

Adventures in international financial markets

Citation for published version (APA):

Huisman, R. (1999). *Adventures in international financial markets*. [Doctoral Thesis, Maastricht University]. Thela Thesis. <https://doi.org/10.26481/dis.19990225rh>

Document status and date:

Published: 01/01/1999

DOI:

[10.26481/dis.19990225rh](https://doi.org/10.26481/dis.19990225rh)

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.

ADVENTURES IN INTERNATIONAL FINANCIAL MARKETS

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan
de Universiteit Maastricht,
op gezag van de Rector Magnificus, Prof. dr. A.C. Nieuwenhuijzen Kruseman,
volgens het besluit van het College van Decanen,
in het openbaar te verdedigen
op donderdag 25 februari 1999 om 14.00 uur

door

Ronald Huisman

geboren op 17 januari 1972 te 's Gravenhage

Promotores:

Prof. dr. C.G. Koedijk

Prof. dr. C.J.M. Kool

Prof. dr. F.C. Palm

Beoordelingscommissie:

Prof. dr. P.C. Schotman (voorzitter)

Prof. dr. R. Lyons (University of California at Berkeley, USA)

Dr. J.R.Y.J. Urbain

Prof. dr. C.G. de Vries (Erasmus Universiteit Rotterdam)

PREFACE

Adventures in International Financial Markets is the result of three creative years spent at the Finance section at Maastricht University. I owe a great debt of gratitude to my supervisors Kees Koedijk, Clemens Kool, and Franz Palm. The trust, time, support, and help you gave me is gratefully acknowledged. A special 'thanks' goes to Kees: the completion of this dissertation is for a great deal due to the investments you made, your willingness to cooperate and share, and the interests you have. This was beyond any value.

I would also like to thank the colleagues at Maastricht University and the following people in particular who helped me with the completion of this dissertation (alphabetically): Piet Eichholtz, Mark Flood, Paul van Homelen, Nathalie Jansen, Jeanine Kippers, Richard Lyons, Ronald Mahieu, François Nissen, Ailsa Röell, Peter Schotman, Mark Schweitzer, Jean Pierre Urbain, Casper de Vries, Carina Wijnands, and PUC. I would also like to thank my friends for their support.

Above all, I would like to thank my parents and my brother for their love and support. I never could have completed these adventures without them.

R.H.

APPENDIX

APPENDIX A
APPENDIX B
APPENDIX C

The first part of the appendix contains the results of the first experiment. The second part contains the results of the second experiment. The third part contains the results of the third experiment. The fourth part contains the results of the fourth experiment. The fifth part contains the results of the fifth experiment. The sixth part contains the results of the sixth experiment. The seventh part contains the results of the seventh experiment. The eighth part contains the results of the eighth experiment. The ninth part contains the results of the ninth experiment. The tenth part contains the results of the tenth experiment.

The first part of the appendix contains the results of the first experiment. The second part contains the results of the second experiment. The third part contains the results of the third experiment. The fourth part contains the results of the fourth experiment. The fifth part contains the results of the fifth experiment. The sixth part contains the results of the sixth experiment. The seventh part contains the results of the seventh experiment. The eighth part contains the results of the eighth experiment. The ninth part contains the results of the ninth experiment. The tenth part contains the results of the tenth experiment.

The first part of the appendix contains the results of the first experiment. The second part contains the results of the second experiment. The third part contains the results of the third experiment. The fourth part contains the results of the fourth experiment. The fifth part contains the results of the fifth experiment. The sixth part contains the results of the sixth experiment. The seventh part contains the results of the seventh experiment. The eighth part contains the results of the eighth experiment. The ninth part contains the results of the ninth experiment. The tenth part contains the results of the tenth experiment.

CONTENTS

1	GENERAL INTRODUCTION	8
2	INTRODUCTION TO THE 1ST PART	14
3	EXTREME SUPPORT FOR UIP	16
3.1	Data	17
3.2	Single equation results	19
3.2.1	Standard time series tests	19
3.2.2	Outlier tests	21
3.3	The panel methodology	24
3.4	Empirical results	27
3.4.1	Results from standard panels	27
3.4.2	The importance of outliers	28
3.4.3	Outliers and EMS realignments	31
3.5	Concluding remarks	32
4	FAT TAILS IN SMALL SAMPLES	34
4.1	Methodology	36
4.1.1	The conventional Hill (1975) estimator	36
4.1.2	An alternative approach	39
4.1.3	Econometric considerations	40
4.1.4	Point estimates	40
4.1.5	Standard errors	41
4.1.6	Testing whether two tail-indexes are equal	43
4.2	Simulation results	44
4.3	Results on stock and exchange rate returns	52
4.3.1	Returns on the S&P 500	53
4.3.2	Returns on foreign exchange rates	55
4.4	Concluding remarks	58
5	THE TAIL FATNESS OF FX RETURNS	62
5.1	Data	63
5.2	Results	65
5.2.1	Tail-index estimates	65
5.2.2	The tail-index of EMS rates versus floating rates	67

5.2.3	A comparison: HKKP estimates versus the literature	68
5.3	Distributional characteristics	69
5.4	Conclusions and practical implications	74
5.5	Appendix: Goodness-of-fit test	75
6	INTRODUCTION TO THE 2 ND PART	78
6.1	Market microstructure	79
6.2	Objectives	84
7	THE EXPERIMENTAL ENVIRONMENT	86
7.1	Overall	86
7.2	The microstructure	86
7.3	Dealer objectives and the liquidation value of the asset	88
7.4	The trading process	89
7.5	The behavior of investors	90
7.6	The trading mechanism	91
8	QUOTE DISCLOSURE	93
8.1	Experimental Design	95
8.1.1	Payoff to the subjects	95
8.2	Data	96
8.3	Results	98
8.3.1	Volume	98
8.3.2	Spreads	99
8.3.3	Price efficiency	101
8.4	Summary and concluding remarks	106
9	POST-TRADE TRANSPARENCY	108
9.1	Experimental design	113
9.1.1	Payoff to the subjects	113
9.2	Data	114
9.3	Results	116
9.3.1	Price efficiency	116
9.3.2	Profits	120
9.3.3	Spreads faced by the different agents	123
9.3.4	Spreads	127
9.4	Conclusions	130

10	SEARCH COSTS	132
10.1	Effective Spreads: Four Leading Models	133
10.2	Experimental Design	135
10.2.1	Information Structure	135
10.2.2	Trade Transparency, Payoffs and Summary Statistics	136
10.3	Results	138
10.3.1	Serial correlation in changes in transaction prices	138
10.3.2	Consistency across effective spreads models	139
10.3.3	Effective spreads and the role of search costs	140
10.4	Concluding remarks	140
11	FINANCIAL MARKET COMPETITION	147
12	OVERALL CONCLUDING REMARKS	150
13	REFERENCES	154
14	NEDERLANDSE SAMENVATTING	164
14.1	Deel 1: Wisselkoersen	164
14.2	Deel 2: De microstructuur van financiële markten	166
	CURRICULUM VITAE	169

1 General Introduction

'An economist is someone who, instead of telling us why the world is the way it is, tells us why it isn't the way his theories suggest it should be.'

This is the first sentence with which Lawrence Copeland starts his handbook on exchange rates and international finance¹. Although one might argue that the above description holds for many economists in various research areas, it typically describes international financial economists. This rather pessimistic view on international finance originated from frequent rejections of theoretical predictions concerning the behavior of international financial markets. For example, Cumby and Obstfeld (1981, 1984) and Meese and Rogoff (1983a, 1983b, 1988) were among the first to conclude that popular theoretical exchange rate models did not outperform the random walk in predicting nominal and real exchange rates. Lewis (1995) formulates two research questions in her extensive review of this literature: 1) 'What can explain the behavior of domestic relative to foreign returns and can these explanations suggest ways in which models of the rest of the economy are either succeeding or failing?' and 2) 'Why do domestic investors appear to ignore potential gains to foreign investment opportunities and does the answer imply necessary modifications to our views about international capital market equilibrium?'

Both questions concern solving the puzzles in international finance that explain the empirical deviations from theoretical predictions; why isn't it the way theories suggest it should be?

The puzzles in international finance have become increasingly important as international financial markets have undergone tremendous growth and as they are changing rapidly in nature as trade-barriers disappear (EMU, NAFTA and Mercosur) and communications improve. Activity (fund-raising) on international financial markets increased by about 350% from \$400 billion in 1990 to \$1.8 trillion in 1997². The markets even grew by 10% in 1997 despite the Asian financial crises that started in October 1997, making 1997 the fourth consecutive record year. The trends to liberalize capital markets in the world are likely to feed extra growth in the future. This thesis focuses on some of the existing puzzles in international financial markets. It consists of two parts. Part 1 presents innovations in modeling exchange rate behavior. Part 2 examines the learning behavior of agents in financial markets that are typical for the foreign exchange market and their influences on the overall market performance and on the assets that are being traded their.

The first chapter in part 1 concentrates on the Fisher (1930) interest parity. According to this theory the expected returns in one country should be equalized through speculation to the returns in another country once converted to the same exchange rate. Many studies have attempted to test the theory empirically and most failed to find confirming results³. Explanations have been proposed, mostly based on the view presented by Fama (1984) that traditional tests are biased due to time varying risk premiums and systematic forecast errors made by market agents. However, accounting for time varying premiums and forecast errors has shown not to produce empirical results that are more in line with Fisher's theory. Lewis concludes that many of the explanations presented ignored the other potential explanations, which suggests that more than one disturbing factor⁴ could be apparent at a certain point in time. In chapter 3 it is argued that the traditional tests that are based on only one exchange rate at a time do not take account of these potentially disturbing factors. We provide a more general approach that pools the information obtained from a multiple of exchange rates at the same time. The proposed model accounts for random economic shocks that might

¹ Copeland (1994).

² Source: OECD Financial Market Trends, February 1988.

³ Lewis (1995) overviews this literature.

⁴ Disturbing in the sense that it is not impounded by the original theoretical model.

influence one or more exchange rates at the same time. The results show that the deviation from Fishers interest parity are less severe and even indicate that under special market conditions the parity almost perfectly holds. At least one of the puzzles could partly be explained by misspecification. The other two chapters in part 1 focus on the unconditional distribution of exchange rate returns. The issue of the distribution characteristics of exchange rate returns is of crucial importance in exchange rate risk management and for proper hypothesis testing. Although it is frequently observed that the distribution of returns on exchange rates have fatter tails than the normal distribution, there is little consensus on the exact magnitude of the tail fatness and therefore on the exact structure of the distribution. Westerfield (1977), Rogalski and Vinso (1978), Boothe and Glassman (1987), and McCulloch (1997) have proposed the Student-t distribution, a mixture of two normal distributions and stable distributions as potential candidates, but mostly show weak evidence that these distributions provide good fits of the unconditional distribution of exchange rate returns. Again, we argue that the weakness is due to potential measurement errors. Due to small sample problems the parameters that are related to the tail structure of the distributions can often not be determined without a bias. We propose a method that corrects for the small sample biases in tail index estimates. This gives us a method to correctly assess tail fatness and model the distribution of the tails of exchange rate returns adequately. The results indicate that the Student-t distribution provides a good fit for most tails (and often for the whole) of the distribution of monthly returns on all floating exchange rates and on most fixed rates. The good fit of the Student-t brings an easy tool to analyze the downside risk faced on exchange rates as appropriate quantiles can easily be calculated using the Student-t structure.

