .

Entry by unemployed workers (R5). Demand has increased to $L^{D^{III}}$ or $L^{D^{IV}}$ such that incumbent laborers are not able to satisfy market demand ($L^{D^I}_1 > L^{ni}_0$) if $w_1 = w_{\min}$.

¹ Recall that (i) the definition of intertemporal sustainability is presented in Chapter 6 and (ii) a sustainability analysis of the barrier market is offered in Chapter 13.

² Recall that unemployed potential laborers who have invested sc_0 *ex ante* face the following payoff *ex post* if they do not enter: $\pi_1 = u_i - sc_0$ [equation (17.4)]. It is now easy to see that unemployed potential entrants prefer to enter for any $w_{\min} < w_1 < w_1^*$ (where $w_1^* = w_0 + sc_0$ is the zero-returns wage rate). Hence, the incumbents' entry-detering wage rate demand amounts to $w_1 = w_{\min}$, which is associated with negative net returns to the human capital improvement (conditional upon, of course, $w_{\min} < w_1^*$). For further discussion, see Section 20.7. This result follows from the observation that the sunk schooling cost incurred in period $t = 0$ is ancient history in period $t = 1$. If there are equally equipped unemployed potential entrants, then the market is, *ceteris paribus*, perfectly contestable in period $t = 1$! Wage rates are then deflated to their reservation level (Proposition 17.1).

Entry by unemployed (but schooled) potential workers reduces the equilibrium wage rate to the reservation level as the *ex post* effective entry threat ($L_{ni_0} < L_{ni_1} < L_N$) predicts that only $w_1 = \min w$ can prevent exit.

Entry from barrier labor markets (R6). The increase in demand to L^{DV} is associated with excess demand ($L_{D_1} > L_{ni_0}$) so that entry is invited. Entry from perfect barrier labor markets restores equilibrium at the exit-avoiding zero-returns wage rate [as the entry threat is effective *ex post* ($L_{ni_0} < L_{ni_1} < L_N$)].

Non-Walrasian wage rigidity (R7). Demand has increased to L^{DV} beyond L_N for $w_1 = w_1^*$. The non-Walrasian assumption of wage rigidity (the wage rate being fixed at the zero-returns level) implies that firms are rationed in the labor market. Firms lack incentives to compete for the scarce labor inputs, because this will only generate cost disadvantages (through higher wage rates). Excess demand (L^{ED}) remains unsatisfied as the wage rate is not adjusted upwards. Entry is insufficient to keep in pace with increasing demand ($L_{ni_0} < L_{ni_1} < L_{D_1}$).

Figure 20.3
Regimes in the Commodity Market with Homogeneity

Walrasian wage rate flexibility (R8). The increase in demand to L^{DV} induces excess demand at the zero-returns wage rate. If wage rates are not inert (which is a Walrasian rather than non-Walrasian assumption), workers can benefit from their monopoly power by increasing the wage rate to a higher level (for example, the Walrasian market clearing rate). At this point the labor supply curve may take two shapes, depending upon whether labor supply is rigid (L^S with market clearing wage rate $BE w_1$) or flexible [with normal good features (L^S with market clearing wage rate $BE w_1'$)] beyond L_N (in Figure 20.2 flexibility is represented "as if" more economic agents decide to offer their labor services

in the market.)

Note that *price* (or, in the case of the labor market, wage rate) adjustments may set in only after the pool of entrants has been exhausted: up to this point *quantity* adjustments (through entry and exit movements) coordinate market trade. So, up to L^N the supply curve is horizontal, whereas beyond L^N the supply schedule takes the usual increasing shape.

Broadly speaking, in the commodity market too eight regimes can occur (Figure 20.3, where Q denotes aggregate supply).

The regimes in the commodity market resemble those in the labor market. If $q_1 \cdot F^N > C_{D_1} (L^{DI}, L^{DII}, L^{DIII}$ and $L^{DIV})$ firms are rationed in the commodity market ($F^N > F_{ni_1}$), whereas $q_1 \cdot F^N < C_{D_1} (L^{DV})$ can indicate rationing of labor in the commodity market ($F^N < F_{ni_1}$). Note, again, the metaphorical nature of L^{DI} , L^{DIV} and L^{DV} . The regimes' equilibrium outcomes depend upon assumptions with regard to the source of potential entry, exit opportunities, location of demand and/or price (in)flexibility. For the sake of illustration, two sources of potential entry are considered: (i) potential entrants are new firms which have anticipated future entry opportunities by undertaking R&D (supply curve CS^A)¹ or (ii) potential entry ensues from nearby or related perfect barrier commodity markets (supply curve CS^B).

Exit into barrier markets (R1). Demand has declined to C_{DI} or remained at C_{DII} . Exit, and so intertemporal unsustainability, occurs in order to retain an equilibrium number of firms in the market (as $q_1 \cdot F_{ni_0} > C_{D_1}$). Exit enhances the free entry *ex post* feature of the commodity market, since the number of equally equipped potential entrants increases ($F_{ni_1} < F_{ni_0} < F^N$). If expelled firms can enter into other perfect barrier markets, entry deterrence requires average total cost pricing if the entry threat comes from nearby or related barrier markets.

Average variable cost pricing (R2). The average variable cost price gives excess supply ($q_1 \cdot F_{ni_0} > C_{D_1}$) if demand is located at C_{DI} . Expelled incumbent firms and new firm potential entrants are willing to accept prices *below* the average total cost price p_1 (at the limit down to avc_1) if they are *not* able to recover elsewhere the sunk costs incurred. Contestability indicates that the entry threat from idle capacity ($F_{ni_1} < F_{ni_0} < F^N$) induces exit-avoiding average variable cost pricing.

Contestable sustainability (R3). Potential entry ensues from new firms, whereas expelled incumbent firms cannot enter any other market. Demand C_{DII} is associated with sustainability for $p_1 = avc_1$ ($q_1 \cdot F_{ni_0} = C_{D_1}$). The new firm entry threat is effective ($F_{ni_1} = F_{ni_0} < F^N$). Contestability gives entry-detering average variable cost prices.

Barrier market sustainability (R4). The increase in demand to C_{DIII} precisely permits incumbent firms to stay in the market for $p_1 = p_1$ if potential entry derives from adjacent perfect barrier markets ($F_{ni_1} = F_{ni_0} < F^N$). Perfect barrier market theory predicts the entry-detering average total cost price in equilibrium.

¹ Recall that these firms are inclined to enter at any price above average variable cost avc_1 as this price permits the recovery of (part of) the sunk cost (Section 11.2).

Entry by new firms (R5). Incumbent firms cannot serve increased demand C_{D1}^{III} or C_{D1}^{IV} ($q_1 \cdot F_{ni0} < C_{D1}$) without inducing entry by new firms. New firm entry reduces price to average variable cost, since the entry threat by new firms is effective *ex post* ($F_{ni0} < F_{ni1} < F_N$). Only $p_1 = avc_1$ can avoid exit.

Entry from barrier commodity markets (R6). Demand has increased to C_{D1}^{IV} . The entry threat comes from adjacent barrier markets. Excess demand ($q_1 \cdot F_{ni0} < C_{D1}$) invites entry. Perfect barrier market entry restores equilibrium against the average total cost price, since the entry threat is effective *ex post* ($F_{ni0} < F_{ni1} < F_N$).

Non-Walrasian price rigidity (R7). The increase in demand to C_{D1}^V has moved beyond $q_1 \cdot F_N$: the entry threat is ineffective *ex post*. A non-Walrasian assumption of price inertia (price being fixed at the average total cost level) implies that workers are rationed in the commodity market (excess demand C_{ED}). Entry is insufficient to retain market clearance ($q_1 \cdot F_N < C_{D1}$).

Walrasian price flexibility (R8). The pool of potential entrants contains an insufficient number of potential entrants to satisfy residual demand after the increase in demand to C_{D1}^V ($q_1 \cdot F_N < C_{D1}$). The commodity market is imperfectly competitive *ex post*. With a Walrasian assumption of price flexibility imperfect competition is likely to be associated with a price increase. At the limit this gives the market clearing level ($^{BE}p_1$) with an inelastic supply curve (CS). Alternatively, with blockaded entry firms can decide to supply beyond q_1 without provoking entry (CS), which may give $^{BE}p_1$.

By way of conclusion it is argued that an effective entry threat *ex post* in the labor ($F_N > F_{ni1}$) and commodity ($^L N > ^L ni_1$) markets is *incompatible* with rationing of firms in the labor market and rationing of laborers in the commodity market respectively. The reason is that the rationing potential induces equilibrium-restoring entry. This implies that rationing of firms in the commodity market and rationing of laborers in the labor market is (likely to be) associated with intertemporal unsustainability of market configurations. Firms' rationing in the commodity market induces a tendency to increased *concentration*, simply because it is likely (in particular for $sc_0 > 0$) that the equilibrium number of firms decreases ($F_{ni1} < F_{ni0}$). Laborers' rationing in the labor market implies that the market is perfectly *contestable*, which is associated with the deflation of the wage rate to its reservation level. It may be argued that entry and exit movements (partly) take over the co-ordinating role of the auctioneer in contestable (Jacquemin 1987) and barrier markets.¹ Of course, the list of scenarios in this section is not exhaustive. However, it points to the fact that the market features which are central in this thesis - conditions of entry and exit in dynamic competition - can be important determinants of the equilibrium outcomes in an economy where the interaction between the labor and the commodity markets cannot be ignored.

20.5 HETEROGENEITY OF FIRMS AND PRICE EQUALIZATION

Since the model in Chapter 15 endogenizes process innovations, heterogeneity of firms is

¹ So is the opposite case of firms' rationing in the commodity market, growing demand, likely to be associated with entry [if $D(p_1)/F_{ni0} > q_1$ and $F_N > F_{ni1}$].

traced to diverging cost levels. Assume that ${}^1\text{matc} < {}^2\text{matc} < {}^3\text{matc}$, etc.. Heterogeneity of firms may change the results significantly when rationing of labor in the commodity market occurs. Assume again that decisions on investment, price and output in period $t = 0$ were made on the basis of an effective entry threat *ex ante*. For the moment, it suffices to discuss the following case: $F_{ni0} < {}^1N_0$, ${}^1N_1 < F_{ni1}$ and ${}^1N_1 + {}^2N_1 > F_{ni1}$ ($F_N = {}^1N + {}^2N + \dots$). The condition that $F_{ni0} < {}^1N_0$ implies that there was an effective entry threat *ex ante* such that market demand did only allow superior firms to be incumbent. However, ${}^1N_1 < F_{ni1}$ shows that in period $t = 1$ the number of superior potential entrants is insufficient to satisfy aggregate demand. Hence, there are entry opportunities for inferior firms. However, the inferior entrants face an effective entry threat *ex post*, because ${}^1N_1 + {}^2N_1 > F_{ni1}$. For the sake of convenience, the implications of firms' heterogeneity are only explored for the case where potential entry derives from adjacent perfect barrier markets.

Superior firms are able to set prices such that ${}^1p_1 = {}^2\text{matc}_1 > {}^1\text{matc}_1$ without evoking exit, simply because there are no equally equipped potential competitors waiting for profitable entry opportunities. Profitable entry of "second-best" potential competitors in order to push aside a superior incumbent firm requires ${}^2p_1 < {}^2\text{matc}_1$. Hence, it is likely that superior firms perform (profit-maximizing) limit pricing (Chapter 7), which leads to ${}^1p_1 = {}^2\text{matc}_1$ and associated supply 1q_1 .¹ If a superior firm were to set ${}^1p_1 > {}^2\text{matc}_1$, then a "second-best" potential entrant would be able to push aside the incumbent firm profitably by offering ${}^2\text{matc}_1 < {}^2p_1 < {}^1p_1$. However, despite the exit-avoiding limit price strategy of the superior firms, entry occurs, because by assumption ${}^1N_1 \cdot {}^1q_1 < D({}^2\text{matc}_1)$. Inferior " ${}^2\text{matc}_1$ " potential competitors enter the market in order to serve residual demand. The inferior entrants adopt an entry-detering strategy, because there are equally equipped potential competitors (${}^1N_1 + {}^2N_1 > F_{ni1}$ by assumption). Hence, prices are set such that ${}^2p_1 = {}^2\text{matc}_1$.

Heterogeneity of firms is associated with an inflation of the price level that partly offsets the deflation associated with the process innovations of the superior firms. (Of course, the opposite case of declining demand can be associated with exit movements that lead to a further deflation of the price level.) Superior firms adapt their price level to the average total cost price of the inferior entrants. Equivalently, the equilibrium price is ${}^3p_1 = {}^3\text{matc}_1$ if ${}^1N_1 + {}^2N_1 < F_{ni1}$ and ${}^1N_1 + {}^2N_1 + {}^3N_1 > F_{ni1}$, and so forth. Superior firms are able to benefit from a cost advantage in line with the predictions of the literature on limit prices [Chapter 7; see also Silberberg's competitive equilibrium with nonidentical firms (1974, p. 740)]. The cost differentials ${}^1\text{matc}_1 - {}^2\text{matc}_1$, ${}^2\text{matc}_1 - {}^3\text{matc}_1$, ${}^1\text{matc}_1 - {}^3\text{matc}_1$, etc., reflect *entry barriers* (Bain 1956, Stigler 1968, Baumol *et al.* 1982a and Stiglitz 1987).

¹ Here it is tacitly assumed that ${}^1p_1 = {}^2\text{matc}_1$ is associated with profit maximization. In fact, however, subtle arguments enter the superior firms' decision making. If, on the one hand, a superior firm correctly foresees the type and implications of the regime in period $t = 1$, then the decision on R&D efforts in period $t = 0$ is used to *anticipate* the maximization of period $t = 1$'s profits. This means that 1sc_0 can be chosen such that ${}^1\pi_1 = ({}^1p_1 - {}^1\text{atc}_1) \cdot {}^1q_1$ is maximized subject to ${}^1p_1 \leq {}^2\text{matc}_1$, where ${}^1\text{atc}_1 = {}^1A_0({}^1q_1) - R({}^1sc_0) + {}^1sc_0/{}^1q_1 - LR(pw_1)$ and ${}^1q_1 = D({}^1p_1)/{}^1N_1$. If, on the other hand, a superior firm does not correctly foresee that it will face imperfect competition in period $t = 1$, then the decision on R&D efforts in period $t = 0$ follows from the entry-detering decision rule as described in Chapter 15. Then, the superior firm does *not* anticipate the opportunity to capture positive profits in period $t = 1$, but instead maximizes profits in period $t = 1$ on the basis of a given cost function ${}^1\text{atc}_1$. In the text it is simply assumed that the corner solution (${}^1p_1 = {}^2\text{matc}_1$) solves either maximization procedure. If, however, the optimal solution is a ${}^1p_1 < {}^2\text{matc}_1$, then *price heterogeneity* results, although a *tendency* to price equalization is sustained.

The implication of firms' heterogeneity is that the supply curve takes the form of a stair which climbs from the left-hand to the right-hand side. Figure 20.4 depicts the dichotomy case, where ${}^i_*Q_1 = {}^i_*q_1 \cdot {}^iN_1$ for $i = 1, 2$ and the blockaded entry scenario resembles regime R8 in Figure 20.3.

The conclusion is, again, that rationing of labor in the commodity market is *incompatible* with an effective entry threat *ex post*. Initial laborers' rationing in the commodity market sets in motion entry movements of successive inferior firms, which is associated with countervailing price increases. Figure 20.4 describes three cases as regards to different locations of demand (C_D^I , C_D^{II} and C_D^{III} respectively). Entry and exit movements show a coordinating role in establishing (temporary) equilibrium after a demand shift. Moreover, in addition to the exit-avoiding and entry-detering motive, the positive profits captured by superior firms (following the creation of entry barriers) introduce an *extra incentive* to carry out (expected individually) optimal R&D programs.

Figure 20.4
Regimes in the Commodity Market with Heterogeneity

A remark is in order with respect to the (common) assumption that potential entrants only enter for a ${}^e p_1 > {}^i e matc_1$ (Chapter 11). That is to say, it is supposed that potential entrants face a zero opportunity cost of entry. With binding capacity constraints this means that they can only capture a zero profit in a home market (${}^e p_1 = {}^i e matc_1$). Dropping this assumption may change the results in two ways (Chapter 11). First, if there are some potential entrants around receiving ${}^e p_1 < {}^i e matc_1$, then these potential entrants are inclined to enter for any ${}^i p_1 > {}^e p_1$. Hence, the incumbents' entry-detering price level is ${}^i p_1 = {}^e p_1 < {}^i e matc_1$, which is associated with *negative* profits. Second, if all potential entrants capture an ${}^e p_1 > {}^i e matc_1$ in home markets, then the entry-detering price yields *positive* profits. In both cases there are repercussions to the co-existing markets in which the superior firms operate. Here these complications are ignored. Section 20.7 offers a

treatment in the context of the labor market. For the moment the assumption is that all potential entrants operate in perfect barrier markets.

20.6 HETEROGENEITY OF LABORERS, RIGID WAGE RATES AND REGIMES

Suppose that laborers are heterogeneous in the sense that there are groups of workers offering *different* $mp - w$ gaps. Let 1pwg denote the largest $mp - w$ differential, 2pwg the second largest, *etc.* Of course, firms preferences amount to ${}^1pwg > {}^2pwg > {}^3pwg > \dots$. Assume that laborers in period $t = 0$ have decided on the basis of an effective entry threat *ex ante*. For the moment, it suffices to focus on a ${}^1pwg - {}^2pwg$ dichotomy. Figures 20.5 and 20.6 describe the regimes that now may occur.

Figure 20.5
Heterogeneity of Laborers and Regimes

Regime	Type of Laborers	
	1pwg	2pwg
1	$S > D$	$S > D = 0$
2	$S = D$	$S > D = 0$
3	$S < D$	$S > D$
4	$S < D$	$S = D$
5	$S < D$	$S < D$

(Recall the assumption that ${}^iL_1 = 1$ for $i = 1, 2$.) Of course, subsequent heterogeneity multiplies the number of regimes. The heterogeneity of laborers introduces the possibility that *differentiated* (un)employment occurs. The implications of differentiated employment depend on whether wage rate adjustments occur. Here the analysis is restricted to the cases where neither laborers nor firms induce adjustments of the wage rates. This situation corresponds to an assumption of *rigid wage rates*, as employed in models with temporary non-Walrasian equilibria (fixprice models). In effect, it is assumed that the wage rates are fixed at their zero-returns level ($w_1 = w_0 + sc_0$).

Regimes 1 (${}^LD^1$) and 2 (${}^LD^2$) are associated with a sufficient number of superior workers to satisfy firms' demand for labor. Regime 1 reflects the case where the threat of entry is effective *ex post*, whereas regime 2 shows full entry of superior laborers. In the latter case, superior laborers face imperfect competition *ex post*, since there is no threat of entry by equally equipped potential competitors. However, with fixed wage rates superior workers cannot exploit the opportunities offered by the effective entry barriers (Section 20.7 describes the flexprice case). In the absence of adjustments of the wage rate, inferior laborers cannot enter the market, because firms prefer to hire " 1pwg_1 "

workers. In effect, hiring inferior laborers constitutes a self-destructive strategy. If an incumbent firm hires " ${}^2\text{pwg}_1$ " laborers, then a potential entrant is able to undercut the incumbent firm's average total cost price by hiring " ${}^1\text{pwg}_1$ " workers. The potential entrant is then able to benefit from a cost advantage. Hence, regimes 1 and 2 show an employment-unemployment dichotomy, since only superior workers are employed to the detriment of inferior laborers.

Figure 20.6
Regimes in the Labor Market with Heterogeneity

Regimes 3 to 5 (LD^3 to LD^5) are associated with a shortage of superior laborers. Hence, firms are forced to hire inferior workers, in this way introducing differentiated (un)employment. All superior workers are hired, whereas inferior laborers may be confronted with unemployment (regime 3). Regimes 4 and 5 describe full employment, but only in regime 5 are firms rationed in the labor market. Superior laborers are confronted with imperfect competition *ex post*. The same holds for the inferior workers in regimes 4 or 5. If the two types of labor are allocated *proportionally* to the incumbent firms, then price homogeneity results, because all firms benefit from the same cost advantages. However, if the two types of labor are distributed *unequally* among firms, then *price heterogeneity* is inevitable. Firms' " ${}^1\text{pwg}_1/{}^2\text{pwg}_1$ " labor ratios diverge, so that cost differentials (adapted to average total cost pricing) introduce price heterogeneity. However, as long as all superior workers are employed, no potential entrant can show a better performance by hiring other laborers. Moreover, there is actually no incentive for employed workers to switch from one firm to another when wage rate adjustments are conspicuous by their absence. In the absence of adjustments of the wage rates, price heterogeneity is, *ceteris paribus*, compatible with sustainability of the commodity market configuration, conditional upon the assumption that all superior workers are employed.

20.7 TENDENCY TO INCREASED WAGE RATE DISPERSION

If wage rate adjustments are permitted, the implications of differentiated employment change radically, since, on the one hand, superior workers are inclined to benefit from their higher efficiency, whereas, on the other, unemployed laborers are willing to accept lower wage rates. Again, it suffices to focus on the ${}^1\text{pwg} - {}^2\text{pwg}$ dichotomy case. In period $t = 0$ market participants have decided on investment efforts on the basis of an effective entry threat *ex ante*. For the sake of simplicity of notation, henceforth time indices are suppressed.

The basic feature of the model is that different types of labor offer different opportunities to achieve efficiency improvements. The larger the wedge between the value of the marginal product and the wage rate, the larger the incremental reduction of average total cost a firm achieves. The difference between superior " ${}^1\text{pwg}$ " and inferior " ${}^2\text{pwg}$ " workers boils down to a cost differential ${}^1\Delta\text{atc} - {}^2\Delta\text{atc}$, where ${}^1\Delta\text{atc}$ and ${}^2\Delta\text{atc}$ follow from the $\text{mp} - \text{w}$ gaps offered by " ${}^1\text{pwg}$ " and " ${}^2\text{pwg}$ " laborers respectively [according to equation (20.1)]. Let q^L denote the quantity of commodities a laborer produces per working hour.¹ In the absence of wage rate adjustments a firm benefits, *ceteris paribus*, [$({}^1\Delta\text{atc} - {}^2\Delta\text{atc}) \cdot q^L$] from hiring a superior rather than an inferior worker. This benefit precisely points to the incremental wage rate a firm can afford in order to replace an inferior worker by a superior laborer (${}^1\Delta\text{w}$) without introducing a rise of the initial cost level. So define

$$(20.2) \quad {}^1\Delta\text{w} = ({}^1\Delta\text{atc} - {}^2\Delta\text{atc}) \cdot q^L.$$

Depending on whether the firms or the laborers induce wage rate adjustments, two cases need to be distinguished.

Firstly, suppose that only *firms* induce wage rate adjustments through competition for superior workers *via* wage offers. It is easy to show that wage rate adjustments will smooth cost differentials up to the point where the incumbent firms' cost levels are equalized. The assumption of wage rate adjustments opens the door to firms to compete for superior workers by offering higher wage rates.² In a barrier commodity market the fear of potential entry induces competition such that all superior workers receive an incremental wage rate equal to ${}^1\Delta\text{w}$. The reason is straightforward. A regime with differentiated unemployment implies that (some) incumbent firms employ " ${}^1\text{pwg}$ " as well as " ${}^2\text{pwg}$ " workers. Let L^1 and L^2 denote the number of superior and inferior laborers respectively employed by a firm. Assume an incumbent firm and potential entrant which have identical L^1/L^2 ratios. If the incumbent firm offers superior workers an ${}^1\Delta\text{w} < {}^e\Delta\text{w}$, then a potential entrant is able to attract the superior workers profitably by bidding an incremental wage rate such that ${}^1\Delta\text{w} < {}^e\Delta\text{w} \leq {}^1\Delta\text{w}$. Both firms' L^1/L^2 ratios change such that ${}^1L^1/L^2 < {}^eL^1/L^2$. Hence, ${}^1\Delta\text{atc} > {}^e\Delta\text{atc}$ so that ${}^1p > {}^ep$. The potential entrant

¹ Given the assumption that employment is associated with $L_0 = L_1 = 1$ per period per worker (Chapter 17), this means that q^L denotes the number of commodities produced by a worker per period.

² It is important to note that in this chapter principal-agent problems are not included (see, for example, Radner 1981 and 1986, Rogerson 1985 and Abreu 1986). The concept of attractiveness introduces principal-agent problems if attractiveness is *not* observable (to firms and competing laborers). Then, the incentive structure, probability of moral hazard, reputation, etc., have to be taken into account. Hence, in contestable and barrier market models either attractiveness (and, therefore, the value of the marginal product) is *public knowledge* or a *credible* incentive not to cheat (e.g. through reputation considerations) exists.

benefits from the improved L^1/L^2 ratio to the detriment of the incumbent firm. An incumbent firm is thus forced to offer ${}^1\Delta w = {}^1\Delta w$ to superior laborers in order to deter entry. So, in equilibrium the superior laborers' incremental wage rate is such that workers' attractiveness is equalized: ${}^1pwg = {}^1mp - {}^1w = {}^2pwg = {}^2mp - {}^2w$.

By way of intermediate conclusion it is argued that competition in a barrier commodity market evokes equalization of the attractiveness of the laborers employed, simply because an effective entry threat induces incremental wage rates offered to superior workers up to the point where the cost reductions reached by incumbent firms are equalized.¹ So, a necessary condition for sustainability of a barrier commodity market configuration when differentiated employment flourishes and wage rate adjustments are permitted, is that incumbent firms offer superior workers an incremental wage rate such that ${}^1\Delta w = {}^1\Delta w$. This result points to some interesting implications of differentiated employment in a two-barrier market model.

First, it appears that the initial and ultimate distribution of types of laborers over firms is *irrelevant*, since the threat of entry forces equalization of employed workers' attractiveness. Second, the fact that superior workers receive an incremental wage rate constitutes an additional *incentive* to introduce (expected individually) optimal schooling investments in the labor market. In effect, superior laborers can capture positive net returns to investments. This result is in line with the predictions of the economics of human capital (Becker 1964 and Siebert 1985) and the theory of entry barriers (Chapter 7). Third, the tendency to cost equalization introduces additional (welfare) effects: (dispersion of) wage rates (is) are increased, and cost levels and prices are less decreased.² Attractiveness of employed laborers is deflated to the level of inferior workers.

The implications of differentiated employment are, again, the result of the fact that the entry threat *ex ante* fails to be effective *ex post*. From the shortage of superior laborers it follows that superior workers do not face equally equipped potential competitors waiting for profitable entry opportunities (${}^1N < L_{ni}$). In regime 3 the opposite holds for inferior workers (${}^1N + {}^2N < L_{ni}$). In regimes 4 and 5 a further shortage of laborers (${}^1N + {}^2N < L_{ni}$) may lead to entry by " 3pwg " workers (${}^1N + {}^2N + {}^3N > L_{ni}$). The incremental wage rate of " 1pwg " and " 2pwg " laborers are then determined as follows: ${}^1\Delta w = ({}^1\Delta atc - {}^3\Delta atc) \cdot q^L$ and ${}^2\Delta w = ({}^2\Delta atc - {}^3\Delta atc) \cdot q^L$. A further entry of less attractive laborers generates a further deflation of the attractiveness of the workers employed.

Secondly, suppose that only laborers induce wage rate adjustments through competition for jobs *via* wage demands. Section 20.4 has shown that unemployed laborers are inclined to accept wage rate reductions up to the point where they receive the reservation wage rate ${}^{min}w$. Hence, in the dichotomy case " 2pwg " workers are forced to adopt the entry-detering wage rate demand ${}^{min}w$ if they face a credible entry threat *ex post* by unemployed " 2pwg " laborers (${}^1N + {}^2N < L_{ni}$). Therefore, the attractiveness of " 2pwg " workers is increased according to ${}^2mp - {}^{min}w$, which is associated with an upward adjustment of ${}^2\Delta atc$ [equation (20.1)]. Superior " 1pwg " laborers are forced to adapt

¹ Of course, a decrease of the wage rate of inferior workers generates equivalent results. However, it is unlikely that a decline of the wage rate is acceptable to inferior laborers, since a wage rate decrease is associated with negative net returns to schooling investments.

² Hence, the deflation in the commodity market is partly neutralized by the inflation in the labor market. A threat of "hit-and-run" entry in the commodity market generates both deflationary (process innovations and human capital improvements) and inflationary (wage rate increases) tendencies. However, the very reason for competition for (superior) laborers (*i.e.*, entry deterrence) implies that the net result is deflation.

downward their entry-deterring incremental wage rate demand ${}^1\Delta w$ accordingly [equation (20.2)]. The implications of the perfect contestability of the market for " 2pwg " workers to the equilibrium features of the market for " 1pwg " laborers depend on the scale of the differences between both type of laborers. Here three cases can be distinguished. Let ${}^1w^R$ denote the zero-returns wage rate demand of superior workers.

First, if ${}^1mp - {}^1w^R > {}^2mp - \min w$, then superior laborers are still able to demand an incremental wage rate ${}^1\Delta w$ that is associated with *positive* net returns to schooling investments without provoking dismissal. Second, for ${}^1mp - {}^1w^R = {}^2mp - \min w$ the market for " 1pwg " laborers shows all the characteristics of a (perfect) barrier market, since the entry-deterring wage rate yields *zero* net returns to superior workers. Third, if ${}^1mp - {}^1w^R < {}^2mp - \min w$, then superior laborers are forced to accept *negative* net returns in order to avoid unemployment, which points to an intermediate case between a contestable and a barrier market. In this case the very incentive to be engaged in schooling investments is, of course, undermined.

20.8 CONCLUSION

Although the analysis of commodity-labor market interactions in the context of general equilibrium models points to the importance of vertical intermarket spillover effects, many crucial aspects of dynamic economics are not yet covered. It is here where industrial organization shows the relevance of (entry-deterring) investment. Chapter 20 has explored the interface between the debate about the microfoundations of macroeconomics and industrial organization by introducing innovative activity in a free entry framework. Firms carry out R&D investment in order to reach cost reductions, whereas laborers undertake schooling investment in order to improve human capital.

It is argued that interesting interactions occur when firms in a free entry commodity market hire workers in a free entry labor market. It appears that rationing of firms or laborers enhances the likelihood of unsustainability, in the sense that exit and/or entry movements occur in order to retain temporary equilibrium. Firms' rationing in the commodity market in a free entry framework is associated with a tendency to concentration in the commodity market, whereas rationing of laborers in the labor market induces perfect contestability.

Heterogeneity of firms in a free entry framework induces price equalization, since superior firms deploy limit pricing in order to profit from entry barriers (cost differentials). Equivalently, if heterogeneity of laborers enters the picture and if only firms are allowed to induce wage rate movements, then differentiated (un)employment introduces firms' cost equalization through wage rate adjustments. This is associated with a tendency to increased wage rate dispersion. Moreover, the attractiveness of laborers is deflated to the level of inferior workers. If only laborers induce wage rate adjustments, the result depends on the scale of the differences between the types of workers involved. The result that price, wage rate and cost levels tend to equalize in an environment with free entry is, of course, hardly new.¹ It has been well-known since, at least, Smith (1776; for a recent examination see Novshek and Sonnenschein 1987). However, an

¹ The equalization postulate of the classical theory generally applies to the long run (Semmler 1984). By and large, this chapter focuses on the short run. The productive capacity and production technology of firms are predetermined in the period $t = 1$ under consideration. The productive capacity of a market is, however, endogenously determined via entry and exit movements. It is these horizontal intermarket spillovers that (partly) take over the coordinating role of the auctioneer.

advanced model in which both static (price, wage and quantity setting) and dynamic (investment) behavior are *endogenously* determined is still conspicuous by its absence. The model above offers a first step.¹ However, two obvious flaws of the model are, for example, that the modeling of the decision making of the (quantity) rationed economic agents on spillover demands and supplies is ignored and formal (existence) proofs are missing.

¹ Of course, this chapter contains preliminaries only. Obvious extensions are the introduction of other aspects of dynamic competition (for example, improvements of product quality, Van Witteloostuijn and Van Wegberg 1988), firms' specific human capital (Becker 1964), an *n*-market framework (compare with general equilibrium models, Arrow and Hahn 1971 and Benassy 1975) and further "simulation" dynamics (Chapter 23).

 

CHAPTER

21

INDUSTRIAL ORGANIZATION AND STRATEGIC MANAGEMENT: Entry(-Deterring) Strategies and Types of Competition

21.1 INTRODUCTION: Strategic Management's Perspective

Intermarket spillovers are a crucial element in the working of the economy. On the one hand, an important contribution of general non-Walrasian equilibrium theory to economics is the modeling of vertical intermarket (rationing) spillovers from commodity to labor markets and *vice versa* (Chapter 20). The key point is that a partial perspective fails to take account of the observation that the (dis)equilibrium features in one market can be of influence on the (dis)equilibrium characteristics of another market. For example, the sales opportunities of firms in the commodity market depend critically upon the purchasing power which buyers can derive from the labor market in their quality as workers.

Traces of multimarket competition can also be found in the literature on industrial organization (Chapter 19). Multimarket competition operates in a set of related markets. In particular, the role of multimarket competition is the cornerstone of the arguments on the source and credibility of potential entry (Part V). This chapter briefly elaborates on the extra dimension which multimarket competition (Van Wegberg and Van Witteloostuijn 1990a and 1990b, and Van Witteloostuijn and Van Wegberg 1990) can add to the theory of competition in industrial organization and strategic management by tracing the specifics of different competitive scenarios. The argument suggests a multimarket framework which permits barrier market theory to be viewed in a broader context.

So, for the sake of variation, interfaces with the field of strategic management are indicated. Here the subareas of strategy analysis and strategy choice are particularly relevant. Johnson and Scholes (1989) point out that

[s]trategic analysis is concerned with understanding the strategic position of the organisation. What changes are going on in the environment and how will they affect the organisation and its activities? What is the resource strength of the organisation in the context of these changes?" (Johnson and Scholes 1989, pp. 11-12).

Strategy choice focuses on the generation, evaluation and selection of strategic options in anticipation of or in reaction to the threats and opportunities reflected in the organization's environment (Johnson and Scholes 1989, pp. 14-15). From these

definitions it is clear that strategic management and industrial organization can benefit from reciprocal entry into each other's field of research as both disciplines are concerned with the analysis of firms' (strategic) behavior in markets (Teece 1984 and Shapiro 1989b).

Strategic management (particularly strategy analysis and choice) cannot be restricted to diagnosis of internal market conditions only. The general case is that competition moves beyond industry boundaries. Incumbency does not necessarily offer protection against potential rivalry. Ever since the 1950s the literature on industrial organization has been unraveling the implications of potential competition (Schmalensee 1988). The models of entry deterrence are particularly worth noting (Schmalensee and Willig 1989). Models of entry-detering strategies abound in the literature (Chapter 7). Moreover, the incumbent firms' very instruments of entry deterrence can also be used as entry strategies by potential entrants (Yip 1982). Porter's (1980) emphasis on potential competition is the benchmark for many scholars and practitioners of strategic management.

This chapter explores the implications of the multimarket perspective for strategic (content) management by identifying antecedents (credibility of entry threat, sources of potential entry, reciprocal entry threats, focus of rivalry and intermarket spillover effects) of strategy choice from the incumbent firm's perspective. The increasing importance of global and multiproduct competition fortifies the force of multimarket rivalry. Strategic content research can benefit from a multimarket perspective. Implications of multimarket competition abound (Karnani and Wernerfelt 1982). This chapter goes into questions that can be raised if strategy choice is motivated by multimarket competition.

So, the argument summarizes results from the literature on industrial organization and strategic management. In doing so, this chapter uses a heuristic device to summarize the arguments on barrier market scenarios by putting barrier market theory in a broader context. The following sections present an analysis of the credibility of the entry threat (Section 21.2) and a heuristic device to explore the implications of sources of potential entry for strategy formulation in an environment in which multimarket competition cannot be ignored (Section 21.3).¹ Moreover, a specific example of multimarket competition (with excess productive capacity) is presented by way of illustration (Section 21.4). By way of summary final remarks on types of competition conclude the chapter (Section 21.5).

21.2 CREDIBILITY OF THE ENTRY THREAT

21.2.1 Entry Conditions

Potential competition is only effective if the entry threat is credible. Empty threats cannot be used to discipline rivals (Schelling 1960). An analysis of entry conditions facilitates the evaluation of the credibility question. Figure 21.1 indicates six essential conditions which determine the (non)credibility of the entry threat and the associated strategy choice. Roman numbers correspond to cases of competition (and associated strategy choice), whereas Arabic numbers indicate questions. Both the cases and questions are subsequently presented.

¹ Sections 21.2 and 21.3 draw heavily upon Van Witteloostuijn (1990d) and Van Wegberg and Van Witteloostuijn (1990a and 1990b), and Van Witteloostuijn and Van Wegberg (1990).

Figure 21.1
Credibility of Entry Threat

The key point is that the relative importance of external conditions limits the set of strategy options. For the sake of the argument, suppose for the moment that the markets show a uniform price level. That is, price heterogeneity is absent (provided that there is no entry threat). Moreover, unless indicated otherwise, the argument is restricted to the two-market case and the incumbent firms' pricing strategy for reasons of space limitation.

21.2.2 Internal Competition

QUESTION 1 (internal market condition): Does internal competition dominate over external conditions?

Strategy choice depends critically on the type of competition that is ruling the market. The first question to be answered is whether internal or external conditions dominate rivalry. Shepherd (1984) opposes contestability by arguing that this theory ignores the dominance of internal over external competition in daily practice (Chapter 6). This certainly can be true if perfect or Bertrand competition defines actual rivalry, since both types of competition give a zero profit so that potential competition does not matter (provided that an assumption of firms' homogeneity is satisfied). That is, internal conditions dominate over external ones. Moreover, in this case incumbent firms are unable to undertake costly entry-detering activities as zero-profit internal competition imposes binding financial restrictions upon the incumbent firms' investment behavior. However, if internal market conditions predict a positive profit, the case can be different.

21.2.3 Pool of Potential Entrants

QUESTION 2 (pool of potential entrants): Is the pool of potential entrants empty?

The entry threat is of course only credible if the pool of potential entrants is not empty. Apart from new firms, existing firms can exert a credible entry threat (Chapters 11 and 19). The key point is that existing firms are not forced to start from scratch. This not only reduces the entry lag, but in addition permits the exploitation of intermarket spillover effects.

Two sources of potential rivalry from existing firms are particularly worth noting (Chapter 11). First, potential entrants probably are existing firms which contest not by investing in the entry market, but rather by exporting goods. Global competition, even if potential, can inspire the incumbent firms' strategy choice (Porter 1986). Import and export strategies can introduce reciprocal entry threats. Second, multiproduct firms constitute a major source of potential entry (Andrews 1949 and 1959, Hines 1957, Berry 1973, Andrews and Brunner 1975, Caves and Porter 1977 and Cairns and Mahabir 1988). Cairns and Mahabir (1988) summarize this familiar argument by pointing out that

[a]ltering the product set requires latent fungibility as well as latent capacity. The latter may be provided by previously sunk expenditures for capital; the former requires discovering the latent fungibility, through R and D and other expenditures that are necessarily sunk, as well as the use of the goodwill of the firm. Clemens (1951), and later Caves and Porter (1977), also view the seeking of new outlets through R and D, and hence an endogenous product set, as important features of corporate strategy, and hence of the policy of entry. ... Product differentiation is another important element of corporate strategy. ... Both Andrews (Andrews and Brunner, 1975) and Hines (1957) note that product differentiation may be an aid to a prospective entrant (Cairns and Mahabir 1988, pp. 273-274).

Gort (1962) and Rumelt (1974)

state [that] multimarket operation of firms will speed redeployment of resources in response to profitable opportunities (Scott 1982, p. 368).

Potential rivalry from multiproduct firms can be associated with both horizontal and vertical multimarket competition. For example, Brunner (1961) supplements Bain's (1956) theory of horizontal entry barriers by pointing to entry through (backward or forward) vertical integration by existing firms, whereas Porter (1980) refers to the horizontal rivalry from new potential entrants and substitute markets next to the vertical bargaining power of suppliers and buyers. For the sake of convenience, the argument in this chapter focuses mainly on horizontal multimarket competition.

21.2.4 Retaliation Lag

QUESTION 3 (retaliation lag): *Can incumbent firms immediately react upon entry?*

An entry threat is not effective if incumbent firms are able to retaliate before entrants can attract consumers (Brock 1983, Schwartz and Reynolds 1983, Farrell 1986 and Schwartz 1986). The point is that if incumbent firms can retaliate immediately, successful entry can never materialize as incumbent firms are inclined to lower price before any consumer is attracted by a potential entrant. Buyers' inertia favor incumbent firms as these rigidities slow down the attraction of consumers by potential entrants. The key point is that potential competition can only discipline the incumbent firms' strategy if potential entrants have the expectation that for a period sufficiently long to warrant entry buyers switch to supplier(s) which offer(s) the lowest price (or a preferred product) before the higher priced firm is able to respond or retaliate (Baumol 1982, p. 4).

21.2.5 Simultaneous Entry Threat

QUESTION 4 (simultaneous entry threat): *Does the threat of simultaneous entry reduce the entry threat to zero?*

Nti (1989) notes that the

research problem raised by the interaction of several mutually aware potential entrants has received little attention in the strategic behavior literature. Yet this is a problem area which raises some important policy issues, especially in antitrust work. A basic assumption in the antitrust policy is that competition is enhanced when there are more potential entrants (Nti 1989, p. 47).

The rivalry among potential entrants cannot be ignored, since the literature suggests that

the probability of independent entry as well as the probability of at least one entry decline with the number of potential entrants (Nti 1989, p. 48).

This result follows from the observation that

[e]ach potential entrant's profit depends not only on the response of existing firms but also on whether other firms enter as well. Multiple entry could impose losses on all (Sherman and Willett 1967, p. 400).

The argument is that too much entry gives a negative payoff to entrants. Hence, the threat of simultaneous entry can imply that the force of external competition is ineffective if potential entrants discount a high probability of too much entry. Incumbent firms can

afford higher prices if the multiplicity of potential rivals reduces the entry threat.

Case I (internal competition). If Question 1, 2, 3 and/or 4 are/is associated with an affirmative answer, the entry threat is zero. Violation of even one of the entry conditions which underlie Question 1 to 4, is sufficient to undermine the credibility of the entry threat. That is, incumbent firms can adopt a strategy that only takes account of internal market conditions. This implies that the blockaded entry price is within reach, where the blockaded entry price denotes the price that would occur in the markets if there were no entry threat.

21.2.6 Scale of Entry

QUESTION 5 (scale of entry): Is full-scale entry unfeasible?

If potential entrants are not able to undertake full-scale (or total) entry, the entry threat does not need to be fully effective. Full-scale entry constitutes a hazard to the very survival of incumbent firms. In the face of small-scale (or partial) entry the incumbent firm may well decide to accommodate niche entry (Section 7.5). The literature on industrial organization shows that in evaluating strategies in response to an entry threat the relative size of the (potential) entrant is of crucial importance (Dixit 1979 and Encaoua *et al.* 1986).

Case II (small-scale entry accommodation). If Question 5 is positively answered, incumbent firms may benefit from entry accommodation (depending on the specifics of the demand rationing scheme). Is so, small-scale entry permits incumbent firms to behave as Stackelberg leaders. Entry can only take place at a scale which still leaves a residual demand to be satisfied by incumbent firms. This means that profit calculation may indicate that the incumbent firm prefers low residual demand at a higher price (entry accommodation) to high total demand at a lower price (entry deterrence).

21.2.7 Profit Motivation

QUESTION 6 (profit motivation): Do potential entrants ignore (small) profit opportunities?

If potential entrants are not inclined to enter for any, however small, profit opportunity, the entry threat is only credible when the entry profit promises to be beyond a particular positive level. For example, potential entrants may ignore (small) entry profit opportunities if they decide on the basis of a set of goals in which other objectives dominate over profit. Risk-averse potential entrants offer a point in case as they are likely to set a risk premium on the entry price so as to evaluate the attractiveness of entry opportunities. Selten (1975) describes a model in which small departures from strict profit maximization rule firms' decision making. This description is supported by the behavioral theory of organizations (Simon 1945 and Cyert and March 1963).

Case III (signal price competition). A positive answer to Question 6 implies that the incumbent firm expects only a price which signals a specific positive profit to induce entry. This means that sustainable profit opportunities can be exploited by setting a price which just allows the critical profit level to be captured. This price can be called the

Figure 21.2
Sources of Potential Entry and (Price) Strategies

signal price.

21.3 SOURCES OF POTENTIAL ENTRY AND (PRICE) STRATEGIES

21.3.1 Multimarket Competition and Strategy

If the evaluation of entry conditions indicates that the entry threat can be fully credible, the next step is to consider the response opportunities open to the incumbent firm: entry deterrence, entry accommodation or exit. The analysis of the sources of potential entry can be used to select strategy. Figure 21.2 illustrates the evaluation of entry sources.

If potential competition cannot be ignored, the diagnosis of sunk, entry and opportunity cost indicates feasible pricing policies. On the basis of the scale of sunk and entry cost and the prices that would occur in the markets if there were no entry threat at the short-run profit-maximizing price (blockaded entry price), types of noncooperative (potential) competition and strategy choice can be distinguished.¹ Figures 21.1 and 21.2 together with the argument in subsequent subsections put the barrier market scenarios of Chapter 11 in a multimarket competitive context. That is, the heuristic device underlying Figures 21.1 and 21.2 and the associated questions clarify under what conditions which type of competition may occur. In particular, Cases VIII, XIII and XV indicate barrier market outcomes. Of course, this type of analysis needs further exploration in future research.

21.3.2 (Non)contestability

QUESTION 7 [(non)contestability]: Is the market perfectly contestable?

QUESTION 8 (opportunity cost): Are potential entrants earning a zero profit if they refrain from entry?

The current extreme case of potential competition in the literature on industrial organization is perfect contestability (Chapter 6). Contestability theory assumes perfect homogeneity of incumbent firms and potential entrants. The literature usually identifies contestability with a condition of zero sunk cost. Average total cost matches average variable (and fixed) cost. Excess productive capacity enhances the credibility of the entry threat (Chapter 11). A further increase of the credibility of the entry threat follows from the condition that in one or both markets the opportunity cost of entry is zero.

¹ A remark is in order with regard to the mode of entry. In effect, the impact of entry cannot be considered without reference to the entry process. In particular, the consequences of entry for the conditions of post-entry competition depend upon the entry mode (Hines 1957, pp. 143-145). This observation is clearly recognized in the literature on multinational enterprise. For example, Caves (1982) points out that "[b]y definition, greenfield entry adds another seller to the market in question, whereas acquisition initially leaves concentration unchanged. Entry by acquisition can have a pro-competitive significance if the MNE [multinational enterprise] picks up and revives a failing business or uses its nonproduction assets to make the acquired company more effective Still, the competitive significance of greenfield entries should be greater" (Caves 1982, p. 102). The concern with the entry mode is related to the interest in, for instance, mergers (Mueller 1989) and direct investment (Dunning 1988). This chapter ignores this issue by focusing on net entries.

Case IV (imperfect contestability). If Question 7 is associated with a positive answer but Question 8 is answered in the negative, contestability cannot be perfect. The point is that the potential entrants' opportunity cost of entry is positive as actual entry is associated with a positive profit forgone. Hence, potential entrants are only inclined to enter if the entry profit exceeds the opportunity cost of entry. This means that incumbent firms can increase price beyond average cost up to the point where this price yields a profit equal to the potential entrants' opportunity cost of entry, which generally follows from the potential entrants' home market price (Chapter 11).

