

Speculative markets dynamics : an econometric analysis of stock market and foreign exchange market dynamics

Citation for published version (APA):

Nieuwland, F. (1993). *Speculative markets dynamics : an econometric analysis of stock market and foreign exchange market dynamics*. [Doctoral Thesis, Maastricht University]. Datawyse / Universitaire Pers Maastricht. <https://doi.org/10.26481/dis.19930324fn>

Document status and date:

Published: 01/01/1993

DOI:

[10.26481/dis.19930324fn](https://doi.org/10.26481/dis.19930324fn)

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.

Speculative Markets Dynamics

An Econometric Analysis of Stock Market and
Foreign Exchange Market Dynamics

Speculative Markets Dynamics

**An Econometric Analysis of Stock Market and
Foreign Exchange Market Dynamics**

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Limburg te Maastricht, op
gezag van de Rector Magnificus, Prof. Mr. M.J. Cohen,
volgens het besluit van het College van Dekanen,
in het openbaar te verdedigen op woensdag,
24 maart 1993 om 16.00 uur

door

Frederik Gertruda Maria Carolus Nieuwland

geboren te Hoensbroek



Promotor :

Prof. dr. C.C.P. Wolff

Beoordelingscommissie :

Prof. dr. J.G. Backhaus (Voorzitter)

Prof. dr. Th. E. Nijman

Dr. P.C. Schotman

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Nieuwland, Frederik Gertruda Maria Carolus

Speculative markets dynamics : an econometric analysis of
stock market and foreign exchange market dynamics /

Frederik Gertruda Maria Carolus Nieuwland. - Maastricht :
Universitaire Pers Maastricht. - Ill., fig., tab.

Thesis Maastricht. - With ref.

ISBN 90-5278-069-2

NUGI 682/683

Subject headings: asset pricing model / GARCH models /
risk premia.

This dissertation is dedicated to the memory of :

Maria Nieuwland-Neff
Gertruda Wiertz-Hollands
Leonard Wiertz

Without them I would not have been what I am today.

Preface

This dissertation is the result of four years of learning, reading, and doing research. I have enjoyed every minute of those four years and I have especially learned to appreciate the freedom I had in setting my own goals and pursuing my own research interests. To me this dissertation is not the final result of a research project, which upon completion should end up in a filing cabinet. It is rather, an opening statement in a hopefully long-lasting discourse. Four years ago I did not expect to be as fascinated by doing research as I am today, but in retrospect the preparation and writing of this dissertation has been one big thrill.

Acknowledgements

I could not have completed this dissertation without the general support I have received over the last four years. First of all I gratefully acknowledge research support by grant 450-227-009 from the Foundation for the Promotion of Research in Economic Sciences (Ecozoek), and financial assistance by the Staal Stichting 1966 to print this dissertation. I owe an unrepayable debt to many people that have never seized to encourage me. First in line is Christian Wolff, who has been much more than just a thesis advisor. Like a shepherd watching his flock, he always made sure that I did not wander off too much from the right track. Any good shepherd has a well-trained sheep-dog. This characterization fits the description of my room-mate Willem Verschoor. Whenever I strayed behind or had another crazy idea, like many sheep tend to have, he summoned me to hurry up, get my act together, and to get back in line. This sheep really needed that. Thanks to both of you!

I would also like to thank my colleagues at the Finance Department of the University of Limburg, who provided me with a first-class working environment, and made me feel like a happy sheep in a happy flock. A special word of thanks goes out to Peter Schotman, Kees Koedijk, Art Selender, Franz Palm, Carina Wijnands and Nathalie Jansen, each of them knows why. The efforts of the members of my dissertation committee consisting of Jürgen Backhaus, Theo Nijman, and Peter Schotman are also greatly appreciated.

I welcome this opportunity to thank all my friends for the interest they have taken, in one form or another, in my research, and for providing me with a joyful social environment. Finally I would like to express my gratitude and love to my brother and parents, to whom I owe the largest of the unrepayable debts.

Fred Nieuwland
January 1993 Maastricht

Table of Contents

	page
<i>Table of Contents</i>	1
<i>List of Tables</i>	4
<i>List of Figures</i>	5
1 Introduction, Motivation, and Setting	9
1.1 The History of Financial Economics	9
1.2 The Role of Financial Economics within Economics	12
1.3 Time-Varying Risk Premia in Financial Economics	16
1.4 GARCH as a Tool for Time-Varying Risk	17
1.5 Research Objectives	18
2 Conditional and Unconditional Distributions for Stock Returns	23
2.1 Introduction	23
2.2 Data and Summary Statistics	26
2.3 Conditional Heteroskedasticity Tests	30
2.4 GARCH Models and the Choice of a Conditional Distribution	37
2.5 Estimation Procedures and Results	42
2.6 Diagnostic Checks	51
2.7 The Effects of Temporal Aggregation	54
2.8 Conclusions	65
Appendix 2.A : Aggregation Results	66

3	An Intertemporal Three-Moment Capital Asset Pricing Model	69
3.1	Introduction	69
3.2	The Linearization of a Non-Linear Budget Constraint	72
3.3	The Utility and Optimal Behavior Framework	76
3.4	The Elimination of Consumption	85
3.5	Empirical Implementation	97
3.6	Conclusions	98
Appendix 3.A	: The Derivation of the MRS	100
Appendix 3.B	: The Linearization of the Euler Equations	103
Appendix 3.C	: The Conditional Variance of Consumption Growth as a Linear Function of the Conditional Variance of the Market Return	104
Appendix 3.D	: Revisions in the Expected Future Values of an ARMA Process	108
4	Asymmetry In Stock Returns : a New Approach	111
4.1	Introduction	111
4.2	Data and GARCH-in-mean Results	117
4.3	The Model and its Assumptions	122
4.4	Calibration Results	125
4.5	Conditional Third Moment Specification	128
4.6	Estimation	130
4.7	Empirical Results	132
4.8	Conclusions	137
5	Stochastic Trends and Jumps In EMS Exchange Rates	139
5.1	Introduction	139

5.2 Random Walks or Mean Reversion	142
5.3 Modelling EMS Exchange Rates	152
5.3.1 The Lognormal Diffusion Process	153
5.3.2 The Mixed Jump-Diffusion Process	153
5.3.3 The Diffusion-(G)ARCH Process	154
5.3.4 The Jump-Diffusion-(G)ARCH Process	155
5.3.5 The Jump-Diffusion-(G)ARCH-t Process	156
5.4 Empirical Results	157
5.5 Conclusions	168
Appendix 5.A : Unit Root Tests	170
6 EMS Exchange Rate Expectations and Time-Varying Risk Premia	175
6.1 Introduction	175
6.2 A Brief Review of Theory	177
6.3 The Survey Data and their Summary Statistics	180
6.4 Time-Varying Exchange Risk Premia	186
6.5 Modelling Time-Varying Risk Premia	190
6.6 Empirical Results	193
6.7 Conclusions	204
7 Summary and Discussion	207
7.1 Summary and Discussion	207
<i>References</i>	213
<i>Nederlandse Samenvatting (Summary in Dutch)</i>	227
<i>Curriculum vitae</i>	231

List of Tables

Table		Page
2.1	Summary Statistics Weekly Returns	28
2.2	Summary Statistics Monthly Returns	29
2.3.a	ARCH Tests for Weekly Returns	36
2.3.b	ARCH Tests for Monthly Returns	36
2.4	GARCH(1,1) Normal Estimates, Weekly Returns	45
2.5	GARCH(1,1) Normal Estimates, Monthly Returns	46
2.6	GARCH(1,1) Student-t Estimates, Weekly Returns	47
2.7	GARCH(1,1) Student-t Estimates, Monthly Returns	48
2.8	GARCH(1,1) GED Estimates, Weekly Returns	49
2.9	GARCH(1,1) GED Estimates, Monthly Returns	50
2.10	Modified Likelihood Ratio Tests	52
2.11.a	Kurtosis Values for Standardized Weekly Residuals	53
2.11.b	Kurtosis Values for Standardized Monthly Residuals	53
2.12	Direct Monthly Estimates Compared to Implied Estimates	55
2.13	Optimal Choice of m , through Monte Carlo Simulations	59
2.14	Tail Index Estimates for Weekly Raw Stock Returns	61
2.15	Tail Index Estimates for Weekly Rescaled Innovations	62
2.16	Tail Index Estimates for Monthly Rescaled Innovations	62
2.17	Tests for Equality of Tail Index Estimates	64
4.1	GARCH(1,1)-M Normal Estimates, Weekly Returns	120
4.2	GARCH(1,1)-M Student-t Estimates Weekly Returns	121
4.3	Normal Implied Values for Three-Moment CAPM	127
4.4	Student-t Implied Values for Three-Moment CAPM	127
4.5	Parameter Estimates for GARCS Process	129
4.6	GMM Estimates for Three Moment CAPM	134
4.7	Correlations between Variance and Third Moment	135
4.8	Combined Parameters of Three Moment CAPM	136
5.1	Summary Statistics of Weekly Log Price Changes	144
5.2	Conditional Heteroskedasticity Tests	148
5.3.a	Tests for Unit Roots in Log Exchange Rates	150
5.3.b	Tests for Unit Roots in Parity Adjusted Rates	151
5.4	Diffusion Model Estimates	158
5.5	Diffusion-ARCH Model Estimates	159

5.6	Diffusion-GARCH Model Estimates	160
5.7	Jump-Diffusion Model Estimates	161
5.8	Jump-Diffusion-ARCH Model Estimates	162
5.9	Jump-Diffusion-GARCH Model Estimates	162
5.10	Generalized Likelihood Ratio Tests	164
5.11	Jump-Diffusion-GARCH-t Model Estimates	166
5.12	Diagnostics	167
6.1.a	Summary Statistics of Expected Depreciation	183
6.1.b	Summary Statistics of Forward Discount	184
6.1.c	Summary Statistics of the Risk Premium	185
6.2	Test of Perfect Substitutability	188
6.3	Conditional Heteroskedasticity Tests	189
6.4	ARCH-in-mean Model Estimates	194
6.5	GARCH-in-mean Model Estimates	195
6.6	Generalized Likelihood Ratio Tests	201
6.7	Diagnostics for ARCH-in-mean Model	202
6.8	Diagnostics for GARCH-in-mean Model	203

List of Figures

Figure

5.1	French Franc/Deutschmark Exchange Rate	146
5.2	Weekly Returns FF/DM Exchange Rate	146
5.3	Italian Lira/Deutschmark Exchange Rate	147
5.4	Weekly Returns IL/DM Exchange Rate	147
6.1.a	Actual Premium and Estimated ARCH Premium DG/DM .	198
6.1.b	Actual Premium and Estimated GARCH Premium DG/DM	198
6.2.a	Actual Premium and Estimated ARCH Premium BF/DM .	199
6.2.b	Actual Premium and Estimated GARCH Premium BF/DM	199
6.3.a	Estimated ARCH and GARCH Premia DG/DM	200
6.3.b	Estimated ARCH and GARCH Premia BF/DM	200

Chapter 1

Introduction, Motivation, and Setting

1.1 The History of Financial Economics

Research in Finance prior to 1950 can be characterized as using predominantly a descriptive and institutional approach. Analytical rigour was strikingly absent, although a few major contributions containing formal analyses, were made as early as 1900, e.g. Bachelier (1900) and Fisher (1930). Unfortunately, the recognition of these contributions as truly seminal did not occur until the mid-fifties and sixties. Hirshleifer (1958) presented a rigorous treatment of the net present value rule along the lines of Fisher (1930). Bachelier's Theory of the Random Walk for security price changes, which leads to normally distributed returns is one of the cornerstones of the Modern Portfolio Theory as developed by Markowitz (1952, 1959) and Tobin (1958). Mandelbrot (1963) and Fama (1965) have subsequently refuted the Random Walk Hypothesis as they showed that security price changes are not independently distributed.

The aforementioned contributions marked, what Hamada (1988) calls : "The Start of Modern Finance". Markowitz and Tobin analytically addressed an individual's investment decision in a portfolio context, and by recognizing and characterizing portfolio risk, statistics entered into the Theory of Finance. At the same time Economics were introduced into Finance by a series of articles by Modigliani and Miller (1958,1961,1963). They incorporated the roles of capital markets and general equilibrium in an uncertain

world into the study of central finance topics like capital structure, cost of capital, dividend policy, valuation and growth. The introduction of statistics and economics into finance led Hamada (1988) to conclude that :

"The study of finance was destined never to be the same"

Meanwhile the world entered into the computer era, and the development of better and faster computers facilitated the empirical testing of financial theories through the availability of computer-accessible financial databases (e.g. the CRSP database). This led to the influential studies on the behavior of security prices by Fama (1965), Mandelbrot (1963) and Samuelson (1965). Using the implications of the major empirical studies of that time, Fama (1970) introduced the efficient market hypothesis which is still central to modern finance theory. Sharpe (1964) and Lintner (1965) picked up on the work by Markowitz, Modigliani and Miller and developed an equilibrium model of security prices, the ubiquitous CAPM, hereby characterizing the risk-return tradeoff for individual securities in an efficient market. The consequences of all of these studies are described in Hamada (1988) :

"With the preceding developments, finance had a paradigm that tied together in one unifying theory both the traditional investment and portfolio decisions made directly by individuals (or through a financial intermediary) and the traditional corporation finance issues, such as dividend, capital structure, and capital budgeting policy. When this was recognized, finance was renamed *financial economics*".

In two decades Finance Theory had evolved from a purely descriptive discipline to a rigorous economic theory subjected to careful empirical testing. At a theoretical level, single-period models were generalized to multi-period models, models set in discrete time were transformed into models set in continuous time (which was fundamental to the development of contingent claim analysis) and as a complement to equilibrium pricing the notion of arbitrage pricing was developed. Seminal contributions are Merton (1969, 1971, 1973), Ross (1976) and Black and Scholes (1973). While generating greater realism than single-period discrete models, multi-period continuous models also demand greater mathematical complexity. Dynamic programming and stochastic calculus have become indispensable tools in the theory of finance. Today the theory of financial economics is a highly analytical subject.

At an empirical level, research in financial economics has kept a close eye on econometric developments and extensively uses econometric procedures which were not necessarily developed for financial applications. Nowadays, econometric concepts like seemingly unrelated regressions, vector auto-regressions, GMM estimation, non-parametric estimation and autoregressive conditional heteroskedasticity are as familiar to the empirical financial economist as they are to the econometrician. Furthermore, because of the still increasing quality and quantity of financial data, the environment of finance presents an obvious biotope to "experiment" with new econometric procedures.

Finally at a practical level, one can observe that the topics addressed by financial economics have found their profitable way into the business community. Perhaps finance is the only field within economics in which practical relevance is so immediately obvious. This point is convincingly illustrated by Merton (1990) :

"Yet, for all its mathematical complexity, finance theory has had a direct and significant influence on finance practice. A casual comparison of current practices with those of 20 years ago is enough to note the impact of efficient market theory, portfolio selection, risk analysis, and contingent-claim pricing theory on money management, financial intermediation, investment banking, and corporate financing and capital budgeting procedures."

In its evolution financial economics has absorbed a vast amount of ideas, concepts and techniques from other disciplines (both within as well as outside the realm of economics), to attain its current position as a vigorous and exhilarating economic subdiscipline. In the next section we will examine the current status of financial economics within economics and its influence on other economic subdisciplines.

1.2 The Role of Financial Economics Within Economics

The current position of financial economics within economics is quite distinctive, to say the least. Ross (1987) even uses adjectives like "isolated" and "eccentric". Obviously financial economics and other economic subdisciplines are interrelated, but there may be as many dissimilarities, as similarities.

First of all a difference in focus between financial economics and general economics can be noted. Financial economists tend to be concerned with the relation between the prices of different assets, and financial models tend to focus on implications across assets. In contrast, general economists tend to be concerned with the relation between the prices of assets and other economic variables and consequently macro model tend to concentrate on the implications across time for the price of a given asset. This point can be clarified by the brilliant example of ketchup economics provided by Summers (1985) :

"There are two groups of researchers concerned with ketchup economics. Some general economists study the market for ketchup as part of the broader economic system. The other group is comprised of ketchup economists located in Department of Ketchup where they receive much higher salaries than do general economists. Each group has a research program. General economists are concerned with the fundamental determinants of prices and quantities in the ketchup market. They attempt to examine various factors affecting the supply and demand for ketchup such as the cost of tomatoes, wages, the prices of ketchup substitutes and consumer incomes. They examine a number of different types of data in an effort to explain fluctuations in ketchup prices. The models that are estimated have some successes in explaining price fluctuations but there remain puzzles. Ketchup economists reject out of hand much of this research on the ketchup market. They believe that the data used is based on almost meaningless accounting information and are quick to point out that concepts such as costs of production vary across firms and are not accurately measurable in any event. They believe that ketchup transactions prices are the only hard data worth studying. Nonetheless ketchup economists have an impressive research program, focusing on the scope for excess opportunities in the ketchup market. They have shown that two quart bottles of ketchup invariably sell for twice as much as one quart bottles of ketchup except for deviations traceable to transactions

costs, and that one cannot get a bargain on ketchup by buying and combining ingredients once one takes account of transactions costs. Nor are there gains to be had from storing ketchup, or mixing together different quality ketchups and selling the resulting product. Indeed, most ketchup economists regard the efficiency of the ketchup market as the best established fact in empirical economics."

Secondly, there is a difference in methodology. In general economics the concepts of demand and supply are still fundamental to economic modelling, and models are more often than not closed by setting prices such that supply equals demand. In financial economics the perspective is a different one; the intuition here is the force of price and not the forces of demand and supply. Financial economists are fortunate enough to have at their disposal many comparable data on nearly all financial assets (the "goods") and for all frequencies. These financial assets essentially trade in the same market and the data are primarily price data. Furthermore financial markets are highly liquid and

"as close to our purely competitive ideal as one can find in the real world. In such an environment prices determine actions and quantities are secondary". Ross (1987).

These considerations can explain why the theory of financial economics is one of price determination, often assuming inelastic supply. The force of price owes its importance to the notion that assets, or combinations of assets, serve as nearly perfect substitutes for each other, resulting in perfectly elastic demand curves. If the price is not the equilibrium price, there will be an infinite difference between supply and demand, an arbitrage situation. To summarize the preceding discussion I quote Ross (1987) once again :

"The ordinary demand and supply curves in competitive economies are drawn under the traditional assumption that other prices are held constant. In neoclassical finance the resulting demand curves are horizontal and perfectly elastic and the supply curves are either perfectly elastic or perfectly inelastic, depending on the problem being studied. What matters in such a situation is not movements along the curves in response to changes in price - such "responses" are unbounded - but, rather, where the curves are in the price-quantity picture. Unlike what occurs when elasticities are in the normal ranges, *everything* of interest is underneath the supply and demand picture and the picture is meaningless. The forces of supply and demand have no meaning, since if the price is not the equilibrium price, then the difference between supply and demand is

infinite. This is precisely what is meant by an arbitrage situation, and it is so qualitatively different from the economist's usual picture of demand and supply as to require a different approach."

Although financial economics has a different focus and employs a different methodology from general economics it remains a subdiscipline of economics as it studies economic phenomena (see section 1). The finance approach to financial markets should be viewed as a complement to the approach taken by general economists, and as complementary approaches neither one of them has a sole claim on being "the real McCoy". Rather each approach has something to offer to the other. Financial economics obviously has borrowed heavily from other economic subdisciplines, but over time the reverse has occurred as well. One could, for instance argue that the early work on efficient markets within finance, provided macroeconomics with a starting point for the development of the concept of rational expectations, beyond Muth's (1961) original article. Contingent claims analysis and the general arbitrage theory were developed within finance and are now generally accepted concepts within economics. Furthermore financial economics can provide extensive databases to test hypotheses, not necessarily developed within finance. Finally both approaches share a common interest in explaining asset prices and the use of econometric methods. I think a case can be made for narrowing the gap between economics and finance, in order to increase the number of fruitful cross fertilizations, see Gibbons (1987), but not for closing the gap, in order to guarantee financial economics its isolated and eccentric position which has been so successful, see Ross (1987).

There is at least one area though, in which financial economics has unjustifiably strayed behind general economics. In their preoccupation with the interrelationships between the prices of different financial assets, financial economists have, until recently, ignored the time-series properties of asset prices. As Summers (1985) points out :

"It would surely come as a surprise to a layman that virtually no mainstream research in the field of finance in the last decade has attempted to account for the stock market boom of the 1960's or the spectacular decline in real stock prices during the mid-1970s."

Gibbons (1987) takes this argument several steps further and critically comments :

"Similarly, financial economists have not tried to explain the dramatic changes in the real rates of interest that have been observed in the 1980's. The study of the properties of these time-series is a missed opportunity. While recent empirical work is attempting to rectify this oversight by modeling the changing conditional moments of the distribution of asset returns, it is difficult to understand why the work did not start sooner. I have only two explanations. First, financial theory has only recently provided satisfactory general equilibrium models where the stochastic process for asset prices is an endogenous outcome from assumptions about consumption and production (for example, see John Cox et al., 1985). Until recently, econometric analysis lacked theoretical guidance on how to model changing conditional moments in a multi-period model of uncertainty. Second, I believe, but cannot prove, that a large number of empirical researchers in finance operated as if efficient markets require conditional expected returns to remain constant. Certainly, they paid lip service to the notion of a joint hypothesis of efficient markets and a correct equilibrium model, but most tests in practice assumed constant conditional expected returns as the equilibrium model. The uncovering of "anomalies" has forced empirical researchers to entertain the notion that conditional expected returns may change in an efficient market. Hopefully, future empirical research in finance will give as much attention to the equilibrium time-series properties of asset prices as past research has given to the cross-sectional characteristics."

Gibbons' argument exactly conveys the first basic motivation and objective for this dissertation. A large part of the focus of this dissertation will be on the time series properties of certain financial assets, where conditional moments of the assets' distributions are allowed to vary over time.

Financial decisions, nearly always entail a degree of risk. This risk is an integral part of the financial decision making process. The most prominent measures of risk used in asset pricing theories are the variance of the return of an asset, which is used in option pricing theory, and the covariances between the future return on an asset and various benchmark stochastic variables.

1.3 Time-Varying Risk in Financial Economics

In Section 1.1 it was mentioned that financial economists already recognized the time variation in the uncertainty in speculative prices, as early as the 1960's. Fama (1965), Mandelbrot (1963) and Samuelson (1965) have shown that variances and covariances of security price changes are changing through time. Consequently asset pricing theories must account for time-varying second moments, i.e. time-varying risk. If the risk of a certain speculative asset is changing through time, so should the compensation for bearing that risk, the expected return on that asset. The importance of incorporating time-varying first-, second-, and possibly higher-order moments into asset pricing theories cannot be over-emphasized. Time-varying second moments imply, time-varying betas, time-varying hedge ratios, option pricing models with stochastic volatility, and time-varying risk premia in general. The obvious question that presents itself is : How to accomplish this task ? However before time-varying moments can be fruitfully incorporated into asset pricing models - which is the second objective of this dissertation - we will have to develop statistical models that are capable of describing the evolution over time of these moments. One could argue that models for time varying second moments should reflect the empirically observed volatility clustering in financial data. As noted by Mandelbrot (1963) :

"Large changes tend to be followed by large changes - of either sign - and small changes tend to be followed by small changes."

One could also argue that the success of modern time series lies in the use of conditional rather than unconditional forecasts. Forecasts based on recent or all available information are more efficient than forecasts which do not use this information. This means that we should opt for models that use recent information in generating volatility forecasts. The Autoregressive Conditional Heteroskedasticity (ARCH) model, developed by Engle (1982), meets both requirements. Since its introduction the ARCH model and its generalized version, the GARCH model developed by Bollerslev (1986) and Taylor (1986), have gained immense popularity. The list of articles applying a (G)ARCH structure to financial time series is immense, see Bollerslev, Chou and Kroner (1992) for an excellent survey. It is the most prominent tool to describe the evolution of conditional variances in financial time series.

1.4 (G)ARCH : A Tool for Time-Varying Risk

In this section we will briefly outline the parameterization of GARCH processes, as a convenience to the reader. This dissertation does not focus on theoretical issues surrounding GARCH and its extensions, but rather will perform a critical analysis of the empirical implementation of GARCH processes in financial economics. Where necessary, theoretical issues will be dealt with in the individual chapters. Basically GARCH processes are discrete time stochastic processes $\{\epsilon_t\}$ of the following form :

$$\epsilon_t = z_t \sigma_t \quad , \quad z_t \text{ i.i.d.} \quad , \quad z_t \sim D(0,1) \quad 4.1$$

$$\sigma_t^2 = \omega + \sum_{i=1}^q \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 \quad 4.2$$

The conditional variance of ϵ_t depends, as can be seen from equation 1.2, on lagged squared realizations of ϵ_t and lagged conditional variances, and D is a symmetric distribution with mean zero and unit variance. This is the linear GARCH(p,q) model, which will appear in all analyses comprising this dissertation. One obtains the ARCH(q) formulation whenever $p=0$. It must be noted that although a GARCH model is set in discrete time, it will still produce the correct estimate of the conditional variance when the underlying process is a diffusion process with stochastic volatility, if the time interval is short enough, see Nelson (1992). This is a reassuring fact since much of modern finance theory is cast in continuous time, but financial time series are only available at discrete time intervals. To ensure a well defined process all parameters in the infinite-order AR representation :

$$\sigma_t^2 = \phi(L) \epsilon_t^2 = (I - \beta(L))^{-1} \alpha(L) \epsilon_t^2 \quad 4.3$$

must be non-negative, where it is assumed that the roots of the polynomial $\beta(\lambda)=I$ lie outside the unit circle, see Nelson and Cao (1991) and Drost and Nijman (1991). Basically the GARCH(p,q) is a model for financial volatility,

but it can be readily extended¹ into a (G)ARCH-in-mean model, developed by Engle, Lilien and Robins (1987), which is often used in an asset pricing context :

$$y_t = f(x_{t-1}, \sigma_t^2; b) + \varepsilon_t \quad 4.4$$

Equations 1.1, 1.2, and 1.4 form the GARCH-in-mean process. We see that the conditional mean of $\{y_t\}$ is given by a function of variables in the information set at time $t-1$, denoted by x_{t-1} , the conditional variance of the process, and a parameter vector b . The GARCH-in-mean model is often used in issues surrounding the risk-return trade-off, which is central to many theories in finance.

1.5 Research Objectives

In this dissertation we aim to provide a critical evaluation of the potential benefits from using GARCH models. We will discuss the use of a GARCH model as a statistical tool to describe conditional heteroskedasticity in the stock market and in the foreign exchange market. We will also analyze the use of GARCH models as a tool to determine time-varying risk premia in an asset pricing context. Again the focus will be on the stock and the foreign exchange market. Finally we will suggest new routes on which empirical research in financial economics, using GARCH models could embark.

More specifically, in Chapter 2 we will focus on the choice of an appropriate conditional distribution for a GARCH process for a broad set of weekly and

¹ Other extensions and alternative parameterizations include for example the multivariate ARCH and GARCH models by Kraft and Engle (1983) and Bollerslev, Engle and Wooldridge (1988) respectively. The asymmetric, exponential GARCH model by Nelson (1990), the threshold ARCH model by Zakoian (1990), the qualitative threshold ARCH model by Gouriéroux and Montfort (1992), and the structural ARCH model by Harvey, Ruiz and Sentana (1992). In addition to these parametric extensions, semi-parametric and non-parametric estimation techniques have appeared in the literature, see e.g. Gallant and Tauchen (1989), Engle and Gonzalez-Rivera (1991), and Pagan and Ullah (1988).

monthly stock indices returns. We will employ several conditional distributions and we will show that for high and moderate frequency returns conditional normality is simply inappropriate. This result will have severe implications for multivariate analyses, which until now have maintained the hypothesis of conditional normality. We also address the questions whether conditional heteroskedasticity of the GARCH form will disappear when the data are aggregated, and whether the unconditional distribution of rescaled innovations will resemble a normal distribution when the data are aggregated more and more. Aggregation from a weekly to a monthly level appears to be insufficient to provide an affirmative answer to both questions. Unconditional normality for financial returns is a relict from the past.

Chapter 3 contains the theoretical analysis in this dissertation and considers the incorporation of a statistical model for conditional heteroskedasticity, i.e. GARCH, in an intertemporal asset pricing model for stock returns. The result is quite surprising. We obtain a model reminiscent of the three moment CAPM, and not the familiar GARCH-in-mean model. The resulting model suggests that we should not only study time-varying second moments but should also allow for time-varying third order moments. Risk averting agents do not only demand a higher expected return for bearing higher risk, but are also willing to give up part of the expected return in the light of positive skewness in the distribution of returns.

A first empirical analysis for the model developed in Chapter 3 is performed in Chapter 4. We will contrast the empirical results for this model, based on the same dataset that was used in Chapter 2, with the results for a GARCH-in-mean model upon which it tries to improve. In order to allow for time-varying conditional third moments we introduce the Generalized Auto Regressive Conditional Skewness (GARCS) model. The results of our empirical analysis show that our version of the three moment CAPM, as opposed to the GARCH-in-mean model, generates significant estimates for the coefficient of relative risk aversion. This can be explained by the lack of intertemporal components in the GARCH-in-mean model. Although the preliminary results are encouraging, much work remains to be done. Nevertheless, the analysis in Chapter 4 provides a basis for future research into the relationship and trade-off between variance and skewness.

Another area in which numerous GARCH applications have found their way into the literature pertains to foreign exchange returns. International portfolio management depends on expected exchange rate movements through time. Policy issues concerning the impact of the exchange rate on different macroeconomic variables are equally important. Therefore a thorough understanding of the exchange rate dynamics is required, and hence the popularity of GARCH applications. Most studies, concerned with these issues, focus on U.S. Dollar exchange rates. Relatively little attention has been paid to bilateral European exchange rates. In order to partially correct for this oversight we will focus on EMS exchange rates. Chapter 5 has more or less the same objective as Chapter 2 had. We will study the statistical properties of exchange rate returns within the Exchange Rate Mechanism of the EMS. As opposed to free-floating currencies, EMS exchange rates contain the risk of realignment. Ignoring this risk factor, which we will measure by a jump process, leads to a misinterpretation of the true conditional risk, and hence to the mispricing of assets. We find that EMS exchange rate returns are characterized by jumps, conditional heteroskedasticity of the GARCH form and conditional fat tails. These features are simultaneously captured by a combined jump-GARCH model with conditionally Student t-distributed innovations.

Chapter 6 is, as Chapter 3 was, concerned with time-varying risk premia for financial asset returns, and considers the conditional variance of foreign exchange rate returns, given by a GARCH process, as a determinant of the risk premium in the forward foreign exchange market. In the international financial economics literature it is agreed upon that the forward rate is not an unbiased predictor of the corresponding future spot rate. This rejection of forward market efficiency can be explained by the irrationality of market participants, by the existence of a time-varying risk premium or by a combination of both phenomena. A survey data set in which direct expectations of future exchange rates by market participants were formed, relieves us from the rational expectations assumption. Several theoretical models have been put forward which generate risk premia in forward foreign exchange markets. In general these risk premia, in some form or another, depend on the conditional probability distribution of the future spot rate. This provides the rationale for considering the conditional variance as a determinant of a time-varying risk premium, i.e. considering the GARCH-in-mean model. Our

results, for a wide variety of EMS exchange rates and for three different forecast horizons, indicate that the GARCH-in-mean model is reasonably successful in explaining the nature of the time-varying risk premium. Chapter 7 will put the results of our analyses in perspective and provides summarizing, concluding and forward-looking remarks.

Chapter 2

Conditional and unconditional distributions for stock returns

2.1 Introduction

Empirical distributions of asset returns have been studied for almost three decades now. Probably the most important factor that has generated this considerable interest is the fact that these distributional properties have a direct impact on the validity of theoretical models in financial economics.

The foundations for the literature on asset returns distributions that is currently available, have been laid in Mandelbrot (1963) and Fama (1965). Both authors have suggested the stable Paretian class of distributions to describe the distributional properties of asset returns. Although this class of distributions exhibits some attractive features such as time-additivity and leptokurtosis (caused by an infinite population variance), the empirical evidence pertaining to the validity of these distributions for asset returns is mixed. Mandelbrot (1963) and Fama (1965), for stock returns, and Westerfield (1977) and McFarland, Pettit, and Sung (1982), for exchange rate returns, have argued in favor of the class of stable distributions as the distributions to use in describing asset returns. Blattberg and Gonedes (1974), however, find that a Student t-distribution provides a better description of the empirical distribution of stock returns than the stable distributions. Akgiray and Booth (1988b) also reject the stable law model for stock

returns, finding that empirical tail shapes are thicker than normal but thinner than stable tails, which suggests a fat-tailed finite variance distribution. For exchange rate returns Boothe and Glassman (1987) find that a Student t-distribution and a mixture of two normal distributions both provide a better fit than a stable distribution. The opposite conclusions for EMS exchange rates are reached by Koedijk, Schafgans and de Vries (1990), who use extreme value theory. These findings are not inconsistent with Friedman and Vandersteel's (1982) suggestion that returns are generated by a time-varying distribution. This notion provides the link to a second branch in the literature on returns distributions.

To account for the observed leptokurtosis in the data, this second view proposes a conditional distribution that is time varying, instead of an unconditional fat-tailed distribution that is fixed over time. Furthermore this view incorporates the Mandelbrot (1963) observation that there appear to be clusters of high and low absolute values in the return data, and therefore clusters of high and low volatility. This of course means that the data are not independently generated.

The first model to capture this clustering phenomenon, while simultaneously generating unconditional fat tails, is the Autoregressive Conditional Heteroskedasticity (ARCH) model, introduced by Engle (1982) and generalized (GARCH) by Bollerslev (1986). The essential feature of a GARCH process is a time varying conditional variance, parameterized as a linear function of past squared innovations and lagged conditional variances. This conditional approach has gained a great deal of popularity with researchers in econometrics and finance, because of its ability to generate the stylized facts observed when studying asset returns distributions. A recent survey of the state of the art in GARCH modelling in finance is given in Bollerslev, Chou, and Kroner (1992). As indicated above, the choice of the appropriate unconditional return generating distribution remains an unresolved issue.

This problem carries over to the conditional or GARCH approach where a choice regarding the conditional distribution has to be made. For high frequency returns it seems that a conditionally normal distribution can be safely rejected. This because of its inconsistency with the conditional leptokurtosis in high frequency returns, see e.g. Hsieh (1989). Bollerslev (1987)

advocates a scaled t-distribution, in order to separate conditional heteroskedasticity from a conditional leptokurtic distribution, either of which could explain the unconditional leptokurtosis in the return data. A wide variety of conditional distributions has been proposed in the empirical literature, but has predominantly been applied to exchange rate returns. Baillie and Bollerslev (1989) successfully use a Student t- and a general error distribution. Hsieh (1989) also uses a normal-Poisson and normal-lognormal mixture distribution, of which the former is also employed by Jorion (1988) and Nieuwland, Verschoor and Wolff (1991). For stock returns the most commonly used conditional distributions are the normal and the scaled t-distribution. Notable exceptions are again Jorion (1988) and Nelson (1991) who uses a generalized error distribution.

This chapter is concerned with alternative distributional specifications in GARCH processes for returns on stock indices. It will be shown that conditional normality can be rejected both for weekly as well as monthly data, against a scaled t- or a generalized error distribution (GED). Furthermore it is concluded that monthly stock index returns do not converge to normality, as is commonly assumed for low frequency data. Finally the scaled t-distribution must be considered to be the preferred choice when selecting a conditional distribution. The second objective in this chapter is to discuss the choice of an unconditional distribution for stock index returns. As we know that GARCH generates unconditional fat tails, we will have to consider unconditional fat-tailed distributions. We restrict ourselves to two alternatives: a stable law and a scaled Student t-distribution with few degrees of freedom. Based on our analysis of tail indices using extreme value theory, we opt for a scaled Student t-distribution with few degrees of freedom.

The chapter now proceeds as follows : Section 2.2 presents the data and some summary statistics. Tests for conditional heteroskedasticity are discussed in section 2.3. The models to be estimated are presented in section 2.4, and the actual estimation results are contained in section 2.5. Section 2.6 provides some diagnostic checks, section 2.7 discusses the effects of temporal aggregation and the appropriate unconditional distribution, and section 2.8 presents a conclusion.

2.2 Data and summary statistics

In this chapter stock indices from five countries, with well-developed stock markets will be used. The countries are the United Kingdom (UK), the United States (USA), the Netherlands (NET), Germany (GER) and Japan (JAP). The indices used are the DATASTREAM¹ Total Market Indices, which contain all stocks quoted on the respective exchanges². These indices are value weighted and are adjusted for stock dividends, capital modifications and the like. The data were collected on a weekly ($T=923$) and monthly ($T=211$) basis, for a period starting on January 2nd 1973 and ending on September 18th 1990, spanning more than 17 years. The quotes used are Tuesdays' closing prices. Continuously compounded returns are used, and are defined as the difference in logarithmic value of two consecutive observations.

$$R_t = \ln(P_t) - \ln(P_{t-1}) \quad 2.1$$

In Tables 2.1 and 2.2 summary statistics for respectively weekly and monthly returns are presented. One can observe that the mean returns are significantly different from zero at the 5% level for the U.K. and Japanese weekly returns, and at the 1% level for the Japanese monthly returns. All returns series display substantial skewness and kurtosis. Under normality, and assuming no autocorrelation, these measures would have asymptotic distributions of $N(0,6/T)$ and $N(3,24/T)$ ³. Further evidence against normally distributed returns is given by the Bera-Jarque test-statistic, see Bera and Jarque (1982), which is a joint test using both skewness and kurtosis, and the Kiefer-Salmon Normality tests for skewness and kurtosis, see Kiefer and Salmon (1983)⁴. In order to test whether returns are autocorrelated or not, the first 50 autocorrelations were computed, together with the Ljung-Box

¹ DATASTREAM is a U.K. incorporated data service company.

² London , New York , Amsterdam , Frankfurt , Tokyo.

³ T is the total number of observations.

⁴ The Bera-Jarque test is asymptotically $\chi^2(2)$ distributed, and the Kiefer-Salmon tests are asymptotically $\chi^2(1)$ distributed.

(1978) statistic for serial correlation :

$$LB(p) = T(T+2) \sum_{\tau=1}^p (T-\tau)^{-1} \hat{\rho}^2(\tau) \stackrel{a}{\approx} \chi^2(p) \quad 2.2$$

As the choice of the appropriate laglength is somewhat arbitrary, both LB(50) and LB(25) were computed. Diebold (1988) showed that in the presence of ARCH effects, the Ljung-Box test has a larger empirical size than a nominal test size of 5%, because the asymptotic variance of the autocorrelations under ARCH is larger than under the null of Gaussian white noise⁵. Diebold proposes a variant on the Ljung-Box statistic which is adjusted for conditional heteroskedasticity and which preserves the proper size.

$$D(p) = T(T+2) \sum_{\tau=1}^p (1 + (\gamma^2(\tau)/\sigma^4))^{-1} (T-\tau)^{-1} \hat{\rho}^2(\tau) \stackrel{a}{\approx} \chi^2(p) \quad 2.3$$

Inspection of the results in the tables reveals that any evidence of serial correlation using the standard Ljung-Box test vanishes, when allowing for conditional heteroskedasticity⁶. The first 5 individual autocorrelations with corresponding robust standard errors are also given. Only for the U.K. and German weekly returns there is some evidence of the presence of first order autocorrelation, but not for higher order autocorrelation. In contrast to these results, the Ljung-Box statistics for the absolute ($|R|$) and squared (R^2) returns are highly significant for the weekly data (except for the U.S.), indicating strong non-linear dependence, conditional heteroskedasticity and the aforementioned clustering phenomenon. For the monthly data one can only detect significant serial autocorrelation for the absolute returns of The U.K., Germany and Japan.

⁵ $\text{var}(\hat{\rho}(\tau)) = (T-\tau)/(T(T+2))$ under Gaussian white noise ; under ARCH this variance needs to be multiplied by $(1 + (\gamma^2(\tau)/\sigma^4))$ where $\gamma^2(\tau)$ is the τ -th order autocovariance of the squared process.

⁶ None of the D(25) and D(50) values is above the 95% critical values of 37.65 and 67.50.

Table 2.1 : Summary Statistics Weekly Returns (1/2/73 - 9/18/90)

	UK	USA	NET	GER	JAP
mean ¹	1.717	1.003	1.077	1.018	1.688
st.dev. ²	2.731	2.348	2.152	2.035	2.057
skewness	-0.188*	-1.734**	-1.489**	-1.382**	-1.456**
kurtosis	16.40**	23.49**	18.83**	13.69**	13.86**
BJ-test	6912**	16617**	9977**	4688**	4858**
KS-1	5.48*	462.64**	341.42**	293.69**	326.09**
KS-2	6907**	16154**	9635**	4394**	4532**
LB(50) R	90.08**	61.80	42.95	75.51*	63.69
D(50) R	53.36	51.33	35.45	41.39	45.92
LB(50) R	351.94**	157.63**	225.45**	624.80**	545.37**
LB(50) R ²	116.62**	14.38	92.26**	260.67**	125.66**
LB(25) R	69.73**	38.05*	18.58	48.63**	30.07
D(25) R	34.42	28.56	11.82	21.43	17.85
LB(25) R	323.88**	96.19**	186.52**	454.97**	378.25**
LB(25) R ²	113.97**	9.32	86.65**	224.49**	95.29**
$\rho(1)$	0.158 (0.051)	-0.005 (0.034)	0.014 (0.047)	0.109 (0.067)	-0.002 (0.042)
$\rho(2)$	0.115 (0.059)	-0.020 (0.050)	0.009 (0.047)	0.069 (0.052)	0.038 (0.048)
$\rho(3)$	0.093 (0.053)	0.038 (0.038)	0.090 (0.071)	0.116 (0.078)	0.087 (0.060)
$\rho(4)$	0.011 (0.048)	-0.023 (0.030)	-0.016 (0.069)	0.031 (0.049)	0.008 (0.037)
$\rho(5)$	0.014 (0.045)	-0.048 (0.036)	-0.006 (0.035)	-0.030 (0.036)	0.035 (0.041)

¹ mean return $\times 10^3$, ² standard deviation $\times 10^2$, * denotes significance at the 5% level.

** denotes significance at the 1% level, BJ-test denotes the Bera-Jarque test for normality; KS-1 and KS-2 denote the Kiefer-Salmon Normality test for respectively skewness and kurtosis. LB(p) denotes the Ljung-Box test for serial correlation using p lags. D(p) denotes the Diebold test for serial correlation using p lags. $\rho(p)$ denotes p-th order autocorrelation, robust standard errors are given in parentheses.

Table 2.2 : Summary Statistics Monthly Returns (1/2/73 - 9/18/90)

	UK	USA	NET	GER	JAP
mean ¹	8.187	4.649	4.869	4.724	8.019
st.dev. ²	6.960	4.674	4.691	4.798	4.521
skewness	0.652**	-0.700**	-0.783**	-0.797**	-0.573**
kurtosis	12.82**	7.02**	9.99**	8.71**	5.61**
BJ-test	863**	159**	452**	309**	71**
KS-1	14.96**	17.24**	21.59**	22.39**	11.54**
KS-2	848**	142**	430**	287**	60**
LB(50) R	41.98	34.35	52.56	64.44	53.95
D(50) R	32.48	34.98	48.96	52.80	41.97
LB(50) R	68.14*	44.82	45.19	85.71**	105.29**
LB(50) R ²	15.01	17.45	15.96	37.49	65.29
LB(25) R	29.67	21.33	35.49	40.17*	20.63
D(25) R	19.43	21.19	30.55	32.24	17.60
LB(25) R	62.87**	26.77	25.20	55.34**	60.22**
LB(25) R ²	13.88	11.85	10.82	26.19	30.24
$\rho(1)$	0.132 (0.115)	0.036 (0.089)	0.101 (0.080)	0.107 (0.097)	0.105 (0.093)
$\rho(2)$	-0.095 (0.107)	-0.032 (0.077)	0.009 (0.054)	-0.045 (0.061)	-0.038 (0.089)
$\rho(3)$	0.053 (0.099)	0.037 (0.077)	0.020 (0.063)	0.105 (0.085)	-0.032 (0.079)
$\rho(4)$	0.072 (0.085)	-0.020 (0.075)	-0.105 (0.090)	-0.003 (0.093)	-0.047 (0.080)
$\rho(5)$	-0.107 (0.089)	0.054 (0.062)	0.045 (0.059)	-0.067 (0.076)	0.082 (0.096)

¹ mean return $\times 10^3$, ² standard deviation $\times 10^2$, * denotes significance at the 5% level.
 ** denotes significance at the 1% level, BJ-test denotes the Bera-Jarque test for normality; KS-1 and KS-2 denote the Kiefer-Salmon Normality test for respectively skewness and kurtosis. LB(p) denotes the Ljung-Box test for serial correlation using p lags. D(p) denotes the Diebold test for serial correlation using p lags. $\rho(p)$ denotes p-th order autocorrelation, robust standard errors are given in parentheses.

Furthermore there seems to be no indication of conditional heteroskedasticity in the monthly returns⁷. To remove possible first order autocorrelation from the returns series and to permit a distinction of correlation effects from heteroskedasticity effects, the following OLS regression was estimated:

$$R_t = \varphi_0 + \varphi_1 R_{t-1} + \varepsilon_t \quad 2.4$$

Inspection of the residuals series reveals that all autocorrelation, both individually as well as serially, has been eliminated. The remaining statistics for the $\{\varepsilon_t\}$ series are very similar to the raw data statistics, therefore and for reasons of space these statistics are not reported.

2.3 Conditional heteroskedasticity tests

In order to test whether ARCH effects are present in the data two different approaches are employed. First the standard Lagrange Multiplier-tests are performed and secondly a test using homogeneous Markov chains is applied. Both tests are easy to calculate, only require estimation under the null hypothesis of no ARCH and are appropriate under all distributional assumptions⁸. However, as will become clear, they lead to quite different results. Consider the following specification :

⁷ Note however that in most cases the LB(50) R^2 statistic does not differ much from the LB(25) R^2 statistic, so that longer lags do not contribute much to the value of the statistic. This begs the question whether including fewer lags would lead to significant results. Experiments with $p=12$ and $p=6$ in general do not lead to different results, except for the U.S. LB(6) $|R|$ and the Japanese LB(12) and LB(6) R^2 statistics, which are significant at the 5% level.

⁸ Weiss (1986a) has shown that the proposed LM-test is appropriate for non-normal distributions, provided some moment conditions are satisfied. The Markov-chain test is completely distribution free.

$$\begin{aligned}
 R_t | \psi_{t-1} &\sim D(\varphi_0 + \varphi_1 R_{t-1}, h_t) \\
 h_t &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \dots + \alpha_p \varepsilon_{t-p}^2 \\
 \varepsilon_t &= R_t - \varphi_0 - \varphi_1 R_{t-1}, \quad t = 1, \dots, T
 \end{aligned} \tag{3.1}$$

Where ψ_{t-1} is the information set available at time $t-1$, and D is a symmetric continuous distribution with finite second and fourth moments. This is the ARCH regression model. The LM test for ARCH effects can be constructed along the lines of the following procedure. First, estimate 2.4 by OLS and save the residuals. Then regress the squared residuals on a constant and p lags. Finally, compute TR^2 , where R^2 is the coefficient of determination of the squared residuals regression. Under the null of homoskedasticity (i.e. no ARCH)⁹, TR^2 is asymptotically χ^2 distributed with p degrees of freedom. Although this version of the LM test is appropriate for any distributional assumption, see Koenker (1981) and Weiss (1986a), its statistical properties under other conditional distributions than the Normal¹⁰ are far from clear.

As an alternative approach to construct a test for ARCH effects, a non-parametric test based on finite-state Markov chains is considered here. Gregory (1989), using Monte Carlo experiments, concludes that this test is superior to the LM test for various distributions other than the Normal, in terms of better finite sample properties like size and power. The construction of the test¹¹ moves along the following lines: the first step is to obtain the squared residuals from the estimation of 2.4. The next step is to apply a discretization rule, by which the squared residuals are divided into different categories or states. The rule applied here is to mark the residuals as being

⁹ $H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_p = 0$

¹⁰ The finite-sample properties of this LM test under conditional normality are investigated in Engle, Hendry and Trumble (1985), and appear to be reasonable in terms of empirical size and power.

¹¹ This test is equivalent to a standard test for independence in a two-way contingency table, and could also be used to test for more general forms of conditional heteroskedasticity.

high or low, with the sample mean¹² as the boundary. This results in a two-state definition: the residuals are either low (state 1) or high (state 2)¹³. Next, it is assumed that $\{e_t^2\}$, which values are now either 1 or 2, possesses the Markov property, that says that the probability distribution of e_t^2 conditional on its entire past equals the probability distribution conditional on its first previous value only:

$$P(e_t^2 = j \mid e_{t-1}^2 = i, e_{t-2}^2, e_{t-3}^2, \dots) = P(e_t^2 = j \mid e_{t-1}^2 = i) = \lambda_{ijt}$$

$$i, j = 1, 2 \quad t = 2, \dots, T \quad 3.2$$

λ_{ijt} is the probability of being in state j at time t , coming from state i at time $t-1$. These transition probabilities are assumed to be time-independent¹⁴, so now $\{e_t^2\}$ is defined as a homogeneous first order Markov chain. Besides a first order chain also a second order homogeneous Markov chain will be used to test for second order ARCH effects. In this case the transition probability is defined as :

$$\lambda_{ijk} = P(e_t^2 = k \mid e_{t-1}^2 = j, e_{t-2}^2 = i) \quad i, j, k = 1, 2 \quad t = 3, \dots, T \quad 3.3$$

λ_{ijk} is the probability of being in state k at time t , coming from state j at time $t-1$ and from state i at time $t-2$ ¹⁵. For the first order and second order Markov Chains one can define the following transition probability matrices :

¹² The sample mean is the ML estimate of the variance under the assumption of no ARCH. This classification implies that under normality 68% of the observations should classify as low.

¹³ Other discretization rules leading to N-state definitions are of course possible, see Gregory (1989).

¹⁴ Note that in the remainder of the text the subscript t is deleted.

¹⁵ The Markov property can be retained by defining composite states (i, j) at time $t-1$ and (j, k) at time t .

						1	2
	1	2				λ_{111}	λ_{112}
1	λ_{11}	λ_{12}	and			λ_{121}	λ_{122}
2	λ_{21}	λ_{22}				λ_{211}	λ_{212}
						λ_{221}	λ_{222}

Where the previous states are given on the vertical axis and the current states on the horizontal axis. In order to estimate these Markov chains one needs the log likelihood functions in terms of the transition probabilities. For the first order Markov chain the likelihood is taken conditional on the first observation, and for the second order Markov chain conditional on the first and second observation¹⁶. The likelihood function for the first order case becomes :

$$LLM1 = \sum_{i=1}^2 \sum_{j=1}^2 n_{ij} \ln(\lambda_{ij}) \quad 3.4$$

The Maximum Likelihood Estimator of λ_{ij} is easily acquired and is $\hat{\lambda}_{ij} = n_{ij} / (n_{i1} + n_{i2})$; n_{ij} is the number of times that a transition from i to j is observed. Define $N_i = n_{i1} + n_{i2}$, as the number of times state i is observed, then the log likelihood value at the MLE becomes :

$$\overline{LLM1} = \sum_{i=1}^2 \sum_{j=1}^2 n_{ij} \ln(n_{ij} / N_i) \quad 3.5$$

The likelihood function for the second order Markov chain is :

$$LLM2 = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^2 n_{ijk} \ln(\lambda_{ijk}) \quad 3.6$$

¹⁶ Ignoring the initial states is of no significant consequence when the total number of observations is large, as is the case considered here, see Neftci (1984). Furthermore by ignoring the initial states a non linear optimization problem is simplified to a linear one.