A completely different methodology is taken to study the puzzles in financial markets in the second part of this thesis. As the first part is based on empirical models to study the behavior in financial markets, the second part is based on experiments. Experiments are used to examine closely the way demand and supply for a financial asset comes together and how price-formation takes place. More specifically, we shall focus on how different trading protocols directly affect price formation in international financial markets. This *market microstructure* approach 'enhances our ability to understand both the returns to financial assets and the process by which markets become efficient' (see O'Hara (1995) page 1) and thus offers an alternative route to study the puzzles in international financial markets. Research in market micro structure, see O'Hara for a first

extensive overview, has shown that the actual design of the trading process installed on a financial market, the microstructure, has a significant impact on the efficiency and trading costs in financial markets. The micro structural environment directly affects the characteristics of the assets being traded. It has been shown that trading mechanisms have their own impact on price-formation and this explains why many different trading structures co-exist worldwide. Not only do we observe differences across financial markets, it is frequently the case that the same security trades in different microstructures, sometimes even in the same room. For example, the New York Stock Exchange runs a call market at the morning opening, and continuous specialist trading throughout the day, along with a separate upstairs market for large transactions. Many European financial markets are also built on a multiple of trading mechanisms. One example is the Amsterdam Stock Exchange where small orders are channeled to a public order book via a screen based continuous auction system. Larger orders can be traded via direct deals, via a quote driven screen based trading system (ASSET) or via a automatic interdealer market (AIDA). An important example internationally is the foreign exchange market that operates simultaneously as a decentralized interbank direct market and a more centralized brokered market.

Many of the market structures described above operate as multiple dealer markets; markets where more than one market maker provide liquidity by setting the bid and ask prices against either outside customers or other market makers can trade. Examples of such markets are NASDAQ, London SEAQ (International), AIDA in Amsterdam and the foreign exchange market. Although a significant amount of transactions are channeled through the above multiple dealer market, the theoretical microstructure literature has typically concentrated on specialist markets where only one market maker is active. The reason for this imbalance is not a lack of interest but complexity. It is hard to implement a multiple of market makers in a theoretical model without restricting the behavioral heterogeneity across these market makers. Some empirical attempts have been done to tackle multiple dealer markets, such as Gemmill (1996) who focused on the London SEAQ International market and Lyons (1995, 1996, 1998) who focused on the foreign exchange market. Empirical studies suffer from other potential problems, as it is hard to disentangle the effects of the microstructures across different exchanges from various economic differences between the countries of origin. Another caveat arises from the data available for the foreign exchange market, as either databases with quotes are available (like those provided by Olsen and Associates) or databases with only

transaction prices (like the one used by Evans (1998)). There are no databases that contain both quotes and transactions that would describe optimally the multiple dealer market. This thesis adopts the experimental methodology to examine the effects of the design of the trading process on multiple dealer financial markets. Although many other potential problems may arise, the experimental environment allows us to study the market without restricting the heterogeneity among market makers and provides us with a database that consists of both quotes and transactions. The experimental design is presented and discussed in chapter 7.

In chapters 8 through 10 the experiments are used to examine the effects of different trading environments on the performance of international financial markets. Basically, two characteristic variables are distinguished: pre-trade and post-trade transparency. The concept of transparency deals with the amount of trading information that is provided to agents on financial markets and generally two types of such information exist. First, pre-trade transparency refers to information on quotes that agents observe before they decide with whom to transact. Second, post-trade transparency refers to the amount of information provided to the agents after a transaction has been cleared. Although one would expect that more information would lead to more efficient markets, it is shown that this is not generally the case. A high level of pre-trade transparency blocks prices to converge to a full information value. The opposite is true for post-trade transparency; the more transaction information is provided to market makers the more price efficient the market is. A reversed pattern is found in terms of liquidity and trading costs. A high level of pre-trade transparency leads to narrower spreads and more (opening) liquidity, whereas the opposite effect holds for post-trade transparency. In chapter 11 the results are summarized and used in a framework of financial market competition. As the EMU will increase the competition for order flow between European exchanges the question is what impact this will have for those European exchanges. As the EMU starts with 11 participating countries, more than 11 exchanges will compete for order flow and none of them wants to lose market share to other exchanges as happened after the revolutionary changes in the trading process on the London Stock Exchange in 1986. It is likely that some of the smaller exchanges will lose market share to other exchanges if the current trend towards more similar trading structures across European exchanges continues. However, it is shown that they could regain market share by profiting from the opposite effects of pre-trade and post-trade transparency on market performance.

ADVENTURES

I

EXCHANGE RATE RETURNS

2 Introduction to the 1st part

The impact of the foreign exchange market in the economy is enormous. For example, the announcement of the sex scandal between President Clinton and the former white house trainee Monica Lewinsky in January 1998 caused a drop of the dollar of more than 3% against its most important counterpart currencies in three days time. This announcement caused an external loss in the value of the dollar of more than \$120 billion, measured in terms of the total money supply in the U.S. To dramatize the impact, the amount of loss is more than the total funding of the International Monetary Fund to the countries involved in the Asia crises in the fall of 1997⁵. The volume, growth and impacts of the foreign exchange market reflect the importance of studying the risks that are faced by agents that actively participate in this market. The average daily volume of transactions in the spot foreign exchange market equaled \$1.2 trillion in 1995, which is the double of the average daily volume in 1989⁶.

The first part of this thesis focuses on studying models that are inputs to foreign exchange rate risk managers. Chapter 3 discusses the predictability of future returns on exchange rates based on the Fisher (1930) relation between exchange rate returns and interest rate differentials between countries. If this relation holds, managers would have a tool they can use to build their future scenarios regarding the profitability of their internationally operating firms. Chapters 4 and 5 concentrate on the current exposure towards exchange rate downside risk and provide information on the characteristics of the unconditional monthly distribution of exchange returns. Knowing these characteristics analysts can calculate their correct (downside) risk exposure in for example Value at Risk analyses or again in future scenario analyses.

⁵ Source: NRC Handelsblad, January 27, 1998.

⁶ Source: Bank of International Settlements survey 1995.

The contribution of this thesis to the existing literature is that it proposes new techniques to study exchange rate behavior as it is argued that measurement and modeling errors have strongly influenced our perception regarding the behavior of international financial markets. In chapter 3 it is argued that the relation between future changes in exchange rates has been tested frequently using time series models based on only one exchange rate. Keeping in mind the illustration of the Clinton - Lewinsky scandal and its impact on the value of the dollar against both the Yen and the Deutsch Mark block currencies, it is shown that these common factors are not accounted for in the traditional time series tests. Therefore, a panel approach is proposed that extends our time series view from one exchange rate at a time to a panel of more exchange rates. It is shown that such panel models account for worldwide shocks in the economy that may act as common factors in exchange rates. The results show that previous conclusions regarding the relation between future exchange rate returns and interest rate differences over countries should be relaxed as it turns out that the deviations from the relation are less severe than was commonly thought. The chapters 4 and 5 concentrate on the tails of the unconditional distribution of exchange rates. It is a well-known stylized fact that the distribution of returns on many financial assets, including exchange rates, exhibit fatter tails than the normal distribution. It is argued in those two chapters that our perception towards the tail fatness of exchange rate returns has been based on estimates of tail fatness that were derived from estimators that are seriously biased in small samples. Mostly, the results overstated the amount of tail fatness. In chapter 4 a new technique is developed that accounts for the estimators' bias in small samples and thus presents a method to estimate the tail parameters more correctly in small samples. In chapter 5 the technique is used to learn more about the tail characteristics of exchange rate returns. It is shown there that the Student-t distribution provides a good fit for the tails of the unconditional distribution of exchange rate returns where previous attempts in the literature failed, again due to errors in estimating the tail parameters of the Student-t distribution.

3 Extreme Support for UIP⁷

Uncovered interest parity (UIP) is one of the fundamental theoretical building blocks for understanding the behavior of returns in international financial markets. If UIP holds, the return on a domestic currency deposit equals the expected return from converting the domestic currency into the foreign currency, investing it in a foreign deposit and then converting the proceeds back into the domestic currency at the expected future exchange rate. Consequently, expected excess returns in the foreign exchange market equal zero on average. UIP implies that the forward premium is an unbiased predictor of the expected change in the spot exchange rate. UIP has been investigated for many different currencies and many different time periods and has been rejected in almost every study.

Various explanations have been proposed for this rejection. Fama (1984) distinguishes between time-varying risk premiums on the one hand and (systematic) market forecast errors on the other. Both sources of rejection have been subject of intensive research over the past two decades, but none of these studies have found satisfactory explanations. Lewis (1995) evaluates the various explanations and concludes: 'examining each of these explanations in isolation might lead to the conclusion that predictable excess returns remain a complete mystery. However, each of these explanations has ignored the other explanations.' Lewis suggests implicitly that a multiple of co-existing biasing factors exist. These factors may show up in more than one currency at the same time and therefore bias the estimates of the slope coefficient in a time series regression of spot changes on observed forward premiums. Baillie and Bollerslev (1997) show that both neglected persistent autocorrelation in the forward premium and the relatively small sample size bias the standard time series UIP test.

⁷ This chapter is based on: R. Huisman, K. Koedijk, C. Kool, and F. Nissen 'Extreme Support for Uncovered Interest Parity' *Journal of International Money and Finance*, 1998, 17, 211-228.

In light of the above, UIP may be valid but might be undetectable for many reasons. Our approach is complementary to the above literature by taking a different approach. The small-sample problem referred to by Baillie and Bollerslev is circumvented by a panel methodology that tests the validity of UIP for more than one exchange rate simultaneously. Obviously, the efficiency of the estimation is increased by using both the cross-sectional and time series dimension of the system of multiple exchange rates. The existence of a multiple of biasing factors as suggested by Lewis is accounted for by the inclusion of random time effects in the panel model. These time effects capture economic shocks that influence all exchange rates whereby the sensitivity to each shock is unique for each exchange rate. In addition, the panel method has the attractive feature of being invariant to the choice of the numeraire currency. The results are then independent of the specific characteristics of the chosen numeraire currency.

Another biasing factor arises from the differences between 'normal' and 'abnormal' observations. Bilson (1981) was the first to make this distinction and found that, due to market inefficiencies, outlier observations (in terms of the size of the forward premium) have worse forecasting power than normal ones. Flood and Rose (1994) and Flood and Taylor (1996) also present evidence on similar differences in forecasting power. In contrast with Bilson, they find that abnormal periods outperform normal periods in forecasting power⁸. In this chapter, we define 'normal' and 'abnormal' observations in terms of the average size of the forward premiums over all exchange rates. We concentrate on the importance of these extreme observations by testing the UIP hypothesis in which we explicitly account for different impacts over both types of observations.

3.1 Data

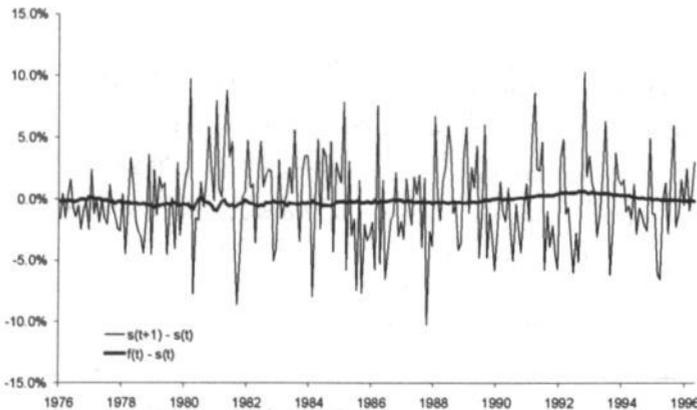
This study uses end-of-the-month spot and forward exchange rates, which cover the period January 1979 until March 1996. The forward maturity is one month. All data are London closing mid prices against the pound sterling for fifteen countries: Austria, Belgium, Canada, Denmark, France, Germany, Ireland, Italy, the Netherlands, Norway, Spain, Sweden, Switzerland, the United Kingdom and the United States.⁹

⁸ Flood and Rose (1994) test for UIP and define outlier periods as those in which EMS realignments occurred. Flood and Taylor (1996) test for UIP and define outliers in terms of the size of the interest differential. Schotman, Straetmans and de Vries (1997) also find important time differences in their test results.