Case V (perfect contestability). The answer to Question 7 is yes if production in the market is not associated with a sunk cost. If Question 8 is answered in the affirmative, potential entrants can undertake hit-and-run entry as entry is free, exit is costless and the opportunity cost of entry is zero. Entry can only be deterred (and exit avoided) by setting the average total cost price, since any price in excess of average total cost signals an entry profit (Chapter 6).

21.3.3 Leapfrogging

QUESTION 9 (sunk entry cost): Is the incremental sunk cost which potential entrants have to bear upon entry positive?

QUESTION 10 (leapfrogging): Are potential entrants motivated by the belief that the probability of leapfrogging is positive?

QUESTION 11 (entry lag): Is the entry lag sufficiently positive?

Imperfect competition implies that a potential entry threat can be effective. Imperfect competition and entry barriers can be closely related. Sunk costs yield a barrier to entry if they are asymmetric. Bain's (1956) theory of entry barriers predicts that entry deterrence requires limit pricing. A credible entry threat can derive from alert (new or unrelated) potential entrants which will prepare for future entry if profitable entry opportunities are observed (Section 11.2). Potential competitors sink cost in reaction to the incumbent firms' price signals, while the incumbent firms have already sunk this cost: sunk cost is asymmetric in favor of incumbent firms. This means that the answer to Question 9 is yes. The key point is that profitable entry requires recovery of the sunk entry cost. *Ceteris paribus*, the sunk entry cost offers incumbent firms the opportunity to set a limit price which gives a positive profit (Section 11.2).

New or unrelated firms can nevertheless constitute a credible entry threat if the opportunity to leapfrog motivates the potential entrants' investment strategy. The point is that in this case potential entrants' can observe an entry opportunity even if the established firms adopt average total cost pricing and an aggressive R&D policy (which implies entry deterrence as far as is feasible). The crucial condition is that the (perceived) probability of leapfrogging is positive, which follows from firms' heterogeneity (Subsection 11.2.3.4).¹

¹ This source of potential entry (new or unrelated firms) is, strictly speaking, not a case of multimarket competition (which operates in a set of related markets). However, this case cannot be ignored while analyzing a firm's environment. It is important from the perspective of both industrial organization

Case VI (effective entry barrier). If Question 10 is given a negative answer, entry barriers are effective. Potential entrants are not motivated by leapfrogging arguments. Hence, as the well-established limit price theory predicts, incumbent firms can benefit from the potential entrants' cost disadvantage by increasing price above average total cost up to the point where the incumbent firms' price just fails to warrant entry (Chapter 7).

Case VII (lagged leapfrogging threat). A positive answer to both Question 10 and Question 11 gives a lagged leapfrogging threat. Heterogeneity of incumbent and potential competitors in favor of the latter is unsustainable if superior potential rivals actually undertake sunk R&D (Section 11.2.3). If incumbent firms are aware of the existence of superior potential entrants, their strategy can be to accept exit in the entry period but to cream off temporary pre-entry profits on the one hand and to refrain from sunk R&D on the other. If incumbents fail to anticipate entry by lower-cost entrants, forced exit in the entry period goes hand in hand with the exit cost that follows from unamortized sunk investment cost. An entry lag facilitates the profitability of the creaming-off strategy as the incumbent firm can temporarily benefit from internal market conditions by setting the blockaded entry price.

Case VIII (immediate leapfrogging threat). If the answer to Question 10 is yes but Question 11 is answered in the negative, the entry threat is effective immediately. Potential entrants can leapfrog over the incumbent firms' innovation without delay. Minimization of price (within the bounds of feasibility) minimizes the probability of leapfrogging entry as a low price signals reduced profit opportunities. This argument points to average total cost pricing. Moreover, if incumbent firms rely on their (R&D) capabilities, an intensive R&D policy can be used to keep ahead of potential entrants. This type of R&D competition is described in the literature on patent races (Chapter 8).

21.3.4 Dumping

QUESTION 12 (excess productive capacity): Do potential entrants have disposal of excess productive capacity?

QUESTION 13 (dumping threat): Can potential entrants announce a credible dumping threat?

QUESTION 14 (reciprocal dumping threat): Are incumbent firms able to signal a credible threat of reciprocal dumping?

QUESTION 15 (large-scale dumping): Are potential entrants capable of large-scale entry?

Although a positive sunk cost can be an exit barrier, it does not necessarily reflect a barrier to entry. To be precise, if incumbent firms and potential entrants both face the same scale of sunk cost and production technology, then there is a zero cost differential. Cairns and Mahabir (1988) therefore explore Clemens' (1951) argument by suggesting that the entry threat ensues from firms with excess productive capacity in related markets (Section 11.3). This indicates a positive answer to Question 12.

If firms face nonbinding productive capacity restrictions (excess capacity), dumping can occur (Brander and Krugman 1983 and Pinto 1986). If a potential entrant has excess capacity, then this firm is, *ceteris paribus*, inclined to use this excess capacity in order to supply in the incumbents' market for any price above marginal entry cost (Subsection 11.3.3). This indicates a positive answer to Question 13. Moreover, an expelled incumbent firm may retaliate by reciprocal entry into the entrant's home market (Bulow *et al.* 1985a and Callem 1988). The (non)credibility of the reciprocal dumping threat indicates two broad classes of noncooperative dumping competition.

Firstly, take the case where the threat of reciprocal dumping is empty. Dumping by expelled incumbent firms does not undermine the credibility of the entry threat if, in particular, (i) the potential entrants' home market is safe against entry by expelled incumbent firms (for example, because the incumbents' response lag is large enough to invalidate the retaliatory dumping threat), (ii) the incumbent firms are disadvantaged (for example, in terms of size, product or cost) relative to the potential entrants and (iii) the incumbent firms' alternative replies (for example, entry into a third market) are more profitable than reciprocal dumping. If the reciprocal entry threat is empty (a negative answer to Question 14), the incumbent firms' strategy choice depends on the dumping scale.

Case IX (large-scale dumping). If dumping can occur on large scale (a yes-answer to Question 15), the very survival of the incumbent firm is at risk. A price in excess of the potential entrants' marginal entry cost induces dumping such that the incumbent firm is expelled from the market. The incumbent firm can either decide to minimize the entry incentive by setting the average variable cost price (which is the incumbent firms' lowest feasible price) or to accommodate entry by exit (which, if feasible, can be accompanied with entry into a third market).

Case X (small-scale dumping). A negative answer to Question 15 permits dumping accommodation. Small-scale dumping implies that the entrants cannot serve total market demand by using excess productive capacity only. Residual demand reflects a market niche which incumbent firms can still exploit by setting an above average variable cost price. The incumbent firm can set price equal to the potential entrants' home market level (which indicates the opportunity cost of entry at a scale beyond idle capacity) so as to deter entry at a scale in excess of the dumping volume.

Secondly, the case is different if both incumbent firms and potential entrants have sustainable excess capacities which can credibly be used to exert a dumping threat (Subsection 11.3.3). Symmetric excess productive capacities can introduce a reciprocal dumping threat (Pinto 1986), which points to a positive answer to Question 14 if (i) the potential entrants' home market is unprotected, (ii) the incumbent firms' response lag is sufficiently short, (iii) the incumbent firms and potential entrants face parity and (iv) reciprocal dumping is the incumbent firms' preferred reply.

Case XI (zero dumping). Reputation considerations can call into question the sustainability features of the case where both firms are dumping (Subsection 11.3.3). The key point is that firms can use trigger strategies (Friedman 1986 and Pinto 1986): if you decide to dump, I punish by undertaking reciprocal dumping. Moreover, firms can (tacitly) agree upon not dumping. The dumping game gives (tacitly) cooperative rather than

noncooperative outcomes.¹

21.3.5 Hit-and-Run

QUESTION 16 (protected home market): Can potential entrants benefit from a home market which is safe against entry?

QUESTION 17 (opportunity cost of entry): Does the incumbent firms' blockaded entry price exceed the potential entrants' home market price?

Home market arguments can be facilitated by a restriction to one-sided entry (threats). That is, the potential entrants' home market is safe against entry (a yes-answer to Question 16). This means that entrants can anticipate an easy return to the home market after hit-and-run entry (Subsection 11.4.2). For example, Calem (1988) introduces the assumption of one-sided entry (threats) by arguing that the incumbent firms' entry cost is sufficiently large to keep it from entering the potential entrants' market (Calem 1988, p. 175) or, alternatively, by supposing that legal or regulatory barriers exist which prevent incumbent firms from being potential entrants into the rivals' market (Calem 1988, p. 182, note 5).²

Hit-and-run strategies imply that potential entrants deploy temporary cross-entry. The hit-and-run strategy is credible if, on the one hand, the entrant is able to run after hitting the incumbents' market and, on the other, the incumbent firms retaliation lag is sufficiently long to warrant a temporary entry profit.

Case XII (dominant entry opportunity cost). Entry is only profitable if the entry profit exceeds the opportunity cost of entry. With a safe home market and zero entry cost this means that the entry market price has to be larger than the potential entrants' home market price. If this is not the case (Question 17 is given a negative answer), entry leaves the entrant worse off. Hence, the entry threat is empty, which permits the incumbent firm to exploit internal market conditions by setting the blockaded entry price.

Case XIII (safe home market). The entry threat is credible if the answer to Question 17 is yes: the blockaded entry price exceeds the potential entrants' home market price. Hence, if the potential entrants' home market is safe against entry, entry deterrence requires a

¹ The result that reciprocal dumping (entry) threats facilitate the sustainability of noncompetitive strategy is supported by multimarket contact (or mutual forbearance) theory (Edwards 1955, Adams 1974, Heggstad and Rhoades 1978, Scott 1982 and Feinberg 1984 and 1985). Multimarket contact theory is a quest for empirical support for Edwards' (1955) hypothesis that "when sellers meet in several markets, their recognition of the interdependence of their operations may blunt the vigor of their competition with each other" (Scott 1982, 369). Feinberg (1985) specifies Edwards' hypothesis by arguing that "companies meeting rivals in more than one market will be able to facilitate collusion in one or all of those markets" (Feinberg 1985, p. 238), which supports the cooperative result of Case XI. Kantarelis and Veendorp's (1988) model provides a theoretical foundation of the multimarket contact hypothesis. The argument in this chapter supplements multimarket contact theory by arguing that potential multimarket contacts also may affect the rivals' strategy choice.

² Willard and Savara (1988) argue that Japanese firms can benefit from one-sided entry as Japan's government protects the home market via legal measures. Yoffie and Milner (1989) even point to the fact that firms seek strategic trade policy. That is, firms may demand trade barriers for the home market from legislators.

strategy of setting price equal to the level in the potential entrants' home market.

21.3.6 Hit-and-Stay

QUESTION 18 (reciprocal entry threat): *Can the incumbent firm exert a credible threat of reciprocal entry?*

Unprotected home market arguments rely on the assumption that both incumbent firms and potential entrants undertake sunk investment in their home markets. This points to the possibility of reciprocal entry threats (Bulow *et al.* 1985a, Pinto 1986 and Calem 1988). If firms have to take account of binding capacity constraints (with fungible capital), leaving the home market can induce entry (Cairns and Mahabir 1988). Question 16 is associated with a negative answer. Since entry affects home market profit, a potential entrant takes the implications for home market demand into account (Subsection 11.4.2).¹

It may be that the expelled incumbent firm enters another (perhaps even the entrant's) market. If incumbent firms can enter the potential entrant's market, there is a reciprocal entry threat. Porter (1980) points to cross-parry. That is,

[w]hen one firm initiates a move in one area and a competitor responds in a different area with one that affects the initiating firm, the situation can be called a *cross-parry*. This situation occurs not infrequently when firms compete in different geographic areas or have multiple product lines that do not completely overlap. For example, an East-Coast-based firm entering the western market may see a western firm entering the eastern market (Porter 1980, p. 84; Porter's italics).

Watson (1982) suggests counter-competition, which implies that

pursuit of a foreign competitor's domestic markets can help protect the threatened company's own home market sales (Watson 1982, p. 40).

Karnani and Wernerfelt (1985) also point to entry into the potential entrant's market as a way to discipline potential entrants. Their analysis suggests that in order to have a short response lag the firm must already have a foothold in the potential entrant's market. Building up a foothold represents a commitment necessary to exert a credible entry threat. The authors call this a mutual foothold equilibrium in which

each firm maintains a foothold in the other firm's market and thus has a stick with which to discipline the other firm. ... This equilibrium offers little advantage to a firm which decides to attack first, because the other firm can counterattack quickly (Karnani and Wernerfelt 1985, p. 90).

Karnani and Wernerfelt's mutual foothold equilibrium resembles the hostage strategy in the theory of multinational enterprise (Caves 1982, p. 107). An ambiguity surrounding foothold and hostage equilibria is that the creation of a foothold can also be a first step in

¹ This implication is of course associated with the particulars of the entry mode (shifting capacity from home to entry market). That is, if entry takes place through investment, the home market can be served as before. However, in this case the reciprocal entry threat can still be effective. For example, Encarnation (1987) points to the importance of cross-investment.

the entry process (Willard and Savara 1988).

So, reciprocal entry (threat) complicates the entry strategy. By shipping goods to an entry market, entrants without excess productive capacity withdraw supply from their home market. The alternative option can be to adopt a strategy of hit-and-stay entry. A hit-and-stay strategy is feasible if the entrant can keep the expelled incumbent firm from recapturing its market share. The re-entry-detering price depends on the expelled firm's opportunity cost. If the incumbent firm faces idle capacity after exit, re-entry (after the retaliation lag) is profitable for any price in excess of average variable and fixed cost. So, as with reciprocal dumping, a threat of reciprocal entry can be credible if lag, parity and profitability conditions are satisfied. At this point the profitability condition can be clarified with reference to Eaton and Lipsey (1980) and Judd (1985). The re-entry-detering price can be increased if an expelled firm is able to benefit from an alternative use of its productive technology. For example, the incumbent firm can enter a third (perhaps even the entrant's) market after exit. The mirror argument is that the credibility of the entry threat increases if the incumbent firms' exit barrier decreases.

Case XIV (credible reciprocal entry threat). If the incumbent firm's reciprocal entry threat is credible (a yes-answer to Question 18: for example, Bulow *et al.* 1985 and Calem 1988), the firm has disposal of an additional entry-detering instrument (provided that the answer to Question 17 is yes). The incumbent firm's ability to strike back in the potential entrants' home market may constitute a motivation to refrain from entry in the first place (as reciprocal entry increases the opportunity cost of entry). The precise content of the incumbent firm's strategy is determined by the rules of the game. For example, if the incumbent firm is only able to undertake small-scale entry, entry deterrence is associated with a lower price (in the extreme the potential entrants' home market level) than when the incumbent firm can credibly announce a threat of reciprocal large-scale entry (which in the extreme permits the blockaded entry price).¹

Case XV (noncredible reciprocal entry threat). If the incumbent firms can only exert empty reciprocal entry threats, the credibility of the hit-and-stay entry threat is facilitated. Entry deterrence requires the incumbent firm to adopt the strategy of setting price equal to the potential entrants' home market level (which indicates the potential entrants' opportunity cost of entry).

21.4 AN EXAMPLE: (Excess) Productive Capacity

21.4.1 Bertrand Competition, Capacity Constraints, Identified Entry and Multimarket Modeling

This section illustrates the implications of multimarket competition for strategy choice and market performance in a two-market - two-period model with reciprocal entry (threats) and investment in productive capacity (Van Witteloostuijn and Van Wegberg 1989). This means that the focus is on a specific intermarket spillover effect: joint economies of scale.

¹ Multimarket contact theory also stipulates the implications of size. For example, Scott (1982) emphasizes the role of concentration by suggesting that "firms in markets where both seller concentration and multimarket contact are high will recognize their interdependence and not compete away their own profits" (Scott 1982, p. 372), which supports the argument of Case XIV.

This section illustrates the (reciprocal) dumping dilemma [(Sub)sections 11.3 and 21.3.4]. An important feature of an industrial organization model is the assumption on entry. It is in this respect that two modeling strategies dominate the literature. On the one hand, the (tacit) assumption that entry barriers impede entry implies an exclusive focus on actual competition among incumbent firms. On the other hand, the assumption of *free entry* introduces exclusive attention to potential competition. Assumptions on entry can be critical to the implications for market performance that the model yields. The literature on Bertrand competition can be used to illustrate this point (Parts III and V).

Bertrand's (1883) criticism of Cournot (1838) induced (after a considerable delay) the modeling of games that indicate the implications of pervasive price competition for market performance in settings without (the threat of) entry. The textbook interpretation of Bertrand competition focuses on the case where rivals face nonbinding capacity constraints (for example, Waterson 1984, p. 25). If equally equipped firms face nonbinding capacity restrictions, Bertrand competition yields average cost pricing even in a market structure with two firms. This result follows from the discontinuity in the individual firm's demand function, which indicates that buyers switch to the firm which offers the lowest price (or preferred product) before a higher-priced firm is able to respond or retaliate (Chapter 6).

Recently, however, Edgeworth' (1925) argument that results are different if the rivals have binding capacity constraints has been studied (Kreps and Scheinkman 1983, Davidson and Deneckere 1986 and Osborne and Pitchik 1986). Broadly speaking, binding capacity restrictions induce a deviation from the static efficient market performance that occurs if capacity is unlimited. If capacity is limited, a lower-priced rival cannot distract all the customers from a higher-priced firm. Contingent on the specifications of demand and the rationing scheme, the result can be identical to the one-stage Cournot equilibrium outcome. Under these conditions, the one-stage Cournot game, where a firm has no control over the market clearing price other than by the choice of its output level, turns out to be shorthand for a two-stage Cournot-Bertrand game in which firms choose capacity before they compete over prices. The (Cournot-)Bertrand games with nonbinding and binding capacity restrictions are, however, modeled in a setting without (the fear of) entry. This means that, first, market structure (in terms of the number of firms) is predetermined and, second, that market behavior is only disciplined by actual rivalry.

The literature has not yet explored the implications for market performance of two-stage Cournot-Bertrand games with *free entry*. The theory of perfectly contestable markets (Baumol 1982 and Baumol *et al.* 1982a) combines the assumptions of Bertrand competition and free entry. In the single-product case this means that incumbent firms are forced to set (minimum) average cost prices so as to seek a position of exit avoidance and entry deterrence. (The set of) market structure(s) is endogenized, since the sustainable number(s) and size(s) of incumbent firms follow from cost and demand properties (Chapter 6). Although firms can face binding capacity constraints, the industry as a whole is associated with unlimited capacity (through entry).

The essential assumption of perfect contestability is that entry is free and exit is costless (Baumol 1982, p. 3). If firms are, however, allowed to choose capacities, the intertemporal sustainability of the market configuration is in danger (Baumol *et al.*

1982a).¹ Entry barriers may result if expansion of productive capacity is associated with a sunk investment. Existing contestability models, however, assume sunk exit cost to be zero (Shepherd 1984 and Stiglitz 1987). The literature suggests that even a small sunk exit cost undermines the favorable implications of contestability (Baumol and Willig 1981, Grossman 1981, Baumol *et al.* 1982, Brock 1983, Schwartz and Reynolds 1983, Schwartz 1986, Stiglitz 1987 and Dasgupta and Stiglitz 1988a; see also Section 11.2).

The argument in Section 21.4 resembles that of Bulow *et al.* (1985a and 1985b) in particular. It appears that multimarket models with entry threats introduce a wide spectrum of research opportunities. However, this section does not offer a well-worked out model, but gives a preliminary impetus. Rather it points out the opportunities and difficulties offered by multimarket modeling in industrial organization without being exhaustive. The sections below explore the features of a model that permits the study of the implications of an entry threat in a multimarket setting. This means that the potential entrants can be identified. Although Calem (1988) describes a two-stage Cournot game, from the perspective of market performance, however, it can be interesting to focus on a two-stage Cournot-Bertrand game (Chapters 6 and 11). Three reasons can be put forward to defend that (second-stage) price competition is to be preferred to quantity competition. First, real world competition shows that price setting is an important competitive weapon of firms. Second, Cournot models have to rely on an auctioneer in order to explain the determination of prices. Third, competition over prices can yield radically different results (as regards market performance) from competition over quantities only.

21.4.2 Demand, Capacity Choice, Price Setting and Entry Deterrence

To keep matters simple and transparent, the model features are illustrated in a two-market setting. Markets A and B are characterized by the fact that demand is localized. That is, the customers in market A cannot buy commodities in market B and *vice versa*. This assumption can be plausible in an international context. In markets A and B initially operate A_{ni} and B_{ni} number of firms respectively. The key point of the two-market model is that the firms in both markets imply a *reciprocal* entry threat. That is to say, incumbent firms in market A assume that potential entrants in market B are alert to any profitable entry opportunity from market B to A and *vice versa*. For the sake of convenience, unless indicated otherwise, market A is generally referred to as the incumbents' market, whereas the firms in market B are regarded as the potential entrants.²

From the demand side both markets are independent. That is, only firms are mobile between markets. Intermarket demand spillovers are ignored. Products in both markets are perfect substitutes. Three assumptions characterize market demand: (i) the demand function; (ii) the demand growth; and (iii) the rationing scheme. For the sake of simplicity, demand (development) is determined exogenously. Moreover, income effects

¹ This is particularly clear from Chapter 14 in Baumol *et al.*'s 1982 book, where it is stated that "[u]nsustainability can come closer to being the rule rather than the exception. ... We begin our discussion with the source of the difficulties that appear to be most general and most serious - economies of scale in sunk construction costs" (Baumol *et al.* 1982a, pp. 405-406).

² Calem (1988) points out that the "distinction between 'incumbent firm' and 'potential entrant' that is commonly employed in models of entry will not always be applicable in the case of a penetrable market. This is because the rivalry between a firm in a penetrable market and a firm threatening to enter the market may be reciprocal. As these firms operate in related markets, each may be a potential entrant into the other's market" (Calem 1988, p. 171).

are ignored. The model is presented in a two-period setting. First, market demand in a period t follows from

$$(21.1) \quad M D_t = D(p_t),$$

where D is twice-continuously differentiable, $\partial D / \partial p_t < 0$, $M = A, B$, $t = 0, 1$ and p denotes price. Second, market demand growth is simply characterized by a growth parameter ϵ . This parameter can of course be different for both markets. That is,

$$(21.2) \quad M D_1 = (1 + M \epsilon_1) \cdot M D_0,$$

for $M = A, B$. Firms can face (market) demand uncertainty. So, the growth rate $M \epsilon_{t+1}$ can be an expectations variable. This is particularly relevant in a model where firms can expand their productive capacity.

Third, the results can also depend critically upon the rationing scheme: "[d]ifferent rationing rules will clearly lead to different equilibria" (Davidson and Deneckere 1986, p. 405). Three examples of rationing schemes are worth noting. First, Kreps and Scheinkman (1983) use the rationing rule that leaves the worst contingent demand curve for higher-priced firms. Davidson and Deneckere (1986) use an alternative scheme: the Beckmann (1965) rationing rule. The Beckmann rule yields the best contingent demand curve for the higher-priced firms. Davidson and Deneckere (1986) point out that the Kreps-Scheinkman and Beckmann rule are the extremes of a large set of possible rationing schemes. Different intermediate types of rationing schemes can be assumed. Both the Kreps-Scheinkman and Beckmann rationing rule (and the intermediate schemes) are, however, associated with an extreme Bertrand assumption: up to its demand or capacity constraint a lower-priced firm attracts the buyers from higher-priced rivals. This implies an assumption of perfect information in the sense that customers can locate lower-priced firms before prices are allowed to change. The fact of the matter is, however, that search costs or informational deficiencies can impede the functioning of perfect rationing rules. If the history of the market is taken into account, it can then be assumed that only a fraction of the customers of the higher-priced firms are attracted to the lower-priced firms. For the sake of clarity, henceforth the Kreps-Scheinkman rule is implicitly assumed.

A firm has to decide on three strategic variables: quantity, price and capacity. Moreover, a firm can choose to transfer supply to the rivals' market. That is, a firm can decide on entry. The model is framed as a two-stage game. In the first stage (period $t = 0$) firms decide on capacity; in the second stage (period $t = 1$) firms choose (home and entry) quantities and prices on the basis of the predetermined scale of productive capacity. Contestability theory assumes that incumbent firms have a first-mover advantage. That is, firms first set prices in their home market. Entry decisions follow only *after* information on the rivals' home market prices has been received (Section 10.3).

A firm has to choose productive capacity while taking into account expected demand in its home market and a potential entry market. As a point of reference the decision making is examined of a firm which assumes a credible entry threat without taking into account its own entry opportunities. First, the expansion technology is discussed. Expansion of productive capacity is subject to a one-stage lag. A sunk setup cost sc_0 in the first stage is associated with a growth of productive capacity Δk_1 in the second stage (suppressing firm and price indices). So,

$$(21.3) \quad \Delta k_1 = K(sc_0),$$

where $K(0) = 0$, K is twice-continuously differentiable, $\partial K/\partial sc_0 > 0$ and $\partial^2 K/\partial sc_0^2 < 0$. The assumption that capacity growth is subject to decreasing returns facilitates the existence of a unique cost-minimizing capacity choice.

For $q \leq k$ (suppressing indices) production is subject to constant returns:

$$(21.4a) \quad avc = L_{avc},$$

where q is the quantity produced by the firm and avc denotes the unit variable (and fixed) cost of production.¹ For $q > k$ it can be assumed that further production is either impossible [case (i)] or associated with decreasing returns [case (ii)]. Case (i) implies that the capacity restriction is strictly binding: $avc = \infty$ for $q > k$. Case (ii) implies that capacity can be binding because further production is increasingly costly:

$$(21.4b) \quad R_{avc} = C(q) \text{ for } q > k,$$

where $R_{avc} > L_{avc}$. C is twice-continuously differentiable, $\partial C/\partial q > 0$ (and $\partial^2 C/\partial q^2 > 0$).

The assumption of Bertrand competition introduces a difficulty if capacity growth is expensive. The following argument illustrates this point. Suppose that the incumbent firms in market A anticipate in the first stage a growth of demand in market A in the second stage by expanding their productive capacities to $\Delta k_1 = \Delta d_1$, which is of course associated with a sunk setup cost Δsc_0 . A zero profit implies an average total cost price: $\Delta p_1 = \Delta avc_1 + \Delta sc_0/\Delta d_1$ (ignoring discounting and assuming that the sunk cost has to be recovered in the second stage). Rivals in market B can, however, capture positive profits by anticipating the developments in market A. Take the following limit case. If firms in market B abandon the decision to expand productive capacity in the first stage, they face the zero-profit price $\Delta p_1 = \Delta avc_1$ in the second stage, since sunk cost is zero. The contestability assumption that all firms have access to the same pool of production technologies ($\Delta avc_1 = \Delta avc_1$), implies profitable entry opportunities for the firms in market B. In a contestable market a firm which abandons capacity expansion is *always* able to profitably underbid an expanding rival.

Three remarks are in order. First, if after full entry market demand is still only partially satisfied (*i.e.*, entrants face binding capacity constraints), it can be that the expanding incumbent firms are able to recover the sunk cost with positive profit notwithstanding the entry of rivals. This can be plausible if market B is small relative to A so that the firms in market B operate as fringe competitors (which implies small-scale entry). Second, firms in both markets face the risk of being pushed aside by entrants. This means that the expansion - nonexpansion game has to be fully analyzed in order to find out the existence and properties of the associated perfect equilibrium (or equilibria). For example, the case where none of the firms decides to enlarge productive capacity can prove to be an equilibrium. This can imply that demand growth remains unsatisfied. Third, the precise (time) structure of the game is crucially important. For example, if the firms in market A which are outperformed by entrants from market B are allowed to

¹ Note that contestability theory is generally illustrated for the special case with a flat-bottomed average cost curve. For the sake of simplicity, here an J-shaped curve is assumed.

respond by entry into market B, then the strategy space (and probably the associated equilibria) is (are) completely different.¹

However, the argument in this section takes another path. To be precise, the assumption is introduced that the (costly) growth of productive capacity is connected with scale economies. This implies that the average variable (and fixed) production cost decreases with scale. So,

$$(21.5) \quad avc_t = S(k_t),$$

where S is twice continuously differentiable, $\partial S/\partial k_t < 0$ and $t = 0, 1$. Scale economies are an incentive to increase productive capacity. Average variable (and fixed) production cost decreases with scale.²

Equations (21.3) and (21.5) imply that a capacity exists which yields a minimum level of unit average total cost in the second stage. Suppose that the firm anticipates that it will sell only in the home market in the second stage. Contestability theory shows that entry deterrence (and exit avoidance) requires (minimum) average cost pricing. In a dynamic setting this implies that average cost has to be minimized (Baumol *et al.* 1982a and Chapter 15). The analysis of this case provides a reference point for further analysis in the next subsections. The decision on capacity in the first stage depends on whether or not the firm expects to face a demand constraint in the second stage.

The *unconstrained* minimization of the second-stage average total cost (atc) follows from the expectation that production at full capacity can be sold in the second stage ($d_1 \geq k_1$). So,

$$(21.6) \quad \underset{sc_0}{\text{Min atc}}_1 = avc_1 + sc_0/k_1 = S[k_0 + K(sc_0)] + sc_0/[k_0 + K(sc_0)],$$

which solves for an unconstrained (superscript U) cost-minimizing enlargement of productive capacity: $^U k_1$. Assume that equation (21.6) gives a unique solution: $^U sc_0$ yields a $^U k_1$ that is associated with a minimum average total cost level $^U atc_1$. Equation (21.6) implies that the minimization of average total cost proceeds unconstrained in the sense that the (tacit) assumption is that $d_1 \geq ^U k_1$.

If the firms expect that in the second stage they will face a demand constraint ($d_1 < ^U k_1$), the situation is different. That is, the *constrained* minimization of second-stage average total cost follows from

$$(21.7) \quad \underset{sc_0}{\text{Min atc}}_1 = avc_1 + sc_0/d_1 = S[k_0 + K(sc_0)] + sc_0/d_1,$$

which solves for a constrained (superscript C) cost-minimizing expansion of productive capacity: $^C k_1$. Assume that $^C sc_0$, $^C k_1$ and $^C atc_1$ are unique. Of course, $^C k_1 < ^U k_1$ and $^C atc_1 > ^U atc_1$.

¹ This means that an essential assumption of the contestability theory, incumbent firms' first-mover advantage, is dropped. After the potential entrants' response to incumbent firms' prices further moves can follow. Of course, in equilibrium it can be that actions are restricted to just one move.

² Here this is formally equivalent to the assumption that an expansion of productive capacity goes hand in hand with a process innovation. It is plausible to think that there may be an increased efficiency of management and organization (for example, through reduced X-inefficiency) if the scale of production is enlarged.

The result of the constrained minimization yields two difficulties. First, $C_{p1} > U_{p1}$, which implies profitable entry opportunities for firms which introduce the unconstrained enlargement of productive capacity. The key point is that intermarket cost spillovers (i.e., joint economies of scale) can enter the picture. This is discussed further in the next subsection below. Second, constrained minimization of average total cost can yield excess capacity if the scale economies dominate the investment cost. It is excess capacity in connection with positive sunk costs that introduces the danger of (reciprocal) dumping. This point has been examined in Section 11.3.

21.4.3 Scenarios of Market Interactions

In order to illustrate the subtle strategic interactions that can occur in a multimarket framework, this subsection explores some of the possible implications of different scenarios in the two-market model. That is to say, the list of strategic interactions presented below does not pretend to be exhaustive. It only illustrates the need for the analysis of multimarket games. If $C_{p1} > U_{p1}$ (Subsection 21.4.2), profitable entry opportunities may exist for firms which introduce the unconstrained enlargement of productive capacity so that they can benefit from joint scale economies through entry. Three categories of scenarios can be distinguished: potential aggregate (or world) excess capacity (I); potential aggregate capacity matching (II); and potential aggregate capacity shortage (III).

The different scenarios ensue from different demand developments in both markets: that is, $A_{\epsilon_1} \neq B_{\epsilon_1}$. Here it is assumed that firms foresee the growth rates of demand in both markets. So, in deciding on capacities firms take notice of the second-stage demand regimes. Firms in both markets are assumed to have access to the same pool of technological (expansion) opportunities; that is, $A_{U_k} = B_{U_k} = U_k$ and $A_{U_{atc}} = B_{U_{atc}} = U_{atc}$. For the sake of clarity, time (or stage) indices are suppressed.

I. Potential aggregate excess capacity

It may be that both markets taken together show excess capacity if all firms adopt unconstrained cost minimization: $(A_{ni} + B_{ni}) \cdot U_k > A_D(A_p) + B_D(B_p)$, where $A_p = B_p = U_{atc}$. This case can be differentiated as to the capacity-demand ratios in both markets in isolation.

(1) Potentially, either market A faces a shortage of productive capacity or productive capacity in market A precisely matches market demand, whereas market B shows excess capacity (or vice versa): $A_{ni} \cdot U_k \leq A_D(A_p)$ and $B_{ni} \cdot U_k > B_D(B_p)$. Firms in market A can either deviate from or comply with unconstrained cost minimization. In both cases entry can be induced. An enlargement of productive capacity beyond the cost-minimizing scale yields higher average total cost prices, whereas the unconstrained cost-minimizing capacity either leaves contingent demand unfulfilled at the average total cost price or is associated with above-average total cost prices that tend to the market clearing level. The options are different for the firms in market B. On the one hand, they can decide to

introduce excess capacity in order to enter into market A.¹ On the other hand, they can adopt the constrained cost-minimizing scale of capacity. However, the latter strategy yields a cost disadvantage, so that average total cost pricing can induce entry: $A_p = U_{atc} < B_p = C_{atc}$.

(2) Both market A and market B potentially show excess capacity: $A_{ni} \cdot U_k > A_D(A_p)$ and $B_{ni} \cdot U_k > B_D(B_p)$. This means that firms in both markets can decide to introduce either the unconstrained or the constrained cost-minimizing scale of productive capacity. So, three cases can be distinguished: (i) A_{Uk} and B_{Uk} ; (ii) A_{Ck} and B_{Uk} (or *vice versa*); and (iii) A_{Ck} and B_{Ck} . Case (i) implies that both firms encounter the opportunity to use the excess capacity in order to enter into the rivals' market. So, case (i) can yield reciprocal dumping. If productive capacity is more excessive in one of the two markets, the benefit (or damage) from dumping is asymmetric. Case (ii) is associated with entry of the firms in market B into market A. Not only can the firms in market B make use of excess capacity, but they can also benefit from a cost (and price) advantage. Case (iii) can be connected with either (reciprocal) entry or (reciprocal) entry deterrence. For example, if both markets show a different magnitude of the quantity demanded from the firms (and so different constrained cost-minimizing capacities), $A_p \neq B_p$. This induces profitable entry opportunities for the lower-priced firms.²

II. Potential aggregate capacity matching

A second category of scenarios occurs when both markets taken together show a precise match of productive capacity and aggregate demand if all firms introduce the unconstrained cost-minimizing capacity: $(A_{ni} + B_{ni}) \cdot U_k = A_D(A_p) + B_D(B_p)$, where $A_p = B_p = U_{atc}$. This case can also be differentiated as to the capacity - demand ratios in both markets in isolation.

(1) In both market A and market B the market capacity that follows from unconstrained cost minimization precisely matches market demand: $A_{ni} \cdot U_k = A_D(A_p)$ and $B_{ni} \cdot U_k = B_D(B_p)$. This implies that in both markets the incumbent firms are inclined to introduce the productive capacity which yields minimum average total cost. Reciprocal entry deterrence requires $A_p = A_{Uatc} = B_p = B_{Uatc}$. Reciprocal entry deterrence (which yields a zero profit to each firm) is an equilibrium outcome: none of the firms benefits from entry, since successful entry is associated with below-average total cost pricing and so a negative payoff.

(2) Potentially, in market A firms face a shortage of productive capacity, whereas market B is associated with excess capacity (or *vice versa*): $A_{ni} \cdot U_k < A_D(A_p)$ and $B_{ni} \cdot U_k >$

¹ At this point it should be noted that the (tacit) assumption here is that firms can and are permitted to anticipate future entry opportunities while deciding on the enlargement of productive capacity. This introduces an extra strategic dimension in the games.

² At this point the implications of the assumption that the game terminates after the potential entrants' second-move response to the incumbent firms' first-move price setting are obvious. Entry is after all associated with a withdrawal from the home market, which leaves unfulfilled demand that can be satisfied by further entry if a third move is permitted. This threat of responsive entry can keep the potential entrant from entry in the first place (Cairns and Mahabir 1988). This argument does, of course, not apply to entry with excess capacity (Chapter 11).

$B_D(B_p)$. Since the summation of the productive capacities in both markets can precisely match aggregate demand (at $A_p = B_p = U_{atc}$), entry from market B into A yields an equilibrium. Firms in market B can make use of the excess capacities so as to satisfy contingent demand in market A. If the incumbent firms in market A set prices above average total cost (whether or not associated with capacities in excess of the cost-minimizing scale) in an attempt profitably to clear the market, underbidding by potential entrants from market B is induced. Alternatively, firms in market B can choose to introduce the constrained cost-minimizing productive capacity. The contingent demand in market A can then induce several (anticipative) responses. For example, the incumbent firms in market A can raise prices to the market clearing level (A_{Lp}). This only induces entry from market B if $A_{Lp} > B_{Cp}$.

III. Potential aggregate capacity shortage

A last category of scenarios indicates that both markets taken together imply shortage of productive capacity if all firms adopt unconstrained cost minimization: $(A_{ni} + B_{ni}) \cdot U_k < A_D(A_p) + B_D(B_p)$, where $A_p = B_p = U_{atc}$. This case can also be differentiated as to the capacity - demand ratios in both markets in isolation.

(1) Potentially, either market A faces excess capacity or productive capacity in market A precisely matches market demand, whereas market B shows a shortage of capacity (or vice versa): $A_{ni} \cdot U_k \geq A_D(A_p)$ and $B_{ni} \cdot U_k < B_D(B_p)$. If firms in both markets A and B retain the unconstrained cost-minimizing productive capacity and the associated average total cost prices, then firms in market A can make use of the excess capacities (if these are present) so as partly to satisfy the contingent demand in market B. However, by assumption a part of market B's demand remains unfulfilled. The (anticipative) reply to this nonmarket clearing disequilibrium can go different ways. Two examples may be indicated. First, the incumbent firms in market B can increase prices to the market clearing level (probably associated with an expansion of productive capacity beyond the unconstrained cost-minimizing scale). This can, however, induce entry from market A. Second, the firms in market B can enter with market clearing prices that exceed the level of average total cost.

(2) In both markets A and B the market capacity that follows from unconstrained cost minimization implies a capacity shortage: $A_{ni} \cdot U_k < A_D(A_p)$ and $B_{ni} \cdot U_k < B_D(B_p)$. This means that firms in both home markets can either decide to raise prices to the market clearing level (with or without an adjustment of productive capacity) or choose to leave contingent demand unsatisfied. Both strategies can induce entry from the rivals' market. For example, it can be that the market clearing level in market A exceeds the one in market B: $A_{Lp} > B_{Lp}$. An equilibrium occurs if (by accident) $A_{Lp} = B_{Lp}$. Alternatively, unsatisfied demand implies an the opportunity for rivals to enter with prices above average total cost.

The above argument is associated with two assumptions in particular: (i) firms cannot respond to entry before productive capacities are adjusted once again; and (ii) they have access to the same pool of technological opportunities. Asymmetry between the firms in both markets is likely to introduce unsustainability of the market configuration in one of the two or in both markets. Different growth rates of demand in both markets is an obvious source of asymmetry. The increased likelihood of unsustainability in a dynamic

setting is stipulated by Baumol *et al.* (1982a). An important example of a source of intertemporal unsustainability is the fact that firms have to choose between constrained and unconstrained cost minimization if they (expect to) face excess capacity when they opt for the latter strategy. The game that is associated with this dilemma is equivalent to the dumping game (Section 11.3).

Except for the solutions described in Chapter 11, there are two ways out of this dilemma. First, the flexibility of the firms can be increased by assuming that there is an interval of enlargements of productive capacity that yield second-stage cost minimization. This procedure closely resembles to Baumol *et al.*'s (1982a) flat-bottomed average cost curve. Baumol *et al.* (1982a) show in a static context that an increased length of the flat bottom of the average variable (and fixed) cost curve facilitates the sustainability of market configurations.¹ Equivalently, multiplicity of the solutions to the cost minimization procedure (21.5) increases the likelihood that market configurations will prove to be sustainable in a dynamic framework.² Second, entry costs - which by definition impede entry - facilitate the stability of markets. This brings in the role of the assumptions of the entry and exit conditions that rule the markets (see Section 21.5 below).

21.5 TYPES OF COMPETITION

By way of summary this section classifies three types of competition. The argument has so far been based upon the assumption of zero entry cost, as this condition is inextricably bound up with the perfect contestability theory. A general specification of a firm's cost function must, however, include the possibility that the transfer of output to the rivals' market is associated with entry costs (Calem 1988). Suppressing time and firm indices and assuming that $A_d = A_q$ and $B_d = B_q$, that is

$$(21.8) \quad \begin{cases} A_{atc} = (A_q + B_q) \cdot S(A_k) + A_{sc}/(A_q + B_q) + B_q \cdot A_{ec}, \text{ and} \\ B_{atc} = (A_q + B_q) \cdot S(B_k) + B_{sc}/(A_q + B_q) + A_q \cdot B_{ec}, \end{cases}$$

where $ec \geq 0$ denotes the entry cost. On the one hand, if ec has to be incurred only once, it is the setup cost of productive capacity $k = A_q$ and $k = B_q$ in markets A and B respectively. On the other hand, if ec re-occurs each time firms in market A sell output in market B (and *vice versa*), then ec reflects a transport cost.

The distinction between sunk and entry costs is not trivial. Although a positive sunk cost can be an exit barrier, it does not necessarily reflect a barrier to entry. To be precise, if incumbent firms and potential entrants *both* face the *same* scale of sunk cost, then there is no cost differential (Chapter 11). As Stigler (1968) points out, the essential feature of an entry barrier is that potential entrants have a cost disadvantage relative to

¹ Moreover, in a multimarket framework without any positive sunk cost a flat-bottomed average cost curve can imply the existence of a credible entry threat, since a firm which produces at the left hand side of the upper bound of the constant returns interval, has disposal of excess capacity that can be used to undertake "hit-and-run" entry. Equivalently, a monopolist which is producing at the left hand side of the lower bound of the constant returns interval, constitutes a credible entry threat (Section 11.3).

² The other side of the coin is that the multiplicity of cost-minimizing scales of productive capacity enhances the possibilities to install excess capacity so as to anticipate entry opportunities.

incumbent firms. Entry costs are therefore an entry barrier by definition. All this means that sunk and entry costs are crucial determinants of (the type of) potential competition.

On the basis of equation (21.8) and the equilibrium prices that would occur in market A and B if there were no entry threat (A_{Ep} and B_{Ep} respectively), three types of noncooperative (potential) competition can be distinguished. For the sake of the argument, suppose for the moment that the markets show a unique equilibrium price. That is, price heterogeneity is absent (provided that there were no entry threat).

(1) *(Perfect) contestability.* The contestability theory assumes perfect homogeneity of incumbent firms and potential entrants. Baumol (1982) uses the definition that "contestable market is one into which entry is absolutely free, and exit is absolutely costless" (Baumol 1982, p. 3; Baumol's italics). In existing contestability models this means that $A_{sc} = B_{sc} = A_{ec} = B_{ec} = 0$. Firms in both markets face neither entry nor exit barriers, which implies that cost differentials are absent: $A_{atc} = A_{avc} = B_{atc} = B_{avc}$ in both markets. Average total cost matches average variable (and fixed) cost, since sunk and entry costs are zero. This means that the threat of dumping is not credible. Excess capacity enhances the credibility of the entry threat (Van Wegberg and Van Witteloostuijn 1989), since excess capacity is used in order to enter into the rivals' market for any price in excess of average variable (and fixed) cost. A further increase of the credibility of the entry threat follows from the condition that $M_{Ep} = M_{atc}$ for $M = A$ and/or $M = B$. For this condition implies that a potential entrant is inclined to transfer home market output to the rivals' market for any price above $M_{Ep} = M_{atc}$.

(2) *Barrier market theory.* The concept of the barrier market modifies the theory of perfect contestability by allowing sunk cost to be positive without violating the free entry condition (Assumptions 10.1 and 10.2). This means that in both markets incumbent firms and potential entrants face the same average total cost: $A_{atc} = B_{atc}$ in markets A and B. A reciprocal entry threat follows from the condition that $A_{sc} = B_{sc} > 0$ and $A_{ec} = B_{ec} = 0$.¹ The threat of dumping can, however, be credible, since sunk costs are positive. Only if the dumping threat is not credible and if $M_{Ep} = M_{atc}$ for $M = A$ and/or $M = B$, can the (reciprocal) entry threat yield static efficiency (perfect barrier market). If, however, the threat of dumping is noncredible and $A_{Ep} > B_{Ep} > A_{atc} = B_{atc}$ (or vice versa), then entry deterrence in market B requires that $A_p = B_{Ep}$ (imperfect barrier market).

(3) *Barriers to entry.* Entry barriers follow from a cost disadvantage for potential entrants relative to incumbent firms: $A_{atc} > B_{atc}$ in one or both markets (or vice versa). Entry barriers can ensue from many sources (Chapter 7). *Ceteris paribus*, entry costs introduce entry barriers by definition. The other two types of costs yield a barrier to entry if they are asymmetric: $A_{avc} \neq B_{avc}$ and/or $A_{sc} \neq B_{sc}$. For $M_{sc} > 0$ (where $M = A$ and/or $M = B$) the threat of dumping can be credible. The entry decision depends on the profit opportunities offered in both markets. It is here where the prices come in. For example, if $A_{Ep} > B_{atc}$ (including an entry cost), then entry is induced. Ignoring the possibility of

¹ For $A_{sc} = B_{sc} = 0$ there is perfect contestability. For $A_{sc} \neq B_{sc}$ and $A_{ec} \neq B_{ec}$ the entry threat is likely to be asymmetric. Firms in either market A or market B have a cost advantage relative to their rivals if the entry cost is ignored. For example, suppose that $A_{sc} > B_{sc}$. From the assumptions of the barrier market it follows then that $A_{ec} < B_{ec}$. This implies that $A_{atc} = B_{atc}$ only holds for particular scales of entry. Otherwise, the entry threat is unidirectional.

dumping, entry deterrence requires limit pricing: $A_p = B_{atc}$ (including entry cost).

The typology of forms of competition can be connected with the scenarios described in Section 21.4.3. Two examples illustrate this point. First, scenarios I and III can be associated with entry barriers. If firms in the two markets decide differently on productive capacity, a cost differential is generated. Scenario I.1 shows that the lower-cost firms can benefit from the entry barriers that the higher-cost rivals are facing by either increasing prices in the home market or by undertaking profitable entry. Scenario I.2 points to the trap of reciprocal dumping, which leaves all firms worse off. Scenarios III.1 and III.2 indicate that partial entry can mitigate the tendency to high prices and positive profits. Not surprisingly, excess capacity induces a tendency to lower-priced entry or entry deterrence (scenarios I), whereas a shortage of productive capacity generally induces higher prices, since entry can only be partial (scenarios III).