Now n_{ijk} is the number of transitions from i to j to k . The MLE of λ_{ijk} is, in analogy to the first order case defined as : $\hat{\lambda}_{ijk} = n_{ijk} / (n_{ij1} + n_{ij2})$. Define $N_{ij} = n_{ij1} + n_{ij2}$ as the number of times that a transition from i to j occurred. The log likelihood value at the MLE is :

$$\overline{LLM2} = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^2 n_{ijk} \ln(n_{ijk} / N_{ij}) \quad 3.7$$

These likelihood values are necessary to develop tests for the null hypothesis of homoskedasticity. If this hypothesis were true , then there would be no dependence in the squared residuals and the transition probabilities would be without any structure. This means that the probability of observing a current state k , λ_k , is independent of previous realizations (i and j). This translates into :

$$H_0 : \lambda_{ij} = \lambda_j \quad 3.8$$

When testing independence against a first order Markov chain and into :

$$H_0 : \lambda_{ijk} = \lambda_k \quad 3.9$$

When testing independence against a second order Markov chain. The MLE under both null hypotheses is : $\hat{\lambda}_j = N_j / N$; N is the total number of observations used in the estimation procedure¹⁷. The log likelihood value at the MLE is:

$$\overline{LLI} = \sum_{i=1}^2 N_i \ln(N_i / N) \quad 3.10$$

Three Likelihood Ratio tests can now be constructed : The first one is a test of independence against a first order Markov chain :

¹⁷ $N = T - 2$.

$$LR (I / M1) = -2 (\overline{LLI} - \overline{LLMI}) \stackrel{a}{\sim} \chi^2(1) \quad 3.11$$

The second is a test of independence against a second order Markov chain :

$$LR (I / M2) = -2 (\overline{LLI} - \overline{LLM2}) \stackrel{a}{\sim} \chi^2(3) \quad 3.12$$

The last one is a test of a first order against a second order Markov chain :

$$LR (M1 / M2) = -2 (\overline{LLMI} - \overline{LLM2}) \stackrel{a}{\sim} \chi^2(2) \quad 3.13$$

The results for the LM tests for ARCH and for the Markov chain tests are given in Tables 2.3.a and 2.3.b. We observe that the tests in some cases lead to conflicting inferences regarding the independence of the squared residuals. In 9 out of 10 cases independence is rejected against a second order Markov chain, while the LM tests only reject independence in 5 cases. Furthermore in 9 out of 10 cases, a first order Markov chain is rejected against a second order Markov chain¹⁸. The evidence presented is consistent with Gregory's (1989) observation that the LM test is biased towards the null hypothesis of no ARCH for non-normal conditional distributions. Relying solely on an LM test appears to be inappropriate.

¹⁸ Stability tests for the Markov chains show that the first order Markov chain is stable in 7 out of 10 cases and the second order chain in 6 out of 10 cases.

Table 2.3.a : ARCH tests for weekly data

	UK	US	NET	GER	JAP
LM(1)	2.21	0.02	3.34*	74.81***	2.14
LM(2)	17.94***	3.08	6.25**	75.00***	9.38***
LM(5)	29.73***	3.43	71.37***	174.5***	37.63***
LRIM1	3.37*	3.37*	3.87**	17.03***	11.52***
LRIM2	28.76***	11.08**	11.55***	25.37***	23.19***
LRM1M2	25.40***	7.71**	7.69**	8.34**	21.67***

Table 2.3.b : ARCH tests for monthly data

	UK	US	NET	GER	JAP
LM(1)	0.56	2.03	0.02	1.01	11.31***
LM(2)	3.51	2.34	0.23	1.37	13.93***
LM(5)	4.59	2.90	1.63	5.63	22.07***
LRIM1	3.99**	0.00	0.16	2.76	7.69***
LRIM2	11.53***	9.96**	8.22**	3.41	17.69***
LRM1M2	7.55**	9.96***	8.06**	0.65	10.00***

* , ** , *** , denote significance at respectively the 10%, 5% and 1% level. The LM(p) tests are computed as TR^2 from a regression of squared OLS residuals on a constant and p lags, and are asymptotically $\chi^2(p)$ distributed. LRIM1 is a Likelihood Ratio test of independence against a first order Markov Chain, and is distributed as $\chi^2(1)$. LRIM2 is a Likelihood Ratio test of independence against a second order Markov Chain, and is distributed as $\chi^2(3)$. LRM1M2 is a Likelihood Ratio test of a first order against a second order Markov Chain, and is distributed as $\chi^2(2)$.

2.4 GARCH models and the choice of a conditional distribution

The results of the analysis in the previous section indicate that weekly and monthly stock indices returns suffer from severe leptokurtosis, possibly generated by conditional heteroskedasticity, are approximately uncorrelated, and exhibit the clustering phenomenon, as is shown by the Markov chain tests. These three stylized facts set the stage for the following class of GARCH(p,q) regression models :

$$R_t = \varphi_0 + \varphi_1 R_{t-1} + \varepsilon_t \quad , \quad \varepsilon_t = \eta_t h_t^{1/2} \quad , \quad \eta_t \sim D(0,1) \quad 4.1$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \quad , \quad \alpha_0 > 0 \quad , \quad \alpha_i, \beta_j \geq 0 \quad 4.2$$

Here $D(0,1)$ can be any symmetrical continuous distribution with a zero mean and unit variance. The conditional variance, h_t is a linear function of squared lagged residuals and lagged conditional variances. This is an elegant means of modelling the occurrence of clusters in the conditional variances. Also any GARCH model can generate unconditional fat tails, see Engle (1982) and Bollerslev (1986). There are at least two ways to determine the optimal orders, p and q , of the GARCH process. The first one uses the notion that 4.2 can be written as an ARMA process in squared residuals. Define :

$$\xi_t = \varepsilon_t^2 - h_t \quad 4.3$$

which has mean zero and is serially uncorrelated, and rewrite 4.2 as :

$$\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^m (\alpha_i + \beta_i) \varepsilon_{t-i}^2 - \sum_{j=1}^p \beta_j \xi_{t-j} + \xi_t \quad 4.4$$

A GARCH(p,q) process can be viewed as an ARMA(m,p) process for the squared residuals, where $m = \max\{p,q\}$, see Pantula (1986). A traditional

Box-Jenkins (1976) analysis may now be performed to determine m , p , and possibly q if $m > p$. The alternative is to estimate 4.1 and 4.2 by Maximum Likelihood, under a wide number of specifications, i.e. different values for p and q , and choose the optimal orders according to a suitable selection criterion. This is the approach used here.

To implement ML estimation, D is chosen to be the standard normal distribution¹⁹. Specifications up to the order (2,2) were estimated for every returns series²⁰. For nested specifications the Likelihood Ratio Test (LRT) was used as a selection criterion and for non-nested specifications the Schwarz (1978) criterion (SC) was used.

$$LRT = -2 (\max LL(\theta_s) - \max LL(\theta_c)) \stackrel{a}{\sim} \chi^2(d) \quad 4.5$$

Here θ_s is the parameter vector of the simpler of the two specifications, and consequently θ_c is the parameter vector of the more complicated specification. This all under the assumption that the simple version is correctly specified, and d is the difference in the number of parameters of the two specifications.

$$SC = -2 \max LL(\theta_n) + n \ln(T) \quad 4.6$$

θ_n is a parameter vector containing n elements. The Schwarz criterion was chosen because it puts a large penalty on the inclusion of additional parameters and the selection rule is simple : choose the specification that gives the lowest value of SC. The results for both selection rules are hardly surprising: Homoskedasticity ($p=q=0$) is overwhelmingly rejected and the popular and parsimonious GARCH(1,1) specification²¹ is consistently

¹⁹ Misspecification errors are dealt with later on in the paper.

²⁰ Starting values for ϵ_1^2 , ϵ_2^2 , h_1 , and h_2 were set equal to the ML estimate of the variance under the assumption of homoskedasticity. Incorrect starting values however are no major problem, see Nelson (1991).

²¹ The list of articles using a GARCH(1,1) specification is almost endless; examples are Chou (1988), Akgiray (1989), and Baillie & DeGennaro (1990).

chosen as the appropriate specification²². Because of this consistency and also for reasons of space, these results are not reported. The model now reduces to :

$$R_t = \varphi_0 + \varphi_1 R_{t-1} + \varepsilon_t \quad , \quad \varepsilon_t = \eta_t h_t^{1/2} \quad , \quad \eta_t \sim D(0,1) \quad 4.7$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad 4.8$$

This model will be estimated under three alternative assumptions for the conditional error distribution. First of all the standard normal distribution is chosen, as it is the most commonly used in GARCH applications. Furthermore in the case of conditional normality, the conditional standard deviation can be interpreted as the stochastic volatility of the process, see Taylor (1990). This however has the implication that the volatility during a time period is known at the end of the previous time period, reducing the impact of unexpected news during the period on volatility to zero. There is also one important empirical drawback to the assumption of conditional normality. Standardized residuals from estimated GARCH-Normal processes often appear to be leptokurtic, which they should not be if normality were appropriate, see again Taylor (1990) and the references therein. The conditional density function for ε_t under normality is :

$$f_n(\varepsilon_t | \psi_{t-1}) = (2\pi h_t)^{-1/2} \exp(-\varepsilon_t^2/2h_t) \quad 4.9$$

$$\psi_{t-1} = \{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots\} \quad 4.10$$

By the prediction error decomposition, the associated log likelihood function becomes:

²² Note that these selection exercises are performed under the questionable assumption of conditional normality and should therefore be interpreted as giving some, but no absolute justification for the use of the GARCH(1,1) specification. At first glance it may seem strange to test for ARCH effects, and then estimate a GARCH model. Lee (1991) however has shown that a LM test for testing white noise disturbances against GARCH disturbances is equivalent to the LM test against ARCH disturbances in a linear regression model.

$$LL_n(\theta) = \frac{1}{2} \left(-T \ln 2\pi - \sum_{t=1}^T \ln h_t - \sum_{t=1}^T \varepsilon_t^2 / h_t \right) \quad 4.11$$

$$\theta = (\varphi_0, \varphi_1, \alpha_0, \alpha_1, \beta_1) \quad 4.12$$

Bollerslev (1987) proposed a scaled Student t-distribution, which has some theoretical and empirical advantages over a normal distribution. First of all it is conditionally fat tailed, a fact that - next to conditional heteroskedasticity - can explain unconditional fat tails. Secondly the scaled t-distribution is identical to a continuous variance mixture of normals, where the mixing variable follows an inverted Gamma-1 distribution depending on the degrees of freedom of the scaled t-distribution. A mixture distribution is desirable when information arrives randomly during a time period, see Tauchen and Pitts (1983), because exactly this random character of information is conveyed by the unobservable mixing variable. This is rather fortunate because now the stochastic volatility for period t is dependent on unexpected news, and is in general not equal to the conditional standard deviation for period t , which is known at the end of period $t-1$. Finally the scaled t-distribution is in contrast to the normal distribution consistent with the observed conditional leptokurtosis of GARCH residuals, see e.g. Baillie and DeGennaro (1990). The conditional density function for ε_t depends on a degrees-of-freedom parameter ν and is given by :

$$f_\nu(\varepsilon_t | \psi_{t-1}) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi} \Gamma\left(\frac{\nu}{2}\right)} ((\nu-2)h_t)^{-\frac{1}{2}} \left(1 + \frac{\varepsilon_t^2}{h_t(\nu-2)}\right)^{-\frac{(\nu+1)}{2}}, \nu > 2 \quad 4.13$$

The associated log-likelihood function reads :

$$LL_v(\theta) = T \left(\ln \Gamma\left(\frac{v+1}{2}\right) - \ln \Gamma\left(\frac{v}{2}\right) - \frac{1}{2} \ln(\pi(v-2)) \right) + \sum_{t=1}^T \left(-\frac{1}{2} \ln h_t - \left(\frac{v+1}{2}\right) \ln \left(1 + \frac{\epsilon_t^2}{(v-2) h_t} \right) \right) \quad 4.14$$

$$\theta = (\varphi_0, \varphi_1, \alpha_0, \alpha_1, \beta_1, v) \quad 4.15$$

The scaled t-distribution approaches normality when $v \rightarrow \infty$. Another distribution that contains normality as a special case, and that is used here as an alternative to the scaled t-distribution, is the General Error Distribution (GED) or the exponential power distribution, described in Box and Tiao (1973). The GED is also a continuous variance mixture of normals, and depending on the value of the characteristic parameter β , can be leptokurtic or platykurtic. The GED has been applied in GARCH processes for exchange rate returns by Hsieh (1989) and Baillie and Bollerslev (1989) and for stock returns by Nelson (1990). The conditional density function for ϵ_t under a GED reads :

$$f_\beta(\epsilon_t | \psi_{t-1}) = \omega(\beta) h_t^{-1/2} \exp\left(-c(\beta) |\epsilon_t h_t^{-1/2}|^{2/(1+\beta)}\right) \quad 4.16$$

where

$$c(\beta) = \{ \Gamma(\frac{3}{2}(1+\beta)) \}^{\frac{1}{1+\beta}} \{ \Gamma(\frac{1}{2}(1+\beta)) \}^{-\frac{1}{1+\beta}} \quad 4.17$$

$$\omega(\beta) = \{ \Gamma(\frac{3}{2}(1+\beta)) \}^{\frac{1}{2}} \{ \Gamma(\frac{1}{2}(1+\beta)) \}^{-\frac{3}{2}} (1+\beta)^{-1} \quad 4.18$$

this all under the condition that $-1 < \beta \leq 1$. When $\beta = 0$ the distribution is Normal; when $\beta = 1$ the double exponential distribution is obtained, and

when $\beta \rightarrow -1$ the distribution tends to the uniform distribution. The associated log-likelihood function is :

$$LL_{\beta}(\theta) = T \ln(\omega(\beta)) - \frac{1}{2} \sum_{t=1}^T \ln h_t - c(\beta) \sum_{t=1}^T \left(\frac{\varepsilon_t^2}{h_t} \right)^{\frac{1}{1+\beta}} \quad 4.19$$

$$\theta = (\varphi_0, \varphi_1, \alpha_0, \alpha_1, \beta_1, \beta) \quad 4.20$$

These are the conditional distributions that will be used and compared. The actual estimation procedures and results will be discussed in the following section.

2.5 Estimation procedures and results

We now proceed along the following lines: we do not assume that any of the suggested conditional distributions is the true conditional distribution. This implies that the log likelihood functions derived in the previous section have to be interpreted as quasi log likelihood functions, which leaves the models subject to distributional misspecification. However as Weiss (1984,1986a,b) has shown, the QML parameter estimates, obtained from a possibly incorrect likelihood function, will still be consistent and asymptotically normal, provided that the conditional mean and variance equations are correctly specified, and some regularity conditions are satisfied. Weiss derived his results with the conditional normal density giving a possible incorrect log likelihood function. We assume that these results remain valid under other distributional specifications. The assumption of "everything holds" is common practice in empirical work using GARCH models, because the asymptotic properties of ML estimates in GARCH models are far from clear.

The Berndt, Hall, Hall and Hausman (1974) algorithm (BHHH), with numerical derivatives, was chosen to obtain maximum likelihood estimates.²³ This algorithm is in fact an iterative estimation procedure and only uses first

²³ Pre sample values for ε_t^2 and h_t were set to sample means.

derivatives. Define l_t as the log likelihood contribution of a single observation. The parameter estimates after the $(i+1)$ th iteration, $\hat{\theta}^{(i+1)}$, are obtained from $\hat{\theta}^{(i)}$ by :

$$\theta^{(i+1)} = \theta^{(i)} + \lambda_i \left(\sum_{t=1}^T \frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta'} \right)^{-1} \Bigg| \theta^{(i)} \sum_{t=1}^T \frac{\partial l_t}{\partial \theta} \Bigg| \theta^{(i)} \quad 5.1$$

Where λ_i is a variable steplength. Now Define Matrices A and B as:

$$A = -E \left(\frac{\partial^2 l_t}{\partial \theta \partial \theta'} \right) \quad 5.2$$

A is the negative expectation of the Hessian matrix, or the information matrix.

$$B = E \left(\frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta'} \right) \quad 5.3$$

B is the cross-product matrix of first derivatives. If the conditional distribution were truly normal, the information matrix equivalence would hold, and one would have: $A=B$, and then the asymptotic variance of $\hat{\theta}^{(n)}$ equals A^{-1} or for that matter B^{-1} , where n denotes the final BHHH iteration. A consistent estimate of the asymptotic covariance matrix is given by:

$$\frac{1}{T} \left(\sum_{t=1}^T \frac{\partial l_t}{\partial \theta} \frac{\partial l_t}{\partial \theta'} \right)^{-1} \Bigg| \hat{\theta}^{(N)} = \hat{B}^{-1} \quad 5.4$$

It is obtained from the last BHHH iteration, and will be called the BHHH covariance matrix. Without Normality, $\hat{\theta}^{(n)}$ is still consistent for θ_0 , the true parameter vector, and asymptotically normal but now with a different asymptotic covariance matrix.

totic covariance matrix: $A^{-1}BA^{-1}$. The estimated covariance matrix, $\hat{A}^{-1}\hat{B}\hat{A}^{-1}$ ²⁴ will be called the robust covariance matrix, see Bollerslev (1986) and Weiss (1984, 1986 a,b).

In tables 2.4 and 2.5, the parameter values under a conditionally normal distribution are given, with both robust as well as BHHH standard errors, which are given for illustration purposes only. In general the BHHH standard errors are smaller than the robust standard errors for the weekly estimates, and the opposite holds for the monthly data. Furthermore for the weekly results very significant²⁵ GARCH parameters are found. This is even the case for the US returns, where the standard LM test for ARCH effects (see Table 2.3.a) indicated otherwise²⁶. The evidence for the monthly returns is mixed, giving significant β_1 estimates for all countries, whereas α_1 is not significant for every country. Note that $\alpha_1 + \beta_1$, is near unity for the German and Japanese weekly and the German monthly returns, which suggests that an integrated GARCH specification may be appropriate.

Tables 2.6 and 2.7 present the estimation results under a conditional t-distribution. For the weekly data the GARCH parameters are roughly similar to the estimates under the normal distribution. Again weekly estimates of the GARCH parameters are highly significant, whereas monthly estimates are much less significant if at all. However these results are certainly not as dramatic as indicated by the LM test (see table 2.3.b), but are in fact much more in line with the Markov Chain Tests. The degrees of freedom parameter, ν , is always significant, and contrary to what one might expect, does not increase when moving from a weekly to a monthly time interval, in four out of five cases. This is surprising as the general notion in the literature suggests that conditional distributions for financial returns, when aggregated, approach normality. If this were indeed true we would expect to see a significant increase in the value of ν . This however is not the case. In section 7 we test the hypothesis that ν is the same for weekly and monthly returns.

²⁴ A can also be consistently estimated by taking the sample mean evaluated at the final estimates.

²⁵ By significant I mean significantly different from zero at the 5% level.

²⁶ Diebold (1988) observed the same phenomenon when fitting an ARCH (12) process to the weekly French Franc / Dollar exchange rate.

Table 2.4 : GARCH(1,1)-Normal Estimates , Weekly Returns January 2 1973 - September 18 1990

	UK	USA	NET	GER	JAP
φ_0	0.001416 (0.000864) [0.000923]	0.000959 (0.000809) [0.000744]	0.001454 (0.000599) [0.000573]	0.000638 (0.000604) [0.000496]	0.002076 (0.000516) [0.000539]
φ_1	0.109904 (0.034735) [0.042291]	0.021378 (0.056107) [0.038185]	0.053189 (0.041430) [0.037546]	0.107404 (0.038291) [0.032598]	0.027863 (0.036404) [0.036880]
α_0	0.000056 (0.000028) [0.000012]	0.000011 (0.000013) [0.000004]	0.000038 (0.000015) [0.000009]	0.000004 (0.000003) [0.000002]	0.000004 (0.000003) [0.000001]
α_1	0.079412 (0.018457) [0.012724]	0.057525 (0.017237) [0.015258]	0.218690 (0.078124) [0.035350]	0.086660 (0.025219) [0.011610]	0.092061 (0.027046) [0.011916]
β_1	0.840325 (0.040420) [0.027777]	0.929216 (0.025725) [0.014114]	0.724620 (0.044078) [0.044616]	0.908177 (0.026264) [0.011546]	0.906243 (0.028137) [0.010572]
LL	2090.93	2179.33	2310.38	2405.84	2393.33

Heteroskedasticity Consistent Standard Errors are given in parentheses. BHHH Standard Errors are given in brackets. LL denotes the Log-Likelihood Value.

The GED results in Tables 2.8 and 2.9 give the same unanticipated result. Remember that $\beta=0$ means normality, and $\beta>0$ gives a conditionally leptokurtic distribution. Here we see that β which is always significant doesn't change towards zero as we shift from weekly to monthly data. Again the weekly GARCH parameters estimates are highly significant, which is not always true for the monthly estimates.

Table 2.5 : GARCH(1,1)-Normal Estimates , Monthly Returns January 2 1973 - September 18 1990

	UK	USA	NET	GER	JAP
φ_0	0.009465 (0.005543) [0.004307]	0.004552 (0.003229) [0.003565]	0.004336 (0.003277) [0.003444]	0.002921 (0.002589) [0.003036]	0.009302 (0.002382) [0.002463]
φ_1	-0.040146 (0.096720) [0.102421]	0.023517 (0.076654) [0.079932]	0.103124 (0.087033) [0.089910]	0.116366 (0.073110) [0.076458]	0.061225 (0.072961) [0.086012]
α_0	0.001148 (0.000617) [0.000474]	0.000903 (0.000486) [0.001348]	0.000306 (0.000393) [0.000753]	0.000018 (0.000032) [0.000021]	0.000205 (0.000115) [0.000100]
α_1	0.361316 (0.201620) [0.131818]	0.056729 (0.047463) [0.051832]	0.029998 (0.103489) [0.069573]	0.096207 (0.033328) [0.033963]	0.322958 (0.148928) [0.081299]
β_1	0.454537 (0.151928) [0.172780]	0.529515 (0.193152) [0.659219]	0.833730 (0.217581) [0.387149]	0.906772 (0.028691) [0.028331]	0.613991 (0.126019) [0.097988]
LL	281.41	349.37	346.77	358.76	374.72

Heteroskedasticity Consistent Standard Errors are given in parentheses. BHHH Standard Errors are given in brackets. LL denotes the Log-Likelihood Value.

Table 2.6 : GARCH(1,1)-t Estimates , Weekly Returns January 2 1973 - September 18 1990

	UK	USA	NET	GER	JAP
φ_0	0.002156 (0.000691) [0.000693]	0.001560 (0.000681) [0.000661]	0.001608 (0.000558) [0.000551]	0.001198 (0.000534) [0.000511]	0.002273 (0.000477) [0.000478]
φ_1	0.100707 (0.032448) [0.035391]	-0.015277 (0.026233) [0.034084]	0.018337 (0.033613) [0.035599]	0.096903 (0.034743) [0.033234]	0.012899 (0.033428) [0.033183]
α_0	0.000052 (0.000016) [0.000018]	0.000029 (0.000011) [0.000014]	0.000042 (0.000017) [0.000015]	0.000007 (0.000004) [0.000003]	0.000006 (0.000004) [0.000003]
α_1	0.115816 (0.029756) [0.032323]	0.057240 (0.020075) [0.023527]	0.167536 (0.042458) [0.044258]	0.084867 (0.022052) [0.020061]	0.086509 (0.031371) [0.020813]
β_1	0.799286 (0.042974) [0.047835]	0.881533 (0.035707) [0.043796]	0.744105 (0.060094) [0.060424]	0.897648 (0.027117) [0.021968]	0.904648 (0.034443) [0.020741]
ν	6.301665 (1.082417) [0.753193]	8.261054 (1.697263) [0.951720]	5.445289 (1.150821) [0.825730]	8.205810 (0.866257) [1.475907]	4.953811 (0.716123) [0.802932]
LL	2168.54	2240.66	2359.50	2428.33	2437.83

Heteroskedasticity Consistent Standard Errors are given in parentheses. BHHH standard errors are given in brackets. LL denotes Log-Likelihood value.

Table 2.7 : GARCH(1,1)-t Estimates , Monthly Returns January 2 1973 - September 18 1990

	UK	USA	NET	GER	JAP
φ_0	0.011621 (0.004307) [0.003596]	0.005599 (0.002980) [0.002819]	0.005019 (0.002858) [0.002684]	0.003861 (0.002850) [0.002569]	0.009390 (0.002521) [0.002394]
φ_1	0.084115 (0.085641) [0.071840]	0.035070 (0.067558) [0.071240]	0.061682 (0.076743) [0.058923]	0.104292 (0.070916) [0.071440]	0.026598 (0.047712) [0.079209]
α_0	0.000957 (0.000598) [0.000770]	0.000703 (0.000313) [0.001012]	0.000161 (0.000122) [0.000489]	0.000130 (0.000159) [0.000119]	0.000098 (0.000168) [0.000085]
α_1	0.179300 (0.123898) [0.159016]	0.073967 (0.046117) [0.080931]	0.013188 (0.020727) [0.032129]	0.141493 (0.093875) [0.085343]	0.208940 (0.198667) [0.090318]
β_1	0.606806 (0.189854) [0.276529]	0.599366 (0.141598) [0.518440]	0.904581 (0.056635) [0.267137]	0.809296 (0.143298) [0.096095]	0.775242 (0.230744) [0.108000]
ν	4.305197 (1.365484) [1.094897]	5.036322 (2.109797) [1.563170]	4.621613 (1.441824) [1.162927]	5.075227 (1.830360) [1.695907]	5.435939 (1.879397) [2.613694]
LL	299.81	361.09	366.45	366.86	378.46

Heteroskedasticity Consistent Standard Errors are given in parentheses. BHHH Standard Errors are given in brackets. LL denotes the Log-Likelihood Value.

Table 2.8 : GARCH(1,1)-GED Estimates , Weekly Returns January 2 1973 - September 18 1990

	UK	USA	NET	GER	JAP
φ_0	0.002141 (0.000504) [0.000672]	0.001594 (0.000834) [0.000651]	0.001836 (0.000712) [0.000525]	0.001016 (0.000538) [0.000509]	0.002229 (0.000672) [0.000463]
φ_1	0.099288 (0.036880) [0.032970]	-0.002563 (0.041856) [0.033320]	0.001764 (0.041288) [0.034054]	0.103847 (0.032733) [0.032845]	0.008825 (0.044782) [0.032078]
α_0	0.000053 (0.000019) [0.000019]	0.000022 (0.000008) [0.000013]	0.000042 (0.000015) [0.000016]	0.000006 (0.000003) [0.000003]	0.000005 (0.000003) [0.000002]
α_1	0.103513 (0.029807) [0.030107]	0.054854 (0.016373) [0.024110]	0.180470 (0.048047) [0.049962]	0.085484 (0.023636) [0.017652]	0.086416 (0.027800) [0.019517]
β_1	0.815469 (0.043891) [0.048861]	0.904570 (0.025976) [0.038814]	0.734510 (0.051168) [0.067932]	0.903256 (0.026759) [0.018044]	0.906111 (0.030014) [0.019243]
β	0.608313 (0.077119) [0.050403]	0.466128 (0.115560) [0.045044]	0.600529 (0.092656) [0.061320]	0.346486 (0.020101) [0.060005]	0.583381 (0.109123) [0.080173]
LL	2150.23	2218.74	2351.55	2419.86	2429.76

Heteroskedasticity Consistent Standard Errors are given in parentheses. BHHH Standard Errors are given in brackets. LL denotes the Log-Likelihood Value.

Table 2.9 : GARCH(1,1)-GED Estimates , Monthly Returns January 2 1973 - September 18 1990

	UK	USA	NET	GER	JAP
φ_0	0.013712 (0.002337) [0.003491]	0.005368 (0.004100) [0.002809]	0.005347 (0.001741) [0.002574]	0.002819 (0.004032) [0.002373]	0.009652 (0.002777) [0.002326]
φ_1	0.043690 (0.071119) [0.072135]	0.031539 (0.099262) [0.068596]	0.051222 (0.086967) [0.057923]	0.166228 (0.106770) [0.065141]	0.038439 (0.063176) [0.078366]
α_0	0.001008 (0.000493) [0.000723]	0.000754 (0.000357) [0.001352]	0.000151 (0.000223) [0.000658]	0.000056 (0.000062) [0.000062]	0.000146 (0.000110) [0.000110]
α_1	0.242727 (0.139218) [0.199993]	0.063920 (0.043066) [0.076591]	0.007907 (0.027254) [0.030684]	0.098559 (0.056871) [0.058495]	0.269388 (0.148811) [0.105995]
β_1	0.545673 (0.147766) [0.292378]	0.577482 (0.164833) [0.691506]	0.915610 (0.120203) [0.349943]	0.882126 (0.069055) [0.059390]	0.693724 (0.138910) [0.127996]
β	0.728418 (0.274528) [0.178284]	0.542972 (0.229441) [0.155811]	0.699722 (0.246301) [0.151466]	0.704886 (0.350778) [0.214788]	0.513527 (0.190306) [0.235658]
LL	295.00	357.82	361.99	366.74	378.95

Heteroskedasticity Consistent Standard Errors are given in parentheses. BHHH Standard Errors are given in brackets. LL denotes the Log-Likelihood Value.

2.6 Diagnostic checks

The normalized residuals, $\hat{\epsilon}_t/h_t^{1/2}$ from the estimated GARCH models show no evidence of any serial correlation or heteroskedasticity unaccounted for by the models. Furthermore, Wald tests for the hypothesis of conditional homoskedasticity, $H_0 : \alpha_1 = \beta_1 = 0$, clearly reject this hypothesis in all cases. Note that this hypothesis implies that the true parameter point is on the boundary of the feasible parameter space. This implies that the standard Wald test is a conservative test procedure, and when the Wald test leads to rejection, then a test procedure using the true asymptotic distribution will surely reject, see Gouriéroux et al. (1982) and Kodde and Palm (1986).

As the Normal distribution is a special case of both the Student t-distribution as well as the GED distribution, LR tests can be readily applied to confront a conditionally Normal distribution with either the Student t or GED specification. The smallest LR value for the weekly data is 28.04, when comparing the Normal to the GED specification for the German data, and for the monthly data the smallest value is 7.48, obtained by comparing the Normal to the t specification for the Japanese data. Even these minimum values exceed the $\chi^2(1)$ critical value at any reasonable significance level. Thus conditional Normality can be dismissed altogether.

Standard Likelihood ratio Tests cannot be employed to discriminate between a conditional t and a conditional GED specification, since they are clearly non-nested. Vuong (1989) however developed a Model Selection test for strictly non-nested models based on the Likelihood Ratio (LR) statistic. This test does not require that any of the two competing specifications is correctly specified, it casts model selection for non-nested models in a probabilistic setting and tries to determine which of the two specifications is closer to the true, unknown and unnecessary, data generating process, (DGP). Under the provision that all regularity conditions are satisfied, see Vuong (1989, assumptions A1-A6), the test can, for our purposes, be constructed along the following lines. Define:

$$LR_T(\hat{\theta}_T^t, \hat{\theta}_T^{GED}) \equiv LL_v(\hat{\theta}_T^t) - LL_\beta(\hat{\theta}_T^{GED}) \quad 6.1$$

The usual LR statistic. Furthermore define:

$$\hat{\omega}_T^2 \equiv \frac{1}{T} \sum_{t=1}^T (l_t'(\theta_T^t) - l_t^{GED}(\theta_T^{GED}))^2 - (\frac{1}{T} LR_T(\theta_T^t, \theta_T^{GED}))^2 \quad 6.2$$

where subscripts denote time and superscripts denote the distribution. Vuong (1989, Theorem 5.1) establishes the following:

$$T^{-1/2} LR_T(\theta_T^t, \theta_T^{GED}) / \hat{\omega}_T \stackrel{a}{\sim} N(0,1) \quad 6.3$$

under the null hypothesis that both specifications are equivalent, which means that they are equidistant from the true data generating process. If the value of the statistic exceeds a predetermined critical value c from the standard Normal Distribution, then a t-distribution is better than a GED, i.e. closer to the true DGP; if it is smaller than $-c$, a GED is better. In all other cases the test is inconclusive. The results of this test are contained in Table 2.10. We chose the 95% critical value from the standard Normal distribution 1.645, as the value of c^{27} . The t-distribution is superior to the GED in 6 cases. For the remaining cases the test is inconclusive. The GED is never preferred.

Table 2.10 : Modified Likelihood Ratio tests to discriminate between a conditional t-distribution and a GED.

	UK	US	NET	GER	JAP
WEEKLY	1.61*	1.71*	1.07	2.20*	2.09*
MONTHLY	2.14*	1.37	1.84*	0.06	-0.41

* denotes a value of the statistic that exceeds the 5% critical value (1.645) from a Standard Normal distribution, implying that a conditional t distribution is better than a GED.

²⁷ Remember we have two one-sided alternatives

As a final diagnostic check we inspect the skewness and kurtosis values of the standardized residuals. From Jensen's inequality it follows that standardized residuals from GARCH models should demonstrate less absolute skewness and should be thinner-tailed than their unconditional raw data counterparts. Any strong violation of this rule should be regarded as evidence of misspecification, see Hsieh (1989).

From Tables 2.11.a and 2.11.b we can learn that for the UK and US weekly data this violation actually occurs. Skewness comparisons are not presented because they show exactly the same picture. Note that standardized residuals from a GARCH (1,1)-t specification always have the highest kurtosis values.

Table 2.11.a : Kurtosis Values for standardized weekly residuals.

	UK	US	NET	GER	JAP
summary	16.4	23.5	18.8	13.7	13.9
normal	24.0	21.3	10.9	5.8	6.5
Student-t	29.1	26.1	12.3	5.9	6.8
GED	27.4	24.5	12.2	5.9	6.7

Table 2.11.b : Kurtosis Values for standardized monthly residuals.

	UK	US	NET	GER	JAP
summary	12.8	7.0	10.0	8.7	5.6
normal	8.8	7.6	10.0	5.5	3.9
Student-t	10.4	8.0	10.3	8.2	4.4
GED	9.8	7.8	10.3	6.3	4.1

2.7 The effects of temporal aggregation

In this section we address two questions regarding the effects of temporal aggregation. First of all: "Does conditional heteroskedasticity disappear when the data are aggregated more and more ?" and secondly "Does the distribution of rescaled innovations resemble a Normal distribution when the data are aggregated more and more ?".

Recently Drost & Nijman (1991) have derived low frequency (i.e. aggregated) models that are implied by an assumed high frequency GARCH models. In their derivation they adopt three definitions of GARCH. In the case of strong GARCH, rescaled innovations are independent, they are uncorrelated in the semistrong GARCH case, and in Weak GARCH h_t is the best linear predictor of ϵ_t^2 based only on past innovations. Their conclusions are that symmetric weak GARCH is closed under temporal aggregation and that (semi)strong GARCH aggregates to weak GARCH. Henceforth, if high frequency rescaled innovations are assumed to be independently distributed, low frequency rescaled innovations will not be independent. Important is that conditional heteroskedasticity of the GARCH form in a high frequency model implies that the low frequency model exhibits conditional heteroskedasticity of the GARCH form as well.

In one of their examples they show how parameter estimates of a high frequency GARCH (1,1) model can be used to obtain estimates of the low frequency GARCH (1,1) parameters, an approach that could result in more efficient estimates than direct estimates from low frequency data. Salient features of this methodology are first of all that the GARCH parameters in the low frequency model depend not only on the GARCH parameters in the high frequency model but also on the unconditional kurtosis of the rescaled innovations of the high frequency model. Secondly the sum of the GARCH parameters in the low frequency case equals the sum of the GARCH parameters in the high frequency case taken to the power of the aggregation level (m):

$$\alpha_m + \beta_m = (\alpha + \beta)^m , \quad \alpha + \beta < 1$$

7.1

This means that conditional heteroskedasticity eventually disappears when we aggregate more and more, but the level of aggregation has to be quite large if the original high frequency process is close to IGARCH, to cause conditional heteroskedasticity of the GARCH form to disappear. For more details, see Drost & Nijman (1991) and Appendix 2.A. Thirdly the value of α tends to rise and β tends to fall for low to moderate aggregation levels. If the aggregation process continues, both α and β decrease in value. In table 2.12 we have confronted our monthly GARCH parameter estimates with the estimates implied by the Drost-Nijman methodology. We only report results for the conditional Student t-distribution since this is our preferred alternative and provides the best results. In deriving these results we have used the measured unconditional kurtosis instead of the implied value (k_y):

$$k_y = k_c \frac{1 - (\beta + \alpha)^2}{1 - (\beta + \alpha)^2 - (k_c - 1)\alpha^2} \quad 7.2$$

Which can easily become negative (e.g. if $\alpha=0.15$ $\beta=0.8$ and $k_c=6$ then $k_y=-39$)

Table 2.12 : Comparison of Direct Monthly Estimates to estimates implied by weekly data (Student t-distribution)

	weekly estimates			monthly estimates			implied monthly		
	α	β	k_y	α	β	k_y	α	β	k_y
UK	0.115	0.799	29.06	0.179	0.607	10.43	0.172	0.507	19.61
USA	0.057	0.882	26.09	0.074	0.599	8.02	0.118	0.644	12.81
NET	0.167	0.744	12.34	0.013	0.905	10.38	0.184	0.485	11.80
GER	0.085	0.897	5.97	0.141	0.809	8.21	0.127	0.798	6.35
JAP	0.087	0.905	6.78	0.209	0.775	4.48	0.120	0.846	8.88
NET2	0.167	0.744	12.34	0.183	0.498	9.20	0.180	0.645	12.31

We can observe that the implied values are quite close to the direct estimates except for the Dutch data. One explanation for this fact is that there is no longer any conditional heteroskedasticity present in the monthly dutch stock returns. This however is partly contradicted by the results for bi-weekly Dutch data (under NET 2 in table 2.12) where conditional heteroskedasticity is still clearly present, and where the direct estimates are not too far off the implied estimates²⁸. This probably illustrates the fact that the implied estimates may be more efficient than the direct estimates. For the datasets analyzed we can conclude that aggregation from the weekly to the monthly level in general does not cause conditional heteroskedasticity of the GARCH form to disappear.

In section 6 we gave some informal results for the characteristic parameters of the t and GED distribution and suggested the absence of convergence to normality. We now turn to a more formal analysis.

Diebold (1988) asserts that if some time series follows an ARCH process, then the aggregated series follows an unconditional normal distribution when the level of aggregation approaches infinity. Below we will show that an aggregation for stock returns from the weekly to the monthly level is insufficient to generate this result. Furthermore we will show that both the raw weekly and monthly series as well as the rescaled innovations from the GARCH processes could be generated by Student t-distributions with few degrees of freedom.

To establish these results, tail index estimates based on extreme value theory are employed²⁹. The use of tail index estimators is particularly appealing for the following reasons: first of all the hypothesized distributions often appear as non-nested alternatives (e.g. a sum-stable and a Student t-distribution) invalidating standard likelihood ratios tests. Tail indices can be estimated in such a way that the different hypotheses appear as nested alternatives. Thus estimates do not rely on one of the alternatives as a

²⁸ For the Dutch data the direct bi-weekly estimates are very close to the implied monthly estimates.

²⁹ The author thanks Kees Koedijk for suggesting this idea and for making the simulation program available to him.

maintained hypothesis. Secondly tail index estimates can be used to test for parameter invariance over subperiods and the effects of temporal aggregation on the distribution. For a detailed discussion on the construction and use of tail index estimators we would like to refer to Koedijk, Schafgans & De Vries (KSV) (1990) and Koedijk and Kool (1992). The following discussion only covers basic concepts and results. Consider the random variable, M_t , defined as:

$$M_t = \max (x_1, \dots, x_t) \quad 7.3$$

where $\{x\}_1^\infty$, is a sequence of i.i.d. random variables with probability distribution F . The probability that M_t is below a given level x is given by:

$$P \{ M_t < x \} = F^t(x) \quad 7.4$$

Extreme Value Theory studies the limiting distributions of the order statistic M_t appropriately scaled. Required are normalizing constants $a_t > 0$, b_t such that:

$$P \{ a_t (M_t - b_t) \leq x \} \xrightarrow{w} G(x) \quad 7.5$$

Where w , stands for weak convergence. The asymptotic distribution $G(x)$, if it exists, is called a max-stable distribution (see KSV (1990), Definition 1 and Theorem 1, or Mood, Graybill, Boes (1988), Ch. 6., Theorems 15 and 16), and can be represented as:

$$G_\gamma(x) = e^{-(1+\gamma x)^{-1/\gamma}}, \quad 1 + \gamma x \geq 0 \quad 7.6$$

The Characteristic Parameter of this distribution, γ , equals the inverse of the tail index, α . In general for a distribution with tail index α only moments of order smaller than α exist. Now it becomes clear how estimates of α can discriminate between competing distributions. For sum stable distributions

the variance does not exist, so this class of distributions requires $\alpha < 2$, the Student t-class of distributions allow for $\alpha \geq 2$, and finally the normal distribution, as a special case requires α to be infinite. In the remainder we will assume $\gamma > 0$ or $0 < \alpha < \infty$, since this is the relevant case for our purposes. For a symmetric sum stable distribution α can be interpreted as the characteristic exponent and for the class of Student t-distributions, α may be viewed as the number of degrees of freedom.

The proposed estimator, also used by KSV, can be established as follows: let x_1, \dots, x_t be a sequence of stationary i.i.d. observations with distribution $F \in D(G(x), \gamma > 0)$ (D denotes the Domain of attraction). Define $x_{(1)}, \dots, x_{(t)}$ as the ascending order statistics from the above sample. The estimator is given by :

$$\hat{\gamma} = 1/\hat{\alpha} = \frac{1}{m} \sum_{i=1}^m (\ln x_{(t+1-i)} - \ln x_{(t-m)}) \quad 7.7$$

Goldie and Smith (1987) have shown that $(\hat{\gamma} - \gamma)m^{1/2}$ is asymptotically normal with mean zero and variance γ^2 . This also means that $\hat{\alpha}$ is asymptotically normal with mean α and variance α^2/m . The independency assumption could be relaxed as dependency affects the normalizing constants but not the tail index, see KSV (1990). The above estimator may be used in several ways. In its form given above it only uses information in the right tail of the distribution. A tail index estimator based on information in the left tail can be obtained by forming a sequence of descending order statistics and taking absolute values. Conditional on the right and left tail having the same index, one can combine the information in both tails by taking absolute values of the observations first and then forming an ascending sequence of order statistics. This procedure significantly enhances the precision of the tail index estimates. The number of order statistics, m , to consider in the estimation procedure is generally unknown. By means of a Monte Carlo study we select an optimal m for a given sample size and α , using the minimal MSE criterion which is appropriate given the asymptotic normality of the estimator.

Each Monte Carlo experiment consists of 300 replications of T draws from five Student t-distributions with degrees of freedom, or α , ranging from 1 to 5. Separate experiments are conducted for using one tail or both tails and for

$T=923$ or $T=211$, the number of weekly and monthly observations respectively. Optimal m levels are reported in Table 2.13. This table presents two notable facts. The optimal number of tail observations to be used for estimation, m , is inversely related to α . This is not so surprising: For lower α and consequently fatter tails, more observations are contained in the tails, and more tail observations should be used. Secondly both the theoretical as the empirical minimal MSE are smaller when using both tails instead of only one tail. This characterizes the greater precision in using both tails.

Table 2.13 : Optimal Choice of m through Monte Carlo experiments by minimizing MSE.

	$\alpha = 1$	$\alpha = 2$	$\alpha = 3$	$\alpha = 4$	$\alpha = 5$
$n = 923$					
one tail	106	60	27	18	17
MSE	0.0100	0.0073	0.0063	0.0080	0.0072
γ^2/m	0.0094	0.0042	0.0041	0.0035	0.0024
both tails	227	82	46	23	19
MSE	0.0060	0.0045	0.0045	0.0047	0.0054
γ^2/m	0.0044	0.0030	0.0024	0.0027	0.0021
$n = 211$					
one tail	38	21	10	10	5
MSE	0.0335	0.0228	0.0184	0.0208	0.01186
γ^2/m	0.0263	0.0119	0.0111	0.0063	0.0080
both tails	62	28	21	14	9
MSE	0.0165	0.0120	0.0109	0.0127	0.0112
γ^2/m	0.0161	0.0089	0.0053	0.0045	0.0044

Where γ^2/m is the theoretical MSE value in the i.i.d. case.

We estimated tail indices for all the series of raw returns under the various hypotheses of α being 1, 2, 3, 4 or 5. The results are described in Table 2.14, where α^+ and α^- denote tail index estimates for right and left tails respectively. Asymptotic standard errors are given in parentheses. $\hat{\alpha}$ denotes the tail index estimator for both tails. We can test for the equality of the tail indices of the right and left tail of the distributions, using the asymptotic

normality of the tail index estimates. The asymptotic variance of $(\alpha^+ - \alpha^-)$ is now given by $((\hat{\alpha}^+)^2/m + (\hat{\alpha}^-)^2/m)$ and the test boils down to a regular t-test. Although the results are not reported, equality could be rejected only in 5 out of 25 cases, out of which 4 rejections can be disregarded because the hypothesis $\alpha = i$ ($i=1\dots 5$) could be rejected. For the monthly raw returns equality was never rejected (not reported). Furthermore we can conclude from Table 2.14 that the hypotheses $\alpha=1$ and $\alpha=2$ are rejected for all series. This means we can reject the sum stable distribution for weekly stock returns which is consistent with Akgiray & Booth (1988b). Finally the hypothesis $\alpha=3$ is never rejected, $\alpha=4$ is rejected for Germany and $\alpha=5$ for the US, the UK and Germany. Weekly stock returns seem to be generated by Student t-distributions with few degrees of freedom ($\alpha=3$, $\alpha=4$) and a possible infinite kurtosis (if $\alpha \leq 4$). For the monthly returns the same results emerge: $\alpha=3$ or $\alpha=4$ is not rejected, while, $\alpha=1$, $\alpha=2$ and $\alpha=5$ are clearly rejected.

We now turn to the results for the rescaled innovations from the estimated GARCH processes using the conditional Student t-distribution, to check whether these innovations follow unconditional Student t-distributions. Again we tested for the equality of the left and right tail index, and this equality was only rejected in 3 out of 25 cases for the weekly data and no rejections occurred for the monthly data. The results are contained in tables 2.15 and 2.16. For weekly rescaled innovations $\alpha=1$ and $\alpha=2$ are always rejected, so we can rule out the sum stable hypothesis. Also $\alpha=3$ is rejected for the Dutch data, whereas $\alpha=5$ is rejected for the UK the US and Germany. The hypothesis that $\alpha=4$ is never rejected. Again the results indicate that the data analyzed are generated by a Student t-distribution with few degrees of freedom. For monthly rescaled innovations the results are somewhat surprising since the hypothesis that $\alpha=2$ cannot be rejected for the Dutch, German and British data. So for these countries we cannot reject the hypothesis of a sum-stable distribution, with infinite unconditional variance. At the same time however, $\alpha=3$ is rejected nowhere and $\alpha=4$ is also not rejected for the Dutch and German Data.

In order to clarify these results the following tests are performed. Firstly we test whether the data for all stock indices can be characterized as having the same tail index, by means of pairwise comparison. We employ the same stat-

Table 2.14 : Tail Index Estimates for Weekly Raw Stock Returns

	US	UK	NET	GER	JAP
$\alpha = 1$					
$\alpha+$	3.06 (0.30)	2.61 (0.25)	2.48 (0.24)	3.02 (0.29)	2.61 (0.25)
$\alpha-$	2.60 (0.25)	2.02 (0.20)	1.91 (0.19)	2.01 (0.20)	1.62 (0.16)
α	2.62 (0.17)	2.19 (0.14)	2.11 (0.14)	2.31 (0.15)	2.05 (0.13)
$\alpha = 2$					
$\alpha+$	3.63 (0.47)	2.84 (0.37)	3.59 (0.46)	3.13 (0.40)	3.51 (0.45)
$\alpha-$	3.14 (0.41)	2.53 (0.33)	2.34 (0.30)	2.23 (0.29)	1.74 (0.22)
α	3.28 (0.36)	2.69 (0.30)	3.22 (0.36)	3.05 (0.34)	2.98 (0.33)
$\alpha = 3$					
$\alpha+$	3.35 (0.65)	3.06 (0.59)	4.21 (0.81)	3.99 (0.77)	4.10 (0.79)
$\alpha-$	3.47 (0.72)	2.73 (0.53)	3.44 (0.66)	2.56 (0.49)	2.19 (0.42)
α	3.55 (0.52)	2.97 (0.44)	3.81 (0.56)	2.95 (0.43)	2.92 (0.43)
$\alpha = 4$					
$\alpha+$	3.22 (0.76)	2.79 (0.66)	4.20 (0.99)	3.69 (0.87)	4.72 (1.11)
$\alpha-$	3.83 (0.90)	3.49 (0.82)	3.37 (0.79)	2.52 (0.59)	2.76 (0.65)
α	3.24 (0.67)	3.19 (0.67)	4.09 (0.85)	2.76 (0.58)	3.63 (0.76)
$\alpha = 5$					
$\alpha+$	3.08 (0.75)	2.65 (0.64)	4.03 (0.98)	3.63 (0.88)	4.59 (1.11)
$\alpha-$	3.66 (0.88)	3.37 (0.82)	3.33 (0.81)	2.41 (0.58)	2.82 (0.68)
α	3.44 (0.79)	3.18 (0.73)	3.79 (0.87)	2.70 (0.62)	3.72 (0.85)

asymptotic standard errors are given in parentheses.

Table 2.15 : Tail Index Estimates for Weekly Rescaled Innovations from GARCH-t Processes

	US	UK	NET	GER	JAP
$\alpha = 1$	2.71 (0.18)	2.51 (0.17)	2.38 (0.16)	2.77 (0.18)	2.31 (0.15)
$\alpha = 2$	2.88 (0.32)	2.92 (0.32)	3.79 (0.42)	3.75 (0.41)	3.02 (0.33)
$\alpha = 3$	3.86 (0.57)	4.04 (0.59)	4.49 (0.66)	3.50 (0.52)	3.35 (0.49)
$\alpha = 4$	3.79 (0.79)	3.16 (0.66)	4.37 (0.91)	3.15 (0.66)	3.84 (0.80)
$\alpha = 5$	3.50 (0.80)	3.18 (0.73)	4.69 (1.07)	2.99 (0.69)	4.86 (1.11)

Table 2.16 : Tail Index Estimates for Monthly Rescaled Innovations from GARCH-t Processes.

	US	UK	NET	GER	JAP
$\alpha = 1$	1.97 (0.25)	2.22 (0.28)	2.31 (0.29)	2.23 (0.28)	2.08 (0.26)
$\alpha = 2$	3.39 (0.64)	2.75 (0.52)	2.30 (0.44)	2.95 (0.56)	3.36 (0.64)
$\alpha = 3$	4.08 (0.89)	2.73 (0.60)	2.48 (0.54)	3.25 (0.71)	3.59 (0.78)
$\alpha = 4$	3.44 (0.92)	2.23 (0.60)	3.07 (0.82)	3.35 (0.89)	2.97 (0.79)
$\alpha = 5$	3.33 (1.11)	2.47 (0.82)	2.72 (0.91)	3.00 (1.00)	4.18 (1.39)

istic that is also used to test the equality of the left and right tail index. We perform tests for weekly and monthly data, for raw returns and for rescaled innovations from the GARCH-t processes, and for $\alpha=3$ and $\alpha=4$. Secondly we test whether temporal aggregation affects the tail index, or in other words whether the tail index for the weekly data of a stock index is statistically

significantly different from the tail index for the monthly data of that same stock index.