⁹ We also conducted the analysis with bid and ask prices and find that the empirical results do not lead to different conclusions as in Bekaert and Hodrick (1992). We left Japan out of the analysis since no sufficient data on forward prices was available.

The data is sampled following the procedure as described in Bekaert and Hodrick (1992) in order to use the exact delivery dates of the forward foreign exchange contracts. For a one-month forward contract, the settlement on the spot market is two business days in the future. The delivery takes place on that calendar day in the next month, which corresponds to the current settlement date in the spot market, under the condition that the delivery date is a business day. If not, delivery takes place on the next business day given that it is in the same month. If the latter condition is not fulfilled, delivery takes place on the first previous business day.

Let $s_{i0}(t)$ be the *log*-price of the foreign currency i in units of the numeraire currency 0 at time t and $f_{i0}(t)$ be the *log*-forward price at t of the exchange rate with a one month maturity ($i = 1..N$; $t = 1..T$). The continuously compounded rate of depreciation of the foreign currency relative to the numeraire currency in the period from t until the forward delivery date $t+1$ is defined as $\Delta s_{i0}(t+1)$, where Δ is the one-period backward difference operator. All values are expressed as annualized percentages. In Table 1 we present averages of forward rate prediction errors defined as $s_{i0}(t+1) - f_{i0}(t)$ against the US Dollar and the German Mark expressed in percent per annum. Relative to the US Dollar, seven exchange rates have biases smaller than one percent on average and none of the average differentials is significantly different from zero. Relative to the German Mark, average differentials are slightly larger and significantly different from zero for Austria, Denmark, France and Ireland. Also, the variability of spot exchange rate changes exceeds the variability of the forward premium for all the exchange rates as noted already by Bilson (1981). This comes forward clearly in the following time-series figure for the DeutschMark - US Dollar exchange rate.



1 Actual returns versus forward premium DEM/USD

This makes the forward premium a very imprecise predictor of the future exchange rate change. For the EMS exchange rates expressed against the German Mark, the variance of spot changes is small relative to floating exchange rates, while the variance of the forward premiums is comparable. From Table 1, it is clear that the choice of numeraire currency is not irrelevant.

1 Summary Statistics

These summary statistics reflect the stylized facts over the exchange rates over the sample period from January 1979 until March 1996. All units are annualized percentages.

Currency	US Dollar			DEM		
	μ	se(μ)	$\sigma_{\Delta s}/\sigma_{fp}$	μ	se(μ)	$\sigma_{\Delta s}/\sigma_{fp}$
Austria	-0.16	(2.96)	12.44	-0.67*	(0.32)	4.64
Belgium	-0.31	(2.97)	13.22	-0.82	(0.73)	3.05
Canada	0.30	(1.12)	8.91	-0.82	(3.01)	13.25
Denmark	-1.15	(2.89)	10.35	-1.66*	(0.80)	2.81
France	-1.27	(2.88)	10.12	-1.78*	(0.75)	2.68
Germany	0.51	(2.97)	11.25	--	--	--
Ireland	-1.68	(2.76)	8.85	-2.19*	(0.97)	2.94
Italy	-2.05	(2.88)	10.71	-2.57	(1.67)	5.31
Netherlands	0.17	(2.95)	12.15	-0.33	(0.33)	4.21
Norway	-1.34	(2.54)	9.02	-1.85	(1.29)	4.41
Spain	-2.01	(2.76)	8.39	-2.52	(1.61)	4.40
Sweden	-0.31	(2.73)	10.70	-0.82	(1.96)	9.71
Switzerland	1.74	(3.26)	10.67	1.23	(1.21)	10.77
United Kingdom	-0.56	(2.91)	13.89	-1.08	(2.32)	10.80
United States	--	--	--	-0.51	(2.97)	11.25

μ , se(μ) and $\sigma_{\Delta s}/\sigma_{fp}$ denote the mean differential from UIP, the standard error of the differentials and the ratio of the spot change variance and the variance of the forward premium.

* significant at the 5% level.

3.2 Single equation results

In this section we replicate tests of UIP. The first part concentrates on the bilateral time series regressions. The second part of this section is devoted to testing UIP over different time periods.

3.2.1 Standard time series tests

In standard time series tests of the unbiasedness hypothesis, the change in the spot rate of currency i expressed against the numeraire currency 0 is regressed on the forward premium:

$$(1) \Delta s_{i0}(t+1) = \alpha_i + \beta_i (f_{i0}(t) - s_{i0}(t)) + \varepsilon_{i0}(t),$$

where ϵ_{0t} is a rational expectations, IID error term. Assuming risk-neutral efficient markets and rational expectations, the null hypothesis for the forward premium to be an unbiased estimator of the future change in spot rate is that the slope coefficient β_1 equals unity and that the constant α_1 equals zero. In Table 2 we report estimated slope coefficients β_1 from (1). Both the US Dollar and the German Mark have been used as numeraire currency.

2 Bilateral time series slope estimates

Slope coefficient estimates for the bilateral time series test of UIP (1) for three sample periods and two numeraire currencies. A slope coefficient that equals 1 supports the null hypothesis that the unbiasedness hypothesis holds.

Currency	US Dollar			DEM		
	1979 1996	1979 1986	1987 1996	1979 1996	1979 1986	1987 1996
Austria	-0.61 (0.89)	-1.47 (1.96)	-0.51 (1.34)	-0.09 (0.43)	0.10 (0.58)	0.05 (0.83)
Belgium	-0.46 (0.80)	-0.87 (0.96)	0.01 (1.27)	0.25 (0.35)	0.26 (0.38)	-1.49 (0.83)
Canada	-1.63 (0.64)	-2.05 (0.86)	-1.02 (1.11)	-1.46 (0.96)	-2.05 (2.33)	-2.03 (1.27)
Denmark	-0.53 (0.68)	-1.32 (1.06)	0.22 (0.76)	-0.26 (0.48)	-0.03 (0.35)	-1.45 (0.76)
France	0.11 (0.91)	0.10 (1.10)	0.04 (1.30)	0.85 (0.43)	1.05 (0.60)	-0.92 (0.53)
Germany	-0.83 (0.84)	-3.14 (2.19)	-0.24 (1.18)	--	--	--
Ireland	0.58 (0.96)	-0.87 (1.08)	1.19 (0.97)	0.83 (0.41)	-0.11 (0.60)	1.08 (0.60)
Italy	0.98 (0.87)	-0.05 (0.91)	3.21 (2.06)	-0.05 (0.35)	0.06 (0.56)	-1.59 (1.22)
Netherlands	-1.71 (0.84)	-5.09 (1.52)	-0.69 (1.30)	0.08 (0.47)	0.00 (0.61)	-0.02 (0.26)
Norway	-1.14 (0.68)	-0.95 (0.81)	-0.85 (1.25)	0.24 (0.45)	1.97 (0.75)	-0.78 (0.46)
Spain	0.91 (0.58)	0.26 (0.51)	2.23 (1.46)	0.98 (0.51)	1.22 (0.78)	-0.19 (0.66)
Sweden	0.43 (0.84)	-1.64 (1.27)	2.15 (0.91)	0.13 (0.56)	0.08 (1.15)	-0.14 (0.52)
Switzerland	-1.57 (0.78)	-3.17 (1.64)	-1.69 (1.42)	-1.41 (0.71)	-1.50 (0.84)	-1.58 (1.53)
United Kingdom	-2.57 (1.05)	-3.50 (1.26)	-1.03 (1.91)	-1.23 (0.79)	-4.13 (1.50)	-0.52 (0.91)

The coefficients are univariate time series estimates of β_1 in equation (1) for exchange rates expressed in terms of the US\$ and the DEM respectively. Robust standard errors are given in parentheses.

From Table 2, we observe that the estimated slope coefficients range from -5.09 to +3.21 and are rarely close to unity. Rejection of the null hypothesis (beta equals one) is not always possible because of the relative large standard errors. However, the same holds for rejecting the hypothesis that beta equals zero. A second observation is that beta estimates show strong instability over the sub-periods. An interesting example is Sweden, which has a beta of -1.64 in the first sub-period and of +2.15 in the second period, against the US Dollar. Overall, we conclude that the standard time series tests generally do not offer significant support for UIP, conform the results found in many previous studies as for example Bilson (1981), Fama (1984), Bekaert and Hodrick (1992) and Lewis (1995).¹⁰ The choice of the numeraire currency appears to be of little importance with respect to this overall qualitative conclusion. However, the exact quantitative results strongly depend on the choice of the numeraire currency.

3.2.2 Outlier tests

Bilson (1981) was the first to focus on the possibility that outlier observations on forward premiums may have different forecasting power than 'normal' ones due to disorderly market conditions. He defines period t as an outlier if the annualized forward premium observed at time t is larger than 10% in absolute terms. Bilson proposes to adjust regression equation (1) as follows:

$$(2) \Delta s_{i0}(t+1) = \alpha_i + \beta_1^N (f_{i0}(t) - s_{i0}(t)) I_{(t \in N)} + \beta_1^O (f_{i0}(t) - s_{i0}(t)) I_{(t \in O)} + \varepsilon_{i0}(t),$$

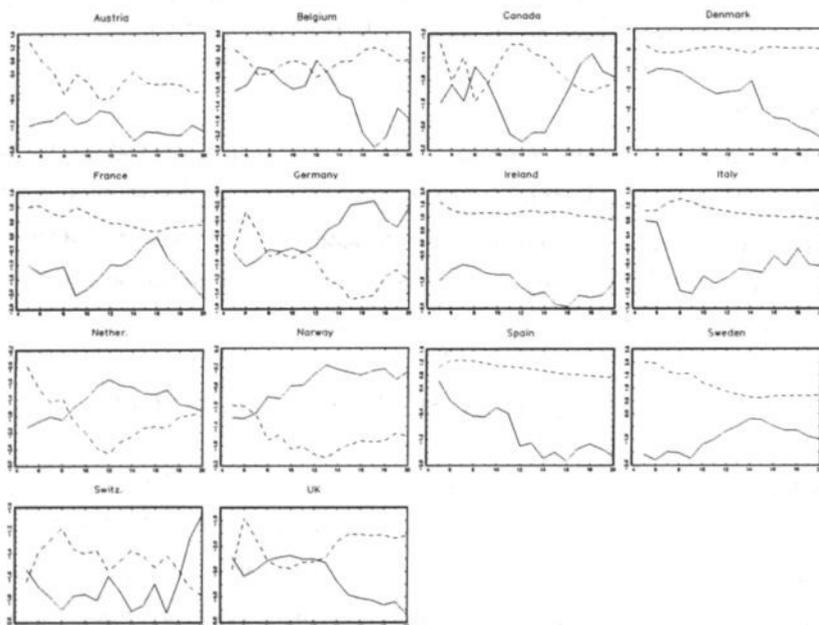
where the superscripts 'N' and 'O' refer to normal-sized and outlier observations of the forward premium respectively, $I_{(t \in N)}$ is an indicator function that equals one (zero) if period t is a normal (outlier) observation, $I_{(t \in O)}$ is defined vice versa and all other symbols have the same interpretation as in equation (1). Bilson defines the observation from period t as an outlier if the forward premium observed at t is larger than 10%. Bilson provides evidence for differences in forecasting power over different time periods. He finds that the forward premium is a better forecast of the future change in the spot rate in normal periods than in outlier periods.

Flood and Taylor (1996) conduct a similar test for UIP. They rank their data with respect to the size of the absolute interest differential and divide their sample over four quartiles depending on the ranking. They show that the deviation from UIP is inversely related to the size of the absolute interest differentials. Interestingly, this is the opposite of Bilson's result. Flood and Rose

¹⁰ We also replicated this test and the tests in the next sections using (overlapping) data sampled at the daily and weekly frequency and for various forward maturities. The results from these studies are qualitatively similar to the results reported here.