Second, the category of scenarios II indicates circumstances that can be associated with a barrier market. Scenario II.1 describes a barrier market without entry, whereas Scenario II.2 can yield barrier market outcomes with entry. In both cases there is a tendency to static efficiency. The third type of potential competition - (perfect) contestability - is not dealt with in this section, since sunk cost is assumed to be positive. Van Wegberg and Van Witteloostuijn (1989) offer an examination of perfect contestability in a multimarket framework.

The argument in this chapter points out that the modeling of potential competition in a multimarket setting offers rich opportunities for future research. The game on (non)expansion can be analyzed in different settings. In particular, the influence of (i) the relative sizes of the markets, (ii) the threat of reciprocal entry, and (iii) the (time) structure of the game can be examined. The preliminary investigation of two-stage Cournot-Bertrand games shows that a large number of scenarios can occur, depending on exogenous (demand) developments and the strategic choices of the rivals. In effect, the type of (potential) competition (that is, entry barriers and the credibility of the entry threat) can appear as a result rather than an assumption of the models, which offers a point of contact to study the evolution of competition. Such a dynamic multimarket setting seems to be the natural frame of reference to the analysis of the static - dynamic efficiency dilemma and the contestability and sustainability of markets.

சென்னை நகரில் உள்ள பஞ்சவர்ணம்

சென்னை நகரில் உள்ள பஞ்சவர்ணம்

சென்னை நகரில் உள்ள பஞ்சவர்ணம்

சென்னை நகரில் உள்ள பஞ்சவர்ணம்

சென்னை நகரில் உள்ள பஞ்சவர்ணம்

சென்னை நகரில் உள்ள பஞ்சவர்ணம்

சென்னை நகரில் உள்ள பஞ்சவர்ணம்

PART

X

EVOLUTION

Competitive Dynamics and Intertemporal Welfare

KEY WORDS. *Parts IV, V, VI, VII, VIII and IX have presented models which are basically static: that is, they describe the one-shot strategy choice of firms or workers (although both take future considerations into account). Part X supplements the argument by putting the barrier market theory in a dynamic and welfare-theoretic perspective. Chapter 22 introduces the issue of dynamic economics by reviewing intertemporal frameworks. Moreover, the pros and cons of simulation techniques are briefly examined. Chapter 23 offers a simulation experiment of competitive dynamics that gives a first impetus to the study of a dynamic barrier market. Chapter 24 reviews the literature on welfare norms so as to evaluate barrier market theory. Chapter 25 summarizes the basic message of this thesis.*

 

[illegible]

CHAPTER

22

CHANGES OVER TIME: Dynamics and Welfare

22.1 EVOLUTION

Parts VI and VII have described two-period decision models of economic agents in a barrier market setting. Therefore, the focus is on one or two-shot games. The same applies to existing models of perfectly contestable markets (Chapter 6) and most models of entry barriers (Chapter 7) and innovation (Chapter 8). Studies of dynamic adjustment processes are scarce. An exception that confirms the rule is the work in the Schumpeterian evolutionary tradition (Nelson and Winter 1982). The state of the art in general (dis)equilibrium theory is not much different. With the exception of the neoclassical theories of business cycles and growth, models of sequence economies and classical models of intersectoral competition, the analysis of intertemporal dynamics is generally missing. This is an important omission as an essential element in the working of an economy are changes over time. The *evolution* of the economy in a competitive environment requires the study of *competitive dynamics*.

The plan of the chapter is as follows. Section 22.2 introduces the essential elements of competitive dynamics, which characterizes evolution in a competitive environment. Section 22.3 evaluates the intertemporal frameworks that are put forward in the literature. Section 22.4 argues that the features of competitive dynamics are difficult to analyze mathematically (Nelson and Winter 1977a). Therefore, the use of simulation techniques is suggested. If changes over time are considered, welfare evaluations become increasingly tedious. Section 22.5 offers concluding remarks on intertemporal welfare.

22.2 COMPETITIVE DYNAMICS

In industrial organization it is *competitive* dynamics that matters. The focus of competitive dynamics is on interactive competitive struggles over time. The key features of competitive struggles are heterogeneous agents which not only act, but also *react* in an environment ruled by uncertainty. Therefore, the *nonsimultaneity* of decision making of *heterogeneous* agents grasps the essentials of competitive dynamics. In the context of competitive dynamics assumptions of representative or identical agents imply that the baby is thrown away with the bath water, simply because competition in a market is based upon interactions among agents which differ.

The essentials of competitive dynamics are described by, for instance, neo-Austrian (for example, Hayek 1946 and Kirzner 1985) and Schumpeterian (for example, Schumpeter 1943 and Nelson and Winter 1982) economists. Kirzner argues that

[w]e see the market as made up, during any period of time, of the *interacting* decisions of consumers, entrepreneur-producers, and resource owners. ... The market *process*, then, is set in motion by the results of the initial *market-ignorance* of the participants. The process itself consists of the systematic *plan changes* generated by the flow of market information released by market participation - that is, by testing of plans in the market (Kirzner 1973, pp. 9-10; italics added).

Hence, *expectational errors* drive the economic system forward. However, neither neo-Austrian economists - for principal reasons - nor Schumpeterian theorists - because of conceptual and technical difficulties - have (yet) succeeded in developing a formal analysis of competitive dynamics.

The importance of economic agents' interactions is of course recognized in the literature on industrial organization (Part III). For example, models of entry-detering strategies focus by definition on nonsimultaneous decision making, because it is argued that incumbent firms in a market have a first-mover advantage relative to potential entrants (Spulber 1981, Gilbert and Newberry 1982, Fudenberg *et al.* 1983 and Gilbert 1986). However, although these models introduce heterogeneous agents and nonsimultaneous decisions, they generally focus on one, two or three-stage games. So, these models

fail to situate [the firms' activities] in a long-term evolutionary process of industrial development (Iwai 1984a, p. 160).

What matters is the evolution of competitive economies. The next section examines the intertemporal frameworks in the literature which aim to capture dynamic economics.

22.3 INTERTEMPORAL FRAMEWORKS

Hicks argues that dynamics can be introduced by bringing a succession of periods into operation (Hicks 1956, reprinted in 1982, p. 218). A sequence analysis consists of two elements,

in the first place, what we may call *single period theory*, theory which is concerned with the determination of what happens in a single period ..., and secondly what we may call *continuation theory*, which is concerned with the effects of the events of a first period upon the expectations and plans which themselves determine the events of its successors (Hicks 1956, p. 223).

Hicks' methods of dynamic economics are characterized by the features of the single period theory. The essential features of a single period theory are basically associated with the assumptions of equilibrium that are imposed: is the Hicksian (dis)equilibrium Walrasian or non-Walrasian in nature?

A further differentiation follows from distinguishing the short from the long run. Although Marshall (1920) relates the short and long run to calendar time, the division between short and long periods is known as a Marshallian analytical device. Marshall argues that

[a]s regards *normal* prices, when the term Normal is taken to relate to *short* periods of a few month or a year, Supply means broadly what can be produced for the price in question with the existing stock of plant, personal and impersonal, in the given time. As regards *normal*

prices, when the term Normal is to refer to *long* periods of several years, Supply means what can be produced by plant, which itself can be remuneratively produced and applied within the given time (Marshall 1920, pp. 314-315; Marshall's italics).

This can be interpreted as the short run being characterized by constant productive capacity and the long run being associated with variable productive capacity.¹ Broadly speaking, it is the assumptions of the variability and mobility of productive capacity that determine the short or long-run nature of a model. In the short run only price and/or quantity adjustments can take part in adjustment processes, whereas in the long run changes in (the location, scale and/or composition of) productive capacity enter the scene. The literature's intertemporal frameworks can be evaluated in Walrasian, Hicksian and Marshallian terms (Van Witteloostuijn 1990e).

(i) *Economics of the choice-theoretic framework.* The established pillar of neoclassical economics is general equilibrium theory, both in the Walrasian and non-Walrasian variant (Section 20.2).² General Walrasian equilibrium theory is clearly based upon Hicks' temporary equilibrium method. In a unit period a price *tâtonnement* smoothes the adjustment to a temporary market clearing equilibrium state. Temporary expectations are confirmed. The temporary Walrasian equilibrium models focus on the short run, since it is only prices that take part in the adjustment to temporary equilibrium. A sequence of temporary Walrasian equilibria can be used to describe the long-run evolvments. Here intertemporal equilibrium models and (real and monetary) business cycle theory (either by introducing rational expectations and price rationing or by applying nonlinear dynamics) are worth mentioning (Lucas 1972 and 1975, Gale 1973, Benhabib and Nishimura 1979 and 1985, Barro 1980 and 1986, Day 1982 and 1983, Grandmont 1985, Boschen and Mills 1988, Benhabib and Laroque 1988, Plosser 1989 and Mankiw 1989).

Two well-known manifestations of Keynesian type economics are IS-LM and fixprice models. IS-LM and fixprice models seek to derive Keynesian results (deficient demand and unemployment) in a macroeconomic and microeconomic framework respectively. In both cases the key argument is that price and wage rigidities prevent price adjustments from being complete. That is, quantity adjustments take part in the burden of adjustment. Fixprice models are based upon Hicks' fixprice method.³ A quantity *tâtonnement* takes care of the adjustment to a temporary non-Walrasian equilibrium state. That is, markets can fail to clear. Temporary fixprice models focus on the short run, since only quantities are adjusted to reach a temporary equilibrium state. (A similar argument applies to the models that assume price setting along with temporary non-Walrasian equilibria.) A sequence of temporary non-Walrasian equilibria can capture the intermediate and long run. This is clear from disequilibrium growth theory (Lorie 1978, Böhm 1978,

¹ Marshall clearly recognizes the drawbacks of his period structure. For example, "[o]f course there is no hard and sharp line of division between 'long' and 'short' periods" (Marshall 1920, p. 314). However, in relation to the short run he points out that "[i]t is true that we do treat variables *provisionally* as constants. But it is also true that this is the only method by which science has ever made any great progress in dealing with complex and changeful matter" (Marshall 1920, p. 315; Marshall's italics).

² Van Ees and Kuipers (1989) offer a survey of the post-war development of macroeconomics in which the neoclassical synthesis, and the controversy regarding it, is crucial.

³ For example, Malinvaud (1977) indicates that his model "is going to be a 'fix-price' theory according to the denomination proposed by J. Hicks" (Malinvaud 1977, p. 12). However, Hoogduin and Van Witteloostuijn (1987b) argue that Hicks (1965) introduces *temporary uncertainty* in his fixprice method in line with Swedish economics.

Honkapohja 1979, Ito 1980, Malinvaud 1980, Kuipers 1981, Honkapohja and Ito 1982 and Eckwert and Schittko 1988).

(ii) *Swedish economics*. The Swedish Stockholm school flourished in the first four decades of this century, but died a silent death in post-war years (Siven 1987). An important contribution of the Swedish school is the description of a *disequilibrium* sequence analysis.¹ Lindahl (1939) criticizes the equilibrium method by pointing out that the

advantage of this method is that the entire static apparatus may be employed in the analysis of a dynamic sequence.² It thus bridges the gap between statics and dynamics. The cumbersome *ex ante* and *ex post* terminology become superfluous, for the individuals are assumed to have knowledge at the beginning of the period of all the transactions and of the relevant prices valid for the period. ... In a real dynamic development such equilibria are probably found only sporadically (Lindahl 1939, p. 69).³

Therefore, Swedish economists have analyzed a series of Hicksian temporary disequilibria by means of the introduction of temporary uncertainty:

The foregoing analysis of the pricing process as a series of disequilibria, rested on the realistic assumption that prices quoted in the market are regarded as the supply prices of sellers (or in certain exceptional cases as the demand prices of buyers). These prices are, it is true, based on sellers' anticipations of the magnitude of demand at different prices, but the anticipations are often more or less false. It are the deviations between the transactions anticipated by sellers and those activities carried out, and the associated changes in stocks and orders, which are the most important factors influencing the decisions of sellers to alter their prices from one period to the next (Lindahl 1939, pp. 64-66).

So, the basic element of the Swedish period analysis is a Hicksian temporary disequilibrium which can be Walrasian and non-Walrasian in nature. Temporary expectations are generally disappointed, which can go hand in hand with nonmarket clearance. The temporary disequilibrium deals with the short run: it is the prices and quantities that determine the nature of the temporary disequilibrium. A sequence of temporary disequilibria permits the study of the long run. For example, Lundberg (1937) offers a treatment of business cycle and growth theory. Recently, the Swedish method of dynamic economics has been employed by Snippe (1985) and Duménil and Lévy (1987a and 1987b).

¹ Ohlin (1937) and Siven (1987) even argue that the Swedish contribution to economics offers theories of macroeconomics and unemployment that parallel Keynes'. For example, Siven points out that 'Lundberg's (and consequently the Stockholm School's) independence of, and methodological superiority to Keynes was subsequently emphasized by Schumpeter (1954, pp. 1173-1174)' (Siven 1987, p. 579).

² Along the same lines Arrow (1978) argues that Hicks 'introduced the device of regarding the commodities at different dates as different commodities, which permits the already developed tools of static analysis to be applied to intertemporal decisions' (Arrow 1978, p. 158).

³ Before the Swedish period analysis reached its ultimate form, as described in the first chapter of Lindahl's *Studies in the Theory of Money and Capital* (1939), many Swedish economists contributed to its development. Worth mentioning are Wicksell (cumulative process, 1936), Myrdal (*ex ante* - *ex post* terminology, 1939), Hammarskjöld (definition of the unit period, 1932 and 1933) and Lundberg (sequence analysis, 1937).

(iii) *Classical economics*. Ever since the publication of Garegnani's (1976) treatment of the equilibrium notion and Nikaido's models (1977 and 1983) of Marxian competition, the classical notion of competition has gone through a revival (Duménil and Lévy 1985 and Semmler 1986). The classical theory of particularly Smith (1776), Ricardo (1817) and Marx (1863-1867) is modeled with the help of modern (mathematical and simulation) techniques. The key argument is that, *ceteris paribus*, there is a long-run equilibrium state that operates as a center of gravitation (which is, however, generally not reached). That is,

[t]he natural prices are determined by their long-run components. These components, however, are regarded as independently determined from the supply and demand mechanism. Yet, it is assumed that there is a tendency to equalise the rates of return on the factors used up in production, enforced by the possibility of the factors to move from areas of low to high returns (Semmler 1984, p. 133).

So, the mobility of productive capacity is the driving force behind the tendency toward a long-run equilibrium (center of gravitation) state. To be precise,

[t]he intersectoral competition of capitals results in investment flows, bringing about a tendency to equalise the rates of profit across industries (Semmler 1984, p. 136).¹

Price adjustments can be induced by a difference between either productive capacity and actual sales (Nikaido 1983) or expected and actual sales (Duménil and Lévy 1987a and 1987b) in a period. The first price adjustment procedure can be associated with market clearing and the confirmation of expectations, whereas the second price adjustment rule can imply a non-Walrasian disequilibrium. The latter dynamic method closely resembles the Swedish period analysis: a sequence of short-run temporary non-Walrasian disequilibria (that is, nonmarket clearing and disappointed expectations) converges to a long-run intertemporal equilibrium position, where firms act as price setters which adjust prices in response to (unexpected) inventory movements (for example, Duménil and Lévy 1987a, p. 146).

(iv) *Post-Keynesian economics*. Post-Keynesian economics argues that neo-Keynesian theory ignores the essentials of Keynes' theory by offering a framework that only offers disguised forms of uncertainty and dynamics. Broadly speaking, two branches of post-Keynesian economics can be distinguished, one focusing on long-run gravity centers and

¹ It can be argued that the assumed contradiction between classical and neoclassical theory is not as drastic as modern classical economists suggest, since the neoclassical framework permits the introduction of price adjustments and intersectoral competition through capital mobility (see, for example, Walras, 1874). In this respect, the neoclassical theory of industrial organization (Part III) speaks for itself. Moreover, the modern classical models now and then take the shape of limit cycle studies by using nonlinear dynamics (Semmler 1986). The main difference between the classical and neoclassical notion of competition lies in the role of supply and demand in determining the long-run constellation of relative prices. The usual classical assumption that the long-run equilibrium prices are independent of demand contradicts neoclassical theory. However, this standpoint can again be modified, since the nonsubstitution theorem indicates that the classical position is a special case of neoclassical theory (Eijgelshoven 1982).

the other concentrating on the short run (Carvalho 1984).¹

Gravity center models are only concerned with long-run equilibrium positions.² The argument is that the short run is unimportant, because the short run only exerts accidental influences. That is, the crucial belief is that short-run influences are only temporary. The gravity center argument can be applied to given equilibrium states (Garegnani 1976, 1978 and 1979 and Eatwell 1977) or growth paths (Kaldor 1970 and Pasinetti 1970 and 1983).³ Dynamics comes in through the study of the implications of structural changes to the long-run equilibrium. This is what Eichner and Kregel call comparative dynamics (1975, pp. 1295-1296). So, the gravity center approach focuses on the Marshallian long run only. The series of short-run (dis)equilibria that lead to the long-run position is ignored. In so far as uncertainty plays a role, the state of long-run expectations is a datum: that is, it is one of the determinants of the long-run equilibrium. This suggests that the long-run equilibrium is of a Hicksian nature.

The second branch of post-Keynesian economics takes the opposite position by focusing (almost) exclusively on the short run (Shackle 1968, 1969 and 1972, Davidson 1972 and 1982, Minsky 1975 and 1982 and Kregel 1976 and 1982). This follows from an interpretation of Keynes' contribution to economics that lays a strong emphasis on the implications of (fundamental) uncertainty to the economic process.⁴ For example, Kregel (1976) argues that the essential influence of uncertainty comes in through the impact of the state of expectations on Keynes' three crucial variables: propensity to consume, marginal efficiency of capital and liquidity preference. Since (particularly long-run) expectations are volatile, the economic system is subject to instability. That is, the convergence to a long-run equilibrium state is absent. The future cannot be known, because the economic process is nonergodic.

(v) *Neo-Austrian and Schumpeterian economics.* The Austrian school started with the work of Menger (1871) and his associates Von Wieser (1884 and 1898) and Von Böhm-Bawerk (1889). Neo-Austrian economics took shape in the first half of this century through the contributions of the Austrian economists Von Mises (1912 and 1949) and Hayek (1944 and 1949). Prominent modern neo-Austrian economists are Lachmann (1969 and 1977), Kirzner (1973, 1979 and 1985) and Rothbard (1962, 1963 and 1977). The distinctive features of neo-Austrian economics are the focus on the dispersion of knowledge, ignorance, creative and alert entrepreneurship and dynamic discovery processes.

¹ Moreover, there is a branch of post-Keynesian economics that focuses on topics of industrial organization (Eichner 1976 and 1979). The focus is on oligopolistic market structures in which firms seek to grow by determining markups that generate internal funds for investment. This indicates the assumption that firms act as price setters. Here these topics are ignored.

² Post-Keynesian economics resembles the classical view on competition and long-run equilibrium. However, the question that is central in classical economics, the stability of the long-run equilibrium state, is generally ignored (or just solved by assumption) in post-Keynesian economics [an exception being Pasinetti (1982)].

³ Two remarks can be made. First, apart from Keynes' the work of Sraffa (1960) often constitutes a source of inspiration to post-Keynesian economists. Second, Kalecki (1943, 1954 and 1971) takes an intermediate stance by focusing on both cyclical disturbances and long-run trends. That is, "[i]n fact, the long-run trend is but a slowly changing component of a chain of short-run situations" (Kalecki 1971, p. 165).

⁴ For an account of the post-Keynesian interpretation of Keynesian uncertainty (and the implication to the role of money) see Hoogduin (1987) and Hoogduin and Van Witteloostuijn (1987a) and the references therein.

A clear statement as regards the neo-Austrian attitude toward the notion of equilibrium is offered by Hayek (1937). Hayek argues that

any change in the relevant knowledge of the person, that is, any change which leads him to alter his plan, disrupts the equilibrium relation between his actions taken before and those taken after the change in his knowledge ... In other words, the equilibrium relationship compromises only his actions during the period in which his anticipations prove correct (Hayek 1949, p. 36)

That is, Hayek clearly refers to a Hicksian equilibrium notion. However, although Hayek asserts that the equilibrium concept has a clear meaning if applied to a single economic agent (1949, p. 35), this meaning is useless in a general framework with many interacting agents: different economic agents have different knowledge¹ so that the Hicksian equilibrium notion is of no use in a general setting.

The dispersion and imperfection of knowledge are the very reasons why neo-Austrian economics frames competition as a dynamic discovery process in which alert entrepreneurs are the crucial economic agents. Kirzner (1973) points out that

[n]ot all the decisions in a given period can be carried out, since many of them may erroneously anticipate and depend upon other decisions which are in fact not being made. ... In short, ignorance of the decisions which others are in fact about to make may cause decision-makers to make unfortunate plans - either plans that are doomed to disappointment or plans which fail to exploit existing market opportunities. During the given period of time, exposure to the decisions of others communicate some of the information these decision-makers originally lacked. ... This newly acquired information concerning the plans of others can be expected to generate, for the succeeding period of time, a *revised set of decisions* (Kirzner 1973, pp. 9-10; Kirzner's italics).

Hence, Kirzner describes the market process in terms that closely resemble the Swedish period analysis: a series of temporary Hicksian disequilibria², where the entrepreneurs can generally set the prices. Neo-Austrian economists do not bother much about the distinction between Hicksian and Walrasian equilibrium features. As already mentioned, the tendency of equilibrium is induced by entrepreneurial activity. The entrepreneurial process can be of both a short and long-run character (Kirzner 1985, pp. 68-69).

Although neo-Austrian and Schumpeterian economics share a clear family resemblance, Schumpeterian economics departs from the neo-Austrian view by focusing on the disequilibrating activity of entrepreneurs (Kirzner 1979, p. 131). The Schumpeterian notion of competition has been reviewed in Chapter 8, so here the treatment of Schumpeterian economics can be brief. The modern Schumpeterian view on dynamic methods is clearly expressed by Winter (1984):

¹ Knowledge is not only dispersed, but it is also of a subjective nature and subject to fundamental uncertainty. This view is clearly expressed by Lachmann (1977). It is particularly at this point that post-Keynesian and neo-Austrian economics show a close resemblance. As O'Driscoll and Rizzo (1985) note: "It is evident that there is much common ground between post Keynesian subjectivism and Austrian subjectivism. Cross-fertilization between these two schools is, however, exceedingly rare, although the possibilities for mutually advantageous interchange seem significant" (O'Driscoll and Rizzo 1985, p. 9).

² Broadly speaking, neo-Austrian economists are averse to the use of mathematics (Shand 1984, pp. 15-31). Littlechild and Owen (1980) present a rare model of the neo-Austrian entrepreneurial market process.

Evolutionary economic models are, of course, intertemporal models. More specifically, they typically fall under the rubric of 'temporary equilibrium' models. Some economic processes are conceived as working very fast, driving some of the model variables to (temporary) equilibrium values within a single period. ... For example, in the model used in this paper, a short-run equilibrium price of output is established in every time period. Slower working processes of investment, and of technological and organizational change, operate to modify the data of the short-run equilibrium system from period to period (Winter 1984, p. 288).

To be precise, Schumpeterian economics models a series of temporary Walrasian equilibria so as to describe long-run intertemporal adjustments. The long-run state is one of disequilibrium. Intertemporal uncertainty drives the economic process forward:

Information imperfections, and informational differences among actors, are not complications of the basic structure, but are central to the Schumpeterian scheme. ... The gains obtainable by guessing better and acting sooner ... are the crucial motive power and adaptive mechanism of a system that is permanently in [intertemporal] disequilibrium. And, because it arises from a continual unfolding of unanticipated possibilities, the disequilibrium is disequilibrium in the fundamental sense: Expectations are not being realized, mistakes are being made and corrections attempted (Nelson and Winter 1977a, p. 271).

22.4 SIMULATION TECHNIQUES

Section 22.3 indicated that the sequential method of dynamic economics dominates the literature. The question is, however, how far the methods of dynamic economics provide tools that permit the study of the kinds of problems which are central to industrial organization. The essentials of competitive processes - expectational errors, heterogeneity of agents and nonsimultaneous decision making (Section 22.2) - throw up technical difficulties when only *mathematical* techniques are employed. Of course, this is not to say that mathematics is an inappropriate tool to analyze some aspects, or simple cases, of competitive processes. However, in *addition* theorists have introduced *simulation* techniques. Formal models of the course of economic processes over time in which expectational errors drive the system forward, are the exception rather than the rule in the economic literature. Three rare examples are Snippe's (1985) Swedish period analysis, Nelson and Winter's (1982) Hicksian sequence analysis and Duménil and Lévy's (1987a) classical dynamics. However, these frameworks prove to be too complicated to be studied with mathematical tools only, so that simulation techniques are introduced in order to further the analysis of the dynamics of the models. For illustrative purpose, Nelson and Winters' argument is examined more closely.

Nelson and Winter's (for example, 1974, 1982 and 1984) contribution to Schumpeterian economics represents a well-worked out attempt to model and analyze the dynamics of competitive processes. Nelson and Winter adapt the evolutionary framework to modeling by adopting a Hicksian equilibrium sequence analysis. In order to bypass the technical difficulties that arise when the dynamics of Schumpeterian evolutionary processes are analyzed mathematically, they adopt simulation techniques. The analysis of (intertemporal) expectational dynamics is difficult to accomplish with the help of mathematical tools only.¹ The technical complexities arise from the essentials of ongoing

¹ For instance, Kuipers (1981, p. 101) and Kamien and Schwartz (1982, p. 199) also point to the limitations imposed by mathematical analysis.

competitive processes; expectational errors are the driving force of continual change by generating incentives to revise plans and adjust behavior. Nelson and Winter (1977a) argue that it

is plausible that the taste of developing formal models is intrinsically more difficult in Schumpeterian theory than in modern orthodoxy ... the most powerful abstractions and simplifications of orthodoxy are inappropriate or ineffective in the Schumpeterian context. As we have argued elsewhere, the absence of formal theory probably accounts for the relative neglect of Schumpeterian ideas (Nelson and Winter 1977a, p. 271).

Nelson and Winter's key argument is that the modification of assumptions so as to make the derivation of analytic results possible with the help of mathematical tools, is generally inferior to running appealing simulations:

However, in many such cases it is impractical or impossible to obtain analytic results from the model the analyst has in mind. Rather than change assumptions and then analyze the wrong model, he may prefer to simulate the right one (Nelson and Winter 1977a, p. 272).

Moreover, simulation techniques offer clear advantages over mathematical analysis:

Freedom from the tractability constraints of available analytical techniques remains the major advantage of simulation; this freedom can be exploited in a preliminary exploration of a variety of alternative model formulations. Simulation also makes possible the sort of instruction of one's intuition that comes from careful scrutiny of detailed quantitative examples; this instruction is highly important when the models under study are stochastic processes, for such models often behave in ways that violate uninstructed intuition. Study of a range of examples may also lead to the perception of unanticipated regularities, and thus the development of new hypotheses about the behaviour of the model (Nelson and Winter 1977a, p. 272).

This is not to say that the use of mathematics must be avoided. A major advantage of mathematical over simulation techniques is that the former permit the derivation of more general and definite results.¹ Of course, mathematical and simulation techniques are complementary rather than contradictory.² The traditional opposition against simulation techniques centers upon the casuistic character of simulation experiments, because functions have to be specified and parameter values have initially to be set. However, a

¹ As Nelson and Winter (1977a) put it: "Of course there are costs involved in working with a simulation rather than an analytical model. The most obvious of these is that the results are of uncertain generality. If there is a large domain of interesting independent variables and parameters to explore, it is virtually impossible to explore all parts of it. ... In addition, it is sometimes argued that the very freedom of assumption afforded by simulation modeling invites sloppiness, whereas the analytic approach imposes a need to make judgments as to which considerations are of central importance. In our view, the most serious problem with many simulation models, although the characteristic is shared by many analytic models, is lack of transparency. The model yields results that are not easy to understand" (Nelson and Winter 1977a, p. 272). The lack of transparency can, however, be avoided if a careful simulation strategy is followed. That is, the model has to be evolved from simple to complex on the basis of a well-designed step-by-step procedure; for example, only one variable, parameter or condition should be varied per simulation run.

² "It would be a mistake, however, to pose the issue in terms of either-or. In the context set by specific questions about a specific model, the two approaches may be substitutes, but they are clearly complementary in the sort of exploratory effort [i.e., Schumpeterian evolutionary competition] we have under study" (Nelson and Winter 1977a, p. 272).

stand against this critique can be put into words with the help of two additional arguments. First, leaving functions unspecified and parameter values unchosen easily results in overlooking important phenomena and difficulties (like the modeling of information flows, diverging expectations and allocation mechanisms).¹ Second, a sensitivity analysis of a model with regard to the function specifications and parameter values employed can bring general results to the surface. The simulation experiments have then to be submitted to statistical tests.²

22.5 A COMPLICATION: Intertemporal Welfare

This chapter emphasizes the importance of the study of competitive dynamics. The evolution of a competitive economy is inextricably bound up with changes over time. Chapter 23 describes a first example of the competitive dynamics in a barrier market by making use of simulation techniques. However, changes over time complicate welfare analysis. During the process of evolution some market participants improve their position, whereas other rivals become worse off. A normative judgment of the welfare aspects of evolution requires the examination of *intertemporal* welfare (changes). The question of welfare norms which take account of distributional shifts is central to much literature on welfare theory. Chapter 24 evaluates this literature.

¹ Here Nelson and Winter (1977a) argue that "[s]imulation, on the other hand, can be a useful adjunct to an analytical approach. It can establish, with the same finality as a theorem, the logical consistency of the model's assumptions with a set of propositions about its behavior. And while it offers a way around the tractability constraints of analytic methods, it imposes its own constructive discipline in the modeling of dynamic systems: the program must contain a complete specification of how the system state at $t + 1$ depends on that at t and exogenous factors, or it will not run" (Nelson and Winter 1977a, p. 272).

² A nice introduction to simulation techniques and the associated statistical tests is offered by Mitrani (1982). Mitrani stipulates the crucial importance of statistical testing by pointing out that "it is not enough to write a correct and efficient program; one should also know how to use it to answer questions about the system being simulated. This is where the science comes in. A simulation study is essentially a statistical experiment, requiring statistical tools for the collection and analysis of data" (Mitrani 1982, p. vii). The simulation experiments presented in this thesis have, by the way, not (yet) been submitted to statistical tests, since these experiments are only presented with the purpose to illustrate the opportunities offered by simulation techniques (in order to undertake the study of competitive dynamics). The execution of detailed and worked-out simulation experiments is postponed to a future date (as yet unknown): it is, however, prominently listed on the research agenda.

CHAPTER

23

COMPETITIVE DYNAMICS: Diverging Expectations, Entry and Exit Movements and Barriers

23.1 INTRODUCTION: A Simulation Experiment

Chapter 20 introduced intermarket spillovers and temporary dynamics (through enter and exit movements) in a model with a barrier commodity and labor market. This chapter moves the analysis one step further by examining *intertemporal* dynamics (in the commodity market only). It appears that entry and exit movements are also crucial to coordination in an intertemporal context. Besides, the essential elements of *competitive* dynamics are introduced. Section 22.2 argued that the features of competitive dynamics are difficult to analyze mathematically (Nelson and Winter 1977a). Therefore, here simulation techniques are employed.¹

For illustrative purposes, this chapter describes five simple simulation experiments: (i) an unsustainable contestable market with increasing demand; (ii) an unsustainable contestable market with decreasing demand; (iii) a sustainable contestable monopoly with costless expansion of productive capacity in an environment with perfect foresight; (iv) an unsustainable contestable market with costless expansion or reduction of productive capacity in an environment with uncertainty; and (v) an unsustainable perfect barrier market with costly process innovation and technical uncertainty (Van Witteloostuijn and Maks 1988). In advance three assumptions (or abstractions) have to be made clear. First, the simulation experiments focus on the commodity market only. Vertical intermarket spillover effects are therefore ignored (Chapter 20). However, in the last experiment horizontal intermarket spillovers arise through joint R&D economies (Chapter 21). Second, multimarket competition is described in an elementary framework by leaving potential entrants unidentified and focusing on one-sided entry (Chapters 11 and 19). Third, the models assume temporary market clearing equilibria (Chapter 20). Expectational dynamics can only enter the picture through intertemporal uncertainty [experiments (iv) and (v)].

The plan of the chapter is as follows; Section 23.2 specifies a two-period decision model along the lines of Part VII; Section 23.3 describes experiments (i) and (ii); Section 23.4 examines experiments (iii) and (iv); Section 23.5 presents the results of experiment (v); and Section 23.6 offers concluding remarks. It must be noted in advance that the simulation experiments are extremely simple and that most of the results are trivial. The

¹ The simulation package used is DEMOS (Birtwistle 1979).

experiments, however, only seek to show that the methodology employed here permits the analysis of some crucial elements of dynamic competition that are generally ignored in the (mathematical) literature. The techniques described below can be used to analyze increasingly complicated models from a (competitive) dynamic perspective. In this respect, simulation experiment (v) is the most interesting one.

23.2 PRICE, QUANTITY AND INVESTMENT

This section describes the decision rules of a firm in a barrier market at a particular point in time, given its knowledge and subjectively certain point expectations [of demand growth (Section 23.4) and technological opportunities (Section 23.5) in particular]. The two-period decision model of firms focuses on a current period t and a future period $t+1$. A firm is engaged in temporary and intertemporal decision making, the former focusing on current period price and quantity setting and the latter on investment in productive capacity that yields results in the future period.

In the context of contestability theory, Baumol *et al.* (1982a) take note of the special case of the flat-bottomed average cost curve. Assume a Cobb-Douglas production function that is associated with switching of the type of returns:

$$(23.1) \quad q_t = \epsilon_t \cdot L_t^\alpha \cdot M_t^\beta, \text{ with}$$

$$(23.1.a) \quad \text{if } q_t < L_{q_t}, \text{ then } \alpha + \beta > 1,$$

$$(23.1.b) \quad \text{if } L_{q_t} \leq q_t \leq R_{q_t}, \text{ then } \alpha + \beta = 1, \text{ and}$$

$$(23.1.c) \quad \text{if } q_t > R_{q_t}, \text{ then } 0 < \alpha + \beta < 1,$$

where L_t denotes units of labor, M_t units of material (for example, energy), α labor elasticity and β material elasticity of production and where $L_t, M_t \geq 0$. From equation (23.1) the average variable cost function can be derived:

$$(23.2) \quad \text{avc}_t = \text{mavc}_t \cdot q_t^\sigma, \text{ with}$$

$$(23.2.a) \quad \text{mavc}_t = (\alpha + \beta) \cdot [(w_L^\alpha \cdot w_M^\beta) / (\epsilon_t \cdot \alpha^\alpha \cdot \beta^\beta)]^{1/(\alpha + \beta)},$$

$$(23.2.b) \quad \begin{cases} \text{if } \alpha + \beta \neq 1, \text{ then } \sigma = (1 - \alpha - \beta) / (\alpha + \beta), \text{ and} \\ \text{if } \alpha + \beta = 1, \text{ then } \sigma = 0, \end{cases}$$

where w_L and w_M are the prices per unit of labor and material respectively.¹ Both are assumed to be determined outside the model and to be known to the firms.

For the moment, ignore a positive sunk cost. In the current period price and quantity decisions are determined by cost and demand features. *Ceteris paribus*, the quantity demanded from a firm follows from (suppressing indices) $d = D(p)/n_i$. For (i) $L_{q_t} \leq d_t \leq R_{q_t}$ and $n_i > 1$ and (ii) $d_t \leq R_{q_t}$ and $n_i = 1$ the firm equates d_t and q_t so as to offer

¹ Of course, the variable cost function can also capture a fixed component.

a price equal to the (minimum) average cost level, since a higher price induces entry and a lower price yields losses (Proposition 14.2).¹ So,

$$(23.3) \quad q_t = d_t, \text{ and}$$

$$(23.4) \quad p_t = \text{avc}_t.$$

The case where $d_t < L_{q_t}$ and $n_t > 1$ is associated with unsustainability (Section 23.3).

If $d_t > R_{q_t}$, then the firm cannot decide on a price - quantity combination that simultaneously guards against losses, deters entry and for which equation (23.3) applies. Therefore, a part of individual demand is not satisfied by deciding to offer R_{q_t} , because then the firm admits entry without suffering any loss. That is,

$$(23.5) \quad q_t = R_{q_t}, \text{ and}$$

$$(23.6) \quad p_t = \text{mavc}_t.$$

Entry is induced, since residual demand ($d_t - R_{q_t} > 0$) offers profit opportunities (Section 23.3). Equation (23.5) reflects a specific assumption which bypasses the indeterminateness of the entry-detering quantity if the cost function is flat-bottomed (Chapter 6). A rationale for assumption (23.5) is that a plausible secondary goal (that is, after the primary goal of profit maximization is satisfied) of firms may be to maximize size or sales (Baumol 1959 and 1962).

The simulation experiments focus on two types of investment: (i) a costless expansion or reduction of productive capacity; and (ii) a costly process innovation. First, the need to expand (or reduce) capacity can occur if demand grows (or shrinks) from period t to period $t+1$. Suppose, for the sake of simplicity, that market demand is price inelastic:

$$(23.7) \quad D_{t+1} = (1+g) \cdot D_t,$$

where g denotes the growth rate of market demand. If demand grows such that $d_{t+1} = D_{t+1}/n_t > R_{q_t}$, then entry deterrence requires expansion of productive capacity such that $R_{q_{t+1}} \geq d_{t+1}$. Assume that the expansion is associated with an unchanging shape of the production and cost curve, which simply means that the average cost curve shifts to the right.

Second, entry deterrence requires a sunk investment if the firm faces opportunities to reduce average cost from period t to period $t+1$ (Chapter 15). For the sake of simplicity, take the following specification of the innovation technology [equation (15.4)]:

$$(23.8) \quad \Delta \text{avc}_{t+1} = h_t \cdot \text{sc}_t^{\mu_t} = R(\text{sc}_t),$$

where $h_t > 0$ and $0 < \mu_t < 1$.² Recall that the amortization procedure in period $t+1$ follows the zero-profit markup rule [function $M(\text{sc}_t, q_t)$ in equations (13.1) and (13.2); Proposition 14.2].

¹ This result follows of course from the theory of perfectly contestable markets (Chapter 6).

² This means that, for the sake of simplicity, the reduction of average cost is not modeled via an increase of ϵ_t to ϵ_{t+1} (Chapter 15).

The firm's innovation decision is given by equation (15.6), which means that the future price is minimized so as to reach a position of exit avoidance [Proposition (15.2)]. Substituting equation (13.2), a discount rate r and (23.8) into (15.4) gives the solution (1sc_t in Chapter 15)

$$(23.9) \quad sc_t = \{(1+r)/[\mu_t \cdot h_t \cdot q_{t+1}]\}^{1/(\mu_t-1)},$$

where r is equal for and known to all firms.¹ Suppose that firms are able to introduce a process innovation to the entire productive capacity immediately at the beginning of period $t+1$. The firm's price and quantity decision in period $t+1$ follows from the equations (23.3) to (23.6) together with the markup rule (13.2). That is, price and average total cost are determined by

$$(23.10) \quad p_t = atc_{t+1} = avc_t - \Delta avc_{t+1} + sc_t/q_{t+1}.^2$$

23.3 INCREASING AND DECREASING DEMAND IN A CONTESTABLE MARKET

This section presents a simulation experiment that illustrates the unsustainability of contestable markets in a dynamic environment (Baumol *et al.* 1982a). The dynamics reveals itself in a changing magnitude of market demand. In a contestable market all firms are identical and sunk cost is zero. That is, all firms (are aware of the fact that they) face the same production technology (23.1) which is constant over time. The simulation is run with $\alpha = 0.7$ and $\beta = 0.7$ for $q < L_q$, $w_L = 1$, $w_M = 1$, $L_q = 10$ and $R_q = 14$ (suppressing time indices). Two simulation experiments are run so as to illustrate the role of entry and exit movements if market demand grows respectively falls from period to period.

A ten-period simulation with *increasing* market demand ($D_0 = 10$ and $g_t = 0.2$ for $t = 1, \dots, 10$) shows that a contestable market is unsustainable (Table 23.1).

Not surprisingly, increasing demand is associated with entry so as to restore equilibrium. For example, from period $t = 4$ to $t = 5$ market demand grows from 20 to 24. So, the two incumbent firms in period $t = 4$ are able to satisfy market demand in period $t = 5$ without inducing entry, since $L_q = q_4 = 10$ and $q_5 = 12 < R_q = 14$. However, the situation is different in period $t = 6$. From $t = 5$ to $t = 6$ market demand grows from 24 to 29, so that the two incumbent suppliers cannot meet market demand without provoking entry: $2 \cdot R_q = 28 < D_6 = 29$. Both incumbent firms decide to offer $q_6 = R_q = 14$ [equation (23.5)]. Excess demand 1 implies a profit opportunity to a potential entrant. In response to the incumbent firms' failure to satisfy market demand, a potential entrant decides to enter. Equilibrium is restored by an entrant which satisfies excess demand.

Five remarks are in order. First, temporary coordination is (partly) achieved with the help of entry movements. A sufficiently large pool of potential entrants facilitates the attainment of a temporary Walrasian equilibrium (Chapter 20). Second, the scale and date of entry is endogenized, since the potential entrants' decision on entry depends on excess demand. The contestability theory of course cannot predict which potential entrants gain

¹ For example, r is the prevailing interest rate (Chapter 12).

² Of course, if equation (23.9) gives $atc_{t+1} > p_t$, the entry-detering investment is $sc_t = 0$.

by the entry opportunities (Chapter 6). This means that the question of the order of entry (Lambkin 1988) is ignored. Third, market performance is associated with static efficiency, because prices are set at the level of (minimum) average cost.

Table 23.1
Increasing Demand and Entry in a Contestable Market

Period	Market demand	Market supply	Number of incumbent firms
1	12	12	1
2	14	14	1
3	17	17	2
4	20	20	2
5	24	24	2
6	29	29	3
7	35	35	3
8	42	42	3
9	50	50	4
10	60	60	5

Fourth, the experiment shows the difficulty that is associated with temporary (un)sustainability: for example, $D_6/n_{i6} = 29/3 < L_q = 10$. Here this is solved by assuming that a last entrant is allowed to satisfy demand at a price in excess of the minimum average cost level (Table 23.2, where $mavc_t = 1.04$ for $t = 1, \dots, 10$).

Table 23.2
Increasing Demand and Prices in a Contestable Market

Period	Price incumbent firms	Price last entrant
1	1.04	-
2	1.04	-
3	1.04	1.46
4	1.04	1.20
5	1.04	1.04
6	1.04	2.00
7	1.04	1.15
8	1.04	1.10
9	1.04	1.35
10	1.04	1.64

The last entrant acts as a contestable monopolist. That is, satisfying residual demand implies producing at an increasing returns level. The entrant satisfies excess demand by offering the associated average cost price, because a higher price is likely to induce further entry.

Fifth, the indeterminacy of a contestable market structure in which firms face a flat-bottomed cost function remains under cover through the "secondary sales maximization" assumption (23.5). Without this assumption a multitude of equilibrium numbers of firms can be calculated. For example, in period $t = 9$ five as well as four incumbent firms are able to deter entry, which implies that the market offers place to two rather than one newcomer. Moreover, feasible residual production assumes a favorable rationing scheme of demand for the niche entrant. That is, the consumers not yet served by lower-priced incumbent firms must have reservation prices above the average cost of niche production.

A ten-period simulation with *decreasing* market demand ($D_0 = 120$ and $g_t = -0.2$ for $t = 1, \dots, 10$) also indicates the unsustainability of contestable markets in a dynamic setting (Table 23.3).

Table 23.3
Decreasing Demand and Exit in a Contestable Market

Period	Market demand	Market supply	Number of incumbent firms
1	120	120	9
2	96	96	7
3	77	77	6
4	62	62	5
5	50	50	4
6	40	40	3
7	32	32	3
8	26	26	2
9	21	21	2
10	17	17	2

Decreasing market demand is associated with exit of incumbent firms if $D_{t+1}/n_i < Lq$. For example, from period $t = 7$ to $t = 8$ market demand falls from 32 to 26. Therefore, in $t = 8$ only two incumbent firms can satisfy market demand by producing at a constant returns level: $D_8/n_{i7} = 26/3 < Lq = 10$.

This experiment induces remarks that resemble the ones associated with the case with increasing demand: (i) temporary coordination is (partly) achieved through exit movements; (ii) scale and date of exit (and so market structure) are endogenized; (iii) contestability yields static efficiency; and (iv) temporary sustainability cannot always be obtained. Remark (ii) deserves further attention, since it remains to be explained which incumbent firms are the victim of unsustainability. None of the incumbent firms can be

indicated as the candidate that has to exit on any *a priori* argument, since all firms are alike. The simulation experiment goes around this inevitable indeterminacy of the contestability theory by adopting randomized exit. That is, the incumbent firms that have to leave the market are stochastically determined. This of course indicates a basic flaw of the theory of perfectly contestable markets.¹

23.4 A CONTESTABLE MARKET WITH COSTLESS CAPACITY EXPANSION AND REDUCTION

The first two simulation experiments abstract from the dynamic aspects of competition. In a perfectly contestable market without the opportunity to expand (or reduce) productive capacity, firms face a fixed capacity constraint (embodied in a constant Rq). This section focuses on an obvious example of dynamic competition: expansion (or reduction) of productive capacity. If incumbent firms can expand (or reduce) productive capacity, they have disposal of a dynamic strategy of exit avoidance and/or entry deterrence (Chapter 7). Two simulations are run: one in which a contestable monopolist has *perfect foresight* and the other in which the firms can face uncertainty of the future demand development.

Three assumptions are worth noting in advance. First, for the sake of simplicity, sunk cost is assumed to be zero. Second, it supposed that productive capacity can be adjusted after a one-period lag. That is, firms can costlessly expand (or reduce) productive capacity from one period to the next so as to keep pace with the growth of (or fall in) market demand. Third, firms try to expand (or reduce) productive capacity such that demand can be precisely met. Strictly speaking, the opportunity to expand productive capacity costlessly is associated with the incentive to introduce a once-and-for-all expansion that yields an infinitely large productive capacity. The reason is straightforward: with (a costlessly acquired) infinite capacity the firm is always able to satisfy individual (and even market) demand without inducing entry (and so unsustainability). Therefore, reasonable modeling of the expansion of productive capacity requires the introduction of an investment (buildup) cost. For that matter, the combination of a first-stage capacity (Cournot) and second-stage price (Bertrand) game is problematic if expansion is costly in a setting with potential entry (Chapter 21).² For the sake of simplicity, here this is ignored.³

The first ten-period simulation experiment assumes that the contestable monopolist has *perfect foresight* as regards to the growth rate of market demand ($D_1 = 12$ and $g_t = 0.2$ for $t = 2, \dots, 10$). Table 23.4 depicts the predictable and trivial results of the simulation run.

From one period to the next the monopolist defends its position by costlessly expanding its productive capacity so as to keep pace with the foreseen growth of market demand: $^Rq_t = D_t$ for $t = 1, \dots, 10$. This permits the contestable monopolist to keep producing at a constant returns level. Hence, entry deterrence by setting the minimum

¹ The fifth remark in Section 23.4 points to the further indeterminacy of the contestability theory which follows from the observation that the flat-bottomed average cost curve can introduce multiple equilibrium market structures (Holler 1985; see Chapter 6).