In order to pursue this analysis we have to rely on a rather indirect method as weekly and monthly samples are interdependent. Therefore we compare tail index estimates for weekly data of one stock index to tail index estimates for monthly data of all the other stock indices. For each test we provide only the largest absolute value of the test statistics. The results are contained in Table 2.17. We can observe that we cannot reject the hypothesis (based on pairwise comparison) that the data for all stock indices are characterized by the same tail index ($\alpha=3$ or $\alpha=4$). Furthermore we can conclude that temporal aggregation in general does not affect the tail index. This is in line with the results in KSV (1990) who find that for EMS exchange rate returns the tail index remains constant in temporal aggregation. These results do not imply that the total distribution remains constant in aggregation, only one characteristic does.

The difference between the estimated tail indices and the degrees of freedom estimated in the GARCH-t processes can be explained by the fact that the former estimates only use information in the tails, whereas the latter estimates use information from the whole distribution.

Table 2.17 : Tests of Equality of Tail Index Estimates

Data	Frequency	α	Hypothesis	Largest t-Value
Raw	Weekly	3	$\hat{\alpha}_{wi} = \hat{\alpha}_{wj}$	1.24 (NET/JAP)
Raw	Weekly	4	$\hat{\alpha}_{wi} = \hat{\alpha}_{wj}$	1.28 (NET/GER)
Raw	Monthly	3	$\hat{\alpha}_{mi} = \hat{\alpha}_{mj}$	0.80 (NET/US)
Raw	Monthly	4	$\hat{\alpha}_{mi} = \hat{\alpha}_{mj}$	1.45 (UK/JAP)
Innovations	Weekly	3	$\hat{\alpha}_{wi} = \hat{\alpha}_{wj}$	1.38 (NET/JAP)
Innovations	Monthly	4	$\hat{\alpha}_{wi} = \hat{\alpha}_{wj}$	1.09 (NET/GER)
Innovations	Monthly	3	$\hat{\alpha}_{mi} = \hat{\alpha}_{mj}$	1.53 (NET/US)
Innovations	Monthly	4	$\hat{\alpha}_{mi} = \hat{\alpha}_{mj}$	1.10 (UK/GER)
Raw	Weekly/Monthly	3	$\hat{\alpha}_{wi} = \hat{\alpha}_{wj}$	1.58 (NET/GER)
Raw	Weekly/Monthly	4	$\hat{\alpha}_{wi} = \hat{\alpha}_{wj}$	1.50 (NET/UK)
Innovations	Weekly/Monthly	3	$\hat{\alpha}_{wi} = \hat{\alpha}_{wj}$	1.94 (UK/NET)
Innovations	Weekly/Monthly	4	$\hat{\alpha}_{wi} = \hat{\alpha}_{wj}$	1.15 (NET/JAP)

The t-value is computed as: $\hat{\alpha}_i - \hat{\alpha}_j / \sqrt{\hat{\alpha}_i^2/m + \hat{\alpha}_j^2/m}$ w=weekly, m=monthly, i,j = UK,US,NET,GER,JAP $i \neq j$.

2.8 Conclusions

In this chapter we used a GARCH (1,1) model under various distributional assumptions to describe the evolution of stock indices returns. In the preliminary analysis we found that the usual LM test for ARCH effects does not always detect these effects, even when they are clearly present, as is seen in the estimation results. A test for ARCH effects based on a Markov-Chain approach appears to be an attractive alternative. Furthermore a conditional Student t-distribution is preferred over a Normal or GED distribution, even for monthly data. This does not mean that we claim that stock returns are conditionally t-distributed, but that it is the best alternative considered. These findings may have considerable consequences for multivariate analyses, which until now have used the assumption of multivariate conditional normality. Here lies a task for future research. Finally we find that conditional heteroskedasticity of the GARCH form does not disappear when moving from weekly to monthly data, there is no convergence to normality, and both weekly as well as monthly stock returns appear to follow an unconditional Student t-distribution with few degrees of freedom.

Appendix 2.A : Aggregation results

In this Appendix the aggregation results obtained by Drost and Nijman (1991) for symmetric weak GARCH (1,1) processes with flow variables (Their example 2) are reproduced. If $\{y_t\}$ symmetric weak GARCH (1,1) with :

$$h_t = \psi + \alpha y_{t-1}^2 + \beta h_{t-1} \quad \text{A.1}$$

and unconditional kurtosis, k_y , then

$$\left\{ y_{(m)m} = \sum_{i=0}^{m-1} y_{m-i} \right\} \quad \text{A.2}$$

is symmetric weak GARCH (1,1) with:

$$h_{(m)m} = \psi_{(m)} + \alpha_{(m)} y_{(m)m-m}^2 + \beta_m h_{(m)m-m} \quad \text{A.3}$$

where

$$\psi_{(m)} = m \psi \frac{1 - (\beta + \alpha)^m}{1 - (\beta + \alpha)} \quad \text{A.4}$$

$$\alpha_{(m)} = (\beta + \alpha)^m - \beta_{(m)} \quad \text{A.5}$$

$$k_{(m)y} = 3 + (k_y - 3)/m + 6(k_y - 1) \frac{\{m - 1 - m(\beta + \alpha) + (\beta + \alpha)^m\}\{\alpha - \beta\alpha(\beta + \alpha)\}}{m^2(1 - \beta - \alpha)^2 (1 - \beta^2 - 2\beta\alpha)} \quad \text{A.6}$$

and $|\beta_{(m)}| < 1$ is the solution of the system:

$$\frac{\beta_{(m)}}{1 + \beta_{(m)}^2} = \frac{a(\beta, \alpha, k_y, m)(\beta + \alpha)^m - b(\beta, \alpha, m)}{a(\beta, \alpha, k_y, m) \{1 + (\beta + \alpha)^{2m}\} - 2b(\beta, \alpha, m)} \quad A.7$$

$$a(\beta, \alpha, k_y, m) = m(I - \beta)^2 + 2m(m - I) \frac{(1 - \beta - \alpha)^2(1 - \beta^2 - 2\beta\alpha)}{(k_y - 1)\{1 - (\beta + \alpha)^2\}} +$$

A.8

$$4 \frac{\{m - I - m(\beta + \alpha) + (\beta + \alpha)^m\}\{\alpha - \beta\alpha(\beta + \alpha)\}}{I - (\beta + \alpha)^2}$$

$$b(\beta, \alpha, m) = \{\alpha - \beta\alpha(\beta + \alpha)\} \frac{I - (\beta + \alpha)^{2m}}{I - (\beta + \alpha)^2} \quad A.9$$

The proof is contained in Drost and Nijman (1991). These results can be readily extended to the strong GARCH case (without implying that strong GARCH is closed under temporal aggregation), where the conditional kurtosis of the rescaled innovations, k_c is given. The unconditional kurtosis can be obtained using:

$$k_y = k_c \frac{I - (\beta + \alpha)^2}{I - (\beta + \alpha)^2 - (k_c - I)\alpha^2} \quad A.10$$

The development of the three-moment CAPM has been driven by the need to improve the performance of the standard CAPM model. This has led to the introduction of more sophisticated models, such as the consumption based CAPM, which attempts to incorporate the effects of consumption on asset returns.

Chapter 3

An Intertemporal Three-Moment Capital Asset Pricing Model

An Analysis using Generalized Isoelastic Preferences and GARCH

3.1 Introduction

The aim of this chapter is to explore particular solutions to the measurement, preference and heteroskedasticity problems associated with the empirical rejection of the consumption based CAPM. The solutions which we propose will lead to an intertemporal three-moment CAPM.

In recent years we have witnessed a still ongoing debate regarding the empirical performance of two of the most widely used asset pricing models in modern finance theory, namely the market-based or static CAPM (Sharpe (1964), Lintner (1965)) and the consumption based or intertemporal CAPM (Merton (1973), Lucas (1978), Breeden (1979)). While the latter should be preferred theoretically, there now is considerable evidence that the former performs not as badly as the latter from an empirical viewpoint, see Mankiw and Shapiro (1986), Attanasio (1991), Grossman, Melino and Shiller (1987). Several explanations for this apparent controversy have been offered. First of all there are several problems associated with measured consumption. Breeden, Gibbons and Litzenberger (1989) mention a summation bias due to

the fact that periodically reported consumption really is an integral of instantaneous consumption rates during the period instead of the consumption measure one needs : the consumption rate on the last day of the period (this, of course, only is a problem in a continuous framework). This summation bias causes consumption betas and the covariance of asset returns with measured consumption to be underestimated. A second, less dramatic, problem is the low frequency at which consumption data are available. This problem can be overcome by using the concept of the maximum correlation portfolio, see Breeden, Gibbons, and Litzenberger (1989).

Furthermore, as Mankiw and Zeldes (1991) have shown, the covariance between stockholders' aggregate consumption and excess stock returns is much higher than the covariance between total consumption and stock returns. This difference is important as it helps to alleviate the so called equity premium puzzle. Mehra and Prescott (1985) showed, in a calibration exercise, that the low covariance of consumption growth with stock returns leads to implausibly high levels of relative risk aversion in order to rationalize the size of the equity premium. The separation of stockholders' consumption from total consumption is only part of the solution : coefficients of relative risk aversion are still too high.

This leads us to the second explanation for the empirical failure of the Consumption Capital Asset Pricing Model. In the expected utility framework that has been used extensively, the coefficient of relative risk aversion equals the reciprocal of the elasticity of intertemporal substitution. Given the fact that a relation between consumption growth and asset returns is solely governed by this elasticity of intertemporal substitution (Hall (1988)), it is possible that the coefficients that e.g. Mehra and Prescott (1985) and Grossman, Melino, and Shiller (1987) obtain, may be interpreted as the reciprocal of the elasticity of intertemporal substitution and do not necessarily reveal anything about true relative risk aversion. Low values of the intertemporal elasticity of substitution, as opposed to high values of relative risk aversion, do not present a puzzle. The concept of time-additive expected utility however is not suited for a clean separation of intertemporal substitution from risk aversion. Therefore the preference assumptions may be at the root of the problem, because in a CCAPM world with standard Von Neumann-Morgenstern preferences risk neutral agents must inevitably have an infinite elasticity of

intertemporal substitution. This is an implausible restriction.

Lastly, one is confronted with the fact that the conditional variance of excess stock returns is variable. This observation has two consequences : First of all in any empirical test one has to take this conditional heteroskedasticity into account, a feature that has been neglected in most of the early tests of the CCAPM but has been extensively used since the development of the class of GARCH models (see Bollerslev et. al. for a recent survey). Secondly as Attanasio (1991) and Hansen and Richard (1987) have pointed out the fact that excess returns are highly variable requires the conditional covariance between excess returns and some benchmark portfolio appearing in a pricing relation to be high and highly variable in order to be consistent with the data. Empirical evidence however suggests that the covariance between consumption growth and stock returns does not fulfil this condition. Hence the rejection of the CCAPM becomes apparent. Summarizing, we have a measurement, a preference and a heteroskedasticity problem.

The objective of this chapter is to explore particular solutions to these problems in a discrete intertemporal asset pricing context. More specifically, the consumption measurement problem will be handled through a simple substitution of consumption out of the derived asset pricing relationship, along the lines indicated by Campbell (1990). This may seem a rather drastic approach, but if an explicit solution for consumption from the intertemporal optimization problem is available, this may be used to eliminate consumption. Such a solution can be obtained by combining a log-linear budget constraint with a second order Taylor approximation to the resulting Euler equation for consumption, see Campbell (1990). Subsequently, this elimination relieves us from actually having to measure consumption. The second problem is addressed by assuming Kreps-Porteus non-expected utility preferences, which provide the desired clean separation of risk aversion from intertemporal substitution, stationarity of preferences and temporal consistency of optimal plans, see Kreps and Porteus (1978,1979a,b), Epstein and Zin (1989,1991), Epstein (1988), Weil (1989) and Giovannini and Weil (1989). The major consequence of this assumption is that equilibrium rates of return will depend on the covariance with both consumption growth and the return on the market. Assuming conditional heteroskedasticity for stock returns and consumption growth will impede the substitution of consumption

but specifying the conditional variance of the market return as evolving according to a GARCH process and using an assumed relation between actual consumption growth and the realized market return will pave the way. Ultimately, after having implemented the proposed modifications, the resulting intertemporal asset pricing model is quite similar to the relatively obscure three-moment CAPM, in which conditional skewness plays a prominent role.

This chapter proceeds as follows : In Section 3.2 the derivation of the approximate log linear budget constraint is described. Section 3.3 contains a discussion of the preference assumptions, the derivation of the Euler equations, and the resulting model for equilibrium returns. Section 3.4 presents the actual consumption substitution, with specific attention to the heteroskedastic case and the final asset pricing model. Section 3.5 considers empirical implementation and finally Section 3.6 presents a summary and conclusions.

3.2 The Linearization of a Non-Linear Intertemporal Budget Constraint

This section draws heavily on Campbell (1990) and is related to the well-known dividend-price ratio linearization of Campbell and Shiller (1988). From the outset the existence of a representative agent is assumed. The justification of this assumption will be discussed in section 3. For the representative agent the intertemporal budget constraint reads:

$$W_{t+1} = R_{m,t+1}(W_t - C_t) \quad 2.1$$

Where wealth at date $t+1$ equals the amount of invested wealth at date t multiplied by the gross return on invested wealth, earned in the time period between date t and date $t+1$. Labour income is entirely abstracted from¹.

¹ The reason for neglecting labour income arised from the difficulty of solving the dynamic problem with an extra term in the budget constraint. Epstein and Zin (1991) attempt to avoid this problem by assuming that the stream of labor income properly

When the limit of discounted future wealth is zero, one can show that current wealth equals current consumption plus the discounted value of future consumption streams or alternatively:

$$W_t - C_t = \sum_{i=1}^{\infty} C_{t+i} \left(\prod_{j=1}^i R_{m,t+j}^{-1} \right) \quad 2.2$$

Invested wealth equals the discounted value of future consumption streams. Note that the return on invested wealth is treated as the return on the market portfolio. The log-linear approximation starts by specifying the growth in wealth as:

$$\frac{W_{t+1}}{W_t} = R_{m,t+1} \left(1 - \frac{C_t}{W_t} \right) \quad 2.3$$

In natural logarithms 2.3 becomes:

$$\Delta w_{t+1} = r_{m,t+1} + \ln(1 - \exp(c_t - w_t)) \quad 2.4$$

$$w_{t+1} = \ln W_{t+1}, \quad r_{m,t+1} = \ln R_{m,t+1} \quad \text{and} \quad c_t - w_t = \ln(C_t/W_t).$$

In general, the second term on the right hand side of 2.4 is a non-linear function of the log consumption wealth ratio. Only when the consumption wealth ratio is a constant, say Π , this term reduces to a constant, and the intertemporal budget constraint is obviously log-linear.

discounted can be seen as a part of the initial wealth if labour income is non stochastic. Labour income could also be seen as a stochastic dividend for a nontraded asset, say human capital. In this view the market portfolio should be extended to include human capital. Neither solution however is satisfactory. The first solution is unrealistic, and the second one requires a problematic estimation of the proportion of human wealth in the economy.

$$\Delta w_{t+1} = r_{m,t+1} + \ln(1-\Pi) , \quad 0 < \Pi < 1 \quad 2.5$$

$$\Delta c_{t+1} = \Delta w_{t+1} = r_{m,t+1} + \ln(1-\Pi) , \quad 0 < \Pi < 1 \quad 2.6$$

From 2.5 and 2.6, it can be seen that in this case the growth in wealth and consumption must be smaller than the return on invested wealth. In general however the consumption growth can be larger than the return on invested wealth. For the non-constant case define π as the mean log consumption wealth ratio and take a first order Taylor expansion of the non-linear term around π . The approximation that results is :

$$\ln(1 - \exp(c_t - w_t)) \approx \ln(1 - e^\pi) - \frac{e^\pi}{1 - e^\pi} (c_t - w_t - \pi) \quad 2.7$$

Note that through Jensen's Inequality : $\pi < \ln \Pi$, if Π now is interpreted as the mean consumption wealth ratio. Only in the constant case $\pi = \ln \Pi$ holds and,

$$\frac{e^\pi}{1 - e^\pi} = \frac{C}{W-C}$$

is the constant ratio of consumption to invested wealth. In the remainder e^π will be substituted by $1-\rho$, a simple rewriting of a coefficient. From 2.4 and 2.7, the loglinear approximation to the intertemporal budget constraint can be written as:

$$\Delta w_{t+1} \approx r_{m,t+1} + \kappa + (1-1/\rho)(c_t - w_t) \quad 2.8$$

where κ is a constant and can be obtained from 2.7. Using the equality,

$$\Delta w_{t+1} = \Delta c_{t+1} + (c_t - w_t) - (c_{t+1} - w_{t+1}) \quad 2.9$$

Together with 2.8 and assuming that the limit of the discounted future log consumption-wealth ratio is zero, this eventually results in:

$$c_t - w_t = \sum_{j=1}^{\infty} \rho^j (r_{m,t+j} - \Delta c_{t+j}) + \frac{\rho \kappa}{1-\rho} \quad 2.10$$

This equation shows that a high (higher than average) consumption wealth ratio today can be achieved if future consumption growth is lower than future returns on invested wealth for several periods. It also shows that one can only sustain equation of future consumption growth to the future return on invested wealth, when the current consumption wealth ratio is way below average. These features are derived from a technical constraint and consequently they have no real economic content. Although equation 2.10 holds ex post as a mechanical construction, it really should be cast in ex ante form, by taking expectations, conditional on time t information:

$$c_t - w_t = E_t \sum_{j=1}^{\infty} \rho^j (r_{m,t+j} - \Delta c_{t+j}) + \frac{\rho \kappa}{1-\rho} \quad 2.11$$

Substituting 2.11 into 2.8 and 2.9 eventually results in:

$$r_{m,t+1} - E_t r_{m,t+1} = (E_{t+1} - E_t) \sum_{j=0}^{\infty} \rho^j \Delta c_{t+1+j} - (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r_{m,t+1+j} \quad 2.12$$

The unexpected return on invested wealth can be unraveled into two parts, the revision, at time $t+1$, of the expected discounted value of the future feasible consumption stream and the revision of the expected discounted value of future returns. Assume for a moment that the realized return is the only source of information our representative agent receives at time $t+1$. If this information does not alter his beliefs about future consumption or future returns, then the unexpected growth in consumption, at $t+1$, equals the unexpected return on invested wealth. This does not allow us to conclude

that consumption growth equals realized return, because we do not know what the expectations at time t were. If this source of positive (negative) information does not change his beliefs about future returns, and current consumption growth is maintained at the expected level, then his expectations about the future feasible consumption stream are revised upwards (downwards) due to a windfall (set-back) today. Similarly if the positive (negative) information does not alter his beliefs about future consumption plans, then his expectations about future returns are revised downwards (upwards). In general it is possible that new information has an effect on both sets of expectations. Equation 2.12 can be written as:

$$c_{t+1} - E_t c_{t+1} = (E_{t+1} - E_t) \sum_{j=0}^{\infty} \rho^j r_{m,t+1+j} - (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j \Delta c_{t+1+j} \quad 2.13$$

Now, similar situations as above can be discussed, but net effects can only be assessed if some model of optimal behavior is used, which will be the subject of the next sections, where also expected future consumption growth will be eliminated from 2.13, leaving us with only current and expected future returns.

3.3 The Utility and Optimal Behavior Framework

In the introduction to this chapter it was already mentioned that the attempts in the empirical literature to explain the behavior of consumption and asset returns over time, using expected utility, representative agent, optimizing models, have encountered some serious difficulties, see Hansen and Singleton (1983), and Mehra and Prescott (1985). A possible explanation for the disappointing results is the rigidity of the expected utility framework. Lucas (1978) noted that with time additive utility there is no way to disentangle intertemporal substitutability from relative risk aversion, two conceptually distinct aspects of preferences. Furthermore the restriction that the product of the elasticity of intertemporal substitution and the coefficient of relative risk aversion is equal to one, implies that agents are indifferent to the timing of

resolution of uncertainty regarding consumption plans. However it is also plausible to claim that agents prefer early resolution because this improves planning. When risk aversion and the elasticity of intertemporal substitution are no longer each others' reciprocal, early (late) resolution is preferred whenever the product of these parameters is larger (smaller) than one. Inducing time non-indifference or relaxing "the axiom of reduction of compound lotteries", for temporal gambles which is a fundamental axiom to the Von Neumann-Morgenstern utility theory, is the key concept in the class of Kreps-Porteus preferences, see Kreps and Porteus (1978,1979a,b).

This class of non expected utility preferences provide the desired separation of risk aversion from intertemporal substitution (allowing time non-indifference), stationarity of preferences and temporal consistency of optimal plans, see Epstein & Zin (1989), and Weil (1987). Kreps & Porteus (1978) show that monotonicity, regularity, consistency and stationarity conditions combined with the independence axiom for irrelevant alternatives are sufficient for the following recursive preference representation to obtain:

$$V_t = U(C_t, E_t V_{t+1}) \quad 3.1$$

V_t denotes utility at time t , E_t is the conditional (on time t information) expected value operator, and U is the so called aggregator function, because it combines current consumption and expected future utility to determine current utility. If the aggregator function would be linear in its second argument, then the expected utility framework would obtain, as can be seen by forward substitution and the law of iterated expectations. Then 3.1 simplifies to:

$$V_t = E_t \sum_{i=0}^{\infty} (U_2)^i U(C_{t+i}) \quad 3.2$$

Current utility is the expected discounted sum of future "felicities", see Giovannini and Weil (1989), where the discount factor is a constant, the derivative of the aggregator function with respect to its second argument,

U_2 . The standard expected utility representation is a special case of Kreps-Porteus preferences. In general the aggregator function can be convex or concave in its second argument leading to a preference for early, respectively late resolution of uncertainty.

The particular parametric form of Kreps-Porteus preferences used in this chapter is based on the isoelastic utility generalizations proposed by Epstein & Zin (1989, 1991) and Weil (1987), a form which is also used in Weil (1989), Giovannini and Weil (1989) and Campbell (1990). This parameterization leads to Generalized Isoelastic Preferences (GIP), and is achieved along the following lines. First introduce a help function H_t and two other functions $R(x)$ and $I(C,x)$ as:

$$H_t = U^{\frac{1}{1-\gamma}} = V_t^{\frac{1}{1-\gamma}} \quad 3.3$$

$$R(x) = x^{1-\gamma} \quad 3.4$$

$$I(C, x) = ((1-\delta) c^{1-\rho} + \delta x^{1-\rho})^{\frac{1}{1-\rho}}, \quad 0 < \delta < 1 \quad 3.5$$

Define $\gamma \geq 0$, $\gamma \neq 1$, and $1/\rho \geq 0$, $\rho \neq 1$, (which differs from the linearization parameter ρ of section 3.2) as respectively the constant coefficient of relative risk aversion and the constant elasticity of intertemporal substitution. Then attitudes towards risk are captured by $R(x)$ and attitudes towards intertemporal substitution are captured by $I(C,x)$. We want current utility to be the aggregate (through $I(C,x)$) of current consumption and the certainty equivalent of future utility (determined by $R(x)$). This can be achieved by setting :

$$H_t = I(C_t, R^{-1}(E_t R(H_{t+1}))) \quad 3.6$$

Through the definition of $I(C, x)$ and $R(x)$ this leads to:

$$H_t = ((1-\delta) C_t^{1-\rho} + \delta (E_t (H_{t+1}^{1-\gamma}))^{\frac{1-\rho}{1-\gamma}})^{\frac{1}{1-\rho}} \quad 3.7$$

which ultimately through the definition of H_t leads to:

$$V_t = U(C_t, E_t V_{t+1}) = \left((1-\delta) C_t^{1-\rho} + \delta (E_t V_{t+1})^{\frac{1-\rho}{1-\gamma}} \right)^{\frac{1}{1-\rho}} \quad 3.8$$

3.8 is exactly the same representation as in Giovannini and Weil (1989). One can observe that γ and $1/\rho$ are distinct parameters, time non-indifference is allowed for, and the expected utility framework obtains as a special case whenever γ equals ρ . Furthermore preferences are seen to be homothetic: assuming that these preferences are common to all agents, and that these agents only differ with respect to their initial endowments of wealth, then the representative agent assumption can be justified, using the Grossman-Shiller (1982) aggregation theorem. One can now determine the optimal consumption and portfolio decisions for the infinitely lived representative agent, in a pure exchange economy, which is similar to the one used in Lucas (1978) except for the agents' preferences. We assume that there are n financial assets in which the agent can invest, and which offer random gross rate of return $R_{i,t+1}$, $i=1,\dots,n$, in the period between date t and date $t+1$. We define the return on invested wealth, or the market return, as a return on an optimally weighted portfolio, where the optimal weights, $\alpha_{i,t}$, $i=1,\dots,n$, are determined at time t :

$$R_{m,t+1} \equiv \sum_{i=1}^n \alpha_{i,t} R_{i,t+1}, \quad \sum_{i=1}^n \alpha_{i,t} = 1 \quad 3.9$$

The objective is to maximize utility, defined by 3.8, by choosing a sequence of consumption plans and portfolio shares, which are subject to 3.9, and the

intertemporal budget constraint 2.1. Denote the information set at time t as S_t which only contains the history of asset returns prior to, but also at, time t . We characterize the resulting optimal program, in terms of a value function $V(W, S)$ which solves the described maximization problem.

$$V(W_t, S_t) = \underset{C_t, \{\alpha_i\}_{i=1}^n}{\operatorname{argmax}} U(C_t, E_t V(W_{t+1}, S_{t+1})) \quad 3.10$$

subject to 3.9 and 2.1. Epstein & Zin (1991) were the first to derive the first order conditions for optimal consumption and portfolio selection. The optimal consumption program can be determined independently from the optimal portfolio problem, because preferences are homothetic. The first order condition for the consumption optimization problem in 3.10 is:

$$U_{1,t} = U_{2,t} E_t(R_{m,t+1} V_{1,t+1}) \quad 3.11$$

Here $U_{i,t}$ denotes the derivative of the aggregator function with respect to its i -th ($i=1,2$) argument evaluated at $(C_t, E_t V_{t+1})$, and V_1 is the derivative of the value with respect to wealth evaluated at (W_t, S_t) . The Envelope condition says that the marginal value of wealth is equal to the marginal utility of consumption, in a consumption optimum or: $U_{1t} = V_{1t}$. Using this in the first order condition in 3.11, leads to:

$$E_t \left\{ R_{m,t+1} U_{2,t} \frac{U_{1,t+1}}{U_{1,t}} \right\} = 1 \quad 3.12$$

This is the fundamental Euler equation for consumption. Considering portfolio selection, the optimization procedure results in the following conditions for any two assets i and j held by the agent :

$$E_t \{ U_{I,t+1} R_{i,t+1} \} = E_t \{ U_{I,t+1} R_{j,t+1} \} \quad 3.13$$

In equilibrium the representative agent holds the market portfolio so 3.13 also applies to the market portfolio. For any asset i , that is held one can now write:

$$E_t \left\{ R_{i,t+1} U_{2,t} \frac{U_{I,t+1}}{U_{I,t}} \right\} = I \quad 3.14$$

This is the Euler equation for optimal portfolio selection. We have obtained the familiar result that in equilibrium, an asset's excess return is determined by its covariance with the marginal rate of substitution.

$$MRS_{t+1} = U_{2,t} \frac{U_{I,t+1}}{U_{I,t}} \quad 3.15$$

This marginal rate of substitution is computed in Appendix 3.A (following Giovannini & Weil (1989), but with slight modifications). The general result is:

$$MRS_{t+1} = \left(\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\rho} \right)^{\frac{1-\gamma}{1-\rho}} (R_{m,t+1})^{\frac{1-\gamma-1}{1-\rho}} \quad 3.16$$

The marginal rate of substitution is a geometric weighted average of the return on the market portfolio and the consumption growth rate. Equilibrium rates of return are thus determined by both their covariance with consumption growth as well as their covariance with the market return. The resulting asset pricing equation contains static and intertemporal CAPM elements. As we shall see, the static CAPM and the consumption CAPM

only hold as special cases to this more general theory. Substituting 3.16 into 3.12 and 3.14 leads to the following Euler Equations.

$$E_t \left\{ \delta \left(\frac{C_{t+1}}{C_t} \right)^{-\rho} R_{m,t+1} \right\}^{\frac{1-\gamma}{1-\rho}} = 1 \quad 3.17$$

and

$$E_t \left\{ \left(\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\rho} \right)^{\frac{1-\gamma}{1-\rho}} (R_{m,t+1})^{\frac{1-\gamma-1}{1-\rho}} R_{i,t+1} \right\} = 1 \quad 3.18$$

See also Giovannini and Weil (1989), Weil (1989) and Campbell (1990). In the remainder the following definitions will be used:

$$\theta = (1-\gamma)/(1-\rho) \quad , \quad \sigma = 1/\rho \quad 3.19$$

The Euler Equations can be written in log form, if one is willing to assume that asset returns and consumption growth are jointly lognormally distributed or if one is willing to work with a second order Taylor approximation. The former approach is used in Giovannini and Weil (1989) and Campbell (1990). The latter approach is employed here. At this point in the chapter we do not wish to make any distributional assumptions. The rewriting of the Euler equations in log form is justified in Appendix 3.B. The log version of 3.17 is simply:

$$0 = \theta \ln \delta - \frac{\theta}{\sigma} E_t \Delta c_{t+1} + \theta E_t r_{m,t+1} + \frac{1}{2} \left(\left(\frac{\theta}{\sigma} \right)^2 \sigma_{c,t}^2 + \theta^2 \sigma_{m,t}^2 - \frac{2\theta^2}{\sigma} \sigma_{cm,t} \right) \quad 3.20$$

Lower case letters indicate logarithms, $\sigma_{c,t}^2$, $\sigma_{m,t}^2$ and $\sigma_{cm,t}$ denote respec-

tively, the conditional variances at time t , of consumption growth and the market return at time $t+1$, and the conditional covariance between consumption growth and the market return. Equation 3.20 displays a linear relationship between expected consumption growth and the expected market return. This linear relation, as we shall see, plays an important role in the elimination of consumption. The relationship is:

$$E_t \Delta c_{t+1} = \mu_{m,t} + \sigma E_t r_{m,t+1}, \quad \mu_{m,t} = \sigma \ln \delta + \frac{1}{2} \frac{\theta}{\sigma} \text{Var}_t(\Delta c_{t+1} - \sigma r_{m,t+1}) \quad 3.21$$

This relation is not a new one in finance theory, but whereas previous studies had difficulties in interpreting the slope coefficient, see Hall (1988), the interpretation here is quite clear. The slope coefficient is the inter-temporal elasticity of substitution, σ . The intercept term, $\mu_{m,t}$, may or may not be a constant depending on whether we live in a homoskedastic or heteroskedastic world. In any case it can be interpreted as the autonomously expected growth in consumption and is related to the conditional variance of the residual in an ex-post version of the same relation. The relation 3.21 suggests that the representative agent makes a provision, *ex ante*, for the fact that plans made at time t , will not be realized at time $t+1$. Note that this relation is not a complete description of consumption behavior under uncertainty. It does not describe the actual amount by which consumption changes when new information about asset returns becomes available. Hall (1988) has estimated σ for several U.S. data sets, and he concluded that σ is unlikely to be much above 0.1 and may well be zero. In our setting a low value of σ means that the value and sign of θ are important in determining whether the representative agent plans to accelerate or postpone consumption in response to a high conditional variance of the error term in an ex post relation. The second Euler equation, written in log form, can be used for cross-sectional asset pricing. The log version of 3.18 becomes:

$$0 = \theta \ln \delta - \frac{\theta}{\sigma} E_t \Delta c_{t+1} + (\theta - 1) E_t r_{m,t+1} + E_t r_{i,t+1} + \frac{1}{2} \left(\left(\frac{\theta}{\sigma} \right)^2 \sigma_{c,t}^2 + (\theta - 1)^2 \sigma_{m,t}^2 + \sigma_{i,t}^2 - \frac{2\theta}{\sigma} (\theta - 1) \sigma_{cm,t} - \frac{2\theta}{\sigma} \sigma_{ci,t} + 2(\theta - 1) \sigma_{im,t} \right) \quad 3.22$$

Equation 3.22 also holds, in a simplified form, for an asset with a risk-free real return, $r_{f,t+1}$, where $\sigma_{c,t}^2 = \sigma_{cf,t} = \sigma_{fm,t} = 0$. Subtracting the risk free version from the general version (3.22) results in :

$$E_t r_{i,t+1} - r_{f,t+1} = -\frac{1}{2} \sigma_{i,t}^2 + \frac{\theta}{\sigma} \sigma_{ci,t} + (1 - \theta) \sigma_{im,t} \quad 3.23$$

For the expected excess return on the market portfolio one obtains :

$$E_t r_{m,t+1} - r_{f,t+1} = \frac{\theta}{\sigma} \sigma_{cm,t} + \left(\frac{1}{2} - \theta \right) \sigma_{m,t}^2 \quad 3.24$$

From 3.23 it is immediately obvious that an asset's expected excess return is determined by its own conditional variance (due to a Jensen's Inequality effect, see Appendix 3.B), and a weighted combination of the asset's conditional covariance with consumption growth divided by σ , and its conditional covariance with the return on invested wealth. The first covariance receives a weight of θ and the second a weight of $(1-\theta)$. A few special cases to this general asset pricing model can be discerned. Whenever γ equals ρ (implying $\theta=1$), one returns to a VNM expected utility setting and the consumption CAPM emerges as a special case, as can be seen from the Euler equations 3.17 and 3.18. However when $\gamma=\rho=1$, logarithmic utility with myopia in consumption and in portfolio selection obtains, and the static CAPM is recovered. The static CAPM also emerges as a special case when $\gamma=1$ and γ is not equal to ρ . For a more general treatment see Giovannini and Weil (1989). Epstein and Zin (1991) obtain GMM estimates for the relevant preference parameters using a system of four different equity returns based on 3.23. They conclude that in general the elasticity of intertemporal substitution is smaller than one, and relative risk aversion is close to one. Their results however, are sensitive to the choice of the consumption measure. Weil (1989) concludes that simply separating risk aversion from intertemporal substitution will not solve the equity premium puzzle.

Recent work by Sundaresan (1989) and Constantinides (1990) on habit formation using time non-separable preferences, suggests that habit formation may explain the equity premium puzzle. The class of preferences that is explored in both articles however, exhibits payoff history dependence, i.e. today's preferences depend on previously realized consumption levels, or preferences are backward looking. The class of preferences at hand in this paper assume payoff history independence, and thus preferences are forward looking. Giovannini and Weil (1989) corroborate the findings of Epstein and Zin (1991) regarding the preference parameters, and conclude that portfolio choice seems to conform mainly to static considerations, lending further empirical support for the static CAPM.

Attanasio (1991) proposed a test that nests the static and the consumption CAPM, albeit under an expected utility framework, and which relies on the existence of conditional heteroskedasticity. His results support the static CAPM as well, a finding that is probably caused by the low conditional covariance of consumption growth with excess asset returns. Although, allowing for time and state non-separable preferences and obtaining a separation for preference parameters seems an enterprise worth undertaking, it will not solve all problems associated with consumption and asset pricing theory. The problems related to measured consumption and the possible conditional heteroskedasticity are still present. An attempt to solve these problems is presented in the next section.

3.4 The Elimination of Consumption

Given the problems associated with measured consumption, mentioned in the introduction to this chapter, it seems legitimate to try and substitute or eliminate consumption from the analysis. This elimination is achieved by combining the linear relationship between expected consumption growth and the return on invested wealth, derived from the Euler equation for consumption, with the results obtained in section 2 on the intertemporal budget constraint. This procedure is entirely valid for the homoskedastic case. Heteroskedasticity appears to be more difficult to handle. We'll start however with the homoskedastic case in order to illustrate the elimination

procedure, which is adapted from Campbell (1990). In a homoskedastic setting, equation 3.21 simply becomes :

$$E_t \Delta c_{t+1} = \mu_m + \sigma E_t r_{m,t+1}, \quad \mu_m = \sigma \ln \delta + \frac{1}{2} \frac{\theta}{\sigma} \text{Var}(\Delta c_{t+1} - \sigma r_{m,t+1}) \quad 4.1$$

The intercept term now is a constant, as the variance term in 4.1 is a constant. The first step in the substitution procedure is to substitute 4.1 into 2.11 to obtain :

$$c_t - w_t = (1-\sigma) E_t \sum_{j=1}^{\infty} \rho^j r_{m,t+j} + \frac{\rho(\kappa - \mu_m)}{1-\rho} \quad 4.2$$

From 4.2 one can observe that if expected returns rise and $\sigma < 1$, ($\sigma > 1$), the representative agent is reluctant (willing) to substitute intertemporally; the income effect of higher expected returns dominates (is dominated by) the substitution effect, thereby raising (lowering) today's consumption wealth ratio. Whenever $\sigma = 1$, the consumption wealth ratio is a constant. Substituting 4.1 into 2.13 gives the innovation in consumption as a function of the current market return and the revision in expectations about future returns :

$$c_{t+1} - E_t c_{t+1} = r_{m,t+1} - E_t r_{m,t+1} + (1-\sigma)(E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r_{m,t+1+j} \quad 4.3$$

Again one can observe that consumption can rise (decline) in response to an increase in expected future returns depending on σ being smaller (larger) than one. The current unexpected return on invested wealth has a one-to-one effect on unexpected consumption. We have derived an expression for the innovation in consumption that does not contain any reference to consumption itself. Note that in this setting, a possible revision of expected future returns can only occur due to new information received at time $t+1$, i.e. $r_{m,t+1}$. Given 4.3 we are able to determine any asset's covariance with

consumption growth :

$$\text{Cov}(r_{i,t+1}, \Delta c_{t+1}) = \sigma_{ic} = \sigma_{im} + (1-\sigma) \sigma_{ih} \quad 4.4$$

$$\sigma_{ih} = \text{Cov}\left(r_{i,t+1}, (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r_{m,t+1+j}\right) \quad 4.5$$

σ_{ih} is the covariance of asset i 's return with upward revisions in expected future market returns. It can be interpreted as a covariance with a hedge portfolio that represents changes in the stochastic investment opportunity set, see Merton (1973). Substituting 4.5 into 3.23 and using the definitions of θ and σ results in an asset pricing equation from which consumption has completely vanished :

$$E_t r_{i,t+1} - r_{f,t+1} = -\frac{1}{2} \sigma_i^2 + \gamma \sigma_{im} + (\gamma - 1) \sigma_{ih} \quad 4.6$$

Equilibrium asset returns are determined by their own variance, their covariance with the return on invested wealth, and their covariance with a "hedge" portfolio. This last term appears because first moments are allowed to vary, and we have thus created a stochastic investment opportunity set. The only preference parameter appearing in 4.6 is γ , the coefficient of relative risk aversion. Together with consumption, the elasticity of intertemporal substitution disappears from the asset pricing equation. One explanation for this disappearance is the dual role that σ plays in theory. Suppose σ is high, which means that the agent is willing to substitute intertemporally, so he does not mind large fluctuations in consumption. This leads to a low required risk premium to compensate for such fluctuations (see 3.23). A high value of σ leads to high anticipated fluctuations in consumption but at the same time lowers the required risk premium. The net result of these offsetting effects is that σ drops out of the asset pricing equation. Note that whenever σ equals one, even the asset's covariance with the hedge portfolio drops out. Equilibrium asset returns are, apart from their

own variances, completely determined by their covariance with the market return, as in the static CAPM.

In essence 4.6 decomposes the expected risk premium into two terms. One appearing with coefficient γ and the other with coefficient $\gamma-1$. When the agent is highly risk averse, $\gamma > 1$, he not only requires a high rate of return for holding assets that have a positive covariance with the market return, but also for assets that have a positive covariance with the news on future market returns. This phenomenon occurs because holding these assets reduces the agent's ability to hedge against a deterioration in investment opportunities; this in spite of the fact that holding these assets would enable the agent to profit from a surge in investment opportunities. When $\gamma < 1$, matters are reversed and the positive "profitability" effect dominates the negative "non-hedging" effect. The borderline case : $\gamma=1$, causes both effects to cancel out, and leaves us with a static CAPM result. The static CAPM version to 4.6 emerges in a few special cases. The cases $\sigma=1$ and $\gamma=1$ were already mentioned. Furthermore under logarithmic utility ($\gamma=\sigma=1$), when the investment opportunity set is constant ($\sigma_{ih} = 0$), and when the market return follows a univariate stochastic process², the static CAPM is the result, (see Campbell (1990)).

Unfortunately we live in a heteroskedastic world, and consequently the analysis above is of limited value. It is however of some use as it helps to highlight the consequences of incorporating heteroskedasticity into the analysis. One consequence is immediately apparent. By allowing conditional second moments to vary, the stochastic opportunity set is in a way extended. Now the stochastics are not only represented by time-varying conditional first moments but also by time-varying second moments. This should be reflected in the resulting asset pricing equation. The inclusion of time-varying second moments also changes the linear relationship between expected consumption growth and the expected return on invested wealth, 4.1, back to its original formulation 3.21.

² In this case news about future returns is perfectly correlated with the current return. Hence there is a coefficient ϕ such that $\sigma_{ih} = \phi\sigma_{im}$ for all assets i, see Campbell (1990).

$$E_t \Delta c_{t+1} = \mu_{m,t} + \sigma E_t r_{m,t+1} , \quad \mu_{m,t} = \sigma \ln \delta + \frac{1}{2} \frac{\theta}{\sigma} \text{Var}_t(\Delta c_{t+1} - \sigma r_{m,t+1}) \quad 4.7$$

The intercept term now changes over time, as the conditional variance of consumption growth minus σ times the market return changes over time. Now one can proceed as before and substitute 4.7 into 2.11. The log consumption wealth ratio is determined by :

$$c_t - w_t = (1 - \sigma) E_t \sum_{j=1}^{\infty} \rho^j r_{m,t+j} - E_t \sum_{j=1}^{\infty} \rho^j \mu_{m,t+j-1} + \frac{\rho \kappa}{1 - \rho} \quad 4.8$$

One extra term appears in 4.8 that was not present in 4.2. This term causes the log consumption wealth ratio to decline, *ceteris paribus*, whenever an increase in the uncertainty of consumption growth relative to the market return is expected. This reflects prudent behavior by the agent. The innovation in consumption also inherits an extra term :

$$\begin{aligned} c_{t+1} - E_t c_{t+1} &= r_{m,t+1} - E_t r_{m,t+1} + (1 - \sigma)(E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r_{m,t+1+j} \\ &\quad - (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j \mu_{m,t+j} \end{aligned} \quad 4.9$$

In 4.9 one can observe the same kind of prudent behavior, as discussed for 4.8. The major problem in 4.9 is the fact that the right hand side still refers to consumption through the variable $\mu_{m,t}$, this in contrast to the homoskedastic case 4.3, where $\mu_{m,t}$ was just a constant and consequently disappeared³. Therefore we cannot use 4.9 directly to eliminate consumption from the

³ In the heteroskedastic case $\mu_{m,t}$ can also become a constant when $\gamma=1$ (because then θ is zero), or when $\sigma=1$ ($\theta=\infty$) because the associated variance term must be zero to guarantee that expected consumption growth remains finite.

intertemporal asset pricing model. We have to resort to an additional assumption on the functional specification for $\mu_{m,t}$. Campbell (1990) assumes that the intercept term is a linear function of the expected market return :

$$\mu_{m,t} = \mu_0 + \psi E_t r_{m,t+1} \quad 4.10$$

Although an appealing assumption, it obscures the linear relationship 4.7 and ignores the variance term defining $\mu_{m,t}$. This assumption causes the right hand side of 4.9 to bear no reference to consumption anymore, and one could once more easily eliminate consumption, see again Campbell (1990). To justify this assumption Campbell needs the conditional variance of the market return to be a linear function of the expected market return, and all other conditional second moments to be linear functions of the conditional variance of the market return. Obviously this implies that the expected excess market return can be written as a linear function of its conditional variance :

$$E_t r_{m,t+1} - r_{f,t+1} = \beta_0 + \beta_1 \sigma_{m,t+1}^2 \quad 4.11$$

It should be noted however that the market price of conditional volatility, β_1 , cannot be interpreted as the coefficient of relative risk aversion, as is often assumed. The current trend in financial econometrics is to estimate 4.11 in combination with an (E)GARCH specification for the conditional variance of the market return, see e.g. Bollerslev, Engle and Wooldridge (1988), French, Schwert, and Stambaugh (1987), Baillie and DeGennaro (1990), and Nelson (1991) to name but a few. The condition that led to 4.11 was :

$$\sigma_{m,t+1}^2 = \alpha_0 + \alpha_1 E_t r_{m,t+1} \quad 4.12$$

Which does not correspond to a GARCH (p,q) specification like :

$$\sigma_{m,t+1}^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{m,t+1-i}^2 + \sum_{j=1}^p \beta_j \sigma_{m,t+1-j}^2 \quad 4.13$$

The point we are trying to illustrate here is that specifying the conditional variance of the market return as a GARCH process, which is common practice, may have consequences for the resulting asset pricing model, causing 4.11 not to hold any longer. Therefore the immediate combination of 4.11 and 4.13 and the subsequent joint estimation may only reveal useful information about the GARCH parameters. The consequences of adopting a GARCH process for the conditional variance of the market return one stage earlier in the analysis will become clear in the remainder of this section. We will replace assumption 4.10 by the following :

$$\mu_{m,t} = \mu_0 + \psi E_t \varepsilon_{m,t+1}^2 \quad , \quad E_t (\varepsilon_{m,t+1}) = 0 \quad 4.14$$

The intercept term is a linear function of the conditional variance of the market return. This implies that the conditional variance of consumption growth and the conditional covariance between consumption growth and the market return are linear functions of the conditional variance of the market return. This assumption can be justified in several ways, all of which are contained in Appendix 3.C. Having made this assumption it is also possible to characterize ψ as :

$$\psi = \frac{1}{2} \frac{\sigma}{\theta} (1 - \gamma)^2 \quad 4.15$$

This would be identically true in a few special cases (see Appendix 3.C). Here it is treated as a convenient approximation. Combining 4.14 and 4.15 and substituting the result in 4.9 leads to :

$$\begin{aligned}
 c_{t+1} - E_t c_{t+1} &= r_{m,t+1} - E_t r_{m,t+1} + (1-\sigma) (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r_{m,t+1+j} \\
 &\quad - \frac{1}{2} \frac{\sigma}{\theta} (1 - \gamma)^2 (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j \varepsilon_{m,t+1+j}^2
 \end{aligned} \tag{4.16}$$

The agent's decision to actually accelerate or postpone consumption is determined by the unexpected market return, the revision in expectations about future returns, the revision in expected future uncertainty regarding the market return and his preference parameters. The next step is to specify the evolution of the conditional variance of the market return. We assume that this conditional variance follows a GARCH(p,q) process, where the conditional variance is a linear function of past squared error terms and past conditional variances :

$$E_t (\varepsilon_{m,t+1}^2) = \sigma_{m,t+1}^2 = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{m,t+1-i}^2 + \sum_{j=1}^p \beta_j \sigma_{m,t+1-j}^2 \tag{4.17}$$

By now there is a vast amount of literature available regarding the theory and applications of GARCH processes. A discussion on the advantages and disadvantages of this class of stochastic processes falls beyond the scope of this chapter ; for a recent and excellent survey see Bollerslev et. al. (1992). At this point it suffices to note that a GARCH process constitutes a simple and elegant vehicle to describe the evolution of conditional volatility. Pantula (1986) has pointed out that 4.17 can be rewritten as an ARMA process in squared residuals :

$$\varepsilon_{m,t+1}^2 = \omega + \sum_{i=1}^m (\alpha_i + \beta_i) \varepsilon_{m,t+1-i}^2 + (\varepsilon_{m,t+1}^2 - \sigma_{m,t+1}^2) - \sum_{j=1}^p \beta_j (\varepsilon_{m,t+1-j}^2 - \sigma_{m,t+1-j}^2) \tag{4.18}$$

Where $m = \max\{p, q\}$, $\alpha_i \equiv 0$ for $i > q$, and $\beta_i \equiv 0$ for $i > p$. Equation 4.18 shows that the squared residuals follow an ARMA(m, p) process with serially uncorrelated innovations. In appendix 3.D it is shown that according to a well known result for the discounted sum of revisions in the expected future values of an ARMA process, the last term in 4.16 can be rewritten as (see Flavin (1981)) :

$$(E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j \varepsilon_{m,t+1+j}^2 = \lambda (\varepsilon_{m,t+1}^2 - \sigma_{m,t+1}^2) \quad 4.19$$

Where λ depends on the order of the GARCH process. For a GARCH(1,1) process λ is given by :

$$\lambda = \frac{\rho \alpha_1}{1 - \rho (\alpha_1 + \beta_1)} \quad 4.20$$

Now 4.16 can be rewritten as :

$$\begin{aligned} c_{t+1} - E_t c_{t+1} &= r_{m,t+1} - E_t r_{m,t+1} + (1 - \sigma) (E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j r_{m,t+1+j} \\ &\quad - \frac{1}{2} \frac{\sigma}{\theta} (1 - \gamma)^2 \lambda (\varepsilon_{m,t+1}^2 - \sigma_{m,t+1}^2) \end{aligned} \quad 4.21$$

Comparing 4.21 to 4.16 shows that an infinite sum of revised expectations of future conditional variances has been reduced to the product of today's squared surprise minus its conditional expectation and a parameter that indexes the persistence of changes to the conditional variance. Consequently the conditional covariance between the return on any asset i and consumption growth can be specified as :

$$\sigma_{ic,t} = \sigma_{im,t} + (1 - \sigma) \sigma_{ih,t} - \frac{1}{2} \frac{\sigma}{\theta} (1 - \gamma)^2 \lambda \sigma_{ie^2,t} \quad 4.22$$

where $\sigma_{im,t}$ and $\sigma_{ih,t}$ are defined as before, and :

$$\sigma_{ie^2,t} = Cov_t(r_{i,t+1}, \varepsilon_{m,t+1}^2) = E_t((r_{i,t+1} - E_t r_{i,t+1})(r_{m,t+1} - E_t r_{m,t+1})^2) \quad 4.23$$

Substituting 4.22 into 3.23 and using the definitions of σ and θ gives :

$$E_t r_{i,t+1} - r_{f,t+1} = -\frac{1}{2} \sigma_{i,t}^2 + \gamma \sigma_{im,t} + (\gamma - 1) \sigma_{ih,t} - \frac{1}{2} (1 - \gamma)^2 \lambda \sigma_{ie^2,t} \quad 4.24$$

This specification for equilibrium asset returns shows that asset returns can be determined without a direct reference to their covariance with consumption growth, using instead their conditional covariances with the market return, news about future market returns, and the squared unexpected market return. This last term appears because of the heteroskedastic setting and the assumptions that were made on the evolution of heteroskedasticity. The only preference parameter appearing in 4.24 is γ , the coefficient of relative risk aversion ; as before σ , the elasticity of intertemporal substitution has completely vanished. In a situation of unit relative risk aversion, asset returns are only determined by their covariance with the market return, as in the static CAPM. In general one can see that the agent not only seeks protection against a deterioration in investment opportunities (a downward mean shift), which was already recognized above, but also against changes in the variability of investment opportunities (a variance shift). The degree to which assets can offer such a protection is given by their conditional covariance with the squared unexpected market return, which we will henceforth name the conditional coskewness (see Ingersoll (1987) p.100). If the conditional coskewness is positive, then the asset under consideration can provide protection against an increase in the variability of investment opportunities, which, ceteris

ceteris paribus, is attractive to the agent. Therefore the asset will sell at a relatively high price and thus will earn a lower expected rate of return (λ is positive by construction). For an extremely risk averse agent this effect is very important and it can even be fortified when a variance shift is persistent (indexed by λ). An asset i 's coskewness is positive if the return on that asset, r_i , tends to be above its conditionally expected value whenever the return on the remainder of the market portfolio (asset i is excluded) is exceptionally high or low, and below its expected value when the return on the remainder of the market portfolio is close to its expected value. This means that the density function for the market return tends to shift to the right for very low and very high values and to the left for intermediate values, which leads to positive skewness. Alternatively, assets with a negative coskewness do not provide any protection against a variance shift, causing the agent to require a higher rate of return, *ceteris paribus*, for holding these assets. For the market return 4.24 simplifies to :

$$E_t r_{m,t+1} - r_{f,t+1} = \left(\gamma - \frac{1}{2} \right) \sigma_{m,t}^2 + (\gamma - 1) \sigma_{mh,t} - \frac{1}{2} (1-\gamma)^2 \lambda \mu_{3,m,t} \quad 4.25$$

$$\mu_{3,m,t} = E_t ((r_{m,t+1} - E_t r_{m,t+1})^3) \quad 4.26$$

Here $\mu_{3,m,t}$ is the conditional third central moment, or the unnormalized conditional skewness. The intuition here is somewhat simpler than above. Positive skewness is looked upon favorably by the agent, which induces him to require, *ceteris paribus*, a lower rate of return. The reverse is true for negative skewness. We have derived an asset pricing relationship reminiscent of the three moment CAPM, through specific assumptions on the relation between actual consumption growth and the realized market return, the assumption of a GARCH process for the conditional variance of the market return and the subsequent elimination of consumption. Going one step further and specifying :

$$\sigma_{ih,t} = \varphi \sigma_{im,t} \quad 4.27$$

which can be derived by giving the market return an ARMA specification and using the aforementioned trick once more, the following relation can be derived by dividing 4.24 by 4.25 and using 4.27 :

$$\frac{\ln E_t (R_{i,t+1}/R_f)}{\ln E_t (R_{m,t+1}/R_f)} = \alpha_t \frac{\sigma_{im,t}}{\sigma_{m,t}^2} + (1-\alpha_t) \frac{\sigma_{is^2,t}}{\mu_{3,m,t}} \quad 4.28$$

or:

$$\beta_{i,t}^T = \alpha_t \beta_{i,t}^m + (1-\alpha_t) \beta_{i,t}^s \quad 4.29$$

$\beta_{i,t}^T$ is the true conditional beta of asset i . $\beta_{i,t}^m$ is the conditional market beta of asset i , and $\beta_{i,t}^s$ is the systematic skewness of asset i , see Tan (1991).

$$\alpha_t = \frac{(\gamma + \varphi(\gamma-1)) \left(\begin{array}{c} \sigma_{m,t}^2 \\ \mu_{3,m,t} \end{array} \right)}{(\gamma + \varphi(\gamma-1)) \left(\begin{array}{c} \sigma_{m,t}^2 \\ \mu_{3,m,t} \end{array} \right) - \frac{1}{2}\lambda(1-\gamma)^2} \quad 4.30$$

The relations 4.28 and 4.29 can be interpreted as conditional three moment CAPM relations. Note that if $\gamma = 1$ then the true conditional beta equals the conditional market beta. In general however the true time-varying beta is a linear combination of the conditional market beta and the systematic skewness. As we shall see in the next section, the estimation of relations like 4.24, 4.25, and 4.29 is not at all straightforward.