(1994) relate outliers to realignment days in the EMS and find that omitting these observations reduces the slope coefficient by 0.35 in a pooled version of the time series test (1).

We replicate the Bilson test for differences in forecasting power of the forward premium between normal and outlier periods by estimating the parameters in equation (2) for each currency. Outliers are defined as the x percent largest absolute forward premiums over the sample, with x ranging from 5 through 20 percent. In Figure 2, the path of the estimates for the slope coefficients over x in both normal and outlier periods are presented for currencies denoted in US Dollars, while Figure 3 provides similar information with the German Mark being the numeraire.

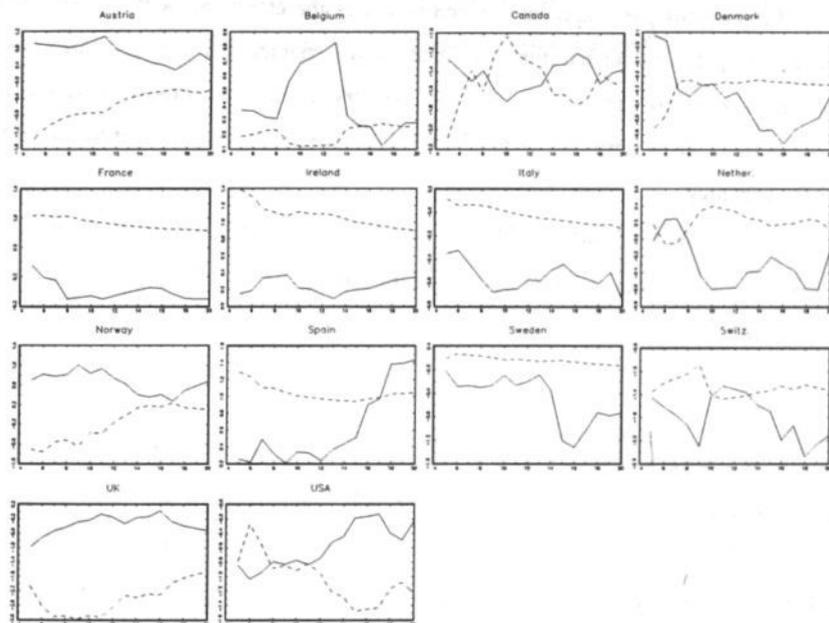


2 Time series slope coefficients estimates – US Dollar

Slope coefficient estimates from the time series model (2) with the US Dollar as numeraire. The drawn line represents the slope coefficient estimates for normal period whereas the dashed line represents the slope in outlier periods. On the x-axis, the percentage outliers in the regression is given.

In each country-specific graph, the drawn line gives the estimates of the slope coefficient in (2) for the normal periods, while the dotted line represents the estimated slope coefficients in outlier periods for each value of x (on the horizontal axis). The figures show a large variation of

results both between countries for the same numeraire and between numeraires for the same country. A few examples may illustrate this. In the case of the French Franc relative to the German Mark, the outlier coefficient is stable and close to 1 over most of the x-range; simultaneously the normal slope coefficient is close to zero. Relative to the US Dollar, both coefficients are negative though close to zero. A comparison between France and Norway for the German Mark as numeraire reveals that Norway's results are roughly the opposite of France: for Norway, the normal slope is close to one and the outlier coefficient is negative. Especially for the German Mark numeraire in Figure 3, most EMS-currencies appear to display similar coefficient patterns. We find that especially the outlier coefficients have a tendency towards unity. Interestingly, this is opposite to Bilson's results. If anything, he finds that UIP holds better in normal periods than in outlier periods. On the other hand, our results are in accordance with Flood and Rose (1994) who do find a 'realignment' effect that supports UIP to some extent and with Flood and Taylor (1996).



3 Time series slope coefficients estimates – German Mark

Slope coefficient estimates from the time series model (2) with the German Mark as numeraire. The drawn line represents the slope coefficient estimates for normal period whereas the dashed line represents the slope in outlier periods. On the x-axis, the percentage outliers in the regression is given.

A well-known caveat applies to the above results. Large standard deviations exist, perhaps due to a number of potential misspecifications given by Lewis (1995) and the small-sample problem referred to by Baillie and Bollerslev (1997). Therefore, a more efficient simultaneous equation variant is in order, where observations are pooled both across time and countries. Bilson, Flood and Rose and Flood and Taylor all increase the efficiency of their tests by conducting results from a pooled dataset. However, none of these models directly account for misspecifications arising from peso-problems, risk premiums, and other explanations as discussed by Lewis (1995). Moreover, the numeraire dependency of our results deserves additional attention. We return to these issues in the next section, where we apply panel models to test UIP.

3.3 The panel methodology

As motivated in the previous section a random time effects panel model is developed that allows for various possibly time-varying and co-existing factors that bias the standard time series tests as discussed in the previous paragraphs and which increases the efficiency of the estimates. In its most general form, the model has three main attractive characteristics. First, currencies may differ in their response to such factors. Second, country-specific and time-specific effects are included as stochastic error terms (random effects). Third, the model is invariant to the choice of the numeraire currency.

We follow Mahieu and Schotman (1994) to decompose the error term $\varepsilon_{i0}(t)$ in the standard time series regression (1) in an error for currency i specifically ($i = 1 \dots n$) and one error term $\varepsilon_0(t)$ for the numeraire:

$$(3) \quad \varepsilon_{i0}(t) = \varepsilon_i(t) - \varepsilon_0(t).$$

The currency-specific error term $\varepsilon_i(t)$ is modeled as the sum of $M+1$ economic, possibly unknown, factors that are mutually independent and orthogonal to the forward premium observed at time t . The first M factors $w_m(t)$ are common to all currencies and each currency i reacts with its own intensity γ_{im} to the m^{th} common shock ($m = 1 \dots M$). The last factor we add, is a currency-specific factor $e_i(t)$. Given this framework, $\varepsilon_i(t)$ becomes:

$$(4) \quad \varepsilon_i(t) = \sum_{m=1}^M \gamma_{im} w_m(t) + e_i(t).$$

Substituting the expression for the currency specific error (4) into the error term (3), the latter is rewritten as:

$$(5) \quad \varepsilon_{i0}(t) = \sum_{m=1}^M (\gamma_{im} - \gamma_{0m}) w_m(t) + e_i(t) - e_0(t).$$

Note from (5) that all exchange rates, expressed against the same numeraire currency 0, contain a common factor $e_0(t)$, which is the currency-specific shock that is linked with the numeraire currency. This common numeraire factor is the basis for the numeraire invariance of the model to test UIP. In the panel literature, this setup is referred to as a random time effects panel model.¹¹ If the error terms in the regression equations (1) or (2) are replaced by expression (5), it can be seen that the resulting model accounts for time varying shocks $w_m(t)$. Without being specific, these shocks may reflect the market risk premiums, realignment effects occurring and possibly misperceptions of future economic developments arising at time t . Since (5) does not analytically model any of these events, it is capable of controlling for multiple events that might co-exist at time t . Note that not all currencies are necessarily affected to the same level by the shocks, since each shock $w_m(t)$ enters the error term with unique sensitivity $(\gamma_{im} - \gamma_{0m})$.

Within the framework of the error term (5), we normalize the variances of the common shocks $w_m(t)$ to unity. Under the assumption that each currency-specific shock $e_i(t)$ has a variance equal to λ_i ($i = 0 \dots n$), the variance of the error term $\varepsilon_{i0}(t)$ equals (the time script is suppressed):

$$(6) \quad \text{var}(\varepsilon_{i0}) = \sum_{m=1}^M (\gamma_{im} - \gamma_{0m})^2 + \lambda_i + \lambda_0,$$

and the covariance between the errors $\varepsilon_{i0}(t)$ and $\varepsilon_{j0}(t)$ of two currencies i and j equals

$$(7) \quad \text{cov}(\varepsilon_{i0}, \varepsilon_{j0}) = \sum_{m=1}^M (\gamma_{im} - \gamma_{0m})(\gamma_{jm} - \gamma_{0m}) + \lambda_0.$$

Consequently, the $(n \times n)$ covariance matrix $\Sigma_0(t)$ of the error term $\varepsilon_{i0}(t)$ can be written as:

$$(8) \quad \Sigma_0 = \sum_{m=1}^M \mathbf{v}_{0m} \mathbf{v}_{0m}' + \Lambda_0 + \lambda_0 \mathbf{1}\mathbf{1}',$$

where \mathbf{v}_{0m} is a $(n \times 1)$ vector with the value $(\gamma_{im} - \gamma_{0m})$ on row i and Λ_0 is a $(n \times n)$ matrix with the variances of the country specific shock λ_i as diagonal elements and zeros elsewhere and $\mathbf{1}$ is a $(n \times 1)$ vector with ones.

In the following, we present the maximum likelihood estimator for the slope coefficients in the discussed random time effects panel model to test the unbiasedness hypothesis. Converting

11 For more detailed information on panels, we refer to Baltagi (1995).

in matrix notation, we define $\Delta S_0(t+1)$ as the $(n \times 1)$ vector with the change in the spot rate of currency i $\Delta s_{i0}(t+1)$ in position i and $F_0(t)$ as the $(n \times 1)$ vector with the forward premium for currency i observed at t in position i . Let $U_0(t)$ be the $(n \times 1)$ vector containing the error term $\varepsilon_{i0}(t)$ defined in (5) in position i . Then in matrix notation, the random time effects model reads:

$$(9) \quad \Delta S_0(t+1) = \beta_0 F_0(t) + U_0(t),$$

in which we assume that β_0 is the same for each currency i expressed in terms of the numeraire currency 0. The maximum likelihood estimator of β_0 equals (again suppressing time subscript):

$$(10) \quad \beta_0 = \sum_{t=1}^T (F_0' \Sigma_0^{-1} F_0)^{-1} (F_0' \Sigma_0^{-1} S_0).$$

One characteristic of the maximum likelihood estimator for the slope coefficient (10) is its invariance to the choice of numeraire currency. To see this, we introduce a $(n \times n)$ transformation matrix P_k with minus ones on row k , ones along the other diagonal elements and zeros elsewhere. If we go from numeraire currency 0 to numeraire k , the transformation is:

$$(11) \quad \Delta S_k(t) = P_k \Delta S_0(t).$$

For the forward premium, we apply a similar transformation to go to numeraire currency k . It can be shown that:

$$(12) \quad P_k' \Sigma_0 P_k = \sum_{m=1}^M v_{km} v_{km}' + \Lambda_k + \lambda_k u' = \Sigma_k.$$

If we rewrite the maximum likelihood estimator in terms of numeraire k and substitute Σ_k by the expression on the left in equation (12), we have the following estimator for β_k :

$$(13) \quad \beta_k = \sum_{t=1}^T (F_k' (P_k' \Sigma_0 P_k)^{-1} F_k)^{-1} (F_k' (P_k' \Sigma_0 P_k)^{-1} S_k).$$

Given the transformation (11) for going from one numeraire to another and given the fact that for matrix P_k holds that $P_k = P_k^{-1}$, we have that expression (13) turns into the estimator for β_0 in (10). Therefore, β_k is equal to β_0 or, otherwise stated, the obtained slope estimate is invariant to the choice of the numeraire currency.

In order to apply the maximum likelihood estimator (10), we distinguish four different structures of the covariance matrix Σ_0 . First, we estimate Σ_0 without any restrictions as in (8), allowing for an unknown but differential impact of the common factors on the n currencies.