In particular, excess capacity introduces opportunities to dump profitably (Chapter 11).

³ It is therefore not surprising to note that the literature shows either first-stage Cournot and second-stage Bertrand models without potential entry (Kreps and Scheinkman 1983, Osborne and Pitchik 1986 and Davidson and Deneckere 1986) or first and second-stage Cournot games with potential entry (Spence 1979, Dixit 1980 and Calem 1988).

average cost price remains feasible [equations (23.5) and (23.6)]. For example, from period $t = 2$ to period $t = 3$ entry deterrence is obtained by keeping track with the growth of market demand from $D_2 = 14$ ($= R_{q_2}$) to $D_3 = 17$ ($< R_{q_2}$) by expanding productive capacity from $R_{q_2} = 14$ to $R_{q_3} = 17$ and setting the associated minimum average cost price $p_2 = p_3 = \text{mavc}_2 = \text{mavc}_3 = 1.04$. This result illustrates, first, the Chicago-thesis in a dynamic setting and, second, the intertemporal sustainability of a contestable monopolist both in an extreme case (costless expansion of productive capacity and perfect foresight), since a monopolistic market configuration is sustained over time without any violation of static efficiency.

Table 23.4

Increasing Demand and Sustainability of a Contestable Monopolist

Period	Market demand	Market supply	Number of incumbent firms	Price
1	12	12	1	1.04
2	14	14	1	1.04
3	17	17	1	1.04
4	20	20	1	1.04
5	24	24	1	1.04
6	29	29	1	1.04
7	35	35	1	1.04
8	42	42	1	1.04
9	50	50	1	1.04
10	60	60	1	1.04

The result is of course completely different if (some) firms face *uncertainty* rather than perfect foresight. Take the following illustrative scenario which brings in the two remaining features of competitive dynamics: expectational errors and heterogeneity. Assume that market demand is subject to stochastic growth. The growth rate of market demand in a period is determined by a draw from a normal distribution with mean 0.3 and standard deviation 0.7. At the beginning of the period t the growth rate of market demand in period $t+1$ is drawn. Firms can be uncertain of the precise value of g_t . Heterogeneity of firms is introduced by assuming that half of the suppliers are well-informed of the development of market demand, whereas the other half of the firms face uncertainty. The uncertain firms form expectations about the growth rate of market

demand in accordance with Keynes' convention.¹ That is, they employ static expectations by assuming that the next period's growth rate will be equal to the current growth rate: $E(g_{t+1}) = g_t$, where E is the expectational operator.

The rationing rule of market demand is assumed to be simple. At the beginning of a period market demand is equally divided among the incumbent firms of the previous period. All firms are fully aware of this allocation mechanism. At the beginning of period t firms anticipate, on the basis of their certain knowledge or uncertain expectations, the growth of demand in the next period by costlessly expanding or reducing their productive capacity in order to be able to meet individual demand in period $t+1$ by producing at minimum average cost. This is necessary so as to deter entry. The simulation experiment is run with 20 firms. The suppliers 1, 2, 5, 8, 9, 11, 14, 16, 18 and 19 are equipped with uncertain expectations. Moreover, $L_{q1} = 3$, $R_{q1} = 4$ and $D_1 = 20$.

The results of a seven-period simulation run are depicted in Tables 23.5 and 23.6.²

Table 23.5
Increasing Demand and Unsustainability of a Contestable Market

Period	Market demand	Individual demand	Market supply	Number of incumbent firms
1	20	3	20	7
2	32	4	32	11
3	32	2	32	13
4	67	8	58	14
5	118	8	80	10
6	233	23	230	10
7	369	36	360	10

After five periods all firms that face uncertainty are expelled from the market by the firms that have perfect foresight. This competitive struggle reveals itself of course in entry and exit movements. For example, from period $t = 2$ to $t = 3$ market demand remains constant. In period $t = 2$ the number of incumbent firms has, however, increased from 7 to 11, because the incorrect anticipatory expansion of productive capacity by uncertain incumbent firms induced entry. So, at the beginning of period $t = 3$ stationary market demand is allocated to an increased number of firms. Where many have to share, nobody will get much: individual demand decreases from 4 to 2. This implies that D_3/n_{i3}

¹ Of course, many other types of expectations procedures can be introduced. The simple case of static expectations is only introduced for illustrative purposes. The key point is that *expectational errors* and *heterogeneity* of firms are captured.

² The report of the simulation is restricted to a seven-period run, because after period $t = 5$ neither entry nor exit occurs.

$= 32/13 < L_{q_2} = 3$. The firms that have perfect foresight correctly foresee this development and anticipate in period $t = 2$ by reducing productive capacity such that individual demand in period $t = 3$ can be satisfied at the minimum average cost price without incurring losses. The uncertain suppliers expected, however, that market demand would grow from period $t = 2$ to period $t = 3$. They therefore failed to reduce productive capacity, so that individual demand in period $t = 3$ forces them to produce at an increasing returns level. The associated average cost price [equations (23.3) and (23.4)] indicates profit opportunities to potential entrants. Lower-priced entrants outperform the higher-priced incumbent firms.

Table 23.6
Exited Firms

Period	Exit firm number
3	1, 2, 5, 8, 9 and 11
5	14, 16, 18 and 19

Five remarks can be added to those already presented in Section 23.3. First, the example makes clear that potential entrants are also assumed to anticipate future (demand) developments (equally well as incumbents). Recall that it can be plausible to think that there can be entry from closely related industries or nearby regions in which firms build up (or destroy) productive capacity for reasons particular for their own market (Chapter 11). Potential entrants too are assumed to face either perfect foresight or static expectations procedures. This means that entrants can show a different scale of entry. A basic flaw of the simulation experiments described in this chapter is that the potential entrants' markets remain unidentified (Chapters 11 and 21).

Second, the experiment ignores the problems that are associated with the exit of firms.¹ That is, it is not specified what happens to expelled firms. Two interpretations can be put forward to illustrate this point. On the one hand, one scenario predicts that the expelled firms are able to enter into another market. This implies that the productive capacity that is associated with the outperformed firms is still used. On the other hand, exit can go hand in hand with bankruptcy. A natural extension of the model is to introduce the financial aspects of competition. This brings financial constraints, debts, creditors, capital markets, *etcetera*, into the picture.²

Third, it is (tacitly) assumed that firms produce and supply an integer number of

¹ This is a common attitude in the literature on industrial organization that deals with exit (for example, Nelson and Winter 1982, Jovanovic 1982 and Fudenberg and Tirole 1986).

² Lippman and Rumelt (1982) and Ghemawat and Nalebuff (1985) offer a model of exit, whereas Jovanovic (1982) and Jovanovic and Lach (1989) present selection frameworks in which both entry and exit take part in the explanation of (the evolution of) market structure. General equilibrium theory shows that the modeling of bankruptcy and credit is not easy (Grandmont 1977).

commodities. So, the rationing rule can cause three distinct situations: market clearing, excess demand and excess supply. The events in period $t = 2$ can illustrate the nonmarket clearing result. On the one hand, if firms choose to round up their equilibrium scale of production to 3, market supply (33) will exceed market demand (32), which leaves one firm with a loss. On the other hand, if suppliers decide to round down their production to 2, market supply (22) falls short of market demand (32). Here the latter rounding off rule is adopted so as to stipulate the coordinating role of entry movements. This means that the number of entrants can exceed the number of expelled firms.

Fourth, a (tacit) assumption of the experiment is that firms fail to anticipate the future exit of rivals. If firms with perfect foresight are able to do so, they can themselves outperform higher-priced rivals by expanding their productive capacity accordingly. The abstraction from this type of anticipatory expansion of productive capacity, together with the assumption of a limited number of potential entrants, causes nonmarket clearing in period $t = 5$ (and, to a lesser extent, in period $t = 4$). In period $t = 5$ market supply falls short of market demand, because higher-priced firms exit from the market but are replaced by lower-priced entrants with insufficient productive capacity (or, alternatively, an insufficient number of lower-priced entrants). This means that from period 5 to 7 the market is no longer contestable.¹ Nonmarket clearing (and noncontestability) can be avoided by modifying the assumptions. To name five possible modifications: (i) incumbent firms can anticipate future exits; (ii) the pool of potential entrants is sufficiently large; (iii) retired firms are allowed to re-enter; (iv) higher-priced firms can stay in the market if entry does not occur; and (v) (some) incumbent firms are permitted to produce at a decreasing returns level. In the simulation experiment under consideration none of these assumptions is adopted in order to stipulate the adaptive capability of the remaining incumbent firms in the market to make corrections in subsequent periods.

Fifth, after six periods all firms with perfect foresight are incumbent, whereas the firms facing uncertainty are expelled from the market. The exit of the latter are the result of expectational errors. This introduces an imperfect fit between market demand and supply, due to the same reasons as mentioned under the third and fourth remarks. Moreover, the imperfect fit illustrates the integer problem (Baumol *et al.* 1982a; see also Chapter 13). For example, in period $t = 7$ residual demand 9 cannot be satisfied with efficient scale production.

23.5 A BARRIER MARKET WITH COSTLY PROCESS INNOVATION

This section examines the competitive dynamics in a perfect barrier market with costly process innovation. That is, nonsimultaneous decision making, heterogeneity of firms and expectational errors all find a place in the simulation experiment. Nonsimultaneous decision making follows from the distinction between incumbent firms and potential entrants. In order to illustrate the first-mover advantage a period is divided up into three stages. In the first stage incumbent firms decide on price and investment. At the end of the first stage potential entrants observe the incumbent firms' prices. In the second stage potential entrants also decide on price and investment. Moreover, they decide whether or not to enter during the second stage. Entry occurs whenever a potential entrant is able to outperform an incumbent firm. It is in the third stage where transactions take place.

¹ Note that the incumbent firms keep behaving as if the market is contestable. Chapter 21 described a different model.

The heterogeneity of agents is introduced by assuming that diverging mechanisms of expectations formation prevail. Firms' expectations are differentiated by the parameter h_t in equation (23.8). That is to say, firms face technical uncertainty. Assume that half the firms form expectations according to Keynes' convention (*i.e.*, static expectations): $E(h_{t+1}) = h_t$. Let the other half employ a simple autoregressive model: $E(h_{t+1}) = 1/2 \cdot h_t + 1/2 \cdot h_{t-1}$. That is, intertemporal uncertainty is introduced. The importance of distinguishing procedures of static and autoregressive expectations is that this introduces *diverging* expectations.¹ Evolutionary selection and competitive struggles are a sequel to heterogeneity (Chapter 4). Diverging expectations are an important source of heterogeneity.

Table 23.7
Process Innovation, Prices and Entry

Period	Lowest feasible price	Actual price surviving incumbents	Actual price retired incumbents	Actual price entrants
1	100.00	100.00	-	-
2	93.46	93.46	-	-
3	88.96	88.96	-	-
4	76.93	82.14	83.49	82.14
5	64.61	73.18	74.63	73.18
6	52.40	60.17	-	-
7	42.64	49.58	-	-
8	31.79	38.91	-	-

An eight-period simulation experiment is run ($t = 1, \dots, 8$).² The actual value of h_{t+1} is determined by a draw at the beginning of period t from a normal distribution with mean 3 and standard deviation 0.3. At the same time, suppliers calculate the expected values of h_{t+1} . In order to generate the initial situation the firms are kept well-informed of h_1 and h_2 . On the basis of the expected h_t firms decide on the scale of R&D outlays to be spent in period t [equation (23.9)]. The values of Δavc_t , mu_t and p_t follow from equations (23.8), (23.2) and (23.10) respectively. Retired incumbent firms are allowed to re-enter one period following their exit. The market consists of five incumbent firms that face seven potential entrants. Firms 3, 4, 7, 10 and 12 deploy Keynes' convention, whilst the

¹ Of course, again other models of expectations formation can be introduced as long as they permit expectations to diverge.

² The report of the experiment is restricted to an eight-period run, because only periods $t = 4$ and $t = 5$ show entry and exit movements.

other firms use an autoregressive expectations procedure. In order to focus attention on the competitive dynamics, it is assumed that the quantity sold per incumbent firm is constant over time: $q_t = 5$ for $t = 1, \dots, 8$. This supply level allows production at a constant returns scale. Moreover, $D_1 = 100$, $r_t = 0.03$, $\mu_t = 0.5$ and $g_t = 0$ for $t = 1, \dots, 8$.

Table 23.8
Innovative Opportunities

t	h_t	$^*sc_{t-1}$	$^*\Delta avc_t$	*mavc_t	$^*\mu_t$	*p_t
4	3.39	21.73	15.80	72.46	4.48	76.93
5	3.07	10.77	10.07	62.39	2.22	64.61
6	3.28	16.89	13.47	48.92	3.48	52.40
7	2.88	7.31	7.79	41.13	1.51	42.64
8	3.22	14.86	12.40	28.73	3.06	31.79

Table 23.9
Decisions and Expectations of Firm 3 and 8

Parameters and variables	$t = 4$		$t = 5$	
	Firm 3	Firm 8	Firm 3	Firm 8
$E(h_t)$	2.31	2.61	3.39	2.85
sc_{t-1}	2.41	4.27	21.73	6.88
$E(\Delta avc_t)$	3.59	5.40	15.80	7.48
$E(mavc_t)$	84.67	82.86	67.19	73.78
μ_t	0.50	0.88	4.48	1.42
$E(p_t)$	85.17	83.74	71.67	75.20

The results of the experiment are depicted in Table 23.7.

After period $t = 2$ suppliers no longer receive advance information of the actual development of innovative opportunities. Therefore, they are unable to keep pace with the time path of the lowest feasible prices. In period $t = 4$ and $t = 5$ higher-priced incumbent firms are outperformed by lower-priced entrants.

The competitive dynamics are illustrated with the help of the life history of incumbent firm 3 and potential entrant 8. The development of the values of the relevant parameters and variables as would be generated by a well-informed supplier with perfect foresight can be employed as a point of reference (Table 23.8). Firm 3 and 8 base their R&D

decision on static and autoregressive expectations respectively, leading to the following results in the turbulent periods $t = 4$ and $t = 5$ (Table 23.9). (After period $t = 5$ neither entry nor exit occurs: the constellation of incumbent firms is stationary.)

In period $t = 3$ both supplier 3 and 8 underestimated the actual innovative opportunities. However, supplier 3's expectation diverges from reality to a greater extent than that of supplier 8. Based upon their estimates, both firms decided on the R&D budget in period $t = 3$. The results in period $t = 4$ are presented in Table 23.10.

In period $t = 4$ firm 8 achieves a larger reduction of average variable cost than firm 3. Firm 3 is unable to compensate for the innovative inferiority by tolerating a lower markup, so that firm 8's price is lower than firm 3's. Hence, in the second stage of period $t = 4$ firms employing the autoregressive expectations procedure are able to outperform those firms which use Keynes' convention. Therefore, the former stay in the market and the latter are expelled (among these is firm 3) by lower-priced entrants (among these is firm 8). At the end of the first stage potential entrants receive information of the incumbents' prices and quantities. Potential entrants employing autoregressive expectations recognize profit opportunities.¹ In the second stage the prices of incumbent firms 3 and 4 are undercut by entrants 8 and 9.² In the third stage transactions take place and market demand is satisfied.

In period 5 the opposite results occur. Firm 3 underestimated and firm 8 overestimated the actual innovative opportunities. Therefore, in period $t = 4$ supplier 2 invested insufficiently and supplier 8 overstated its investment. Based upon the under respectively overshooting of the R&D outlays the following results occur in period $t = 5$ (Table 23.11).

Table 23.10
Results in Period $t = 4$

Parameters and variables	Firm 3	Firm 8
Δavc_4	5.62	7.00
$mavc_4$	83.00	81.62
μ_4	0.50	0.88
p_4	83.49	82.14

¹ Strictly speaking, the entrants fail to seize these profits, since they enter with their average total cost price. This strategy may seek to deter further entry. This result can of course easily be modified, so that an underbidding entrant is allowed to capture a (temporary) positive profit. In effect, the simulation experiment does not model Baumol *et al.*'s (1982a) "hit-and-run" proposition, because entrants apparently intend to stay. The (intention of the) entrants' behavior therefore resembles a "hit-and-stay" strategy (Schwartz 1986).

² With respect to entry the "first come first served" rule applies, since the demand side shows no preference for either equally-priced entrant. Strategic interactions among potential entrants (which can, for example, result in excessive or zero entry) are ignored. This is the common attitude in the literature (Schwartz and Reynolds 1983 and Nti 1989).

Table 23.11
Results in Period $t = 5$

Parameters and variables	Firm 3	Firm 8
Δavc_5	14.30	8.04
$mavc_5$	68.70	73.21
μu_5	4.48	1.42
p_5	73.18	74.63

Firm 3 with static expectations is able to re-enter the market by outperforming a higher-priced incumbent firm 8 which employs the autoregressive expectations mechanism. Incumbent firm 8 is surprised by a *larger* reduction of average cost than expected. The *opposite* holds for potential entrant 3. Nevertheless, potential entrant 3 is able to undercut incumbent firm 8's price, because the latter's lower markup does not outweigh the former's larger average variable cost reduction. At the end of the first stage potential entrants receive information of the incumbents' prices and quantities. The potential entrants with static expectations notice profit opportunities. Therefore, in the second stage firm 3 *re-enters* the market by outbidding an incumbent which employs autoregressive expectations. In the second stage the incumbent firms 1, 2, 5, 8 and 9 are pushed aside by the lower-priced entrants 3, 4 (both re-entering), 6, 7 and 12. In the third stage transactions take place and market demand is satisfied.

In the periods thereafter the suppliers with static expectations are able to survive the competitive process at the expense of the suppliers with autoregressive expectations. Thus, the competitors which use only the most recent information in their forecasting procedure show a superior performance. This result depends of course on the particulars of the model and parameter values employed. The key point is, however, that the simulation experiment permits the analysis of competitive dynamics by focusing on the implications of nonsimultaneous decision making, heterogeneity of firms and expectational errors in cohesion.

The simulation experiment with costly innovation in a barrier market adds one important result to those presented in Sections 23.3 and 23.4: static efficiency (average total cost pricing) and dynamic improvements (cost reductions) are *both* induced. Of course, since there are winners as well as losers, Pareto criteria cannot be used as a matter of course to judge the quality of intertemporal competition. Chapter 24 discusses this in more detail. Furthermore, the simulation experiment captures strategic and nonsimultaneous interactions between rivals along with expectational errors and heterogeneity of firms. Therefore, it reflects a first step toward the analysis of competitive dynamics.

23.6 COMPETITIVE DYNAMICS AND MARKET PERFORMANCE

The argument in this chapter serves two purposes: first, the analysis of competitive dynamics with the help of simulation techniques is illustrated; and, second, market

performance of contestable and barrier markets is revealed from a dynamic perspective. Broadly speaking, three conclusions are worth summarizing:

- (1) the simulation experiments provide a point of contact where the three essential features of competitive dynamics can be analyzed in cohesion: nonsimultaneity of decision making, heterogeneity of agents, and expectational errors;
- (2) the static efficient features and intertemporal unsustainability of contestable markets are illustrated; and
- (3) the innovation-inducing incentives in a barrier market in combination with technical uncertainty introduce evolutionary selection: static efficiency (zero profits) and dynamic improvements (cost reductions) go hand in hand.

Moreover, the simulation experiments illustrate the kind of problems one encounters while modeling and analyzing competitive dynamics. Five examples can be named: (i) which firms are allowed to enter? (ii) how is the problem of temporary unsustainability solved? (iii) where do entrants come from? (iv) what happens to expelled firms? and (v) in which way are the information flows modeled? Of course, the simulation experiments bypass many additional difficulties and complexities. An example of an important topic ignored in the simulation experiments is the evaluation of winners and losers in the evolutionary selection process from a welfare theoretical perspective. However, the key argument is that simulation techniques permit the analysis of the many topics associated with economic (competitive) dynamics.

WELFARE IMPLICATIONS AND POLICY INTERVENTIONS

24.1 INTRODUCTION: Welfare Evaluations in Economics

The welfare evaluation of market performance is not an easy matter. The sections on market performance in Part III illustrate this point by indicating some of the problems that interfere with reaching clearcut judgments. Here the tradeoff between static and dynamic efficiency of market behavior, which is so strongly suggested by the literature on industrial organization, is telling. To make things worse, the implications of dynamic competition are not favorable by definition. For example, (the speed or rate of) innovation can be both insufficient and excessive (Chapter 8). Broadly speaking, each type of dynamic competition (that is, each source of barriers to entry) has to be evaluated on its own terms (Chapter 7).

In general, the literature on industrial organization seeks a way out of the difficulties by making use of a *partial welfare* function in which the consumers' and producers' surplus are the arguments. However, this approach is characterized by two crucial flaws. First, it is well-known that partial welfare requirements need not to be favorable in a *general* framework. Second, the economic surplus apparatus takes insufficient notice of the *distributional* effects that go hand in hand with dynamic competition. That is, winners and losers are simply assumed to cancel each other out. *Welfare economics* is a branch of economic theory that focuses on these types of difficulties. This chapter briefly reviews the tools that the history of thought in welfare economics has brought about¹ with the aim to decide on the yardstick which can be employed to evaluate dynamic barrier market competition.

The plan of this chapter is as follows. Section 24.2 focuses on the informational requirements that are a condition for well-founded policy interventions. Section 24.3 examines the welfare norms that have been put forward to introduce the measurement, general and distributional dimensions of welfare. Section 24.4 offers preliminary observations on the welfare features of the (perfect nonstrategic) barrier market. Since the welfare evaluation of dynamic barrier market competition is only preliminary, Section 24.5 points to future research.

¹ Of course, this chapter cannot be more than a caricature of an encompassing review of the literature on welfare economics. The key point is, however, that the argument illustrates the prominent achievements and flaws of welfare economics. For example, Ng (1979), Mishan (1981a) and Cornes and Sandler (1986) offer introductions to welfare economics.

24.2 INTERVENTION AND INFORMATION

The evaluation apparatus of welfare economics - Pareto optimality, (potential) Pareto improvement, economic surplus, compensation principles, second-best rules and superfairness criteria (Section 24.3) - suggests that market performance, if suboptimal, can be improved by policy intervention: that is, if the welfare criteria indicate that the market fails to pass the tests, policy intervention can be used to bring market performance closer to the welfare-theoretic ideal. Welfare criteria can be used as a guide for policy. The key point is of course to indicate under what circumstances policy intervention can be justified; and if so, what are the appropriate measures to take. This section briefly reviews the arguments that can be used to examine these topics (Van Witteloostuijn 1988d). Here the focus is on intervention in markets, since industrial organization is central to this thesis.¹

Pre-war decades show no doubt an increasing popularity of government intervention in markets. The number of institutional (entry and exit) barriers in modern economies is very large. It is illustrative to note that Quirk and McDougall (1981) observe that there are 47 governmental institutions in the United States which devote all their energy to the regulation of the economy. Regulation is manifest in many forms, varying from quality requirements for tires of tractors to prohibition of cartels. The economics of regulation seeks to study the effects of institutional barriers on market performance. In this respect, four topics deserve further attention: the arguments for and against regulation; the theoretical effects of regulation on specific markets or the economy as a whole; the results and consequences of regulation in practice; and the difficulties that are associated with the implementation of regulatory measures.

The economics of regulation suggests that regulation is only justified if specific conditions apply (see also Part III). For example, Breyer (1984, p. 234) points out that the nonregulated market is the norm. That is, the onus of proof lies with the proponents of regulation. Here (at least) four economic arguments can be put forward that point to useful government intervention (Friedman 1985): (i) market power; (ii) externalities; (iii) imperfect information; and (iv) public goods.

The argument of *market power* is particularly directed to (natural) monopolies. It is argued that the absence of sufficient competition permits monopolies to set (too) high prices and to produce (too) small quantities.² *Externalities* occur if the (production) activities of one agent induce unintended costs and/or benefits to one or more other agents. Here environmental pollution is an illuminating example. *Imperfect information* can give rise to government intervention. For example, consumers may not be able to observe the qualities and dangers of products without the legislative provision of additional information. Thus it is plausible that information on security and health can be useful. *Public goods* lack the features that induce profitable private production. This is particularly the case with commodities that allow consumers to benefit from their services without any compensating effort or expenditure. Defence is the classic example.

Besides economic reasons, there can of course be noneconomic arguments in favor of regulation. In this respect, the most important motivation for intervention in modern

¹ The regulation issue is central in industrial organization. This is, for example, clear from the four chapters which the *Handbook of Industrial Organization* (Schmalensee and Willig 1989) devotes to this topic. This section can only offer a sidelong glance at this complicated matter.

² The question of the regulation of (natural) monopolies goes beyond this simple "market power" statement. However, space limitations dictate brevity. Braeutigam (1989) offers a recent survey of the literature on "optimal policies for natural monopolies".

economies is the promotion of a fair *distribution of income*. Moreover, reference can be made to the protection of *vital industries* in the economy (for example, the provision of energy) or the guarantee of *employment*.

The *benefits* of regulation are of course in accordance with the intended goals: the reduction of market power, a decrease or settlement of externalities, the provision of good information, the production of public commodities, the promotion of a fair income distribution, the protection of vital industries and/or the guarantee of employment. However, side effects which are associated with *costs* generally thwart these intended benefits of regulation. Here three arguments can illustrate this point.

First, the cost of regulation are in accordance with the predictions of industrial organization (Part III; in particular, Chapter 7). Regulation and putting up institutional entry barriers are merely two sides of the same coin. The restriction of entry opportunities is even an important policy instrument: for example, in the form of a license system (Quirk and McDougall 1981). Much of the regulation measures induce cost increases. Environmental prescriptions, quality requirements, security conditions, *etcetera*, increase the cost of producing and selling commodities. So, it is not surprising that regulation measures go hand in hand with the restriction of the entry opportunities of potential entrants (Field 1983). Behind the protecting institutional entry barriers firms are less inclined to undertake costly R&D activities and to introduce efficiency improvements (Reynolds 1985). Moreover, if price and quantity controls are not used as regulatory instruments, price and quantity effects are likely to occur. That is, reduced competition as a result of raised barriers to entry creates the condition for monopoly behavior, which leads to increased prices and decreased production (Scherer 1980).¹

Second, it is not clear on *a priori* grounds that the public authorities only and always follow the public interest: that is, it can be that *private interests* influence regulatory decision making. The theory of pressure groups (Stigler 1971 and Posner 1984) indicates that regulatory measures of public authorities can be the playground of private interest groups. The institutional barriers raised by regulation activities are only to the benefit of those who are being regulated (Phillips 1975).² A similar warning derives from the theory of public choice (Mueller 1979). Public interest is not the only guide to public policy. As Olsen (1965) points out, not only individual agents but also groups seek to satisfy private interests. For example, it can be that politicians strive for vote maximization (Downs 1957), whereas bureaucracies aim to maximize their budgets (Niskanen 1971).³

Third, the *implementation* of regulatory measures runs into serious informational difficulties (Kahn 1970): To what extent is there too much market power? How large are the externalities involved? What price level can be considered reasonable? In which way is the quality of products to be measured and controlled? And so on.⁴ The informational requirements are further increased through the second-best problem (Section 24.3).

¹ So, as Field (1983) points out, regulation can induce inflation.

² The theory of rent seeking offers a further case in point by stipulating the fact that private agents try to explore public regulatory activities. A key argument is that the rent-seeking activities of private agents can reduce the community's welfare (for example, Tullock 1967 and 1980, Krueger 1974, Posner 1975, Bhagwati and Srinivasan 1980 and Appelbaum and Katz 1986a, 1986b and 1987).

³ Here Phillips (1975) points out that regulation induces even more regulation. That is, the danger is that regulatory bureaucracies start to operate as self-amplifying systems. Moreover, keeping alive and working large regulation institutions is very costly.

⁴ This argument points out that policy intervention often asks for (a scale and type of) information that is not available (at the date of intervention).

Intervention in one market of the economy induces spillovers to other markets. The second-best theory suggests that notice can be taken of spillover effects by making use of second-best rules. Ng's (1979) theory of third-best rules can illustrate the restrictions imposed by informational limitations.¹

Ng (1979) argues that the policy rules which can be adopted to enhance welfare depend critically on the scale of the relevant information that is available. Ng distinguishes three informational regimes:

(1) *Informational Poverty* where the available information is insufficient to provide a reasonable probabilistic judgment regarding: (a) the direction and extent of divergence of the second-best optimum from that resulting from the application of the first-best rule in the presence of second-best distortion; (b) the shape and skewness of the relation curve apart from its general concavity;

(2) *Information Scarcity* where the available information is sufficient for such a judgment but is not perfect;

(3) *Informational Abundance*, i.e. perfect information (Ng 1979, p. 231; Ng's italics).

The relation curve relates the value of the objective (community's welfare) function to the direction and degree of divergence from the first-best rule for the variable (commodity or market) under consideration. The relation curve permits account to be taken of administrative (information gathering and processing) cost.

On the basis of his classification of informational conditions Ng recommends (i) applying first or second-best rules (depending on the observed restrictions) if informational abundance rules the world (provided that administrative cost is sufficiently low), (ii) adopting first-best rules if informational poverty is present (provided that administrative cost is negligible) and (iii) behaving in accordance with third-best rules in a situation of informational scarcity. That is,

[t]he third-best policies depend on the amount of information as well as on the administration costs involved. With informational scarcity, third-best policies converge with first-best ones. With perfect information and negligible costs of administration, third-best policies converge with second-best ones. Hence both first-best and second-best policies are special, polar cases of the general third-best policies (Ng 1979, p. 233).²

¹ The theoretical arguments which indicate the costs of (too much) regulation are supported by the results of empirical studies. This type of study generally focuses on the effects of (de)regulation on particular industries: for example, electricity services (Primeaux Jr. 1975 and Weiss 1975), road transport (Moore 1975, Field 1983, Reynolds 1985 and Simon 1986), shipping (Larner 1975), telecommunication (Waverman 1975), financial institutions (Mann 1975 and Phillips 1975) local taxi companies (Reynolds 1985), car insurance (Smallwood 1975) and air transport (Eads 1975, Field 1983 and Albers 1986). The purport of the literature on international integration is in accordance with the predictions of the theoretical and empirical economics of regulation. The demolition of international institutional barriers generally facilitates welfare. This result can be derived on the basis of theoretical arguments (Viner 1950, Meade 1955, Gehrels 1956, Lipsey 1957 and 1970, Scitovsky 1958, Balassa 1961, Leibenstein 1966, Bhagwati 1971, Chacholiades 1978, Corden 1981 and Vanheukelen 1985) as well as empirical arguments (Mayes 1971 and 1978, Magee 1972, Verdoorn and Schwartz 1972, Aitken 1973, Balassa 1974, Resnick and Truman 1974, Deardorff and Stern 1979, Baldwin *et al.* 1980, Baldwin 1984 and Pelkmans 1985).

² The emphasis on informational restrictions motivates neo-Austrian economics to take the argument one step further by stipulating that public authorities have to be discouraged from intervention in markets, since the type of information upon which policy measures have to be based cannot be present (or acquired) at a central level

24.3 WELFARE CRITERIA

24.3.1 Pre-Paretian and Paretian Welfare Criteria¹

In order to be able to judge whether market intervention is appropriate, a welfare criterion cannot be missed. Moreover, if market intervention is justified, a welfare criterion is needed to guide (advice regarding) regulatory policies. Section 24.3.1 reviews the standard welfare norm: Pareto optimality. Section 24.3.2 goes on to evaluate additional criteria which have been developed so as to assimilate the shortcomings of Pareto optimality.

Broadly speaking, pre-Paretian welfare economics is based upon two axioms: (i) the law of diminishing marginal utility²; and (ii) human beings have an equal capacity for enjoyment. Axiom (i) implies the assumption that an extra unit of (physical or nonphysical) commodities offers less utility than the preceding unit, whereas axiom (ii) means that all economic agents can be associated with identical utility functions. A crucial auxiliary assumption is that interpersonal utility comparisons can be carried out. This has the important implication that the *cardinal* approach to utility can be used to internalize externalities and distributional effects. Pigou (1912) is a representative example of the classical tradition in welfare economics.

However, in modern welfare economics the argument on interpersonal comparison goes both ways. First, (a degree of) interpersonal comparison of individual utilities can be retained (Roth 1979). Second, the *ordinal* approach to utility (Hicks and Allen 1934) takes the stand that utility is *not* interpersonally comparable. This means that distributional effects are distortions to the theory. As Mishan (1981a) summarizes:

Differences between the conclusions reached by the older tradition associated with Marshall, Pigou, Robertson, and to some extent Lerner (1946), as distinguished proponents of the 'Cambridge' or 'neoclassical' School - which built on a foundation of diminishing marginal utility and a belief in the basic similarity of human beings - and the more modern approach associated with the names of Vilfredo Pareto (1909), Hicks (1939), Samuelson (1950), and many others, turn out to be significant only with respect to distributional propositions With respect to the purely allocative analysis, there is general agreement, although, with the passage of time, there is greater appreciation today of its limitations (Mishan 1981a, p. 7).

Pareto (1909) is of course the founding father of modern welfare economics. Pareto's central proposition is what nowadays is called the Pareto principle: a change is desirable if it makes some agent(s) better off without making (any) other(s) worse off. This rule defines a welfare improvement. Moreover, Pareto derives three first-order conditions for a welfare optimum from marginal utility theory:

- (1) *Exchange condition.* The marginal rate of substitution between any two goods is equal for all agents.
- (2) *Production condition.* The marginal rate of substitution between any two factors of

as a result of the very nature of market competition (for example, Hayek 1949 and Kirzner 1985).

¹ Schumpeter (1954) examines the history of the economic analysis of welfare. He traces the track of welfare economics thinking back to the scholastics. The modern variants of welfare economics have clearly Benthamian roots: "So far as this goes, modern welfare economists merely revive the Benthamite tradition" (Schumpeter 1954, p. 1069).

² This indicates that Dupuit (1844) and Gossen (1854) are forerunners of modern welfare economics. Here Dupuit's notion of consumers' surplus is particularly worth noting.

production is equal for all products.

- (3) *Top level condition.* The relation between products and preferences is characterized by the equality for any two goods between the marginal psychological rate of substitution and the marginal technical rate of substitution.

An auxiliary assumption is that the production process is efficient. These conditions can be illustrated with the help of such well-established tools as the Edgeworth box, indifference curves, utility functions and production and utility possibility frontiers. The first-order conditions only hold in the case of (i) divisibility of commodities, (ii) continuity of the functions and (iii) absence of externalities in consumption and production if internalization cannot be obtained.¹ Moreover, two second-order conditions have to be satisfied so as to indicate a welfare optimum: both the marginal rates of substitution and the marginal rates of transformation for goods and factors are a diminishing function of quantity. Particularly the second feature can be in conflict with reality, because it rules out increasing returns in production technologies. An economy that complies with Pareto's welfare conditions and features is called a first-best world.

The theory of first, second, and third-best rules is based upon the assumption that a welfare criterion can be determined by means of maximizing a Bergsonian community's welfare function (Bergson 1938)², subject to one or more constraints. A community's welfare function can take account of distributional effects. Such a welfare function looks like

$$(24.1) \quad wc = WC(wm_1, \dots, wm_{nc}),$$

where wc denotes the community's welfare and wm_i ($i = 1, \dots, nc$) represents the welfare of the i 'th agent (generally in terms of utility) in a community of nc members. Function (24.1) is a Paretian one if condition (24.2) holds:

$$(24.2) \quad \partial WC / \partial wm_i > 0 \text{ for } i = 1, \dots, nc.$$

The maximization of the objective function (24.1) subject to condition (24.2) yields the conditions that give social welfare optimality. This procedure can be applied to a first-best world and is therefore called the first-best rule. The first-best rule implies the equalization of all marginal rates of substitution and transformation: that is, prices are equal to marginal costs. A first-best world is characterized by efficient production.

The theory of Pareto is associated with (at least) three basic shortcomings. First, it is unclear in which way the measurement of welfare (changes) has to take place. Second, the Pareto principle implies that the theory cannot be applied in cases where some agent(s) is

¹ For example, as Coase (1960) shows, under particular conditions undisturbed competition can guarantee the achievement of optimality. In particular, Coase's theory can be applied to the cases where distortions from the first-best world occur as a result of externalities. Coase's theorem implies that, in the absence of transaction cost but with full information, a Pareto optimal outcome will be reached if the parties involved can agree on compensating arrangements (irrespective of whether the affecting or affected party has the liability to pay compensation).

² The literature raises doubts as regards the logical possibility of deriving a community's welfare function on the basis of the agents' preferences in the form of ordinal orderings. Arrow (1951) denies this possibility in his well-known impossibility theorem. The same, under different conditions, is done by Kemp and Ng (1976) and Parks (1976). The theory of social choice (for example, Sen 1982 and Elster and Hylland 1986) focuses on ways out of this dilemma.

(are) made better off, whereas some other(s) is (are) made worse off. That is, Pareto's theory has to be adjusted so as to take account of distributional effects and externalities. Third, the features which characterize the first-best world are never met in actual economies. Four theories can be examined to illustrate the responds of welfare economics to this critique: (i) economic surplus conditions; (ii) compensation principles; (iii) second-best rules; and (iv) superfairness criteria.

24.3.2 Economic Surplus Conditions, Compensation Principles, Second-Best Rules and Superfairness Criteria

This section successively examines economic surplus conditions, compensation principles, second-best rules and superfairness criteria that have been proposed in order to deal with measurement problems, general multimarket economies, distributional effects and/or externalities.

(i) *Economic surplus conditions.* The theory of consumers' and producers' surplus emphasizes that a money proxy is a proper indication of welfare. This theory is based upon the measurement of the effects on the money surplus of consumers and producers following price and/or quantity changes. Broadly speaking, the consumers' surplus follows from the area under the demand curve and above the market price, whereas the producers' surplus is defined as profit.¹ For the sake of brevity, here the focus is on the consumers' surplus.

Marshall (1890) defines the consumers' surplus as

the excess of the price which he [the consumer] would be willing to pay rather than go without the thing, over that which he actually does pay (Marshall 1890, p. 124).

Hicks (1943) points out that the consumers' surplus can be measured in (at least) four ways. Ng (1979) neatly summarizes the Hicksian measures:

- (1) Compensation variation (CV) of a change in prices (or other variables, if we are not confined to changes in prices) is the amount of compensation (usually in monetary terms) that can be taken from an individual while leaving him just as well off as *before* the change;
- (2) Compensation surplus (CS) is the amount of compensation that can be taken from the individual while leaving him just as well off as *before* the change if he were constrained to buy at the new price the quantity of the commodity he would buy in absence of compensation;
- (3) Equivalent variation (EV) is the amount of compensation that has to be given to the individual, in absence of change, to make him as well off as he would be with the change; and
- (4) Equivalent surplus (ES) is the amount of compensation that has to be given to the individual, in the absence of the change, to make him as well off as he would be with the change if he were constrained to buy at the old price the quantity of the commodity he would buy in absence of compensation (Ng 1979, p. 85; Ng's italics).

In addition to the four Hicksian definitions, Ng (1979) indicates two further measures:

- (1) The Laspeyre cost difference of a price change is the amount of compensation that could be taken from the consumer while leaving him just able (not necessarily willing) to buy the original bundle of goods he bought before the change.

¹ Strictly speaking, the term producers' surplus is misleading, since it deals with the welfare that accrues to the owners of the factors of production. For that reason, Mishan (1981a) refers to the rent surplus.

(2) The Paasche cost difference is the amount the consumer would have to be paid to have just enough money to buy the new bundle of goods at the original prices (Ng 1979, p. 89).

The choice of the appropriate measure of the consumers' surplus depends on the type and quality of the data and the nature of the problem at hand. The theory of economic surplus is, however, associated with (at least) three important flaws. First, money (income) is used as the indication of welfare. Second, distributional effects are ignored. Third, the concept of surplus is partial: that is, it fails to take account of the spillover effects in multimarket economies.

(ii) *Compensation principles.* Kaldor (1939) is the founding father of the theory of compensation principles. Compensation principles can be applied to cases where some agent(s) is (are) made better off, whereas some other(s) is (are) made worse off, without, however, making use of interpersonal utility comparisons. Hicks (1939) supported and extended Kaldor's principle. The Kaldor-Hicks compensation criterion implies that a welfare improvement is reached if (1) all those who suffer as a result of a change, can be fully compensated for their loss, while the rest of the community will still be better off than before (Kaldor's criterion) *or* (2) losers cannot profitably bribe the gainers to oppose the change (Hicks' criterion). A further point is that the Kaldor-Hicks principle is about *hypothetical* rather than actual compensation.

The hypothetical nature of the compensation has the important implication that the Kaldor-Hicks criterion can lead to contradictions (Scitovsky 1941). That is, it can be that Kaldor's criterion applies, whereas Hicks' principle is violated (or *vice versa*). To avoid such contradictions, Scitovsky (1941) introduces the reversal test. The reversal test indicates that a change only implies a welfare improvement if it passes *both* Kaldor's and Hicks' criterion. A crucial flaw of the Kaldor-Hicks principle and Scitovsky's reversal test is that both are about *potential* welfare improvements: that is, both fail to take account of distributional effects.

Little's (1949) compensation principle takes notice of changes of income distribution. Little's criterion is based upon two premises: (1) a change implies a welfare improvement if someone is made better off without anyone being made worse off (Pareto principle) *and* (2) *ceteris paribus*, a change implies a welfare improvement if the income distribution is made better.¹ The second condition introduces, of course, an additional difficulty: what makes one distribution better than another? In this respect, Mishan (1981a) notes the following:

For one thing, a direct utility or welfare comparison between each of the persons is not essential for the judgment that one distribution is more equal than another. It is possible to make such judgments by attention only to the *structure* of the distribution in each position and irrespective, therefore, of the welfare of each of the particular persons.

For another, the distributional judgments can be, and in practice are, made by reference to the alternative structure of *money* incomes. It may be thought that the use of distributions of money incomes as a proxy of distributions of 'real' incomes, or real welfare, is itself an arbitrary assumption. But it may be plausibly argued that society as a whole, at least in affluent countries, is prone to think of distribution of income primarily in terms of money (Mishan 1981a, pp. 317-318; Mishan's italics).

¹ It can be argued that Little's compensation principle imposes a distribution condition on Scitovsky's reversal test.

So, Little's second condition is rendered concrete by narrowing down distributional effects to changes in the distribution of money incomes. However, a normative (political) judgment is still needed to compose a rank ordering of "good" and "bad" income distributions (measured in terms of money).¹

(iii) *Second-best rules.* Starting with the seminal contribution of Lipsey and Lancaster (1956) a large literature on second-best rules has been produced. This literature focuses on the problems that arise as a result of distortions from the first-best world: That is, what can be said about welfare evaluations if the conditions of Pareto optimality are violated in a multimarket economy? The current pillar of the literature on second-best welfare economics is Ng (1979).² The key question is concerned with which way the first-best rule has to be changed if distortions from the first-best world are unavoidable.

Ng (1979) indicates four important sources of distortions from the first-best world: (1) differential degrees of monopoly power; (2) uncorrected externalities; (3) taxation; and (4) government interventions. These distortions impose additional constraints on the maximization of the community's welfare function (24.1). Consider an objective function F ,

$$(24.3) \quad F(x_1, \dots, x_n),$$

which is subject to constraint G ,

$$(24.4) \quad G(x_1, \dots, x_n) = 0.$$

Maximizing F subject to G gives the necessary first-order conditions

$$(24.5) \quad F_i/F_n = G_i/G_n,$$

where $F_i = \partial F/\partial x_i$, $G_i = \partial G/\partial x_i$ and $i = 1, \dots, n$. Function F can be interpreted as a community's welfare function. Distortions from the first-best world give rise to additional constraints in the form such as

$$(24.6) \quad F_i/F_n = \alpha \cdot (G_i/G_n),$$

where $\alpha \neq 1$. Restriction (24.6) introduces the impossibility to equate all marginal rates of substitution and transformation. Hence, the first-best rule cannot be (fully) applied; that is, Pareto optimality or improvement cannot be guaranteed.

The second-best theory argues that the second-best rule is to maximize the community's welfare function subject to first and second-best constraints. That is to say, if one market is

¹ The (familiar) key argument is that welfare evaluations cannot do without normative judgments. As Ng (1979) points out, even the Pareto principle implies "a value judgment, but it is a very weak one in the sense that most people will accept it and in the sense that many other value judgments subsume the Pareto-principle and yet also contain something more" (Ng 1979, p. 30). These arguments raise the question of the positive or normative (and scientific) character of welfare economics. For example, to what extent is welfare objectively measurable? Here these topics are ignored (see, for example, the disagreement between Mishan and Hennipman 1984).

² A random sample from the population of articles on this topic is Sandler (1978), Guesnerie (1980), Mishan (1981a and 1981b), Walsh (1982) and Keenan and Keenan (1984). As pointed out in Chapter 9, it can be argued that Clark's (1940) concept of workable competition foreshadowed the theory of second-best (Reid 1987). Ng (1979) elaborates on the theory by introducing third-best rules (Section 24.2).

associated with inescapable imperfect competition, it can be that second-best optimality requires a degree of imperfect competition in other markets as well. As Lipsey and Lancaster (1956) put it:

If there is introduced into a general equilibrium system a constraint which prevents the attainment of one of the Paretian conditions, the other Paretian conditions although still attainable, are, in general, no longer desirable (Lipsey and Lancaster 1956, p. 11).

This has far-reaching implications for the recommendations of welfare economics. For example, Sandler (1978) points out that

[s]econd best theory has held some important implications for welfare economics and, more specifically, piecemeal policy in which policy removes some, but not all, distortions in an economic system ... In all likelihood, economic reality is a second best rather than a first best world owing to immovable distortions that are founded on technological, legal and political rigidities (Sandler 1978, p. 331).

A world which is characterized by distortions from the first-best world is called a second-best world. Second-best conditions are generally very complicated, since they depend on, first, the degrees of complementary and substitutability between commodities in the constrained markets and those in the unconstrained markets and, second, the effects of increased (or decreased) production of a commodity on the marginal cost of producing other goods.¹

(iv) *Superfairness criteria*. The superfairness theory seeks to reach tools that can be used to take explicit notice of equity considerations in welfare economics. Important contributions to the superfairness theory are Foley (1967), Schmeidler (1969), Schmeidler and Vind (1972), Varian (1974, 1975, 1976a, 1976b and 1980), Pazner and Schmeidler (1974 and 1978), Feldman and Kirman (1974), Crawford (1977, 1979a, 1979b and 1980) and Baumol (1978). Thompson and Varian (1985) and Baumol (1986) offer reviews of the literature. Broadly speaking, a

distribution is called (nonstrictly) *superfair* if each class of participants prefers its own share to the share received by another group, that is, if no participant *envies* the other (Baumol 1986, p. 15; Baumol's italics).²

Superfairness theory can be illustrated with the help of the Edgeworth box apparatus. The superfairness criterion involves a number of connections with the Paretian norms (Pareto optimality and Pareto improvement). The key point is, however, that the concept of envy permits the introduction of fairness considerations. Baumol (1986) defines envy as follows:

A distribution of n commodities is said to involve envy by individual 2 of the share obtained by

¹ Some conditions concerning the relation between commodities in constrained and unconstrained markets can give rise to less complicated second-best rules. Worth noting are (i) upper bound constraints, (ii) separability of objective and constraint functions (which implies that all commodities are independent) and (iii) the observation that the functions are partially weakly separable (which means that the degree of complementarity or substitutability is equal between all pairs of commodities). These three conditions permit the derivation of the first-best rules for all but the constrained markets.