3.5 Empirical Implementation

Theoretically, the relations derived in the previous section have intuitive appeal. For instance, the modified version of 4.25,

$$E_t r_{m,t+1} - r_{f,t+1} = \left(\left(\gamma - \frac{1}{2} \right) + \varphi (\gamma - 1) \right) \sigma_{m,t+1}^2 - \lambda (1 - \gamma)^2 \mu_{3,m,t+1}$$

$$\sigma_{m,t+1}^2 = E_t (\varepsilon_{m,t+1}^2) = \omega + \sum_{i=1}^q \alpha_i \varepsilon_{m,t+1-i}^2 + \sum_{j=1}^p \beta_j \sigma_{m,t+1-j}^2$$

$$\mu_{3,m,t+1} = E_t (\varepsilon_{m,t+1}^3) \quad 5.1$$

shows that the equity premium, for let us say a broad stock market index, increases when the conditional variance increases and when the conditional unnormalized skewness becomes more and more negative. Large equity premia could therefore be (partially) explained by a large negative skewness parameter. In other words the occasional occurrence of extremely negative returns causes the agent to increase his required return. Empirical evidence suggests that stock index returns are in fact negatively skewed (see e.g Campbell and Hentschel (1992), Nelson (1991) and Nieuwland (1991)).

Before one can turn to the actual estimation and testing of 5.1 several problems have to be solved. First of all there is an identification problem, which can be solved by restricting λ , choosing an appropriate value for ρ . More fundamentally there is a problem concerning the value and persistence of $\mu_{3,m,t+1}$. If one would make the so called strong GARCH assumption, see Engle (1982), Bollerslev (1986) and Drost and Nijman (1990), where the error term, $\varepsilon_{m,t+1}$, conditionally follows a symmetric distribution, then $\mu_{3,m,t+1}$ would be equal to zero. One would then return to the GARCH-in-mean specification of Bollerslev, Engle and Wooldridge (1988). However the semi-strong GARCH assumption (C.2 in Appendix 3.C and the second formula in 5.1) in the spirit of Weiss (1984, 1986) and Drost and Nijman

(1990) allows higher order moments to be time varying, and thus $\mu_{3,m,t+1}$ is allowed to wander over time. Tan (1991) asserts that if systematic skewness does not persist, it will not be able to explain very much about realized returns. This means that if 5.1 is to have any real empirical importance, the skewness should be persistent and one consequently needs a time series representation for $\mu_{3,m,t+1}$ generating this persistence (if present).

A cross-sectional test for 4.24 can be obtained by adopting a multivariate GARCH process. Restricting correlations to be constant then gives the evolution of the covariances as functions of the respective conditional variances, and a similar approach can be used for the conditional coskewness (if present and persistent).

These considerations also apply to 4.28, the asset pricing model most reminiscent of the three moment CAPM, where we have the additional problem that the weighting parameter, α_t , generally is non-constant, unless the ratio of the conditional second and third central moment of the market return is constant. Although beyond the scope of this chapter, the issues mentioned above are subject of future research.

3.6 Conclusions

In this chapter it is shown how consumption can be eliminated from an intertemporal asset pricing model. The primary reason for pursuing this elimination stems from the problems one encounters in actually measuring consumption. The elimination is obtained, in a homoskedastic setting, by combining an approximately log linear budget constraint and a log linear Euler equation for consumption. The linearization of this Euler equation, under Kreps-Porteus preferences, results in a linear relationship between expected consumption growth and the expected return on invested wealth. The linearization of the Euler equation for portfolio allocation, which also owes its special form to the assumption of Kreps-Porteus - non-expected utility - preferences results in an asset pricing model containing both static as well as intertemporal CAPM elements. The combination mentioned above gives an explicit solution for the consumption innovation, which can in

covariance form be inserted in the derived asset pricing relation. As a result, equilibrium asset returns can be determined without any reference to their covariance with consumption growth, using instead their covariances with the market return and a hedge portfolio, indexing changes in the stochastic opportunity set. In a heteroskedastic setting matters are complicated, and one has to rely on particular assumptions regarding the evolution of the intercept term in the relation between consumption growth and the market return. More specifically it is assumed that this intercept term depends linearly on the conditional variance of the market return, which is specified as a GARCH process. These assumptions allow consumption to be eliminated again. Furthermore it is shown in this chapter that assuming a GARCH process for the conditional variance of the market return has a surprising consequence for the resulting asset pricing model. Not the familiar GARCH-in-mean equations but an asset pricing model reminiscent of the three moment CAPM is the result. It should be noted that the results, although dependent on a number of assumptions, are derived without any distributional assumption. The empirical implementation of the results derived in this chapter will be the subject of the next chapter.

Appendix 3.A : The Derivation of the Marginal Rate of Substitution

This appendix contains the derivation of the Euler equations presented in the text as equations 3.17 and 3.18. We need to compute the marginal rate of substitution along an optimal consumption path. To be able to do this we need to characterize the value function that is the solution to the functional equation 3.10. We start by recognizing that through the homotheticity of the aggregator function, the isoelasticity of preferences and the separation of the consumption and portfolio problems, the value function is homogeneous of degree $1-\gamma$ in wealth and that it can be written as :

$$V(W_t, S_t) = \Phi(S_t) W_t^{1-\gamma} \quad \text{A.1}$$

where $\Phi(S_t)$ is an unknown function of current information. We guess that consumption is linear in wealth, where the marginal propensity to consume is state dependent :

$$C_t = \mu(S_t) W_t \quad \text{A.2}$$

By using A.1 and A.2 in the maximization of 3.10, it can be seen that $\Phi(S_t)$ and $\mu(S_t)$ are related as follows :

$$(1-\delta)(\mu(S_t))^{-\rho} = \delta \phi_t(1 - \mu(S_t))^{-\rho} \quad \text{A.3}$$

$$\Phi(S_t) = (1-\delta)^{\theta} (\mu(S_t))^{-\rho\theta} \quad \text{A.4}$$

$$\phi_t = (E_t (\Phi(S_{t+1})(R_{m,t+1})^{1-\gamma}))^{\theta^{-1}} \quad A.5$$

$$\theta = (1-\gamma)/(1-\rho) \quad A.6$$

The expressions A.3 and A.4, when combined, yield a functional equation in $\Phi(S_t)$, which has an explicit solution in only a few special cases, like logarithmic utility, lognormal asset returns or i.i.d. uncertainty (see Giovannini and Weil (1989)). The relations A.2, A.3, A.4 and A.5 along with the budget constraint 2.1 imply that along an optimal consumption path :

$$\frac{\phi_t^\theta}{\Phi(S_{t+1})R_{m,t+1}^{1-\gamma}} = \delta^{-\theta} \left(\frac{C_{t+1}}{C_t} \right)^{\rho\theta} R_{m,t+1}^{-\theta} \quad A.7$$

This expression will be used below. From the definition of the marginal rate of substitution:

$$MRS_{t+1} = \frac{U_{1,t+1}}{U_{1,t}} U_{2,t}$$

it is straightforward to show that along an optimal path :

$$MRS_{t+1} = \delta \left(\frac{C_{t+1}}{C_t} \right)^{-\rho} \left(\frac{E_t V_{t+1}}{V_{t+1}} \right)^{\frac{1-\theta}{\theta}} \quad A.8$$

Furthermore, by using the budget constraint 2.1, A.1 and A.2, we can show that :

$$\frac{E_t V_{t+1}}{V_{t+1}} = \frac{\phi_t^\theta}{\Phi(S_{t+1}) R_{m,t+1}^{1-\gamma}} \quad A.9$$

Combining A.7, A.8, and A.9 yields :

$$MRS_{t+1} = \left(\delta \left(\frac{C_{t+1}}{C_t} \right)^{-\rho} \right)^\theta R_{m,t+1}^{\theta-1} \quad A.10$$

This characterization of the marginal rate of substitution can be used to yield the Euler equations presented in the text.

Appendix 3.B : The Linearization of the Euler Equations

In this appendix we will show how the Euler equations can be linearized without assuming lognormality. We start from the Euler equation for consumption (the same procedure can be followed for the portfolio Euler equation).

$$E_t \left(\left\{ \delta \left(\frac{C_{t+1}}{C_t} \right)^{-\rho} R_{m,t+1} \right\}^\theta \right) = 1 \quad B.1$$

Next, define :

$$Z_t \equiv \left\{ \delta \left(\frac{C_{t+1}}{C_t} \right)^{-\rho} R_{m,t+1} \right\}^\theta \quad B.2$$

From Jensen's Inequality, using a second order Taylor expansion we know that :

$$\ln(E_t(Z_t)) \approx E_t(\ln Z_t) + 0.5 (E_t(Z_t))^{-2} Var_t(Z_t) \quad B.3$$

We also know that (Mood, Graybill and Boes (1988) p.181):

$$Var_t(\ln Z_t) \approx (E_t(Z_t))^{-2} Var_t(Z_t) \quad B.4$$

Combining B.3 and B.4 leads to :

$$\ln E_t(Z_t) \approx E_t(\ln Z_t) + 0.5 Var_t(\ln Z_t) \quad B.5$$

Substituting B.2 into B.5 leads to the desired result.

Appendix 3.C : The Conditional Variance of Consumption Growth as a Linear Function of the Conditional Variance of the Market Return

This appendix contains several possible ways to justify assumption 4.14 in the text. We start by defining :

$$\varepsilon_{c,t+1} \equiv \Delta c_{t+1} - E_t \Delta c_{t+1} , \quad E_t (\varepsilon_{c,t+1}) = 0 \quad C.1$$

$$\varepsilon_{m,t+1} \equiv r_{m,t+1} - E_t r_{m,t+1} , \quad E_t (\varepsilon_{m,t+1}) = 0 \quad C.2$$

Substituting C.1 and C.2 into 4.7 leads to :

$$\Delta c_{t+1} - \sigma r_{m,t+1} = \mu_{m,t} + \varepsilon_{c,t+1} - \sigma \varepsilon_{m,t+1} \quad C.3$$

Hall (1988) asserted that actual movements of consumption differ from planned movements by a completely unpredictable random variable that indexes all the information available at time $t+1$ that was not incorporated in the planning process at time t . We can use:

$$\Delta c_{t+1} - E_t \Delta c_{t+1} = \alpha \varepsilon_{m,t+1} + u_{t+1} , \quad E_t (u_{t+1}) = 0 , \quad E_t (u_{t+1}^2) = \sigma_u^2 \quad C.4$$

In C.4, u_{t+1} is a homoskedastic error term, incorporating exogenous information, which is uncorrelated with the unexpected market return. Sundaresan (1989) and Constantinides (1990), employing habit formation models find a constant ratio of the variance of consumption growth to the variance of wealth changes. Because in our situation the growth in wealth is not exactly proportional to the market return, C.4 may still be plausible. It allows for a less than perfect correlation between consumption growth and

the market return, and the conditional variance of consumption growth can be substantially smaller ($\alpha < < 1$) than the conditional variance of the market return (the empirically observed smoothing of consumption). Substituting C.4 into C.3 gives :

$$\Delta c_{t+1} - \sigma r_{m,t+1} = \mu_{m,t} + (\alpha - \sigma) \varepsilon_{m,t+1} + u_{t+1} \quad \text{C.5}$$

$$Var_t(\Delta c_{t+1} - \sigma r_{m,t+1}) = (\alpha - \sigma)^2 E_t \varepsilon_{m,t+1}^2 + \sigma_u^2 \quad \text{C.6}$$

Which, given 4.7, leads to :

$$\mu_{m,t} = \mu_0 + \frac{1}{2} \left(\frac{\theta}{\sigma} \right) (\alpha - \sigma)^2 E_t \varepsilon_{m,t+1}^2, \quad \mu_0 = \sigma \ln \delta + \frac{1}{2} \left(\frac{\theta}{\sigma} \right) \sigma_u^2 \quad \text{C.7}$$

Another justification for assumption 4.14 is given by analyzing the log version of equation A.7 from Appendix 3.A. We know that along an optimal consumption path :

$$\frac{E_t(\Phi(S_{t+1})R_{m,t+1}^{1-\gamma})}{\Phi(S_{t+1})R_{m,t+1}^{1-\gamma}} = \delta^{-\theta} \left(\frac{C_{t+1}}{C_t} \right)^{\rho\theta} R_{m,t+1}^{-\theta} \quad \text{C.8}$$

Taking logarithms in equation C.8 calls for the same approximation as outlined in Appendix 3.B. Subsequently taking conditional expectations for the log version of equation C.8 results in :

$$E_t \Delta c_{t+1} = \sigma \ln \delta + \sigma E_t r_{m,t+1} + \frac{1}{2} \frac{\sigma}{\theta} Var_t(\ln \Phi(S_{t+1}) + (1 - \gamma)r_{m,t+1}) \quad \text{C.9}$$

Comparing C.9 to 4.7 reveals that :

$$\frac{1}{2} \frac{\sigma}{\theta} \text{Var}_t (\ln \Phi(S_{t+1}) + (1-\gamma)r_{m,t+1}) = \frac{1}{2} \frac{\theta}{\sigma} \text{Var}_t (\Delta c_{t+1} - \sigma r_{m,t+1}) \quad \text{C.10}$$

The result C.9 would enable us to eliminate consumption in a way similar to the elimination in the homoskedastic setting, if we would know what $\ln \Phi(S_{t+1})$ is. Unfortunately this cannot be determined in general, unless $\ln \Phi(S_{t+1})$ is a constant⁴ and C.9 reduces to :

$$E_t \Delta c_{t+1} = \sigma \ln \delta + \sigma E_t r_{m,t+1} + \frac{1}{2} \frac{\sigma}{\theta} (1-\gamma)^2 \text{Var}_t(r_{m,t+1}) \quad \text{C.11}$$

It is, however possible to show from the log version of C.8 and C.9 that :

$$\Delta c_{t+1} - E_t \Delta c_{t+1} = r_{m,t+1} - E_t r_{m,t+1} - \frac{\sigma}{\theta} (\ln \Phi(S_{t+1}) - E_t (\ln \Phi(S_{t+1}))) \quad \text{C.12}$$

Which can be rewritten as (using A.4) :

$$\Delta c_{t+1} - E_t \Delta c_{t+1} = r_{m,t+1} - E_t r_{m,t+1} + (\ln \mu(S_{t+1}) - E_t (\ln \mu(S_{t+1}))) \quad \text{C.13}$$

Remember that $\mu(S_{t+1})$ is the marginal propensity to consume, which depends on ϕ_{t+1} (see A.3). We also know that ϕ_{t+1} depends on the expectation at $t+1$ of the market return in period $t+2$. The last term in C.13 thus represents a correction term indicating that the current market return not only has an immediate one-to-one effect on current consumption growth, but also that the current return causes expectations to be revised, and thereby raising or lowering consumption growth. Again we can use C.4 to characterize unexpected consumption growth, leading to C.7, which is identical to assumption 4.14 in the main text.

⁴ As Campbell (1990) assumes joint lognormality, which is one situation in which $\Phi(S_{t+1})$ is a constant, his assumption 4.4 could be fruitfully replaced by assuming C.11, which is the same as assuming 4.14.

Appendix 3.D : Revisions in the Expected Future Values of an ARMA Process

This appendix is intended to show how equation 4.19 in the text is derived. We start by rewriting 4.18 as :

$$\begin{aligned}\varepsilon_{m,t+1}^2 &= \omega + \sum_{i=1}^m \varphi_i \varepsilon_{m,t+1-i}^2 + v_{m,t+1} + \sum_{j=1}^p \theta_j v_{m,t+1-j} \\ \varphi_i &= \alpha_i + \beta_i, \quad \theta_j = -\beta_j, \quad v_{m,t+1} = \varepsilon_{m,t+1}^2 - \sigma_{m,t+1}^2\end{aligned}\quad \text{D.1}$$

D.1 can also be written as an infinite order MA process :

$$\begin{aligned}\varepsilon_{m,t+1}^2 &= \hat{\omega} + \sum_{j=0}^{\infty} \psi_j v_{m,t+1-j} \\ \psi_0 = \theta_0 &= 1, \quad \psi_j = \theta_j + \sum_{i=1}^j \varphi_i \psi_{j-i}\end{aligned}\quad \text{D.2}$$

The infinite series of moving average parameters can be characterized as :

$$\sum_{j=0}^{\infty} \psi_j = \frac{\sum_{j=0}^{\infty} \theta_j}{1 - \sum_{i=1}^{\infty} \varphi_i} = \frac{1 + \sum_{j=1}^p \theta_j}{1 - \sum_{i=1}^m \varphi_i}\quad \text{D.3}$$

See Flavin (1981). It is also true that :

$$\sum_{j=0}^{\infty} \rho^j \psi_j = \frac{I + \sum_{j=1}^p \rho^j \theta_j}{I - \sum_{i=1}^m \rho^i \varphi_i} \quad D.4$$

Furthermore it is straightforward to establish from D.2 that :

$$(E_{t+1} - E_t) \varepsilon_{m,t+1+j}^2 = \psi_j v_{m,t+1} \quad , \quad j \geq 0 \quad D.5$$

Using D.2, D.4, and D.5 we can derive that :

$$(E_{t+1} - E_t) \sum_{j=1}^{\infty} \rho^j \varepsilon_{m,t+1+j}^2 = \sum_{j=1}^{\infty} (E_{t+1} - E_t) \rho^j \varepsilon_{m,t+1+j}^2 = \sum_{j=1}^{\infty} \rho^j \psi_j v_{m,t+1} = \\ \left(\sum_{j=0}^{\infty} \rho^j \psi_j - I \right) v_{m,t+1} = \frac{\sum_{j=1}^p \rho^j \theta_j + \sum_{i=1}^m \rho^i \varphi_i}{I - \sum_{i=1}^m \rho^i \varphi_i} v_{m,t+1} = \lambda (\varepsilon_{m,t+1}^2 - \sigma_{m,t+1}^2) \quad D.6$$

D.6 is the result we wanted to derive. The specification of λ depends on the order of the GARCH process, and the value of ρ . For example in the GARCH(1,1) case:

$$\lambda = \frac{\rho \theta_1 + \rho \varphi_1}{I - \rho \varphi_1} = \frac{\rho \alpha_1}{I - \rho(\alpha_1 + \beta_1)} \quad D.7$$

Note that ρ in this case denotes the linearization parameter of section 3.2 and not the inverse of the elasticity of intertemporal substitution.

Chapter 4

Asymmetry in Stock Returns : A New Approach

4.1. Introduction

The objective of this chapter is to perform a first empirical analysis for the model developed in the previous chapter, taking into account the criticisms raised in the literature in the context of the traditional three-moment CAPM, and to contrast the results with the results for a GARCH-in-mean model upon which it tries to improve.

Theoretically the importance of skewness in an asset pricing context in discrete time is indisputable. Pioneers in this area are Jean (1971, 1973), Arditti (1967, 1971) and Levy (1969), who established the fact that any risk averse investor with any non-quadratic concave utility function will accept a lower expected return from an investment with a higher positive skewness of returns as opposed to an investment with the same variance but lower skewness. This is a logical consequence of the observation that any risk averse investor is reluctant to undertake any investment that is subject to the possibility, however small, of a large loss and only a limited gain. Skewness is a measure of this asymmetry factor, which induces a risk averter to abhor negative skewness and cherish positive skewness. However in continuous time when asset prices follow diffusion processes and portfolios are rebalanced continuously, mean-variance approximation can lead to an exact optimal strategy and the continuous time CAPM provides a framework which resolves the multi-period dimensions as well. Merton (1971) provides exact

solutions for consumption and portfolio rules under the assumptions of log-normality of prices and HARA utility functions. When rebalancing of portfolios is restricted to some time interval, mean-variance approximation may become inadequate and moments of higher order than the variance of portfolio returns become relevant to portfolio choice. Introducing skewness of returns may add the dimension needed to improve the approximation by the mean and the variance, see Kane (1982). Arditti (1967), Arditti and Levy (1975) and Kraus and Litzenberger (1976) claim that, in fact, all the information contained in any distribution is in its first three moments, and the economic importance of the role played by higher moments is very dubious as no a priori behavioristic arguments for investor attitudes towards the fourth and higher moments have been made. Aversion towards variance and preference for positive skewness, however, are general characteristics of all investors having utility functions displaying the desirable behavioristic attributes of decreasing marginal utility of wealth and non-increasing absolute risk aversion.

A second important theoretical contribution in three-moment asset pricing was made by Jean (1973) and Ingersoll (1975). They independently showed that in a three moment world the investor's feasible set is a cone originating from the point $(R_f, 0, 0)$, and the efficient set will be that portion of the cone with maximum expected return (E) for given standard deviation (σ) and skewness (M), minimum σ for given E and M , and maximum M for given E and σ . As a consequence the general separation theorem will not apply in this three dimensional world, without further assumptions, since investors may hold different combinations of risky assets, depending on their relative preference for skewness. This assumed preference could explain why, as opposed to the obvious rationale for diversification, investors hold only a limited number of assets. Simkowitz and Beedles (1978) pointed out that adding a security to an existing portfolio may increase or decrease portfolio skewness, depending on whether the security's skewness is negative or positive. Consequently if positive skewness is a desirable feature of return distributions, then the fact that diversification may destroy skewness is an explanation of investors holding only a few assets. Conine and Tamarkin (1981) established that the optimal number of assets is obtained at the point where the marginal increase in expected utility from a decrease in variance is equal to the marginal decrease in expected utility from the reduction in skewness.

Despite the theoretical importance of skewness, Beedles (1979) concluded that as an empirical matter, the study of skewness is appropriately characterized as being highly controversial. This conclusion still stands. As a first empirical problem, the question how skewness should enter an equilibrium asset pricing model had to be solved. Ingersoll (1975) and Kraus and Litzenberger (K&L) (1976) were the first to solve this problem. By assuming utility functions with positive and decreasing marginal utility for wealth and non-increasing absolute risk aversion (which implies a preference for positive skewness), they solved for the capital market equilibrium conditions which state that an investor's portfolio (p) contains safe and risky assets (with returns R_f and R_i) in proportions such that the expected excess return on each risky asset equals the sum of his marginal rate of substitution between expected return and standard deviation ($-U_2/U_1$) times the asset's marginal contribution to the portfolio's standard deviation ($\beta_{ip}\sigma_p$) plus his marginal rate of substitution between expected return and skewness ($-U_3/U_1$) times the asset's marginal contribution to the portfolio's skewness ($\gamma_{ip}m_p$) :

$$\bar{R}_i - R_f = -(U_2/U_1)\beta_{ip}\sigma_p - (U_3/U_1)\gamma_{ip}m_p \quad 1.1$$

Note that, as indicated above, the optimal portfolios for different investors may differ. In order to obtain a model of market equilibrium K&L (1976) need to assume homogeneous probability beliefs and linear risk tolerance (HARA utility) with the same cautiousness parameter for all investors. In Ingersoll's (1975) terminology this would assure that all investors are on the same ray. These assumptions are necessary and sufficient to establish that the composition of each investor's optimal portfolio of risky assets is the same. The market equilibrium model is denoted by :

$$\bar{R}_i - R_f = b_1\beta_i + b_2\gamma_i , \quad \beta_i = \sigma_{iM}/\sigma_M^2 , \quad \gamma_i = m_{iMM}/m_M^3 \quad 1.2$$

Here b_1 and b_2 are estimable coefficients attached to the market beta of risky asset i and the market gamma or systematic skewness of risky asset i

respectively. This model has gained considerable attention in the empirical literature, in contrast to the consumption oriented three-moment CAPM which K&L (1983) developed a couple of years later. The empirical evidence presented in K&L (1976) is consistent with the predictions ($b_1 > 0$, $b_2 < 0$ if $m_M > 0$, $R_M - R_f = b_1 + b_2$) of their three moment extension of the traditional CAPM. Furthermore zero beta, zero systematic skewness portfolios earn the risk free rate. They conclude that empirical findings inconsistent with the traditional CAPM are the result of misspecification of the CAPM by omission of systematic skewness. These results were challenged by Friend and Westerfield (F&W) (1980) who claimed that K&L's (1976) results were largely a result of the particular time period and estimation procedures they had used. F&W (1980), in their analysis, find that the importance of systematic skewness in explaining asset returns is highly sensitive to the different market indices used and to the sign of the excess market return. Furthermore they find the estimated risk free rate to be significantly higher than the actual risk free rate, a result which is insensitive to their various test constructions.

Several criticisms to both the K&L (1976) as well as the F&W (1980) approach are in order. First, as Barone-Adesi (1985) has pointed out, the differing conclusions of the two studies may be due to the high collinearity of their estimated portfolio parameters (β , γ). Both studies use these estimated parameters as regressors in the estimation of risk premia. The estimation procedures are thus affected by the errors-in-variables problem, and further complicated by multicollinearity. The resulting estimators are biased, but both the size and the direction of their bias is unknown. Using Gibbons' (1982) multivariate approach, which removes the errors-in-variables problem, Barone-Adesi (1985) tests a quadratic market specification, based on an arbitrage argument, which appears to be consistent with the K&L model and predictions. A second criticism was offered by Sears and Wei (1985), who noted that the K&L empirical specification of their model (1.2) does not reveal all of the information in the theoretical model. Their model also implies :

$$\bar{R}_M - R_f = b_1 + b_2 \quad 1.3$$

That is, 1.2 and 1.3 impose restrictions between the market risk premium and the coefficients. Dividing 1.2 by 1.3 results in :

$$\bar{R}_i - R_f = [(\bar{R}_M - R_f)/(1 + K_3)]\beta_i + [K_3(\bar{R}_M - R_f)/(1 + K_3)]\gamma_i, \quad K_3 = b_2/b_1 \quad 1.4$$

This indicates that the market risk premium is implicit in each of the model's coefficients and that empirical tests of the three-moment model may be sensitive to the market risk premium¹. Failure to separate the market risk premium from K_3 , which is indicative of the importance of skewness, may result in incorrect inferences regarding the importance as well as the sign of risk and skewness. The third criticism was provided by Singleton and Wingender (1986) who found that skewness in individual stock returns and portfolio returns does not persist. And if skewness does not persist it should not explain much about the pricing of stocks. To measure the order of skewness persistence Singleton and Wingender (1986) use Spearman's Rank Order Correlation (SROC) for skewness measures in adjacent time periods. The low correlation values they obtain suggest a lack of skewness persistence. The usefulness of SROC in this case however should be doubted, as one can easily construct examples for which SROC indicates no persistence, while in fact there is persistence in skewness². Nevertheless, the notion that skewness, or more appropriately coskewness, should persist in a proper way (to be defined below), to be of any importance is still valid.

¹ F&W (1980) discovered this sensitivity when they separated periods with a positive excess market return from periods with a negative excess market return.

² Consider, for example, ten portfolios with skewness measures ranging from $-\alpha$ to $+\alpha$ and with difference in skewness between two ranked portfolios equal to $\alpha/5$ ($2\alpha/5$ for portfolios 5 and 6). Assume that for oddly ranked portfolios, the skewness measure in the following period is the exact opposite of the previous value, and that for evenly ranked portfolios the skewness stays the same. Rank again and compute SROC, which is 0.015. This suggests no persistence while in half of the cases there is persistence.

Recently, Lim (1989) and Tan (1991) reestimated the K&L model in light of the criticisms mentioned above. Lim (1989) tests the Sears and Wei (1985) version of the K&L model using Hansen's (1982) Generalized Method of Moments (GMM). The GMM is an appropriate methodology to test the K&L model as it avoids the measurement error problem and requires no distributional assumptions. His tests provide some evidence that systematic skewness is priced, and that there is a trade-off between risk and systematic skewness. Tan's (1991) results based on a cross-sectional regression analysis of the Sears and Wei model are less encouraging and are more consistent with the F&W (1980) results. Tan (1991) is particularly troubled by the sign of the estimated coefficient for systematic skewness, which is negative. However, average systematic skewness for all time periods in his study is positive, so a negative sign for the market price of gamma should come as no surprise. Theoretically this sign should always be negative.

In the preceding chapter we have developed an asset pricing model in which conditional skewness³ plays an important role. One of the intermediate results in that development was that the two components of what is generally known as the GARCH-in-mean model may not be compatible (see chapter 3). This was really the original objective in that chapter. The fact that the subsequent analysis led to a model similar to a three moment CAPM could be viewed as pure serendipity. Although the model is similar to the three-moment CAPM it has a number of additional features.

First of all the model is of a conditional nature as the expected conditional variance of the market return and the expected conditional coskewness appear in the pricing relation. Secondly the conditional variance of the market return is allowed to vary over time and is assumed to follow a GARCH process. Thirdly, even if the conditional coskewness or the conditional third moment for the market return would be constants, the model differs from the GARCH-in-mean model - upon which it tries to improve - as the parameter linked to the conditional variance of the market return is not the coefficient of relative risk aversion (as is assumed in the GARCH-in-mean model). Lastly the model is obtained without any distributional as-

³ Although the term "skewness" is usually reserved for the normalized third moment, I will use the terms skewness and third moment interchangeably.

sumptions. As was stated earlier, the objective of this chapter is to provide an empirical analysis for this model, taking into account the criticisms raised to the traditional three moment CAPM where necessary, and contrast the results with the results for a GARCH-in-mean model.

The organization of this chapter is as follows : Section 2 will discuss the data used in this analysis and present the results for a GARCH-in-mean model applied to these data. Section 3 will review the proposed alternative model and its assumptions. A calibration exercise to gauge the model's ability to provide sensible results will be presented in Section 4. Before moving to the actual estimation procedure in Section 6, Section 5 will discuss the formation of conditional skewness expectations. Section 7 contains the empirical results and a comparison to the GARCH-in-mean results. Section 8 concludes.

4.2. Data and GARCH-in-mean Results

The dataset used in this analysis is partially the same as the set used in chapter 2. Stock market indices from five countries will be used. The countries are the United Kingdom (UK), the United States (USA), the Netherlands (NET), the Federal Republic of Germany (GER) and Japan (JAP). The indices used are the DATASTREAM Total Market Indices, which contain all stocks quoted on the respective exchanges. These indices are value weighted and are adjusted for stock dividends, capital modifications and the like. The data were collected on a weekly basis, for a period starting on January 2nd 1973 and ending on September 18th 1990, spanning more than 17 years. The total number of observations equals 923. The quotes used are Tuesdays' closing prices. Continuously compounded returns are used, and are defined as the difference in logarithmic value of two consecutive observations. Summary statistics for these data can be found in Table 1 of Chapter 2. As a risk free rate measure, we have used Treasury Bill rates, which were also obtained from DATASTREAM⁴.

⁴ DATASTREAM is U.K. incorporated data service company. The DATASTREAM mnemonics for the data obtained are : TOTMKUK, TOTMKUS, TOTMKNL, TOTMKJP and TOTMKBD for the indices and UKTRSBL%, USTRSBL%, NLTRSBL%,

Under the assumptions of the intertemporal CAPM, see Merton (1973), Merton (1980) has derived an exact relation between the expected market risk premium and the (conditional) variance of the market return :

$$E_t r_{m,t+1} - r_{f,t+1} = \gamma \sigma_{m,t+1}^2 \quad 2.1$$

In Merton's derivation 2.1 is the solution of a continuous time optimization and γ is the harmonic mean of individual investors' Pratt-Arrow measures of Relative Risk Aversion (CRRA). In discrete time 2.1 is an approximate linear relationship, which can only be obtained by assuming quadratic utility or normally distributed returns. The (G)ARCH-in-mean model, initially developed by Engle, Lilien and Robins (1987), provides an appropriate tool for estimation of this relationship. In fact the list of studies using the GARCH-in-mean has become almost uncountable, see the survey article by Bollerslev, Chou and Kroner (1992).

Prominent examples of applications of this model to stock index returns are French, Schwert and Stambaugh (1987), Chou (1988), and Baillie and DeGennaro (1990), and Chou, Engle and Kane (1992). The GARCH-in-mean model that will be used here is :

$$r_{m,t+1} - r_{f,t+1} = \varphi_0 + \gamma \sigma_{m,t+1}^2 + \varepsilon_{m,t+1} \quad 2.2$$

$$\sigma_{m,t+1}^2 = E_t(\varepsilon_{m,t+1}^2) = \alpha_0 + \alpha_1 \varepsilon_{m,t}^2 + \beta_1 \sigma_{m,t}^2, \quad \varepsilon_{m,t+1} \sim D(0, \sigma_{m,t+1}^2) \quad 2.3$$

Although the model has been extensively used, it is not without criticism. Pagan and Ullah (1988) pointed out that the estimates for the parameters in the conditional mean equation are not asymptotically independent from the estimates of the parameters in the conditional variance equation. Mispecification of the GARCH part of the model therefore leads to biases and

BDTRSBL% and JPTRSBL% for the treasury bill rates. The monthly treasury bill rates were converted into weekly returns.

inconsistencies in the mean part of the model. Furthermore the linear relationship in 2.2 might not be constant, as is evidenced by Chou, Engle and Kane (1992). To aggravate matters, the relationship may not even be significant. Baillie and DeGennaro (1990) obtained insignificant estimates for the CRRA, in their analysis of daily and monthly returns using both normal as well as Student t-distributions. The most severe criticism however pertains to the theoretical and economic underpinnings of the model. These are virtually non-existent, apart from Merton's continuous-time result. It is a statistical model and should be used as such. The model is estimated here, solely in order to contrast the empirical results with the results from a model that is deemed more appropriate. The results are contained in Tables 4.1 and 4.2. We have used a conditional normal as well as a conditional Student t-distribution. The results are entirely consistent with the results obtained by Baillie and DeGennaro (1990). As expected the GARCH parameters are highly significant as is the degrees of freedom parameter for the t-distribution. The constant term in the mean equation is never significant, which is also non-surprising as the excess market return should only be determined by the market risk premium. The point estimates of the coefficient of relative risk aversion, although sometimes negative (which in itself is puzzling), on average take on reasonable values. These estimates are however not significantly different from zero. Based on these results serious doubts arise about the usefulness of a univariate GARCH-in-mean process⁵.

⁵ For the sake of brevity, and because the model serves only illustrative purposes, no residual analysis or misspecification tests have been performed.

Table 4.1 : GARCH(1,1)-M-Normal Estimates , Weekly Returns January 2 1973 - September 18 1990

	UK	USA	NET	GER	JAP
φ_0	-0.000845 (0.002006)	-0.001422 (0.001113)	0.000310 (0.000859)	-0.000341 (0.000785)	0.001307 (0.000681)
α_0	0.000055 (0.000027)	0.000011 (0.000012)	0.000038 (0.000015)	0.000004 (0.000003)	0.000004 (0.000003)
α_1	0.084390 (0.041600)	0.057067 (0.016969)	0.221646 (0.083117)	0.084953 (0.024380)	0.090752 (0.026355)
β_1	0.838921 (0.060784)	0.929076 (0.024881)	0.723762 (0.042579)	0.910957 (0.025973)	0.907940 (0.026875)
γ	0.719727 (2.463213)	1.820271 (2.129728)	0.219108 (1.515409)	0.795010 (1.497292)	-0.368270 (1.776769)
LL	2086.38	2178.70	2308.46	2400.02	2392.30

Heteroskedasticity Consistent Standard Errors are given in parentheses. LL denotes Log-Likelihood value.

Table 4.2 : GARCH(1,1)-M-t Estimates, Weekly Returns January 2 1973 - September 18 1990

	UK	USA	NET	GER	JAP
φ_0	0.000413 (0.000816)	-0.001596 (0.001553)	0.001173 (0.000998)	0.000290 (0.000796)	0.001060 (0.000676)
α_0	0.000054 (0.000017)	0.000028 (0.000011)	0.000044 (0.000018)	0.000007 (0.000004)	0.000005 (0.000004)
α_1	0.117903 (0.032615)	0.056586 (0.019691)	0.170028 (0.042493)	0.085836 (0.022327)	0.085148 (0.029240)
β_1	0.795299 (0.047742)	0.884137 (0.032920)	0.739106 (0.060441)	0.896359 (0.027126)	0.906373 (0.031742)
γ	-0.018015 (0.714621)	3.610432 (3.046097)	-1.918812 (2.355048)	0.703816 (1.898754)	1.203111 (1.780296)
ν	6.247790 (1.658295)	8.212930 (2.920091)	5.359445 (1.114297)	8.237479 (2.174330)	4.917913 (0.722414)
LL	2163.24	2239.95	2358.98	2423.10	2437.11

Heteroskedasticity Consistent Standard Errors are given in parentheses. LL denotes Log-Likelihood value.

4.3. The Model and its Assumptions

The model developed in the previous chapter, is an asset pricing equilibrium model which states that asset returns can be determined by their conditional covariances with the market return, (m), with the news about future market returns, (h), and with the squared unexpected market return (ϵ^2):

$$E_t r_{i,t+1} - r_{f,t+1} = -\frac{1}{2} \sigma_{i,t+1}^2 + \gamma \sigma_{im,t+1} + (\gamma-1) \sigma_{ih,t+1} - \frac{1}{2} (1-\gamma)^2 \lambda \sigma_{i\epsilon^2,t+1} \quad 3.1$$

$$\sigma_{i\epsilon^2,t+1} = Cov_t(r_{i,t+1}, \epsilon_{m,t+1}^2) = E_t((r_{i,t+1} - E_t r_{i,t+1})(r_{m,t+1} - E_t r_{m,t+1})^2) \quad 3.2$$

For the market return the model can be rewritten into :

$$E_t r_{m,t+1} - r_{f,t+1} = \left(\gamma - \frac{1}{2} \right) \sigma_{m,t+1}^2 + (\gamma-1) \sigma_{mh,t+1} - \frac{1}{2} (1-\gamma)^2 \lambda \mu_{3,m,t+1} \quad 3.3$$

$$\mu_{3,m,t+1} = E_t((r_{m,t+1} - E_t r_{m,t+1})^3) \quad 3.4$$

The model described in equations 3.1 to 3.4 is derived without any distributional choices and is based on the following assumptions. First of all, the non-linear intertemporal budget constraint (2.1 in chapter 3) can be linearized by a log-linear approximation and a first order Taylor expansion of the non-linear term around the mean log consumption-wealth ratio. This gives an approximate solution for consumption growth (2.13 in chapter 3), and it is assumed that this approximation is reasonably accurate, see Campbell (1990). Furthermore we assume Generalized Isoelastic Preferences (GIP), introduced by Epstein and Zin (1989) and Weil (1990), which are based on Kreps-Porteus (1978,1979a,b) non-expected utility preferences. A parameterization according to GIP allows one to discriminate between the constant coefficient of relative risk aversion (CRRA) and the elasticity of intertemporal substitution (EIS), which would just be each other's reciprocals in

an expected utility framework. Solving for a representative agent's intertemporal maximization problem results in a set of stochastic Euler equations, that can again be linearized, by a second order Taylor approximation. In this fashion we obtain a linear relationship between expected consumption growth and the expected market return (3.21 in chapter 3), and an asset pricing equation (3.23 in chapter 3) that says that the expected excess return on an asset is determined by its own conditional variance (due to a Jensen's inequality effect), and by its covariances with consumption growth and the market return. Whenever the product of CRRA and EIS equals one, the traditional CCAPM is recaptured, and when CRRA equals one the static CAPM is obtained as a special case.

In order to eliminate consumption from the asset pricing equation, we employ two further assumptions. We assume that the conditional variance of consumption growth and the conditional covariance between consumption growth and the market return are linear functions of the conditional variance of the market return. This assumption is justified in Appendix C of chapter 3. If we would have assumed joint log-normality, then this assumption would hold exactly. Now it is treated as a convenient approximation. Finally we have assumed that the conditional variance of the market return follows a GARCH process (λ is a function of the GARCH parameters). These assumptions and the derivations provided in chapter 3 ultimately result in the model described above. Recently, Restoy (1992) has shown that, under GIP and assuming a VAR-GARCH process for realized and conditionally expected market returns :

$$E_t r_{m,t+1} - r_{f,t+1} = \left(\gamma - \frac{1}{2} \right) \sigma_{m,t+1}^2 + (\gamma - 1) \sigma_{mh,t+1} \quad 3.5$$

However, Restoy arrives at this specification by assuming conditional normality, in which, by assumption, conditional odd moments are zero. By assuming a conditionally symmetric distribution, our specification would reduce to Restoy's specification. This is the main reason why we have not used any distributional assumptions. Another special case would obtain whenever CRRA, γ in our model, would equal one. Then the model would

collapse into a static CAPM model, or more appropriately a GARCH-in-mean model. In its present form (3.1 to 3.4) our version of the three-moment CAPM is rather different from the traditional specification. Not only is it a conditional model, rather than an unconditional model, but it is also derived through an intertemporal optimization procedure, instead of a one-period optimization. This intertemporal character is most vividly conveyed by the covariance between assets returns and the news about future market returns, appearing in the pricing equations. In order to achieve a closer resemblance between the two specifications we assume that the covariance between asset returns and the news about future market returns is a linear function of the covariance between asset returns and the current market return⁶.

$$\sigma_{ih,t+1} = \varphi \sigma_{im,t+1} \quad 3.6$$

Dividing 3.1 by 3.3 using 3.6, and ignoring Jensen inequality effects, results in :

$$\frac{E_t r_{i,t+1} - r_{f,t+1}}{E_t r_{m,t+1} - r_{f,t+1}} = \alpha_{t+1} \frac{\sigma_{im,t+1}}{\sigma_{m,t+1}^2} + (1 - \alpha_{t+1}) \frac{\sigma_{ie^2,t+1}}{\mu_{3,m,t+1}} \quad 3.7$$

Where α_{t+1} is given by 4.30 in chapter 3. This result bears a close resemblance to the Sears and Wei (1985) version of the three moment CAPM. When α_{t+1} equals one (i.e. when CRRA is one) the traditional CAPM is again the result. The weights assigned to systematic covariance and systematic coskewness are determined by the variance/skewness ratio, see 4.30 in chapter 3. For our empirical analysis we will only use the pricing equation for the expected market return for the following reasons. First of all this is a first empirical test of the model, and we want to keep matters as simple as possible. We will use a time-series rather than a cross-sectional approach. Secondly we want to be able to compare the results with the results for the GARCH-in-mean model of section 2.

⁶ Specifying the market return as an ARMA process, and using the ARMA-trick of chapter 3 will give the desired result.

4.4. Calibration Results

In this section we try to assess our model's capability of producing sensible results. From the GARCH-in-mean results we can infer implied values for relevant parameters in the three moment market model. The procedure that is followed moves along the following lines. First consider the GARCH-in-mean process again :

$$r_{m,t+1} - r_{f,t+1} = \varphi_0 + \gamma \sigma_{m,t+1}^2 + \varepsilon_{m,t+1} \quad 4.1$$

$$\sigma_{m,t+1}^2 = E_t(\varepsilon_{m,t+1}^2) = \alpha_0 + \alpha_1 \varepsilon_{m,t}^2 + \beta_1 \sigma_{m,t}^2, \quad \varepsilon_{m,t+1} \sim D(0, \sigma_{m,t+1}^2) \quad 4.2$$

and compare it to the three-moment market model, the alternative in this chapter :

$$r_{m,t+1} - r_{f,t+1} = ((\gamma - 1/2) + \varphi_1(\gamma - 1)) \sigma_{m,t+1}^2 - \lambda/2(1 - \gamma)^2 \mu_{3,m,t+1} + \varepsilon_{m,t+1} \quad 4.3$$

$$\sigma_{m,t+1}^2 = E_t(\varepsilon_{m,t+1}^2) = \alpha_0 + \alpha_1 \varepsilon_{m,t}^2 + \beta_1 \sigma_{m,t}^2 \quad 4.4$$

$$\mu_{3,m,t+1} = E_t((r_{m,t+1} - E_t r_{m,t+1})^3) = E_t(\varepsilon_{m,t+1}^3) \quad 4.5$$

If the conditional skewness, or better the expected conditional third moment would be a constant then the two models would not be that different. The constant term that now appears in the three moment market model, because skewness is assumed constant, can be equated to the constant term in the GARCH-in-mean model⁷ :

⁷ If the conditional third moment would be zero, than the constant term in the GARCH-in-mean model should be zero. It is however argued in this chapter that the conditional third moment is not equal to zero.

$$\hat{\varphi}_0 = -\frac{\lambda}{2}(1 - \gamma_{3M})^2 \bar{\mu}_3 \quad 4.6$$

We are now able to gauge the three-moment market model's ability to provide sensible results for γ and φ_1 . We use the GARCH parameter estimates to determine λ (the definition of λ is given in Appendix 3.D), save the residuals from the GARCH-in-mean model, and calculate the third moment for these residuals. From the estimated GARCH parameters λ can be computed :

$$\lambda = \frac{\rho \hat{\alpha}_I}{1 - \rho(\hat{\alpha}_I + \hat{\beta}_I)} \quad 4.7$$

where ρ is a discount factor⁸, representing an annual discount rate of five percent. Now γ_{3M} can be computed from 4.6. Finally, φ_1 can be determined through :

$$\hat{\gamma}_{GIM} = (\gamma_{3M} - 1/2) + \varphi_1(\gamma_{3M} - I) \quad 4.8$$

This procedure allows us to obtain "calibrated" values for relevant parameters in the three moment market model from the parameter estimates of the GARCH-in-mean model. The results are contained in Tables 4.3 and 4.4. All in all the results indicate that skewness is present in GARCH residuals and that the average value of the third moment is large enough to generate plausible values for the coefficient of relative risk aversion. Furthermore the negative values of φ_1 suggest a negative correlation between current and future market returns, see equation 3.6.

⁸ ρ can be interpreted as the average annual ratio of invested wealth to total wealth. $1-\rho$ can similarly be interpreted as the average annual consumption wealth ratio, see Chapter 3. As we use weekly data we will use the 52nd root of ρ , where ρ is set equal to 0.95, see Campbell (1990).

Table 4.3 : Parameter values for three moment CAPM implied from GARCH-in-mean Normal model

	UK	USA	NET	JAP	GER
$\hat{\mu}_3$	-4.5×10^{-6}	-2.3×10^{-5}	-1.5×10^{-5}	-1.2×10^{-5}	-1.2×10^{-5}
skewness	-1.71	-1.68	-1.04	-0.60	-0.70
λ	1.09	3.84	3.99	39.54	16.73
γ	14.14	5.04	3.28	2.63	2.30
φ_1	-0.98	-0.67	-1.12	-1.53	-0.77

Table 4.4 : Parameter values for three moment CAPM implied from GARCH-in-mean Student-t model

	UK	USA	NET	JAP	GER
$\hat{\mu}_3$	-3.7×10^{-6}	-2.3×10^{-5}	-1.4×10^{-5}	-1.3×10^{-5}	-1.2×10^{-5}
skewness	-2.39	-2.33	-1.17	-0.64	-0.77
λ	1.34	0.94	1.85	9.00	4.56
γ	10.10	9.62	7.80	4.02	3.30
φ_1	-1.06	-0.64	-1.36	-0.77	-0.91

4.5. Conditional Third Moment Specification

In the introduction to this chapter we have already mentioned that skewness, or in this case the expected conditional third moment, ought to be time-varying and persistent in order to be able to explain anything about asset returns. As far as we know, no attempt has yet been made to model time-varying third moments. Any time-series representation trying to capture possible persistence, unavoidably will be of an ad hoc nature. Furthermore the empirical validity of such a representation, in our view, will depend on the time period considered. If one were to consider the crash period of 1987, for instance, then it would probably not be difficult to find skewness persistence, no matter what the chosen model looked like⁹. The basic idea of time-series analysis really is that past observations on a particular variable can provide useful information in predicting its future value. In the tentative specification we offer here we will assume that the current expected conditional third moment depends on the previous expected conditional third moment and the previous realized cubic residual :

$$\mu_{3,m,t+1} = E_t(\varepsilon_{m,t+1}^3) = \delta_0 + \delta_1 \varepsilon_{m,t}^3 + \delta_2 \mu_{3,m,t} \quad 5.1$$

This specification is similar to a GARCH(1,1) model for the conditional variance. We will dub this model the Generalized AutoRegressive Conditional Skewness or GARCS model. Contrary to the GARCH(1,1) model, where the conditional variance has to be positive, we do not need to impose restrictions on the signs of the parameters, as the conditional third moment can be both positive and negative. While the strengths and weaknesses of GARCH-like models have been carefully documented, we really do not

⁹ The crash period is included in our dataset as well. To make sure that our results are not solely driven by this period, we split the total period into four equal subperiods. In the fourth period all skewness values are obviously significantly negative. In the first three periods the skewness for the United Kingdom changes from significantly positive to negative. For Japan skewness is always significantly negative. For the other indices the skewness values do change signs over the various subperiods but are not always significant.

know much about the evolution of skewness, let alone theoretical foundations and regularity conditions for time-series models for conditional skewness. As a stationarity condition we will assume that $\delta_1 + \delta_2$ is less than one. Therefore a specification like 5.1 should be considered as one possible way to model conditional skewness. In order to get an impression of the usefulness of this specification we regressed the cubic residuals from the GARCH-in-mean model, on their one period lags and on their lagged conditional expectation, which is given in 5.1. The results are given in Table 4.5.