Therefore, the unrestricted covariance matrix is an $(n \times n)$ matrix with $\frac{1}{2}(n+1)n$ free parameters. The second covariance matrix only allows for one single common factor ($m = 1$) in (8), that influences the n currencies with different intensities. For the third covariance matrix, we assume that all currencies react with the same intensity to the common factors, thus $(\gamma_{im} - \gamma_{0m}) = (\gamma_{jm} - \gamma_{0m}) = 0$ and therefore $\nu_{0m} = 0$ in (8). In effect, the exchange rates are not affected by the common shocks, but are only influenced by idiosyncratic noise. The covariance matrix reduces to

$$(14) \quad \Sigma_0 = \Lambda_0 + \lambda_0 \mathbf{u}'.$$

The fourth and last covariance matrix we implement is the one motivated by Koedijk and Schotman (1990). They impose the additional restriction on (14) that all country specific shocks $\epsilon_i(t)$ have equal variance such that $\lambda_0 = \lambda_1 = \dots = \lambda_n = \sigma^2/2$. Given this restriction the covariance matrix (14) reduces to:

$$(15) \quad \Sigma_0 = \sigma^2 (\mathbf{I} + \mathbf{u}') / 2.$$

where \mathbf{I} is a $(n \times n)$ identity matrix. It can be seen from (15), that this covariance structure is equal for any choice of numeraire currency. Since (15) is completely specified up to one parameter, the GLS squares estimator is directly applicable.

3.4 Empirical results

3.4.1 Results from standard panels

Table 3 contains estimates for the slope coefficient β_0 of the random effects panel model for the above four specifications of the structure of the covariance matrix, using the maximum likelihood estimator given in (10). The slope estimates across the four specifications of the covariance matrix range from 0.43 through 0.55 but do not differ significantly. We conclude that in all cases the estimated slope is quite insensitive to the restrictions on the covariance matrix so that we feel rather confident with respect to the robustness of the results. Though UIP still has to be rejected, these results show a significant improvement of the slope coefficient. It confirms the idea that a time series approach is relatively inefficient in coping with possible biases from the existence of risk premiums or systematic forecast errors. Also, it suggests that such biases may be the result of several simultaneously existing and interrelated factors and of small samples. The presented panel model accounts for such effects by modeling them as random common

factors or shocks, without the need to make identifying assumptions on their structural characteristics.

3.4.2 *The importance of outliers*

As a further step in the analysis of the deviations from UIP, we again distinguish between large and small forward premiums and estimate the numeraire invariant panel version of equation (2). Again, we estimate slope coefficients both for periods with normal-sized forward premiums and abnormal sized or outlier forward premiums. The selection of outliers is based on the cross-sectional standard deviation of the forward premium. In this respect, we diverge from Bilson (1981) since he selects outliers for each currency individually. The advantage of our approach is that the cross sectional standard deviation is numeraire invariant and will therefore generate equal slope coefficients for every choice of the numeraire currency.¹² For the estimation, we use the most restricted covariance structure (15). The first row in Table 4 contains both slope coefficients, where the outlier observations cover 5% of the data.

3 Panel slope coefficients

Estimated slope coefficients (10) for the random effects panel model. We distinguish between four different covariance matrices. 1 corresponds with an unrestricted covariance matrix (8). 2 imposes a restriction on (8) that m equals 1. 3 corresponds with the matrix (14) in which all currencies are assumed to react with the same intensity to a shock. 4 corresponds with (15) that imposes the extra restriction that all country specific shocks have equal variances.

Covariance structure	1	2	3	4
β	0.50 (0.06)	0.54 (0.06)	0.55 (0.06)	0.43 (0.10)
LogLik	-1724	-1940	-2275	-3198

LogLik is the log likelihood. Robust standard errors correcting for heteroskedasticity and autocorrelation are presented in parentheses (see Baltagi (1995)).

The results in Table 4 show that the relation between the forward premium and the future spot change differs between periods with normal observations and periods indicated as outliers.

¹² The cross-sectional standard error is not one-to-one related with the level of the average absolute forward premium, although both are highly correlated. The absolute forward premium however is not numeraire invariant. Results are similar when we consider outliers of the absolute forward premium for individual countries.

UIP cannot be rejected in outlier periods; the estimated slope coefficient is not significantly different from one. In contrast with the estimates from the standard time series tests reported in Table 2, the slope coefficient is significantly different from zero as well. In normal periods, the slope coefficient is 0.24 and both significantly different from zero and one. This is consistent with the results found by Flood and Rose (1994) who find a reduction of 0.35 as soon as EMS realignments are deleted from the sample. Flood and Taylor (1996) similarly find that large absolute interest differentials lead to a slope estimate of approximately 0.55. A significantly negative slope coefficient equal to -2.99 results in periods with small absolute interest differentials. Interestingly, both the results of Flood and Rose and Flood and Taylor are opposite to what was found by Bilson (1981), who only studied data from 1974 through 1980.

4 Outlier and normal estimates

UIP and 5% outliers.

	Normal period	'outlier' period
EMS realignments are included		
β	0.24 (0.11)	1.04 (0.20)
EMS realignments are excluded		
β	-0.06 (0.11)	0.58 (0.21)

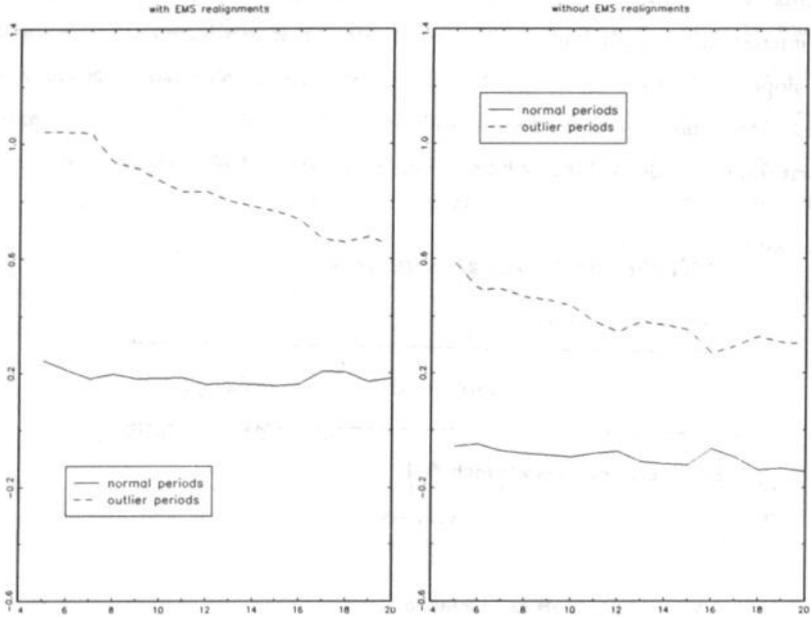
Robust standard errors are in parentheses.

Outliers are defined as the 5% largest observations in terms of the cross sectional average absolute value of the forward premium over the time period January 1997 through March 1996.

The selected 5% outlier observations are July 1979, August 1981, September 1981, March 1982, October 1987, November 1982, December 1982, September 1992, December 1992 and January 1993 for the sample where all EMS outliers are included. The selected 5% outlier observations are July 1979, August 1981, March 1982, October 1982, November 1982, December 1982, September 1992, November 1992, December 1992 for the sample where all EMS outliers are included.

Like in Figures 2 and 3, we graphically display the slope coefficients of both the normal observations and the outlier slope coefficient as a function of the percent outliers used denoted as x , where x ranges from 5 to 20 percent. The left graph in Figure 4 presents these slope estimates. As can be seen from Figure 4, the coefficient is close to unity and UIP seems to hold,

when only the largest observations are taken into account. The slope coefficient on the remaining normal observations is small. The results from Table 4 and Figure 4 indicate that the relationship between the forward premium and the future change in the spot rate, as suggested by UIP, tends to be particularly strong in periods with large forward premiums, which are highly correlated with market volatility.



4 Normal and outlier slope coefficients

Both figures show the slope estimates of the numeraire invariant random time effects panel model (10) both for normal periods and periods defined as outliers (on the y-axis) as a function of the number of outliers used (in percentages on the x-axis). A period is an outlier if the cross-sectional average absolute forward premium belongs to the 5% largest over the whole sample period. In the left figure, we present the estimates over the whole sample from January 1979 through April 1996. In the right picture, we excluded all EMS realignment observations from the sample.

Several explanations are worth mentioning. The first explanation would be in line with Baldwin (1990). He argues that the presence of small transaction costs may lead to a large inactivity band within which no arbitrage takes place. Baldwin implements his theoretical model by making various assumptions on the values of the model's parameters to arrive at the conclusion that interest differentials between one and four percent on an annual basis may

persist without introducing arbitrage flows. A second explanation would be the existence of peso-problems. The recent experience of exchange rates within the EMS is a case in point. Because of systematically expected devaluations, peso-problems may be a frequent phenomenon there. As discussed, Flood and Rose (1994) test UIP with a pooled regression and show that the estimated slope coefficient equals approximately 0.6 in normal periods and decreases to almost 0.2 as soon as EMS realignments are excluded. They conclude that the peso-problem for EMS exchange rates leads to a bias of 0.35.

3.4.3 *Outliers and EMS realignments*

To explore the EMS phenomenon in more detail, we repeat the exercise from the previous paragraph with realignments removed from the sample. However, we first recalculate the average deviations from UIP from Table 1 without the realignments and present them in Table 5. The average deviations are generally larger (though not always significantly) in magnitude than those in Table 1. Especially for the German Mark as numeraire, the significance of average UIP deviations increases. This well-known result is another stylized representation of the peso-problem. In non-realignment periods systematically negative forward premiums exist for many EMS currencies against the German Mark due to an expected average devaluation over a longer period of time. However, this devaluation only occurs infrequently at realignments. At that time, the deviation from UIP is considerably positive and should, according to UIP, approximately compensate for the earlier negatives.

Following Flood and Rose (1994), we now investigate whether the previous 'outlier' results can be solely attributed to EMS realignments. On the right of Figure 4, we again graph the slope coefficients for normal days and for outlier observations as a function of different outlier percentages. The line that contains the coefficients of the outlier observations lies substantially above the line that contains the 'normal' coefficients. From Table 4 we observe clearly that even without EMS realignments the slope coefficient on the outlier observations at the 5 percent level is 0.58. The coefficient on the 'normal' observations drops to -0.06. These results indicate that large forward premiums predict exchange rate changes and these large premiums are not solely related to EMS realignments.

5 UIP differentials

Average differentials from UIP obtained from a sample in which all EMS realignment observations are excluded.

Currency	US Dollar		DEM	
	μ	se(μ)	μ	se(μ)
Austria	1.18	(3.06)	-0.68*	(0.32)
Belgium	0.56	(3.01)	-1.31*	(0.66)
Canada	-0.67	(1.11)	-2.54	(3.16)
Denmark	-0.42	(2.98)	-2.29*	(0.78)
France	-0.93	(2.94)	-2.80*	(0.65)
Germany	1.87	(3.08)	--	--
Ireland	-1.28	(2.82)	-3.16*	(0.83)
Italy	-2.69	(2.88)	-4.56*	(1.51)
Netherlands	1.53	(3.04)	-0.33	(0.31)
Norway	-0.81	(2.65)	-2.68*	(1.35)
Spain	-3.06	(2.80)	-4.92*	(1.53)
Sweden	-0.74	(2.71)	-2.61	(1.92)
Switzerland	3.74	(3.35)	1.86	(1.25)
United Kingdom	-0.51	(2.95)	-2.37	(2.32)
United States	--	--	-1.87	(3.08)

* Rejected at the 5% level.

μ and se(μ) denote the mean differential from UIP and the standard error of the differentials and the ratio of the spot change variance. All units are annualized percentages.