² This type of argument resembles Tinbergen's exchange theory (1957), which pleads for a fair income distribution.

individual 1 if 2 would rather have the bundle of commodities received by 1 under this distribution than the bundle the distribution assigns to 2 (Baumol 1986, p. 19).

This definition points out that superfairness theory does not involve interpersonal utility comparisons, since agents evaluate the commodity bundles of others in terms of their own utility measure.

The superfairness notion can be used as an incremental, partial or sequential criterion, which broadens the scope of applicability of theory. Incremental superfairness only compares the vector of increments in the bundles of goods of agents, partial superfairness focuses on a partition of the total bundle of commodities and sequential superfairness refers to iterative replications of the superfairness analysis so as to narrow down the set of fair distributions.¹ The key contribution of the superfairness criteria is, however, that they evaluate the fairness of (changes in) distributions from the viewpoint of the (groups of) agents: that is, is a (change in) distribution associated with envy?²

24.4 WELFARE FEATURES OF THE BARRIER MARKET: FIRST REMARKS

The brief account in this chapter of the current state of the art in welfare economics gives rise to two conclusions. First, the judgments of the performance of (perfect nonstrategic) barrier markets, as these are presented in Parts IV and V of this thesis, can be fleshed out by applying (some of the) criteria which are offered by welfare economics. This section only presents some preliminary thoughts on this topic. Second, a well-founded policy intervention is associated with the need for a type and scale of information that is difficult to obtain (in time). Here this point is ignored.

A welfare evaluation of the performance of the barrier market can be carried out at three levels by making use of partial, general and intertemporal welfare criteria. The preliminary conclusion of parts IV, V, VII and VIII was that the free entry conditions of the barrier market induce incentives *both* to follow careful pricing and to undertake cost-minimizing and/or consumers' surplus-maximizing innovation. *Ceteris paribus*, this means that from one period to another Pareto improvements are achieved.³ In a *partial* context the welfare apparatus traditionally in use in industrial organization can do the job: *consumers' and producers' surpluses*. Tirole (1988) presents the rationale by pointing out that for the case where the goods and industries considered

represent only a small share of consumer expenditure[, p]rice changes are ... likely to generate small income effects, and it may be appropriate to assume that demand is downward sloping and that the consumer surplus is a good approximation of welfare (Tirole 1988, pp. 11-12).

Perhaps even more important is the *intertemporal* dimension of the welfare evaluation of the performance in (a) barrier market(s). As soon as heterogeneity of economic agents is

¹ Superfairness theory is related to sophisticated game-theoretic tools, such as the core, the Shapley value and the nucleolus, which focus on the fairness in pricing (in relation to the allocation of cost). Here these concepts are ignored. Schotter and Schwödiauer (1980) offer a survey.

² Space limitations leave no room for a further examination and illustration of superfairness theory. Baumol (1986) offers applications of superfairness analysis to rationing of scarce commodities, compensation of victims of externalities, pricing for conservation, industry pricing and peak pricing and congestion.

³ The argument in this section proceeds along the lines of the commodity market, but can be applied to the labor market as well.

introduced, competitive dynamics goes hand in hand with winners and losers. The market selects those firms which adopt better innovation strategies. Inferior incumbent firms are outperformed by superior entrants (Chapter 23). This introduces a distributional dimension in the welfare evaluation of market performance. Compensation principles can be used to examine whether there is a potential Pareto improvement. That is, are the winners (consumers and superior firms) willing to offer the losers (inferior firms) a (hypothetical) compensation for their loss while still being better off themselves? Employing (*potential*) compensation arguments can offer a first answer to this question. This means that, in the first instance, Tirole's (1988) argument applies again. That is,

[t]he 'compensation principle' of Hicks (1940) and Kaldor (1939) holds that we need only be concerned about efficiency; if total surplus increases, the winners can compensate the losers and everyone is made better off. The classic drawback of this approach is that the distribution branch may not function, and compensation need not occur (Samuelson 1947). This caveat should be borne in mind in all our welfare conclusions (Tirole 1988, p. 12).

A first answer to the welfare question can be provided by tracing the surplus changes from period t to $t+1$ of the five groups of market participants involved for the models of Chapter 15 (process innovation) and 16 (product innovation) if it is assumed that the market is associated with a credible entry threat *ex ante* (in period t) and *ex post* (in period $t+1$):

- (i) From the exit-avoiding and entry-permitting strategies of incumbent firms respectively (potential) entrants it follows immediately that the consumers benefit from dynamic competition in the barrier market. Firms either set lower prices after the introduction of a cost reduction or offer a quality improvement (and associated zero-profit price) that increases the consumers' utility after undertaking a product innovation (positive surplus change).
- (ii) Superior incumbent firms are able to avoid exit after the introduction of the process or product innovation. Since exit avoidance in the barrier market requires, *ceteris paribus*, average total cost pricing, the zero profit level is sustained (zero surplus change).
- (iii) Inferior incumbent firms are expelled from the market, because in the barrier market their inefficient innovation induces entry. This means that these firms are clearly confronted with a loss. For example, the sunk cost that was incurred in order to innovate is not (fully) recovered (negative surplus change).
- (iv) Superior potential entrants can outperform inferior incumbent firms by entering profitably. The very reason for entry into the barrier market implies that superior entrants benefit from the competitive dynamics that is induced by innovation (positive surplus change).
- (v) The potential entrants which fail to enter¹, stay, *ceteris paribus*, where they are. So, the change in profit of these potential entrants, if there is any, can be ignored while evaluating the barrier market under consideration (zero surplus change).

By way of preliminary conclusion two observations are worth noting. First, the specifics of

¹ These potential entrants can be either inferior or superior. The barrier market assumptions imply that inferior potential entrants cannot enter profitably. If the pool of superior potential entrants is such that not all can enter (given the number and size of inferior incumbent firms), then, *ceteris paribus*, there are superior potential entrants which refrain from entry.

the barrier market combine two favorable outcomes: competition in a barrier market induces (near) average total cost pricing *and* the introduction of dynamic economies (innovations). Second, the first observation ignores distributional effects: it can be argued that the surplus increase of the two groups of winning participants may or may not exceed the surplus decrease of the one group of losing participants.

24.5 FUTURE RESEARCH

This chapter only summarizes the state of the art regarding welfare-economic yardsticks and presents a preliminary argument on the welfare features of a dynamic barrier market. Whether the competitive dynamics in the (perfect nonstrategic) barrier market is, *ceteris paribus*, associated with a potential Pareto improvement from a period t to a period $t+1$, needs careful consideration in future research. Future modeling should take into consideration cost of barriers, economic surpluses, potential compensation schemes, *etcetera*. The crucial question is whether there is a (hypothetical) compensation of the victims' losses that still leaves the winners better off. In a partial context the economic surplus and compensation notions can do the job, but in a general setting second- and third-best considerations cannot be ignored.

Even in this partial two-period framework complications have to be taken into account. In so far as R&D investment is concerned, it is important to take explicit notice of the externalities that can interfere with *potential* Pareto-improving market performance (Chapter 8). To be precise, the features of appropriability and competition can induce insufficient or excessive innovation. For example, it can be that the (perfect nonstrategic) barrier market is associated with excessive R&D expenditures, because the pervasive entry threat can induce duplication of R&D outlays. Moreover, it can be interesting to introduce other sources of dynamic competition in a barrier market model (for example, advertizing). However, the literature on entry barriers makes clear that the implications of dynamic competition for the consumers' surplus depend critically on the source of investment (Chapter 7). Moreover, in a general multi-barrier market model spillover effects introduce further complexities that have to be taken into account (Chapters 20 and 21). This implies that second-best and third-best problems most certainly enter the picture.

CHAPTER

25

MULTIMARKET COMPETITIVE DYNAMICS

This thesis is about three important economic concepts: rationality, competition and dynamics. The argument in Part II indicates that the rationality and maximizing postulate are chosen as the point of departure. The key point is that a review of the literature suggests that the (rationality and) maximizing principle is (are) able to assimilate many of the anomalies as these were (and are) brought up by opponents. In the absence of a credible alternative, abandoning the use of the maximizing argument (either in the form of marginalist optimization techniques or decision-analytic game theory) is likely to end up with throwing away the baby with the bath water. Therefore, the subsequent parts of the thesis frame agents' (decision making) behavior in terms of maximizing a goal function subject to (a) restriction(s).

This chapter summarizes the argument by briefly indicating a future research agenda. Of course, space limitations only permit the presentation of a priority list. In effect, at many points in the preceding chapters weaknesses and lacunas in the arguments in the current literature and this thesis are revealed. For example, the barrier market theory can be improved upon by exploring further scenarios and strategies and the models of competition can be imposed to a sophisticated welfare theoretic evaluation. This chapter focuses, however, on three major topics (without further reference to the literature) by way of concluding statement: (i) intermarket spillovers and the microfoundation of macroeconomics; (ii) multimarket competition and the microfoundation of competition; and (iii) competitive dynamics and the evolution of competition.

The microfoundation of macroeconomics is an evergreen in the repertoire of the economists' choir. Patinkin set into motion an impressive quest for a grand theory in economics which houses micro as well as macroeconomics. By using general equilibrium theory as the frame of reference, the interactions among many agents in many markets are analyzed with reference to macroeconomic outcomes. Walrasian macroeconomics is reached if prices (and wage rates) are fully flexible so that markets always clear, whereas Keynesian macroeconomics is associated with conditions which keep prices (and wage rates) from moving sufficiently fast implying that nonmarket clearing equilibria can be obtained. The microeconomics associated with Keynesian macroeconomics is based upon vertical intermarket rationing spillovers. The microfoundation of the price inertia can be reached by introducing market imperfection. The literature over the last decade stipulates the role of imperfect competition (with its characteristics of price setting, externalities, coordination failures, imperfect information, *etcetera*).

The study of (im)perfect competition is the central focus of industrial organization. Recent contributions to industrial organization can be used to enrich the theory of the microfoundation of macroeconomics. One specific opportunity to achieve synergy strikes the

eye: dynamic competition. Ever since, at least, Schumpeter's contribution to economics the crucial role of the dynamics of competition has been recognized. This is particularly clear from the literature on entry barriers and innovations. Industrial organization's theories on dynamic competition can improve upon the microfoundation of macroeconomics, since this mixture permits the study of essential elements of the working of the economy from a general perspective. For example, investment in innovation and human capital cannot be missed in a modern microfoundation of macroeconomics.

The synergy is not one-sided. The study of industrial organization's models of competition in a general equilibrium framework broadens the horizon. To be precise, a general framework permits the study of the implications of multimarket competition. The multimarket perspective is essential as incumbent firms in a market not only compete with companions in (mis)fortune, but also have to take into account forces from outside the market. Here the impressive literature on entry and entry deterrence is telling. The key point is that market performance can be different if the threat of potential entry is credible. The microfoundation of the (non)credibility of the entry threat can only take place satisfactorily in a multimarket framework.

The multimarket perspective opens the door to the study of intermarket (demand, cost and strategic) spillovers which imposes an extra dimension of the theory of competition. Joint (dis)economies of scale and/or scope, cross-parry, indirect competition (through straddlers), *etcetera*, complicate but enrich the theory of competition. In effect, the type of competition should be endogenous in multimarket theory. The incumbent firms' and potential entrants' strategies give shape to the economy's competitive features through the outcomes of the incumbents against incumbents, incumbents against entrants and entrants against entrants games.

The multimarket competitive framework can be brought into motion by shifting attention to competitive dynamics. Heterogeneity of firms and competitive dynamics are associated with selection. If the investment strategies of (incumbent and potential) suppliers diverge, a repeated (Cournot)-Bertrand game can give entry and exit movements. Both external (market discipline) and internal (organizational learning) selection can be elements of competitive dynamics. Lower-priced (or favored-product) firms survive to the detriment of lower-priced (or unfavored-product) rivals. The model of internal selection can be introduced into the simulation of competitive dynamics.

In line with Nelson and Winter's work the simulation of repeated (Cournot-Bertrand) games with external and internal selection can indicate the evolution of competition over time. As divergent strategy choices (of incumbent firms and potential entrants) can induce entry and exit movements, market structure and the type of competition can fluctuate from period to period. For example, if entry and exit bring about a market in which only lower-priced (or favored-product) firms are incumbent, entry barriers can introduce imperfect competition. If neither incumbent firms nor potential entrants decide to invest a sunk cost *ex ante*, the market is perfectly contestable *ex post*. The evolution of rivals' strategies exert influence on the evolution of price, cost, product, productive capacity, *etcetera*.

AFTERWORD

This afterword says goodbye to the patient and persistent readers who have gone through this thesis' lengthy argument. The import of the argument is economic with attention to both the positive and welfare-theoretic perspective. The barrier market theory is the hard core of the thesis. The barrier market, as workable competition, is defined as to its implications for market performance. An important element of the barrier market theory is to examine what conditions can give favorable market performance.

The almost uncountable number of cases and scenarios in this thesis [particularly Chapters 11, 19, 20 and 21 do a good (or bad) job in this respect] is likely to puzzle and confuse the reader (and writer!). Where is the wood between all those trees? However, in my opinion this is what industrial organization (and economics) is all about. Here Fisher's and Shapiro's recent discussion in the *Rand Journal of Economics* (1989, 20, 113-137) is telling: Fisher takes the stance that

the status of the theory of oligopoly¹ is that of exemplifying theory. We know that a lot of different things *can* happen. We do not have a full, coherent, formal theory of what *must* happen ... Such a theory is presently lacking. At present, oligopoly theory consists of a large number of stories, each one an anecdote describing what might happen in some particular situation. ... We have a large and increasing number of formal anecdotes in which the outcome appears to depend heavily on the context. The role of a true generalizing theory of oligopoly is to tell us the nature of that dependence (Fisher 1989, pp. 118-119; Fisher's italics).

The point is, however, that Fisher's pretension - a "generalizing" theory of oligopoly - is, as far as I can see, a utopia.

Fisher defines generalizing and exemplifying theory as follows:

Generalizing theory proceeds from wide assumptions to inevitable consequences. It speaks in terms of what *must* happen, given background circumstances. ... Exemplifying theory does not tell us what *must* happen. Rather it tells us what *can* happen. In a good exemplifying-theory paper, the model is stripped bare, with specializing assumptions made so that one can concentrate on the phenomena at issue (Fisher 1989, pp. 117-118; Fisher's italics).

The strategic (and expectational) dimension of oligopoly behavior implies, however, that only exemplifying theory has something to say on the real issues. Shapiro correctly observes that

[m]ost fundamentally, I think that Professor Fisher is misguided in seeking a single "generalizing" theory of oligopoly. First, a very general theory may be of quite limited usefulness if it fails to match real-world conditions. ... Second, a search for a single generalizing theory of oligopoly denies the very richness of business behavior that makes I.O.

¹ Fisher identifies industrial organization with oligopoly theory. This is too narrow a view as a sidelong glance at the *Handbook of Industrial Organization* (Schmalensee and Willig 1989) reveals.

[i]ndustrial organization] so interesting. We do have some general themes emerging from the theory of business strategy - the importance of timing and the role of commitment - but to seek a single theory of strategic behavior is both unwise and fruitless (Shapiro 1989b, p. 132).

Furthermore,

[i]ndustrial organization economists are both blessed and cursed: our field encompasses a wide range of business behavior that is a rich arena in which to apply economic principles, but the very richness of business strategy defies simple and general theories. ... A metatheme that emerges from this research is that the predictions of the model concerning the character of the equilibrium of the game tend to be quite sensitive to the exact specification of the firms' strategies and the timing of actions.¹ ... There is no reason to expect or strive for a single unified oligopoly theory that would deliver unique predictions to armchair theorists, independent of the particulars of how competition is played out in a given industry (Shapiro 1989b, pp. 125-126).

This case-by-case approach inevitably gives a large set of cases and scenarios. However, enlarging this set significantly contributes to our understanding of industrial organization:

The diversity of predictions in different game-theoretic models reflects our broadening understanding of business strategy. With our game-theoretic tools, we can carefully analyze a much wider range of competitive strategy than was previously possible. The theory tells us the conditions under which different outcomes occur (Shapiro 1989b, p. 126).

Shapiro goes on to argue that

[e]ven the term 'oligopoly theory' is in need of replacement. ... The explosion of game-theoretic work in I.O. is better described as the 'theory of business strategy' (Shapiro 1989b, p. 125).

In particular, the strategy dimension of (market) competition can be applied to the strategy content theory (strategy analysis and choice) of business administration. The application of the economic theory of industrial organization to strategic management, particularly the strategy content dimension, is at the top of my research agenda's priority

¹ Fisher and Shapiro's discussion centers upon the role and usefulness of game theory. Fisher observes critically that "[f]irst, game theory in normal form is an inconvenient language with which to prepare the way for the generalizing theory of oligopoly. Second, while game theory in its extensive form may very well be a highly useful language in this regard, most of the existing literature has concentrated on one-shot games, and these are not the interesting ones" (p. 121). Apart from the fact that this is not an inherent critique (but rather a reproach of current practice), Shapiro points out that "[m]ost of Fisher's attacks are directed at 'static, one-shot oligopoly theory.' This is a straw man. ... in my view Fisher's emphasis on the shortfalls of static oligopoly theory is quite out-of-date and entirely out-of-touch with the theoretical advances of the past fifteen years. The whole point of the theory of business strategy is that static models are inadequate, that commitments and timing matter a great deal, that imperfect information has a profound impact on firms' behavior, and that we can study an enormous range of business behavior, not just pricing and output competition, by using a single set of game-theoretic tools (though not in a single model)" (p. 132). In effect, Shapiro rightly indicates that "[u]sing extensive-form games to model strategic interactions has the virtue of forcing the analyst to think carefully and to be quite precise about the specific nature of competition. At this time, game theory provides the only coherent way of logically analyzing strategic behavior" (p. 125).

list (Van Witteloostuijn 1990f).¹ For example, the multimarket oligopoly theory with identified potential entrants offers a point of contact to explore the strategy space of firms from the perspective of business administration.

¹ In this respect, Shapiro presents the opinion that "the more applied field of corporate strategy has a long and rich history. Economists have much to learn from scholars who have studied corporate behavior and corporate strategy in detail. I hope the learning will be mutual, as economists undertake more case studies to test their theories and as corporate strategy experts integrate game-theoretic insights into their work" (p. 126, note 1). This is not to say that no mutual learning is going on at the moment. The recent special issue of the *Strategic Management Journal* (1988, 9, 1-142) on strategy content research, Teece's (1984) plea for exploring interfaces and existing (and well-known) case studies in the industrial organization tradition such as Schmalensee (1978b) and Ghemawat (1984) are illustrative in this respect.

A. Minimum Average Total Cost

Proof Proposition 13.1. It remains to be shown that the straightforward intuition that the introduction of sunk costs is associated with an average total cost function with a *unique minimum*, is correct. For the sake of simplicity of notation, indices are suppressed. The first-order derivative of equation (13.3) is

$$(A1) \quad \partial C / \partial q = \partial A / \partial q + \partial M / \partial q.$$

Condition (13.1a) shows that $\partial M / \partial q < 0$. Conditions (13.4a) and (13.4b) indicate that for $q \leq R_q$ $\partial A / \partial q \leq 0$. Hence, for $q \leq R_q$ $\partial C / \partial q < 0$. Given that in equilibrium $D(p)/n_i$ is equal to q , differentiation of equation (13.2) gives

$$(A2) \quad \partial M / \partial q = -sc/q^2 < 0.$$

Hence, $\partial M / \partial q$ is a monotonic increasing function of q . For $q > R_q$ $\partial A / \partial q$ is a monotonic increasing function of q as well [condition (13.4c)]. However, for $q > R_q$ $\partial A / \partial q > 0$ and $\partial M / \partial q < 0$. It is to be shown that there exists *only one* q^* where $\partial C / \partial q = 0$. Take a $q_1 > R_q$ with $\partial A / \partial q = \beta > 0$. For all $q > q_1$ $\partial A / \partial q > \beta$. Equation (13.6) shows that $\partial M / \partial q$ limits to zero for $q \rightarrow \infty$. Hence, a q_2 exists with $\partial M / \partial q = -\beta$. For all $q > q_2$ $\partial M / \partial q > -\beta$. So, for all $q > \max\{q_1, q_2\}$ $\partial C / \partial q$ exceeds zero. Hence, since $\partial C / \partial q$ is an *increasing* function of q with $\partial C / \partial q < 0$ for $q = R_q$, there exists a value of q (q^*) such that $q^* > R_q$ and $\partial C / \partial q = 0$. Since $\partial A / \partial q$ and $\partial M / \partial q$ are both increasing functions of q , $\partial C / \partial q$ is strictly increasing in q . Hence, q^* is unique. For q^* $\partial C / \partial q = 0$, so $C(q^*)$ gives the unique minimum cost and price level and q^* the associated unique level of output. *Q.E.D.*

B. Decreasing Returns R&D Technology

Proof Proposition 15.1. It is to be shown under what conditions equation (15.4) gives Δavc_1 as a concave and increasing function of sc_0 .

$$(B1) \quad \partial R / \partial sc_0 = -\partial A / \partial \epsilon_1 \cdot \partial E / \partial sc_0 > 0.$$

Condition (B1) holds, because $\partial A / \partial \epsilon_1 < 0$ [condition (14.1a)] and $\partial E / \partial sc_0 > 0$ [condition (15.2a)]. Hence, Δavc_1 is a (monotonic) *increasing* function of sc_0 . Moreover,

$$(B2) \quad \partial^2 R / \partial sc_0^2 = -\partial^2 A / \partial \epsilon_1^2 \cdot \partial E / \partial sc_0 - \partial A / \partial \epsilon_1 \cdot \partial^2 E / \partial sc_0^2.$$

In order to preserve concavity (decreasing returns) $\partial^2 R / \partial sc_0^2$ must be negative. The second term of condition (A2) ($-\partial A / \partial \epsilon_1 \cdot \partial^2 E / \partial sc_0^2$) is negative, because $\partial A / \partial \epsilon_1 < 0$

[condition (15.1a)] and $\partial^2 E / \partial sc_0^2 < 0$ [condition (15.2b)]. Besides, $\partial E / \partial sc_0 > 0$ [condition (15.2a)]. Hence, concavity of equation (15.4) requires that

$$(B3) \quad \partial^2 A / \partial \epsilon_1^2 > 0, \text{ or}$$

$$(B4) \quad \partial^2 A / \partial \epsilon_1^2 = 0 \text{ or}$$

$$(B5) \quad \begin{cases} \partial^2 A / \partial \epsilon_1^2 < 0, \text{ and} \\ \partial^2 A / \partial \epsilon_1^2 \cdot \partial E / \partial sc_0 < \partial A / \partial \epsilon_1 \cdot \partial^2 E / \partial sc_0^2. \end{cases}$$

The interpretation runs as follows. The influence of investment outlays on the reduction of average nonsunk cost works out through two steps. First, R&D outlays determine directly the change in the state of the technology (technology effect). Second, a shift of the state of the technology induces a change in average nonsunk cost (cost effect). If condition (B3) holds, both the technology as well as the cost effect are subject to decreasing returns. Hence, the decreasing returns property of the technology effect is enhanced by the one of the cost effect. Condition (B4) states that the cost effect shows constant returns. Then, the decreasing returns property of the technology effect influences the relation between investment outlays and reduction of average nonsunk cost without hindrance. When condition (A5) holds, the cost effect shows increasing returns. However, the cost effect is dominated by the technology effect, which, on balance, yields a decreasing returns property of the influence of investment outlays on the reduction of average nonsunk cost. The opposite occurs when neither condition (B3) nor (B4) nor (B5) holds. *Q.E.D.*

C. Existence Optimal R&D Outlays

Proof Proposition 15.3. The existence of a solution to the minimization problem (15.6) requires that the first-order derivative equals zero, whilst the second-order derivative is positive. Hence,

$$(C1) \quad \partial P / \partial sc_0 = -\partial R / \partial sc_0 + 1/q_1 = 0,$$

which is associated with

$$(C2) \quad \partial R / \partial sc_0 = 1/q_1.$$

Moreover,

$$(C3) \quad \partial^2 P / \partial sc_0^2 = -\partial^2 R / \partial sc_0^2 > 0.$$

Condition (C3) holds when either condition (B3), (B4) or (B5) is valid (Proposition 15.1). *Q.E.D.*

D. Uniqueness Optimal R&D Outlays

Proof Proposition 15.4. The existence of a unique solution requires that $\partial P / \partial sc_0 = -$

$\partial R/\partial sc_0 + 1/q_1$ only once intersects the zero-axis. First, it appears that the following conditions must hold:

$$(D1) \quad \lim_{sc_0 \rightarrow 0} \partial P/\partial sc_0 = -\partial R/\partial sc_0 + 1/q_1 < 0, \text{ and}$$

$$(D2) \quad \lim_{sc_0 \rightarrow \infty} \partial P/\partial sc_0 = -\partial R/\partial sc_0 + 1/q_1 > 0.$$

In order to preserve condition (D1) $\partial R/\partial sc_0$ must exceed $1/q_1$ if sc_0 approaches zero. Condition (D2) requires that

$$(D3) \quad \lim_{sc_0 \rightarrow \infty} \partial R/\partial sc_0 < 1/q_1.$$

It is very likely that condition (D3) holds, because the occurrence of negative average nonsunk cost is economically nonsensical. Hence,

$$(D4) \quad \lim_{sc_0 \rightarrow \infty} R(sc_0) = \max \Delta avc_1 \leq avc_1, \text{ where } \max \Delta avc_1 > 0.$$

So,

$$(D5) \quad \lim_{sc_0 \rightarrow \infty} \partial R/\partial sc_0 = 0 < 1/q_1.$$

Hence, provided that condition (D5) is plausible, the existence of a unique solution requires that $\partial R/\partial sc_0 > 1/q_1$ for $sc_0 = 0$. Condition (D5) holds if learning-by-doing is not excessive. If learning-by-doing emerges such that the optimal reduction of average cost is reached without any investment outlay (and therefore $sc_0 = 0$ gives $\Delta avc_1 = \max \Delta avc_1$), then the markup $\mu_1 = 1/q_1$ equals zero, so that condition (D5) is violated.

Second, it appears that

$$(D6) \quad \partial^2 P/\partial sc_0^2 = -\partial^2 R/\partial sc_0^2 > 0,$$

if either condition (A3), (A4) or (A5) holds (Proposition 13.1). Hence, the fact that $\partial P/\partial sc_0 = -\partial R/\partial sc_0 + 1/q_1$ reflects a continuously monotonic increasing function proves that under the conditions mentioned there exists a *unique* solution to the minimization problem (15.6). *Q.E.D.*

E. Existence Unique Optimal Human Capital Investment

Proof Proposition 18.3. The existence of a solution requires that

$$(E1) \quad \partial H/\partial sc_0 = \partial F/\partial sc_0 - \partial G/\partial sc_0 = 0, \text{ and}$$

$$(E2) \quad \partial^2 G/\partial sc_0^2 = \partial^2 F/\partial sc_0^2 - \partial^2 G/\partial sc_0^2 < 0.$$

The observation that maximum future attractiveness is associated with zero net returns, because pwg_1 can be increased for any $\Delta w_1 > (1+r) \cdot sc_0$ by lowering the incremental wage rate to $\Delta w_1 = (1+r) \cdot sc_0$ without violating the feasibility condition, implies that $\Delta w_1 = (1+r) \cdot sc_0$ and $\partial G/\partial sc_0 = (1+r)$. Provided that $\partial G/\partial sc_0 = (1+r)$, condition (E2) holds [equation (18.5)]. The existence of a *unique* solution requires that $\partial H/\partial c_0$ only once intersects the zero-axis. It appears that the following conditions must hold:

$$(E3) \quad \lim_{sc_0 \rightarrow 0} \partial H/\partial sc_0 = \partial F/\partial sc_0 - (1+r) > 0, \text{ and}$$

$$(E4) \quad \lim_{sc_0 \rightarrow \infty} \partial H/\partial sc_0 = \partial F/\partial sc_0 - (1+r) < 0.$$

In order to preserve condition (E3) $\partial F/\partial sc_0$ must exceed $(1+r)$ if sc_0 approaches zero [condition (i)]. Condition (E4) requires that

$$(E5) \quad \lim_{sc_0 \rightarrow \infty} \partial F/\partial sc_0 < (1+r).$$

It is very likely that there is an upper bound to human capital improvement so that condition (ii) holds:

$$(E6) \quad \lim_{sc_0 \rightarrow \infty} F(sc_0) = \max \Delta mp_1, \text{ where } \max \Delta mp_1 \geq 0,$$

which implies that

$$(E7) \quad \lim_{c_0 \rightarrow \infty} F(sc_0) = 0 < (1+r).$$

Condition (E7) holds if learning-by-doing is not excessive [condition (iii)]. If learning-by-doing occurs such that maximal attractiveness is reached without any investment outlay, then the incremental wage rate equals zero, so that condition (E7) is violated ($sc_0 = 0$ and $\Delta mp_1 = \max \Delta mp_1$). *Q.E.D.*

REFERENCES

- Abreu, D. - *Repeated Games with Discounting*, 1983, Ph.D. Thesis, Princeton, N.J.: Princeton University.
- Abreu, D. - "Extremal Equilibrium of Oligopolistic Supergames," 1986, *Journal of Economic Theory*, 39, 191-225.
- Adams, W.J. - "Market Structure and Corporate Power: The Horizontal Dominance Hypothesis Reconsidered," 1974, *Columbia Law Review*, 74, 1276-1297.
- Addison, J.T. and W.S. Siebert - *The Market for Labor: An Analytical Treatment*, 1979, Santa Monica, Cal.: Goodyear.
- Afriat, S.N. - "The System of Inequalities $a_{rs} > X_s - X_r$," 1963, *Proceedings of the Cambridge Philosophical Society*, 59, 125-133.
- Afriat, S.N. - "The Construction of Utility Functions from Expenditure Data," 1967, *International Economic Review*, 7, 67-77.
- Aghion, P. and P. Bolton - "Contracts as a Barrier to Entry," 1987, *American Economic Review*, 77, 388-401.
- Ahmad, S. - "On the Theory of Induced Invention," 1966, *Economic Journal*, 76, 344-357.
- Aitken, N.D. - "The Effect of the EEC and EFTA on European Trade: A Temporal Cross-Section Analysis," 1973, *American Economic Review*, 63, 881-892.
- Akerlof, G.A. - "A Theory of Social Custom, of which Unemployment may be One Consequence," 1980, *Quarterly Journal of Economics*, 94, 749-775.
- Akerlof, G.A. - "Loyalty Filters," 1983, *American Economic Review*, 73, 54-63.
- Akerlof, G.A. - *An Economic Theorist's Book of Tales*, 1984, Cambridge: Cambridge University Press.
- Akerlof, G.A. and W.T. Dickens - "The Economic Consequences of Cognitive Dissonance," 1982, *American Economic Review*, 72, 307-319.
- Akerlof, G.A. and J.L. Yellen - "A Near-Rational Model of the Business Cycle, with Wage and Price Inertia," 1985a, *Quarterly Journal of Economics*, 100, 823-838.
- Akerlof, G.A. and J.L. Yellen - "Can Small Deviations from Rationality Make Significant Differences to Economic Equilibria?," 1985b, *American Economic Review*, 75, 708-720.
- Akerlof, G.A. and J.L. Yellen - *Efficiency Wage Models of the Labor Market*, 1986, Cambridge: Cambridge University Press.
- Akerlof, G.A. and J.L. Yellen - "Rational Models of Irrational Behaviour," 1987, *American Economic Review*, 77, 137-142.
- Albers, I. - "De liberalisatie van de luchtvaart in Europa," 1986, *Economisch statistische berichten*, 71, 721-726.
- Alchian, A.A. - "Uncertainty, Evolution, and Economic Theory," 1950, *Journal of Political Economy*, 58, 211-221.
- Alchian, A.A. - "Information Costs, Pricing and Resource Unemployment," 1969, reprinted in Phelps E.S. (ed.) - *Microeconomic Foundation of Employment and Inflation Theory*, 1970, New York: Norton.

- Allais, M. - "Le Comportement de l'Homme Rationnel Devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine," 1953, *Econometrica*, 21, 503-546.
- Amihud, U. and H. Mendelson - "Price Smoothing and Inventory," 1983, *Review of Economic Studies*, 87-98.
- Andrews, P.W.S. - *Manufacturing Business*, 1949, London: McMillan.
- Andrews, P.W.S. - "Industrial Analysis in Economics," in: Wilson, T. and P.W.S. Andrews (eds.) - *Oxford Studies in the Price Mechanism*, 1959, Oxford: Clarendon Press.
- Andrews, P.W.S. and E. Brunner - *Studies in Pricing*, 1975, London: McMillan.
- Angelmar, R. - "Market Structure and Research Intensity in High-Technological-Opportunity Industries," 1985, *Journal of Industrial Economics*, 34, 69-79.
- Appelbaum, E. and E. Katz - "Transfer Seeking and Avoidance: On the Full Social Costs of Rent Seeking," 1986a, *Public Choice*, 48, 175-181.
- Appelbaum, E. and E. Katz - "Rent Seeking and Entry," 1986b, *Economics Letters*, 23, 207-212.
- Appelbaum, E. and E. Katz - "Seeking Rents by Setting Rents: The Political Economy of Rent Seeking," 1987, *Economic Journal*, 97, 685-699.
- Appelbaum, E. and C. Lim - "Contestable Markets under Uncertainty," 1985, *Rand Journal of Economics*, 16, 28-40.
- Arrow, K.J. - *Social Choice and Individual Values*, 1951, New York: John Wiley.
- Arrow, K.J. - "Toward a Theory of Price Adjustment," 1959, in: Abramovitz, A. (ed.) - *Allocation of Economic Resources*, Stanford, Cal.: Stanford University Press.
- Arrow, K.J. - "Economic Welfare and the Allocation of Resources for Invention," in: Nelson, R.R. (ed.) - *The Rate and Direction of Inventive Activity*, 1962, Princeton, N.J.: Princeton University Press.
- Arrow, K.J. - "The Future and the Present in Economic Life," 1978, *Economic Inquiry*, 16, 117-169.
- Arrow, K.J. and G. Debreu - "Existence of Equilibrium for a Competitive Economy," 1954, *Econometrica*, 22, 265-290.
- Arrow, K.J. and F.H. Hahn - *General Competitive Analysis*, 1971, Amsterdam: North-Holland.
- Arrow, K.J. and L. Hurwicz - "An Optimality Criterion for Decision-Making under Ignorance," in: Carter, C.F. and J.L. Ford (eds.) - *Uncertainty and Expectations in Economics*, 1972, Clifton, N.J.: A.M. Kelley.
- Arvan, L. - "Sunk Capacity Costs, Long-Run Fixed Costs, and Entry Deterrence under Complete and Incomplete Information," 1986, *Rand Journal of Economics*, 17, 105-121.
- Aumann, R.J. - "Survey of Repeated Games," in: Aumann et al. - *Essays in Game Theory*, 1981, Mannheim: Bibliographisches Institut.
- Azariadis, C. - "Implicit Contracts and Unemployment Equilibria," 1975, *Journal of Political Economy*, 83, 1183-1202.
- Azariades, A. and J.E. Stiglitz - "Implicit Contracts and Fixed-Price Equilibria," 1983, *Quarterly Journal of Economics*, 98, 1-22.
- Baily, M.N. - "Wages and Employment under Uncertain Demand," 1974, *Review of Economic Studies*, 41, 37-50.
- Bain, J.S. - "A Note on Pricing in Monopoly and Oligopoly," 1949, *American Economic Review*, 448-464.
- Bain, J.S. - "Workable Competition in Oligopoly," 1950, *American Economic Review*,

- 41, 35-47.
- Bain, J.S. - "Relation of Profit Rate to Industry Concentration: American Manufacturing, 1936-1940," 1951, *Quarterly Journal of Economics*, 65, 293-324.
- Bain, J.S. - *Barriers to New Competition*, 1956, Cambridge, Mass.: Harvard University Press.
- Balassa, B. - *The Theory of Economic Integration*, 1961, Homewood, Ill: Irwin.
- Balassa, B. - "Trade Creation and Trade Diversion in the European Market," 1974, *Economic Journal*, 57, 1-17.
- Baldwin, R.E. - "Trade Policies in Developed Countries," in: Jones, R.J. and P.B. Kenen (eds.) - *Handbook of International Economics*, 1984, Amsterdam: North-Holland.
- Baldwin, R.E., J. Mutti and J.D. Richardson - "Welfare Effects on the United States of a Significant Multilateral Tariff Reduction," 1980, *Journal of International Economics*, 10, 405-423.
- Ball, L, N.G. Mankiw and D. Romer - "The Keynesian Economics and the Output-Inflation Trade-off," 1988, *Brookings Papers on Economic Activity*, 19, 1-65.
- Barro, R.J. - "Rational Expectations and the Role of Monetary Policy," 1976, *Journal of Monetary Economics*, 2, 1-32.
- Barro, R.J. - "A Capital Market in an Equilibrium Business Cycle Model," 1980, *Econometrica*, 48, 1391-1417.
- Barro, R.J. - "Comment on 'Do Equilibrium Real Business Cycles Theories Explain Postwar U.S. Business Cycles?'," in: Fisher, S. (ed.) - *NBER Macroeconomics Annual 1986*, Cambridge, Mass.: MIT Press.
- Barro, R.J. and H.I. Grossman - "A General Disequilibrium Model of Income and Employment," 1971, *American Economic Review*, 61, 82-93.
- Barro, R.J. and H.I. Grossman - "Suppressed Inflation and the Supply Multiplier," 1974, *Review of Economic Studies*, 41, 87-104.
- Barro, R.J. and H.I. Grossman - *Money, Employment and Inflation*, 1976, Cambridge: Cambridge University Press.
- Bartel, A.P. and Borjas, G.J. - "Specific Training and its Effects on the Human Capital Investment Profile," 1978, *Southern Economic Journal*, 44, 333-341.
- Barzel, Y. - "Optimal Timing of Innovations," 1968, *Review of Economics and Statistics*, 50, 348-355.
- Baumol, W.J. - *Business Behavior, Value and Growth*, 1959, New York: McMillan.
- Baumol, W.J. - "On the Theory of the Expansion of the Firm," 1962, *American Economic Review*, 52, 1078-1087.
- Baumol, W.J. - *Economic Dynamics: An Introduction*, 1970, 3rd edition, New York: McMillan.
- Baumol, W.J. - "Equity vs. Allocative Efficiency: Toward a Theory of Distributive Justice," 1978, *Atlantic Economic Journal*, 6, 8-16.
- Baumol, W.J. - "Contestable Markets: An Uprising in the Theory of Industry Structure," 1982, *American Economic Review*, 72, 1-15.
- Baumol, W.J. - *Superfairness: Applications and Theory*, 1986, Cambridge, Mass.: MIT Press.
- Baumol, W.J., J.C. Panzar and R.D. Willig - *Contestable Markets and the Theory of Industry Structure*, 1982a, New York: Harcourt, Brace, Jovanovich.
- Baumol, W.J., J.C. Panzar and R.D. Willig - "On the Theory of Perfectly Contestable Markets," 1982b, printed in Stiglitz, J.B. and G.F. Mathewson (eds.) - *New*

- Developments in the Analysis of Market Structure*, Cambridge, Mass.: MIT Press.
- Baumol, W.J., J.C. Panzar and R.D. Willig - "Contestable Markets: An Uprising in the Theory of Industry Structure: Reply," 1983, *American Economic Review*, 73, 491-496.
- Baumol, W.J. and R.E. Quandt - "Rules of Thumb and Optimally Imperfect Decisions," 1964, *American Economic Review*, 54, 23-46.
- Baumol, W.J. and R.D. Willig - "Fixed Costs, Sunk Costs, Entry Barriers, and Sustainability of Monopoly," 1981, *Quarterly Journal of Economics*, 101, 405-432.
- Baumol, W.J. and R.D. Willig - "Contestability: Developments since the Book," 1986, *Oxford Economic Papers*, 38, 9-36.
- Beath, J., Y. Katsoulacos and D. Ulph - "The Game-Theoretic Analysis of Innovation: A Survey," 1989, *Bulletin of Economic Research*, 41, 163-184.
- Becker, G.S. - "Investment in Human Capital," 1962, *Journal of Political Economy*, 70, S9-S49.
- Becker, G.S. - *Human Capital: A Theoretical and Empirical Analysis*, 1964, 2nd edition, 1975, New York: National Bureau of Economic Research.
- Becker, G.S. - "A Theory of the Allocation of Time," 1965, *Economic Journal*, 75, 493-517.
- Becker G.S. - "A Theory of Marriage," in: Schultz, T.W. (ed.) - *Economics of the Family*, 1975, Chicago: University of Chicago Press.
- Becker, G.S. - "Altruism, Egoism, and Genetic Fitness: Economics and Sociobiology," 1976, *Journal of Economic Literature*, 14, 817-826.
- Beckmann, M. - "Edgeworth-Bertrand Duopoly Revisited," 1965, in: R. Henn (ed.) - *Operations Research-Verfahren*, III, Meisenheim: Verlag Anton Hein.
- Bell, D. - "Regret in Decision Making under Uncertainty," 1982, *Operations Research*, 30, 961-981.
- Bell, D. - "Disappointment in Decision Making under Uncertainty," 1985, *Operations Research*, 33, 1-27.
- Benassy, J.P. - "Neo-Keynesian Disequilibrium in a Monetary Economy," 1975, *Review of Economic Studies*, 42, 502-523.
- Benassy, J.P. - "The Disequilibrium Approach to Monopolistic Price Setting and General Monopolistic Equilibrium," 1976, *Review of Economic Studies*, 43, 69-81.
- Benassy, J.P. - "A Neo-Keynesian Model of Price and Quantity Determination," in: Schwödiauer, G. (ed.) - *Equilibrium and Disequilibrium in Economic Theory*, 1977, Boston: Reidel.
- Benassy, J.P. - *The Economics of Market Disequilibrium*, 1982, New York: Academic Press.
- Benhabib, J. and G. Laroque - "On Competitive Cycles in Productive Economies," 1988, *Journal of Economic Theory*, 45, 145-170.
- Benhabib, J. and K. Nishimura - "The Hopf Bifurcation and the Existence and Stability of Closed Orbits in Multisector Models of Optimum Economic Growth," 1979, *Journal of Economic Theory*, 21, 421-444.
- Benhabib, J. and K. Nishimura - "Competitive Equilibrium Cycles," 1985, *Journal of Economic Theory*, 35, 284-306.
- Benoit, J.P. and V. Krishna - "Finitely Repeated Games," 1985, *Econometrica*, 53, 905-922.
- Ben-Porath Y. - "The Production of Human Capital and the Life Cycle of Earnings," 1967, *Journal of Political Economy*, 75, 352-365.
- Bergson, A. - "A Reformulation of Certain Aspects of Welfare Economics," 1938,

- Quarterly Journal of Economics*, 52, 310-334.
- Bernheim, D. - "Strategic Deterrence of Sequential Entry into an Industry," 1984, *Rand Journal of Economics*, 15, 1-11.
- Berry, C.H. - "Corporate Diversification and Market Structure," 1973, *Bell Journal of Economics*, 4, 196-204.
- Bertrand, J. - "Théorie mathématique de la richesse sociale," 1883, *Journal des savants*, 499-508.
- Besen, S.M. and L.L. Johnson - *Compatibility Standards, Competition, and Innovation in the Broadcasting Industry*, 1986, Santa Monica, Cal.: RAND Publication Series.
- Bhagwati, J. - "Trade Diverting Customs Unions and Welfare Improvement: A Clarification," 1971, *Economic Journal*, 81, 580-587.
- Bhagwati, J.N. and T.N. Srinivasan - "Revenue Seeking: A Generalization of the Theory of Tariffs," 1980, *Journal of Political Economy*, 88, 1069-1087.
- Bhattacharya, G. - "Learning and the Behavior of Potential Entrants," 1984, *Rand Journal of Economics*, 15, 281-289.
- Bhattacharya, G. - "Strategic Learning and Entry-Equilibrium," 1985, *Journal of Economic Dynamics and Control*, 9, 195-223.
- Bierings, H.B.A. and A. van Witteloostuijn - "Werkloosheid en vacatures: inzichten vanuit de economische theorie," 1989, *Maandschrift economie*, 53, 382-400.
- Biggadike, E.R. - *Corporate Diversification: Entry Strategy and Performance*, 1976, Cambridge, Mass.: Harvard University Press.
- Binswanger, H.P. - "A Microeconomic Approach to Induced Innovation," 1974, *Economic Journal*, 84, 940-958.
- Binswanger, H.P. - "The Microeconomics of Induced Technical Change," in: Binswanger, H.P. and V.W. Ruttan (eds.) - *Induced Innovation: Technology, Institutions and Development*, 1978, Baltimore: Johns Hopkins University Press.
- Birtwistle, G.M. - *Discrete Event Modelling on Simula*, 1979, London: McMillan.
- Blanchard, O.J. and N. Kiyotaki - "Monopolistic Competition and the Effects of Aggregate Demand," 1987, *American Economic Review*, 77, 647-666.
- Blaug, M. - "The Empirical Status of Human Capital Theory: A Slightly Jaundiced Survey," 1976, *Journal of Economic Literature*, 14, 827-855.
- Blaug, M. - *The Methodology of Economics, or, How Economists Explain*, 1980, Cambridge: Cambridge University Press.
- Blinder, A.S. - "Inventories and Sticky Prices: More on the Microfoundations of Macroeconomics," 1982, *American Economic Review*, 72, 334-348.
- Böhm, V. - "Disequilibrium Dynamics in a Simple Macroeconomic Model," 1978, *Journal of Economic Theory*, 17, 179-199.
- Böhm-Bawerk, E. von - *Kapital und Kapitalzins*, 1889, English translation, *Capital and Interest*, 1959 (three volumes), South Holland, Ill.: Libertarian Press.
- Boland, L.A. - "On the Futility of Criticizing the Neoclassical Maximization Hypothesis," 1981, *American Economic Review*, 71, 1031-1036.
- Bonanno, G. - "Location Choice, Product Proliferation and Entry Deterrence," 1987, *Review of Economic Studies*, 54, 37-45.
- Bond, E.W. and L. Samuelson - "Durable Goods, Market Structure and the Incentives to Innovate," 1987, *Economica*, 54, 57-67.
- Boschen, J.F. and L.O. Mills - "Tests of the Relation Between Money and Output in the Real Business Cycle Model," 1988, *Journal of Monetary Economics*, 22, 355-374.
- Bosworth, D. and T. Westaway - "The Influence of Demand and Supply Side Pressures

- on the Quantity and Quality of Inventive Activity," 1984, *Applied Economics*, 16, 131-146.
- Braeutigam, R.R. - "Optimal Policies for Natural Monopolies," in: Schmalensee, R. and R.D. Willig (eds.) - *Handbook of Industrial Organization*, 1989, Amsterdam: North Holland.
- Brander, J.A. and J. Eaton - "Product Line Rivalry," 1984, *American Economic Review*, 74, 323-326.
- Brander, J.A. and P.R. Krugman - "A 'Reciprocal Dumping' Model of International Trade," 1983, *Journal of International Economics*, 15, 313-321.
- Brander, J.A. and P.J. Spencer - "Trade Warfare: Tariffs and Cartels," 1984, *Journal of Industrial Economics*, 16, 227-242.
- Branson, W.H. - *Macroeconomic Theory and Policy*, 1979, 2nd edition, New York: Harper and Row.
- Breyer, S. - "Analyzing Regulatory Failure: Mismatches, Less Restrictive Alternatives, and Reform," in: Ogus, A.I. and C.G. Veljanovski (eds.) - *Readings in the Economics of Law and Regulation*, 1984, Oxford: Clarendon Press.
- Brock, W.A. - "Contestable Markets: An Uprising in the Theory of Industry Structure: A Review Article," 1983, *Journal of Political Economy*, 91, 1055-1066.
- Brock, W.A. - "Pricing, Predation, and Entry Barriers in Regulated Industries," in: Evans, D.S. (ed.) - *Breaking up Bell*, 1983, New York: North-Holland.
- Brozen, Y. - "Competition, Efficiency and Anti-Trust," 1969, reprinted in Brozen, Y. (ed.) - *The Competitive Economy: Selected Readings*, 1975, Morristown, N.J.: General Learning Press.
- Brunner, E. - "A Note on Potential Competition," 1961, *Journal of Industrial Economics*, 9, 248-250.
- Bulow, J.I., J.D. Geanakoplos and P.D. Klemperer - "Multimarket Oligopoly: Strategic Substitutes and Complements," 1985a, *Journal of Political Economy*, 93, 488-511.
- Bulow, J.I., J.D. Geanakoplos and P.D. Klemperer - "Holding Idle Capacity to Deter Entry," 1985b, *Economic Journal*, 95, 178-182.
- Cagan, P. - "The Monetary Dynamics of Hyperinflation," 1956, in: Friedman, M. (ed.) - *Studies in the Quantity Theory of Money*, Chicago: Chicago University Press.
- Cairns, R.D. and D. Mahabir - "Contestability: A Revisionist View," 1988, *Economica*, 55, 269-276.
- Calem, P.S. - "Entry and Entry Deterrence in Penetrable Markets," 1988, *Economica*, 55, 171-183.
- Call, G.D. and T.E. Keeler - "Airline Deregulation, Fares, and Market Behavior: Some Empirical Evidence," in: Daughety, A.F. (ed.) - *Analytical Studies in Transport Economics*, 1984, Cambridge: Cambridge University Press.
- Capozza, D.R. and R. van Order - "A Generalised Model of Spatial Competition," 1982, *American Economic Review*, 68, 896-908.
- Carlton, D.W. - "The Theory and the Facts of How Markets Clear: Is Industrial Organization Valuable for Understanding Macroeconomics," in: Schmalensee, R. and R.D. Willig (eds.) - *Handbook of Industrial Organization*, 1989, Amsterdam: North-Holland.
- Carr, J.L. and G.F. Mathewson - "Unlimited Liability as a Barrier to Entry," 1988, *Journal of Political Economy*, 96, 766-784.
- Carvalho, F. - "Alternative Analyses of Short and Long Run in Post Keynesian Economics," 1984, *Journal of Post Keynesian Economics*, 7, 214-234.