Table 4.5 : Parameter estimates for GARCS process :

$$\varepsilon_{m,t+1}^3 = \delta_0 + \delta_1 \varepsilon_{m,t}^3 + \delta_2 \mu_{3,m,t} + \zeta_{m,t+1}, \quad \zeta_{m,t+1} \sim N(0, \sigma_\zeta^2)$$

	UK	USA	NET	JAP	GER
δ_0	-1.0×10^{-6} (5.0×10^{-6})	-1.2×10^{-5} (6.3×10^{-5})	-4.0×10^{-6} (9.0×10^{-6})	-2.0×10^{-6} (7.0×10^{-6})	-2.0×10^{-6} (2.0×10^{-6})
δ_1	0.039 (0.031)	-0.003 (0.024)	0.012 (0.024)	0.005 (0.022)	0.053 (0.020)
δ_2	0.762 (0.288)	0.464 (2.684)	0.707 (0.505)	0.800 (0.547)	0.795 (0.073)
μ_3	-5.0×10^{-6}	-2.2×10^{-5}	-1.4×10^{-5}	-1.0×10^{-5}	-1.3×10^{-5}

Standard errors are given in parentheses. $\mu_{3,m,t+1}$ is given by 5.1.

The results are truly mixed. Only in two cases is δ_2 significantly different from zero at the 5% level; for Germany and the United Kingdom. Only for Germany, δ_1 is significant. The estimate for the unconditional third moment, which is computed by :

$$\mu_3 = \frac{\delta_0}{I - \delta_1 - \delta_2} \quad 5.2$$

is, however rather close to the unconditional third moment for the GARCH-in-mean residuals. Finally the assumption that the error term in this regression is normally distributed is questionable. Although the results above are not convincing, the specification for the expected conditional third moment given in 5.1 will be used in the estimation for the full model, as no plausible alternative is available. Distributional considerations will not pose a problem as we will use Hansen's (1982) Generalized Method of Moments.

4.6. Estimation

Simultaneous estimation of the three equations is performed using Hansen's (1982) generalized method of moments (GMM) procedure as follows. Define the q -dimensional parameter vector of interest by β , its true value by β_0 and the system's p -dimensional innovation vector by :

$$\iota_{t+1}(\beta_0) \equiv \{ \iota_{1,t+1}(\beta_0), \iota_{2,t+1}(\beta_0), \dots, \iota_{p,t+1}(\beta_0) \}' \quad 6.1$$

Let $z_t(\beta_0)$ be an m -dimensional vector of time t information, uncorrelated with $\iota_{t+1}(\beta_0)$, to serve as instrumental variables. The vector of innovations can be thought of as a vector of prediction errors, resulting in the following family of orthogonality conditions, which are used in the estimation procedure.

$$E(\iota_{t+1}(\beta_0) \otimes z_t(\beta_0)) = \underline{0} \quad 6.2$$

The vector of zeros has dimension mp . The GMM estimator of β_0 , b_T , minimizes the following quadratic criterion function :

$$\varphi(\beta) = \left[\frac{1}{T} \sum_{t=0}^{T-1} (\iota_{t+1}(\beta) \otimes z_t(\beta)) \right]' (S_T)^{-1} \left[\frac{1}{T} \sum_{t=0}^{T-1} (\iota_{t+1}(\beta) \otimes z_t(\beta)) \right] \quad 6.3$$

$$S_T = \frac{1}{T} \sum_{t=0}^{T-1} (\iota_{t+1}(b) \iota_{t+1}(b)' \otimes z_t(b) z_t(b)') \quad 6.4$$

Where b is a consistent estimator of β_0 . Hansen (1982) has shown that :

$$\sqrt{T}(b_T - \beta_0) \xrightarrow{d} N(0, Q) \quad , \quad Q = (D' S^{-1} D)^{-1} \quad 6.5$$

$$D = E [\partial(\iota_{t+1}(\beta_0) \otimes z_t(\beta_0)) / \partial \beta'] \quad 6.6$$

$$S = E [\iota_{t+1}(\beta_0) \iota_{t+1}(\beta_0)' \otimes z_t(\beta_0) z_t(\beta_0)'] \quad 6.7$$

The choice of S_T as a weighting matrix yields a heteroskedasticity-consistent estimate of the covariance matrix of b_T . The matrix D can be consistently estimated by:

$$D_T = \frac{1}{T} \sum_{t=0}^{T-1} [\partial(\iota_{t+1}(b_T) \otimes z_t(b_T)) / \partial \beta'] \quad 6.8$$

Because of the recursive structure in the innovations vector, analytical derivatives are very difficult to obtain, and numerical derivatives are used both in the minimization of the criterion function as well as in the calculation of the covariance matrix. In estimation initial values for the innovations are set to their expected values of zero. The actual estimation process requires a two-step procedure. In the first step an initial guess for β is used to construct an initial weighting matrix S_T , and the criterion function is minimized. The resulting first-step estimator, although consistent, does not have the proposed asymptotic distribution. Final estimates can be obtained by performing the minimization once again and using the first-step estimates in the construction of the weighting matrix. The advantage of the GMM procedure is that it generates a robust estimator, without any distributional assumptions. One

disadvantage is that it is not, in general, asymptotically efficient, see Mark (1988). The chosen specification incorporates a joint hypothesis that includes the appropriateness of the model, the information structure, rational expectations, the specifications for the conditional expectations, and the appropriateness of the data used, see again Mark (1988). Hansen's (1982) specification test is used to test this joint hypothesis. The first order conditions of the estimation procedure set q (the number of parameters) linear combinations of the mp orthogonality conditions to zero. This means that the model is exactly identified when $q=mp$. When $mp>q$, then there are $mp-q$ linearly independent orthogonality conditions that should be close to zero if the model is correctly specified. Hansen has shown that $T\varphi(b_T)$ is asymptotically chi-square distributed with $mp-q$ degrees of freedom under the null hypothesis. This chi-square test is used to test the model's overidentifying restrictions.

4.7 Empirical Results

The following system of equations, corresponding to 4.3, 4.4, and 5.1, is considered for estimation. In 7.1 we have included a constant term as the conditional third moment is now time-varying.

$$r_{m,t+1} - r_{f,t+1} = \varphi_0 + ((\gamma - 1/2) + \varphi_1(\gamma - 1)) \sigma_{m,t+1}^2 - \lambda/2(1-\gamma)^2 \mu_{3,m,t+1} + \varepsilon_{m,t+1} \quad 7.1$$

$$\sigma_{m,t+1}^2 = E_t(\varepsilon_{m,t+1}^2) = \alpha_0 + \alpha_1 \varepsilon_{m,t}^2 + \beta_1 \sigma_{m,t}^2 \quad 7.2$$

$$\mu_{3,m,t+1} = E_t(\varepsilon_{m,t+1}^3) = \delta_0 + \delta_1 \varepsilon_{m,t}^3 + \delta_2 \mu_{3,m,t} \quad 7.3$$

The innovations vector contains three elements :

$$\iota_{t+1}(\beta_0) = \{\varepsilon_{m,t+1}(\beta_0), \eta_{m,t+1}(\beta_0), \zeta_{m,t+1}(\beta_0)\}' \quad 7.4$$

$$\eta_{m,t+1} = \varepsilon_{m,t+1}^2 - \sigma_{m,t+1}^2, \quad \zeta_{m,t+1} = \varepsilon_{m,t+1}^3 - \mu_{3,m,t+1} \quad 7.5$$

The specification for the conditionally expected third moment, $\mu_{3,m,t+1}$, reflects, as mentioned before, that past observations on a variable provide useful information in predicting its future value. In determining an appropriate set of instruments the econometrician is granted considerable latitude. The only requirement that the instruments need to fulfil is that they be "predetermined" as of time $t+1$, they need not be "econometrically exogenous", see Hansen and Singleton (1982). This means that the set of legitimate instruments is quite large. There are however three reasons to keep the chosen set small. First of all, any expansion of the set beyond the set of absolutely necessary instruments imposes a greater computational burden. Secondly it is conceivable that with a long enough search over a large set of instruments, the econometric specification can ultimately be rejected. This however is not a desirable procedure as it provides little evidence on the usefulness of the model. Thirdly, Tauchen (1986) provides evidence, based on simulation results, that expansion of the instrument set introduces serious bias to the GMM estimator in small samples. On the other hand the number of instruments to be considered is bounded from below. Remember that we need to estimate nine parameters, and that the innovation vector contains three elements. This means that we need at least four instruments to be able to use the proposed specification test. In principle any variable that is in the econometrician's information set at time $t+1$ could be allowed to enter the set. Moreover this is the only real guidance one has in choosing the instruments. The set of instruments that is used contains a constant, the expected conditional variance, the expected conditional third moment, and last period's squared innovation to the excess market return.

$$\zeta_t(\beta_0) = \{I, \sigma_{m,t+1}^2(\beta_0), \mu_{3,m,t+1}(\beta_0), \varepsilon_t^2(\beta_0)\}' \quad 7.6$$

With this set we obtain 12 orthogonality conditions and three overidentifying restrictions. The estimation results are contained in Table 4.6.

Table 4.6 : GMM estimates for 3m-CAPM Model

	UK	USA	NET	JAP	GER
$\phi_0 * 10^{-4}$	-7.18 (19.48)	-8.88 (8.12)	-33.99 (34.07)	-16.78 (15.32)	-0.73 (15.27)
ϕ_1	-1.236 (2.868)	-0.404 (0.055)	1.592 (0.260)	1.233 (0.143)	0.035 (0.099)
γ	2.044 (0.287)	1.621 (0.257)	3.540 (0.258)	3.806 (0.141)	1.808 (0.400)
$\alpha_0 * 10^{-4}$	0.35 (0.46)	1.53 (6.09)	3.87 (1.31)	2.12 (2.67)	2.89 (3.92)
α_1	0.048 (0.084)	0.035 (0.112)	0.065 (0.133)	0.071 (0.141)	0.220 (0.469)
β_1	0.896 (0.138)	0.676 (1.137)	0.099 (0.041)	0.385 (0.458)	0.042 (0.534)
$\delta_0 * 10^{-6}$	-9.00 (47.0)	-16.0 (4.2)	-12.0 (20.0)	-2.00 (6.00)	-13.0 (21.0)
δ_1	0.365 (0.441)	0.177 (0.447)	-0.016 (0.022)	0.134 (0.097)	0.128 (0.291)
δ_2	-0.405 (1.569)	0.041 (0.142)	0.228 (0.030)	0.532 (0.100)	-0.533 (0.151)
$T\phi(b_T)$	2.05 [0.56]	0.80 [0.85]	0.45 [0.93]	1.55 [0.67]	0.79 [0.85]

Standard errors are given in parentheses. P-values for $T\phi(b_T) \sim \chi^2(3)$ are given in brackets.

The test of the overidentifying restrictions, given by the chi-square statistic with three degrees of freedom does not reject the model in any case. A more important test involves the hypothesis that the coefficient of relative risk aversion equals one. In that case the whole model collapses into the GARCH-in-mean model. This hypothesis can be rejected at a five percent significance level, for all estimates of γ in Table 4.6. Comparing the estimation results for γ in Table 4.6 to the GARCH-in-mean results, we can observe that in the former case γ is estimated much more precisely.

Another striking feature of the GMM results is that the GARCH parameters are often insignificantly different from zero, which is in contrast with the GARCH-in-mean results. Furthermore the persistence in the conditional variance (if measured by $\alpha_1 + \beta_1$) appears to decline in the GMM results. This declining persistence in the conditional variance especially occurs when one or more of the parameters in the expected conditional third moment is significant, and the GARCH parameters themselves are not significant. It appears that in some cases we are confronted with an either/or situation. Either one or more of the GARCH parameters is significant or one or more of the GARCS parameters is significant. This is the case for the United Kingdom, Japan, and Germany. For the United States neither the GARCS nor the GARCH parameters are significant. For the Netherlands one of the GARCH parameters and one of the GARCS parameters is significant. These results suggest that there not only exists an economic but also a statistical tradeoff between skewness and variance. In Table 4.7 we have tabulated the correlations between the conditional variance and the conditional third moment for all the indices :

Table 4.7 : Correlations between $\mu_{3,m,t+1}$ and $\sigma^2_{m,t+1}$.

	UK	USA	NET	JAP	GER
correlation	0.01	-0.67	0.89	-0.86	-0.66

Irrespective of the sign, we can observe that there is a large correlation between the conditional variance and the conditional third moment, except for the United Kingdom. We can furthermore observe that when the sign of δ_1 is positive, the correlation between the conditional third moment and the conditional variance becomes negative. A negative error then has a positive effect on the conditional variance and a negative effect on the conditional third moment. When negative errors occur more frequently than positive errors, this can explain the negative correlation. In the same fashion a negative value for δ_1 induces a positive correlation. In a way the conditional variance specification in our model is concerned with the magnitude effect of shocks to asset returns, and the conditional third moment specification is concerned with the sign effect. Which is more important has to be determined empirically. The importance of separating magnitude effects from sign

effects was first noted by Nelson (1991) in his EGARCH model. Contrary to the EGARCH model we have not included a sign or asymmetry factor in the conditional variance specification, but we have modeled this asymmetry factor explicitly and included it in the pricing equation. Finally we will examine the values and significance of the combined parameters appearing in the pricing equation in Table 4.8.

Table 4.8 : Combined parameters of 3m-CAPM model

	UK	USA	NET	JAP	GER
λ	0.86 (1.15)	0.12 (0.62)	0.08 (0.16)	0.13 (0.39)	0.30 (1.03)
$-1/2\lambda(1-\gamma)^2$	-0.47 (0.37)	-0.03 (0.09)	-0.25 (0.49)	-0.52 (1.46)	-0.10 (0.24)
$(\gamma-1/2)+\varphi_1(\gamma-1)$	0.25 (3.06)	0.87 (0.17)	7.08 (1.31)	6.77 (0.72)	1.34 (0.49)

standard errors are given in parentheses.

The rather low values and insignificance of λ can be attributed to the reduced persistence in the conditional variance and the frequent insignificance of the GARCH parameters. As a consequence the coefficient for the conditional third moment is always insignificant. This is the most severe drawback of our model. The economic importance of the conditional third moment is dependent on the persistence in the conditional variance. The total coefficient for the conditional variance is significant in four out of five cases. This significance largely depends on whether φ_1 , which is indicative for the correlation between the current and future market returns, is significant. This intertemporal component is absent in the GARCH-in-mean model, and this absence may cause the insignificance of the γ estimates in the GARCH-in-mean model.

4.8. Conclusions

The objective of this chapter was to perform a first empirical analysis of the asset pricing model that was developed in chapter 3, which is similar to a three-moment CAPM and which tries to improve upon a GARCH-in-mean model. In order to be able to compare the results of this analysis to the estimation results under a GARCH-in-mean specification, we restricted our attention to the market version of our model. The major results of this analysis will now be summarized and suggestions for further research will be given. First of all a calibration exercise has shown that the proposed model can generate plausible values for the parameters of interest. An area in which the current model can be improved is the time-series representation of the expected conditional third moment. The specification that was chosen here is not totally without merit, given the estimation results, but should be considered only as a first attempt to model time-varying conditional third moments. As opposed to the estimation of GARCH-in-mean model, the estimation of the alternative model results in significant estimates for the coefficient of relative risk aversion. This may be due to the fact that the GARCH-in-mean specification does not contain an intertemporal component. Two important drawbacks of the alternative model are that the economic importance of the conditional third moment is dependent on the persistence in the conditional variance and that conditional variance and conditional third moment are highly correlated. The first problem can probably be solved by free estimation of λ , (if we can solve the identification problem), or restricting λ in another way. The second problem is rather common in the existing literature on the three moment CAPM, but a solution is not immediately obvious. Finally the true usefulness the model we have proposed here should be examined for individual assets. This analysis has provided a basis for future research into the relationship and trade-off between variance (risk) and skewness.

Chapter 5

Stochastic Trends and Jumps in EMS Exchange Rates

5.1 Introduction

Since its inception in 1979, the European Monetary System (EMS) has provided an interesting example of a formal Exchange Rate Mechanism (ERM) and a framework for international policy coordination. Intra-European exchange rates are allowed to fluctuate only within official bilateral limits, and this commits national monetary authorities to maintain exchange rates within these limits through exchange market interventions. The ERM is a semi-fixed exchange rate system or alternatively an exchange rate target zone with narrow bands. The description of exchange rate behavior within a target zone has been formalized over the past few years, see Bertola and Svensson (1990), Dumas and Svensson (1991), Flood and Garber (1989), Krugman (1991), and Krugman and Rotemberg (1990), for instance.

The general starting point in this literature is that the presence of a band should have an influence on the movement of an exchange rate, which is induced by forward looking rational expectations. By adopting a continuous process to drive the fundamentals, assuming normal innovations, and assuming perfectly credible target zones, Krugman (1991) shows that exchange rates within a semi-fixed system follow a regulated Brownian

motion process¹. A drawback of the Krugman non-linear target zone model is that it abstracts from realignments. However eleven realignments of central parities occurred during the life of the Exchange Rate Mechanism of the EMS up through 1987. The occurrence of realignments, until very recently, seemed to be limited to the pre-1987 period. One could argue that since 1987 the ERM has behaved like a credible target zone, see Frankel and Phillips (1991). Recent events however in which the United Kingdom and Italy were forced to suspend their membership of the ERM, invalidate this argument. Dumas and Svensson (1991) and Krugman and Rotemberg (1990) have extended the Krugman (1991) model to allow for speculative attacks on the band once official reserves are sufficiently low. Bertola and Caballero (1990) allow for repeated realignments of the central parity, and interventions within the band are allowed for by Lewis (1990). Empirical evidence supporting the target zone literature is, at least to our knowledge, still limited. Evidence against the validity of target zone models is provided by Meese and Rose (1990) and Flood, Rose and Mathieson (1991) who found no support for the postulated S-shape of exchange rate behavior.

Despite the lack of convincing empirical evidence, the target zone literature provides us with a theoretical framework that can explain several stylized facts that have emerged from the empirical exchange rate literature. In the various target zone models, the instantaneous standard deviation of exchange rate returns varies with the position of the fundamental in the band. The standard deviation is largest when the fundamental is at its central parity, because then the exchange rate has a larger potential to move in either direction, than it would have near a band. Consequently a target zone generates conditional heteroskedasticity, and a non-normal unconditional exchange rate return distribution endogenously. If the bandwidth would go to infinity (the free-float case), the conditional heteroskedasticity would disappear. The stylized facts of conditional heteroskedasticity and non-normality therefore should be present in EMS exchange rate returns.

¹ For tests of target zone credibility see e.g. Bertola and Caballero (1990), Giovannini (1990), and Weber (1991). Frankel and Phillips (1991) have independently applied the Bertola-Svensson (1990) and Rose-Svensson (1991) methodology to evaluate the credibility of EMS exchange rates.

Most academic studies on the EMS concentrate on macroeconomic relationships. [See, e.g., Giavazzi, Micossi and Miller (1988), Masera and Triffin (1984) and Rogoff (1984)]. Typically, these studies focus on monetary convergence, capital controls and economic growth. Interestingly, this type of literature pays little attention to observed statistical distributions of EMS exchange rates. Exchange rates involving the US Dollar have been subjected to much closer scrutiny. As is well documented, empirical distributions on US Dollar exchange rate returns exhibit leptokurtic behavior and clusters of high and low volatility. The ARCH class of models, introduced by Engle (1982) and generalized (GARCH) by Bollerslev (1986), are successful in accounting for most of the conditional heteroskedasticity and the inherent clustering phenomenon; this is because ARCH processes lead to "fat-tailed" unconditional densities, even though their conditional densities are normal - see the survey on ARCH modelling in finance by Bollerslev, Chou, and Kroner (1992). Studies by Baillie and Bollerslev (1989), Boothe and Glassman (1987), Hsieh (1989), Jorion (1988), Meese and Rogoff (1983) and Wolff (1987b) provide extensive statistical evidence on US Dollar exchange rates. Overall, the findings overwhelmingly favor the conclusion that the assumption of conditional normality does not capture all the excess kurtosis (i.e. in excess of a normal value of 3) observed in high frequency exchange rate returns. Several alternative conditional distributions have consequently been employed in the literature, for instance the Student-t, normal-Poisson, generalized error, and normal-lognormal distributions, see, e.g., Akgiray and Booth (1988a), Baillie and Bollerslev (1989), Jorion (1988), and Hsieh (1989) and also Chapter 2.

In the current chapter we aim to provide extensive evidence for EMS currencies. We will step aside from the target zone literature and will follow another route to document conditional heteroskedasticity, non-normality (in a similar fashion as we did for stock indices returns in Chapter 2) and the effects of jumps (due to intramarginal interventions and realignments) in EMS exchange rate returns. An important question is whether EMS exchange rates follow random walks or not. It is conceivable that the exchange rate coordination agreement introduces a certain degree of mean reversion

into the actual statistical distributions². We attempt to answer this question and model the stochastic processes followed by EMS exchange rates, in order to arrive at satisfactory descriptions of the time paths followed by these currencies. The observed leptokurtosis may be explained by several classes of models. Therefore, alternative time-series processes characterizing EMS exchange rates are considered. We combine normal distributions with time-varying parameters (i.e we allow for GARCH-type conditional heteroskedasticity) and a stochastic jump process to account for the observed leptokurtosis, which may be due to the existence of discontinuities, or jumps, in EMS data. In addition to exploring normal conditional densities, the Student-t distribution, which is conditionally fat tailed, is employed. The results will form a body of evidence which can serve as a frame of reference for further research.

This chapter is organized as follows. In Section 5.2 we address the question whether EMS exchange rates exhibit some degree of mean reversion, after describing our dataset and providing summary statistics. The models employed to describe the patterns followed by EMS exchange rates are explained in Section 5.3. Section 5.4 presents the main empirical results of our analysis and Section 5.5 contains our concluding comments.

5.2 Random Walks or Mean Reversion ?

The provisions of the EMS provide for participating countries to maintain their exchange rates within bilateral limits of ± 2.25 percent [± 6 percent for Italy (until January 1990), Spain, and the United Kingdom, who joined in July 1989 and October 1990, respectively]³. From time to time realignments

² Rose and Svensson (1991) provide strong evidence of statistically significant mean reversion of EMS exchange rates within the band.

³ In January 1990 the bands for the Italian Lire were narrowed from $\pm 6\%$ to the norm of $\pm 2.25\%$. The transition was accomplished by lowering the upperband limit and leaving the lower limit unchanged. In September 1992, which is not included in our sample period, the United Kingdom and Italy were forced to suspend their membership of the ERM, and realignments for the Spanish Peseta and Portuguese Escudo were effected.

are undertaken. No fewer than eleven realignments occurred during the first eight years and they always have taken the form of devaluations, in varying degrees, against the Deutschmark. With regard to exchange rates involving the US Dollar, many authors conclude that these rates can be described fairly accurately by random walks, see Meese and Rogoff (1983) for an early example. The EMS exchange rate mechanism could very well lead to mean reversion in bilateral EMS exchange rates if the system succeeds in maintaining longer-run target rates or central rates. In the period after January 1987, when the last realignment occurred, the maintenance of central rates has actually been accomplished, until September 1992, see footnote 3.

In this section we will present tests of mean reversion. Our database contains weekly bilateral exchange rates (Thursday closing rates) for currencies of countries that became EMS members in March 1979. For a number of years, prior to the formal adherence to the EMS, the Spanish Peseta was informally kept within a ± 6 percent band with respect to EMS currencies. This consideration leads us to the inclusion of the Spanish Peseta. For reasons of comparison, the British Pound is also included. The data were obtained from Datastream. Even though daily rates are available, we choose to employ weekly data in order to avoid issues surrounding the day-of-the-week effect with regard to exchange rate volatility (on which, see Hsieh, 1988). Our sample includes 677 weekly observations, ranging from 15 March 1979 through 27 February 1992. In Table 5.1 we present summary statistics on weekly log price changes. For the period analyzed (March 15th, 1979 through February 27th, 1992) the mean weekly log price changes are all positive, indicating that the EMS currencies depreciated against the German Mark. Furthermore, the distribution of exchange rate returns is highly skewed to the right, which may be a result of asymmetric movements in the parity adjustments. In order to assess the distributional properties of EMS exchange rate returns, the Bera-Jarque (1982) Normality test and the Kiefer-Salmon (1983) Lagrange multiplier normality test are reported in Table 5.1, where the former is a joint test using both skewness and kurtosis and the latter is an LM test for normal skewness (KS-1) and normal kurtosis (KS-2), respectively⁴. Overall, the evidence presented suggests a consistent rejection

⁴ The Bera-Jarque test is asymptotically Chi-square (2) distributed, and the Kiefer-Salmon normality tests are asymptotically Chi-square (1) distributed, see also chapter 2.

Table 5.1 : Summary Statistics of Weekly Log Price Changes: $\ln(S_t/S_{t-1})$, March 15, 1979 through February 27, 1992: 675 Observations

	BF/DM	BP/DM	DG/DM	DK/DM
Mean	0.0004	0.0004	0.0001	0.0005
St.dev	0.0076	0.0113	0.0024	0.0042
T-test	1.26	0.93	0.71	3.08
Skewness	0.49	0.49	1.09	1.49
Kurtosis	9.80	5.23	13.97	15.02
BJ-test	1326***	167***	3515***	4313***
KS-1	26.7***	26.9***	132***	249***
KS-2	1299***	140***	3382***	4064***
D(25) R	11.96	34.97*	41.74**	17.92
D(50) R	37.06	51.84	69.59**	43.63
LB(25) R ²	645***	81.37***	99.04***	14.11
LB(50) R ²	871***	110***	184***	21.21
	FF/DM	IL/DM	IP/DM	SP/DM
Mean	0.0006	0.0008	0.0005	0.0008
St.dev	0.0045	0.0057	0.0046	0.0078
T-test	3.35	3.42	2.92	2.60
Skewness	6.04	1.83	4.54	2.24
Kurtosis	67.76	33.27	53.96	21.99
BJ-test	122070***	26141***	75368***	10706***
KS-1	4110***	376***	2323***	564***
KS-2	117960***	25764***	73044***	10142***
D(25) R	33.54	23.27	35.51*	21.89
D(50) R	58.89	43.28	52.60	44.91
LB(25) R ²	6.54	63.48	4.36	8.00
LB(50) R ²	139***	191***	5.32	12.89

Legend : BF = Belgian Franc; BP = British Pound; DG = Dutch Guilder; DK = Danish Krone; DM = Deutschmark; FF = French Franc; IL = Italian Lire; IP = Irish Pound; SP = Spanish Peseta. The BJ-test denotes the Bera Jarque test for normality; KS-1 and KS-2 pertain to the Kiefer Salmon Normality test for respectively skewness and kurtosis; D(p) and LB(p) denote the Diebold and Ljung-Box test for serial correlation using p lags; * (**) [***] denotes rejection at the 10% (5%) [1%] level of the normality hypothesis.

of the normality hypothesis. In addition, Table 5.1 reports the Diebold (1988) and Ljung-Box (1978) test statistics for kth-order serial correlation in $\ln(S_t/S_{t-1})$ and $[\ln(S_t/S_{t-1})]^2$, respectively⁵. The squared exchange rate returns exhibit substantially more autocorrelation than the raw data, which is indicative of strong conditional heteroskedasticity.

In Graphs 1 and 3 two examples of exchange rate paths and official bilateral intervention limits are presented (the FF/DM and IL/DM exchange rates). Graphs 2 and 4 display corresponding graphs of log differences in these exchange rate levels, corresponding (approximately) to percentage changes in the levels. The shifts of the EMS bands due to realignments can be clearly detected in these figures. Note that the occurrence of a realignment is not always associated with the exchange rate reaching a barrier.

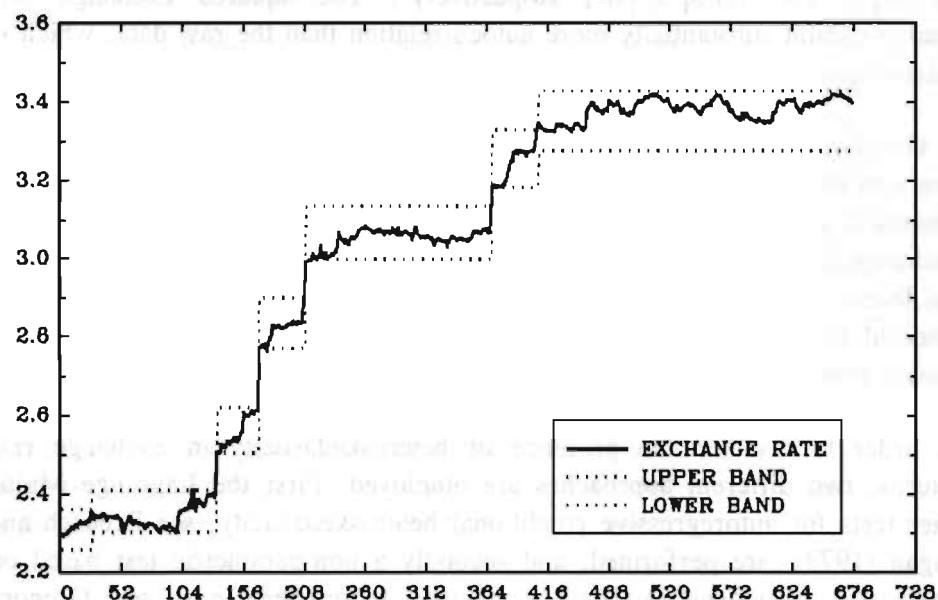
In order to test for the presence of heteroskedasticity in exchange rate returns, two different approaches are employed. First the Lagrange Multiplier tests for autoregressive conditional heteroskedasticity, see Breusch and Pagan (1979), are performed, and secondly a non-parametric test based on finite-state homogeneous Markov chains, is applied, e.g. see Gregory (1989), and Nieuwland and Verschoor (1992). Using Monte Carlo analysis Gregory (1989) concludes that under other distributions than the Normal the LM test is biased towards the null hypothesis of no ARCH, and that the Markov Chain test is superior to the LM test in terms of better finite sample properties. Both tests only require estimation under the null hypothesis of no heteroskedasticity. The results of the LM and Markov chain tests for the presence of heteroskedasticity are given in Table 5.2. Overall, the evidence presented suggests a fairly consistent rejection of the hypothesis of no heteroskedasticity. However, it is interesting to note that the results for the Danish Kroner, French Franc and Irish Pound lead to conflicting inferences regarding the independence of exchange rate return series⁶.

⁵ Diebold (1988) proposes an adjusted Ljung-Box test statistic to allow for heteroskedasticity. He showed that in the presence of ARCH effects, the Ljung-Box test has a larger empirical size than a nominal test size of 5%, because the asymptotic variance of the autocorrelations under ARCH is larger than under the null of Gaussian white noise, for the definitions of these test statistics see Chapter 2 p.26.

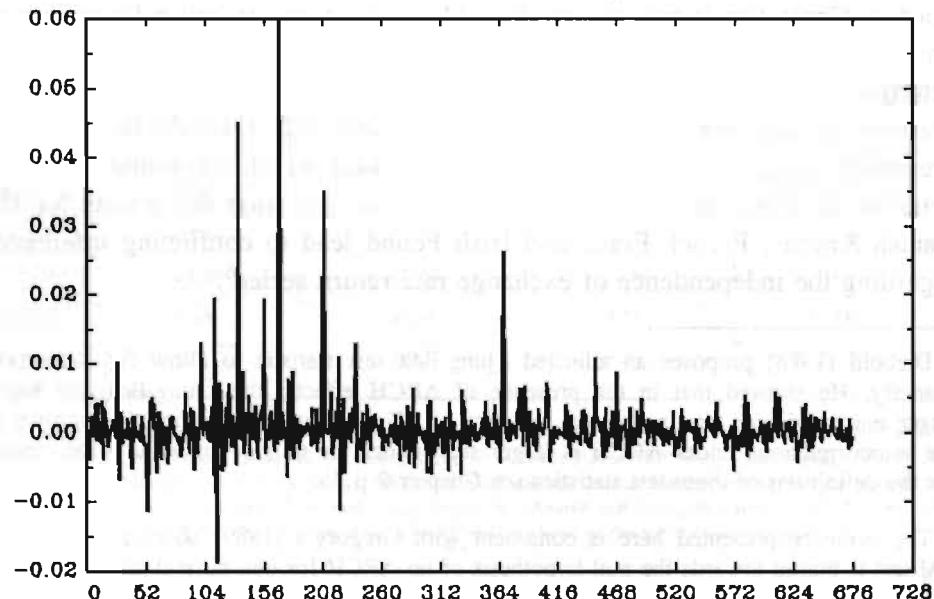
⁶ The evidence presented here is consistent with Gregory's (1989) observation that the LM test is biased towards the null hypothesis of no ARCH for non-normal distributions.

Graph 1: FRENCH FRANC / DEUTSCHMARK EXCHANGE RATE

march 15, 1979 – february 27, 1992

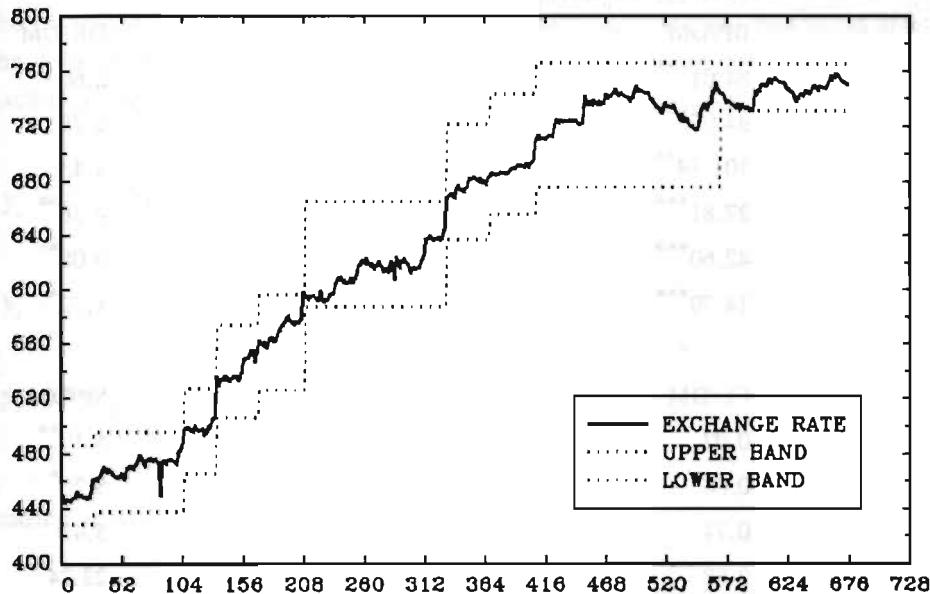
**Graph 2: WEEKLY LOG PRICE CHANGES FF/DM EXCHANGE RATE**

march 15, 1979 – february 27, 1992



Graph 3: ITALIAN LIRA/DEUTSCHMARK EXCHANGE RATE

march 15, 1979 – february 27, 1992

**Graph 4: WEEKLY LOG PRICE CHANGES IL/DM EXCHANGE RATE**

march 15, 1979 – february 27, 1992

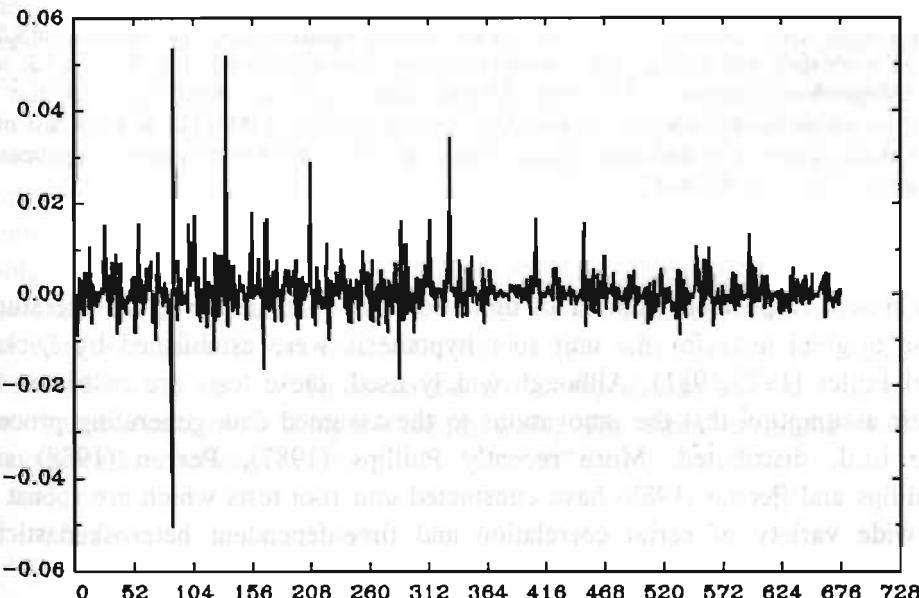


Table 5.2 : Heteroskedasticity tests of Weekly Log Price Changes : $\ln(S_t/S_{t-1})$

	BF/DM	BP/DM	DG/DM	DK/DM
LM(1)	84.01***	13.22***	29.76***	2.66
LM(2)	94.97***	13.29***	30.68***	2.75
LM(5)	101.88***	17.53***	31.41***	3.43
LRIM1	27.81***	13.23***	28.38***	4.36**
LRIM2	42.60***	18.70***	42.58***	9.08**
LRM1M2	14.79***	5.46*	14.20***	4.72*
	FF/DM	IL/DM	IP/DM	SP/DM
LM(1)	0.02	59.49***	0.04	5.19**
LM(2)	0.68	65.53***	0.13	5.25*
LM(5)	0.74	65.62***	0.22	5.43
LRIM1	9.88***	7.21***	22.21***	22.24***
LRIM2	17.55***	12.18***	30.24***	23.65***
LRM1M2	7.67**	4.97*	8.02**	1.42

The LM(p) test is given by TR^2 from a regression of squared log price changes, $\ln(S_t/S_{t-p})$ on a constant and p lags, and is asymptotically $\chi^2(p)$ distributed. LRIM1 is a LR test of independence against a first order Markov Chain ($\chi^2(1)$); LRIM2 is a LR test of independence against a second order Markov Chain ($\chi^2(3)$); LRM1M2 is a LR test of a first order against a second order Markov Chain ($\chi^2(2)$). * (**) [***] denotes significance at the 10% (5%) [1%] level.

By now a respectable number of unit root tests is available in the literature. The original tests for the unit root hypothesis were established by Dickey and Fuller (1979,1981). Although widely used, these tests are restricted by their assumption that the innovations to the assumed data generating process are i.i.d. distributed. More recently Phillips (1987), Perron (1988) and Phillips and Perron (1988) have conducted unit root tests which are robust to a wide variety of serial correlation and time-dependent heteroskedasticity

structures for the series under consideration. Given the evidence presented earlier in this paper that the exchange rates series we consider are clearly not independent, we apply the Phillips-Perron tests to our data in order to assess the presence or absence of unit roots. The test-statistics are based on one of the following three regressions, where y_t denotes the logarithm of the spot exchange rate:

$$y_t = \tilde{\mu} + \tilde{\beta}(t - T/2) + \tilde{\alpha}y_{t-1} + \tilde{u}_t \quad 2.1$$

$$y_t = \mu^* + \alpha^*y_{t-1} + u_t^* \quad 2.2$$

$$y_t = \delta y_{t-1} + u_t \quad 2.3$$

leading to six different test statistics:

$$Z(\Phi_1), Z(\Phi_2), Z(\Phi_3), Z(t_{\tilde{\alpha}}), Z(t_{\alpha^*}), Z(t_{\hat{\alpha}})$$

The simple test statistics based on the t-values of the autoregressive parameters in the respective regressions test for the presence of a unit root only. The joint statistics test for both presence of a unit root as well as for absence of constants and time trends (where appropriate). Further details on the construction of these statistics and the testing strategies used are given in Appendix 5.A. For comparison purposes we have also performed the Augmented Dickey-Fuller (ADF) test. The results of these tests are contained in Table 5.3.a. The overall conclusion that can be drawn is that we cannot reject the unit root hypothesis for seven out of the eight currencies analyzed. The only exception is the Dutch Guilder, where mean reversion is clearly present. From the remaining seven currencies the British Pound is the only one characterized by a simple random walk. The other currencies follow random walks with drift terms.

Table 5.3.a : Test for Unit Roots in the Logarithms of Spot Exchange Rates: $\ln(S_t)$
March 15, 1979 through February 27, 1992: 675 Observations

	BF/DM	BP/DM	DG/DM	DK/DM
$Z(t_\alpha^\wedge)$	1.60	-0.86	0.61	2.75
$Z(t_\alpha^*)$	-2.44	-0.81	-3.16**	-3.44***
$Z(\Phi_1)$	4.40*	0.62	5.37**	12.92***
$Z(t_\alpha^\sim)$	-1.43	-3.00	-4.10***	-2.42
$Z(\Phi_2)$	2.90	3.32	6.08**	8.56***
$Z(\Phi_3)$	2.98	4.70	8.84***	7.10**
$Z(t_{ADF})$	-1.54	-2.70	-3.56**	-2.46
<hr/>				
	FF/DM	IL/DM	IP/DM	SP/DM
$Z(t_\alpha^\wedge)$	3.08	4.06	-3.01***	2.01
$Z(t_\alpha^*)$	-2.04	-2.29	-1.39	-1.95
$Z(\Phi_1)$	8.51***	12.33***	5.11**	4.34*
$Z(t_\alpha^\sim)$	-0.31	-0.45	-0.89	-0.66
$Z(\Phi_2)$	5.91**	7.98***	3.41	3.02
$Z(\Phi_3)$	2.25	2.60	1.08	2.00
$Z(t_{ADF})$	-0.38	-0.61	-0.64	-0.57

The seven different statistics all test for a unit root in the univariate time-series representation for each of the eight spot rates, S_t , against a stationary or trend-stationary alternative. The exact forms of the tests are given in Appendix 5.A. * (**) [***] denotes significance at the 10% (5%) [1%] level. Critical values for these statistics can be found in Fuller (1976) and Dickey and Fuller (1981).

Formally we should not refer to a random walk but to a martingale since the innovations series are not independent, see Tables 5.1 and 5.2. These results are somewhat surprising since one of the objectives of the exchange rate mechanism of the EMS is to keep bilateral exchange rates within a well defined band or target zone, which consequently would lead to some kind of mean reversion to the agreed parity rates. The fact that we do not discover

this mean reversion obviously has to have a cause; our results can be reconciled with the assumption of mean reversion if we take parity revisions into account. Perron (1989) has pointed out the dangers in testing for unit roots in series that contain a "trend-break". If one would not take this break into account one could get the false impression that a unit root is present and that random shocks are permanent, whereas in reality the series is trend-stationary and random shocks are transitory.

The test-statistics Perron derives allow for a one-time trend-break, but in the ERM semi-fixed exchange rate system we are confronted with several of these breaks or interventions, leading to parity adjustments. Therefore we have performed the same unit root tests on the currencies now divided by the appropriate parities. The results are described in Table 5.3.b and clearly indicate mean reversion⁷.

Table 5.3.b : Tests for Unit Roots in the Logarithm of Parity Adjusted Spot Exchange Rates: $\ln(S_t/P_t)$ March 15, 1979 through February 27, 1992

	BF/DM	DG/DM	DK/DM	FF/DM	IL/DM	IP/DM
$Z(t_{\alpha}^{\wedge})$	-2.61***	-6.06***	-5.52***	-5.10***	-3.66***	-6.51***
$Z(t_{\alpha}^{*})$	-3.70***	-6.23***	-5.71***	-5.11***	-3.84***	-6.56***
$Z(\Phi_1)$	6.85***	19.43***	16.39***	13.09***	7.47***	21.55***
$Z(t_{\alpha}^{\sim})$	-4.85***	-6.32***	-6.42***	-6.47***	-4.92***	-6.70***
$Z(\Phi_2)$	7.89***	13.36***	13.83***	14.02***	8.13***	15.04***
$Z(\Phi_3)$	11.84***	20.03***	20.72***	21.02***	12.15***	22.55***
$Z(t_{ADF})$	-3.98***	-4.51***	-5.51***	-5.90***	-4.38***	-5.30***

For notes see Table 5.3.a

⁷ The British Pound and the Spanish Peseta are excluded, as no realignments involving these currencies occurred since their adherence to the ERM.

Summarizing we can conclude that the time series behavior of ERM exchange rates is characterized by mean reversion between two realignments, but that the occurrence of realignments leads to the non-rejection of unit roots for the overall series⁸.

5.3. Modelling EMS Exchange Rates

In the academic literature it is agreed upon that empirical distributions of exchange rate returns exhibit fatter tails than one expects from a normal distribution. See, for example, Boothe and Glassman (1987) and Hsieh (1989). In Table 5.1 we have provided ample evidence confirming this fact. These fat tails, or observed leptokurtosis, may be explained by several classes of models. See, for example, Jorion (1988) and Vlaar and Palm (1992). In this chapter we will concentrate on three possible explanations. The observed leptokurtosis may be due to the existence of discontinuities, or jumps, in the data. These jumps, especially with EMS data, can occur because of the official realignments of central rates within the EMS and because of intramarginal interventions to maintain bilateral parities. The occurrence of jumps can also explain the observed positive skewness in our EMS data. A second explanation suggests that exchange rate returns can be described by a normal distribution with time-varying parameters, captured by a GARCH process, see Hsieh, (1989). Finally we consider the possibility that EMS exchange rates are generated by a conditionally leptokurtic distribution. These considerations lead us to the maximum-likelihood estimation of the following stochastic processes:

⁸ It is however possible that the mean reversion that we find is spurious. This occurs when the realignments, i.e. the permanent adjustments in the mean, are solely responsible for the apparent "mean reversion". In this case a better description would be : mean adjusting. Inspection of the parity adjusted data however shows that there is some degree of real mean reversion between realignments, accompanied by mean adjustment at the realignment dates.

5.3.1 The Lognormal Diffusion Process

We assume that exchange rate returns follow a lognormal diffusion process: $dS/S = \alpha dt + \sigma dz$, where α is the instantaneous expected return and z is a Wiener process. This assumption implies that in discrete time $\ln(S_t/S_{t-1})$ is normally distributed with mean μ , $\mu = \alpha - \sigma^2/2$, and variance σ^2 . In discrete time the model is:

$$\ln(S_t/S_{t-1}) = \mu + \sigma z_t \quad 3.1$$

where z_t is a standard normal deviate. This process will serve as a benchmark. Given the evidence presented earlier, however, this process is unlikely to be the true data generating process. Given T independent observations the associated log likelihood function can be written as:

$$L_d = -\frac{T}{2} \ln(2\pi) + \sum_{t=1}^T \ln \left\{ \left(\frac{1}{\sigma} \right) \exp \left(\frac{-(\ln(S_t/S_{t-1}) - \mu)^2}{2\sigma^2} \right) \right\} \quad 3.2$$

5.3.2. The Mixed Jump-Diffusion Process

As mentioned above, possible jumps could be the reason for the observed leptokurtosis. The mixed jump-diffusion process considered here models these jumps. We assume that the jumps arrive according to a Poisson process, $P(\lambda)$. This means that during a time interval of length 1 (in our case one week), the mean number of jumps, and also the variance of the number of jumps, is equal to a constant λ , where $\lambda > 0$. Furthermore, we assume that the individual jump size, Y , is independently lognormally distributed, i.e. $\ln(Y) \sim N(\theta, \delta^2)$. Svensson (1989, 1990) provide examples of target zone models with Poisson components. The mixed jump diffusion process can now be described as:

$$\ln(S_t/S_{t-1}) = \mu + \sigma z_t + \sum_{i=1}^{n_t} \ln(Y_i) \quad 3.3$$

where n_t represents the actual number of jumps during the time interval, see Jorion (1988). The total exchange rate return variance is decomposed into the variance of the diffusion process plus the variance of the jump process, where both processes are assumed to be independent. The log likelihood function for this process can be expressed as:

$$L_j = -T\lambda - \frac{T}{2} \ln(2\pi) + \sum_{t=1}^T \ln \left\{ \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \frac{1}{\sqrt{(\sigma^2 + \delta^2 j)}} \exp \left(-\frac{(\ln(S_t/S_{t-1}) - \mu - \theta j)^2}{2(\sigma^2 + \delta^2 j)} \right) \right\} \quad 3.4$$

The infinite sum in the above likelihood function poses a problem in the numerical optimization of this function. The infinite sum has to be truncated at a suitable truncation number, U , so that sufficient accuracy for all parameter values is guaranteed. An upper bound for the truncation error was derived by Ball and Torous (1985). This upper bound can be used to select a suitable U . For all our purposes $U=20$ gives sufficient accuracy. The second possible explanation for the observed leptokurtosis was that exchange rate returns can be normally distributed with time-varying parameters. We will focus on the time variation in second moments and, specifically, we will estimate a diffusion-ARCH(1) and a diffusion-GARCH(1,1) process. We know that GARCH processes with conditionally normal innovations generate unconditional distributions with fat tails.

5.3.3 The Diffusion-(G)ARCH Process

Here we make a conditional distributional assumption: $\ln(S_t/S_{t-1}) \sim N(\mu, h_t)$. Conditional on all available information at time $t-1$, the process can be

described as:

$$\ln(S_t/S_{t-1})|_{t-1} = \mu + \varepsilon_t \quad , \quad \varepsilon_t = h_t^{1/2} z \quad , \quad z \sim N(0,1) \quad 3.5$$

$$h_t = E_{t-1}(\varepsilon_t^2) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad 3.6$$

In the ARCH(1) formulation β_1 is zero. The associated log-likelihood function reads:

$$L_a = -\frac{T}{2} \ln(2\pi) + \sum_{t=1}^T \ln \left\{ \frac{1}{\sqrt{h_t}} \exp \left(\frac{-(\ln(S_t/S_{t-1}) - \mu)^2}{2h_t} \right) \right\} \quad 3.7$$

Both the diffusion-(G)ARCH process and the jump-diffusion process can outperform the standard lognormal diffusion process. It is, however, possible that neither model alone can fully account for the observed leptokurtosis. Therefore, we also investigate a model that combines both a jump process and a (G)ARCH process: the combined jump-diffusion-(G)ARCH model.