3.5 Concluding remarks

Uncovered interest parity has been rejected statistically in most empirical literature. The general consensus is that these rejections do not imply that UIP is invalid. Instead, many believe that UIP holds but that it is undetectable for a wide array of causes. Examples of such arguments are the existence of (time varying) risk premiums, peso-problems, market inefficiencies and neglected persistent autocorrelation in the forward premium. In addition, small sample problems have been blamed for the lack of confirmative results.

In this chapter, we shed new light on the validity of UIP. We adopt an alternative methodology that allows us to avoid most of the above-mentioned problems. Our panel methodology is capable of controlling for the suggested interrelating explanations by modeling them as a random time effect. Moreover, the small sample problem is reduced to a large extent

by pooling across many currencies. In addition, the UIP tests become invariant to the choice of the numeraire currency.

Using the panel approach, we find that the rejection of UIP is not as severe as is commonly found with standard time series tests. The estimated slope coefficients are significantly larger than zero and equal 0.5 on average. However, the estimated slope coefficient is significantly different from 1 and UIP still needs to be rejected.

To examine the remaining source of the rejection, we distinguish between 'normal' and 'abnormal' observations, following Bilson (1981). Using the same panel method again, we conclude that UIP almost perfectly holds for periods where the absolute forward premium is large. The estimated slope coefficient is not significantly different from 1 and UIP cannot be rejected in these abnormal periods. In normal (less volatile) time periods, the estimated slope coefficient is reduced with 0.8 and UIP needs to be rejected strongly.

Flood and Rose (1994) find similar results and attribute their findings to the effect of EMS realignments. We show that UIP is more valid in extreme periods than in normal times, even in samples from which the EMS realignments are excluded. This result indicates that the differences in forecasting power of the forward premium over normal and abnormal periods are more than a realignment effect. These findings point to the potential importance of peso-problems and/or inactivity bands as explanations for the rejection of UIP in the foreign exchange market.

4 Fat Tails in Small Samples¹³

It is a well-known stylized fact that the distributions of returns on many financial assets are fat tailed¹⁴. This implies that large movements in the asset's value occur frequently and the observed tail fatness of financial assets has important implications for asset return models. Typically the normal distribution is assumed to model the expected returns. For instance, the most popular parametric Value at Risk method infers the worst expected loss over a given time horizon from the normal distribution. Furthermore, a distribution dependent measure like the variance is typically used in studies where uncertainty with respect to financial asset movements is measured. If the distribution of returns is heavily non-normal, means and variances do not adequately describe the return distributions; returns and the amount of risk faced on investments in financial assets must reflect the large probability on large fluctuations. The degree of tail fatness contains important information for investors.

In this chapter we concentrate on measuring the degree of tail fatness in such that it can be incorporated explicitly in asset return models. We do not try to fully specify the underlying true parametric distribution in order to describe the tail fatness of asset returns accurately. We take a more modest approach and concentrate on the tail shapes of the empirical distribution by estimating the tail index α . The tail index is a measure of the amount of tail fatness of the distribution under investigation and fits within extreme value theory (EVT). This investigates the distribution of the tail observations. From EVT it is known that the shape of the tails follow a Pareto law for a common class of fat-tailed

¹³ This chapter is based on: R. Huisman, K. Koedijk, C. Kool, and F. Palm, 1997, 'Fat Tails in Small Samples'.

¹⁴ See Jansen and de Vries (1991), Koedijk, Stork and de Vries (1992), Koedijk and Kool (1994), Loretan and Phillips (1994) and Keams and Pagan (1997) for various contributions.

distributions. The limit law for the tails is characterized by the tail index α , which is one to one with the number of existing moments.

Hill (1975) proposed the best-known and most often applied estimator for the tail-index. The Hill estimator can be easily implemented and is asymptotically unbiased. However, it is biased in small samples. Several alternative estimators have been proposed since then. Pictet, Dacorogna, Müller and de Vries (1996) give an overview and study the performance of a number of these estimators. They conclude that many estimators perform rather well for extremely large sample sizes, but that all suffer severely from small sample bias. As a result, the empirical applicability of the current class of tail-index estimators is limited to cases where a large sample is available, either in the form of high-frequency data or in the form of an extremely long sampling period. In many practical cases, neither of these two conditions is fulfilled. Moreover, even when a long sample is available, it may be interesting to split the sample and analyze whether the tail structure of the sample has changed over time. The relevance of a method to obtain correct information about tail behavior in small samples is clear.

An important part of the bias in these estimators stems from the selection of the exact number of tail observations to include in the estimation process. If one includes too many observations, one reduces the variance of the estimate but takes too much center observations (i.e. non-tail information) into account, resulting in a selection bias. If one includes too few observations, the variance of the estimate becomes too large. Hill and other estimators do not provide an optimal number valid for all sample sizes and some have used heuristic methods for approximation¹⁵. Beirlant, Vynckier and Teugels (1996, BVT) were the first to present a method to select the optimal number of tail observations to include in the Hill estimator. Their method thus eliminates the selection bias.

This chapter presents an alternative methodology to estimate the tail-index. Contrary to Hill and BVT the method does not condition its tail-estimate on one specific number of tail observations. The method exploits information obtained from a multiple of Hill estimates, each conditioned on a different number of tail observations. The result is a weighted average of a set of conventional Hill estimators, with weights obtained by simple least squares techniques.

¹⁵ See for example Jansen and de Vries (1991).

4.1 Methodology

In this section, we first discuss the conventional Hill estimator and its sensitivity to both the sample size and the number of tail observations (k) used in the estimator. We conclude that the method suffers from severe small-sample bias. Subsequently, we propose a simple alternative method to reduce the small sample bias in tail-index estimates.

4.1.1 The conventional Hill (1975) estimator

Suppose a sample of n observations is drawn from some unknown fat-tailed distribution. The elements of the sample are assumed to be positive. Let the parameter γ be the tail-index of this distribution (per definition, γ equals $1/\alpha$, where α refers to the maximum number of existing finite moments), while x_i is the i^{th} increasing order statistic ($i = 1..n$). Then Hill (1975) proposes the following estimator for the tail-index:

$$(16) \quad \gamma(k) = \frac{1}{k} \sum_{j=1}^k \ln(x_{n-j+1}) - \ln(x_{n-k}),$$

where k is a pre-specified number of tail observations to include ($k = 1 .. n-1$).

Various methods have been used to select k . From a simulation study, Loretan and Phillips (1994) show that k should not exceed 10% of the sample size. They use different values for k to obtain a range of tail estimates for stock and exchange rate returns. Here, a drawback is the fact that it remains unclear which estimate out of the range of tail estimates is the correct one. Other studies obtain k from a Monte Carlo simulation¹⁶. They draw n innovations from a distribution (for example Student- t) with *a priori* known tail-index, and select that k that minimizes the MSE of the Hill estimates.¹⁷ In this case, the results depend on the assumed underlying unknown distribution, so that the results need to be interpreted cautiously.

In general, the choice of k is crucial to obtain unbiased estimates of the tail-index. Consider the following class of distribution functions:

$$(17) \quad F(x) = 1 - ax^{-\alpha} (1 + bx^{-\beta}),$$

¹⁶ See Koedijk, Schafgans and de Vries (1990), Jansen and de Vries (1991), Koedijk, Stork and de Vries (1992) and Koedijk and Kool (1994).

¹⁷ Koedijk and Kool (1994) select k equal to 31 for two-sided tail indices and equal to 20 for single sided indices in a sample consisting of approximately 240 observations.

where α and β are positive and a and b are real numbers. Equation (17) provides the second order asymptotic expansion of the cumulative distribution function (c.d.f.) for almost any fat-tailed distribution. For this class of distribution functions, it can be shown that the asymptotic expected value of the Hill estimator for a given k is approximated by

$$(18) \quad E(\gamma(k)) \approx \frac{1}{\alpha} - \frac{b\beta}{\alpha(\alpha + \beta)} a^{-\frac{\beta}{\alpha}} \left(\frac{k}{n}\right)^{\frac{\beta}{\alpha}}$$

and the asymptotic variance of the Hill estimator by

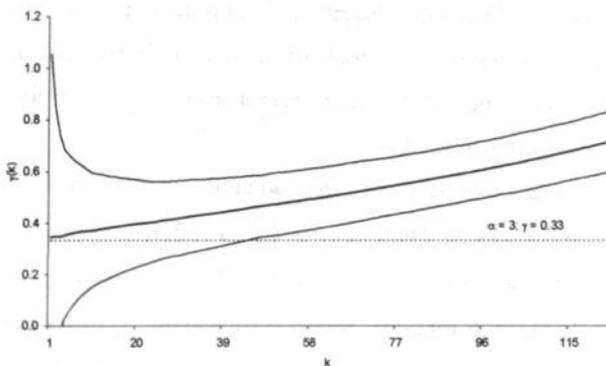
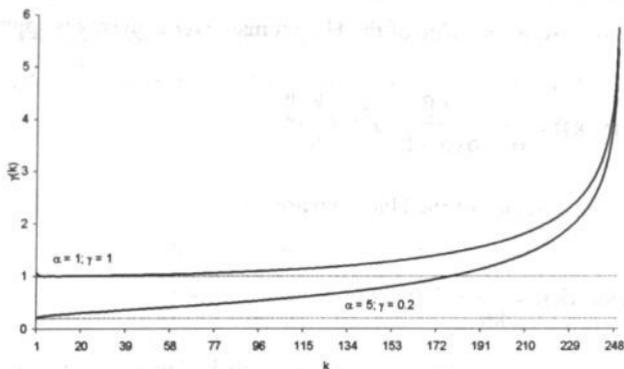
$$(19) \quad \text{Var}(\gamma(k)) \approx \frac{1}{k\alpha^2}.$$

According to equation (18), the bias increases with k . Therefore, a small k is optimal from the perspective of unbiasedness. Equation (19), on the other hand, shows that the estimator's variance decreases with k . A trade-off between unbiasedness and uncertainty is clearly apparent. Minimization of the mean squared error as discussed above trades off some bias against a reduction in variance.

The following simulation experiment offers additional insights in this trade-off. In Figure 5, we present the average Hill estimates (γ) for different values of k over 1500 simulated samples. The value of k ranges from 1 to 249. Each sample consists of 250 observations from Student-t distributions with α equal to 1 and 5. Absolute values of the observations are taken to be able to use both tails simultaneously in the estimation procedure. The tail-index of the distribution γ (equal to $1/\alpha$) now is a priori known since α equals the number of degrees of freedom of the Student-t distribution.

From the top graph in Figure 5, we observe that for small values of k , the Hill estimate $\gamma(k)$ almost equals the true tail-index $1/\alpha$. The bias increases with k , as suggested by (18). For example, the upper line represents the estimates of γ for different k from the Student-t distribution with one degree of freedom. In the limit, when k approaches zero, the line is seen to intersect the vertical axis exactly at 1. The bottom-graph shows the pattern of the γ -estimates for the Student-t distribution with α equal to 3 in more detail. Here the average estimated γ are presented within a band of two standard deviations (measured cross-sectionally) for k ranging from 1 through 125. The trade-off is clearly visible: for small k , accurate (unbiased) estimates of the tail-index result on average. However, individual γ -estimates can diverge considerably from the

average due to the high variance. Increasing k leads to a higher average bias, but less individual outlying estimates.



5 Hill estimates for the Student-t distribution

These graphs present the pattern of the Hill estimates as a function of k . The patterns are the result of a simulation study consisting of 1500 samples with 250 observations drawn from a Student-t distribution. The top graph contains the average Hill estimates for two Student-t samples with α 's equal to 1 and 5. The bottom graph shows the Hill pattern of the Student-t distribution with α equal to 3 within a ± 2 standard deviation interval for $k = 1, \dots, 125$.