- Casas, F.R. - "Monopoly as a Basis for Trade," 1989, *Canadian Journal of Economics*, 22, 195-201.
- Cassidy, P.A. - "Australian Overseas Cargo Liner Shipping," 1982, *Working Paper*, University of Queensland.
- Caves, R.E. - *Multinational Enterprise and Economic Analysis*, 1982, Cambridge: Cambridge University Press.
- Caves, R.E. and R.M. Bradburd - "The Empirical Determinants of Vertical Integration," 1988, *Journal of Economic Behavior and Organization*, 9, 265-279.
- Caves, R.E. and M.E. Porter - "From Entry Barriers to Mobility Barriers: Conjectural Decisions and Contrived Deterrence to New Competition," 1977, *Quarterly Journal of Economics*, 91, 241-261.
- Cayseele, P. van - "Spillovers and the Cost of Multiproject R&D," 1986, *Managerial and Decision Economics*, 7, 133-139.
- Cayseele, P. van - "Economies of Scope in Research and Development," 1987, *Journal of Economics*, 47, 273-285.
- Cayseele, P. van and H. Schreuder - "Strategiebepaling door ondernemingen: een overzicht," 1988, *Economisch statistische berichten*, 73, 1152-1159.
- Cayseele, P. van and H. Schreuder - "Strategische groepen: een overzicht van onderzoek," 1989, *Maandblad voor accountancy en bedrijfshuishoudkunde*, 63, 497-506.
- Chacholiades, M. - *International Trade Theory and Policy*, 1978, New York: McGraw-Hill.
- Chamberlin, E.H. - *The Theory of Monopolistic Competition*, 1933, Cambridge, Mass.: Harvard University Press.
- Champsaur, P. and J.C. Rochet - "Multiproduct Duopolists," 1989, *Econometrica*, 57, 533-557.
- Chang, Y. and H. Thomas - "The Impact of Diversification Strategy on Risk Return Performance," 1989, *Strategic Management Journal*, 10, 271-284.
- Chari, V.V., L.E. Jones and R.E. Manuelli - "Labor Contracts in a Model of Imperfect Competition," 1989, *American Economic Review: Papers and Proceedings*, 79, 358-363.
- Chatterjee, S. and R. Cooper - "Multiplicity of Equilibria and Fluctuations in Dynamic Imperfectly Competitive Economies," 1989, *American Economic Review: Papers and Proceedings*, 79, 353-357.
- Chew, S. and K. McCrimmon - "Alpha-nu Choice Theory: A Generalization of Expected Utility Theory," 1972, *Working Paper 669*, University of British Columbia.
- Clark, J.B. - *Essentials of Economic Theory*, 1915, London: McMillan.
- Clark, J.N. - "Toward a Concept of Workable Competition," 1940, *American Economic Review*, 30, 241-256.
- Clark, J.M. - "Competition: Static Models and Dynamic Aspects," 1955, *American Economic Review*, 45, 450-462.
- Clark, J.M. - *Competition as a Dynamic Process*, 1961, Washington, D.C.: Brookings Institution.
- Clark, W.A.V. - "Recent Research on Migration and Mobility: A Review and Interpretation," 1982, *Progress in Planning*, 18, 1-56.
- Clarke, R.N. - "Duopolists Don't Wish to Share Information," 1983a, *Economics Letters*, 11, 33-36.
- Clarke, R.N. - "Collusion and the Incentives for Information Sharing," 1983b, *Bell*

- Journal of Economics*, 14, 383-395.
- Clarkson, K.W. and R.L. Miller - *Industrial Organization: Theory, Evidence, and Public Policy*, 1983, Singapore: McGraw-Hill.
- Clemens, E.W. - "Price Discrimination and the Multi Product Firm," 1951, *Review of Economic Studies*, 19, 1-11.
- Clower, R.W. - "The Keynesian Counter-Revolution: A Theoretical Appraisal," in: Brechling, F. and F.H. Hahn (eds.) - *The Theory of Interest Rates*, 1965, London: McMillan.
- Clower, R.W. - "A Reconsideration of the Microfoundations of Monetary Theory," 1967, *Western Economic Journal*, 6, 1-9.
- Coase, R.H. - "The Problem of Social Cost," 1960, *Journal of Law and Economics*, 3, 1-44.
- Cohen, W.M. and R.C. Levin - "Empirical Studies of Innovation and Market Structure," in: Schmalensee, R. and R.D. Willig (eds.) - *Handbook of Industrial Organization*, 1989, Amsterdam: North-Holland.
- Cohen, W.M., R.C. Levin and D.C. Mowery - "Firm Size and R&D Intensity: A Re-Examination," 1987, *Journal of Industrial Economics*, 35, 543-565.
- Cohen, W.M. and D.A. Levinthal - "Innovation and Learning: The Two Faces of R&D," 1989, *Economic Journal*, 99, 569-596.
- Collard, D. - *Altruism and Economy*, 1978, Oxford: Martin Robertson.
- Comanor, W.S. and T.A. Wilson - "Advertising, Market Structure and Performance," 1967, *Review of Economics and Statistics*, 49, 423-440.
- Comanor, W.S. and T.A. Wilson - "The Effect of Advertising on Competition: A Survey," 1979, *Journal of Economic Literature*, 17, 453-476.
- Conlisk, J. - "Optimization Cost," 1988, *Journal of Economic Behavior and Organization*, 9, 213-228.
- Connor, K.R. - "Strategies for Product Cannibalism," 1988, *Strategic Management Journal*, 9, 9-26.
- Cool, K. and D. Schendel - "Strategic Group Formation and Performance: The Case of the U.S. Pharmaceutical Industry, 1963-1982," 1987, *Management Science*, 33, 1102-1124.
- Cooper, M.H. and Culyer, A.H. - *Health Economics*, 1973, London: Penguin Books.
- Cooper, R. and A. John - "Coordinating Coordination Failures in Keynesian Models," 1988, *Quarterly Journal of Economics*, 103, 441-463.
- Corden, W.M. - *Inflation, Exchange Rates and the World Economy*, 1981, Oxford: Clarendon Press.
- Cornes, R. and T. Sandler - *The Theory of Externalities, Public Goods, and Club Goods*, 1986, Cambridge: Cambridge University Press.
- Cournot, A. - *Researches into the Mathematical Principles of the Theory of Wealth*, 1897, New York: McMillan, first published in 1838 in French.
- Coursey, D., R.M. Isaac and V.L. Smith - "Natural Monopoly and Contested Markets: Some Experimental Results," 1984a, *Journal of Law and Economics*, 27, 91-113.
- Coursey, D., R.M. Isaac, M. Luke and V.L. Smith - "Market Contestability in the Presence of Sunk (Entry) Costs," 1984b, *Rand Journal of Economics*, 15, 69-84.
- Crawford, V.P. - "A Game of Fair Division," 1977, *Review of Economic Studies*, 44, 235-247.
- Crawford, V.P. - "A Procedure for Generating Pareto-Efficient Egalitarian-Equivalent Allocations," 1979a, *Econometrica*, 47, 49-60.

- Crawford, V.P. - "On Compulsory Arbitration Schemes," 1979b, *Journal of Political Economy*, 87, 131-160.
- Crawford, V.P. - "A Self-Administered Solution of the Bargaining Problem," 1980, *Review of Economic Studies*, 47, 385-392.
- Creedy, J. and K. Whitfield - "The Economic Analysis of Internal Labour Markets," 1988, *Bulletin of Economic Research*, 40, 247-269.
- Cubbin, J. - "Advertising and the Theory of Entry Barriers," 1981, *Economica*, 48, 289-298.
- Cyert, R.M. and J.G. March - *A Behavioural Theory of the Firm*, 1963, Englewood Cliffs, N.J.: Prentice-Hall.
- Dana, R.A. and L. Montrucchio - "Dynamic Complexity in Duopoly Games," 1986, *Journal of Economic Theory*, 40, 40-56.
- Darby, M.R. - "Rational Expectations under Conditions of Costly Information," 1976, *Journal of Finance*, 31, 889-895.
- Dasgupta, P. - "The Theory of Technological Competition," 1986, in: Stiglitz, J.E. and G.F. Mathewson (eds.) - *New Developments in the Analysis of Market Structure*, Cambridge, Mass.: MIT Press..
- Dasgupta, P. and E. Maskin - "The Simple Economics of Research Portfolios," 1987, *Economic Journal*, 97, 581-595.
- Dasgupta, P. and Stiglitz, J. - "Industrial Structure and the Nature of Innovative Activity," 1980a, *Economic Journal*, 90, 266-293.
- Dasgupta, P. and J.E. Stiglitz - "Uncertainty, Industrial Structure and the Speed of R&D," 1980b, *Bell Journal of Economics*, 11, 1-28.
- Dasgupta, P. and J.E. Stiglitz - "Potential Competition, Actual Competition and Economic Welfare," 1988a, *European Economic Review*, 32, 569-577.
- Dasgupta, P. and J.E. Stiglitz - "Learning-by-Doing, Market Structure and Industrial and Trade Policies," 1988b, *Oxford Economic Papers*, 40, 246-268.
- Davidson, C. and R. Deneckere - "Long-Run Competition in Capacity, Short-Run Competition in Price, and the Cournot Model," 1986, *Rand Journal of Economics*, 17, 404-415.
- Davidson, P. - *Money and the Real World*, 1972, London: McMillan.
- Davidson, P. - *International Money and the Real World*, 1982, New York: John Wiley.
- Davidson, P. - "Rational Expectations: A Fallacious Foundation for Studying Crucial Decision Making," 1983, *Journal of Post Keynesian Economics*, 5, 182-198.
- Davies, J.E. - "The Theory of Contestable Markets and its Application to the Linear Shipping Industry," 1986, *Working Paper*, Ottawa-Hull: Canadian Transport Commission.
- Davies, J.E. and F.S. Lee - "A Post Keynesian Appraisal of the Contestability Criterion," 1988, *Journal of Post Keynesian Economics*, 11, 3-24.
- Davies, S. - *The Diffusion of Process Innovations*, 1979, Cambridge: Cambridge University Press.
- Davies, S. - "Concentration," in: Davies, S. and B. Lyons (eds.) - *Economics of Industrial Organisation*, 1988, London: Longman.
- Day, R.H. - "Profits, Learning and the Convergence of Satisficing to Marginalism," 1964, *Quarterly Journal of Economics*, 78, 302-311.
- Day, R.H. - "Irregular Growth Cycles," 1982, *American Economic Review*, 72, 406-414.
- Day, R.H. - "The Emergence of Chaos from Classical Economic Growth," 1983, *Quarterly Journal of Economics*, 98, 201-212.

- Deardorff, A.V. and R.M. Stern - "An Economic Analysis of the Effects of the Tokyo Round of Multilateral Trade Negotiations on the United States and Other Major Industrial Countries, 1979, *MTN Studies* 5, Washington, D.C.: Committee on Finance, U.S. Senate.
- Deaton, A. and J. Muellbauer - *Economics and Consumer Behavior*, 1980, Cambridge: Cambridge University Press.
- DeBondt, R.R. - "Limit Pricing, Uncertain Entry, and the Entry Lag," 1976, *Econometrica*, 44, 939-946.
- DeBondt, R.R. - "Innovative Activity and Barriers to Entry," 1977, *European Economic Review*, 10, 95-109.
- DeBondt, R.R. - "Short-Run Industry Performance and Potential Competition," 1978, *Journal of Industrial Economics*, 26, 267-272.
- DeBondt, R.R. and R. Veugelers - "Strategic Investment with Spillovers," 1989, presented at the *Workshop on Microeconomics/Strategic Management Interfaces*, Brussels: EIASM, November 16-17, 1989.
- Debreu, G. - *Theory of Value*, 1959, New York: John Wiley.
- Demsetz, H. - "Why Regulate Utilities," 1968, reprinted in Brozen, Y. (ed.) - *The Competitive Economy: Selected Readings*, 1975, Morristown, N.J.: General Learning Press.
- Demsetz, H. - "Barriers to Entry," 1982, *American Economic Review*, 72, 47-57.
- Denison, E.F. - *The Sources of Economic Growth in the United States and the Alternatives Before Us*, 1962, New York: Committee for Economic Development.
- Denison, E.F. - *Why Growth Rates Differ: Postwar Experience in Nine Western Countries*, 1967, Washington, D.C.: Brookings Institution.
- Dewatripont, M. - "Entry Deterrence under Trade Unions," 1987, *European Economic Review*, 31, 149-156.
- Dewatripont, M. - "The Impact of Trade Unions on Incentives to Deter Entry," 1988, *Rand Journal of Economics*, 19, 191-199.
- Diamond, P. - "A Model of Price Adjustment," 1971, *Journal of Economic Theory*, 3, 156-168.
- Dixit, A.K. - "A Model of a Duopoly Suggesting a Theory of Entry Barriers," 1979, *Bell Journal of Economics*, 10, 20-32.
- Dixit, A.K. - "The Role of Investment in Entry Deterrence," 1980, *Economic Journal*, 90, 95-106.
- Dixit, A. - "Imperfect Competition and Public Policy: Recent Developments in Oligopoly Theory," 1982, *American Economic Review*, 72, 12-17.
- Dixit, A.K. and J.E. Stiglitz - "Monopolistic Competition and Optimum Product Diversity," 1977, *American Economic Review*, 67, 297-308.
- Dixon, H. - "Oligopoly Theory Made Simple," in: Davies, S. and B. Lyons (eds.) - *Economics of Industrial Organisation*, 1988, London: Longman.
- Doeringer, P. and M. Piore - *International Labour Markets and Manpower Analysis*, 1971, Lexington, Mass: Heath.
- Dosi, G. - "Sources, Procedures, and Microeconomic Effects of Innovation," 1988, *Journal of Economic Literature*, 26, 1120-1171.
- Dosi, G., C. Freeman, R. Nelson, G. Silverberg and L. Soete (eds.) - *Technical Change and Economic Theory*, 1988, London: Pinter.
- Downs, A. - *An Economic Theory of Democracy*, 1957, New York: Harper and Row.
- Drandakis, E.M. and E.S. Phelps - "A Model of Induced Invention, Growth and

- Distribution," 1966, *Economic Journal*, 76, 823-840.
- Drazen, J.H. - "Existence of an Exchange Equilibrium under Price Rigidities," 1980, *Econometrica*, 48, 283-305.
- Drèze, J.H. - "Existence of an Exchange Equilibrium under Price Rigidities," 1975, *International Economic Review*, 16, 301-320.
- Duménil, G. and D. Lévy - "The Classics and the Neoclassicals: A Rejoinder to Frank Hahn," 1985, *Cambridge Journal of Economics*, 9, 327-345.
- Duménil, G. and D. Lévy - "The Dynamics of Competition: A Restoration of the Classical Analysis," 1987a, *Cambridge Journal of Economics*, 11, 133-164.
- Duménil, G. and D. Lévy - "The Macroeconomics of Disequilibrium," 1987b, *Journal of Economic Behavior and Organization*, 8, 377-395.
- Dunning, J.H. - *Multinationals, Technology and Competitiveness*, 1988, London: George Allen and Unwin.
- Dupuit, J. - "On the Measurement of the Utility of Public Works," in: Arrow, K.J. and T. Scitovsky (eds.) - *Readings in Welfare Economics*, 1969, London: George Allen and Unwin, first published in French in 1844.
- Dybvig, P.H. and C.S. Spatt - "Adoption Externalities as Public Goods," 1983, *Journal of Public Economics*, 20, 231-247.
- Eads, G.C. - "Competition in the Domestic Trunk Airline Industry: Too Much or too Little?" in: Phillips, A. (ed.) - *Promoting Competition in Regulated Markets*, 1975, Washington, D.C.: Brookings Institution.
- Earl, P.E. - *The Economic Imagination: Towards a Behavioural Analysis of Choice*, 1983, Armonk, N.Y.: M.E. Sharpe.
- Earl, P.E. - *The Corporate Imagination: How Big Companies Make Mistakes*, 1984, Armonk, N.Y.: M.E. Sharpe.
- Eaton, B.C. and R.G. Lipsey - "The Principle of Minimum Differentiation Reconsidered: Some New Developments in the Theory of Spatial Competition," 1975, *Review of Economic Studies*, 42, 27-49.
- Eaton, B.C. and R.G. Lipsey - "Exit Barriers are Entry Barriers: The Durability of Capital as a Barrier to Entry," 1980, *Bell Journal of Economics*, 11, 721-729.
- Eaton, B.C. and R.G. Lipsey - "Product Differentiation," in: Schmalensee, R. and R.D. Willig (eds.) - *Handbook of Industrial Organization*, 1989, Amsterdam: North-Holland.
- Eaton, B.C. and R. Ware - "A Theory of Market Structure with Sequential Entry," 1987, *Rand Journal of Economics*, 18, 1-16.
- Eatwell, J. - "The Irrelevance of Returns to Scale in Sraffa's Analysis," 1977, *Journal of Economic Literature*, 15, 61-68.
- Eckwert, B. and U. Schittko - "Disequilibrium Dynamics," 1988, *Scandinavian Journal of Economics*, 90, 189-209.
- Edgeworth, F.Y. - *Mathematical Psychics*, 1881, London: Kegan Paul.
- Edgeworth, F.Y. - *The Pure Theory of Monopoly*, 1925, London: McMillan.
- Edwards, C.D. - "Conglomerate Bigness as a Source of Power," in: National Bureau of Economic Research Conference Report, *Business Concentration and Price Policy*, 1955, Princeton, N.J.: Princeton University Press.
- Edwards, W. - "How to Use Multi-Attribute Utility Measurement for Social Decision Making," 1976, *Report 01597-I-T*, University of Southern California.
- Edwards, W. - "How to Use Multi-Attribute Utility Measurement for Social Decision Making," 1977, *IEEE Transactions on Systems, Man and Cybernetics*, 7, 1-30.
- Eichner, A.S. - *The Megacorp and Oligopoly*, 1976, Cambridge: Cambridge University

Press.

- Eichner, A.S. - *A Guide to Post Keynesian Economics*, 1979, White Plains, N.Y.: M.E. Sharpe.
- Eichner, A.S. and J.A. Kregel - "An Essay on Post-Keynesian Theory: A New Paradigm in Economics," 1975, *Journal of Economic Literature*, 13, 1293-1314.
- Eijgelshoven, P.J. - *De kritiek op de neo-klassieke allocatie- en verdelingstheorie*, 1982, Groningen: Wolters-Noordhoff.
- Einhorn, H.J. and R.M. Hogarth - "Decision Making under Ambiguity," 1986, *Journal of Business*, 59, S225-S249.
- Elster, J. and A. Hylland - *Foundations of Social Choice Theory*, 1986, Cambridge: Cambridge University Press.
- Encaoua, D., P. Geroski and A. Jacquemin - "Strategic Competition and the Persistence of Dominant Firms: A Survey," 1986, in: Stiglitz, J.E. and G.F. Mathewson (eds.) - *New Developments in the Analysis of Market Structure*, Cambridge, Mass.: MIT Press.
- Encarnation, D.J. - "Cross-Investment: A Second Front of Economic Rivalry," 1987, *California Management Review*, 24, 20-48.
- Evenson, R. and Y. Kislev - *Agricultural Research and Productivity*, 1975, New Haven, Conn.: Yale University Press.
- Fama, E.F. and A.B. Laffer - "The Number of Firms and Competition," 1972, *American Economic Review*, 62, 670-674.
- Farrell, J. - "How Effective is Potential Competition?," 1986, *Economics Letters*, 20, 67-70.
- Farrell, J. and G. Saloner - "Standardization, Compatibility, and Innovation," 1985, *Rand Journal of Economics*, 16, 70-80.
- Farrell, J. and G. Saloner - "Installed Base and Compatibility: Innovation, Product Preannouncements and Predation," 1986, *American Economic Review*, 76, 940-955.
- Feige, E.L. and D.K. Pearce - "Economically Rational Expectations: Are Innovations in the Rate of Inflation Independent of Innovations in Measures of Monetary and Fiscal Policy?," 1976, *Journal of Political Economy*, 84, 499-522.
- Feinberg, R.M. - "Mutual Forbearance as an Extension of Oligopoly Theory," 1984, *Journal of Economics and Business*, 36, 243-249.
- Feinberg, R.M. - "'Sales-at-Risk': A Test of the Mutual Forbearance Theory of Conglomerate Behavior," 1985, *Journal of Business*, 58, 225-241.
- Feldman, A. and A. Kirman - "Fairness and Envy," 1974, *American Economic Review*, 64, 995-1005.
- Fellner, W. - "Two Propositions in the Theory of Induced Innovation," 1961, *Economic Journal*, 71, 305-308.
- Ferguson, C.E. - *A Macroeconomic Theory of Workable Competition*, 1964, Durham, North Carolina: Duke University Press.
- Festinger, L. - *A Theory of Cognitive Dissonance*, 1957, Evanston, Ill.: Row, Peterson.
- Field, W.P. Jr. - *Basic Economics*, 1983, Boston: Allyn and Bacon.
- Fishburn, P.C. - "Foundations of Decision Analysis: Along the Way," 1989, *Management Science*, 35, 387-405.
- Fisher, F.M. - "Games Economists Play: A Noncooperative View," 1989, *Rand Journal of Economics*, 20, 113-137.
- Fisher, F.M. and K. Shell - "Taste and Quality Change in the Pure Theory of the True Cost-of-Living Index," in: Z. Griliches (ed.) - *Price Indexes and Quality Change*, 1971, Cambridge, Mass.: Harvard University Press.

- Fitoussi, J.P. (ed.) - *Modern Macroeconomic Theory*, 1983, Oxford: Basil Blackwell.
- Foley, D. - "Resource Allocation and the Public Sector," 1967, *Yale Economic Essays*, 7, 45-98.
- Frank, R.H. - "If Homo Economicus Could Choose His Own Utility Function, Would He Want One with a Conscience?," 1987, *American Economic Review*, 77, 593-604.
- Freeman, C. - *The Economics of Industrial Innovation*, 1982, 2nd edition, London: Pinter.
- Frey, B.S. and K. Foppa - "Human Behavior: Possibilities Explain Action," 1986, *Journal of Economic Psychology*, 7, 137-160.
- Fried, D. - "Incentives for Information Production and Disclosure in a Duopolistic Environment," 1984, *Quarterly Journal of Economics*, 99, 367-381.
- Friedman, B.M. - "The Theoretical Non-Debate about Monetarism," 1976, *Kredit und Kapital*, 9, 347-367.
- Friedman, J.W. - "Cooperative Equilibria in Finite Horizon Noncooperative Supergames," 1985, *Journal of Economic Theory*, 35, 390-398.
- Friedman J.W. - *Game Theory: With Applications to Economics*, 1986, Oxford: Oxford University Press.
- Friedman, J.W. and R.W. Rosenthal - "A Positive Approach to Non-Cooperative Games," 1986, *Journal of Economic Behavior and Organization*, 7, 235-251.
- Friedman, M. - *Essays in Positive Economics*, 1953, Chicago: University of Chicago Press.
- Friedman, M. - "The Role of Monetary Policy," 1968, *American Economic Review*, 58, 1-17.
- Friedman, M. and L.J. Savage - "The Utility Analysis of Choices Involving Risk," 1948, *Journal of Political Economy*, 56, 279-304.
- Friedman, S.R. - "Structure, Process, and the Labor Market," 1984, in: Darity, W. (ed.) - *Labor Economics: Modern Views*, Boston: Kluwer-Nijhoff.
- Froeb, L. and J. Geweke - "Perfect Contestability and the Postwar U.S. Aluminium Industry," 1984, *Working Paper*, Tulane University.
- Fudenberg, D., R.J. Gilbert, J.E. Stiglitz and J. Tirole - "Preemption, Leapfrogging and Competition in Patent Races," 1983, *European Economic Review*, 22, 3-31.
- Fudenberg, D. and E. Maskin - "The Folk Theorem in Repeated Games with Discounting and with Incomplete Information," 1983, *Working Paper*, MIT.
- Fudenberg, D. and E. Maskin - "Understanding Rent Dissipation: On the Use of Game Theory in Industrial Organization," 1987, *American Economic Review*, 77, 176-183.
- Fudenberg, D. and J. Tirole - "Capital as a Commitment: Strategic Investment to Deter Mobility," 1983a, *Journal of Economic Theory*, 31, 227-250.
- Fudenberg, D. and J. Tirole - "Learning-by-Doing and Market Performance," 1983b, *Bell Journal of Economics*, 14, 522-530.
- Fudenberg, D. and J. Tirole - "The Fat-Cat Effect, the Puppy-Dog Ploy, and the Lean and Hungry Look," 1984, *American Economic Review: Papers and Proceedings*, 74, 361-366.
- Fudenberg, D. and J. Tirole - "Preemption and Rent Equalization in the Adoption of New Technology," 1985, *Review of Economic Studies*, 52, 383-401.
- Fudenberg, D. and J. Tirole - "A 'Signal-Jamming' Theory of Predation," 1986, *Rand Journal of Economics*, 17, 366-376.
- Fuhrer, J.C. - "Information Gathering and Expectation Formation under Model Uncertainty," 1987, *Southern Economic Journal*, 54, 658-701.

- Furnham, A. and A. Lewis - *The Economic Mind: The Social Psychology of Economic Behaviour*, 1986, Brighton: Wheatsheaf Books.
- Futia, C.A. - "Schumpeterian Competition," 1980, *Quarterly Journal of Economics*, 95, 675-695.
- Gabszewicz, J.J. and J.F. Thisse - "Entry (and Exit) in a Differentiated Industry," 1980, *Journal of Economic Theory*, 22, 327-338.
- Gabszewicz, J.J. and J.F. Thisse - "On the Nature of Competition with Differentiated Products," 1986, *Economic Journal*, 96, 160-172.
- Gabszewicz, J.J. and J.P. Vial - "Oligopoly 'à la Cournot' in a General Equilibrium Analysis," 1972, *Journal of Economic Theory*, 4, 381-400.
- Gailbraith, J.K. - *American Capitalism*, 1952, Boston: Houghton Mifflin.
- Gale, D. - "Pure Exchange Equilibrium of Dynamic Economic Model," 1973, *Journal of Economic Theory*, 6, 12-36.
- Gale, D. - "A Note on Conjectural Equilibria," 1978, *Review of Economic Studies*, 45, 33-38.
- Gallini, N.T. - "Deterrence by Market Sharing: A Strategic Incentive for Licensing," 1984, *American Economic Review*, 74, 931-941.
- Gallini, N.T. and Y. Kotowitz - "Optimal R and D Processes and Competition," 1985, *Economica*, 52, 321-334.
- Gallini, N.T. and R.R. Winter - "Licensing in the Theory of Innovation," 1985, *Rand Journal of Economics*, 16, 237-252.
- Garegnani, P. - "On a Change in the Notion of Equilibrium in Recent Work on Value and Distribution," 1976, in: Brown, M., K. Sato and P. Zarembka (eds.) - *Essays in Modern Capital Theory*, Amsterdam: North-Holland.
- Garegnani, P. - "Notes on Consumption, Investment and Effective Demand: I," 1978, *Cambridge Journal of Economics*, 2, 335-353.
- Garegnani, P. - "Notes on Consumption, Investment and Effective Demand: II," 1979, *Cambridge Journal of Economics*, 3, 63-82.
- Gaskins, D.W., Jr. - "Dynamic Limit Pricing: Optimal Pricing under Threat of Entry," 1971, *Journal of Economic Theory*, 3, 306-322.
- Gehrels, F. - "Customs Union from a Single-Country Viewpoint," 1956, *Review of Economic Studies*, 24, 61-64.
- Gelauff, G.M.M., A.R.M. Wennekers and A.H.M. de Jong - "A Putty-Clay Model with Three Factors of Production and Partly Endogenous Technical Progress," 1985, *De Economist*, 133, 327-351.
- Gerowski, P. - "Competition Policy and the Structure-Performance Paradigm," in: Davies, S. and B. Lyons (eds.) - *Economics of Industrial Organisation*, 1988, London: Longman.
- Ghemawat, H. - "Capacity Expansion in the Titanium Dioxide Industry," 1984, *Journal of Industrial Economics*, 33, 145-162.
- Ghemawat, H. and B. Nalebuff - "Exit," 1985, *Rand Journal of Economics*, 15, 184-194.
- Gilbert, R.J. - "Preemptive Competition," 1986, in: Stiglitz, J.E. and G.F. Mathewson (eds.) - *New Developments in the Analysis of Market Structure*, 1986, Cambridge, Mass.: MIT Press.
- Gilbert, R.J. - "The Role of Potential Competition in Industrial Organization," 1989a, *Journal of Economic Perspectives*, 3, 99-106.
- Gilbert, R.J. - "Mobility Barriers and the Value of Incumbency," in: Schmalensee, R.

- and R.D. Willig (eds.) - *Handbook of Industrial Organization*, 1989b, Amsterdam: North-Holland.
- Gilbert, R.J. and D. Newberry - "Preemptive Patenting and the Persistence of Monopoly," 1982, *American Economic Review*, 72, 3, 514-526.
- Gilbert, R.J. and X. Vives - "Entry Deterrence and the Free Rider Problem," 1986, *Review of Economic Studies*, 53, 71-83.
- Glustoff, E. - "On the Existence of a Keynesian Equilibrium," 1968, *Review of Economic Studies*, 35, 327-334.
- Goldberg, V. and S. Moirao - "Limit Pricing and Potential Competition," 1973, *Journal of Political Economy*, 81, 1460-1466.
- Gort, M. - *Diversification and Integration in American Industry*, 1962, NBER, Princeton, N.J.: Princeton University Press.
- Gort, M. and R.A. Wall - "The Evolution of Technologies and Investment in Innovation," 1986, *Economic Journal*, 96, 741-757.
- Gossen, H.H. - *Entwicklung der Gesetze des menschlichen Verkehrs und der daraus fließenden Regeln für menschliches Handeln*, 1927, Berlin: Prager, first published in 1854.
- Grabowski, H.G. and J.M. Vernon - "Pioneers, Imitators, and Generics: A Simulation Model of Schumpeterian Competition," 1987, *Quarterly Journal of Economics*, 102, 491-525.
- Grandmont, J.M. - "Temporary Equilibrium Theory," 1977, *Econometrica*, 45, 535-572.
- Grandmont, J.M. - *Money and Value: A Reconsideration of Classical and Neo-Classical Monetary Theories*, 1983, Cambridge: Cambridge University Press.
- Grandmont, J.M. - "On Endogenous Competitive Business Cycles," 1985, *Econometrica*, 53, 995-1045.
- Grandmont, J.M. and G. Laroque - "On Keynesian Temporary Equilibria," 1976, *Review of Economic Studies*, 43, 53-67.
- Green, C. - "Industrial Organization Paradigms, Empirical Evidence, and the Economic Case for Competitive Policy," 1987, *Canadian Journal of Economics*, 20, 482-505.
- Green, J.R. - "Temporary General Equilibrium in a Sequential Trading Model with Spot and Future Transactions," 1973, *Econometrica*, 41, 1103-1124.
- Greenwald, G. and J.E. Stiglitz - "Toward a Theory of Rigidities," 1989, *American Economic Review: Papers and Proceedings*, 79, 364-369.
- Grether, D.M. and C.R. Plott - "Economic Theory of Choice and the Preference Reversal Phenomenon," 1979, *American Economic Review*, 69, 623-638.
- Griesinger, D.W. and J.W. Livingston - "Toward a Model of Interpersonal Motivation in Experimental Games," 1973, *Behavioral Science*, 18, 173-188.
- Grossman, S. - "Nash Equilibrium and the Industrial Organization of Markets with Large Fixed Costs," 1981, *Econometrica*, 49, 1149-1172.
- Guesnerie, R. - "Second-Best Pricing Rules," 1980, *Journal of Public Economics*, 13, 51-80.
- Guesnerie, R. and O. Hart - "Welfare Losses Due to Imperfect Competition: Asymptotic Results for Cournot Nash Equilibria with and without Free Entry," 1985, *International Economic Review*, 26, 525-545.
- Guilford, J.P. - "Three Faces of Intellect," 1959, *American Psychologists*, 14, 469-479.
- Guilford, J.P. "Some New Looks at the Nature of Creative Processes," in: Frederiksen, N. and H. Gulliksen (eds.) - *Contributions to Mathematical Psychology*, 1964, New York: Holt, Rinehart and Winston.

- Hagen, O. - "Towards a Positive Theory of Preferences under Risk," in: Allais, M. and O. Hagen (eds.) - *Expected Utility Hypotheses and the Allais Paradox*, 1979, Dordrecht: Reidel.
- Hahn, F.H. - "Excess Capacity and Imperfect Competition," 1955, *Oxford Economic Papers*, 7, 229-240.
- Hahn, F.H. - "On Some Problems of Proving the Existence of an Equilibrium in a Monetary Economy," 1965, in: Brechling, F. and F.H. Hahn (eds.) - *The Theory of Interest Rates*, London: McMillan.
- Hahn, F.H. - "Equilibrium with Transaction Costs," 1971, *Econometrica*, 39, 417-439.
- Hahn, F.H. - *On the Notion of Equilibrium in Economics*, 1973, Cambridge: Cambridge University Press.
- Hahn, F.H. - "Exercises in Conjectural Equilibria," 1977, *Scandinavian Journal of Economics*, 79, 210-226.
- Hahn, F.H. - "On Non-Walrasian Equilibria," 1978, *Review of Economic Studies*, 45, 1-17.
- Hahn, F.H. - *Equilibrium and Macroeconomics*, 1984, Oxford: Basil Blackwell.
- Hammar skjöld, D. - *Konjunkturspridningen*, 1932, Stockholm: P.A. Norstedt.
- Hannan, T.H. - "Prices, Capacity, and the Entry Decision: A Conditional Logit Analysis," 1983, *Southern Economic Journal*, 50, 539-550.
- Hannan, M.T. and J. Freeman - "The Population Ecology of Organizations," 1977, *American Journal of Sociology*, 82,, 929-964.
- Hannan, M.T. and J. Freeman - *Organizational Ecology*, 1989, Cambridge, Mass.: Harvard University Press.
- Hansen, A.H. - *A Guide to Keynes*, 1953, New York: McGraw-Hill.
- Hansen, B. - "Excess Demand, Unemployment, Vacancies and Wages," 1970, *Quarterly Journal of Economics*, 84, 1-23.
- Harrigan, K.R. - "Vertical Integration and Corporate Strategy," 1985, *Academy of Management Journal*, 28, 397-425.
- Harrington, J.E., Jr. - "Limit Pricing When the Potential Entrant Is Uncertain of Its Cost Function," 1986, *Econometrica*, 54, 429-437.
- Harris, C. and J. Vickers - "Perfect Equilibrium in a Model of a Race," 1985, *Review of Economic Studies*, 52, 193-209.
- Harrison, G.W. - "Experimental Evaluation of the Contestable Market Hypothesis," in: Bailey, E. (ed.) - *Public Regulation: New Perspectives on Institutions and Policies*, 1986, Cambridge, Mass.: MIT Press.
- Harrison, G.W. and M. McKee - "Monopoly Behavior, Decentralized Regulation, and Contestable Markets: An Experimental Evaluation," 1985, *Rand Journal of Economics*, 16, 51-69.
- Harrison, G.W., M. McKee and E.E. Rutström - "Experimental Evaluation of Institutions of Monopoly Restraint," in: Green, L. and J.H. Kagel (eds.) - *Advances in Behavioral Economics*, 1987, Norwood: Ablex.
- Harrod, R. - *Economic Essays*, 1953, London: McMillan.
- Hart, O.D. - "Perfect Competition and Optimal Product Differentiation," 1980, *Journal of Economic Theory*, 22, 279-312.
- Hashimoto, M. - "Bonus Payments, On-the-job Training, and Lifetime Employment in Japan," 1979, *Journal of Political Economy*, 87, 1086-1104.
- Hatten, K. and M.L. Hatten - "Strategic Groups, Asymmetrical Mobility Barriers and Contestability," 1987, *Strategic Management Journal*, 8, 329-342.

- Hause, J.C. and G. du Rietz - "Entry, Industry Growth, and the Microdynamics of Industry Supply," 1984, *Journal of Political Economy*, 92, 733-757.
- Haveman, R.H. - "Unemployment in Western Europe and the United States: A Problem of Demand, Structure, or Measurement?," 1978, *American Economic Review*, 68, 44-50.
- Hayek, F.A. - "Das intertemporale Gleichgewichtssystem der Preise und die Bewegungen des 'Geldwertes'," 1928, *Weltwirtschaftliches Archiv*, 28, 33-76.
- Hayek, F.A. - "Economics and Knowledge," 1937, reprinted in *Individualism and Economic Order*, 1949, London: Routledge and Kegan Paul.
- Hayek, F.A. - *The Road to Serfdom*, 1944, Chicago: Chicago University Press.
- Hayek, F.A. - "The Meaning of Competition," 1946, reprinted in *Individualism and Economic Order*, 1949, London: Routledge and Kegan Paul.
- Heggstad, A. and S.A. Rhoades - "Multi-Market Interdependence and Local Market Competition in Banking," 1978, *Review of Economics and Statistics*, 60, 523-532.
- Heiner, R.A. - "The Origin of Predictable Behavior," 1983, *American Economic Review*, 73, 560-595.
- Heiner, R.A. - "Uncertainty, Signal-Detection Experiments, and Modeling Behavior," in: R.N. Langlois (ed.) - *Economics as a Process: Essays in the New Institutional Economics*, 1986, Cambridge: Cambridge University Press.
- Heiner, R.A. - "The Origin of Predictable Dynamic Behavior," 1989, *Journal of Economic Behavior and Organization*, 12, 233-258.
- Helden, G.J. van - "Prijszetting van een naar een aanvaardbare winst strevende onderneming: enkele verschilpunten tussen het marginalisme en de gedragstheorie," 1972, *De Economist*, 120, 329-352.
- Hennipman, P. - "Monopoly: Impediment or Stimulus to Economic Progress?," in: Chamberlin, E.H. (ed.) - *Monopoly and Competition and Their Regulation*, 1954, London: McMillan.
- Hennipman, P. - *Economisch motief en economisch principe*, 1945, Amsterdam: Noord-Hollandsche Uitgevers Maatschappij.
- Hennipman, P. - "Normative or Positive: Mishan's Half-Way House," 1984, *De Economist*, 132, 86-99.
- Herck, G. van - "Entry, Exit and Profitability," 1984, *Managerial and Decision Economics*, 5, 25-30.
- Hey, J.D. - *Uncertainty in Microeconomics*, 1979, Oxford: Martin Robertson.
- Hicks, J.R. - *The Theory of Wages*, 1932, London: McMillan.
- Hicks, J.R. - "Mr. Keynes and the Classics," 1937, *Econometrica*, 5, 147-159.
- Hicks, J.R. - *Value and Capital*, 1939, Oxford: Oxford University Press.
- Hicks, J.R. - "The Four Consumers' Surplus," 1943, *Review of Economic Studies*, 11, 31-41.
- Hicks, J.R. - "Methods of Dynamic Analysis," 1956, reprinted in *The Economics of John Hicks*, 1982, Oxford: Basil Blackwell.
- Hicks, J.R. - "The Process of Imperfect Competition," 1954, *Oxford Economic Papers*, 6, 41-54.
- Hicks, J.R. - *Capital and Growth*, 1965, Oxford: Oxford University Press.
- Hicks, J.R. and R.D.G. Allen - "A Reconsideration of the Theory of Value," 1934, *Economica*, 1 and 2, 52-76 and 196-219.
- Hines, H.H. - "Effectiveness of 'Entry' by Already Established Firms," 1957, *Quarterly Journal of Economics*, 71, 132-150.

- Hirsch, B.T. and Addison, J.T. - *The Economic Analysis of Unions: New Approaches and Evidence*, 1986, Boston: George Allen and Unwin.
- Hirshleifer, J. - "The Expanding Domain of Economics," 1985, *American Economic Review*, 75, 53-68.
- Hirshleifer, J. and Riley, J.G. - "The Analytics of Uncertainty and Information: An Expository Survey," 1979, *Journal of Economic Literature*, 17, 1375-1421.
- Hollander, A. - "On Price-Increasing Entry," 1987, *Economica*, 54, 317-324.
- Holler, M.J. - "The Theory of Contestable Markets: Comment," 1985, *Bulletin of Economic Research*, 37, 65-67.
- Holler, M.J. - "Moral Sentiments and Self-Interest Reconsidered", in: Diekmann, A. and P. Mitter (eds.) - *Paradoxical Effects of Social Behavior: Essays in Honor of Anatol Rapoport*, 1986, Vienna: Physica-Verlag.
- Holler, M.J. - "Strategic Markets: Two Papers on Strategic Market Theory," 1987, *Aarhus Discussion Paper 1987-14*, University of Aarhus.
- Holmström, B. - "Equilibrium Long-Term Labor Contracts," 1983, *Quarterly Journal of Economics*, 98, 23-54.
- Honkapohja, S. - "On the Dynamics of Disequilibria in a Macro Model with Flexible Wages and Prices," in: Aoki, M. and A. Marzollo (eds.) - *New Trends in Dynamic System Theory and Economics*, 1979, London: Academic Press.
- Honkapohja, S. and T. Ito - "Disequilibrium Dynamics with Monetarist Price Expectations," 1982, *Economics Letters*, 9, 69-75.
- Hoogduin, L.H. - "On the Difference between the Keynesian, Knightian and the 'Classical' Analysis of Uncertainty and the Development of a More General Monetary Theory," 1987, *De Economist*, 135, 52-65.
- Hoogduin, L.H. and A. van Witteloostuijn - "On the Fruitfulness of Hicks' Dynamic Methods for the Description of a Monetary Economy: A Reflection on Hicks' 'Methods of Dynamic Economics'," 1987a, *Research Memorandum 216*, University of Groningen.
- Hoogduin, L.H. and A. van Witteloostuijn - "Equilibrium, Uncertainty and Dynamics: Hicks and a Classification of Dynamic Methods," 1987b, *Research Memorandum RM 87-015*, University of Limburg.
- Hotelling, H. - "Stability in Competition," 1929, *Economic Journal*, 39, 446-452.
- Houthakker, H. - "Revealed Preference and the Utility Function," 1950, *Economica*, 17, 159-174.
- Houthakker, H. and L.D. Taylor - *Consumer Demand in the United States 1929-1970*, 1970, 2nd edition, 1st edition 1966, Cambridge, Mass.: Harvard University Press.
- Houtman, M. and J.A.H. Maks - "Determining all Maximal Subsets Consistent with Revealed Preference," 1985, *Kwantitatieve methoden*, 19, 89-104.
- Howard, J. and J. Sheth - *The Theory of Buyer Behavior*, 1969, New York: John Wiley.
- Huber, G.P. - *Managerial Decision Making*, 1980, Glenview, Ill.: Scott, Foresman.
- Hutchison, T.W. - *Revolution and Progress in Economic Knowledge*, 1978, Cambridge: Cambridge University Press.
- Isnard, A.N. - *Traité des richesses*, 1781, London and Lausanne.
- Ito, T. - "Disequilibrium Growth Theory," 1980, *Journal of Economic Theory*, 23, 380-404.
- Ito, T. and S. Honkapohja - "On Macroeconomic Equilibrium with Stochastic Rationing," 1985, *Scandinavian Journal of Economics*, 87, 66-88.
- Iwai, K. - *Disequilibrium Dynamics*, 1981, New Haven, Conn.: Yale University Press.