5.3.4. The Jump-Diffusion-(G)ARCH Process

In this process the conditional exchange rate return consists of an ordinary drift term, a jump component, and a conditional error term.

$$\ln(S_t/S_{t-1})|_{t-1} = \mu + \varepsilon_t + \sum_{i=1}^{n_t} \ln(Y_i) \quad 3.8$$

where the conditional variance function remains as given in equation 3.6. The associated log likelihood function can be formulated as:

$$L_{ja} = -T\lambda - \frac{T}{2} \ln(2\pi) + \sum_{t=1}^T \ln \left\{ \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \frac{1}{\sqrt{(h_t + \delta^2 j)}} \exp \left(-\frac{(\ln(S_t/S_{t-1}) - \mu - \theta j)^2}{2(h_t + \delta^2 j)} \right) \right\} \quad 3.9$$

It would be inappropriate to limit our attention to jump-diffusion-ARCH processes for the following reasons. Consider the situation where a realignment has been effected just recently, which is translated into a large innovation term. In the ARCH model this would cause a sudden large increase in the conditional variance and a subsequent large drop in variance, whereas in the GARCH model the conditional variance would change more gradually. In the former case the large increase in volatility could be confused with an additional jump, causing the jump intensity to move upwards. This is less likely to happen in the latter case. This was pointed out by Vlaar and Palm (1992). Furthermore we know from empirical evidence, e.g. Baillie and Bollerslev (1989), that volatility shocks are much more persistent than is implied by an ARCH(1) model. This suggests that the parameter on the lagged conditional variance in the GARCH(1,1) model will be highly significant.

5.3.5 The Jump-Diffusion-(G)ARCH-t Process

The third explanation we offer for the observed leptokurtosis considers a conditionally fat-tailed distribution for the innovation series. Although GARCH models with conditionally normal innovations and the jump part in our models both generate unconditional fat tails, and are thus in accordance with stylized facts, these two explanations may be subsumed by a conditionally fat-tailed innovation distribution. Recent results on the temporal aggregation of GARCH processes by Nelson (1991), and Drost and Nijman (1991) show that the discrete time GARCH model which approximates the standard continuous time model with time varying volatility can well have leptokurtic rescaled residuals. In order to assess this possibility we propose a

Student-t distribution for the innovations. The model remains as is described in equations 3.6 and 3.8. The associated log likelihood function however changes into:

$$\begin{aligned}
 L_{ja} = & T \left(-\lambda - \frac{1}{2} \ln(\pi(v - 2)) + \ln \Gamma \left(\frac{v+1}{2} \right) - \ln \Gamma \left(\frac{v}{2} \right) \right) + \\
 & \sum_{t=1}^T \ln \left\{ \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \frac{1}{\sqrt{(h_t + \delta^2 j)}} \left(1 + \frac{(\ln(S_t/S_{t-1}) - \mu - \theta j)^2}{(v-2)(h_t + \delta^2 j)} \right)^{-\frac{(v+1)}{2}} \right\} \quad 3.10
 \end{aligned}$$

5.4. Empirical Results

In this section we present empirical results from the estimation procedures outlined in the previous section. First, we present the estimation results for individual models and then the various models are compared and formal hypotheses are tested. Maximum likelihood estimates of the parameters and their heteroskedasticity consistent asymptotic standard errors were obtained by pre- and post-multiplying the cross-product matrix of the first derivatives by the inverse Hessian, using the Berndt, Hall, Hall and Hausman (1974) (BHHH) algorithm⁹, see Chapter 2. In Tables 5.4-5.9, the estimation results are reported for the six stochastic processes described in the previous section. The results for the diffusion model are described in Table 5.4. This is our base case or benchmark model. All estimated values of the mean μ are positive (and five of these statistically significant) indicating that all currencies depreciated vis-a-vis the DM. The standard deviation of the BP/DM exchange rate is much larger than those for other EMS exchange rates.

⁹ All calculations were performed with the software package GAUSS.

Table 5.4 : Diffusion Models March 15, 1979 through February 27, 1992: 675 Observations

	$\hat{\mu} \cdot 10^2$	$\hat{\sigma} \cdot 10^2$	L.L.
BF/DM	0.0370 (0.0294)	0.7628*** (0.0435)	2333.52
BP/DM	0.0405 (0.0435)	1.1313*** (0.0448)	2067.44
DG/DM	0.0065 (0.0092)	0.2391*** (0.0166)	3116.66
DK/DM	0.0494*** (0.0160)	0.4162*** (0.0300)	2742.45
FF/DM	0.0579*** (0.0173)	0.4493*** (0.0707)	2690.75
IL/DM	0.0757*** (0.0221)	0.5740*** (0.0627)	2525.45
IP/DM	0.0521*** (0.0178)	0.4635*** (0.0649)	2669.81
SP/DM	0.0784*** (0.0301)	0.7810*** (0.0689)	2317.54

The heteroskedasticity consistent standard errors of the coefficients are given in parentheses; * (**) [***] denotes significance at the 10% (5%) [1%] level. L.L. denotes the log-likelihood values.

The results for the diffusion-ARCH model are summarized in Table 5.5. All but two of the estimated models result in statistically significant α_1 coefficients, thus supporting the ARCH-specification. The diffusion-GARCH estimation results are reported in Table 5.6. The estimated α_1 and β_1 coefficients are statistically significant in five out of eight cases. However, for the Belgian Franc, Dutch Guilder, French Franc and Irish Pound, the estimates of α_1 and β_1 are greater than one, indicating infinite persistence in the volatility shocks, or IGARCH behavior, see Engle and Bollerslev (1986).

**Table 5.5 : Diffusion-ARCH Models March 15, 1979 through February 27, 1992:
675 Observations**

	$\hat{\mu} \cdot 10^2$	$\hat{\alpha}_0 \cdot 10^2$	$\hat{\alpha}_1$	L.L.
BF/DM	0.0138 (0.0206)	0.0020*** (0.0004)	0.9787*** (0.2543)	2448.50
BP/DM	0.0351 (0.0415)	0.0102*** (0.0011)	0.2171*** (0.0758)	2080.61
DG/DM	0.0062 (0.0063)	0.0003*** (0.0001)	0.5075*** (0.1597)	3193.13
DK/DM	0.0279* (0.0148)	0.0012*** (0.0002)	0.3768* (0.2149)	2771.37
FF/DM	0.0596 (0.0461)	0.0013** (0.0007)	1.2753 (2.1414)	2697.51
IL/DM	0.0377 (0.0281)	0.0018*** (0.0005)	0.7261 (0.6830)	2588.10
IP/DM	0.0461* (0.0243)	0.0010*** (0.0002)	0.9231* (0.4892)	2741.35
SP/DM	0.0582** (0.0262)	0.0035*** (0.0004)	0.4678*** (0.1479)	2382.36

The heteroskedasticity consistent standard errors of the coefficients are given in parentheses; * (**) [***] denotes significance at the 10% (5%) [1%] level. L.L. denotes the log-likelihood values.

As conjectured by Diebold (1986), and Lamoureux and Lastrapes (1990), this may be the result of not accounting for discrete shifts in monetary regimes which affect the level of the unconditional variances, and, therefore can lead to misspecification of the GARCH model¹⁰. In Table 5.7 the estimation results for the mixed jump-diffusion process are presented.

⁹ Lastrapes (1989) finds that persistence of exchange rate volatility decreases when regime shifts are accounted for, diminishing the likelihood of integrated-in-variance processes.

Table 5.6 : Diffusion-GARCH Models March 15, 1979 through February 27, 1992: 675 Observations

	$\hat{\mu}(.10^2)$	$\hat{\alpha}_0(.10^2)$	$\hat{\alpha}_1$	$\hat{\beta}_1$	L.L.
BF/DM	-0.0035 (0.0510)	0.0001* (0.0000)	0.2355*** (0.0629)	0.7839*** (0.0455)	2513.92
BP/DM	0.0382 (0.0376)	0.0004 (0.0003)	0.0821** (0.0219)	0.8866*** (0.0242)	2104.43
DG/DM	0.0014 (0.0057)	0.0000*** (0.0000)	0.03583** (0.0149)	0.9637*** (0.0116)	3254.38
DK/DM	0.0291 (0.0146)	0.0005** (0.0002)	0.4096** (0.1930)	0.3682** (0.1528)	2780.14
FF/DM	0.0648 (0.0556)	0.0006 (0.0009)	0.9968 (1.3579)	0.3583 (0.3564)	2722.00
IL/DM	0.0323 (0.0340)	0.0015** (0.0007)	0.7831 (0.7427)	0.0755 (0.1489)	2588.79
IP/DM	0.0253* (0.0148)	0.0004*** (0.0001)	0.8742** (0.3731)	0.3001*** (0.1028)	2767.60
SP/DM	0.0565** (0.0269)	0.0035*** (0.0006)	0.4777*** (0.1464)	0.0000 (0.0639)	2383.88

We find that four of the estimated jump intensity coefficients, λ , are statistically significant, and λ ranges from 5.8% (the FF/DM exchange rate) to 95.5% (the BP/DM exchange rate). Interpreting the estimated λ coefficient for the FF/DM exchange rate, an estimated jump intensity of 5.8% implies an estimated expected number of jumps equal to 39 over the analyzed period. However, only six realignments involving the FF/DM exchange rate occurred during the life of the ERM. Some of the additional jumps could result from intra-marginal or exchange market interventions to maintain bilateral parities. The estimated average jump size, θ , is always positive (with a maximum of 0.8 percent for the French Franc) and sometimes significantly so, which is in accordance with the positive skewness.

**Table 5.7 : Jump-Diffusion Models March 15, 1979 through February 27, 1992:
675 Observations**

	$\hat{\mu}(.10^2)$	$\hat{\alpha}_0(.10^2)$	$\hat{\lambda}$	$\hat{\theta}(.10^2)$	$\hat{\delta}(.10^2)$	L.L.
BF/DM	0.0009 (0.0214)	0.3605*** (0.0892)	0.3573 (0.2402)	0.1011 (0.1105)	1.0997*** (0.3498)	2451.72
BP/DM	-0.1253*** (0.0479)	0.3955*** (0.1288)	0.9548** (0.3741)	0.1737** (0.0694)	1.0672*** (0.1795)	2126.01
DG/DM	-0.0054 (0.0072)	0.1280*** (0.0253)	0.2669 (0.1874)	0.0477 (0.0458)	0.3755*** (0.1227)	3233.80
DK/DM	0.0153 (0.0142)	0.2436*** (0.0492)	0.2799 (0.2348)	0.1219 (0.1147)	0.5978*** (0.2134)	2833.11
FF/DM	0.0120 (0.0107)	0.2189*** (0.0154)	0.0575* (0.0313)	0.7987* (0.4692)	1.3224** (0.6234)	3023.89
IL/DM	0.0223 (0.0202)	0.3324*** (0.0309)	0.0844 (0.0626)	0.6328* (0.3558)	1.4268* (0.7960)	2724.00
IP/DM	-0.0022 (0.0133)	0.2054*** (0.0281)	0.3835** (0.1630)	0.1417* (0.0812)	0.5428*** (0.1321)	2847.62
SP/DM	-0.0204 (0.0261)	0.4760*** (0.0529)	0.2281* (0.1234)	0.4330* (0.2571)	1.1344*** (0.2648)	2417.88

On average, when restricting our attention to the official realignments, the estimated jump intensity appears to be excessively high, and the estimated jump size is rather low. Looking at Table 5.8, where the results of the combined jump-diffusion-ARCH process are described, two interesting observations emerge. First of all we can see that compared to the results of Table 5.5, the diffusion-ARCH process, the ARCH parameter drops significantly and uniformly. The hypothesis that the process is integrated in variance can be rejected altogether, where it cannot be rejected in four cases in the diffusion ARCH case. Furthermore all ARCH parameters are highly significant.

Table 5.8 : Combined Jump-Diffusion-ARCH Models, March 15, 1979 through February 27, 1992

	$\hat{\mu}(.10^2)$	$\hat{\alpha}_0(.10^2)$	$\hat{\alpha}_1$	$\hat{\lambda}$	$\hat{\theta}(.10^2)$	$\hat{\delta}(.10^2)$	L.L.
BF/DM	-0.0202 (0.0180)	0.0004* (0.0002)	0.5568*** (0.1087)	0.4472* (0.2407)	0.0801 (0.0696)	0.7077*** (0.1772)	2509.35
BP/DM	-0.0825** (0.0357)	0.0010* (0.0002)	0.1965*** (0.0560)	0.8127* (0.1902)	0.1272*** (0.0674)	1.0617*** (0.1172)	3139.94
DG/DM	-0.0062 (0.0056)	0.0001** (0.0004)	0.4033*** (0.0867)	0.2568*** (0.0840)	0.0427 (0.0373)	0.3166*** (0.0570)	3281.69
DK/DM	0.0170 (0.0140)	0.0007*** (0.0000)	0.1197*** (0.0517)	0.1417 (0.1441)	0.2024 (0.2046)	0.7294** (0.3037)	2837.34
FF/DM	0.0105 (0.0085)	0.0004*** (0.0002)	0.2189*** (0.0679)	0.0387** (0.0148)	0.9533** (0.4152)	1.5758*** (0.4669)	3035.91
IL/DM	0.0123 (0.0238)	0.0008*** (0.0000)	0.1762*** (0.0658)	0.1051 (0.1219)	0.3730 (0.2509)	1.1408 (0.8570)	2734.13
IP/DM	-0.0130 (0.0138)	0.0003*** (0.0003)	0.2552*** (0.0832)	0.2722* (0.1595)	0.2001* (0.1196)	0.5417*** (0.1708)	2860.29
SP/DM	-0.0250 (0.0258)	0.0021*** (0.0001)	0.1367*** (0.0521)	0.1691*** (0.0643)	0.5927** (0.2318)	1.1442*** (0.2111)	2425.75

Table 5.9 : Combined Jump-Diffusion-GARCH Models, March 15, 1979 through February 27, 1992

	$\hat{\mu}(.10^2)$	$\hat{\alpha}_0(.10^2)$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\lambda}$	$\hat{\theta}(.10^2)$	$\hat{\delta}(.10^2)$	L.L.
BF/DM	-0.0220 (0.0002)	0.0000 (0.0001)	0.321*** (0.099)	0.567*** (0.113)	0.214* (0.116)	0.169 (0.125)	0.782*** (0.228)	2534.9
BP/DM	-0.0700** (0.0301)	0.0000 (0.0030)	0.106*** (0.027)	0.776*** (0.048)	0.417** (0.099)	0.259** (0.112)	1.215*** (0.137)	2162.3
DG/DM	-0.0018 (0.0052)	0.0000** (0.0000)	0.011*** (0.004)	0.975*** (0.007)	0.058*** (0.019)	0.164 (0.123)	0.624*** (0.108)	3305.6
DK/DM	0.0134 (0.0179)	0.0004 (0.0003)	0.114** (0.052)	0.271 (0.242)	0.186 (0.203)	0.185 (0.173)	0.655** (0.296)	2838.1
FF/DM	0.0093 (0.0084)	0.0002** (0.0000)	0.149* (0.079)	0.490*** (0.145)	0.034*** (0.013)	1.108*** (0.437)	1.570*** (0.463)	3036.3
IL/DM	0.0097 (0.0181)	0.0003** (0.0001)	0.160*** (0.058)	0.514*** (0.107)	0.100 (0.080)	0.366* (0.200)	1.159* (0.690)	2739.4
IP/DM	0.0015 (0.0216)	0.0002* (0.0001)	0.259*** (0.076)	0.294* (0.161)	0.151 (0.225)	0.395 (0.486)	0.648 (0.538)	2868.2
SP/DM	-0.0256 (0.0263)	0.0015 (0.0010)	0.127** (0.062)	0.217 (0.392)	0.162** (0.065)	0.608** (0.248)	1.151*** (0.218)	2426.6

Secondly, we observe that compared to the results of Table 5.7, the jump-diffusion model, the estimated jump intensity drops in 6 out of 8 cases, and on average reaches a higher level of significance. It appears that the combination of a Jump and an ARCH process is more appropriate than modelling the processes individually. Not surprisingly, the same statement can be made regarding combined jump-diffusion-GARCH process. Comparing the results of the diffusion-GARCH model to the results of the jump-diffusion-GARCH model in Table 5.9, we notice that the GARCH parameters are highly significant and that IGARCH behavior, which was a problem in the former case, is no longer present, with the exception of the Dutch Guilder. The estimated jump intensities are consistently smaller, and the average jump sizes are consistently larger in the combined jump-diffusion-GARCH case than in the simple jump-diffusion case. Given our previous remark about the magnitudes of intensities and sizes it seems that a combined jump-diffusion-GARCH process constitutes a considerable improvement over the simple jump-diffusion process. As was mentioned before, it would be inappropriate to limit our attention to jump-diffusion-ARCH processes; jump intensities may have a positive bias. This is confirmed empirically when we compare the estimated jump intensities of the jump-diffusion-ARCH process to the intensities of the jump-diffusion-GARCH process. In the latter case the intensities are substantially smaller. Given the results in Tables 5.4-5.9, it is interesting to compare the relative fit of various models. We employ generalized likelihood ratio tests to compare nested models. Table 5.10 presents the results of these tests.

We can observe that all p-values associated with the chi-square statistics are very close to zero, except the values for the IL/DM and SP/DM exchange rates in the sixth column. In the lower half of table 5.10, all but three of the p-values associated with the chi-square statistics are close to zero. Thus, the generalized likelihood ratio tests reject the simpler model in favor of the more complicated model in most of the cases. It is clear that the combined jump-diffusion-GARCH model performs best. However, in the case of the DK/DM, FF/DM and SP/DM exchange rate, the generalized likelihood ratio tests cannot reject the simpler (the combined jump-diffusion-ARCH) model in favor of the more complicated (the combined jump-diffusion-GARCH) model.

Table 5.10 : Generalized Likelihood Ratio Tests

	[a] $\chi^2_{(1)}$	[b] $\chi^2_{(2)}$	[c] $\chi^2_{(3)}$	[d] $\chi^2_{(4)}$	[e] $\chi^2_{(5)}$	[f] $\chi^2_{(1)}$
BF/DM	229.96 (0.000)	360.08 (0.000)	236.40 (0.000)	351.66 (0.000)	382.76 (0.000)	130.84 (0.000)
BP/DM	26.34 (0.000)	73.98 (0.000)	117.14 (0.000)	145.00 (0.000)	189.72 (0.000)	47.64 (0.000)
DG/DM	152.94 (0.000)	275.44 (0.000)	234.28 (0.000)	330.06 (0.000)	377.88 (0.000)	122.50 (0.000)
DK/DM	57.84 (0.000)	75.38 (0.000)	181.32 (0.000)	189.78 (0.000)	191.30 (0.000)	17.54 (0.000)
FF/DM	13.52 (0.000)	62.50 (0.000)	666.28 (0.000)	690.32 (0.000)	691.10 (0.000)	48.98 (0.000)
IL/DM	125.30 (0.000)	126.68 (0.000)	397.10 (0.000)	417.36 (0.000)	427.90 (0.000)	1.38 (0.240)
IP/DM	143.08 (0.000)	195.58 (0.000)	355.62 (0.000)	380.96 (0.000)	396.78 (0.000)	52.50 (0.000)
SP/DM	129.64 (0.000)	132.68 (0.000)	200.68 (0.000)	216.42 (0.000)	218.12 (0.000)	3.04 (0.081)
	[g] $\chi^2_{(3)}$	[h] $\chi^2_{(4)}$	[i] $\chi^2_{(3)}$	[j] $\chi^2_{(1)}$	[k] $\chi^2_{(2)}$	[l] $\chi^2_{(1)}$
BF/DM	121.70 (0.000)	172.80 (0.000)	41.96 (0.000)	115.26 (0.000)	166.36 (0.000)	51.10 (0.000)
BP/DM	118.66 (0.000)	163.38 (0.000)	115.74 (0.000)	27.86 (0.000)	72.58 (0.000)	44.72 (0.000)
DG/DM	177.12 (0.000)	224.94 (0.000)	102.44 (0.000)	95.78 (0.000)	143.60 (0.000)	47.82 (0.000)
DK/DM	131.94 (0.000)	133.46 (0.000)	115.92 (0.000)	8.46 (0.000)	9.98 (0.007)	1.52 (0.218)
FF/DM	676.80 (0.000)	677.58 (0.000)	628.60 (0.000)	24.04 (0.000)	24.82 (0.000)	0.78 (0.377)
IL/DM	292.06 (0.000)	302.60 (0.000)	301.22 (0.000)	20.26 (0.000)	30.80 (0.000)	10.54 (0.000)
IP/DM	237.88 (0.000)	253.70 (0.000)	201.20 (0.000)	25.34 (0.000)	41.16 (0.000)	15.82 (0.000)
SP/DM	86.78 (0.000)	88.48 (0.000)	85.44 (0.000)	15.74 (0.000)	17.44 (0.000)	1.70 (0.192)

[a] Diffusion against Diffusion-ARCH Model; [b] Diffusion against Diffusion-GARCH Model; [c] Diffusion against Jump-Diffusion Model; [d] Diffusion against Jump-Diffusion-ARCH Model; [e] Diffusion against Jump-Diffusion-GARCH Model; [f] Diffusion-ARCH Model against Diffusion-GARCH Model; P-Values are given in parentheses.[g] Diffusion-ARCH Model against Jump-Diffusion-ARCH Model; [h] Diffusion-ARCH Model against Jump-Diffusion-GARCH Model; [i] Diffusion-GARCH Model against Jump-Diffusion-GARCH Model; [j] Jump-Diffusion Model against Jump-Diffusion-ARCH Model; [k] Jump-Diffusion Model against Jump-Diffusion-GARCH Model; [l] Jump-Diffusion-ARCH Model against Jump-Diffusion-GARCH Model.

Our third explanation for the stylized facts in EMS exchange rates considers conditionally fat-tailed distributed innovation terms. We re-estimated the jump-diffusion-GARCH process, which was considered most appropriate, using a conditional Student t-distribution. The estimation results are contained in Table 5.11. The jump intensities decrease even further and lose a lot of their significance. The lowest jump intensity is 0.5 percent for the Dutch Guilder which corresponds to four jumps over the period analyzed. This is not too far off the actual number of realignments involving the Dutch Guilder, namely two. Most of the other jump intensities, albeit not always significant, hover between two and five per cent. Overall the jump sizes increase considerably, with a largest size for the Belgian Franc of 1.6 per cent. Furthermore we can safely conclude that EMS exchange rate return innovations (excluding the British Pound), even after allowing for possible jumps, are conditionally fat-tailed. In four cases the estimated degrees-of-freedom parameter, ν , is below 4, indicating infinite kurtosis. Regarding the GARCH parameters we can say that for the Danish Kroner and French Franc the volatility persistence has been reduced, whereas IGARCH cannot be ruled out for the Dutch Guilder and the Belgian Franc. Overall, the conclusion must be that each of the three explanations we have offered has its own merit but that the use of a conditionally fat-tailed distribution, largely affects the relevance of jump processes.

In order to determine the adequacy of the statistical specification, the models are subjected to diagnostic checks on the standardized residuals, $z_t = \hat{\epsilon}_t / \sqrt{\hat{h}_t}$, where $\hat{\epsilon}_t$ is the residual from equation (2.8) and \hat{h}_t is the estimated conditional variance from equation (2.6). From Jensen's inequality it follows that the standardized residuals, z_t , should demonstrate less absolute skewness and should be thinner tailed than their unconditional raw data counterparts. Any strong violation of this rule should be regarded as evidence of model misspecification - see Hsieh (1989). The diagnostics for the combined jump-diffusion-GARCH models are presented in Table 5.12. Overall, we find that in most cases the standardized residuals demonstrate less absolute skewness and kurtosis as compared to those reported by Table 5.1, thus supporting our model specifications. However, in the case of the BP/DM exchange rate, skewness and kurtosis are substantially larger than their unconditional raw data counterparts, indicating model misspecification.

Table 5.11 : Combined Jump-Diffusion-GARCH-t Models:

	$\hat{\mu}(.10^2)$	$\hat{\alpha}_0(.10^2)$	$\hat{\alpha}_1$	$\hat{\beta}_1$	$\hat{\lambda}$
BF/DM	-0.0194 (0.0223)	0.0000 (0.0000)	0.3329* (0.1874)	0.7023*** (0.1565)	0.0406 (0.0088)
BP/DM	-0.0732** (0.0308)	0.0000 (0.0000)	0.1149*** (0.0298)	0.7897*** (0.0584)	0.3908*** (0.1176)
DG/DM	-0.0017 (0.0052)	0.0000 (0.0000)	0.0219** (0.0108)	0.9707*** (0.0100)	0.0053 (0.0059)
DK/DM	0.0209 (0.0141)	0.0011** (0.0002)	0.1981** (0.0976)	0.0000** (0.0000)	0.0209 (0.0377)
FF/DM	0.0068 (0.0080)	0.0005*** (0.0001)	0.3224** (0.1248)	0.0000 (0.0000)	0.0214 (0.0148)
IL/DM	-0.0271 (0.0191)	0.0003 (0.0003)	0.2718* (0.1508)	0.4983*** (0.1099)	0.2542 (0.1948)
IP/DM	-0.0053 (0.0111)	0.0006** (0.0002)	0.4781*** (0.1752)	0.1965 (0.1754)	0.0487** (0.0196)
SP/DM	-0.0067 (0.0295)	0.0025** (0.0010)	0.2285** (0.1114)	0.1436 (0.2522)	0.0415 (0.0365)
	$\hat{\theta}(.10^2)$	$\hat{\delta}(.10^2)$	$\hat{\nu}$		L.L.
BF/DM	0.6134 (0.7044)	0.4001 (1.5823)	4.3015** (1.7853)		2545.00
BP/DM	0.2928** (0.1467)	1.1831*** (0.1632)	11.7509 (12.6786)		2162.86
DG/DM	0.8555*** (0.2559)	0.1152 (0.3027)	4.4121*** (0.8638)		3309.19
DK/DM	0.9324 (1.0925)	0.7506** (0.3555)	3.8902*** (0.8691)		2844.31
FF/DM	1.4906 (1.0693)	2.0505** (0.9994)	3.9302*** (1.0486)		3049.25
IL/DM	0.2783 (0.2001)	0.5280*** (0.1293)	3.1657*** (0.8197)		2751.62
IP/DM	0.6357*** (0.0671)	0.0002 (0.0004)	3.4656*** (0.6204)		2882.97
SP/DM	1.5572** (0.7331)	0.6804*** (0.2416)	5.1147*** (1.2689)		2431.53

For notes see table 5.

**Table 5.12 : Diagnostics: March 15, 1979 through February 27, 1992:
675 Observations**

Combined Jump-Diffusion-GARCH Models

	BF/DM	BP/DM	DG/DM	DK/DM
Skewness	0.65	6.25	0.56	0.96
Kurtosis	6.27	101.48	9.59	9.03
F(1)	1.29	66.88***	21.91***	0.74
F(5)	2.76**	22.16***	8.18***	0.35
	FF/DM	IL/DM	IP/DM	SP/DM
Skewness	6.98	-0.17	2.64	0.95
Kurtosis	88.15	29.88	22.96	6.88
F(1)	0.02	0.59	0.66	2.38*
F(5)	0.10	0.23	0.32	0.91

Combined Jump Diffusion-GARCH Models: t-Estimates

	BF/DM	BP/DM	DG/DM	DK/DM
Skewness	0.53	5.58	0.49	0.99
Kurtosis	5.07	87.72	8.56	9.56
F(1)	1.27	49.35***	19.66***	0.39
F(5)	1.85*	18.56***	7.42***	0.45
	FF/DM	IL/DM	IP/DM	SP/DM
Skewness	6.37	-0.22	2.73	0.89
Kurtosis	77.47	29.51	23.56	6.53
LM(1)	0.32	0.60	0.00	1.46
LM(5)	0.14	0.24	0.10	0.58

* (**) [***] denotes significance at the 10% (5%) [1%] level.

Note that a conditional Student-t distribution is preferred to a normal distribution, as evidenced by the smaller skewness and kurtosis values of the standardized residuals. In order to test for remaining heteroskedasticity, a residual-based test of the models may be carried out by regressing $(\hat{\epsilon}_t^2 - \hat{h}_t)/\hat{h}_t$ on $1/\hat{h}_t$, and on one to five lags of the dependent variable and testing whether the estimated coefficients are significantly different from zero by a conventional F-test. The results are reported under F(1) and F(5). These statistics follow F(2,671) and F(6,663) distributions, respectively. For both the combined jump-GARCH normal model and combined jump-GARCH Student-t model, rejection at the 10 % significance level of the null hypothesis of no heteroskedasticity occurs in only two cases (the BP/DM and DG/DM exchange rates). The combined jump-GARCH models appears to be successful at removing conditional heteroskedasticity from weekly EMS exchange rate movements.

5.5 Conclusions

In this chapter we have extensively studied the statistical properties of exchange rate returns within the European Monetary System. These properties, and primarily the time-varying conditional variance as it is often used as a risk measure, are important for dynamic hedging strategies, the pricing of currency options and for international asset pricing models. For the exchange rates of countries participating in the Exchange Rate Mechanism of the EMS an additional risk factor is present, namely the realignment risk. Ignoring this additional risk factor (measured by a jump process), leads to a misinterpretation of the true conditional risk, and hence to the mispricing of assets. Our results demonstrate clearly that a significant degree of mean reversion is present in the patterns followed by EMS exchange rates, although one has to approach the data quite carefully to be able to detect this mean reversion. We have explored three different explanations for the observed leptokurtosis in the distribution of exchange rate returns: (1) the existence of discontinuities, or jumps, in the data ; (2) the data may be generated by distributions with time-varying parameters; and (3) the data may be generated by a conditionally leptokurtic distribution. Our findings indicate that each of the three explanations is relevant. There is considerable interference, however,

between jumps in the distributions, on the one hand, and allowing for fat tails, on the other hand. Allowance for fat tails tends to increase the average jump size, but to reduce the frequency of jumps to more realistic proportions. Most successful in capturing the relevant features of EMS exchange rate returns is a combined jump-GARCH model with conditionally t-distributed innovations.

Appendix 5.A : Unit Root Tests

In this appendix we define the seven test statistics used in the text to test for unit roots. This presentation follows the same ordering as is given by the testing strategy in Perron (1988). Define y_t as the series to be analyzed (where y_0 is given), preferably in logarithms. Then start by running the following, most elaborate regression :

$$y_t = \tilde{\mu} + \tilde{\beta}(t - T/2) + \tilde{\alpha}y_{t-1} + \tilde{u}_t, \quad t = 1, \dots, T \quad \text{A.1}$$

This regression allows for a deterministic trend and a drift term. Now we can use three test statistics to determine the presence or absence of a unit root :

$$Z(t_{\tilde{\alpha}}) = (S_u/S_{Tl})t_{\tilde{\alpha}} - (T^3/\sqrt{48D_x} S_{Tl})(S_{Tl}^2 - S_u^2) \quad \text{A.2}$$

$$Z(\Phi_3) = (S_u^2/S_{Tl}^2)\Phi_3 - \left(\frac{1}{2S_{Tl}^2} \right)(S_{Tl}^2 - S_u^2) \left(T(\tilde{\alpha} - 1) - \left(\frac{T^6}{48D_x} \right)(S_{Tl}^2 - S_u^2) \right) \quad \text{A.3}$$

$$Z(\Phi_2) = (S_u^2/S_{Tl}^2)\Phi_2 - \left(\frac{1}{3S_{Tl}^2} \right)(S_{Tl}^2 - S_u^2) \left(T(\tilde{\alpha} - 1) - \left(\frac{T^6}{48D_x} \right)(S_{Tl}^2 - S_u^2) \right) \quad \text{A.4}$$

These statistics respectively test for the presence of a unit root alone, the presence of a unit root cum absence of a trend, and the presence of a unit root cum absence of a trend and a drift term :

$$H_0^1 : \tilde{\alpha} = 1, \quad H_0^2 : \tilde{\alpha} = 1, \tilde{\beta} = 0, \quad H_0^3 : \tilde{\alpha} = 1, \tilde{\beta} = 0, \tilde{\mu} = 0 \quad \text{A.5}$$

The various components of these statistics are defined as :

$$\Phi_1 = (2S^{*2})^{-1}(TS_0^2 - TS^{*2}) \quad \text{A.6}$$

$$\Phi_2 = (2\tilde{S}^2)^{-1}(TS_0^2 - T\tilde{S}^2) \quad \text{A.7}$$

$$\Phi_3 = (2\tilde{S}^2)^{-1}\left(T\left\{S_0^2 - (\bar{Y} - \bar{Y}_{-I})^2\right\} - T\tilde{S}^2\right) \quad \text{A.8}$$

where :

$$S_0^2 = \frac{1}{T} \sum_{t=1}^T (y_t - y_{t-1})^2 \quad , \quad S^{*2} = \frac{1}{T} \sum_{t=1}^T u_t^{*2} \quad , \quad \tilde{S}^2 = \frac{1}{T} \sum_{t=1}^T \tilde{u}_t^2$$

$$D_x = \det(X^T X) \quad , \quad \bar{Y} = \frac{1}{T} \sum_{t=1}^T y_t \quad , \quad \bar{Y}_{-I} = \frac{1}{T} \sum_{t=1}^T y_{t-1}$$

The statistics in A.6, A.7, and A.8 can be interpreted as regression F-tests. X is the $T \times 3$ matrix of explanatory variables in the OLS regression defined by A.1. Finally we need to establish a consistent estimator of the nuisance parameter :

$$\sigma_u^2/\sigma^2 = S_u^2/S_{Tl}^2 \quad \text{A.9}$$

Consistent estimators for S_u^2 are given by the estimated residual variances corresponding to the chosen regression. Consistent estimators for S_{Tl}^2 , which takes the temporal dependence of the residuals into account are constructed in the following manner :

$$S_{Tl}^2 = \frac{1}{T} \sum_{t=1}^T u_t^2 + \frac{2}{T} \sum_{\tau=1}^l \omega(\tau, l) \sum_{t=\tau+1}^T u_t u_{t-\tau} \quad A.10$$

Here again the u_t 's correspond to the residuals of the estimated regression. The $\omega(\tau, l)$'s denote a weighting scheme to guarantee a nonnegative estimate (e.g. a Newey-West Window). All our estimates were initially positive, so all the weights were set to one. Also l , the truncation lag parameter, was set at 10.

If, based on the results of the first two test statistics, rejection of the unit root hypothesis is possible then we are done. However if rejection is not possible this may be attributable to the poor power properties of these statistics relative to test statistics based on the following regression, see Perron (1988) :

$$y_t = \mu^* + \alpha^* y_{t-1} + u_t^* \quad A.11$$

Test statistics based on this regression are not invariant with respect to μ , so it must be verified, via $Z(\Phi_2)$, that the drift term μ is essentially equal to zero. Regression A.6 cannot be used as a starting point in a testing strategy as it cannot distinguish between a unit root and a linear stationary trend. Only when the absence of a trend is verified, using A.1, statistics based on A.6 can be safely used to test for a unit root :

$$Z(t_\alpha) = (S_u / S_{Tl}) t_\alpha - \left(\frac{1}{2S_{Tl}} \right) (S_{Tl}^2 - S_u^2) \left(\frac{1}{T^2} \sum_{t=1}^T (y_{t-1} - \bar{Y}_{-1})^2 \right)^{-1/2} \quad A.12$$

$$Z(\Phi_I) = (S_u^2 / S_{Tl}^2) \Phi_I - \left(\frac{1}{2S_{Tl}^2} \right) (S_{Tl}^2 - S_u^2) \left(T(\alpha^* - I) - \left(\frac{1}{4} \right) (S_{Tl}^2 - S_u^2) \left(\frac{1}{T^2} \sum_{t=1}^T (y_{t-1} - \bar{Y}_{-1})^2 \right)^{-1} \right) \quad A.13$$

The accompanying null hypotheses are :

$$H_0^4 : \alpha^* = 1 \quad , \quad H_0^5 : \alpha^* = 1, \mu^* = 0 \quad \text{A.14}$$

If rejection of the unit root hypothesis is still not possible, and the absence of a drift term is verified via $Z(\Phi_I)$, then we can turn to the simplest regression, given in A.15 and compute A.16 :

$$y_t = \alpha y_{t-1} + u_t \quad \text{A.15}$$

$$Z(t_\alpha) = (S_u/S_{Tl}) t_\alpha - \left(\frac{1}{2} \right) (S_{Tl}^2 - S_u^2) \left(S_{Tl} \left(\frac{1}{T^2} \sum_{t=1}^T y_{t-1}^2 \right)^{1/2} \right)^{-1}, H_0^6 : \alpha = 1 \quad \text{A.16}$$

Furthermore we have employed an Augmented Dickey Fuller (ADF) test to test for a unit root. The estimated regression reads :

$$\Delta y_t = \mu_0 + \mu_1 t + (\alpha_1 + \alpha_2 - 1)y_{t-1} - \alpha_2 \Delta y_{t-1} + u_t \quad \text{A.17}$$

$$H_0^7 : \alpha_2 + \alpha_1 - 1 = 0$$

The test statistic to be used is the t-value of $\alpha_1 + \alpha_2 - 1$, with critical values tabulated by Fuller (1976). The first six test statistics presented above only require estimation of a first order autoregression by least squares and a correction factor based on the structure of the residuals from this regression. This procedure is valid under a wide range of data generating processes, e.g. the tests are robust against ARCH effects.

Chapter 6

EMS Exchange Rate Expectations and Time-Varying Risk Premia

6.1 Introduction

One of the well established empirical regularities in the international financial economics literature is the finding that the forward discount is a biased predictor of the future change in the exchange rate - see the surveys on the efficiency of the foreign exchange market by Hodrick (1987) and Levich (1985). The rejection of forward market efficiency may be attributable to the irrationality of market participants, as suggested by Bilson (1981), Cumby and Obstfeld (1984) and Longworth (1981), or to the existence of time-varying risk premia, as suggested by Fama (1984), Hodrick and Srivastava (1984), Hsieh (1984) and Wolff (1987a), or to some combination of both of these phenomena. The debate regarding the relative size and variability of the exchange risk premium continues to be an issue of central concern in the financial economics literature. Conditional on the hypothesis that the foreign exchange market is efficient or rational, the existence of time-varying premia has been documented in the literature by Fama (1984), Frankel (1982), Hansen and Hodrick (1980), Hodrick and Srivastava (1984), Hsieh (1984), Korajczyk (1985), and Wolff (1987a).

Alternative methodologies to measure time-varying premia have been explored in the literature. First, models that are based strictly on the time series properties of spot and forward exchange rates and asset prices were examined, see the latent variable model of Hansen and Hodrick (1983) with its extensions by Hodrick and Srivastava (1984), Campbell and Clarida (1987), and Giovannini and Jorion (1987). The presence of conditional

heteroskedasticity in forecast errors prompted Domowitz and Hakkio (1985) to model a time-varying risk premium using the autoregressive conditional heteroskedasticity (ARCH) framework. Korajczyk (1985) noted that the variability of risk premia in theory can be related to variations in expected real interest rates. A second approach is to employ some measurement of market fundamentals in an attempt to test specific theories of the risk premium. Frankel (1982) and Frankel and Engel (1984) examined an asset market equilibrium model based on assets demands derived from a two-period mean-variance maximization problem. The third approach for assessing the risk premium interpretation (the existence of time-varying risk premia) attempts to measure expected depreciation directly, thereby avoiding reliance on inferences from realized depreciations, see Frankel and Froot (1987b), Froot and Frankel (1989), and Cavaglia, Verschoor and Wolff (1991b), for instance. This would not tell us about the economic determinants of risk premia, but it could tell us about the relative importance of risk and market inefficiency in explaining the forward discount bias.

This chapter extends the analysis of Domowitz and Hakkio (1985) to a survey data set of exchange rate expectations covering a wide range of EMS currencies over a different sample period, thereby combining the first and third approach. The principal benefit of using such data is that one obtains a direct measure of agents' beliefs, thus allowing for separate testing of an underlying model of exchange rate determination and a hypothesis about expectations, whereas previous work has been explored **conditional** on the hypothesis that the foreign exchange market is efficient or rational. Our survey data set begins in January 1986 and ends in September 1991 covering a period of US Dollar depreciation (and Deutschmark appreciation) relative to the currencies we review. Cavaglia, Verschoor and Wolff (1991b) have recently examined a survey data set of exchange rate expectations, that includes several EMS currencies. Their findings indicate that for EMS exchange rates relative to the Deutschmark variation in the forward discount primarily reflects changes in risk premia rather than changes in expected depreciation and, thus, that the forward discount bias is primarily attributable to significant variation in the risk premium component. As in Domowitz and Hakkio (1985) and Diebold and Pauly (1988), we employ exchange rate data on a monthly basis. Diebold (1988) and Baillie and Bollerslev (1989) demonstrate that ARCH effects in foreign exchange rates tend to disappear

under temporal aggregation, and although significant daily or weekly ARCH effects may generate a time varying risk premium, this risk premium may appear to be constant with monthly data.¹ However, results covering nearly all EMS currencies provide an interesting complement to previous work that has largely focused on the five most actively traded currencies vis-a-vis the US Dollar. Conditional heteroskedasticity appears to be a prominent feature of exchange rate behavior in the post-EMS period - see Diebold and Pauly (1988b), and Nieuwland, Verschoor and Wolff (1991)², for instance.

This chapter is presented in six sections. Section 6.2 presents a brief review of the theory. In section 6.3., the construction of the exchange rate survey is outlined and summary statistics describing the data are provided. In section 6.4., the presence of time-varying risk premia is examined as well as the presence of heteroskedastic OLS residuals. The models employed to capture the time-varying risk premia are explained in section 6.5. Section 6.6. presents the main empirical results of this investigation and section 6.7. contains our concluding remarks.

6.2 A Brief Review of Theory

Many of the models that have been developed to explain the existence and nature of a risk premium in the forward foreign exchange market share some common theoretical ground. The foundations of these models are imbedded in the Lucas'(1982) model for the international economy. In this economy, rational representative agents are assumed, which maximize their expected utility intertemporally. A general implication of this rational behavior is that the equilibrium price of an asset can be determined by equating the loss in

¹ Indeed, Kendall and McDonald (1989) on using weekly data for the Australian/US Dollar rate obtain a significant estimate for the ARCH-M parameter, as do McCurdy and Morgan (1989) using daily Canadian futures data. However Pagan and Ullah (1988) find strong support for the presence of a time-varying risk premium in the Canadian/US Dollar market with monthly data over the earlier time period 1970-1978.

² The observed leptokurtosis may be due to the existence of discontinuities, or jumps, in the data. These jumps can occur for at least two reasons. First, there are the official realignments of central rates within the EMS, and secondly one can think of intramarginal interventions to maintain bilateral parities.

marginal utility from giving up one unit of the domestic currency at time t to buy the asset, to the expected gain in marginal utility from holding the asset and selling at time $t+1$.

$$U'(C_t) \pi_t = E_t(\rho U'(C_{t+1}) \pi_{t+1} R_{i,t+1}) \quad 2.1$$

$$E_t(Q_{t+1} R_{i,t+1}) = 1 \quad 2.2$$

$$Q_{t+1} = \frac{\rho U'(C_{t+1}) \pi_{t+1}}{U'(C_t) \pi_t} \quad 2.3$$

Q_{t+1} is the intertemporal marginal rate of substitution of a unit of domestic currency between periods t and $t+1$, ρ is a discount factor, π_t is the purchasing power of the domestic currency at time t , and $R_{i,t+1}$ is the gross return at time $t+1$ of an investment of one unit of the domestic currency in asset i at time t . Now assume that an asset with a certain nominal return (denominated in domestic currency) exists, and denote its gross return as $R_{f,t+1}$. The risk free return must also satisfy the Euler equation given in 2.2 :

$$R_{f,t+1} = \frac{1}{E_t(Q_{t+1})} \quad 2.4$$

If we consider investing domestic currency abroad, this requires conversion into foreign currency. Let S_t be the domestic price of a unit of foreign currency, so $s_{t+1} - s_t = \ln(S_{t+1}/S_t)$ is the rate of depreciation of the domestic currency against the foreign currency. Let the foreign currency risk free return be $R_{f,t+1}^*$. Then the return in domestic currency from investing in the foreign money market and bearing the foreign exchange risk is : $R_{f,t+1}^* S_{t+1}/S_t$. This return must also satisfy the Euler equation 2.2. If we assume that s_{t+1} and $q_{t+1} = \ln(Q_{t+1})$ are jointly conditionally normally distributed and define $R_{f,t+1} = \exp(i_t)$ and $R_{f,t+1}^* = \exp(i_t^*)$, then we can derive :

$$i_t = -E_t(q_{t+1}) - 0.5 \operatorname{Var}_t(q_{t+1}) \quad 2.5$$

$$i_t^* + E_t(s_{t+1} - s_t) - i_t = -0.5 \operatorname{Var}_t(s_{t+1}) - \operatorname{Cov}_t(s_{t+1}, q_{t+1}) \quad 2.6$$

The forward foreign exchange rate $F_{t,t+1}$ can be established from a no-arbitrage condition as :

$$F_{t,t+1} = S_t \frac{R_{f,t+1}}{R_{f,t+1}^*} \quad 2.7$$

which is the same as covered interest rate parity. In logarithms this becomes:

$$i_t = i_t^* + f_{t,t+1} - s_t \quad 2.8$$

$$E_t(s_{t+1} - s_t) = (f_{t,t+1} - s_t) - 0.5 \operatorname{Var}_t(s_{t+1}) - \operatorname{Cov}_t(s_{t+1}, q_{t+1}) \quad 2.9$$

The risk premium in the forward foreign exchange market, $E_t(s_{t+1}) - f_{t,t+1}$, is determined by the conditional variance of the log exchange rate and by the conditional covariance between the exchange rate and the intertemporal marginal rate of substitution. Equation 2.9 forms the basis for many econometric models for the risk premium in the forward foreign exchange market. Two examples are given by the models in Kaminsky and Peruga (1990) and Domowitz and Hakkio (1985). Kaminsky and Peruga (1990) start from the one good version of the Lucas model and assume a utility function displaying constant relative risk aversion. Their expression for the risk premium is in essence the same as in 2.9. Domowitz and Hakkio use the two country, two good version of the same model, assume a utility function of the Cobb-Douglas form, and assume that the exogenous variables in the model follow conditionally normal AR(1) processes. In their analysis they obtain the following expression for the risk premium :

$$E_t(s_{t+1}) - f_{t,t+1} = -0.5(h_{m,t+1} - h_{n,t+1}) \quad 2.10$$

The risk premium depends on the relative uncertainty of the domestic money supply ($h_{m,t+1}$) to the foreign money supply ($h_{n,t+1}$). If the domestic and foreign per capita money supplies are equally uncertain then there will be no risk premium, and uncovered interest parity holds. In their empirical analysis Domowitz and Hakkio have replaced their theoretical derivation of the risk premium by a risk premium generated by a GARCH-in-mean model.

6.3 The Survey Data and Their Summary Statistics

Since 1985, Business International Corporation has been conducting a monthly survey of exchange rate expectations covering five currencies relative to the German Mark which are published in its Cross Rates Bulletin. For publication purposes, survey participants are asked a few days prior to month's end to fax three, six and twelve months ahead expectations of a number of currencies with projections being made from the beginning of the following month. Thus, for instance, the three, six and twelve months ahead expected French Franc / German Mark rate recorded on December 27th, 1990 reflect a slightly longer forecast horizon as they represent the expected spot rate on April 1st, 1991, June 1st, 1991 and January 2nd, 1992 respectively³. The dates on which the surveys are conducted have been recorded as well as the spot, three, six, and twelve month ahead forward rates recorded on that particular day. The thirty-odd participants of the survey are treasurers of multinationals and private banks residing in four of the world's continents. Although not all participants will provide their views regarding a particular currency, the response rate is at worst 60 percent. The Cross

³ Although the notation used in this chapter will be presented as if the survey was constructed on December 31 (in the example at hand), care has been exercised throughout the empirical analysis to ensure that conditional expectations are computed on the proper information set. The few days discrepancy between the maturity date of the forward contracts and the maturity date of the expected spot rate is no larger than that of other authors who have examined the efficiency of foreign exchange markets.

Rates Bulletin reports the geometric mean forecast of the responses received, thus minimizing the effect of extreme forecasts. Unfortunately disaggregated survey respondent data are not available, although the standard deviation of the respondents' expectation is reported.

Conditional on market efficiency and rational expectations, the forward exchange rate is equal to the expected future spot rate plus a risk premium. The use of survey data allows the direct measurement of a risk premium from the decomposition of the forward discount into its two components - expected depreciation and the risk premium:

$$F_{t,t+k} - S_t = (E_t S_{t+k} - S_t) + P_t^k \quad 3.1$$

Here S_t is defined as the natural logarithm of the spot exchange rate at time t , $E_t S_{t+k}$ is defined as the expected logarithm of the spot exchange rate at time $t+k$ formed at time t and $F_{t,t+k}$ is defined as the natural logarithm of the forward rate at time t for delivery at time $t+k$ and P_t^k is the associated risk premium. The left-hand side of 3.1 is the forward discount, and the right-hand side is the expected rate of depreciation of the home currency relative to the foreign currency (German Mark - the exchange rates are expressed as units of home currency per unit of foreign currency) plus the risk premium. Because the surveys are direct estimates they do not require us to assume any particular model of expected depreciation or of the risk premium. To give equation 3.1 economic content, a model of international asset pricing that describes the determination of P_t^k is required. Equilibrium models of international asset pricing that provide us with such descriptions are presented, for instance, in Adler and Dumas (1983), Hodrick (1981), Hodrick and Srivastava (1984), Roll and Solnik (1977), and Stulz (1981).

Tables 6.1.a and 6.1.b provide summary statistics for the expected exchange rate depreciation and forward discount across forecast horizon and across currencies. The summary statistics for the risk premium across horizon and across currencies are reported in table 1c. For the period analyzed (January 1st, 1986 through September 1st, 1991) the standard deviations of both the expected depreciation and forward discount across the 3, 6, and 12 month

horizons are generally larger for the Italian Lire relative to the German Mark than the corresponding estimates of the other EMS exchange rates involving the German Mark⁴. Comparing Tables 6.1.a and 6.1.b, one notes that in general the expected rates of depreciation and the forward discount are of the same sign. Thus the currencies that were expected to depreciate were at a forward discount.

This confirms the results of Frankel and Froot (1987a, 1987b). Table 6.1.c suggest the presence of time-varying risk premia, thus regarding domestic and foreign assets as imperfect substitutes. The results differ from summary statistics reported by Frankel and Froot (1987a, 1987b), which demonstrate surprisingly large exchange risk premia in a number of cases. In order to assess the distributional properties of the expected depreciation, forward discount, and risk premia series, the Bera Jarque (1982) Normality test and the Kiefer Salmon (1983) Lagrange Multiplier normality test are reported in Tables 6.1.a, 6.1.b, and 6.1.c, where the former is a joint test using both skewness and kurtosis and the latter is an LM test for normal skewness (KS-1) and normal kurtosis (KS-2), respectively⁵. Overall, the evidence presented suggests a fairly consistent rejection of the normality hypothesis. Failure to reject the null occurs in only seven out of forty-five cases. Thus, in spite of the notion that leptokurtic unconditional densities of ARCH processes approach normality by temporal aggregation - see Diebold (1988) and Baillie and Bollerslev (1989) - it appears that the monthly series used here may be characterized as highly leptokurtic. This is in line with Koedijk, Schafgans and de Vries (1990) who find that for EMS exchange rates ARCH effects become less important in time aggregation whereas fat tails remain important.