It has been proved widely that the distribution of financial returns is generally fat-tailed and may be approximated by a Student-t or Pareto distribution. Most empirical studies on tail fatness of financial returns thus suffer from the above trade-off problem. Generally, a single "optimal" k is selected. As should be clear now, in fact a bias exists for any k exceeding zero. This raises the suspicion that tail fatness of financial returns in the

past has been exaggerated due to poor estimates. One way to deal with this problem is to increase the number of observations. Equation (18) shows that with increasing n the importance of the bias for any given k decreases. This partly explains the increased use of high-frequency data in tail-index research, see for instance Dacorogna, Müller, Pictet and de Vries (1995). In practice, of course, the large number of observations that can be obtained from high frequency data is often lacking.

4.1.2 *An alternative approach*

To improve on the conventional Hill estimator, we propose to exploit an important characteristic of the bias function that is exemplified by figure 5. For values of k smaller than some threshold value κ (for example for k smaller than or equal to $n/2$ in this case), γ -estimates are seen to increase almost linearly in k ¹⁸. Rewriting equation (18) with the bias-term replaced by a bias function $f(k)$ leads to:

$$(20) \quad \gamma(k) = \beta_0 + f(k) + \varepsilon(k),$$

where the true tail-index equals β_0 and $\varepsilon(k)$ is an idiosyncratic error term. Due to the fact that the bias in the Hill estimate can be approximated by a linear function of k , for $k \leq \kappa$, equation (20) can be transformed as follows:

$$(21) \quad \gamma(k) = \beta_0 + \beta_1 k + \varepsilon(k), \quad k = 1 \dots \kappa.$$

We have also estimated nonlinear regression models based on the functional form for the asymptotic bias in (18). The results were much worse than those obtained by the linear specification in (21) for $k \leq \kappa$.

Instead of selecting one optimal k to estimate the tail-index of the distribution under consideration, we propose to compute $\gamma(k)$ for a range of values for k from 1 to κ . Subsequently, the vector of computed $\gamma(k)$'s is used to estimate equation (21). In subsection 4.1.1, we already argued that an unbiased estimate of γ can be obtained only for k approaching zero. Evaluation of equation (21) for k approaching zero, yields an optimal unbiased estimate of γ equal to the intercept β_0 . Applying this procedure circumvents the bias-variance trade-off, since we use the information from a whole range

¹⁸ Unreported simulation results have shown that a similar linear pattern is observed for other fat-tailed distributions as the Pareto and Burr distributions.

of conventional Hill estimates for different values of k to obtain an estimate for the tail-index.

Roughly speaking, the procedure involves fitting a line function through the linear part of the bias function as shown in figure 5. The threshold level κ must be set such that the linearity applies. Although we shall set κ equal to half of the sample, one should check this carefully in practice. Especially in larger samples, setting κ equal to half of the sample size might lead to poor estimates.

4.1.3 *Econometric considerations*

Although the coefficients in (21) can be estimated by Ordinary Least Squares (OLS), two issues complicate the procedure. First, equation (19) indicates that the variance of Hill estimates $\gamma(k)$ is not constant for different k , so that the error term $\varepsilon(k)$ in equation (21) is heteroskedastic. As an alternative, we suggest a Weighted Least Squares (WLS) approach to correct for this form of heteroskedasticity. Second, an overlapping data problem exists due to the construction of $\gamma(k)$. The variables $\gamma(k)$ are correlated, in terms of k , since estimates $\gamma(k)$ and $\gamma(m)$ where $k \neq m$ are based on $1 + \min(k, m)$ common observations, see equation (16). Consequently, traditional standard errors computations both for the OLS and WLS estimates are inappropriate. Therefore, we provide an appropriate alternative to calculate correct standard errors using the asymptotic normality of order-statistics. Notice also that OLS and WLS do not yield fully efficient estimates of β_0 . Feasible GLS would yield a more efficient estimate but it has the drawback of requiring an estimate for the full error covariance matrix.

4.1.4 *Point estimates*

For the OLS based estimator, we first write (21) in the following matrix notation

$$(22) \quad \gamma = Z\beta + \varepsilon,$$

where Z is a $(\kappa \times 2)$ matrix with ones in the first column and the vector $\{1, 2, \dots, \kappa\}'$ in the second. The OLS estimate for β equals

$$(23) \quad b_{OLS} = (Z'Z)^{-1}Z'\gamma,$$

where the estimated tail-index γ equals the first element of the vector b_{OLS} . This OLS procedure does not account for heteroskedasticity in the error term ε . Equation (19)

reveals that the variance of the Hill estimator is inversely related to k . To correct for this form of heteroskedasticity, we propose to apply WLS with a $(\kappa \times \kappa)$ weighting matrix W . W has $\{\sqrt{1}, \sqrt{2}, \dots, \sqrt{\kappa}\}$ as diagonal elements and zeros elsewhere. Transformation of equation (22) through pre-multiplication with matrix W then results in:

$$(24) \quad W\gamma = WZ\beta + W\epsilon,$$

where $W\epsilon$ is homoskedastic. The following WLS-estimate for β then results.

$$(25) \quad b_{WLS} = (Z'W'WZ)^{-1}Z'W'W\gamma,$$

Again, the estimated tail-index γ equals the first element of the vector b_{WLS} .

4.1.5 Standard errors

The Hill estimates $\gamma(k)$ are highly autocorrelated for different k due to the use of common observations. Neither the OLS nor the WLS estimation takes into account the issue of the corresponding autocorrelation in the error term. Consequently, the formulae for the usual standard errors are inappropriate. Here, we derive appropriate standard errors for both OLS and WLS based estimators using the asymptotic normality of order-statistics. Although asymptotic characteristics might not be applicable in small samples we will present evidence that the resulting standard errors are quite accurate in the next section.

Let y be the $((\kappa+1) \times 1)$ vector of increasing order statistics $\{y_{n-\kappa}, \dots, y_n\}$ with $y_i = \ln(x_i)$ for $i = n - \kappa \dots n$. Since the Hill estimator (16) is a linear combination of the y_i 's we can express the vector γ consisting of $\gamma(k)$, $k = 1, \dots, \kappa$, as $\gamma = Ay$ for some $((\kappa \times (\kappa + 1))$ transformation matrix A . Let Σ be the covariance matrix of the order statistics contained in y . Then $\Omega = A\Sigma A'$ is the covariance matrix for the set of Hill estimates in γ . Having Ω , we obtain the standard error of the modified Hill estimates as follows. For the OLS based estimator (23), the covariance matrix of b_{OLS} can be obtained from

$$(26) \quad \text{cov}(b_{OLS}) = (Z'Z)^{-1}Z'\Omega(Z'Z)^{-1}.$$

For the WLS based estimator given in (25), the covariance matrix of b_{WLS} is given by

$$(27) \quad \text{cov}(b_{WLS}) = (Z'W'WZ)^{-1}Z'W'W\Omega W'WZ(Z'W'WZ)^{-1}.$$

To compute the covariance matrix in (26) or (27), the covariance matrix Ω must be specified, or alternatively, A and Σ must be given. The $((\kappa \times (\kappa+1)))$ matrix A can be easily derived from the conventional Hill estimator (16) for the different values of k :

$$(28) \quad A = \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 & -1 & 1 \\ 0 & \cdots & 0 & 0 & -1 & 1/2 & 1/2 \\ 0 & \cdots & 0 & -1 & 1/3 & 1/3 & 1/3 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & \cdots & 1/\kappa & 1/\kappa & 1/\kappa & 1/\kappa & 1/\kappa \end{bmatrix}$$

To obtain the covariance matrix of the order statistics Σ , we use the fact that increasing order-statistics z_i , $i = 1, \dots, n$, are asymptotically multivariate normally distributed¹⁹ with mean μ_i and the covariances between order-statistics z_i and z_j equal to v_{ij} where

$$(29) \quad \mu_i = F_z^{-1}(p_i)$$

and

$$(30) \quad v_{ij} = \frac{p_i(1-p_j)}{nf_z(\mu_i)f_z(\mu_j)} \text{ for } i \leq j.$$

Here, p_i is approximated by i/n , $F_z(z)$ denotes the cumulative distribution function of z and $f_z(z)$ denotes the probability density function of z . Since any fat-tailed distribution is (approximately) Pareto-distributed far in its tails, we propose to use a Pareto-distribution in equations (29) and (30) for our application. That is, we assume that the x_i 's on which the modified Hill estimator is ultimately based, are drawn from a Pareto distribution for $i = n - \kappa, \dots, n$. In that case, the cumulative distribution function of x is given by:

$$(31) \quad F_x(x) = 1 - x^{-\alpha}$$

¹⁹ See e.g. Cox and Hinkley (1974)

for $\alpha > 0$. We know that the Hill estimator is a linear combination of the order-statistics y_i , where y_i equals the natural logarithm of x_i . From (31), it follows that $F_y(y)$ -- the c.d.f. of y -- equals:

$$(32) \quad F_y(y) = 1 - (e^y)^{-\alpha}.$$

The distributional assumptions made here pertain only to the $\kappa+1$ tail observations used in the modified Hill estimator, thus for $i = n - \kappa, \dots, n$. Renumbering such that $i = 1, \dots, \kappa + 1$ and substituting the Pareto specification (32) in the covariance expression (30), we can rewrite (30) as

$$(33) \quad v_{ij} = \frac{i}{\alpha^2 n(n-i)} \text{ for } i \leq j \text{ and}$$

$$(34) \quad v_{ij} = \frac{j}{\alpha^2 n(n-j)} \text{ for } i > j.$$

Given the covariance expressions (33) and (34), Σ - the covariance matrix of the order-statistics y - is fully defined. Subsequently, Ω is obtained from $A\Sigma A'$. Substitution in (26) and (27) leads to appropriate covariance matrixes under OLS and WLS, respectively.

4.1.6 Testing whether two tail-indexes are equal

When one has obtained tail-index estimates an interesting test is to see whether several tail-index estimates differ significantly from each other or not. Two examples will show up in the next chapter. First, one could compare the left tail-index with the right one. If these differ significantly then one couldn't use a tail-symmetric distribution like the Student-t distribution to model the complete empirical distribution. Second, one could examine the tail-indexes of two different distributions. In the next chapter for example, we shall compare tail-index estimates of floating exchange rate regimes with tail-index estimates obtained from non-floating EMS exchange rates.

One way of testing the hypothesis that two tail-indexes are equal is to compute the differences between these tail-index estimates and divide by its asymptotic standard errors. If we assume that the modified Hill estimators (23) and (25) are asymptotically normal, the difference between the two estimates divided by the square root of the variance of this difference is a Student-t distributed test statistic. The single unresolved

issue is to calculate the variance of this difference. We then need the covariance between the two tail-index estimates as input, since the necessity of using order-statistics to obtain tail-index estimates induces a correlation between two tail-index estimates that are obtained from the same sample²⁰.

The covariance between two tail-index estimates can be calculated using the variance specification in section 4.1.5. Let $\text{cov}(b^1, b^2)$ be the covariance-matrix of the elements in the vectors b^1 and b^2 of estimates that are obtained from (23). Given the OLS estimator (23) and the fact that we can write a vector of Hill estimates γ as Ay (see section 4.1.5), we obtain the following expression for the covariance matrix:

$$(35) \quad \text{cov}(b^1, b^2) = (Z'Z)^{-1} Z' A \text{cov}(y^1, y^2) A' Z(Z'Z)^{-1},$$

where the vectors y contain the order-statistics used in the respective tail-index estimates. Note that the matrices Z and A are the same for both estimates if the estimates are obtained using the same threshold level κ in both cases.

4.2 Simulation results

In this section, we test the OLS and the WLS based modified Hill estimators (23) and (25) using simulated data. In particular, we apply the modified Hill estimator to obtain tail-index estimates for relatively small samples with observations drawn from Student-t, a modified Pareto, and Burr distributions respectively.