- Iwai, K. - "Schumpeterian Dynamics: An Evolutionary Model of Innovation and Imitation," 1984a, *Journal of Economic Behavior and Organization*, 5, 159-190.
- Iwai, K. - "Schumpeterian Dynamics Part II: Technological Progress, Firm Growth and Economic Selection," 1984b, *Journal of Economic Behavior and Organization*, 5, 321-351.
- Jacquemin, A. - *The New Industrial Organization: Market Forces and Strategic Behavior*, 1987, Oxford: Clarendon Press.
- Jaffé, W. - "Walras's Economics as Others See It," 1980, *Journal of Economic Literature*, 18, 528-549.
- Jaffé, W. - *William Jaffé's Essays on Walras*, Walker, D.A. (ed.), 1983, Cambridge: Cambridge University Press.
- Jevons, W.S. - *The Theory of Political Economy*, 1981, London: Routledge and Sons.
- Johnson, G. and K. Scholes - *Exploring Corporate Strategy: Text and Cases*, 1989, New York: Prentice-Hall.
- Jones, L.E. - "The Efficiency of Monopolistically Competitive Equilibria in Large Economies: Commodity Differentiation with Gross Substitutes," 1987, *Journal of Economic Theory*, 41, 356-391.
- Jovanovic, B. - "Selection and the Evolution of Industry," 1982, *Econometrica*, 50, 649-670.
- Jovanovic, B. and S. Lach - "Entry, Exit, and Diffusion with Learning by Doing," 1989, *American Economic Review*, 79, 690-699.
- Judd, K.L. - "Credible Spatial Preemption," 1985, *Rand Journal of Economics*, 16, 153-166.
- Judd, K.L. and B.C. Petersen - "Dynamic Limit Pricing and Internal Finance," 1986, *Journal of Economic Theory*, 39, 368-399.
- Kahn, A.E. - *The Economics of Regulation: Principles and Institutions, Vol. 1: Economic Principles*, 1970, New York: John Wiley.
- Kahneman, D. and A. Tversky - "Prospect Theory: An Analysis of Decision under Risk," 1979, *Econometrica*, 47, 263-291.
- Kahneman, D. and A. Tversky - "Rational Choice and the Framing of Decisions," 1986, *Journal of Business*, 59, S251-278.
- Kaldor, N. - "Welfare Propositions of Economics and Intertemporal Comparisons of Utility," 1939, *Economic Journal*, 49, 549-552.
- Kaldor, N. - "A Model of Distribution," 1970, in: Sen, A.K. (ed.) - *Growth Economics*, Harmondsworth: Penguin Books.
- Kalecki, M. - "Costs and Prices," 1943, in: Kalecki, M. - *Studies in Economic Dynamics*, 1943, London: George Allen and Unwin.
- Kalecki, M. - *Studies in Economic Dynamics*, 1943, London: George Allen and Unwin.
- Kalecki, M. - *Theory of Economic Dynamics*, 1954, London: George Allen and Unwin.
- Kalecki, M. - *Selected Essays on the Dynamics of the Capitalist Economies*, 1971, Cambridge: Cambridge University Press.
- Kalish, L., J. Hartzog and H. Cassidy - "The Threat of Entry with Mutually Aware Potential Entrants," 1978, *Journal of Political Economy*, 86, 147-155.
- Kamien, M.I. and N.L. Schwartz - "Limit Pricing and Uncertain Entry," 1971, *Econometrica*, 39, 441-454.
- Kamien, M.I. and N.L. Schwartz - "Timing of Innovations under Rivalry," 1972, *Econometrica*, 40, 43-60.
- Kamien, M.I. and N.L. Schwartz - "Risky R&D with Rivalry," 1974a, *Annals of*

- Economic and Social Measurement*, 3, 276-277.
- Kamien, M.I. and N.L. Schwartz - "Patent Life and R&D Rivalry," 1974b, *American Economic Review*, 64, 183-187.
- Kamien, M.I. and N.L. Schwartz - "Market Structure and Innovative Activity: A Survey," 1975, *Journal of Economic Literature*, 13, 1-37.
- Kamien, M.I. and N.L. Schwartz - "On the Degree of Rivalry for Maximum Innovative Activity," 1976, *Quarterly Journal of Economics*, 1976, 90, 245-260.
- Kamien, M.I. and N.L. Schwartz - "Potential Rivalry, Monopoly Profits and the Pace of Inventive Activity," 1978, *Review of Economic Studies*, 95, 547-557.
- Kamien, M.I. and N.L. Schwartz - "A Generalized Hazard Rate," 1980, *Economics Letters*, 5, 245-249.
- Kamien, M.I. and N.L. Schwartz - *Market Structure and Innovation*, 1982, Cambridge: Cambridge University Press.
- Kantarelis, D. and E.C.H. Veendorp - "Live and Let Live Type Behavior in a Multi-Market Setting with Demand Fluctuations," 1988, *Journal of Economic Behavior and Organization*, 10, 235-244.
- Kaplinsky, R. - "Firm Size and Technical Change in a Dynamic Context," 1983, *Journal of Industrial Economics*, 32, 39-59.
- Kapteijn, A. - "Utility and Economics," 1985, *De Economist*, 133, 1-20.
- Karnani, A. and B. Wernerfelt - "Multiple Point Competition," 1985, *Strategic Management Journal*, 6, 87-96.
- Katz, M.L. and C. Shapiro - "Network Externalities, Competition and Compatibility," 1985, *American Economic Review*, 75, 424-440.
- Katz, M.L. and C. Shapiro - "R&D Rivalry with Licensing or Imitation," 1987, *American Economic Review*, 77, 402-420.
- Keenan, D.W. and D.T. Keenan - "First and Second-Best Pricing with Uncertain Demand," 1984, *Southern Economic Journal*, 51, 80-88.
- Keeney, R.L. and H. Raiffa - *Decisions with Multiple Objectives: Preferences and Value Trade-Offs*, 1976, New York: John Wiley.
- Kemp, M.C. and Y.K. Ng - "On the Existence of Social Welfare Functions, Social Orderings and Social Decision Functions," 1976, *Economica*, 43, 59-66.
- Kennedy, C. - "Induced Bias in Innovation and the Theory of Distribution," 1964, *Economic Journal*, 74, 541-547.
- Kennedy, C. and A.P. Thirlwall - "Surveys in Applied Economics: Technological Progress," 1981, *Economic Journal*, 82, 11-63.
- Kerr, C. - "The Balkanization of Labor Markets," 1954, in: Wight Bakke *et al.* (eds.) - *Labor Mobility and Economic Opportunity*, Cambridge, Mass: MIT Press.
- Keynes, J.M. - *A Treatise on Probability*, 1921, *Collected Writings of John Maynard Keynes: Volume III*, 1973, London: McMillan.
- Keynes, J.M. - *The General Theory of Employment, Money and Interest*, 1936, London: McMillan.
- Keynes, J.M. - "The General Theory of Employment," 1937, *Quarterly Journal of Economics*, 52, 209-223.
- Kihlström, R.E. and J.J. Laffont - "A General Equilibrium Entrepreneurial Theory of Firm Formation Based on Risk Aversion," 1979, *Journal of Political Economy*, 87, 719-748.
- Kihlström, R.E. and J.J. Laffont - "Implicit Labor Contracts and Free Entry," 1983, *Quarterly Journal of Economics*, 98, 55-105.

- Kihlström, R.E. and D. Levhari - "Quality, Regulation and Efficiency," 1977, *Kyklos*, 30, 214-234.
- Killingsworth, M.R. - *Labor Supply*, 1983, Cambridge: Cambridge University Press.
- Kim, J.S. - "Optimal Price-Quality Schedules and Sustainability," 1987, *Journal of Industrial Economics*, 36, 231-244.
- Kirzner, I.M. - *Competition and Entrepreneurship*, 1973, Chicago: University of Chicago Press.
- Kirzner, I.M. - *Perception, Opportunity and Profit*, 1979, Chicago: University of Chicago Press.
- Kirzner, I.M. - *Discovery and the Capitalist Process*, 1985, Chicago: University of Chicago Press.
- Klant, J.J. - *Spelregels voor economen*, 1972, Leiden: Stenfort Kroese.
- Klemperer, P. - "Markets with Consumer Switching Costs," 1985, *Quarterly Journal of Economics*, 100, 375-394.
- Klundert, Th.C.M.J. and R. de Groof - "Induced Technical Progress and Economic Growth," 1977, *De Economist*, 125, 506-524.
- Knieps, G. and I. Vogelsang - "The Sustainability Concept under Alternative Behavioral Assumptions," 1982, *Bell Journal of Economics*, 13, 234-241.
- Knight, F.H. - *Risk, Uncertainty, and Profit*, 1921, London: University of London.
- Koo, A. - "An Empirical Test of Revealed Preference Theory," 1963, *Econometrica*, 31, 646-664.
- Koopmans, T.C. - *Three Essays on the State of Economic Science*, 1957, New York: McGraw-Hill.
- Koutsoyiannis, A. - *Modern Microeconomics*, 1979, New York: McMillan.
- Krattenmaker, T.G. and S.C. Salop - "Anticompetitive Exclusion: Raising Rivals' Cost to Achieve Power over Price," 1986, *Yale Law Journal*, 96, 209-295.
- Kregel, J.A. - "Economic Methodology in the Face of Uncertainty: The Modelling Methods of Keynes and the Post Keynesians," 1976, *Economic Journal*, 86, 209-225.
- Kregel, J.A. - "Money, Expectations and Relative Prices in Keynes' Monetary Equilibrium," 1982, *Economie appliquée*, 35, 449-465.
- Kreps, D.M. and J.A. Scheinkman - "Quantity Precommitment and Bertrand Competition Yield Cournot Outcomes," 1983, *Bell Journal of Economics*, 14, 326-337.
- Kreps, D.M. and R. Wilson - "Reputation and Imperfect Information," 1982, *Journal of Economic Theory*, 27, 253-279.
- Krueger, A.O. - "The Political Economy of the Rent Seeking Society," 1974, *American Economic Review*, 64, 291-303.
- Krugman, P.R. - "Industrial Organization and International Trade," in: Schmalensee, R. and R.D. Willig (eds.) - *Handbook of Industrial Organization*, 1989, Amsterdam: North-Holland.
- Kuipers, S.K. - "Keynesian and Neo-Classical Growth Models: A Sequential Analytical Approach," 1981, *De Economist*, 129, 58-104.
- Kuipers, S.K. and H. van Ees - "Macro-economische theorie en economische politiek," 1990, *Economisch statistische berichten*, 75, 32-43.
- Kuipers, S.K. and A.H. van Zon - "Output and Employment Growth in The Netherlands in the Postwar Period: A Putty-Clay Approach," 1982, *De Economist*, 130, 38-70.
- Kurz, M. - "Equilibrium with Transaction Cost and Money in a Single Market Exchange Economy," 1974, *Journal of Economic Theory*, 7, 418-452.
- Kydland, F.W. and E.C. Prescott - "Rules rather than Discretion: The Inconsistency of

- Optimal Plans," 1977, *Journal of Political Economy*, 85, 473-491.
- Lachmann, L.M. - "The Role of Expectations in Economics as a Social Science," 1943, *Economica*, reprinted in *Capital, Expectations, and the Market Process*, 1977. Kansas City: Sheed, Andrews, and McMeel.
- Lachmann, L.M. - "Methodological Individualism and the Market Process," 1969, in: Streissler et al. (eds.) - *Roads to Freedom: Essays in Honor of Friedrich von Hayek*, London: Routledge and Kegan Paul.
- Lachmann, L.M. - *Capital, Expectations and the Market Process*, 1977, Kansas City: Sheed, Andrews and McMeel.
- Lambkin, M. - "Order of Entry and Performance in New Markets," 1988, *Strategic Management Journal*, 9, 127-140.
- Lancaster, K. - "A New Approach to Consumer Theory," 1966, *Journal of Political Economy*, 74, 132-157.
- Lancaster, K. - *Variety, Equity and Efficiency*, 1979, New York: Columbia University Press.
- Langlois, R.N. - "Rationality, Institutions, and Explanation," in: Langlois, R.N. (ed.) - *Economics as a Process: Essays in the New Institutional Economics*, 1986, Cambridge: Cambridge University Press.
- Larner, R. - "Public Policy in the Ocean Freight Industry," in: Phillips, A. (ed.) - *Promoting Competition in Regulated Markets*, 1975, Washington, D.C.: Brookings Institution.
- Lawson, T. - "Uncertainty and Economic Analysis," 1985, *Economic Journal*, 95, 909 - 927.
- Lea, S.E.G., R.M. Tardy and P. Webley - *The Individual in the Economy: A Survey of Economic Psychology*, 1987, Cambridge: Cambridge University Press.
- Lecraw, D.J. - "Diversification Strategy and Performance," 1984, *Journal of Industrial Economics*, 33, 179-198.
- Lee, T. and L. Wilde - "Market Structure and Innovation: A Reformulation," 1980, *Quarterly Journal of Economics*, 94, 429-436.
- Leibenstein, H. - "Allocative Efficiency versus X-Efficiency," 1966, *American Economic Review*, 56, 392-415.
- Leijonhufvud, A. - *On Keynesian Economics and the Economics of Keynes*, 1968, New York: Oxford University Press.
- Leontief, W.W. - *The Structure of the American Economy, 1919-1939*, 1941, New York: Oxford University Press.
- Lerner, A.P. - *The Economics of Control*, 1946, New York: McMillan.
- Lewin, K. - *Field Theory in Social Science: Selected Theoretical Papers*, 1951, New York: Harper and Row.
- Lieberman, M.B. - "Excess Capacity as a Barrier to Entry: An Empirical Appraisal," 1987, *Journal of Industrial Economics*, 35, 607-627.
- Liebhafsky, H.H. - *American Government and Business*, 1971, New York: John Wiley.
- Liebrand, W.B.G. - "Experimentele psychologie en het energievraagstuk," in: Ester, P. and F.L. Leeuw (eds.) - *Energie als maatschappelijk probleem*, 1981, Assen: Van Gorcum.
- Lier, A. van and A. van Witteloostuijn - "Size Asymmetries and (Non)chaotic Cournot Duopoly," 1990, forthcoming *Research Memorandum*, University of Limburg.
- Lindahl, E. - "Penningspolitiekens Model," 1929, reprinted in *Studies in the Theory of Money and Capital*, 1939, New York: A.M. Kelley.

- Lindbeck, A. and D.J. Snower - "Efficiency Wages *versus* Insiders and Outsiders," 1987, *European Economic Review*, 31, 407-416.
- Lindbeck, A. and D.J. Snower - *The Insider-Outsider Theory of Employment and Unemployment*, 1988, Cambridge, Mass.: MIT Press.
- Lindblom, C.E. - *The Intelligence of Democracy*, 1965, New York: McMillan.
- Lippman, S.A. and J.J. McCall. - "The Economics of Job Search: A Survey, Parts I and II," 1976, *Economic Inquiry*, 14, 155-189 and 347-368.
- Lippman, S.A. and R.P. Rumelt - "Uncertain Imitability: An Analysis of Interfirm Differences in Efficiency under Competition," 1982, *Bell Journal of Economics*, 13, 418-436.
- Lipsey, R.G. - "The Theory of Customs Unions: Trade Diversion and Welfare," 1957, *Economica*, 24, 40-46.
- Lipsey, R.G. - *The Theory of Customs Unions: A General Equilibrium Analysis*, 1970, London: Weidenfeld and Nicolson.
- Lipsey, R.G. and K. Lancaster - "The General Theory of Second-Best," 1956, *Review of Economic Studies*, 24, 11-32.
- Little, I.M.D. - *A Critique on Welfare Economics*, 1949, Oxford: Clarendon Press.
- Littlechild, S.C. and G. Owen - "An Austrian Model of the Entrepreneurial Market Process," 1980, *Journal of Economic Theory*, 23, 361-379.
- Loasby, B.J. - *Choice, Complexity and Ignorance*, 1976, Cambridge: Cambridge University Press.
- Loeb, P.D. - "Further Evidence of the Determinants of Industrial Research and Development Using Single and Simultaneous Equation Models," 1983, *Empirical Economics*, 8, 203-214.
- Loomes, G. and R. Sugden - "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty," 1982, *Economic Journal*, 92, 805-824.
- Lorie, H.R. - "Price-Quantity Adjustments in a Macro Disequilibrium Model," 1978, *Economic Inquiry*, 16, 265-287.
- Loury, G.G. - "Market Structure and Innovation," 1979, *Quarterly Journal of Economics*, 92, 395-410.
- Lucas, R.E. Jr. - "Tests of a Capital-Theoretic Model of Technological Change," 1967, *Review of Economic Studies*, 34, 175-180.
- Lucas, R.E. Jr. - "Expectations and the Neutrality of Money," 1972, *Journal of Economic Theory*, 4, 103-124.
- Lucas, R.E. Jr. - "An Equilibrium Model of the Business Cycle," 1975, *Journal of Political Economy*, 83, 1113-1144.
- Luce, R.D. and H. Raiffa - *Games and Decisions: Introduction and Critical Survey*, 1957, New York: John Wiley.
- Lundberg, E. - *Studies in the Theory of Economic Expansion*, 1937, New York: King.
- Lyons, B. - "Barriers to Entry," 1988, in: Davies, S. and B. Lyons (eds.) - *Economics of Industrial Organisation*, London: Longman.
- Machina, M.J. - "'Expected Utility' Analysis without the Independence Axiom," 1982, *Econometrica*, 50, 277-323.
- Machlup, F. - "Theories of the Firm: Marginalist, Behavioral, Managerial," 1967, *American Economic Review*, 57, 1-33.
- Magee, S.P. - "The Welfare Effects of Restrictions on U.S. Trade," 1972, *Brookings Papers on Economic Activity*, 3, 645-707.
- Maital, S. - *Minds, Markets, and Money: Psychological Foundations of Economic*

- Behavior*, 1982, New York: Basic Books.
- Maks, J.A.H. - "Consistency and Consumer Behaviour in the Netherlands, 1921-1962," 1978, *European Economic Review*, 11, 343-362.
- Maks, J.A.H. - "On the Compatibility of the Welfare State and the Market Economy," 1986, in: Albeda, W. (ed.) - *The Future of the Welfare State*, Maastricht: Presses Interuniversitaires Européennes.
- Malinvaud, E. - *The Theory of Unemployment Reconsidered*, 1977, Oxford: Basil Blackwell.
- Malinvaud, E. - *Profitability and Unemployment*, 1980, Cambridge: Cambridge University Press.
- Malthus, T.R. - *An Essay on Population*, 1798, London: Johnson.
- Malveaux, J. - "Theoretical Explanations of the Persistence of Racial Differentials," 1984, in: Darity, W. (ed.) - *Labor Economics: Modern Views*, Boston: Kluwer-Nijhoff.
- Mankiw, N.G. - "Small Menu Costs and Large Business Cycles: A Macroeconomic Model of Monopoly," 1985, *Quarterly Journal of Economics*, 100, 529-537.
- Mankiw, N.G. - "Real Business Cycles: A New Keynesian Perspective," 1989, *Journal of Economic Perspectives*, 3, 79-90.
- Mankiw, N.G. and M.D. Whinston - "Free Entry and Social Efficiency," 1986, *Rand Journal of Economics*, 17, 48-58.
- Mann, H.M. - "The New York Stock Exchange," in: Phillips, A. (ed.) - *Promoting Competition in Regulated Markets*, 1975, Washington, D.C.: Brookings Institution.
- Mansfield, E. - *Industrial Research and Technological Innovation*, 1968, New York: Norton.
- March, J.G. and H.A. Simon - *Organizations*, 1958, New York: John Wiley.
- Margolis, E. - "Monopolistic Competition and Multiproduct Brand Names," 1989, *Journal of Business*, 62, 199-209.
- Margolis, H. - *Selfishness, Altruism, and Rationality*, 1982, Cambridge: Cambridge University Press.
- Markowitz, H.M. - "The Utility of Wealth," 1952, *Journal of Political Economy*, 60, 151-158.
- Markowitz, H.M. - *Portfolio Selection*, 1959, New York: John Wiley.
- Marshall, A. - *Principles of Economics*, 8th edition 1920, London: McMillan.
- Martin, S. - "Sunk Costs, Financial Markets, and Contestability," 1989, *European Economic Review*, 33, 1089-1113.
- Marx, K. - *Das Kapital*, 1863, English translation, *Capital*, vol. III, 1981, New York: First Vintage Book Edition, first published in German in 1863-1867.
- Mas-Colell, A. - "Noncooperative Approaches to the Theory of Perfect Competition: Presentation," 1980, *Journal of Economic Theory*, 22, 121-135.
- Mas-Colell, A. - "Walrasian Equilibria as Limits of Noncooperative Equilibria. Part I: Mixed Strategies," 1983, *Journal of Economic Theory*, 30, 153-170.
- Mas-Colell, A. - "Notes on Price and Quantity Tatonnement Dynamics," in: Sonnenschein, H. (ed.) - *Models of Economic Dynamics*, 1986, Berlin: Springer-Verlag.
- Maskin, E. and J. Tirole - "A Theory of Dynamic Oligopoly, I: Overview and Quantity Competition with Large Fixed Costs," 1988, *Econometrica*, 56, 549-569.
- Mason, E.S. - "Price and Production Policies of Large-Scale Enterprise," 1939, *American Economic Review*, 29, 61-74.
- Mason, E.S. - "The Current State of the Monopoly Problem in the United States," 1949,

- Harvard Law Review*, 62, 1265-1285.
- Masson, R.T. and J. Shaanan - "Excess Capacity and Limit Pricing: An Empirical Test," 1986, *Economica*, 53, 365-378.
- Mayes, D.G. - *The Effects of Alternative Trade Groupings on the United Kingdom*, 1971, Ph.D. Thesis, University of Bristol.
- Mayes, D.G. - "The Effects of Economic Integration on Trade," 1978, *Journal of Common Market Studies*, 17, 1-25.
- McCallum, B.T. - "Rational Expectations and Macroeconomic Stabilization Policy," 1980, *Journal of Money, Credit, and Banking*, 12, 716-745.
- McCallum, B.T. - "The Development of Keynesian Macroeconomics," 1987, *American Economic Review: Papers and Proceedings*, 77, 125-129.
- McClintock, Ch.G. - "Game Behavior and Social Motivation in Interpersonal Settings," in: McClintock, Ch.G. (ed.) - *Experimental Social Psychology*, 1972, New York: Rinehart and Winston.
- McGee, J. and H. Thomas - "Strategic Groups: Theory, Research and Taxonomy, 1986, *Strategic Management Journal*, 7, 141-160.
- McKenzie, L. - "On the Existence of General Equilibrium for a Competitive Market," 1959, *Econometrica*, 27, 54-71.
- McLean, R.P. and M.H. Riordan - "Equilibrium Industry Structure with Sequential Technology Choice," 1990, forthcoming in the *Journal of Economic Theory*, 37.
- McLeod, W.B. - "Entry, Sunk Costs, and Market Structure," 1987, *Canadian Journal of Economics*, 20, 140-151.
- McNulty, P.J. - "Economic Theory and the Meaning of Competition," 1968, reprinted in Brozen, Y. (ed.) - *The Competitive Economy: Selected Readings*, 1975, Morristown, N.J.: General Learning Press.
- Meade, J.E. - *The Theory of Customs Unions*, 1955, Amsterdam: North-Holland.
- Menger, C. - *Grundsätze der Volkswirtschaftslehre*, 1871, English translation, *Principles of Economics*, 1950, Glencoe, Ill.: Free Press.
- Mierlo, J.G.A. and A. van Witteloostuijn - "De interactie tussen particuliere ondernemingen en publieke instellingen: een bedrijfstakonderzoek naar de markt voor ziektekostenverzekeringen in Nederland," 1990, forthcoming *Onderzoekmemorandum*, University of Limburg.
- Milgrom, P. and J. Roberts - "Predation, Reputation, and Entry Deterrence," 1982, *Journal of Economic Theory*, 27, 280-312.
- Milgrom, P. and J. Roberts - "Informational Asymmetries, Strategic Behavior, and Industrial Organization," 1987, *American Economic Review*, 77, 184-193.
- Mill, J.S. - *Principles of Political Economy with Some of their Applications to Social Philosophy*, 1878, 8th edition, London: Longmans, Green, Reader and Dyer, first published in 1848.
- Miller, R.A. - "Innovation and Reputation," 1988, *Journal of Political Economy*, 96, 741-765.
- Mills, D.E. - "Preemptive Investment Timing," 1988, *Rand Journal of Economics*, 19, 114-122.
- Mincer, J. - *Schooling, Experience and Earnings*, 1974, New York: NBER.
- Minsky, H.P. - *John Maynard Keynes*, 1975, New York: Columbia University Press.
- Minsky, H.P. - *Can 'It' Happen Again?*, 1982, Armonk, N.Y.: M.E. Sharpe.
- Mises, L. von - *Theorie des Geldes und der Umlaufsmittel*, 1912, English translation, *The Theory of Money and Credit*, 1953, New Haven, Conn.: Yale University Press.

- Mises, L. von - *Human Action: A Treatise on Economics*, 1949, New Haven, Conn.: Yale University Press.
- Mishan, E.J. - *Economic Efficiency and Social Welfare*, 1981a, London: George Allen and Unwin.
- Mishan, E.J. - *Introduction to Normative Economics*, 1981b, Oxford: Oxford University Press.
- Mishan, E.J. - "The Implications of Alternative Foundations to Welfare Economics," 1984, *De Economist*, 132, 75-85.
- Mitrani, I. - *Simulation Techniques for Discrete Event Systems*, 1982, Cambridge: Cambridge University Press.
- Modigliani, F. - "Liquidity Preference and the Theory of Interest and Money," 1944, *Econometrica*, 18, 45-88.
- Modigliani, F. - "New Developments on the Oligopoly Front," 1958, *Journal of Political Economy*, 35, 215-232.
- Moore, T.G. - "Deregulating Surface Freight Transportation," in: Phillips, A. (ed.) - *Promoting Competition in Regulated Markets*, 1975, Washington, D.C.: Brookings Institution.
- Morgenstern, O. - "Vollkommene Voraussicht und wirtschaftliches Gleichgewicht," 1935, *Zeitschrift für Nationalökonomie*, 6, 337-357, English translation in Schotter, A. (ed.) - *Selected Writings of Oskar Morgenstern*, 1976, New York: New York University Press.
- Morgenstern, O. and J. von Neumann - *Theory of Games and Economic Behaviour*, 1944, Princeton, N.J.: Princeton University Press.
- Morishima, M. - *Walras' Economics*, 1977, Cambridge: Cambridge University Press.
- Muellbauer, J. en R. Portes - "Macroeconomic Models with Quantity Rationing," 1978, *Economic Journal*, 88, 787-821.
- Mueller, D.C. - *Public Choice*, 1979, Cambridge: Cambridge University Press.
- Mueller, D.C. - *The Corporation Growth, Diversification and Mergers*, 1987, London: Harwood.
- Mueller, D.C. (ed.) - "Special Issue on 'Mergers'," 1989, *International Journal of Industrial Organization*, 7, 1-174.
- Murphy, G. and J.K. Kovach - *Historical Introduction to Modern Psychology*, 1972, London: Routledge and Kegan Paul.
- Muth, J.F. - "Rational Expectations and the Theory of Price Movements," 1961, *Econometrica*, 129, 315-335.
- Muysken, J. and A.H. van Zon - "Employment and Unemployment in the Netherlands, 1960-1985: A Putty-Clay Approach," 1987, *Recherches économiques de Louvain*, 53, 101-133.
- Myrdal, G. - *Monetary Equilibrium*, 1939, London: Hodge, first published in Swedish 1933.
- Nakao, T. - "Cost-Reducing R&D in Oligopoly," 1989, *Journal of Economic Behavior and Organization*, 12, 131-148.
- Nash, J.F. - "Non-Cooperative Games," 1951, *Annals of Mathematics*, 54, 286-295.
- Nathan, A. and E.H. Neave - "Competition and Contestability in Canada's Financial System: Empirical Results," 1989, *Canadian Journal of Economics*, 22, 576-594.
- Neary, J.P. and J.E. Stiglitz - "Toward a Reconstruction of Keynesian Economics: Expectations and Constrained Equilibria," 1983, *Quarterly Journal of Economics*, 98, 199-228.

- Negishi, T - "Monopolistic Competition and General Equilibrium," 1960, *Review of Economic Studies*, 28, 196-201.
- Negishi, T. - *Microeconomic Foundations of Keynesian Macroeconomics*, 1979, Amsterdam: North-Holland.
- Nelson, P. - "The Economic Consequences of Advertising," 1975, *Journal of Business*, 48, 213-241.
- Nelson, R.R. - "The Simple Economics of Basic Scientific Research," 1959, *Journal of Political Economy*, 67, 297-306.
- Nelson, R.R. - "Research on Productivity Growth and Productivity Difference: Dead Ends and New Departures," 1981, *Journal of Economic Literature*, 19, 1029-1064.
- Nelson, R.R. - "Evolutionary Modelling of Economic Change," in: Stiglitz, J.E. and G.F. Mathewson (eds.) - *New Developments in the Analysis of Market Structure*, 1986, Cambridge, Mass.: MIT Press.
- Nelson, R.R. and S.G. Winter - "Neo-Classical vs. Evolutionary Theories of Economic Growth: Critique and Prospectus," 1974, *Economic Journal*, 84, 886-905.
- Nelson, R.R. and S.G. Winter - "Simulation of Schumpeterian Competition," 1977a, *American Economic Review*, 67, 271-276.
- Nelson, R.R. and S.G. Winter - "Dynamic Competition and Technical Progress," in: Balassa, B. and R.R. Nelson (eds.) - *Economic Progress, Private Values and Public Policy*, 1977b, Amsterdam: North-Holland.
- Nelson, R.R. and S.G. Winter - "Forces Generating and Limiting Concentration under Schumpeterian Competition," 1978, *Bell Journal of Economics*, 9, 524-548.
- Nelson, R.R. and S.G. Winter - "Firm and Industry Response to Changed Market Conditions: An Evolutionary Approach," 1980, *Economic Enquiry*, 18, 179-202.
- Nelson, R.R. and S.G. Winter - *An Evolutionary Theory of Economic Change*, 1982, Cambridge, Mass.: Harvard University Press.
- Nelson, R.R., S.G. Winter and H.L. Scheutte - "Technical Change in an Evolutionary Model," 1976, *Quarterly Journal of Economics*, 90, 90-118.
- Neuman, S. - "Religious Observance within a Human Capital Framework: Theory and Application," 1986, *Applied Economics*, 18, 1193-1202.
- Neumann, J. von - "Über ein Okonomisches Gleichungssystem und eine Verallgemeinerung des Brouwerischen Fixpunkt-Satzes," 1937, *Ergebnisse eines Mathematischen Kolloquiums*, 8, 73-83.
- Ng, Y.K. - *Welfare Economics*, 1979, London: McMillan.
- Nickell, S.J. - *The Investment Decision of Firms*, 1978, Cambridge: Cambridge University Press.
- Nikaido, H. - "Refutation of the Dynamic Equalization of Profit Rates in Marx's Scheme of Reproduction," 1977, *Modelling Research Group 7722*, University of Southern California.
- Nikaido, H. - "Marx on Competition," 1983, *Journal of Economics*, 43, 337-362.
- Niskanen, W.A. - *Bureaucracy and Representative Government*, 1971, Chicago: Aldine-Atherton.
- Nordhaus, W.D. - "The Optimal Rate and Direction of Technical Change," in: Shell, K. (ed.) - *Essays on the Theory of Optimal Economic Growth*, 1967, Cambridge, Mass.: MIT Press.
- Novshek, W. - "Cournot Equilibrium with Free Entry," 1980, *Review of Economic Studies*, 47, 473-486.
- Novshek, W. and H. Sonnenschein - "Cournot and Walras Equilibrium," 1978, *Journal*

- of *Economic Theory*, 22, 243-255.
- Novshek, W. and H. Sonnenschein - "Walrasian Equilibria as Limits of Noncooperative Equilibria. Part II: Pure Strategies," 1983, *Journal of Economic Theory*, 30, 171-187.
- Novshek, W. and H. Sonnenschein - "General Equilibrium with Free Entry: A Synthetic Approach to the Theory of Perfect Competition," 1987, *Journal of Economic Literature*, 25, 1281-1306.
- Nti, K.O. - "More Potential Entrants May Lead to Less Competition," 1989, *Journal of Economics*, 49, 47-70.
- Nti, K.O. and M. Shubik - "Noncooperative Oligopoly with Entry," 1981, *Journal of Economic Theory*, 24, 187-204.
- O'Connell, J.F. - "The Labor Market for Engineers: An Alternative Methodology," 1972, *Journal of Human Resources*, 7, 71-86.
- O'Driscoll, G.P. Jr. and M.J. Rizzo - *The Economics of Time and Ignorance*, 1985, Oxford: Basil Blackwell.
- Ohlin, B. - "Some Notes on the Stockholm Theory of Saving and Investment I-II," 1937, *Economic Journal*, 47, 53-69 and 221-240.
- Okun, A.M. - "Inflation: Its Mechanics and Welfare Costs," 1975, *Brooking Papers on Economic Activity*, 6, 351-390.
- Okun, A.M. - *Prices and Quantities*, 1981, Washington, D.C.: Brookings Institution.
- Olsen, M. - *The Logic of Collective Action: Public Goods and the Theory of Groups*, 1965, Cambridge, Mass.: Harvard University Press.
- Ordover, J.A. and G. Saloner - "Predation, Monopolization, and Antitrust," in: Schmalensee, R. and R.D. Willig (eds.) - *Handbook of Industrial Organization*, 1989, Amsterdam: North-Holland.
- Osborne, M.J. and C. Pitchik - "Price Competition in a Capacity-Constrained Duopoly," 1986, *Journal of Economic Theory*, 38, 238-260.
- Ostroy, J.M. - "The Informational Efficiency of Monetary Exchange," 1973, *American Economic Review*, 63, 597-610.
- Ostroy, J.M. and R.M. Starr - "Money and the Decentralization of Exchange," 1974, *Econometrica*, 42, 1093-1114.
- Panzar, J.C. - "Technological Determinants of Firm and Industry Structure," in: Schmalensee, R. and R.D. Willig (eds.) - *Handbook of Industrial Organization*, 1989, Amsterdam: North-Holland.
- Panzar, J.C. and R.D. Willig - "Free Entry and the Sustainability of Natural Monopoly," 1977, *Bell Journal of Economics*, 8, 1-22.
- Pareto, V. - *Manuel d'économie politique*, 1909, Paris: Giraud and Brière, English translation, *Manual of Political Economy*, 1971, London: McMillan.
- Parks, R.P. - "An Impossibility Theorem for Fixed Preferences: A Dictorial Bergson-Samuelson Welfare Function," 1976, *Review of Economic Studies*, 43, 447-450.
- Pashigan, B.P. - "Limit Price and the Market Share of the Leading Firm," 1968, *Journal of Industrial Economics*, 16, 165-177.
- Pasinetti, L. - "Profit and Growth," 1970, in: Sen, A.K. (ed.) - *Growth Economics*, Harmondsworth: Penguin Books.
- Pasinetti, L. - *Structural Change and Economic Growth*, 1983, Cambridge: Cambridge University Press.
- Patinkin, D. - *Money, Interest and Prices*, 1956, Evanston, Ill.: Row, Peterson.
- Pazner, E.A. and D. Schmeidler - "A Difficulty in the Concept of Fairness," 1974, *Review of Economic Studies*, 41, 441-443.

- Pazner, E.A. and D. Schmeidler - "Egalitarian-Equivalent Allocations: A New Concept of Economic Equity," 1978, *Quarterly Journal of Economics*, 92, 1-45.
- Pelkmans, J. - "De interne EG-markt voor industriële producten," 1985, *WRR* v 47, Den Haag: Staatsuitgeverij.
- Perrakis, S. and G. Warskett - "Capacity and Entry under Demand Uncertainty," 1983, *Review of Economic Studies*, 50, 495-511.
- Perrakis, S. and G. Warskett - "Uncertainty, Economies of Scale, and Barrier to Entry," 1986, *Oxford Economic Papers*, 38, 58-74.
- Perry, M. - "Sustainable Positive Profit Multiple-Price Strategies in Contestable Markets," 1984, *Journal of Economic Theory*, 32, 246-265.
- Peters, H.J.M. - *Bargaining Game Theory*, 1986, Meppel: Krips Repro.
- Phelps, E.S. (ed.) - *Microeconomic Foundations of Employment and Inflation Theory*, 1970, New York: Norton.
- Phillips, A. - "Patents, Potential Competition, and Technical Progress," 1966, *American Economic Review*, 56, 301-310.
- Phillips, A. - "Competitive Policy for Depository Financial Institutions," in: Phillips, A. (ed.) - *Promoting Competition in Regulated Markets*, 1975, Washington, D.C.: Brookings Institution.
- Pigou, A.C. - *Wealth and Welfare*, 1912, London: McMillan.
- Pinto, B. - "Repeated Games and the 'Reciprocal Dumping' Model of Trade," 1986, *Journal of International Economics*, 20, 357-366.
- Pissarides, C.A. - *Equilibrium Unemployment Theory*, 1988, Oxford: Basil Blackwell.
- Plosser, C.I. - "Understanding Real Business Cycles," 1989, *Journal of Economic Perspectives*, 3, 51-78.
- Poel, J.H.R. van de - *Judgement and Control: Individual and Organizational Aspects of Performance Evaluation*, 1986, Groningen: Wolters-Noordhoff.
- Poitevin, M. - "Financial Signalling and the 'Deep-Pocket' Argument," 1989, *Rand Journal of Economics*, 20, 26-40.
- Pollak, R.A. - "Endogenous Tastes in Demand and Welfare Analysis," 1978, *American Economic Review*, 68, 374-377.
- Porter, M.E. - "Please Note Location of Nearest Exit," 1976, *California Management Review*, 13, 21-33.
- Porter, M.E. - *Competitive Strategy: Techniques for Analyzing Industries and Competitors*, 1980, New York: Free Press.
- Porter, M.E. (ed.) - *Competition in Global Industries*, 1986, Boston: Harvard Business School Press.
- Posner, R.A. - "The Social Cost of Monopoly and Regulation," 1975, *Journal of Political Economy*, 83, 807-827.
- Posner, R.A. - *The Economic Analysis of Law*, 1977, Boston: Little, Brown.
- Posner, R.A. "Theories of Economic Regulation," in: Ogus, A.I. and C.G. Veljanovski (eds.) - *Readings in the Economics of Law and Regulation*, 1984, Oxford: Clarendon Press.
- Prescott, E.C. and M. Visscher - "Sequential Location among Firms with Foresight," 1977, *Bell Journal of Economics*, 8, 378-393.
- Primeaux, W.J. Jr. - "A Reexamination of the Monopoly Market Structure for Electric Utilities," in: Phillips, A. (ed.) - *Promoting Competition in Regulated Markets*, 1975, Washington, D.C.: Brookings Institution.
- Pyatt, G. - "Profit Maximisation and the Threat of New Entry," 1971, *Economic*

- Journal*, 81, 242-255.
- Quirk, J. and D. McDougall - *Microeconomics*, 1981, Chicago: University of Chicago Press.
- Raa, T. ten - *A Theory of Value and Industry Structure*, 1980, Ph.D. Thesis, New York University.
- Rachlin, H. - "Economics and Behavioral Psychology," in: Staddon, J.E.R. (ed.) - *Limits to Action*, 1980, New York: Academic Press.
- Radner, R. - "Competitive Equilibrium under Uncertainty," 1968, *Econometrica*, 36, 31-58.
- Radner, R. - "Market Equilibrium under Uncertainty: Concepts and Problems," 1974, in: Intriligator, M.D. and P.A. Kendrick (eds.) - *Frontier of Quantitative Economics Vol. II*, Amsterdam: North-Holland.
- Radner, R. - "A Behavioral Model of Cost Reduction," 1975, *Bell Journal of Economics*, 6, 196-215.
- Radner, R. - "Collusive Behavior in Noncooperative Epsilon-Equilibria of Oligopolies with Long but Finite Lives," 1980, *Journal of Economic Theory*, 22, 136-154.
- Radner, R. - "Monitoring Cooperative Agreements in a Repeated Principal Agent Relationship," 1981, *Econometrica*, 49, 1127-1148.
- Radner, R. - "Equilibrium under Uncertainty," 1982, in: Arrow, K.J. and M.D. Intriligator (eds.) - *Handbook of Mathematical Economics Vol. II*, Amsterdam: North-Holland.
- Radner, R. - "Repeating Partnership Games with Imperfect Monitoring and No Discounting," 1986, *Review of Economic Studies*, 53, 43-58.
- Raiffa, H. - *Decision Analysis: Introduction Lectures on Choices under Uncertainty*, 1969, Reading, Mass.: Addison Wesley.
- Ramsey, F.P. - "A Contribution to the Theory of Taxation," 1927, *Economic Journal*, 37, 47-61.
- Rand, D. - "Exotic Phenomena in Games and Duopoly Models," 1978, *Journal of Mathematical Economics*, 5, 173-184.
- Rashid, S. - "Quality in Contestable Markets: A Historical Problem," 1988, *Quarterly Journal of Economics*, 103, 245-249.
- Raubitschek, R.S. - "A Model of Product Proliferation with Multiproduct Firms," 1987, *Journal of Industrial Economics*, 35, 269-279.
- Reid, G.C. - *Theories of Industrial Organization*, 1987, Oxford: Basil Blackwell.
- Reinganum, J.F. - *Dynamic Games of R&D with Rivalry*, 1979, Ph.D. Thesis, Northwestern University.
- Reinganum, J.F. - "Dynamic Games of Innovation," 1981, *Journal of Economic Theory*, 25, 21-41.
- Reinganum, J.F. - "Uncertain Innovation and the Persistence of Monopoly," 1983, *American Economic Review*, 73, 741-748.
- Reinganum, J.F. - "The Timing of Innovation: Research, Development, and Diffusion," in: Schmalensee, R. and R.D. Willig (eds.) - *Handbook of Industrial Organization*, 1989, Amsterdam: North-Holland.
- Resnick, S.A. and E.M. Truman - "The Distribution of West European Trade under Alternative Tariff Policies," 1974, *Review of Economics and Statistics*, 56, 83-91.
- Reynolds, L.G. - *Microeconomics: Analysis and Policy*, 1985, Homewood, Ill.: Irwin.
- Ricardo, D. - *The Principles of Political Economy and Taxation*, 1960, London: Dent, first published in 1817.

- Riordan, M.H. - "Imperfect Information and Dynamic Conjectural Variations," 1985, *Rand Journal of Economics*, 16, 41-50.
- Rizzutto, R. and P. Wachtel - "Further Evidence on the Returns to School Quality," 1980, *Journal of Economic Resources*, 15, 240-254.
- Robbins, L. - *An Essay on the Nature and Significance of Economic Science*, 1932, 2nd edition 1935, London: McMillan.
- Roberts, J. - "A Signaling Model of Predatory Pricing," 1986, *Oxford Economic Papers*, 38, 75-93.
- Roberts, J. - "Battles for Market Share: Incomplete Information, Aggressive Strategic Pricing, and Competitive Dynamics," in: Bewley, T. (ed.) - *Advances in Economic Theory*, 1987, Cambridge: Cambridge University Press.
- Robinson, J. - *Economics of Imperfect Competition*, 1933, London: McMillan.
- Rogerson, W. - "Repeated Moral Hazard," 1985, *Econometrica*, 53, 69-76.
- Rosen, S. - "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," 1974, *Journal of Political Economy*, 82, 34-55.
- Roth, A.E. - *Axiomatic Models of Bargaining*, 1979, Cambridge: Cambridge University Press.
- Rothbard, M.N. - *Man, Economy, and State*, 1962 (two volumes), Los Angeles: Nash.
- Rothbard, M.N. - *America's Great Depression*, 1963, Princeton, N.J.: D. van Nostrand.
- Rothbard, M.N. - *Power and Market: Government and the Economy*, 1977, Kansas City: Sheed, Andrews and McNeel.
- Rothschild, M. - "Models of Market Organization with Imperfect Information: A Survey," 1973, *Journal of Political Economy*, 81, 1283-1308.
- Rothwell, R. and W. Zegveld - *Industrial Innovation and Public Policy: Preparing for the 1980s and the 1990s*, 1982, London: Pinter.
- Rotter, J.B. - *Social Learning and Clinical Psychology*, 1954, Englewood Cliffs, N.J.: Prentice-Hall.
- Rumelt, R.P. - *Strategy, Structure, and Economic Performance*, 1974, Boston: Harvard Business School Press.
- Sah, R.K. and J.E. Stiglitz - "Human Fallibility and Economic Organization," 1985, *American Economic Review*, 75, 292-297.
- Sah, R.K. and J.E. Stiglitz - "The Architecture of Economic Systems: Hierarchies and Polyarchies," 1986, *American Economic Review*, 76, 716-727.
- Sah, R.K. and J.E. Stiglitz - "The Invariance of Market Innovation to the Number of Firms," 1987, *Rand Journal of Economics*, 18, 98-108.
- Sah, R.K. and J.E. Stiglitz - "Committees, Hierarchies and Polyarchies," 1988, *Economic Journal*, 98, 451-470.
- Salant, S. - "Preemptive Patenting and the Persistence of Monopoly: Comment," 1984, *American Economic Review*, 74, 247-250.
- Saloner, G. - "Predation, Mergers, and Incomplete Information," 1987, *Rand Journal of Economics*, 18, 165-186.
- Salop, S.C. - "Strategic Entry Deterrence," 1979, *American Economic Review*, 69, 335-338.
- Salop, S. and D.T. Scheffman - "Raising Rivals' Costs," 1983, *American Economic Review: Papers and Proceedings*, 73, 267-271.
- Salop, S.C., D.T. Scheffman and W. Schwartz - "A Bidding Analysis of Special Interest Regulation: Raising Rivals' Costs in a Rent Seeking Society," in: Rogowsky, R. and B. Yandle (eds.) - *The Political Economy of Regulation: Private Interest in the Regulatory*

- Process*, 1984, Washington, D.C.: Federal Trade Commission.
- Salter, W.E.G. - *Productivity and Technical Change*, 1960, Cambridge: Cambridge University Press.
- Samuelson, P.A. - "The Stability of Equilibrium: Comparative Statics and Dynamics," 1941, *Econometrica*, 9, 97-120.
- Samuelson, P.A. - "Consumption Theory in terms of Revealed Preference," 1948, *Economica*, 15, 113-129.
- Samuelson, P.A. - "Evaluation of Real Income," 1950, *Oxford Economic Papers*, 2, 1-29.
- Samuelson, P.A. - "A Theory of Induced Innovation along Kennedy-Weizsäcker Lines," 1965, *Review of Economics and Statistics*, 47, 343-356.
- Samuelson, P.A. - "St. Petersburg Paradoxes: Defanged, Dissected, and Historically Described," 1977, *Journal of Economic Literature*, 15, 24-55.
- Sandler, T. - "Public Good and the Theory of the Second-Best," 1978, *Public Finance*, 33, 331-334.
- Sargent, T.J. and N. Wallace - "Rational Expectations, the Optimal Monetary Instrument and the Optimal Money Supply Rule," 1975, *Journal of Political Economy*, 83, 241-254.
- Sargent, T.J. and N. Wallace - "Rational Expectations and the Theory of Economic Policy," 1976, *Journal of Monetary Economics*, 2, 169-184.
- Sato, R. and G.S. Suzawa - *Research and Productivity: Endogenous Technical Change*, 1983, Boston: Auburn House.
- Savage, L.J. - *The Foundation of Statistics*, 2nd edition, 1972, New York: John Wiley.
- Say, J.B. - *Traité d'économie politique*, 6th edition, 1827, Brussels: C.J. de Mat fils et H. Remy, first published in 1803.
- Schaffer, M.E. - "Are Profit-Maximisers the Best Survivors? A Darwinian Model of Economic Natural Selection," 1989, *Journal of Economic Behavior and Organization*, 12, 29-45.
- Schelling, T.C. - "An Essay on Bargaining," 1956, *American Economic Review*, 281-306.
- Schelling, T.C. - *The Strategy of Conflict*, 1960, Cambridge, Mass.: Harvard University Press.
- Scherer, F.M. - "Time-Cost Trade-Offs in Uncertain Empirical Research Projects," 1966, *Naval Research Logistics Quarterly*, 13, 71-82, reprinted in Scherer, F.M. - *Innovation and Growth: Schumpeterian Perspectives*, 1984, Cambridge, Mass.: MIT Press.
- Scherer, F.M. - "Research and Development Resource Allocation under Rivalry," 1967, *Quarterly Journal of Economics*, 81, 359-394, reprinted in Scherer, F.M. - *Innovation and Growth: Schumpeterian Perspectives*, 1984, Cambridge, Mass.: MIT Press.
- Scherer, F.M. - "The Microeconomics of Decision Making for Innovation," in: Reynolds, L. (ed.) - *The Economics of Technological Progress*, 1976, reprinted in Scherer, F.M. - *Innovation and Growth: Schumpeterian Perspectives*, 1984, Cambridge, Mass.: MIT Press.
- Scherer, F.M. - "The Welfare Economics of Product Variety: An Application to the Ready-to-Eat Cereals Industry," 1979, *Journal of Industrial Economics*, 28, 113-134, reprinted in Scherer, F.M. - *Innovation and Growth: Schumpeterian Perspectives*, 1984, Cambridge, Mass.: MIT Press.
- Scherer, F.M. - *Industrial Market Structure and Economic Performance*, 1980, 2nd

- edition, Chicago: Rand McNally.
- Scherer, F.M. - *Innovation and Growth: Schumpeterian Perspectives*, 1984, Cambridge, Mass.: MIT Press.
- Schmalensee, R. - "Brand Loyalty and Barriers to Entry," 1974, *Southern Economic Journal*, 40, 579-588.
- Schmalensee, R. - "A Model of Advertising and Product Quality," 1978a, *Journal of Political Economy*, 86, 485-503.
- Schmalensee, R. - "Entry Deterrence in the Ready-to-Eat Breakfast Cereal Industry," 1978b, *Bell Journal of Economics*, 9, 305-327.
- Schmalensee, R. - "Economies of Scale and Barriers to Entry," 1981, *Journal of Political Economy*, 89, 122-138.
- Schmalensee, R. - "Product Differentiation Advantages of Pioneering Brands," 1982, *American Economic Review*, 72, 349-365.
- Schmalensee, R. - "Advertising and Market Structure," in: Stiglitz, J.E. and F. Matthewson (eds.) - *New Developments in the Analysis of Market Structure*, 1986, Cambridge, Mass.: MIT Press.
- Schmalensee, R. - "Industrial Economics: An Overview," 1988, *Economic Journal*, 98, 643-681.
- Schmalensee, R. - "Inter-Industry Studies of Structure and Performance," in: Schmalensee, R. and R.D. Willig (eds.) - *Handbook of Industrial Organization*, 1989, Amsterdam: North-Holland.
- Schmalensee, R. and R.D. Willig (eds.) - *Handbook of Industrial Organization*, 1989, Amsterdam: North-Holland.
- Schmeidler, D. - "The Nucleolus of a Characteristic Function Form Game," 1969, *SIAM Journal of Applied Mathematics*, 17, 1163-1170.
- Schmeidler, D. and K. Vind - "Fair Net Trades," 1972, *Econometrica*, 40, 637-642.
- Schmookler, J. - *Invention and Economic Growth*, 1966, Cambridge, Mass.: Harvard University Press.
- Schoemaker, P.J.H. - "The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations," 1982, *Journal of Economic Literature*, 20, 529-563.
- Schotter, A. - *The Economic Theory of Social Institutions*, 1981, Cambridge: Cambridge University Press.
- Schotter, A. and G. Schwödiauer - "Economics and the Theory of Games: A Survey," 1980, *Journal of Economic Literature*, 18, 479-527.
- Schultz, T.W. - "Investment in Human Capital," 1961, *American Economic Review*, 51, 1-17.
- Schultz, T.W. - *The Economic Value of Education*, 1963, New York: Columbia University Press.
- Schumpeter, J.A. - *Capitalism, Socialism and Democracy*, 1943, London: George Allen and Unwin.
- Schumpeter, J.A. - *History of Economic Analysis*, 1954, New York: Oxford University Press.
- Schwartz, M. - "The Nature and Scope of Contestability Theory," 1986, *Oxford Economic Papers*, 38, 37-57.
- Schwartz, M. and R.J. Reynolds - "Contestable Markets: An Uprising in the Theory of Industry Structure: Comment," 1983, *American Economic Review*, 73, 488-490.
- Schwartz, M. and E.A. Thompson - "Divisionalization and Entry Deterrence," 1986, *Quarterly Journal of Economics*, 101, 305-321.