⁴ The provisions of the EMS provide for participating countries to maintain their exchange rates within bilateral limits of $\pm 2.25\%$ ($\pm 6\%$ for Italy (until January 1990) and Spain (since June 1989)). In September 1992 the United Kingdom and Italy were forced to suspend their membership of the ERM.

⁵ The Bera-Jarque test is asymptotically Chi-square (2) distributed, and the Kiefer-Salmon normality tests are asymptotically Chi-square (1) distributed.

Table 6.1.a : Summary Statistics of Expected Depreciation: $E_t S_{t+k} - S_t$, January 1, 1986 through September 1, 1991 : 69 Observations

	BF/DM	DG/DM	FF/DM	IL/DM	SP/DM
3 months					
Mean	0.0018	0.0028	0.0048	0.0067	0.0121
St.dev	0.0107	0.0098	0.0063	0.0257	0.0147
T-test	1.41	2.43	6.41	2.16	6.84
Skewness	2.15	0.11	0.26	-2.52	0.65
Kurtosis	11.14	10.87	3.48	18.36	4.84
BJ-test	243.4***	178.2***	1.45	750.8***	14.6***
KS-1	52.9***	0.14	0.78	72.9***	4.8**
KS-2	190.5***	178.0***	0.67	677.9***	9.8***
6 months					
Mean	0.0046	-0.0010	0.0077	0.0122	0.0198
St.dev	0.0114	0.0078	0.0125	0.0242	0.0202
T-test	3.35	-1.08	5.11	4.19	8.14
Skewness	1.13	-1.86	0.20	-2.90	0.83
Kurtosis	5.60	11.15	2.76	19.47	4.57
BJ-test	34.1***	230.5***	0.65	876.5***	15.1***
KS-1	14.6***	39.6***	0.48	96.5***	8.0***
KS-2	19.5***	190.9***	0.17	778.0***	7.1***
12 months					
Mean	0.0060	-0.0024	0.0122	0.0168	0.0252
St.dev	0.0147	0.0102	0.0144	0.0306	0.0196
T-test	3.43	-1.93	7.03	4.54	10.67
Skewness	0.79	-0.31	0.16	-1.50	0.60
Kurtosis	4.20	5.07	2.65	10.17	3.32
BJ-test	11.4***	13.4***	0.65	173.6***	4.4**
KS-1	7.3***	1.1	0.30	28.9***	4.1**
KS-2	4.1**	12.3***	0.35	147.7***	0.3

Legend: BF = Belgian Franc; DG = Dutch Guilder; DM = Deutschmark; FF = French Franc; IL = Italian Lire; SP = Spanish Peseta. The BJ-test denotes the Bera Jarque test for normality; KS-1 and KS-2 pertain to the Kiefer Salmon Normality test for respectively skewness and kurtosis; * (**) [***] denotes rejection at the 10% (5%) [1%] level of the normality hypotheses.

Table 6.1.b : Summary Statistics of Forward Discount: $F_{t,t+k} - S_p$, January 1, 1986 through September 1, 1991: 69 Observations

	BF/DM	DG/DM	FF/DM	IL/DM	SP/DM
3 months					
Mean	0.0062	0.0017	0.0080	0.0159	0.0170
St.dev	0.0084	0.0053	0.0050	0.0110	0.0088
T-test	6.09	2.62	13.48	12.01	15.94
Skewness	0.72	-4.69	0.79	4.67	-0.25
Kurtosis	12.57	39.45	3.95	32.08	2.73
BJ-test	269.2***	4072***	9.9***	2681***	0.92
KS-1	6.0**	254***	7.3***	251***	0.71
KS-2	263.2***	3818***	2.6	2430***	0.22
6 months					
Mean	0.0112	0.0027	0.0151	0.0312	0.0314
St.dev	0.0114	0.0059	0.0092	0.0270	0.0167
T-test	8.11	3.78	3.61	9.60	15.63
Skewness	2.19	-3.65	0.66	6.11	-0.65
Kurtosis	18.97	28.01	3.27	46.06	2.94
BJ-test	788.2***	1951***	5.4**	5760***	4.8**
KS-1	54.9***	153***	5.2**	429***	4.8**
KS-2	733.3***	1798***	0.2	5331***	0.0
12 months					
Mean	0.0209	0.0064	0.0285	0.0584	0.0587
St.dev	0.0170	0.0137	0.0168	0.0431	0.0297
T-test	10.21	3.88	14.11	11.25	16.45
Skewness	1.12	4.57	0.41	5.70	-0.76
Kurtosis	11.04	35.94	2.56	42.36	3.10
BJ-test	200.4***	3360***	2.46	4828***	6.6**
KS-1	14.5***	241***	1.89	374***	6.6**
KS-2	185.9***	3119***	0.57	4454***	0.0

Legend: BF = Belgian Franc; DG = Dutch Guilder; DM = Deutschmark; FF = French Franc; IL = Italian Lire; SP = Spanish Peseta. The BJ-test denotes the Bera Jarque test for normality; KS-1 and KS-2 pertain to the Kiefer Salmon Normality test for respectively skewness and kurtosis; * (**) [***] denotes rejection at the 10% (5%) [1%] level of the normality hypotheses.

Table 6.1.c : Summary Statistics of Risk Premium : P_t^k , January 1, 1986 through September 1, 1991: 69 Observations

	BF/DM	DG/DM	FF/DM	IL/DM	SP/DM
3 months					
Mean	0.0044	-0.0011	0.0032	0.0092	0.0049
St.dev	0.0125	0.0085	0.0052	0.0250	0.0144
T-test	2.91	-1.08	5.16	3.07	2.80
Skewness	-0.63	-1.16	0.49	2.96	0.27
Kurtosis	8.37	6.90	4.68	16.21	5.88
BJ-test	87.6***	59.1***	10.8***	602.98***	24.7***
KS-1	4.6**	15.4***	2.7*	100.94***	0.8
KS-2	83.0***	43.7***	8.1***	502.04***	23.9***
6 months					
Mean	0.0065	0.0037	0.0074	0.0190	0.0116
St.dev	0.0133	0.0067	0.0086	0.0329	0.0221
T-test	4.09	4.57	7.08	4.79	4.38
Skewness	1.84	1.01	0.15	5.18	-0.58
Kurtosis	12.94	5.11	3.02	33.01	6.39
BJ-test	323.2***	24.4***	0.27	2898***	36.9**
KS-1	39.0***	11.6***	0.27	308.9***	3.9**
KS-2	284.2***	12.8***	0.00	2590***	33.0***
12 months					
Mean	0.0149	0.0088	0.0163	0.0417	0.0336
St.dev	0.0191	0.0162	0.0121	0.0490	0.0296
T-test	6.46	4.48	11.19	7.06	9.42
Skewness	1.82	4.17	1.03	5.19	-0.16
Kurtosis	11.22	28.84	4.81	37.55	2.64
BJ-test	232.3***	2119***	21.73***	3741***	0.67
KS-1	38.1***	199.8***	12.27	309.5***	0.30
KS-2	1944.2***	1919***	9.47***	3432***	0.37

Legend: BF = Belgian Franc; DG = Dutch Guilder; DM = Deutschmark; FF = French Franc; IL = Italian Lire; SP = Spanish Peseta. The BJ-test denotes the Bera Jarque test for normality; KS-1 and KS-2 pertain to the Kiefer Salmon Normality test for respectively skewness and kurtosis; * (**) [***] denotes rejection at the 10% (5%) [1%] level of the normality hypotheses.

6.4 Time-Varying Exchange Risk Premia

Survey expectations data can be exploited to decompose the forward discount bias into portions attributable to irrational behaviour of economic agents or to the existence of time-varying risk premia - see Frankel and Froot (1987b), Froot and Frankel (1989), and Cavaglia, Verschoor, and Wolff (1991b), for instance. Here we start from the notion that regardless of how expectations are formed, under risk neutrality, the expected rate of depreciation will be equal to the forward discount, or :

$$F_{t,t+k} - S_t = E_t S_{t+k} - S_t \quad 4.1$$

Where E_t denotes the survey expectations. Under risk aversion however a risk premium is required to compensate agents for holding an open position in foreign currency instead of selling the currency forward, see equation 3.1. Assets are viewed as imperfect substitutes.

In order to test whether the existence of time-varying risk premia is the economically important reason for rejection of forward market efficiency, the following equation may be fitted⁶:

$$E_t S_{t+k} - S_t = \alpha + \beta(F_{t,t+k} - S_t) + \epsilon_t \quad 4.2$$

where ϵ_t is a random error term. The null hypothesis of perfect substitutability implies that $\alpha=0$ and $\beta=1$. The degree to which changes in the forward discount reflect changes in the risk premium can be inferred from a regression of expected depreciation on the forward discount (equation 4.2). Under the hypothesis that the correlation of the risk premium with the forward

⁶ Tests of perfect substitutability are usually cast in the form given by 4.2. In compliance with the existing literature, we have estimated this equation using OLS. Equation 4.2 may however suffer from a simultaneity bias. This problem might be solved by fitting the reversed equation, with the expected rate of depreciation as a regressor. This however suffers from an errors-in-variables bias when the expectations are measured with error. Consequently we are stuck between Scylla and Charybdis. The estimation results from the reversed regression are qualitatively the same as the results for the substitutability test, discussed in the main text.

discount is zero (no time-varying risk premia), β will equal one. From equation 4.2 it follows that:

$$\beta = \frac{\text{Cov}(F_{t,t+k} - S_t, E_t S_{t+k} - S_t)}{\text{Var}(F_{t,t+k} - S_t)} \quad 4.3$$

Hence, the decomposition of the forward discount in 3.1 implies that:

$$\beta = 1 - \frac{\text{Cov}(F_{t,t+k} - S_t, P_t^k)}{\text{Var}(F_{t,t+k} - S_t)} \quad 4.4$$

The results of fitting equation 4.1 for each currency and for each forecast horizon ($k = 3, 6$, and 12 months) are reported in Table 6.2. It is interesting to note that - with the exception of the Dutch Guilder at the 3 month horizon - the results provide a consistent rejection of the hypothesis of perfect substitutability ($\alpha=0$ and $\beta=1$ jointly) for all forecast horizons and EMS exchange rates. Rejection of the hypothesis $\beta=1$ (no time-varying risk premia) was obtained in twelve out of fifteen cases, so there is considerable evidence of significant variation in risk premia. Similar results were obtained by MacDonald and Torrance (1989) for survey based tests of uncovered interest parity. In contrast, Frankel and Froot (1987b) and Froot and Frankel (1989) obtain estimates of β that were insignificantly different from one for survey based tests using some of the major currencies relative to the US Dollar covering the period 1981-1985. Thus, as in most models in which sterilized foreign exchange intervention is effective, variation in the forward discount for EMS currencies reflect a statistically significant degree of variation in the risk premium component.

In order to test for the presence of heteroskedastic OLS residuals (ϵ_t from equation 4.1), two different approaches are employed. First the Lagrange Multiplier tests for autoregressive conditional heteroskedasticity - see Breusch and Pagan (1979) - are performed, and secondly the non-parametric test based on finite-state homogeneous Markov chains is applied, see Gregory (1989) and see also Chapter 2, pp 31-35. Using Monte Carlo analysis Gregory (1989) concludes that under other distributions than the Normal the LM test is biased towards the null hypothesis of no ARCH, and

Table 6.2 : Tests of Perfect Substitutability: $E_t S_{t+k} - S_t = \alpha + \beta(F_{t,t+k} - S_t) + e_p$, January 1, 1986 through September 1, 1991 : 69 observations

	3 months		
	$\hat{\alpha}$	$\hat{\beta}$	F
BF/DM	0.0006 (0.0016)	0.2069*** (0.1538)	18.85*** (0.000)
DG/DM	0.0012 (0.0011)	0.9441 (0.1957)	0.62 (0.543)
FF/DM	-0.0013 (0.0012)	0.7629** (0.1236)	18.86*** (0.000)
IL/DM	-0.0035 (0.0053)	0.6413 (0.2739)	5.63*** (0.005)
SP/DM	0.0028 (0.0037)	0.5482*** (0.1917)	6.79*** (0.002)
	6 months		
	$\hat{\alpha}$	$\hat{\beta}$	F
BF/DM	0.0010 (0.0019)	0.3247*** (0.1165)	28.66*** (0.000)
DG/DM	-0.0029*** (0.0009)	0.7217*** (0.1347)	13.08*** (0.000)
FF/DM	-0.0067*** (0.0020)	0.9653 (0.1155)	23.58*** (0.000)
IL/DM	0.0073* (0.0044)	0.1581*** (0.1078)	52.09*** (0.000)
SP/DM	0.0086* (0.0050)	0.3569*** (0.1410)	28.78*** (0.000)
	12 months		
	$\hat{\alpha}$	$\hat{\beta}$	F
BF/DM	0.0011 (0.0027)	0.2400*** (0.1019)	64.14*** (0.000)
DG/DM	-0.0029*** (0.0014)	0.0757*** (0.0907)	77.16*** (0.000)
FF/DM	-0.0048*** (0.0025)	0.6007*** (0.0753)	99.87*** (0.000)
IL/DM	0.0106* (0.0062)	0.1058*** (0.0858)	118.72*** (0.000)
SP/DM	0.0123*** (0.0050)	0.2199*** (0.0761)	164.66*** (0.000)

The standard errors of the coefficients are given in parentheses; * (**) [***] denotes rejection at the 10% (5%) [1%] level for the hypotheses that $\alpha=0$ or $\beta=1$. The F-statistic pertains to the joint hypothesis that $\alpha=0$ and $\beta=1$ (p-Values are given in parentheses).

that the Markov Chain test is superior to the LM test in terms of better finite sample properties. Both tests only require estimation under the null hypothesis of no heteroskedasticity and are appropriate under all distributional assumptions⁷. The development of this test is discussed in Chapter 2.

The results of the LM and Markov chain tests for the presence of heteroskedasticity are given in Table 6.3.

Table 6.3 : Heteroskedasticity tests of OLS residuals: $E_f S_{t+k} S_t = \alpha + \beta(F_{t,t+k} S_t) + \epsilon_p$, January 1, 1986 through September 1, 1991 ; 69 Observations

	BF/DM	DG/DM	FF/DM	IL/DM	SP/DM
3 months					
LM(1)	0.25	1.27	3.29*	1.13	14.01***
LM(2)	0.48	1.44	6.81	4.76	3.52
LM(5)	2.83	3.42	10.13	9.11	29.81***
LRIM1	0.02	5.46**	0.60	0.13	0.38
LRIM2	0.16	6.61*	1.15	3.10	0.39
6 months					
LM(1)	20.77***	0.04	1.65	0.27	0.60
LM(2)	30.56***	0.06	2.66	9.22***	1.40
LM(5)	4.90	5.28	5.21	26.05**	4.19
LRIM1	3.38*	0.02	5.46**	3.38*	1.22
LRIM2	4.27	0.40	7.47*	4.35	2.36
12 months					
LM(1)	14.37***	0.13	1.28	0.03	0.11
LM(2)	10.86***	0.22	1.29	10.31***	2.22
LM(5)	3.36	1.91	2.83	13.08***	6.63
LRIM1	3.38*	0.37	0.38	8.05***	0.02
LRIM2	4.06	2.23	0.55	9.58**	2.96

The LM(p) test is estimated by a regression of squared OLS residuals [equation (2)] on a constant and p lags, and is asymptotically Chi-square(p) distributed. LRIM1 is a Likelihood Ratio test of independence against a first order Markov Chain, and is distributed as Chi-square(1); LRIM2 is a Likelihood Ratio test of independence against a second order Markov Chain, and is distributed as Chi-square(3). * (**) [***] denotes significance at the 10% (5%) [1%] level.

⁷ Weiss, (1986a) has shown that the proposed LM-test is appropriate for non-normal distributions, provided some moment conditions are satisfied. The Markov Chain test is completely distribution free, see also Chapter 2.

Overall, the evidence presented suggests a weak rejection of the hypothesis of no heteroskedasticity. However, it is interesting to note that the results for the Belgian Franc and the Italian Lira at the 6 and 12 month horizon provide a strong rejection of the null hypothesis. Thus, although ARCH effects tend to weaken with less frequently sampled data, in several cases the expected EMS exchange rate changes still display significant ARCH effects at the 3, 6, and 12 months horizons. The evidence presented contrasts with the results of Domowitz and Hakkio (1985), who found no significant ARCH effects, except in the case of Japan.

6.5 Modelling Time-Varying Risk Premia

Hodrick's (1987) and Levh's (1985) reviews of the literature on the efficiency of foreign exchange market suggest that there is overwhelming evidence in favor of the view that forward rates are biased predictors of future spot rates. For the EMS currencies examined, rejection is generally attributed to the presence of a significant time-varying risk premium. Several theoretical models have been put forward which generate risk premia in foreign exchange markets, examples are Hodrick and Srivastava (1984), Domowitz and Hakkio (1985), Diebold and Pauly (1988a) and Kaminsky and Peruga (1990). Most of these theories share Lucas' (1982) model for the international economy as a starting point, see section 6.2. Although this dynamic general equilibrium model provides useful insights into the possible structure of risk premia in the forward foreign exchange markets, direct tests of this model are impossible without further restrictions. This is due to the general stochastic structure of the model. The second common denominator in these models is that in general the risk premium depends on the conditional probability distribution of the future spot rate, which may lead to a time-varying risk premium, if this distribution is time-varying. Empirically many specifications for such a risk premium have been employed which depend on the conditional variance of the spot rate. Nevertheless, in their review Bollerslev, Chou, and Kroner (1992) note that: "A satisfactory model for the time varying risk premium in the forward foreign exchange market has yet to be formulated".

In this chapter we adopt a different approach which is inspired by the availability of survey data. We do not have to assume rational expectations nor do we have to rely on estimation methods which use unobserved variables, see Hodrick and Srivastava (1984, 1986). **Conditional** on the hypothesis that the foreign market is efficient or rational, the modelling of time-varying risk premia has been explored by Domowitz and Hakkio (1985), Diebold and Pauly (1988a), and Kendall and McDonald (1989), for instance. Based on the utility optimizing models of Lucas (1982), Domowitz and Hakkio (1985) present an intertemporal asset pricing model in which the risk premium is a function of the conditional variances of the domestic and foreign money supplies⁸. The methodologies used in these papers usually involve measurement of time-varying risk premia conditional on the hypothesis that exchange rate forecasts are rational. Conclusions about the behavior of premia in the pricing of forward foreign exchange are drawn on the basis of rationality on the part of economic agents. Since the results of previous research favor the conclusion that economic agents exhibit irrational behavior⁹, we propose an alternative approach to measure premia.

This approach implements the use of survey data on exchange rate expectations. The survey data allows the direct measurement of risk premia from the decomposition of the forward discount in 3.1, thereby avoiding reliance on inferences from a hypothesis about expectations. In addition, several currencies display significant ARCH effects. However, as we do not specify a general equilibrium model we do not know the true structure of the covariance matrix and to what variables it is related. A (G)ARCH model is an acceptable alternative because it can be interpreted as a reduced form of a more complicated dynamic structure for the time-varying conditional second order moments. The ARCH-in-mean model developed by Engle, Lilien and Robins (1987) which we propose, can be used in addressing questions regarding the risk-return tradeoff in a time series context. The ARCH-in-

⁸ In particular, the intertemporal capital asset pricing model of Lucas (1982) implies that the risk premium is determined by the conditional covariance between the return on a long position in the forward market and the marginal rate of substitution between current and future consumption.

⁹ See Frankel and Froot (1987a,b) and Cavaglia, Verschoor and Wolff (1991) for instance.

mean model extends the ARCH model to allow the conditional variance to affect the conditional mean directly and is given by:

$$E_t S_{t+k} - S_t = RP_t^k + \beta_1 (F_{t,t+k} - S_t) + \varepsilon_t \quad 5.1$$

$$RP_t^k = \beta_0 + \theta h_t^{1/2} \quad 5.2$$

$$\varepsilon_t | I_{t-1} \sim N(0, h_t) \quad 5.3$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 \quad 5.4$$

where I_{t-1} represents the information set at time $t-1$. The risk premium, RP_t^k , depends directly on the conditional variance of ε_t , that is denoted h_t . The conditional variance of the expected rate of depreciation given time t information is postulated to depend on the realizations of the squared error terms in the previous months. A generalization proposed by Bollerslev (1986) is the GARCH model. For the first order GARCH-in-mean model the conditional variance becomes:

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 h_{t-1} \quad 5.5$$

The degree of persistence in variance is determined by the magnitude of the parameters of the conditional variance equations 5.4 and 5.5 and non-negativity constraints are imposed on these parameters. We restrict our attention to a GARCH (1,1) specification since it has been shown to be a parsimonious representation of conditional variance that adequately fits many economic time series - see Bollerslev (1987). The estimation of the econometric model described above using maximum likelihood methods is not as straightforward as may seem at first sight.

First of all we assume conditional normality without knowing what the true conditional distribution really is. This leaves the model subject to distributional misspecification. Weiss (1986) has shown that the quasi maximum likelihood parameter estimates are still consistent and asymptotically normal but with a modified asymptotic covariance matrix, which is robust to departures from normality. This all under the condition that the model is correctly specified, see also Bollerslev and Wooldridge (1989). Furthermore the information matrix is not block diagonal in the present framework, as a function of the conditional variance enters the mean equation. This means that we cannot use the scoring algorithm with analytical derivatives to obtain the maximum likelihood estimates. Instead we have to rely on numerical maximization of the likelihood function. Finally any misspecification in the conditional variance equation leads to biased estimates in the conditional mean equation, because the parameter estimates in the conditional mean equation are not asymptotically independent from the parameter estimates in the conditional variance equation. Diagnostic tests for the conditional variance equation are imperative upon interpretations of the parameter estimates.

We choose the Berndt, Hall, Hall and Hausman (1974) algorithm (BHHH) with numerical derivatives to obtain maximum likelihood estimates. The robust covariance matrix is calculated by pre- and post-multiplying the inverse of the BHHH covariance matrix by the inverse of the estimated information matrix. The estimation results and diagnostics of both ARCH and GARCH-in-mean models are presented in the next section.

6.6 Empirical Results

Maximum likelihood estimates of the parameters and their heteroskedasticity consistent asymptotic standard errors were obtained by numerical methods using the BHHH algorithm¹⁰. In Tables 6.4 and 6.5, the estimation results are reported for the two stochastic process described in the previous section. Table 6.4 reports the ARCH-in-mean results for each currency and for each

¹⁰ All calculations were performed with the software package GAUSS.

Table 6.4 : ARCH-in Mean Models, January 1, 1986 through September 1, 1991: 69 Observations

	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\theta}$	$\hat{\alpha}_0 \cdot 10^2$	$\hat{\alpha}_1$	L.L.
BF/DM	-0.0128 (0.0151)	0.2632*** (0.1003)	1.3091 (1.4563)	0.0067*** (0.0019)	0.5003** (0.2517)	218.54
DG/DM	-0.0018 (0.0365)	0.8339 (0.5004)	0.3735 (4.3029)	0.0069** (0.0024)	0.0006*** (1.1530)	230.76
FF/DM	0.0125*** (0.0045)	0.7731 (0.1581)	-2.8027*** (-1.0187)	0.0023*** (0.0005)	0.0515 (0.0376)	268.48
IL/DM	-0.0176 (0.0159)	0.8382 (0.8929)	0.5236* (0.3124)	0.0343 (0.0244)	1.2005 (1.1738)	159.19
SP/DM	-0.0146 (0.0136)	0.4874*** (0.1730)	1.3919* (0.8132)	0.0149*** (0.0052)	0.1423 (0.1242)	201.47
6 months						
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\theta}$	$\hat{\alpha}_0 \cdot 10^2$	$\hat{\alpha}_1$	L.L.
BF/DM	-0.0022 (0.0025)	0.5399* (0.2780)	0.1091 (0.2699)	0.0061** (0.0025)	0.6493 (0.6348)	217.83
DG/DM	0.0101 (0.0149)	0.7170 (0.2251)	-2.0323 (2.4406)	0.0036*** (0.0011)	0.1341 (0.3673)	251.15
FF/DM	0.0053 (0.0503)	0.9482 (0.2061)	-1.3584 (6.0928)	0.0065*** (0.0021)	0.0950 (0.2568)	231.13
IL/DM	0.0001 (0.088)	0.3432*** (0.1822)	0.4893* (0.2444)	0.0088*** (0.0018)	1.1044*** (0.2727)	182.84
SP/DM	0.0085 (0.0074)	0.2105*** (0.1333)	0.2588 (0.3319)	0.0155** (0.0081)	0.8533 (0.6892)	179.06
12 months						
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\theta}$	$\hat{\alpha}_0 \cdot 10^2$	$\hat{\alpha}_1$	L.L.
BF/DM	-0.0076 (0.0058)	0.0921*** (0.1101)	0.8288*** (0.2939)	0.0080 (0.0039)	0.7280** (0.3534)	203.24
DG/DM	0.0067*** (0.0013)	-0.0349*** (0.0915)	0.0220 (0.1232)	0.0040*** (0.0014)	1.0041** (0.4549)	222.25
FF/DM	-0.0266*** (0.0070)	0.5944*** (0.0754)	2.1682*** (0.8023)	0.0092*** (0.0017)	0.0646 (0.0603)	220.24
IL/DM	0.0079 (0.0096)	0.1382*** (0.0992)	0.2918*** (0.0929)	0.0098*** (0.0034)	1.3161*** (0.3596)	165.88
SP/DM	-0.0171 (0.0211)	0.2195*** (0.0618)	1.5984* (1.1739)	0.0335*** (0.0071)	0.0043*** (0.0469)	177.98

The heteroskedasticity consistent standard errors of the coefficients are given in parentheses; * (**) [***] denotes significance at the 10% (5%) [1%] level for the hypotheses $\beta_0=0$, $\beta_1=1$, $\theta=0$, $\alpha_0=0$ or $\alpha_1=0$ respectively. L.L. denotes the log-likelihood values.

Table 6.5 : GARCH-in-Mean Models, January 1, 1986 through September 1, 1991: 69 Observations

	3 months						
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\theta}$	$\hat{\alpha}_0(.10^2)$	$\hat{\alpha}_1$	$\hat{\gamma}_1$	L.L.
BF/DM	-0.0135 (0.0051)	0.2122*** (0.1064)	1.3406*** (0.5486)	0.0042 (0.0032)	0.4337 (0.5036)	0.2457 (0.3598)	219.30
DG/DM	0.0074** (0.0032)	0.9548 (0.2771)	-0.7371* (0.3921)	0.0003 (0.0003)	3.3E-9** (1.3E-9)	0.9499*** (0.0507)	232.46
FF/DM	0.0020 (0.0066)	0.7577 (0.1673)	-0.6475 (1.4011)	0.0023*** (0.0005)	0.0790 (0.0700)	7.2E-12 (1.4E-10)	268.58
IL/DM	-0.0222 (0.0063)	2.4866*** (0.4408)	-0.8173*** (0.2054)	0.0018 (0.0013)	0.7323*** (0.2003)	0.4779*** (0.1323)	181.67
SP/DM	-0.0090** (0.0037)	0.2819*** (0.1798)	1.4544*** (0.2920)	0.0005* (0.0003)	1.8E-13 (2.2E-12)	0.9233*** (0.0275)	211.45
	6 months						
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\theta}$	$\hat{\alpha}_0(.10^2)$	$\hat{\alpha}_1$	$\hat{\gamma}_1$	L.L.
BF/DM	-0.0177*** (0.0025)	0.0810*** (0.0371)	2.2182*** (0.2991)	0.0011 (0.0007)	0.1588** (0.0714)	0.6898*** (0.0566)	228.25
DG/DM	0.0083*** (0.0027)	0.7614* (0.1445)	-1.7600*** (0.5177)	0.0008 (0.0005)	0.1693* (0.1032)	0.6315*** (0.1852)	253.87
FF/DM	0.0065 (0.0189)	0.9514 (0.1770)	-1.4950 (2.3751)	0.0060** (0.0027)	0.0899 (0.1302)	0.0753 (0.2780)	231.15
IL/DM	0.0031 (0.0130)	0.3781*** (0.2128)	0.2564 (0.5179)	0.0003 (0.0010)	0.4060*** (0.1246)	0.7236*** (0.0684)	186.74
SP/DM	0.0085 (0.0085)	0.2105*** (0.1383)	0.2587 (0.3474)	0.0155** (0.0073)	0.8536 (0.7672)	0.0000 (0.0885)	179.06
	12 months						
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\theta}$	$\hat{\alpha}_0(.10^2)$	$\hat{\alpha}_1$	$\hat{\gamma}_1$	L.L.
BF/DM	-0.0076 (0.0058)	0.0920*** (0.1101)	0.8288*** (0.2940)	0.0080** (0.0039)	0.7281** (0.3534)	0.0000 (0.0000)	203.24
DG/DM	0.0068 (0.0071)	0.0395*** (0.2347)	-0.7682 (0.9379)	0.0019 (0.0013)	0.4264 (0.6647)	0.4798 (0.4368)	224.15
FF/DM	-0.0204** (0.0114)	0.5969*** (0.1324)	1.6435 (1.2154)	0.0041 (0.0097)	0.0136 (0.3429)	0.5150 (1.3983)	221.84
IL/DM	0.0073 (0.0076)	0.1542*** (0.0873)	0.2353 (0.1705)	0.0038 (0.0026)	0.7959*** (0.2534)	0.3568*** (0.1520)	167.69
SP/DM	0.0215*** (0.0044)	0.2653*** (0.0655)	-0.6997*** (0.2117)	2.2E-10 (2.2E-10)	6.6E-9 (5.3E-9)	0.9863*** (0.0051)	180.19

The heteroskedasticity consistent standard errors of the coefficients are given in parentheses; * (**) [***] denotes significance at the 10% (5%) [1%] level for the hypotheses $\beta_0=0$, $\beta_1=1$, $\theta=0$, $\alpha_0=0$, $\alpha_1=0$ or $\gamma_1=0$ respectively. L.L. denotes the log-likelihood values.

forecast horizon ($k = 3, 6$, and 12 months). Rejection of the hypothesis $\beta_1=1$ (no time-varying risk premia) was obtained in ten out of fifteen cases, thus corroborating the results of Table 6.2, which demonstrated significant time-varying risk premia. In a number of cases, the results provide evidence of both θ and α_1 being insignificantly different from zero. Rejection of the hypothesis $\theta=0$ was obtained in seven out of fifteen cases, whereas rejection of $\alpha_1=0$ was obtained in five cases, suggesting that time-varying risk premia might be explained by the conditional standard deviation of the expected rate of depreciation. The Italian Lira/ German Mark exchange rate appears to be integrated-in-variance, see Engle and Bollerslev (1986), a condition analogous to a unit root in conditional mean. At the $3, 6$, and 12 month forecast horizon, the estimated α_1 coefficient is greater than one which implies that the unconditional distribution of the expected depreciation is extremely fat tailed with an infinite variance, see Table 6.1.a¹¹.

The GARCH-in-mean estimation results for each currency and for each forecast horizon are reported in Table 6.5. It is interesting to note that the generalized specification of the conditional variance significantly improves the fit of the model. The results provide a fairly consistent rejection of the hypothesis $\beta_1=1$ (no time-varying risk premia), suggesting significant variation in the risk premium. The estimated α_1 and γ_1 coefficients are statistically significant in eight out of fifteen cases, thus supporting the GARCH specification. Moreover, all but seven of the estimated models result in statistically significant θ coefficients, suggesting that premia for EMS exchange rates relative to the German Mark follow GARCH (1,1) processes in a number of cases. In the case of the Italian Lire / German Mark exchange rate, the coefficient estimates of $\alpha_1 + \gamma_1$ are greater than one, indicating high persistence in the volatility shocks, or IGARCH behavior. As conjectured by Diebold (1986), and Lamoureux and Lastrapes (1990), this may be the result of shifts in monetary regimes which affect the level of the unconditional variances¹².

¹¹ The conditional distribution, which for most purposes is the relevant distribution, is of course still normal with a finite variance.

¹² Lastrapes (1989) finds that persistence of exchange rate volatility decreases when regime shifts are accounted for, diminishing the probability of finding integrated-in-variance processes.

In Graphs 1.a - 2.b, two examples of the actual (P_t^k from the decomposition of the forward discount in 3.1) and estimated premia of both ARCH and GARCH-in-mean models are presented (the DG/DM and BF/DM exchange rate with 6 months forecast horizon). Graphs 3.a and 3.b display corresponding figures of the ARCH and GARCH premium estimates. The figures indicate substantial volatility in the actual exchange risk premia. The GARCH-in-mean premium estimates indicate substantial movement of the conditional variance of the expected rate of depreciation¹³. These figures are representative for exchange rates and forecast horizons with significant (G)ARCH and significant risk premium parameters.

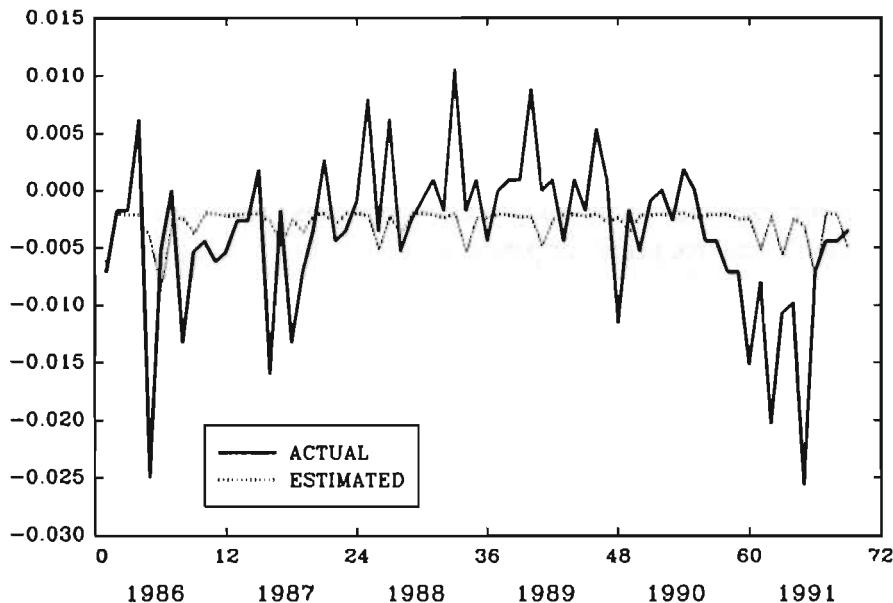
Given these results, it is interesting to compare the relative fit of both models. We employ generalized likelihood ratio tests to compare nested models. This test statistic is based on the difference (Λ) in the log likelihood values of the simple and the elaborate version of two nested models. Under the null, the statistic -2Λ has a χ^2 distribution with degrees of freedom equal to the difference in the number of parameters between the two models, which in this case equals one. Table 6.6 presents the generalized likelihood ratio tests to compare the relative fit of the two models. All but five of the p-values associated with the χ^2 statistics are close to zero. Thus, the generalized likelihood ratio tests reject the simpler (ARCH-in-mean) model in favour of the more complicated (GARCH-in-mean) model in ten out of fifteen cases. In the case of the BF/DM exchange rate at the 12 month horizon and the SP/DM exchange rate at the 6 month horizon, the p-values associated with the chi-square statistics are equal to one, thus strongly supporting the ARCH specification.

In order to determine the adequacy of the statistical specification, the models are subjected to diagnostic checks on the standardized residuals:

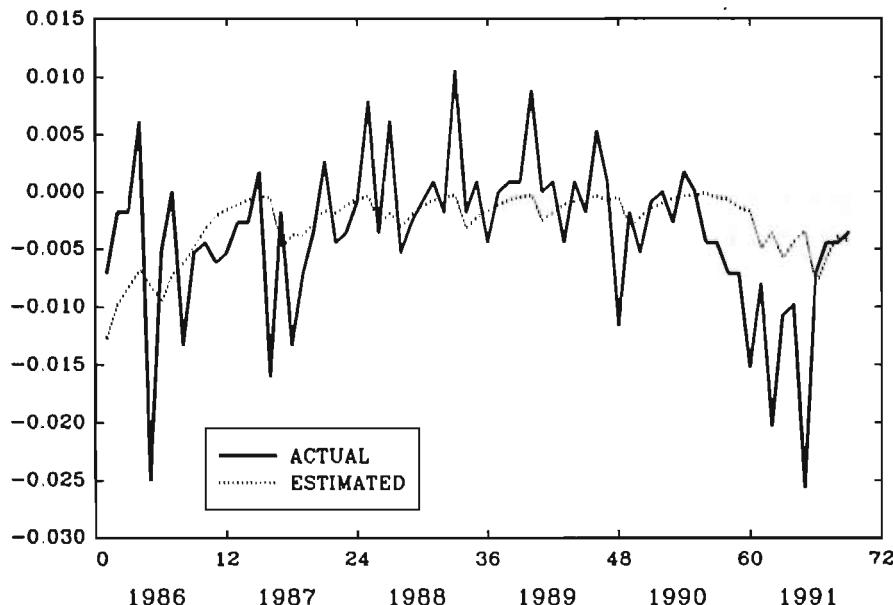
$$z_t = \hat{\epsilon}_t / \hat{h}_t \quad 6.1$$

¹³ If one would assume that our specification fully represents the true time varying risk premium, then the difference between the estimated premium (which now is assumed to be the true premium) and the observed premium can be interpreted as the measurement error on the exchange rate expectations, which would be quite substantial as can be seen from Graphs 1.a - 2.b.

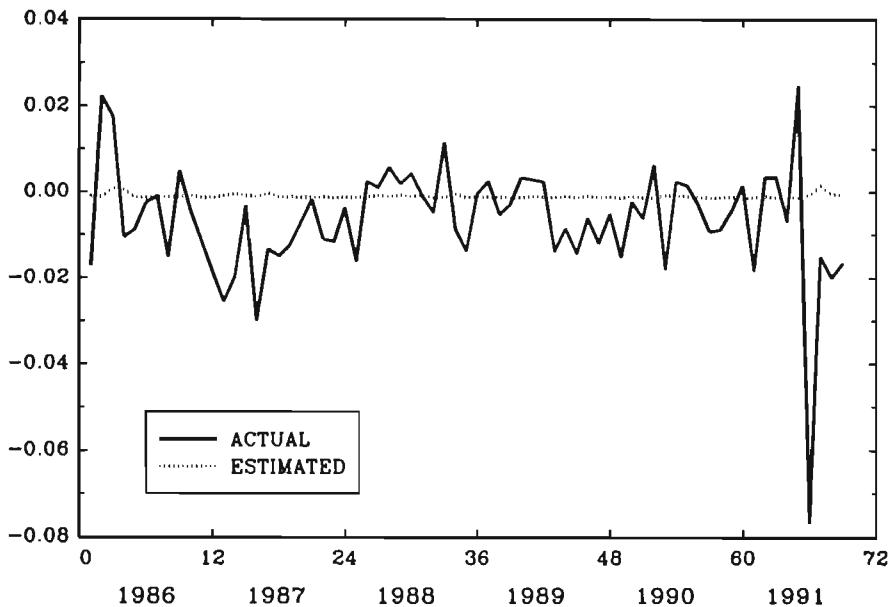
Graph 1a: Actual and Estimated ARCH Premia DG/DM
6 Months Forecasts



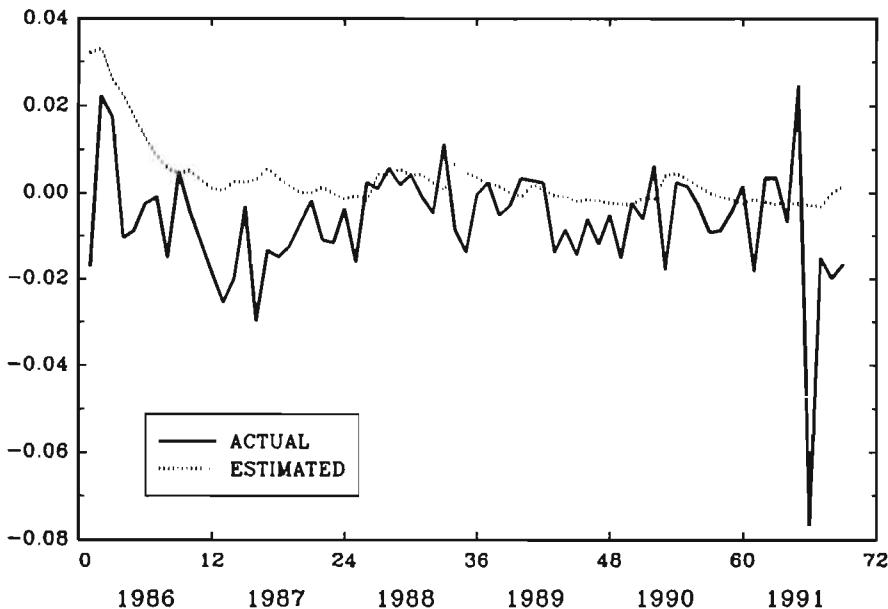
Graph 1b: Actual and Estimated GARCH Premia DG/DM
6 Months Forecasts



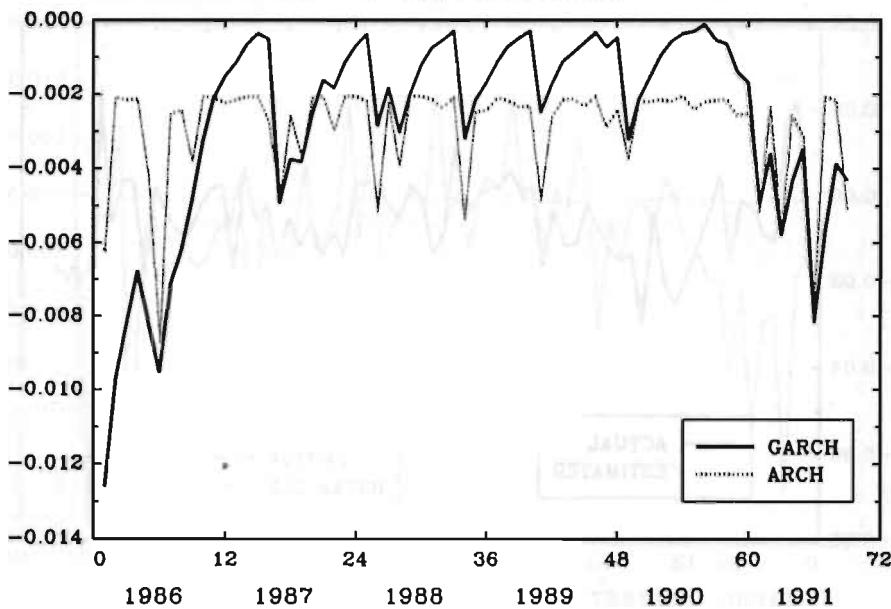
Graph 2a: Actual and Estimated ARCH Premia BF/DM
6 Months Forecasts



Graph 2b: Actual and Estimated GARCH Premia BF/DM
6 Months Forecasts



**Graph 3a: Estimated ARCH and GARCH Premia DG/DM
6 Months Forecasts**



**Graph 3b: Estimated ARCH and GARCH Premia BF/DM
6 Months Forecasts**

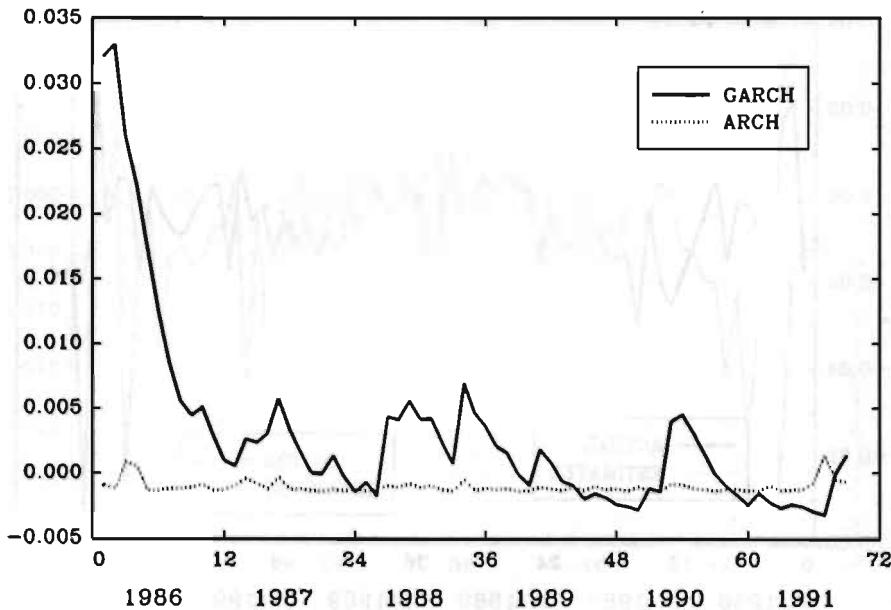


Table 6.6 : Generalized Likelihood Ratio Tests: ARCH-in-mean model against GARCH-in-mean model, January 1, 1986 through September 1, 1991: 69 Observations

	3 months	6 months	12 months
BF/DM	1.52 (0.218)	20.84*** (0.000)	0.00 (1.000)
DG/DM	3.40* (0.065)	5.44** (0.020)	3.80* (0.051)
FF/DM	0.20 (0.655)	0.04 (0.841)	3.20* (0.074)
IL/DM	44.96*** (0.000)	7.80*** (0.005)	3.62* (0.057)
SP/DM	19.96*** (0.000)	0.00 (1.000)	4.42** (0.035)

P-Values are given in parentheses; * (**) [***] denotes rejection at the 10% (5%) [1%] level.

where $\hat{\epsilon}_t$ is the residual from equation 5.1 and \hat{h}_t is the estimated conditional variance from equations 5.4 and 5.8. From Jensen's inequality it follows that the standardized residuals, z_t , should demonstrate less absolute skewness and should be thinner tailed than their unconditional raw data counterparts. Any strong violation of this rule should be regarded as evidence of model misspecification, see Hsieh (1989). The diagnostics for ARCH and GARCH-in-mean models are presented in Tables 6.7 and 6.8. Overall, the evidence presented suggest a less consistent rejection of the normality hypotheses as compared with the results of Tables 6.1a, 6.1.b, and 6.1.c. For the standardized ARCH-in-mean residuals, rejection occurs in ten out of fifteen cases, whereas for the standardized GARCH-in-mean residuals, rejection was obtained in only seven cases. In addition, we find that in most cases the estimated statistics - the BJ-test, KS-1, and KS-2 - are smaller than those reported by Tables 6.1.a, 6.1.b, and 6.1.c, thus supporting our model specifications. In particular, the GARCH-in-mean model is quite successful in removing excess kurtosis and skewness in a number of cases.

Table 6.7 : Diagnostics for ARCH-in-Mean Models, January 1, 1986 through September 1, 1991: 69 Observations

	BF/DM	DG/DM	FF/DM	IL/DM	SP/DM
3 months					
Skewness	1.15	1.31	-0.21	-3.24	0.29
Kurtosis	4.90	7.10	4.05	16.58	6.50
BJ-test	25.45***	67.87***	3.66	650.69***	36.17***
KS-1	15.20***	19.52***	0.49	120.69***	0.99
KS-2	10.35***	48.34***	3.17*	530.29***	35.17***
LM(1)	0.79	0.95	0.01	1.14	0.72
LM(5)	7.81	2.94	3.88	8.40	27.37***
6 months					
Skewness	0.23	-0.66	-0.06	-0.17	0.60
Kurtosis	4.50	3.86	3.08	3.27	3.89
BJ-test	7.06**	7.16**	0.06	0.53	6.47***
KS-1	0.61	5.01**	0.04	0.32	4.19**
KS-2	6.46**	2.16	0.02	0.20	2.29
LM(1)	0.07	0.52	0.10	3.24	0.17
LM(5)	1.04	7.44	3.06	8.21	4.33
12 months					
Skewness	0.41	0.32	-0.59	0.00	0.47
Kurtosis	3.73	4.12	3.80	3.39	3.47
BJ-test	3.49	4.83*	5.89*	0.45	3.19
KS-1	1.92	1.18	4.07*	0.00	2.54
KS-2	1.53	3.64*	1.82	0.47	0.64
LM(1)	0.34	2.08	2.02	3.06	0.16
LM(5)	3.01	5.47	3.02	4.63	6.85

The BJ-test denotes the Bera Jarque test for normality; KS-1 and KS-2 pertain to the Kiefer Salmon Normality test for respectively skewness and kurtosis: * (**) [***] denotes rejection at the 10% (5%) [1%] level of the normality hypotheses.

Table 6.8 : Diagnostics for GARCH-in-Mean Models, January 1, 1986 through September 1, 1991: 69 Observations

	BF/DM	DG/DM	FF/DM	IL/DM	SP/DM
3 months					
Skewness	1.27	1.23	-0.16	-0.23	-0.21
Kurtosis	5.49	7.17	4.00	7.11	3.02
BJ-test	36.24 ***	67.55 ***	3.18	49.24 ***	0.50
KS-1	18.40 ***	17.45 ***	0.30	0.62	0.50
KS-2	17.83 ***	50.09 ***	2.88 *	48.62 ***	0.00
LM(1)	0.82	1.04	0.11	0.94	1.16
LM(5)	7.68	2.91	5.47	0.96	5.18
6 months					
Skewness	0.20	-0.47	-0.07	-0.24	0.60
Kurtosis	2.84	3.97	3.09	3.47	3.89
BJ-test	0.54	5.20 *	0.08	1.30	6.47 **
KS-1	0.47	2.52	0.06	0.66	4.19 **
KS-2	0.07	2.68	0.03	0.64	2.28
LM(1)	0.08	1.81	0.11	1.57	0.17
LM(5)	2.18	5.06	3.12	6.24	4.33
12 months					
Skewness	0.41	0.42	-0.43	-0.05	0.26
Kurtosis	3.73	4.95	3.01	2.80	3.06
BJ-test	3.46	12.97 *	2.13	0.15	0.79
KS-1	1.92	1.98	2.12	0.03	0.77
KS-2	1.54	10.99 ***	0.01	0.11	0.01
LM(1)	0.34	1.35	0.95	0.79	0.26
LM(5)	3.01	4.58	2.70	4.75	6.37

The BJ-test denotes the Bera Jarque test for normality; KS-1 and KS-2 pertain to the Kiefer Salmon Normality test for respectively skewness and kurtosis; * (**) [***] denotes rejection at the 10% (5%) [1%] level of the normality hypotheses.

In order to test for remaining heteroskedasticity, a residual-based test of the models may be carried out by regressing $(\hat{\epsilon}_t^2 - \hat{h}_t^2)/\hat{h}_t^2$ on $1/\hat{h}_t^2$ and on one to five lags of the dependent variable. The results are reported under LM(1) and LM(5), and are Chi-square distributed with one and five degrees of freedom. For the ARCH-in-mean models, rejection of the null hypothesis of no heteroskedasticity occurs in only one case (the SP/DM exchange rate), whereas the GARCH-in-mean models all result in statistically insignificant test statistics.