- Scitovsky, T. - "A Note on Welfare Propositions in Economics," 1941, *Review of Economic Studies*, 9, 77-88.
- Scitovsky, T. - *Economic Theory and Western European Integration*, 1958, Stanford, Cal.: Stanford University Press.
- Scott, J.T. - "Multimarket Contact and Economic Performance," 1982, *Review of Economics and Statistics*, 64, 368-375.
- Selten, R. - "Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit," 1965, *Zeitschrift für die gesamte Staatswissenschaft*, 12, 301-324.
- Selten, R. - "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," 1975, *International Journal of Game Theory*, 4, 25-55.
- Selten, R. - "The Chain-Store Paradox," 1978, *Theory and Decision*, 9, 127-159.
- Selten, R. - "Equity and Coalition Bargaining in Experimental Three-Person Games," in: Roth, A.E. (ed.) - *Laboratory Experimentation in Economics: Six Points of View*, 1987, Cambridge: Cambridge University Press.
- Semmler, W. - "On the Classical Theory of Competition, Value and Prices of Production," 1984, *Australian Economic Papers*, 130-150.
- Semmler, W. (ed.) - *Competition, Instability, and Nonlinear Cycles*, 1986, Berlin: Springer-Verlag.
- Sen, A.K. - *Choice, Welfare and Measurement*, 1982, Oxford: Basil Blackwell.
- Sengupta, J.K., J.E. Leonard and J.P. Vanyo - "A Limit Pricing Model for U.S. Computer Industry: An Application," 1983, *Applied Economics*, 15, 297-308.
- Senior, N.W. - *Outline of the Science of Political Economy*, 1931, London: Library of Economics Reprint, first published in 1836.
- Shackle, G.L.S. - *A Scheme of Economic Theory*, 1968, Cambridge: Cambridge University Press.
- Shackle, G.L.S. - *Decision, Order and Time*, 1969, Cambridge: Cambridge University Press.
- Shackle, G.L.S. - *Epistemics and Economics: A Critique of Economic Doctrines*, 1972, Cambridge: Cambridge University Press.
- Shackle, G.L.S. - *Keynesian Kaleidics*, 1974, Edinburgh: Edinburgh University Press.
- Shaked, A. and J. Sutton - "Relaxing Price Competition through Product Differentiation," 1982, *Review of Economic Studies*, 49, 3-13.
- Shaked, A. and J. Sutton - "Natural Oligopolies," 1983, *Econometrica*, 51, 1469-1483.
- Shand, A.H. - *The Capitalist Alternative: An Introduction to Neo-Austrian Economics*, 1984, Brighton: Harvester Press.
- Shapiro, C. - "Consumer Information, Product Quality, and Seller Reputation," 1982, *Bell Journal of Economics*, 13, 20-35.
- Shapiro, C. - "Premiums for High Quality Products as Returns to Reputation," 1983, *Quarterly Journal of Economics*, 98, 659-680.
- Shapiro, C. - "Exchange of Cost Information in Oligopoly," 1986, *Review of Economic Studies*, 53, 433-446.
- Shapiro, C. - "Theories of Oligopoly Behavior," in: Schmalensee, R. and R.D. Willig (eds.) - *Handbook of Industrial Organization*, 1989a, Amsterdam: North-Holland.
- Shapiro, C. - "The Theory of Business Strategy," 1989b, *Rand Journal of Economics*, 20, 125-137.
- Shaw, R. and S. Shaw - "Patent Expiry and Competition in Polyester Fibres," 1977, *Scottish Journal of Political Economy*, 24, 117-132.
- Sheffrin, S.M. - *Rational Expectations*, 1983, Cambridge: Cambridge University Press.

- Shepherd, W.G. - *The Economics of Industrial Organization*, 1979, London: Prentice-Hall.
- Shepherd, W.G. - "Contestability vs. Competition," 1984, *American Economic Review*, 74, 572-587.
- Sherman, R. and T.D. Willett - "Potential Entrants Discourage Entry," 1967, *Journal of Political Economy*, 75, 400-403.
- Shubik, M. - *Game Theory in the Social Sciences, Vol. II*, 1984, Cambridge, Mass.: MIT Press.
- Siebert, W.S. - "Development in the Economics of Human Capital," 1985, in: Carline *et al.* - *Labour Economics*, New York: Longman.
- Silberberg, E. - "The Theory of the Firm in 'Long-Run' Equilibrium," 1974, *American Economic Review*, 64, 734-741.
- Simon, H.A. - *Administrative Behavior: A Study of Decision Making Process in Administrative Organisation*, 1957, first published in 1945, London: McMillan.
- Simon, H.A. - "Rationality in Psychology and Economics," 1986, *Journal of Business*, 59, S209-S224.
- Siven, C.H. - "The End of the Stockholm School," 1985, *Scandinavian Journal of Economics*, 87, 577-593.
- Sjaastad, L.A. - "The Costs and Returns of Human Migration," 1962, *Journal of Political Economy*, 70, 80-93.
- Skinner, B.F. - *The Behavior of Organisms*, 1938, New York: Appleton Century Crofts.
- Smallwood, D.E. - "Competition, Regulation, and Product Quality in the Automobile Insurance Industry," in: Phillips, A. (ed.) - *Promoting Competition in Regulated Markets*, 1975, Washington, D.C.: Brookings Institution.
- Smiley, R. - "Empirical Evidence on Strategic Entry Deterrence," 1988, *International Journal of Industrial Organization*, 6, 167-180.
- Smith, A. - *The Theory of Moral Sentiments*, 1966, New York: John Wiley, first published in 1759.
- Smith, A. - *The Wealth of Nations*, 1974, Harmondsworth: Penguin, first published in 1776.
- Smith II, R.L. - "Efficiency Gains from Strategic Investment," 1981, *Journal of Industrial Economics*, 30, 1-23.
- Snippe, J. - *Macroeconomic Adjustment Processes: A Sequential Analytical Approach*, 1985, Ph.D. Thesis, University of Groningen.
- Soete, L.L.G. - "Size of Firm, Oligopoly and Research: A Reappraisal," 1977, *Extrait de réseaux*, 35-36.
- Soete, L.L.G. - "Firm Size and Inventive Activity, 1979, *European Economic Review*, 12, 319-340.
- Solow, R.M. - "Technical Change and the Aggregate Production Function," 1957, *Review of Economics and Statistics*, 39, 312-320.
- Solow, R. - "Insiders and Outsiders in Wage Determination," 1985, *Scandinavian Journal of Economics*, 87, 441-428.
- Solow, R.M. and J.E. Stiglitz - "Output, Employment and Wages in the Short Run," 1968, *Quarterly Journal of Economics*, 82, 537-560.
- Sondermann, D. - "Temporary Competitive Equilibrium under the Conditions of Uncertainty," in: Drèze (ed.) - *Allocation under Uncertainty: Equilibrium and Optimality*, 1974, New York: McMillan.
- Sosnick, S.H. - "A Critique of Concepts of Workable Competition," 1958, *Quarterly*

- Journal of Economics*, 72, 380-423.
- Spence, A.M. - "Product Differentiation and Welfare," 1976, *American Economic Review*, 66, 407-414.
- Spence, A.M. - "Entry, Capacity, Investment and Oligopolistic Pricing," 1977, *Bell Journal of Economics*, 8, 534-544.
- Spence, A.M. - "Investment Strategy and Growth in a New Market," 1979, *Bell Journal of Economics*, 10, 1-19.
- Spence, A.M. - "Notes on Advertising, Economies of Scale and Entry Barriers," 1980, *Quarterly Journal of Economics*, 94, 493-504.
- Spence, A.M. - "The Learning Curve and Competition," 1981, *Bell Journal of Economics*, 12, 49-70.
- Spence, A.M. - "Contestable Markets and the Theory of Industry Structure," 1983, *Journal of Economic Literature*, 21, 981-990.
- Spence, A.M. - "Cost Reduction, Competition and Industry Performance," 1986, in: Stiglitz, J.E. and G.F. Mathewson (eds.) - *New Developments in the Analysis of Market Structure*, Cambridge, Mass.: MIT Press.
- Spence, A.M. - "Cost Reduction, Competition and Industry Performance," 1984, *Econometrica*, 52, 101-122.
- Spence, R.W. - "Channels of Monetary Influence: A Survey," 1974, *Federal Reserve Bank of St. Louis Review*, 56, 8-26.
- Spulber, D.F. - "Capacity, Output, and Sequential Entry," 1981, *American Economic Review*, 71, 503-514.
- Spulber, D.F. - "Products Liability and Monopoly in a Contestable Market," 1988, *Economica*, 55, 333-341.
- Sraffa, P. - *Production of Commodities by Means of Commodities*, 1960, Cambridge: Cambridge University Press.
- Stackelberg, H. von - *Grundlagen einer reinen Kostentheorie*, 1932, Vienna.
- Starret, D. - "Inefficiency and the Demand for 'Money' in a Sequence Economy," 1973, *Review of Economic Studies*, 40, 437-448.
- Stigler, G.J. - "The Economics of Information," 1961, *Journal of Political Economy*, 70, 94-105.
- Stigler, G.J. - *The Theory of Price*, 1966, New York: McMillan.
- Stigler, G.J. - *The Organization of Industry*, 1968, Homewood, Ill.: Irwin.
- Stigler, G.J. - "The Theory of Economic Regulation," 1971, *Bell Journal of Economics and Management Science*, 2, 3-21.
- Stigler, G.J. and Becker, G.S. - "De Gustibus Non Est Disputandum," 1977, *American Economic Review*, 67, 76-90.
- Stiglitz, J.E. - "Theory of Competition, Incentives and Risk," in: Stiglitz, J.E. and G.F. Mathewson (eds.) - *New Developments in the Analysis of Market Structure*, 1986, Cambridge, Mass.: MIT Press.
- Stiglitz, J.E. - "Technological Change, Sunk Costs, and Competition," 1987, *Brookings Papers on Economic Activity*, 18, 883-937.
- Stiglitz, J.E. - "Imperfect Information in the Product Market," in: Schmalensee, R. and R.D. Willig (eds.) - *Handbook of Industrial Organization*, 1989, Amsterdam: North-Holland.
- Stiglitz, J.E. and G.F. Mathewson (eds.) - *New Developments in the Analysis of Market Structure*, 1986, Cambridge, Mass.: MIT Press.
- Stiglitz, J.E. and A. Weiss - "Credit Rationing in Markets with Imperfect Information,"

- 1981, *American Economic Review*, 71, 393-410.
- Stiglitz, J.E. and A. Weiss - "Credit Rationing: Reply," 1987, *American Economic Review*, 77, 228-231.
- Stigum, B.P. - "Entrepreneurial Choice over Time under Conditions of Uncertainty," 1969a, *International Economic Review*, 10, 426-442.
- Stigum, B.P. - "Competitive Equilibrium under Uncertainty," 1969b, *Quarterly Journal of Economics*, 83, 533-561.
- Stigum, B.P. - "Resources Allocation under Uncertainty," 1972, *International Economic Review*, 13, 431-459.
- Stohs, M. - "'Uncertainty' in Keynes' General Theory," 1980, *History of Political Economy*, 3, 372-382.
- Stoneman, P. - *The Economic Analysis of Technological Change*, 1983, Oxford: Oxford University Press.
- Sugden, R. - "New Developments in the Theory of Choice under Uncertainty," 1986, *Bulletin of Economic Research*, 38, 1-24.
- Sutton, J. - "Vertical Product Differentiation: Some Basic Themes," 1986, *American Economic Review*, 76, 393-398.
- Sylos-Labini, P. - *Oligopoly and Technical Progress*, 1962, Cambridge, Mass.: Harvard University Press, first published in 1957 in Italian.
- Tang, M.J. - "An Economic Perspective on Escalating Commitment," 1988, *Strategic Management Journal*, 9, 79-92.
- Taubman, P. - "The Determinants of Earnings: Genetics, Family and other Environments: A Study of Male Twins," 1976, *American Economic Review*, 66, 858-870.
- Teece, D.J. - "Economies of Scope and the Scope of the Enterprise," 1980, *Journal of Economic Behavior and Organization*, 1, 223-247.
- Teece, D.J. - "Towards an Economic Theory of the Multiproduct Firm," 1982, *Journal of Economic Behavior and Organization*, 3, 39-63.
- Teece, D.J. - "Economic Analysis and Strategic Management," 1984, *California Management Review*, 26, 87-110.
- Thepot, J. - "The Strategic Investment Problem Revisited," 1989, presented at the *Workshop on Microeconomics/Strategic Management Interfaces*, Brussels: EIASM, November 16-17, 1989.
- Thibaut, J.W. and H.H. Kelly - *The Social Psychology of Groups*, 1959, New York: John Wiley.
- Thompson, W. and H.R. Varian - "Theories of Justice Based on Symmetry," in: Hurwicz et al. (eds.) - *Social Goals and Social Organization: Essays in Memory of Elisha Pazner*, 1985, New York: Cambridge University Press.
- Thünen, J.H. von - *Der isolierte Staat in Beziehung auf Landwirtschaft und National-ökonomie*, 1826.
- Thurow, L.C. - *Investment in Human Capital*, 1970, Belmont: Wadsworth.
- Thurow, L.C. - *Generating Inequality*, 1975, London: McMillan.
- Tinbergen, J. - "Welfare Economics and Income Distribution," 1957, *American Economic Review*, 48.
- Tirole, J. - *The Theory of Industrial Organization*, 1988, Cambridge, Mass.: MIT Press.
- Tullock, G. - "The Welfare Cost of Tariffs, Monopolies, and Theft," 1967, *Western Economic Journal*, 5, 224-232.
- Tullock, G. - "Efficient Rent Seeking," in: Buchanan, J.M., R.D. Tollison and G.

- Tullock (eds.) - *Towards a Theory of Rent Seeking Society*, 1980, College Station: Texas A.M. Press.
- Ulph, A.M. and D.T. Ulph - "Transaction Costs in General Equilibrium Theory: A Survey," 1975, *Economica*, 42, 355-372.
- Ushio, Y. - "Cournot-Nash Equilibria in Large Markets: The Case of Declining Average Cost Curves," 1983, *Review of Economic Studies*, 50, 347-354.
- Vanheukelen, M. - "Niet-tarifaire barrières in de Europese Gemeenschap," 1985, *Tijdschrift voor economie en management*, 30, 107-125.
- Varian, H.R. - "Equity, Envy and Efficiency," 1974, *Journal of Economic Theory*, 9, 63-91.
- Varian, H.R. - "Distributive Justice, Welfare Economics, and the Theory of Fairness," 1975, *Philosophy and Public Affairs*, 4, 223-247.
- Varian, H.R. - "Two Problems in the Theory of Fairness," 1976a, *Journal of Public Economics*, 5, 249-260.
- Varian, H.R. - "On the History of Concepts of Fairness," 1976b, *Journal of Economic Theory*, 13, 486-487.
- Varian, H.R. - "Redistributive Taxation as Social Insurance," 1980, *Journal of Public Economics*, 14, 49-68.
- Varian, H.R. - "Nonparametric Tests of Consumer Behaviour," 1983, *Review of Economic Studies*, 50, 99-110.
- Varian, H.R. - *Microeconomic Analysis*, 1984, New York: Norton.
- Veblen, T. - "Why is Economics not an Evolutionary Science?," 1898, *Quarterly Journal of Economics*, 12, 373-397.
- Verdoorn, P.J. and A.N.R. Schwartz - "Two Alternative Estimates of the Effects of EEC and EFTA on the Pattern of Trade", 1972, *European Economic Review*, 3, 291-335.
- Viner, J. - *The Customs Union Issue*, 1950, New York: Carnegie Endowment for International Peace.
- Vives, X. - "Technological Competition, Uncertainty, and Oligopoly," 1989, *Journal of Economic Theory*, 48, 386-415.
- Waagstein, T. - "Fixed Costs, Limit Pricing and Investments in Barriers to Entry," 1982, *European Economic Review*, 17, 75-86.
- Waagstein, T. - "A Dynamic Model of Entry Deterrence," 1983, *Scandinavian Journal of Economics*, 85, 325-337.
- Wald, A. - "Über einige Gleichungssysteme der Mathematische Ökonomie," 1936, *Zeitschrift für Nationalökonomie*, 7, 637-670.
- Walras, L. - *Éléments d'économie politique pure*, 1974, Lausanne: Corbaz, English translation, *Elements of Pure Economics*, 1954, Homewood, Ill.: Irwin.
- Walsh, J.P. - "On Deriving Pricing Rules in the Theory of Second-Best," 1982, *Public Finance Quarterly*, 10, 499-509.
- Wan, H.Y. Jr. - *Economic Growth*, 1971, New York: Harcourt Brace Jovanovich.
- Ware, R. - "Inventory Holding as a Strategic Weapon to Deter Entry," 1985, *Economica*, 52, 93-101.
- Waterson, M. - *Economic Theory of the Industry*, 1984, Cambridge: Cambridge University Press.
- Watson, C.M. - "Counter-Competition Abroad to Protect Home Markets," 1982, *Harvard Business Review*, 40-42.
- Waverman, L. - "The Regulation of Intercity Telecommunications," in: Phillips, A. (ed.)

- *Promoting Competition in Regulated Markets*, 1975, Washington, D.C.: Brookings Institution.
- Wegberg, M. van and A. van Witteloostuijn - "Multimarket Modeling in Industrial Organization, Part I: Bertrand Games with Nonnegative Entry Cost," 1989, *Research Memorandum RM 89-028*, University of Limburg.
- Wegberg, M. van and A. van Witteloostuijn - "Multimarket Competition: Entry Strategies and Entry Deterrence When the Entrant Has a Home Market," forthcoming in Thépot, J. and R.A. Thièrtart (eds.) - *The Microeconomics/Strategic Management Interfaces*, Brussels: EIASM.
- Wegberg, M. van and A. van Witteloostuijn - "Multimarkt concurrentie: strategische afhankelijkheden bij simultane concurrentie op meerdere markten," 1990b, forthcoming in *Bedrijfskunde: tijdschrift voor modern management*, 62.
- Weick, K.E. - *The Social Psychology of Organizing*, 1979, Reading, Mass.: Addison Wesley.
- Weintraub, E.R. - *Microfoundations: The Compatibility of Microeconomics and Macroeconomics*, 1979, Cambridge: Cambridge University Press.
- Weiss, L.W. - "Antitrust in the Electric Power Industry," in: Phillips, A. (ed.) - *Promoting Competition in Regulated Markets*, 1975, Washington, D.C.: Brookings Institution.
- Weitzman, M. - "Increasing Returns and the Foundations of Unemployment Theory," 1982, *Economic Journal*, 92, 787-804.
- Weitzman, M.L. - "Contestable Markets: An Uprising in the Theory of Industry Structure," 1983, *American Economic Review*, 73, 486-487.
- Weizsäcker, C.C. von - "A Welfare Analysis of Barriers to Entry," 1980a, *Bell Journal of Economics*, 11, 399-420.
- Weizsäcker, C.C. von - *Barriers to Entry: A Theoretical Treatment*, 1980b, Berlin: Springer-Verlag.
- Wenders, J.T. - "Excess Capacity as a Barrier to Entry," 1971, *Journal of Industrial Economics*, 20, 14-19.
- Wicksell, K. - *Geldzins und Güterpreise*, 1898, Jena: G. Fisher, English translation, *Interest and Prices*, 1936, New York: A.M. Kelley.
- Wieser, F. von - *Über den Ursprung und die Hauptgesetze des wirthschaftlichen Werthes*, 1884, Vienna.
- Wieser, F. von - *Der Natürlichen Werth*, 1889, English translation, *Natural Value*, 1893, London: McMillan.
- Willard, G.E. and A.M. Savara - "Patterns of Entry: Pathways to New Markets," 1988, *California Management Review*, 25, 57-76.
- Williamson, O.E. - "Selling Expense as a Barrier to Entry," 1963, *Quarterly Journal of Economics*, 77, 112-128.
- Williamson, O.E. - *Markets and Hierarchies*, 1975, New York: Free Press.
- Williamson, O.E. - "Transaction Cost Economics: The Governance of Contractual Relations," 1979, *Journal of Law and Economics*, 22, 238-246.
- Williamson, O.E. - "Transaction Cost Economics," in: Schmalensee, R. and R.D. Willig - *Handbook of Industrial Organization*, 1989, Amsterdam: North-Holland.
- Winston, G.C. - "Imperfectly Rational Choice: Rationality as the Result of a Costly Activity," 1989, *Journal of Economic Behavior and Organization*, 12, 67-86.
- Winter, S.G. - "Economic 'Natural Selection' and the Theory of the Firm," 1964, *Yale Economic Essays*, 4, 225-272.

- Winter, S.G. - "Optimization and Evolution in the Theory of the Firm," in: Day, R.H. and T. Groves (eds.) - *Adaptive Economic Models*, 1975, New York: Academic Press.
- Winter, S.G. - "Schumpeterian Competition in Alternative Technological Regimes," 1984, *Journal of Economic Behavior and Organization*, 5, 287-320.
- Witt, U. - "Economic Behavior and Biological Evolution: Some Remarks on the Sociobiology Debate," 1985, *Journal of Institutional and Theoretical Economics*, 141, 365-389.
- Witt, U. - "Firms' Market Behavior under Imperfect Information and Economic Natural Selection," 1986, *Journal of Economic Behavior and Organization*, 7, 265-290.
- Witteeloostuijn, A. van - "Rationality in Economics: An Effete Misnomer," 1987, *Research Memorandum RM 87-022*, University of Limburg.
- Witteeloostuijn, A. van - "Maximising and Satisficing: Opposite or Equivalent Concepts?," 1988a, *Journal of Economic Psychology*, 9, 289-313.
- Witteeloostuijn, A. van - "Entry and Exit Barriers: A Note on the Compatibility of Free Entry and Costly Exit," 1988b, *Economic Research Report 88-40*, New York University, C.V. Starr Center For Applied Economics.
- Witteeloostuijn, A. van - "Free Entry, Costly Exit and Amortisation: A Note on the Optimality of Conventional Behaviour," 1988c, University of Limburg, *Research Memorandum RM 88-013*, reprinted in Vanden Abeele, P. (ed.) - *Psychology in Micro and Macroeconomics*, Proceedings of the 13th annual colloquium of the International Association for Research in Economic Psychology, September 28 - October 1, Leuven: Catholic University of Leuven.
- Witteeloostuijn, A. van - "De effecten van de Europese integratiemaatregelen: een overzicht van de empirische en theoretische inzichten," 1988d, *Maandschrift economie*, 52, 265-282.
- Witteeloostuijn, A. van - "Uncertainty and Economic Psychology: A Note on the Opportunities Offered by Economic Concepts," in: Tyszka, T. and P. Gasparski (eds.) - *Homo Oeconomicus: Presumptions and Facts*, 1989a, Proceedings of the 14th annual colloquium of the International Association for Research in Economic Psychology, September 24-27, Warschau: Polish Academy of Sciences.
- Witteeloostuijn, A. van - "Toetredingsbelemmerende strategieën: een overzicht van de theoretische en empirische literatuur," 1989b, printed in *Bedrijfskunde: tijdschrift voor modern management*, 1990, 62, 59-71.
- Witteeloostuijn, A. van - "Investment Contestability and Average Cost Reduction," 1989c, forthcoming in the *European Journal of Political Economy*, 1990, 6.
- Witteeloostuijn, A. van - "Statische en dynamische efficiëntie op de arbeidsmarkt: potentiële concurrentie en marktprestaties," 1989d, *Maandschrift economie*, 53, 278-291.
- Witteeloostuijn, A. van - "Learning in Economic Theory," 1990a, forthcoming in the *Journal of Economic Psychology*, 11.
- Witteeloostuijn, A. van - "Economics and Psychology of Decision Making: Objectivism and Subjectivism," 1990b, forthcoming *Research Memorandum*, University of Limburg.
- Witteeloostuijn, A. van - "Product Innovation and Contestability," 1990c, forthcoming in *De Economist*, 138.
- Witteeloostuijn, A. van - "Multimarket Competition and Strategic Management," 1990d, *Research Memorandum RM 90-005*, University of Limburg.
- Witteeloostuijn, A. van - "Methods of Dynamic Economics: (Non)market Clearing and

- Expectations," 1990e, *Research Memorandum RM 90-004*, University of Limburg.
- Witteloostuijn, A. van - "Externe organisatie en de moderne theorie van concurrentie," 1989f, *Onderzoekmemorandum RM 90-006*, University of Limburg, forthcoming in *Maanblad voor Bedrijfsadministratie en Bedrijfsorganisatie*, 1990, 94.
- Witteloostuijn, A. van and A. van Lier - "Chaotic Patterns in Cournot Competition," 1990, *Research Memorandum RM 90-002*, University of Limburg, forthcoming in *Metroeconomica: An International Review of Economics*, 1990, 41.
- Witteloostuijn, A. van and J.A.H. Maks - "Concurrentie en mededingingsbeleid: de barrière-markt als een micro-fundering van werkbare concurrentie," 1987, *Maandschrift economie*, 51, 448-461.
- Witteloostuijn, A. van and J.A.H. Maks - "Workable Competition and the Barrier Market," 1988, *European Journal of Political Economy*, 4, 117-135.
- Witteloostuijn, A. van and J.A.H. Maks - "Walras: A Hicksian *avant la lettre*," 1989, *Economie appliquée*, 41, 595-608.
- Witteloostuijn, A. van and J.A.H. Maks - "Walras on Temporary Equilibrium and Dynamics," 1990, forthcoming in the *History of Political Economy*, 22.
- Witteloostuijn, A. van and M. van Wegberg - "Product Quality Innovation and Entry Deterrence," 1988, *Research Memorandum RM 88-019*, University of Limburg, paper presented at the 15th annual conference of the European Association for Research in Industrial Economics, August 31 - September 2, 1988, Rotterdam.
- Witteloostuijn, A. van and M. van Wegberg - "Multimarket Modeling in Industrial Organization, Part II: Cournot-Bertrand Games with Sunk Costs," 1989, *Research Memorandum RM 89-029*, University of Limburg.
- Witteloostuijn, A. van and M. van Wegberg - "Multimarket Competition and European Integration", in: Rugman, A.M. and A. Verbeke (eds.) - *Research in Global Strategic Management. Volume II: Europe 1992*, 1990, Greenwich, Conn.: JAI Press.
- Yellen, J.L. - "Efficiency Wage Models of Unemployment," 1984, *American Economic Review*, 74, 200-205.
- Yip, G.S. - *Barriers to Entry*, 1982, Lexington, Mass.: Lexington Books.
- Yoffie, D.B. and H.V. Milner - "An Alternative to Free Trade or Protectionism: Why Corporations Seek Strategic Trade Policy," 1989, *California Management Review*, 26, 111-131.
- Yunker, J.A. - "In Defence of Utilitarianism: An Economist's Viewpoint," 1986, *Review of Social Economy*, 54, 27-79.

In deel I wordt aangegeven dat het onderwerp van het proefschrift betrekking heeft op drie centrale begrippen binnen de economische wetenschap: rationaliteit, concurrentie en evolutie (hoofdstuk 1). De meeste aandacht gaat echter uit naar een analyse van concurrentieprocessen (hoofdstuk 2). In deel II wordt de discussie rond het *rationaliteitsbegrip* (hoofdstuk 3) en het *maximalisatieprincipe* (hoofdstuk 4) onder de loep genomen. Een kritische evaluatie van het debat lijkt uit te wijzen dat de tegenstanders van beide concepten niet in staat zijn gebleken een bruikbaar alternatief te formuleren. Om die reden worden in deze dissertatie de rationaliteitsnotie in het algemeen en het maximalisatieprincipe in het bijzonder als uitgangspunt gehanteerd: verondersteld wordt dat het gedrag van economische agenten beschreven kan worden door een (subjectieve) doelstellingsfunctie te maximaliseren onder randvoorwaarden.

In deel III wordt een overzicht gepresenteerd van de belangrijkste inzichten die de economische wetenschap in het algemeen en de industriële economie in het bijzonder hebben opgeleverd met betrekking tot de werking van *concurrentie*. Hierbij staat de *afweging* tussen statische efficiëntie enerzijds en dynamische voordelen anderzijds centraal (hoofdstuk 5). Kortweg kan worden gesteld dat *statische efficiëntie* wordt bevorderd zodra prijzen tenderen in de richting van het niveau van de (minimale) gemiddelde kosten, terwijl *dynamische voordelen* onder meer voortvloeien uit proces- en produktinnovaties (bij voorbeeld kostenredukties respectievelijk produktverbeteringen). Een drietal theorieën beschrijft statisch efficiënte concurrentie: volkomen mededinging, (homogene) Bertrand concurrentie en betwistbare markten (hoofdstuk 6). De theorie van betwistbare markten benadrukt het belang van krachtige potentiële concurrentie: op een betwistbare markt worden zittende bedrijven gedwongen gemiddelde kostprijzen te stellen onder druk van de toetredingsdreiging door gelijkwaardige potentiële rivalen.

Kenmerkend voor betwistbare markten (en volkomen mededinging) is de aanname dat toetreding *vrij* en uittreding *kosteloos* kan geschieden. Deze veronderstelling impliceert de afwezigheid van toe- en uittredingsbarrières. De introductie van dynamische voordelen gaat echter in het algemeen hand in hand met het opwerpen van drempels. De reden is dat het introduceren van dynamische aspecten van concurrentie verzonken investeringen verlangt. Juist dergelijke investeringen leiden tot het ontstaan van *toe- en uittredingsbelemmeringen* (hoofdstuk 7). *Proces- en produktinnovaties* vormen belangrijke voorbeelden van dynamische voordelen van concurrentie die gepaard kunnen gaan met het ontstaan van drempels (hoofdstuk 8).

In de eerste plaats creëert een verzonken investering door een zittende aanbieder *kostennadelen* voor potentiële rivalen indien de laatste nog tot het verzinken van kosten moeten overgaan. Bij voorbeeld het opbouwen van naamsbekendheid door zittende aanbieders met behulp van advertentie-uitgaven plaatst een onbekende potentiële concurrent op een achterstand: de potentiële rivaal moet immers additionele kosten (in de vorm van een kostbare reclamecampagne) investeren om met de zittende ondernemingen op gelijke voet te kunnen concurreren. In de tweede plaats kunnen verzonken investeringen uittredingsdrempels opwerpen. Verzonken kosten moeten op de markt

geminimaliseerd (hoofdstuk 13). Dit betekent dat de kans groot is dat de marktvraag niet geheel kan worden bevredigd door een natuurlijk veelvoud van deze unieke omvang. Daarenboven determineren fluktuierende vraag- en investeringshoeveelheden een wisselend aantal zittende ondernemingen in het evenwicht. In de tweede plaats kan de noodzaak tot *afschrijving* van investeringskosten tijdelijke toetreding door gelijkwaardige potentiële rivalen mogelijk maken (hoofdstuk 14). Indien zittende en potentiële concurrenten gebruik maken van afwijkende afschrijvingsregels [dat wil zeggen, verschillende (reeksen van) totale kostprijzen calculeren], dan geeft de afschrijvingsprocedure van de mogelijke toetreders op enig moment in de tijd een lagere kostprijs dan die van het zittende bedrijf. Dit schept winstgevend toetredingsmogelijkheden. Het bestaan van afschrijvingsconventies kan echter soelaas bieden.

In deel VII wordt een tweetal uitwerkingen van barrière-goederenmarkten beschreven. Het eerste model betreft *kostenredukties* (hoofdstuk 15). Het voorkomen van toetreding vergt in dit geval het minimaliseren van de toekomstige kostprijs via de introductie van een efficiënte procesinnovatie. Indien de innovatietechnologie aan enkele (gebruikelijke) voorwaarden voldoet, kan een unieke investeringsomvang worden bepaald zodanig dat een maximale kostenreductie (bestaande uit een afneming van de gemiddelde variabele en vaste kosten en een toeneming van de gemiddelde verzonken kosten) wordt bewerkstelligd. Het tweede model heeft betrekking op verbeteringen van de *produktkwaliteit*. Een analoge argumentatie geeft aan dat potentiële rivalen alleen buiten de markt kunnen worden gehouden door produktverbeteringen te introduceren die de afnemers het meeste nut bezorgen. De analyse wordt gecompliceerd door twee effecten die produktinnovaties met zich mee brengen (waarvan de tweede ontbreekt in het geval van procesverbeteringen): aan de ene kant leidt de introductie van verzonken investeringen tot een verschuiving van de efficiënte (dat wil zeggen, kostenminimaliserende) schaal van produktie; aan de andere kant laat een produktverbetering de helling en lokatie van de vraagcurve niet onverlet.

In deel VIII worden de drie categorieën concurrentietheorieën toegepast op de *arbeidsmarkt*. Op de arbeidsmarkt kunnen dynamische voordelen de vorm aannemen van bijvoorbeeld verbeteringen van menselijk kapitaal (die zich manifesteren via de toeneming van de produktiviteit). In samenhang met looneisen kan ook hier een afweging tussen statische en dynamische efficiëntie optreden (hoofdstuk 17). Op een betwistbare arbeidsmarkt zijn zittende aanbieders bang voor de reacties van potentiële rivalen die beschikken over *identiek* menselijk kapitaal. Op een betwistbare arbeidsmarkt worden zittende aanbieders gedwongen tot het stellen van looneisen op het niveau van een minimum (reservatie)loonvoet. Met name de bedreiging vanuit gerelateerde arbeidsmarkten en de potentiële toetreding door gelijkgeschoolde werklozen zijn relevant.

De invloed van potentiële concurrentie kan afnemen ten gevolge van beperkte toetreding tot opleidingen en regionale immobiliteiten. Als zittende werknemers een hogere produktiviteit kunnen bewerkstelligen via verzonken investeringen in specifieke bedrijfsvaardigheden, bieden de ontstane toetredingsbelemmeringen de mogelijkheid hogere looneisen te stellen. Een zittende aanbieder kan een positief investeringsrendement behalen zonder toetreding uit te lokken. Dynamische voordelen (bijvoorbeeld specifieke bedrijfsvaardigheden) worden geïntroduceerd ten koste van de statische efficiëntie van het marktgedrag.

De barrière-markttheorie geldt voor de gevallen waarin de zittende aanbieders aannemen dat de aanwezigheid van verzonken kosten vrije toetreding *niet* (volledig) onmogelijk maakt. Dit impliceert het bestaan van potentiële toetreders die, aan de ene kant, geconfronteerd worden met dezelfde hoogte van verzonken kosten en die, aan de andere

kant, gelijkwaardige verbeteringen van menselijk kapitaal bewerkstelligen (hoofdstuk 18). Hierbij kan worden gedacht aan toetreding(sdreigingen) vanuit een markt waarop een aanverwante categorie arbeid wordt aangeboden of aan toetreding vanuit een nabij gelegen regio. Op een barrière markt worden aanbieders van arbeid gedwongen tot het stellen van gematigde looneisen. Toetreding kan slechts worden voorkomen door looneisen te stellen die een nulrendement op de investeringen in menselijk kapitaal opleveren. De *gematigde looneisen* gaan hand in hand met *investeringen in menselijk kapitaal* die de attractiviteit van de zittende werknemers voor werkgevers maximaliseren.

De discussie over de bronnen van toetreding en barrière-marktscenario's wijst in de richting van de relevantie van het concept *multimarkt concurrentie*. De gedachtenvorming rond multimarkt concurrentie staat centraal in deel IX. Multimarkt concurrentie is gedefinieerd als concurrentie tussen rivalen die opereren in een verzameling aanverwante markten (hoofdstuk 19). Deze invalshoek benadrukt de relevantie van een algemeen in plaats van een partieel perspectief. Centrale concepten zijn intermarkt overloopeffecten en wederzijdse toetreding(sdreigingen). Intermarkt overloopeffecten ontstaan zodra de evenwichtsbeloningen die op de ene markt worden verkregen, van invloed zijn op de beloningen op de andere markt. Voorbeelden zijn gemeenschappelijke schaal- en assortimentsvoordelen. Wederzijdse toetredingsdreigingen impliceren dat de zittende aanbieder op de ene markt een potentiële rivala is van de concurrent op de andere markt en vice versa.

Verticale multimarkt concurrentie kan onder meer optreden tussen rivalen op gerelateerde goederen- en arbeidsmarkten (hoofdstuk 20). Onder andere deze interactie staat centraal in de algemene evenwichtstheorie (inclusief de onevenwichtigheidsvariant) in het algemeen en het debat rond de microfundering van de macro-economie in het bijzonder. Rantsoenering van vraag of aanbod op een markt kan aanleiding geven tot toe- en uittredingsbewegingen. Bij voorbeeld aanbodtekorten bieden potentiële rivalen de mogelijkheid winstgevend toe te treden, terwijl vraagtekorten uittreding onvermijdelijk maken. Deze toe- en uittredingsbewegingen gaan gepaard met een tendens tot gelijkschakeling van prijs-, loon- en/of kostenniveaus. Daarnaast bestaat de mogelijkheid van het gelijktijdig optreden van gedifferentieerde werkloosheid en onvervulbare vacatures.

Horizontale multimarkt concurrentie kan onder andere ontstaan tussen aanbieders op gerelateerde goederenmarkten (hoofdstuk 21). Deze interactie is onderwerp van studie binnen de vakgebieden industriële economie en strategisch management. Potentiële rivaliteit en horizontale multimarkt concurrentie vormen twee zijden van dezelfde medaille. Een gedegen concurrentie-analyse kan niet zonder de beantwoording van vragen met betrekking tot de identiteit van potentiële rivalen, de geloofwaardigheid van de toetredingsdreiging en het belang van intermarkt overloopeffecten. Deze stelling wordt geïllustreerd aan de hand van een model waarin bedrijven op twee gerelateerde markten beslissingen nemen omtrent capaciteitsuitbreiding. Het feit dat beide partijen de mogelijkheid bezitten gezamenlijke schaalvoordelen te benutten door op de andere markt toe te treden, compliceert de analyse in aanzienlijke mate.

Concurrentie vindt niet alleen plaats in de context van meerdere markten, maar evolueert daarnaast over de tijd. Dat wil zeggen, concurrentie is een *dynamisch proces*. In deel X wordt tot slot enige aandacht besteed aan het thema *evolutie*. Allereerst worden de *methoden* van dynamische analyse die in de economische literatuur kunnen worden teruggevonden, op een rijtje gezet (hoofdstuk 22). Dynamische concurrentieprocessen zijn verbonden met drie kenmerken: verwachtingsfouten, heterogeniteit van agenten en sequentiële besluitvorming. Economische agenten met verschillende eigenschappen zijn

verwikkelt in een proces van actie en reactie waarin verwachtingen regelmatig worden gelogenstraft. Een nadere bestudering van dergelijke concurrentieprocessen kan in het algemeen slechts geschieden met behulp van *simulatieve* technieken.

Een simulatieve analyse van eenvoudige modellen van betwistbare en barrière markten kan de complexiteit van dynamische concurrentieprocessen illustreren (hoofdstuk 23). Een eerste groep simulaties laat zien dat betwistbare markten aan de ene kant statisch efficiënt marktgedrag stimuleren maar aan de andere kant intertemporele instabiliteit vertonen. Een laatste experiment simuleert een evolutionair *selectieproces* op een barrière markt. In een omgeving met technische onzekerheid proberen concurrenten te overleven door middel van het introduceren van procesinnovaties. Rivalen die in mindere mate het slachtoffer zijn van verwachtingsfouten, zijn in staat inferieure concurrenten te onderbieden. Het selectieproces gaat gepaard met toe- en uittredingsbewegingen.

Aangezien de barrière-marktheorie een referentiepunt beoogt te beschrijven waarbij statisch en dynamisch efficiënt marktgedrag wordt gestimuleerd, is een beschouwing over welvaartstheoretische implicaties onvermijdelijk (hoofdstuk 24). Een overzicht van de inzichten uit de welvaartstheorie geeft aan dat welvaartstheoretische uitspraken over dynamische concurrentieprocessen met grote *voorzichtigheid* moeten worden omgeven. De cruciale complicatie is dat concurrentieprocessen winnaars *en* verliezers opleveren. Dat wil zeggen, hoewel barrière-marktsenario's een concurrentie-omgeving beschrijven waarin rivalen zich gedwongen voelen voorzichtige prijsstelling (c.q. looneisen) te combineren met de introductie van investeringen in proces- en produktinnovaties (c.q. verbeteringen in menselijk kapitaal), is het niet mogelijk eenduidige welvaartstheoretische conclusies te verbinden aan de uitkomsten van concurrentie op barrière markten.

Het proefschrift wordt afgesloten met een pleidooi voor het uitbuiten van de mogelijkheden tot *kruisbestuiving* tussen onderzoek op verschillende vakgebieden binnen de economische wetenschap (hoofdstuk 25). Hierbij wordt met name gewezen op de vruchten die geplukt kunnen worden bij integratie van de industriële economie enerzijds en de microgefundeerde macro-economie en het strategisch management anderzijds. Twee voorbeelden kunnen deze stelling illustreren. In de eerste plaats kan de theorievorming rond de microfundering van de macro-economie worden verbeterd door aandacht te besteden aan de bijdrage van de industriële economie aan de bestudering van dynamische concurrentiemechanismen. In de tweede plaats kunnen industrieel-economische inzichten worden verrijkt door de geloofwaardigheid van toetredingsdreigingen te microfunderen in een multimarkt raamwerk. Het moge duidelijk zijn: onderzoekers in deze vakgebieden kunnen profiteren van de mogelijkheden tot wederzijdse toetreding tot elkaars onderzoeksterreinen.

CURRICULUM VITAE

Arjen (Adriaan) van Witteloostuijn is op 12 oktober 1960 geboren in Haarlem. In 1979 behaalde hij zijn Atheneum-B diploma aan de RSG Epe (Gelderland). Gedurende de periode september 1979 - januari 1985 studeerde hij Economie en Psychologie (vanaf september 1981) aan de Rijksuniversiteit Groningen. Na het behalen van zijn doctoraal diploma Algemene en Bedrijfseconomie (cum laude) is hij in februari 1985 als Wetenschappelijk Assistent in dienst getreden bij de Economische Faculteit van dezelfde universiteit (zijn studie Psychologie heeft hij medio 1986 voortijdig afgebroken). Het promotie-onderzoek bevond zich op het raakvlak van de micro- en macro-economie. Halverwege het jaar 1986 heeft hij het onderzoek voortgezet aan de Economische Faculteit van de Rijksuniversiteit Limburg (eveneens als Wetenschappelijk Assistent). Van maart tot juni 1988 is hij als gastonderzoeker werkzaam geweest aan de C.V. Starr Center for Applied Economics, Department of Economics van de New York University in de Verenigde Staten. Voor dit bezoek ontving hij een beurs van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek. Na een kortstondige periode als Universitair Docent "Micro-Economie" (van oktober 1988 tot maart 1989) is hij benoemd tot (beoogd) Universitair Hoofddocent "Externe Organisatie en Strategisch Management" bij de Sectie Organisatie van de Vakgroep Bedrijfseconomie binnen Faculteit der Economische Wetenschappen der Rijksuniversiteit Limburg.