6.7 Conclusions

This paper has examined exchange risk premia using survey data for a set of EMS exchange rates relative to the Deutschmark over the 1985-1991 period. The methodologies used in previous empirical research on premia in the pricing of forward foreign exchange usually involve measurement of time-varying risk premia **conditional** on market efficiency or rational expectations. We implemented an alternative approach to measure premia. The approach involves application of survey data to allow the direct measurement of risk premia from the forward discount decomposition into its two components - expected depreciation and the risk premium. We extended the analysis of Domowitz and Hakkio (1985) to model time-varying risk premia in the pricing of forward foreign exchange that do **not** require us to assume rationality on the part of economic agents. We find considerable support for the presence of time-varying risk premia in the pricing of forward foreign exchange. The estimated premium models, ARCH-in-mean and GARCH-in-mean, indicate that time-varying premia can be explained by the conditional standard deviation of the expected rate of depreciation. In particular, the GARCH-in-mean model appears to be reasonable successful in accounting for both time-varying risk premia and conditional heteroskedasticity. The results basically contrast with the results of Domowitz and Hakkio (1985), who found only minimal support for the ARCH-in-mean model for some of the major currencies (British Pound, French Franc, German Mark, Japanese Yen, and Swiss Franc) relative to the United States Dollar covering the period 1973-1982. The evidence presented in this paper is sufficiently strong to suggest further research as to whether foreign exchange risk can success-

fully be approximated by a measure of exchange rate variability. In addition, cross sectional dispersion in survey data might be considered as a measure of uncertainty. The dispersion in exchange rate forecasts provides an indicator of the heterogeneity in exchange rate expectations, and may be an approximation of the fundamental underlying uncertainty. This issue requires further investigation in future work.

Chapter 7

Summary and Discussion

7.1 Summary and Discussion

The focus of this dissertation has been on the time-series properties of certain financial assets, which, until recently, has been a neglected area in Financial Economics, see Gibbons (1987). We have started from the notion that conditional moments emerging from an asset's conditional distribution are time-varying. The pricing of financial assets obviously has to be such that these changing moments are accounted for. One immediate problem that arises then is that although much of modern finance theory is set in continuous time, financial time series are only available at discrete time intervals. This implies that in an empirical assessment of time-varying conditional moments, one has to resort to a discrete time setting. In this dissertation we have subjected the most popular approach to describe time-varying conditional second moments in discrete time, the GARCH approach, to close scrutiny. We have presented substantial evidence in favor of the use of the GARCH process as a econometric tool to describe conditional heteroskedasticity in the stock and foreign exchange market. The second major objective of this dissertation was concerned with the use of a GARCH process as an economic tool in the determination of time-varying risk premia. The evidence regarding this issue is less clear-cut. The major results of the individual chapters are given below.

Chapters 2, 3, and 4 are concerned with the stock market and Chapters 5 and 6 with the foreign exchange market. Chapters 2 and 5 discuss the use of GARCH models as an econometric tool, whereas the remaining chapters discuss the usefulness of GARCH as an economic tool.

In Chapter 2, a GARCH(1,1) model is used, under various distributional assumptions to describe the evolution in the conditional variance for a broad set of returns of weekly and monthly stock indices. The preliminary analysis showed that the usual Lagrange Multiplier test to detect (G)ARCH effects has to be doubted when the underlying distribution is non-normal. A test for conditional heteroskedasticity that is based on homogeneous Markov Chains appears to be an appropriate alternative. The estimation results reveal that for both weekly as well as monthly returns, GARCH effects are clearly present. The choice of the conditional distribution is limited to three alternatives. We employed a conditional Normal, a conditional Student t-distribution, and a Generalized Error or Power Exponential distribution. Generalized Likelihood Ratio tests clearly point in one direction. The Student t-distribution is the preferred alternative both for weekly as well as monthly returns. Conditional normality can be dismissed altogether. Furthermore we conclude that aggregation from a weekly to a monthly level is insufficient to cause conditional heteroskedasticity of the GARCH form to disappear. Our monthly estimates are broadly consistent with the estimates implied by an aggregation theorem developed by Drost and Nijman (1991). Finally the unconditional distribution of rescaled GARCH residuals is definitely non-normal, and the evidence presented by a tail-index analysis, see Koedijk, Schafgans and de Vries (1990), suggests an unconditional Student t-distribution with few degrees of freedom.

Chapter 3 considers the implementation of a GARCH model for the conditional variance of the market return into an intertemporal asset pricing model for stock returns. In the initial analysis of this chapter we have derived an asset pricing model containing both the traditional CAPM as well as Consumption CAPM elements, on the basis of Generalized Isoelastic Preferences. In order to arrive at an "assets-only" model we had to eliminate consumption growth from the model. In a heteroskedastic world this is not at all trivial. It was assumed that the intercept term in the linear relation between expected consumption growth and the expected market return depends

linearly on the conditional variance of the market return, which is modelled by a GARCH process. An intermediate result in the analysis was that modelling the conditional variance as a GARCH process does not result in the familiar GARCH-in-mean model. The resulting asset pricing model is reminiscent of a three moment CAPM, and includes a time-varying conditional third moment. This indicates that besides a risk-return tradeoff, we should also be concerned with a risk-skewness tradeoff.

In Chapter 4 we were not only concerned with time-varying second moments, but also with time-varying third moments. This concern was inspired by the results of Chapter 3. The market version of the model developed in that chapter was subjected to an empirical analysis, using the same dataset as in Chapter 2. In a calibration exercise we have shown that our model is capable of generating plausible values for the parameters of interest. Subsequently we have introduced the Generalized AutoRegressive Conditional Skewness (GARCS) process, in order to model time-varying conditional third moments. The full model was estimated by GMM as we did not wish to make any distributional assumptions, and we compared the results to the results obtained from a GARCH-in-mean model. Our version of the three-moment CAPM generates significant estimates for the coefficient of relative risk aversion, whereas the GARCH-in-mean model does not. This can be explained by the lack of intertemporal components in the GARCH-in-mean model. We have also found a high degree of correlation between the conditional variance and the conditional third moment, which is a common problem in the existing literature on the three moment CAPM. In Chapters 3 and 4 we have shown that incorporating GARCH into an asset pricing model is not at all as straightforward as the GARCH-in-mean model led us to believe, and the issues discussed in these chapters provide sufficient incentives for future research.

The statistical properties of weekly exchange rate returns within the Exchange Rate Mechanism of the EMS were analyzed in Chapter 5. Our preliminary analysis shows that these exchange rate returns are characterized by highly leptokurtic distributions, and high degree of mean reversion between parity adjustments. Without parity adjustments the martingale hypothesis could not be rejected. Again conditional heteroskedasticity is clearly present, but besides the risk factor represented by the conditional

variance of ordinary diffusion another risk factor is present, namely realignment risk. In our analysis this is accounted for by an independent Poisson Jump process. More specifically, the model that is most successful in capturing the relevant features of EMS exchange rate returns, and in generating the most plausible values for the jump parameters, is a combined Jump-GARCH model with conditionally Student t-distributed innovations.

Chapter 6, as Chapters 3 and 4 before, addressed the issue of time-varying risk premia for financial asset returns, and considered the conditional variance of foreign exchange rate returns, given by a GARCH process, as a determinant of the risk premium in the forward foreign exchange market. The rejection of the forward rate as an unbiased predictor of the expected future spot rate can be attributed to the irrationality of market participants or the existence of a time-varying risk premium. The use of a survey dataset in which direct expectations regarding future exchange rates were formed, relieves us from the assumption of rational expectations. Many theoretical models generate a risk premium that depends on the conditional probability distribution of the future spot rate. This provides the rationale for a consideration of the conditional variance of the exchange rate return as a determinant of a time-varying risk premium. Contrary to our findings in Chapter 4 for the stock market, the estimation results in this chapter indicate that the GARCH-in-mean model is reasonably successful in describing a time-varying risk premium.

As a general remark, we should mention that all of our analyses could also be cast in a multivariate framework. This dissertation has focused on a univariate "time-series" approach, while many interesting issues in asset pricing theory require a multivariate or "cross-sectional" approach. This is particularly relevant for the analyses in Chapters 4 and 6. The asset pricing model that was developed in chapter 3 should be examined for a cross-section of individual assets, in order to provide further evidence regarding its strengths and weaknesses. The determination of the risk premia in the forward foreign exchange market could be cast in a multivariate GARCH setting, where a system of exchange rates can be analyzed simultaneously. For such a multivariate GARCH approach, with high frequency and symmetrically distributed data, the use of a conditional multivariate Student t-distribution should be seriously considered, see chapter 2.

Our analyses in Chapters 5 and 6 could be combined to assess the impact of jumps on the risk premium in the forward foreign exchange market. Alternatively the Jump-GARCH models presented in Chapter 5 could be extended to allow for a jump intensity parameter that depends on the position of the exchange rate within the band. Finally it should be possible to check whether our evidence in favor of the GARCH-in-mean model in Chapter 6, where GARCH is treated as a reduced form of a more complicated dynamic structure for time-varying conditional second order moments, still holds up if we would incorporate GARCH directly into an asset pricing model, as was the case in Chapter 3.

References

- Adler, Michael and Bernard Dumas, 1983, "International Portfolio Choice and Corporation Finance: A Synthesis", *Journal of Finance*, 38: 925-84.
- Akgiray, Vedat and G. Geoffrey Booth, 1988a, "Mixed Diffusion-Jump Process Modelling of Exchange Rate Movements", *Review of Economics and Statistics*, 70: 631-37.
- Akgiray, Vedat and G. Geoffrey Booth, 1988b, "The Stable Law Model of Stock Returns", *Journal of Business and Economic Statistics*, 6: 51-57.
- Akgiray, Vedat, 1989, "Conditional Heteroskedasticity in Time Series of Stock Returns: Evidence and Forecasts", *Journal of Business*, 62: 55-80.
- Arditti, Fred D., 1967, "Risk and the Required Return on Equity", *the Journal of Finance*, 22: 19-36.
- Arditti, Fred D., 1971, "Another Look at Mutual Fund Performance", *Journal of Financial and Quantitative Analysis*, 6: 909-912.
- Arditti, Fred D., and Haim Levy, 1975, "Portfolio Efficiency Analysis in Three Moments: the Multiperiod Case", *The Journal of Finance*, 30 : 797-809.
- Attanasio, Orazio P., 1991, "Risk, Time Varying Second Moments and Market Efficiency", *Review of Economic Studies*, 58: 479-494.
- Bachelier, Louis, 1900, "Theorie de la Speculation", Paris, Gaulthier Villars, reprinted in P. Cootner ed., 1964, "The Random Character of Stock Market Prices", Cambridge, MIT Press.
- Baillie, Richard T. and Tim Bollerslev, 1989, "The Message in Daily Exchange Rates : A Conditional Variance Tale", *Journal of Business and Economic Statistics*, 7: 297-305.
- Baillie, Richard T. and Ramon P. DeGennaro, 1990, "Stock Returns and Volatility", *Journal of Financial and Quantitative Analysis*, 25: 203-214.
- Ball, C. and W. Torous, 1985, "On Jumps in Common Stock Prices and Their Impact on Call Option Pricing", *Journal of Finance*, 40: 155-73.
- Barone-Adesi, Giovanni, 1985, "Arbitrage Equilibrium with Skewed Asset Returns", *Journal of Financial and Quantitative Analysis*, 20: 299-313.

- Beedles, William L., 1979, "Return, Dispersion and Skewness: Synthesis and Investment Strategy", *The Journal of Financial Research*, 2: 71-80.
- Bera, Anil K., and C.M. Jarque, 1982, "Model Specification Tests : A Simultaneous Approach", *Journal of Econometrics*, 20: 59-82.
- Berndt, Ernst K., Bronwyn H. Hall, Robert E. Hall and Jerry A. Hausman, 1974, "Estimation and Inference in Nonlinear Structural Models", *Annals of Economic and Social Measurement*, 4: 653-665.
- Bertola, Giuseppe, and Ricardo Caballero, 1989, "Target Zones and Realignments", Working Paper, Princeton University and Columbia University.
- Bertola, Giuseppe and Lars Svensson, 1990, "Stochastic Devaluation Risk and the Empirical Fit of Target Zone Models", Working Paper, Institute for International Economic Studies.
- Bilson, John, F.O., 1981, "The Speculative Efficiency Hypothesis", *Journal of Business*, 54: 435-51.
- Black, Fisher, and Myron S. Scholes, 1973, "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy*, 81: 637-54.
- Blattberg, Robert C. and Nicholas J. Gonedes, 1974, "A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices", *Journal of Business*, 47: 244-280.
- Bollerslev, Tim, 1986, "Generalized Autoregressive Conditional Heteroskedasticity", *Journal of Econometrics*, 31: 307-327.
- Bollerslev, Tim, 1987, "A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return", *Review of Economics and Statistics*, 69: 542-547.
- Bollerslev, Tim, Robert F. Engle and Jeffrey M. Wooldridge, 1988, "A Capital Asset Pricing Model with Time Varying Covariances", *Journal of Political Economy*, 96: 116-131.
- Bollerslev, Tim, and Jeffrey M. Wooldridge, 1989, "Quasi Maximum Likelihood Estimation of Dynamic Models with Time Varying Covariances", unpublished manuscript, Department of Economics, MIT.
- Bollerslev, Tim, Ray Y. Chou, and Kenneth F. Kroner, 1992, "ARCH Modeling in Finance : A Review of the Theory and Empirical Evidence", *Journal of Econometrics*, 52: 5-59.
- Boothe, Paul, and Debra Glassman, 1987, "The Statistical Distributions of Exchange Rates: Empirical Evidence and Economic Implications", *Journal of Financial Economics*, 22: 297-319.
- Box, George E.P. and George C. Tiao, 1973, *Bayesian Inference in Statistical Analysis*, Addison-Wesley, Reading.

- Breeden, Douglas T., 1979, "An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities", *Journal of Financial Economics*, 7: 265-96.
- Breeden, Douglas T., Michael R. Gibbons, and Robert H. Litzenberger, 1989, "Empirical Tests of the Consumption Oriented CAPM", *Journal of Finance*, 44: 231-262.
- Breusch, Thomas S. and Adrian R. Pagan, 1979, "A Simple Test for Heteroskedasticity and Random Coefficient Variation", *Econometrica*, 47: 1287-94.
- Campbell, John Y. and Robert J. Shiller, 1988, "The Dividend Price Ratio and Expectations of Future Dividends and Discount Factors", *Review of Financial Studies*, 1: 195-228.
- Campbell, John Y., and Ludger Hentschel, 1992, "No News is Good News : An Asymmetric Model of Changing Volatility in Stock Returns", *Journal of Financial Economics*, 31: 281-318.
- Campbell, John Y., 1990, "Intertemporal Asset Pricing without Consumption", Discussion Paper no. 107, London School of Economics, Financial Markets Group.
- Campbell, John Y., and Richard Clarida, 1987, "The Term Structure of Euromarket Interest Rates: An Empirical Investigation", *Journal of Monetary Economics*, 19: 25-44.
- Cavaglia, Stefano, Willem F.C. Verschoor and Christian C.P. Wolff, 1991, "On the Biasedness of Forward Foreign Exchange Rates: Irrationality or Risk Premia?", Working Paper, University of Limburg and OECD.
- Cavaglia, Stefano, Willem F.C. Verschoor and Christian C.P. Wolff, 1992, "Further Evidence on Exchange Rate Expectations", *Journal of International Money and Finance*, forthcoming.
- Chou, Ray Y., 1988, "Volatility Persistence and Stock Valuations : Some Empirical Evidence using GARCH", *Journal of Applied Econometrics*, 3: 279-294.
- Conine, Thomas E. Jr., and Maury J. Tamarkin, 1981, "On Diversification Given Asymmetry in Returns", *The Journal of Finance*, 36: 1143-55.
- Constantinides, George M., 1990, "Habit Formation, A Resolution of the Equity Premium Puzzle", *Journal of Political Economy*, 98: 519-543.
- Cumby, Robert, and Maurice Obstfeld, 1984, "International Interest Rate and Price Level Linkages under Flexible Exchange rates: A Review of Recent Evidence", in: *Exchange Rate Theory and Practice*, Chicago: Chicago University Press.
- Diebold, Francis X., and Peter Pauly, 1988a, "Endogenous Risk in a Portfolio Balance Rational Expectations Model of the Deutschmark-Dollar Rate", *European Economic Review*, 32: 27-53.

- Diebold, Francis X., and Peter Pauly, 1988b, "Has the EMS Reduced Member Country Exchange Rate Volatility?", *Empirical Economics*, 1988b, 13: 81-102.
- Diebold, Francis X., 1988, *Empirical Modeling of Exchange Rate Dynamics*, Springer Verlag, New York.
- Dickey, David A., and Wayne Fuller, 1979, "Distributions of the Estimators for Autoregressive Time Series With a Unit Root", *Journal of the American Statistical Association*, 74: 427-32.
- Dickey, David A., and Wayne Fuller, 1981, "Likelihood Ratio Statistics for Autoregressive Time Series With a Unit Root", *Econometrica*, 1981, 49: 1057-72.
- Dominguez, Kathryn M., 1986, "Are Foreign Forecast Rational? New Evidence from Survey Data", *Economics Letters*, 21: 277-281.
- Domowitz, Ian, and Craig Hakkio, 1985, "Conditional Variance and the Risk Premium in the Foreign Exchange Market", *Journal of International Economics*, 19: 47-66.
- Drost, Feike C., and Theo E. Nijman, 1991, "Temporal Aggregation of GARCH processes", Working Paper, Tilburg University.
- Dumas, Bernard, and Lars Svensson, 1991, "How Long Do Unilateral Target Zones Last?", Working Paper, Presented at the EEA Meeting 1991, Cambridge.
- Engle, Robert F., 1982, "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation", *Econometrica*, 50: 987-1007.
- Engle, Robert F., David F. Hendry and David Trumble, 1985, "Small Sample Properties of ARCH Estimators and Tests", *Canadian Journal of Economics*, 18: 66-93.
- Engle, Robert F., and Tim Bollerslev, 1986, "Modelling the Persistence of Conditional Variances", *Econometric Reviews*, 5: 1-50.
- Engle, Robert F., David M. Lilien and Russell P. Robins, 1987, "Estimating Time-Varying Risk Premia in the Term Structure: The ARCH-M Model", *Econometrica*, 55: 391-407.
- Engle, Robert F., and Gloria Gonzalez-Rivera, 1991, "Semiparametric ARCH Models", *Journal of Business and Economic Statistics*, 9: 345-360.
- Epstein, Larry G., 1988, "Risk Aversion and Asset Prices", *Journal of Monetary Economics*, 22: 179-192.
- Epstein, Larry G., and Stanley E. Zin, 1989, "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns : A Theoretical Framework", *Econometrica*, 57: 937-969.
- Epstein, Larry G., and Stanley E. Zin, 1991, "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns : An Empirical Analysis", *Journal of Political Economy*, 99: 263-286.

- Fama, Eugene F., 1965, "The Behavior of Stock Market Prices", *Journal of Business*, 38: 34-105.
- Fama, Eugene F., 1984, "Forward and Spot Exchange Rates", *Journal of Monetary Economics*, 14: 319-338.
- Fisher, Irving, 1930, *The Theory of Interest*, Augustus M. Kelley Publishers, New York.
- Flavin, Marjorie A., 1981, "The Adjustment of Consumption to Changing Expectations about Future Income", *Journal of Political Economy*, 89: 974-1009.
- Flood, Robert P. and Peter M. Garber, 1989, "The Linkage between Speculative Attack and Target Zone Models of Exchange Rates", NBER Working Paper #2918.
- Flood, Robert P., A.K. Rose, and D.J. Mathieson, 1991, "An Empirical Exploration of Exchange Rate Target-Zones", IMF Working Paper.
- Frankel, Jeffrey A., 1982, "In Search of the Exchange Risk Premium: A Six Currency Test Assuming Mean-Variance Optimization", *Journal of International Money and Finance*, 1: 255-74.
- Frankel, Jeffrey A. and Charles Engel, 1984, "Do Assets Demands Optimize over the Mean and Variance of Real Returns? A Six-Currency Test", *Journal of International Economics*, 17: 309-23.
- Frankel, Jeffrey A. and Kenneth A. Froot, 1987a, "Using Survey Data to Test Propositions Regarding Exchange Rate Expectations", *American Economic Review*, 77: 133-153.
- Frankel, Jeffrey A. and Kenneth A. Froot, 1987b, "Short-Term and Long-Term Expectations of the Yen/Dollar Exchange Rate: Evidence from Survey Data", *Journal of the Japanese and International Economies*, 1: 249-274.
- Frankel, Jeffrey A., and Stephen Phillips, 1991 "The European Monetary System: Credible at last?", NBER Working Paper # 3819.
- French, Kenneth R., G. William Schwert, and Robert F. Stambaugh, 1987, "Expected Stock Returns and Volatility", *Journal of Financial Economics*, 19: 3-29.
- Friend, Irwin, and Randolph Westerfield, 1980, "Co-Skewness and Capital Asset Pricing", *The Journal of Finance*, 35: 897-913.
- Friedman, Daniel and S. Vandersteel, 1982, "Short-run Fluctuations in Foreign Exchange Rates", *Journal of International Economics*, 13: 171-186.
- Froot, Kenneth A. and Jeffrey A. Frankel, 1989, "Forward Discount Bias: Is it an Exchange Risk Premium?", *Quarterly Journal of Economics*, 54: 139-61.
- Gallant, A. Ronald, and George Tauchen, 1989, "Semi Non-Parametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset pricing Applications", *Econometrica*, 57: 1091-1120.

- Giavazzi, Francesco, Stefano Micossi, and Marcus Miller, 1988, *The European Monetary System*, Cambridge, Cambridge University Press.
- Giavazzi, Francesco and Alberto Giovannini, 1989, *Limiting Exchange Rate Flexibility: The European Monetary System*, Cambridge, MIT Press.
- Gibbons, Michael R., 1982, "Multivariate Tests of Financial Models", *Journal of Financial Economics*, 10: 3-27.
- Gibbons, Michael R., 1987, "The Interrelations of Finance and Economics: Empirical Perspectives", *American Economic Review*, 77: 35-41.
- Giovannini, Alberto, and Philippe Jorion, 1987, "Interest Rates and Risk Premia in the Stock Market and in the Foreign Exchange Market", *Journal of International Money and Finance*, 6: 107-24.
- Giovannini, Alberto, 1990, "European Monetary Reform: Progress and Prospects", *Brookings Papers on Economic Activity*, 2: 217-91.
- Giovannini, Alberto and Philippe Weil, 1989, "Risk Aversion and Intertemporal Substitution in the Capital Asset Pricing Model", NBER Working Paper # 2824.
- Goldie, C.M, and R.L. Smith, 1987, "Slow Variation with Remainder: Theory and Applications", *Quarterly Journal of Mathematics*, 2: 45-71.
- Gourieroux, Christian, A. Holly, and Alain Montfort, 1982, "Likelihood Ratio Test, Wald Test, and Kuhn-Tucker Test in Linear Models with Inequality Constraints on the Regression Parameters", *Econometrica*, 50: 63-80.
- Gourieroux, Christian, and Alain Montfort, 1992, "Qualitative Threshold ARCH Models", *Journal of Econometrics*, 52: 159-199.
- Gregory, Allan W., 1989, "A Non-Parametric Test for Autoregressive Conditional Heteroskedasticity : A Markov-Chain Approach ", *Journal of Business and Economic Statistics*, 7: 107-115.
- Grossman, Sanford J., and Robert J. Shiller, 1982, "Consumption Correlatedness and Risk Measurement in Economies with Non-Traded Assets and Heterogeneous Information", *Journal of Financial Economics*, 10: 195-210.
- Grossman, Sanford J., Angelo Melino, and Robert J. Shiller, 1987, "Estimating the Continuous-Time Consumption Based Asset Pricing Model", *Journal of Business and Economic Statistics*, 5: 315-327.
- Hall, Robert E., 1988, "Intertemporal Substitution of Consumption", *Journal of Political Economy*, 96: 339-357.
- Hamada, Robert, 1988, "The Role of Statistics in Accounting, Marketing, Finance, and Production", *Journal of Business and Economic Statistics*, 6: 261-70.

- Hansen, Lars P. and Robert J. Hodrick, 1980 , "Forward Exchange Rates as Optimal Predictors of Future Spot Rates: An Econometric Analysis", *Journal of Political Economy*, 88: 829-53.
- Hansen, Lars P., and Kenneth J. Singleton, 1983, "Stochastic Consumption, Risk Aversion and the Temporal Behavior of Asset Returns", *Journal of Political Economy*, 91: 249-265.
- Hansen, Lars P., and Scott F. Richard, 1987, "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models", *Econometrica*, 55: 587-614.
- Harvey, Andrew, Esther Ruiz, and Enrique Sentana, 1992, "Unobserved Component Time Series Models with ARCH Disturbances", *Journal of Econometrics*, 52: 129-157.
- Hirshleifer, Jack, 1958, "On the Theory of Optimal Investment Decision", *Journal of Political Economy*, 66: 329-352.
- Hodrick, Robert J., 1981, "International Asset Pricing with Time-varying Risk Premia", *Journal of International Economics*, 1: 573-77.
- Hodrick, Robert J., 1987, *The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets*, Chur, Harwood Publishers.
- Hodrick, Robert J., and Sanjay Srivastava, 1984, "An Investigation of Risk and Return in the Forward Foreign Exchange Market", *Journal of International Money and Finance*, 3: 5-29.
- Hodrick, Robert J., and Sanjay Srivastava, 1986, "The Covariation of Risk Premiums and Expected Future Spot Exchange Rates", *Journal of International Money and Finance*, 5: 5-21.
- Hsieh, David A., 1984, "Tests of Rational Expectations and No Risk Premium in Forward Exchange Markets", *Journal of International Economics*, 17: 173-84.
- Hsieh, David A., 1989, "Modeling Heteroskedasticity in Daily Foreign Exchange Rates", *Journal of Business and Economic Statistics*, 7: 307-317.
- Ingersoll, Jonathan E. Jr., 1975, "Multidimensional Security Pricing", *Journal of Financial and Quantitative Analysis*, 10: 785-798.
- Ingersoll, Jonathan E. Jr., 1987, *Theory of Financial Decision Making*, Rowman & Littlefield Publishers, New Jersey.
- Jean, William H., 1967, "More on Multidimensional Portfolio Analysis", *Journal of Financial and Quantitative Analysis*, 8: 475-490.
- Jean, William H., 1971, "The Extension of Portfolio Analysis to Three or More Parameters", *Journal of Financial and Quantitative Analysis*, 6: 505-515.

- Jorion, Philippe, 1988, "On Jump Processes in the Foreign Exchange and Stock Markets", *Review of Financial Studies*, 1: 427-445.
- Kaminsky, Graciela and Rodrigo Peruga, 1990, "Can a Time-Varying Risk Premium explain Excess Return in the Forward Market for Foreign Exchange?", *Journal of International Economics*, 1990, 28: 47-70.
- Kane, Alex, 1982, "Skewness Preference and Portfolio Choice", *Journal of Financial and Quantitative Analysis*, 17: 15-25.
- Kendall, Jon D. and A. David McDonald, 1989, "Univariate GARCH-M and the Risk Premium in a Foreign Exchange Market", unpublished manuscript, University of Tasmania.
- Kiefer, N.M. and M. Salmon, 1983, "Testing Normality in Econometric Models", *Economics Letters*, 11: 123-127.
- Kodde, David A. and Franz C. Palm, 1986, "Wald Criteria for Jointly Testing Equality and Inequality Restrictions", *Econometrica*, 54: 1243-48.
- Koedijk, Kees G., Marcia M.A. Schafgans, and Casper G. de Vries, 1990, "The Tail Index of Exchange Rate Returns", *Journal of International Economics*, 29: 93-108.
- Koedijk, Kees G., and Clemens J.M. Kool, 1992, "Tail Index Estimates of East European Exchange Rates", *Journal of Business and Economic Statistics*, 10: 83-96.
- Koenker, Roger, 1981, "A Note on Studentizing a Test for Heteroskedasticity", *Journal of Econometrics*, 17: 107-12.
- Korajczyk, Robert A., 1985, "The Pricing of Forward Contracts for Foreign Exchange", *Journal of Political Economy*, 93: 346-68.
- Kraft, Dennis F., and Robert F. Engle, 1983, "Autoregressive Conditional Heteroskedasticity in Multiple Time Series", Working Paper, UCSD.
- Kraus, Alan, and Robert H. Litzenberger, 1976, "Skewness Preference and the Valuation of Risk Assets", *The Journal of Finance*, 31: 1085-1100.
- Kraus, Alan, and Robert Litzenberger, 1983, "On the Distributional Conditions for a Consumption-Oriented Three Moment CAPM", *The Journal of Finance*, 38: 1381-91.
- Kreps, David M., and Evan L. Porteus, 1978, "Temporal Resolution of Uncertainty and Dynamic Choice Theory", *Econometrica*, 46: 185-200.
- Kreps, David M., and Evan L. Porteus, 1979a, "Temporal Von Neumann-Morgenstern and Induced Preferences", *Journal of Economic Theory*, 20: 81-109.
- Kreps, David M., and Evan L. Porteus, 1979b, "Dynamic Choice Theory and Dynamic Programming", *Econometrica*, 47: 91-100.

- Krugman, Paul, 1991, "Target Zones and Exchange Rate Dynamics", *Quarterly Journal of Economics*, 56: 669-82.
- Krugman, Paul, and Julio Rotemberg, 1990, "Target Zones and Limited Reserves", NBER Working Paper, # 3418.
- Lamoureux, Christopher, and William D. Lastrapes, 1990, "Persistence in Variance, Structural Change and the GARCH Model", *Journal of Business and Economic Statistics*, 8: 225-34.
- Lastrapes, William D., 1989, "Exchange Rate Volatility and U.S. Monetary Policy: An ARCH Application", *Journal of Money, Credit, and Banking*, 21: 66-77.
- Lee, John H.H., 1991, "A Lagrange Multiplier Test for GARCH Models", *Economics Letters*, 37: 265-271.
- Levich, Richard, 1985, "Empirical Studies of Exchange Rates: Price Behavior, Rate Determination, and Market Efficiency", in R. Jones and P. Keenen, eds, *Handbook of International Economics*, Amsterdam: North Holland.
- Levy, Haim, 1969, "A Utility Function Depending on the First Three Moments", *The Journal of Finance*, 24: 715-719.
- Lewis, Karen K., 1990, "Can managed Float Interventions make Exchange Rate Behavior Resemble a Target Zone", Working paper, New York University.
- Lim, Kian-Guan, 1989, "A New Test of the Three-Moment Capital Asset Pricing Model", *Journal of Financial and Quantitative Analysis*, 24: 205-216.
- Lintner, John, 1965, "The Valuation of Risk Assets and the Selection of Risky Investment in Stock Portfolios and Capital Budgets", *Review of Economics and Statistics*, 47: 13-47.
- Ljung, Greta M. and George E.P. Box, 1978, "On a Measure of Lack of Fit in Time Series Models", *Biometrika*, 65, 297-303.
- Longworth, David, 1981, "Testing the Efficiency of the Canadian-US Exchange Market under the Assumption of no Risk Premium", *Journal of Finance*, 36: 43-49.
- Lucas, Robert E., 1978, "Asset Prices in an Exchange Economy", *Econometrica*, 46: 1429-1445.
- Lucas, Robert E., 1982, "Interest Rates and Currency Prices in a Two-Country World", *Journal of Monetary Economics*, 10: 335-60.
- MacDonald, Ronald and Thomas Torrance, 1989, "Some Survey Based Tests of Uncovered Interest Parity" in R. MacDonald and M. Taylor, eds, *Exchange Rates and Open Macroeconomics*, Cambridge: Basil Blackwell.
- Mandelbrot, Benoit, 1963, "The Variation of Certain Speculative Prices", *Journal of Business*, 36: 147-165.

- Mankiw, N. Gregory, and Matthew D. Shapiro, 1986, "Risk and Return, Consumption Beta Versus Market Beta", *The review of Economics and Statistics*, 68: 452-459.
- Mankiw, N. Gregory, and Stephen P. Zeldes, 1991, "The Consumption of Stockholders and Non-Stockholders", *Journal of Financial Economics*, 29, 97-112.
- Mark, Nelson C., 1988, "Time-Varying Betas and Risk Premia in the Pricing of Forward Foreign Exchange Contracts", *Journal of Financial Economics*, 22: 335-54.
- Markowitz, Harry, 1952, "Portfolio Selection", *The Journal of Finance*, 7: 77-91.
- Markowitz, Harry, 1959, *Portfolio Selection: Efficient Diversification of Investment*, Wiley, New York.
- Masera, Rainer, and Robert Triffin (eds.), 1984, *"Europe's Money: Problems in European Monetary Co-ordination and Integration*, Oxford,Clarendon Press.
- McCurdy, Thomas H., and Ieuan Morgan, 1992 , "Evidence of Risk Premia in Foreign Currency Futures Markets", *Review of Financial Studies*, 5: 65-83.
- McFarland, James W., R. Richardson Pettit and Sam K. Sung, 1982, "The Distribution of Foreign Exchange Price Changes: Trading Day Effects and Risk Measurement", *The Journal of Finance*, 37: 693-715.
- Meese, Richard A., and Kenneth Rogoff, 1983, "Empirical Exchange Rate Models of the Seventies: Are Any Fit to Survive?", *Journal of International Economics*, 14: 3-24.
- Mehra, Rajnish, and Edward C. Prescott, 1985, "The Equity Premium, A Puzzle", *Journal of Monetary Economics*, 15: 145-161.
- Merton, Robert C., 1969, "Lifetime Portfolio Selection Under Uncertainty: The Continuous-Time Case", *Review of Economics and Statistics*, 51: 247-57.
- Merton, Robert C., 1971, "Optimum Consumption and Portfolio Rules in a Continuous-Time Model", *Journal of Economic Theory*, 3: 373-413.
- Merton, Robert C., 1973, "An Intertemporal Capital Asset Pricing Model", *Econometrica*, Vol 41. no. 5, 867-887.
- Merton, Robert C., 1980, "On Estimating the Expected Return on the Market: An Exploratory Investigation", *Journal of Financial Economics*, 8: 323-61.
- Merton, Robert C., 1990, *Continuous-Time Finance*, Basil Blackwell, Cambridge, Massachusetts.
- Miller, Merton H., and Franco Modigliani, 1961, "Dividend Policy, Growth, and the Valuation of Shares", *Journal of Business*, 34: 411-33.
- Modigliani, Franco, and Merton H. Miller, 1958, "The Cost of Capital, Corporation Finance, and the Theory of Investment", *American Economic Review*, 56: 178-97

- Modigliani, Franco, and Merton H. Miller, 1963, "Corporate Income Taxes and the Cost of Capital: A Correction", *American Economic Review*, 53: 433-43.
- Mood, Alexander M., Franklin A. Graybill, and Duane C. Boes, 1988, *Introduction to the Theory of Statistics*, third edition, McGraw-Hill.
- Neftci, S.N., 1984, "Are Economic Time Series Asymmetric over the Business Cycle", *Journal of Political Economy*, 92: 307-328.
- Nelson, Daniel B., 1991, "Conditional Heteroskedasticity in Asset Returns : A New Approach", *Econometrica*, 59: 347-370.
- Nelson, Daniel B., 1992, "Filtering and Forecasting with Misspecified ARCH Models I: Getting the Right Variance with the Wrong Model", *Journal of Econometrics*, 52: 61-90.
- Nelson, Daniel B., and Charles Q. Cao, 1992, "Inequality Constraints in the Univariate GARCH Model", *Journal of Business and Economic Statistics*, 10: 229-35.
- Newey, Whitney K., and Kenneth D. West, 1987, "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix", *Econometrica*, 55: 703-708.
- Nieuwland, Frederik G.M.C., Willem F.C. Verschoor and Christian C.P. Wolff, 1991, "EMS Exchange Rates", *Journal of International Financial Markets, Institutions and Money*, 1: 21-42.
- Nieuwland, Frederik G.M.C., 1991 "Conditional Distributions for Stock Returns", Working Paper, University of Limburg.
- Nieuwland, Frederik G.M.C. and Willem F.C. Verschoor, 1992, "EMS Exchange Rate Expectations and Time-Varying Risk Premia", Working Paper, University of Limburg.
- Pagan, Adrian R. and Aman Ullah, 1988, "The Econometric Analysis of Models with Risk Terms", *The Journal of Applied Econometrics*, 3: 87-105.
- Pantula, Sastry G., 1986, "Modelling the Persistence of Conditional Variances : A Comment", *Econometric Reviews*, 5: 71-74.
- Perron, Pierre, 1988, "Trends and Random Walks in Macroeconomic Time Series", *Journal of Economic Dynamics and Control*, 12: 297-332.
- Perron, Pierre, 1989, "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis", *Econometrica*, 6: 1361-1401.
- Phillips, Peter C.B., 1987, "Time Series Regression with a Unit Root", *Econometrica*, 55: 277-301.
- Phillips, Peter C.B., and Pierre Perron, 1988, "Testing for a Unit Root in Time Series Regression", *Biometrika*, 75: 335-46.

- Restoy, Fernando, 1992, "Optimal Portfolio Policies under Time-Dependent Returns", Working Paper, Banco de España.
- Rogoff, Kenneth, 1984, "Can Exchange Rate Predictability be Achieved Without Monetary Convergence? Evidence from the EMS", *European Economic Review*, 28: 93-115.
- Roll, Richard, and Bruno Solnik, 1977, "A Pure Exchange Asset Pricing Model", *Journal of International Economics*, 7: 161-79.
- Rose, A. and L. Svensson, 1991, "Expected and Predicted Realignments: The FF/DM Exchange Rate during the EMS", NBER Working Paper # 3685.
- Ross, Stephen A., 1976, "Arbitrage Theory of Capital Asset Pricing", *Journal of Economic Theory*, 13: 341-60.
- Ross, Stephen A., 1987, "The Interrelations of Finance and Economics: Theoretical Perspectives", *American Economic Review*, 77: 29-34.
- Samuelson, Paul A., 1965, "Proof that Properly Anticipated Prices Fluctuate Randomly", *Industrial Management Review*, 6: 41-49.
- Schwarz, Gideon, 1978, "Estimating the Dimension of a Model", *Annals of Statistics*, 6, 461-464.
- Sears, R. Stephen, and K.C. John Wei, 1985, "Asset Pricing, Higher Moments, and the Market Risk Premium: A Note", *The Journal of Finance*, 40: 1251-1253.
- Sharpe, William F., 1964, "Capital Asset Prices : A Theory of Market Equilibrium under Conditions of Risk", *The Journal of Finance*, 19: 425-442.
- Simkowitz, Michael A., and William L. Beedles, 1978, "Diversification in a Three Moment World", *Journal of Financial and Quantitative Analysis*, 13: 927-941.
- Singleton J. Clay, and John Wingender, 1986, "Skewness Persistence in Common Stock Returns", *Journal of Financial and Quantitative Analysis*, 21: 335-41.
- Stulz, Rene, 1981, "A Model of International Asset Pricing", *Journal of Financial Economics*, 9: 383-406.
- Sundaresan, Suresh M., 1989, "Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth", *Review of Financial Studies*, 2: 73-89.
- Summers, Lawrence H., 1985, "On Economics and Finance", *The Journal of Finance*, 40: 633-35.
- Tan, Kai-Jiaw, 1991, "Risk Return and the Three Moment CAPM: Another Look", *Journal of Banking and Finance*, 15: 449-460.
- Tauchen, George E. and Mark Pitts, 1983, "The Price Variability-Volume Relationship on Speculative Markets", *Econometrica*, 51: 485-505.

- Taylor, Stephen J., 1986, *Modelling Financial Time Series*, Wiley, New York.
- Taylor, Stephen J., 1990, "Modelling Stochastic Volatility", Unpublished Manuscript, University of Lancaster.
- Tobin, James, 1958, "Liquidity Preference as Behavior Towards Risk", *Review of Economic Studies*, 25: 68-85.
- Vlaar, Peter, and Franz C. Palm, 1992, "The Message in Weekly Exchange Rates in the EMS: Mean Reversion, Conditional Heteroskedasticity and Jumps", Working Paper, University of Limburg.
- Vuong, Quang H., 1989, "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses", *Econometrica*, 57, 2, 307-333.
- Weber, A., 1991, "Reputation and the Credibility in the European Monetary System", *Economic Policy*, 12: 57-102.
- Weil, Philippe, 1987, "Non-Expected Utility in Macroeconomics", Working Paper, Harvard University.
- Weil, Philippe, 1989, "The Equity Premium Puzzle and the Risk-free Rate Puzzle", *Journal of Monetary Economics*, 24: 401-421.
- Weiss, Andrew A., 1984, "ARMA Models with ARCH errors", *Journal of Time Series Analysis*, 5: 129-143.
- Weiss, Andrew A., 1986a, "Asymptotic Theory for ARCH Models : Estimation and Testing", *Econometric Theory*, 2: 107-131.
- Weiss, Andrew A., 1986b, "ARCH and Bilinear Time Series Models : Comparison and Combination", *Journal of Business and Economic Statistics*, 4: 59-70.
- Westerfield, Janice M., 1977, "An Examination of Foreign Exchange Risk under Fixed and Floating Rate Regimes", *Journal of International Economics*, 7: 181-200.
- Wolff, Christian C.P., 1987a, "Forward Foreign Exchange Rates, Expected Spot Rates and Premia: A Signal-Extraction Approach", *The Journal of Finance*, 42: 395-406.
- Wolff, Christian C.P., 1987b, "Time-Varying Parameters and the Out-of-Sample Forecasting Performance of Structural Exchange Rate Models", *Journal of Business and Economic Statistics*, 5: 87-97.
- Zakoian, Jean-M, 1990, "Treshold Heteroskedastic Model", Working Paper, INSEE, Paris.

Nederlandse Samenvatting

Het begrip risico speelt een essentiële rol binnen de wereld van de financiële markten. Voor een juiste prijsvorming met betrekking tot beleggingsobjecten is een juiste inschatting van het te lopen risico onontbeerlijk. Immers, in een wereld die bevolkt wordt door risico-averse agenten dient een verhoogd risico gepaard te gaan met een verhoogde verwachte opbrengst, i.e. het verwachte rendement van het beoogde beleggingsobject. Binnen de Financiële Economie wordt het risico van beleggingsobjecten doorgaans gemeten via de variantie van de empirische kansverdeling van historische rendementen, of via de covariantie van toekomstige rendementen met diverse referentie-objecten. Sinds het midden van de jaren zestig is bekend dat deze risicomaatstaven aan veranderingen onderhevig zijn. Dit fenomeen noemt men heteroskedasticiteit en heeft als consequentie dat ten gevolge van een veranderend risico, ook de risicopremie, i.e. de vergoeding die geboden wordt voor het aangaan van dit risico, in de tijd zal variëren.

Gegeven het bovenstaande dienen zich twee vragen aan. Allereerst zal men zich moeten buigen over de vraag hoe deze maatstaven in de tijd evolueren. Vervolgens vereist de opname van tijdsvariërende risicomaatstaven in financiële waarderingsmodellen de nodige aandacht. Dit proefschrift gaat in op beide vragen en concentreert zich daarbij op de aandelenmarkt en de valutamarkt.

Het statistische instrumentarium dat in alle analyses gebruikt en op zijn merites beoordeeld wordt, is een (Gegeneraliseerd) Autoregressief Model ter beschrijving van Conditionele Heteroskedasticiteit ((G)ARCH Model). Dit model is in de Financiële Economie zonder twijfel het meest gebruikte statistische model wanneer het gaat om de modellering van de evolutie van

conditionele varianties in financiële tijdreeksen. Een kritische evaluatie van de voordelen die dit model biedt is zeker op zijn plaats. Verder wordt in dit proefschrift bezien of het model gebruikt kan worden bij het bepalen van tijdsvariërende risicopremies en wordt een eerste aanzet gegeven voor een juiste koppeling tussen dit model en financiële waarderingsmodellen. De belangrijkste bevindingen van de individuele hoofdstukken worden hieronder uiteengezet.

In **Hoofdstuk 2** wordt een GARCH(1,1) model, onder diverse verdelingsassumpties, gebruikt om het tijdspatroon in de conditionele variantie van de rendementen van verscheidene aandelenindices te beschrijven. Deze analyse wordt zowel voor wekelijkse als voor maandelijkse rendementen uitgevoerd. De schattingresultaten wijzen eenduidig op de aanwezigheid van conditionele heteroskedasticiteit, zoals weergegeven door het GARCH model. De standaard Lagrange Multiplier test die doorgaans gebruikt wordt om GARCH effecten op te sporen wijst dit echter niet altijd uit. Een alternatieve test die gebaseerd is op homogene Markov-ketens blijkt betrouwbaarder te zijn. De volgende conditionele verdelingen zijn gebruikt ; een Normale verdeling, een Student t-verdeling en een GED (Generalized Error Distribution). De Student t-verdeling blijkt op basis van "Generalized Likelihood Ratio" toetsen de beste keuze te zijn, en conditionele Normaliteit wordt sterk verworpen. Uit de analyse blijkt verder dat conditionele heteroskedasticiteit volgens de GARCH vorm niet verdwijnt onder tijdsaggregatie, hetgeen wel dikwijls wordt aangenomen, en dat de niet-conditionele verdeling van herschaalde residuen zeker niet Normaal is. Een Student t-verdeling met enkele vrijheidsgraden blijkt een plausibeler alternatief. Deze uitkomsten hebben met name gevolgen voor multivariate analyses waarin men doorgaans nog conditionele normaliteit aanneemt.

Hoofdstuk 3 beschouwt de implementatie van een GARCH model, dat de conditionele variantie van het marktrendement (het referentie-object) beschrijft, in een intertemporeel waarderingsmodel voor aandelenrendementen. Op basis van de aannname van Gegeneraliseerde Isoelastische Preferenties wordt een model afgeleid dat zowel elementen uit het traditionele Capital Asset Pricing Model (CAPM) bevat, als elementen uit het Consumptie CAPM. Daar er diverse bezwaren kleven aan het gebruik van consumptie-maatstaven wordt consumptie uit het model geëlimineerd. In een hetero-

skedastische wereld is dit echter verre van triviaal. De aanname dat het intercept in de lineare relatie tussen de verwachte consumptiegroei en het verwachte marktrendement afhangt van de conditionele variantie van het marktrendement is vereist om eliminatie van consumptie mogelijk te maken. In het vervolg van de analyse wordt duidelijk dat een GARCH model voor de conditionele variantie van het marktrendement niet leidt tot het GARCH-in-mean model dat in deze context vaak gebruikt wordt. Het waarderingsmodel dat uiteindelijk resulteert kan gezien worden als een conditioneel drie-momenten CAPM waarin niet alleen conditionele variantie maar ook conditionele tijdsvariërende scheefheid een rol speelt. Dit model illustreert het belang van een afweging tussen risico versus scheefheid, naast de gebruikelijke afweging tussen rendement versus risico.

In **Hoofdstuk 4** wordt het model uit het voorgaande hoofdstuk aan een empirische toets onderworpen. Daarbij wordt uitgegaan van de marktversie van het ontwikkelde model om vergelijking met het GARCH-in-mean model, waarop het nieuwe model een verbetering is, mogelijk te maken. Er wordt gebruikt gemaakt van de dataset uit hoofdstuk 2. Door middel van calibratiemethoden wordt aangetoond dat het model in staat is om plausibele waarden voor de relevante parameters te genereren. Vervolgens wordt het Gegeneraliseerde Autoregressieve Conditionele Scheefheid (GARCS) proces geïntroduceerd om tijdsvariërende conditionele derde momenten te modelleren. Wanneer men de schattingsresultaten van het conditionele drie-momenten CAPM vergelijkt met de resultaten voor het GARCH-in-mean model dan blijkt dat het eerstgenoemde model realistische en statistisch significante waarden voor de mate van relatieve risico-versuscheefheid oplevert in tegenstelling tot het GARCH-in-mean model. Verder blijkt uit de resultaten dat er sprake is van een risico versus scheefheid afweging. Samengevat kan men stellen dat de hoofdstukken 3 en 4 uitwijzen dat de implementatie van een GARCH model in een economisch waarderingsmodel niet zo eenvoudig is als het GARCH-in-mean model veronderstelt.

Conditionele heteroskedasticiteit speelt ook een grote rol bij de beschrijving van de statistische eigenschappen van valutarendementen. In **Hoofdstuk 5** wordt dit bevestigd voor valutarendementen van landen die deelnemen aan het wisselkoersarrangement van het Europese Monetaire Systeem. Naast het normale diffusierisico is met betrekking tot deze valuta nog een andere

risicofactor aanwezig. De mogelijkheid van een pariteitsherschikking is altijd aanwezig. Dergelijke pariteitsaanpassingen zorgen voor grote sprongen in de bilaterale wisselkoersen. In de empirische analyse wordt hiermee rekening gehouden door een onafhankelijk Poisson sprong proces in de beschouwing op te nemen. Het model dat het best de relevante eigenschappen van EMS valutarendementen weergeeft is een gecombineerd Sprong-GARCH model met innovatietermen die conditioneel Student t-verdeeld zijn.

Hoofdstuk 6, tenslotte, gaat in op de vraag of de conditionele variantie van verwachte (EMS) valutarendementen, beschreven door een GARCH proces, als een determinant voor de risicopremie op de valuta termijnmarkt kan fungeren. In het algemeen is de termijnkoers geen zuivere voorspeller voor de verwachte toekomstige spotkoers. Dit kan verklaard worden door irrationeel gedrag van marktparticipanten of door het bestaan van tijdsvariërende risicopremies, of door het optreden van beide. In dit hoofdstuk wordt het bestaan van tijdsvariërende risicopremies aangetoond, waarbij de beschikbaarheid van gemeten verwachtingen de analyse vrijwaart van de aannname van rationele verwachtingen. Verder blijkt dat de risicopremies redelijk beschreven kunnen worden met behulp van het GARCH-in-mean model.

Curriculum vitae

Frederik G.M.C. Nieuwland is op 31 mei 1965 geboren te Hoensbroek. In 1983 slaagde hij voor het eindexamen Gymnasium-beta aan het Sint-Jan's College te Hoensbroek. Gedurende de periode september 1983 - augustus 1987 studeerde hij Econometrie aan de Katholieke Universiteit Brabant met als specialisatierichtingen Bedrijfseconometrie en Besliskunde. In augustus 1987 behaalde hij zijn doctoraaldiploma. Na het vervullen van zijn militaire dienstplicht trad hij in februari 1989 in dienst van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). Als Onderzoeker in Opleiding (OIO) werkte hij binnen de sectie Financiering van de Faculteit der Economische Wetenschappen van de Rijksuniversiteit Limburg aan een onderzoeksproject waarvan dit proefschrift het resultaat is. Momenteel is hij aan dezelfde faculteit verbonden als Postdoctoral Research Fellow. Per april 1993 zal hij aan de Graduate School of Business van de University of Chicago verbonden zijn als Visiting Scholar.

Frederik G.M.C. Nieuwland, born May 31st 1965 in Hoensbroek, studied Econometrics at Katholieke Universiteit Brabant from 1983 until 1987. In August 1987 he received his Master's degree in Econometrics. After his military service, he became a Research Assistant for the Dutch Organization for Scientific Research (NWO) in February 1988. Stationed at the Department of Finance of the University of Limburg he completed this doctoral dissertation. Currently he is employed as a Postdoctoral Research Fellow at the same university. In April 1993, he will join the Graduate School of Business at the University of Chicago, as a Visiting Scholar.

Datawyse | Universitaire Pers Maastricht
ISBN 90 5278 069 2
NUGI 682/683