In the first simulation study we draw 2000 samples that consist of n Student-t innovations each. The sample size n equals 100, 250, 500 or 1000 observations. Student-t distributions with degrees of freedom ranging from 1 to 5 are used. Note that the tail-index α is equal to the number of degrees of freedom of the generating Student-t distribution. For each of the 2000 samples we calculate the tail-index using both the OLS and WLS based modified Hill estimator. The average over the 2000 estimates is reported in table 6 together with the appropriate standard errors based on the average estimate²¹. For comparison, the cross-sectional standard deviations of the point estimates are

²⁰ If the two tail-index estimates are obtained from two independent samples, this covariance is zero of course.

²¹ Since calculating the standard errors is time consuming, we have not calculated the standard errors for each single estimate. Instead, we provide the standard error based on the average estimate over the simulations. This can be done since the formulae for the standard error presented in paragraph 4.1.5 only depend on the estimated tail-index, the number of observations and κ and not on any other sample characteristics that may vary over the simulation. We performed a simulation study that showed that using this average estimate does lead to an adequate estimate of the standard error.

reported as well. For now, κ equals half the sample size. In a subsequent experiment, we will show that the results are quite robust with respect to the choice of threshold κ for these sample sizes.

Table 6 shows that the remaining bias in the modified Hill estimator is small for Student-t distributions, even in small samples. A decrease in sample size from 1000 to 100 hardly affects the average value of the estimated tail-index. A decreasing sample size does increase the average standard error, however. The bias is smallest when the true α is low, that is, for relatively fat-tailed distributions. For α 's exceeding 4, the bias increases probably due to the assumption of strong tail fatness underlying the Hill estimator. Fortunately, the region in which our estimator exhibits the least bias, coincides with the economically most relevant range. Most reported tail-index estimates for stock and exchange rate returns are in this range.²² The computed standard errors based on (26) and (27) for OLS and WLS respectively are close to the cross-sectional standard deviation of the estimates, which supports the appropriateness of the formulae presented in paragraph 4.1.5 for the standard errors.

In tables 7 and 8, we present average tail-index estimates for samples drawn from the modified Pareto distributions and a Burr distribution, respectively²³. In the latter case, α equals 4 (γ equals 0.25) by definition. For both types of distributions, the performance of the modified Hill estimator is satisfactory, though marginally less than for the Student-t.

²² See Loretan and Phillips (1994) for evidence on both stock and exchange rate returns.

²³ Casper de Vries has shown that de Hill estimator is unbiased for any choice of k for the Pareto distribution. Not reported simulation tests indicate that the modified Hill estimator is also unbiased for these distributions. However, we use a modified Pareto distribution here, which is often also referred to as Pareto. The cdf for the modified Pareto distribution reads $F(x) = 1 - (1+x)^{-\alpha}$.

6 Tail-index estimates for Student-t samples

This table provides γ -estimates of the tails of Student-t samples for four sample sizes and 5 numbers of degrees of freedom (these numbers equal the true value of the tail index α of the Student-t). The numbers presented are obtained from 2000 simulations. In each cell we present the average γ -estimate over the 2000 simulations, the corrected standard error (26) and (27) based on the average estimate in parentheses and the cross-sectional standard deviation of the estimates as a benchmark for the corrected standard error. The first block of numbers represents the OLS based modified Hill γ -estimates (23) and the second block represents the WLS based modified Hill estimator (25). The threshold level κ is set to the half of the sample size for each sample.

true tail-index		sample size n			
γ	α	100	250	500	1000
OLS based modified Hill γ-estimates					
1.000	1	0.960 (0.343)	0.962 (0.239)	0.965 (0.175)	0.974 (0.126)
		0.378	0.247	0.173	0.126
0.500	2	0.480 (0.172)	0.481 (0.119)	0.475 (0.086)	0.478 (0.062)
		0.193	0.125	0.088	0.063
0.333	3	0.329 (0.118)	0.335 (0.083)	0.332 (0.060)	0.333 (0.043)
		0.137	0.087	0.063	0.045
0.250	4	0.272 (0.097)	0.269 (0.067)	0.266 (0.048)	0.268 (0.035)
		0.110	0.069	0.051	0.036
0.200	5	0.239 (0.085)	0.232 (0.058)	0.232 (0.042)	0.232 (0.030)
		0.094	0.062	0.043	0.030
WLS based modified Hill γ-estimates					
1.000	1	0.940 (0.310)	0.938 (0.194)	0.946 (0.136)	0.949 (0.095)
		0.307	0.194	0.140	0.098
0.500	2	0.459 (0.152)	0.460 (0.095)	0.456 (0.066)	0.456 (0.046)
		0.168	0.106	0.076	0.051
0.333	3	0.319 (0.105)	0.321 (0.066)	0.316 (0.045)	0.318 (0.032)
		0.126	0.080	0.058	0.040
0.250	4	0.260 (0.086)	0.258 (0.053)	0.258 (0.037)	0.257 (0.026)
		0.107	0.066	0.047	0.034
0.200	5	0.227 (0.075)	0.222 (0.046)	0.224 (0.032)	0.222 (0.022)
		0.098	0.061	0.031	0.031

7 Tail-index estimates for modified Pareto samples

This table provides γ -estimates of the tails of modified Pareto samples for four sample sizes and 5 numbers of true tail indices α . The numbers presented are obtained from 2000 simulations. In each cell we present the average γ -estimate over the 2000 simulations, the corrected standard error from section 4.1.5 based on the average estimate in parentheses and the cross-sectional standard deviation of the estimates as a benchmark for the corrected standard error. The first block of numbers represents the OLS based modified Hill γ -estimates (23) and the second block represents the WLS based modified Hill estimator (25). For each sample, κ is set to the half of the sample size.

true tail-index		sample size n			
γ	α	100	250	500	1000
OLS based modified Hill γ-estimates					
1.000	1	0.964 (0.345)	0.970 (0.241)	0.979 (0.177)	0.975 (0.126)
		0.385	0.241	0.176	0.124
0.500	2	0.549 (0.196)	0.538 (0.134)	0.537 (0.097)	0.536 (0.069)
		0.209	0.134	0.096	0.067
0.333	3	0.424 (0.152)	0.416 (0.103)	0.410 (0.074)	0.408 (0.053)
		0.163	0.101	0.071	0.052
0.250	4	0.361 (0.129)	0.358 (0.089)	0.353 (0.064)	0.351 (0.045)
		0.140	0.088	0.061	0.042
0.200	5	0.330 (0.118)	0.321 (0.080)	0.321 (0.058)	0.320 (0.041)
		0.127	0.077	0.056	0.039
WLS based modified Hill γ-estimates					
1.000	1	0.959 (0.317)	0.953 (0.197)	0.957 (0.138)	0.954 (0.096)
		0.325	0.203	0.145	0.100
0.500	2	0.529 (0.175)	0.530 (0.109)	0.527 (0.076)	0.529 (0.053)
		0.197	0.122	0.085	0.060
0.333	3	0.412 (0.136)	0.409 (0.084)	0.409 (0.059)	0.408 (0.041)
		0.157	0.095	0.068	0.049
0.250	4	0.359 (0.118)	0.354 (0.073)	0.355 (0.051)	0.352 (0.035)
		0.141	0.087	0.061	0.044
0.200	5	0.329 (0.109)	0.322 (0.066)	0.320 (0.046)	0.321 (0.032)
		0.133	0.082	0.056	0.041

8 Tail-index estimates for Burr samples

This table provides γ -estimates of the tails of Burr samples for four sample sizes (the true tail index α equals 4 for the Burr distribution). The numbers presented are obtained from 2000 simulations. In each cell we present the average γ -estimate over the 2000 simulations, the corrected standard error from section 4.1.5 based on the average estimate in parentheses and the cross-sectional standard deviation of the estimates as a benchmark for the corrected standard error. The first block of numbers represents the OLS based modified Hill γ -estimates (23) and the second block represents the WLS based modified Hill estimator (25). For each sample, κ is set to the half of the sample size.

true tail-index		sample size n			
γ	α	100	250	500	1000
OLS based modified Hill γ-estimates					
0.250	4	0.268 (0.096) 0.104	0.269 (0.067) 0.069	0.269 (0.049) 0.048	0.268 (0.035) 0.034
WLS based modified Hill γ-estimates					
0.250	4	0.268 (0.088) 0.100	0.265 (0.055) 0.062	0.265 (0.038) 0.042	0.265 (0.027) 0.031

To shed more light on the sensitivity of the previous results with respect to the choice of κ (equal to $n/2$), we present evidence for different choices of κ for these sample sizes in table 9. The results presented here are the average tail-index estimates over 2000 samples with observations drawn from a Student-t distribution with 3 degrees of freedom, for seven different choices of κ smaller than $n/2$. The resulting estimates for various κ are only marginally different. We conclude that the modified Hill estimator is quite robust with respect to the choice of κ , for κ smaller than or equal to $n/2$. As a rule of thumb, we choose κ equal to $n/2$ from now on. However, the rule is not general; for different sample sizes one should check the appropriateness of κ .

9 Tail-index estimates for different thresholds κ

This table provides tail-index γ -estimates obtained from the OLS based and WLS based modified Hill estimators (23) and (25) for four different sample sizes n and threshold levels κ as a function of n . The presented estimates are obtained from 2000 simulations. In each simulation, a Student-t sample is drawn consisting of n innovations in each simulation. The true tail index α of the underlying Student-t distribution equals 3. In parentheses we present the corrected standard errors given in section 4.1.5.

κ	Sample size n			
	100	250	500	1000
OLS based modified Hill γ-estimates				
$\frac{n}{3}$	0.351 (0.143)	0.345 (0.088)	0.345 (0.068)	0.342 (0.054)
$\frac{n}{4}$	0.343 (0.150)	0.348 (0.103)	0.343 (0.085)	0.343 (0.062)
$n^{0.75}$	0.343 (0.142)	0.348 (0.103)	0.344 (0.092)	0.342 (0.073)
$\left(\frac{n}{2}\right)^{0.9}$	0.342 (0.139)	0.346 (0.092)	0.343 (0.080)	0.343 (0.060)
$\left(\frac{n}{2}\right)^{0.925}$	0.349 (0.138)	0.346 (0.086)	0.343 (0.075)	0.345 (0.056)
$\left(\frac{n}{2}\right)^{0.95}$	0.346 (0.132)	0.343 (0.080)	0.342 (0.071)	0.343 (0.052)
$\left(\frac{n}{2}\right)^{0.975}$	0.343 (0.127)	0.340 (0.075)	0.338 (0.065)	0.340 (0.047)
WLS based modified Hill γ-estimates				
$\frac{n}{3}$	0.330 (0.134)	0.331 (0.084)	0.330 (0.059)	0.332 (0.041)
$\frac{n}{4}$	0.344 (0.157)	0.348 (0.103)	0.343 (0.071)	0.344 (0.049)
$n^{0.75}$	0.344 (0.143)	0.348 (0.103)	0.345 (0.078)	0.345 (0.059)
$\left(\frac{n}{2}\right)^{0.9}$	0.342 (0.138)	0.344 (0.091)	0.342 (0.066)	0.344 (0.048)
$\left(\frac{n}{2}\right)^{0.925}$	0.342 (0.132)	0.343 (0.085)	0.341 (0.061)	0.344 (0.044)
$\left(\frac{n}{2}\right)^{0.95}$	0.337 (0.123)	0.338 (0.079)	0.337 (0.056)	0.340 (0.040)
$\left(\frac{n}{2}\right)^{0.975}$	0.330 (0.115)	0.331 (0.073)	0.330 (0.051)	0.332 (0.036)

Finally, we compare the performance of our modified Hill estimator to the original Hill estimator in small samples. For this purpose, the same experiment as in table 1 is repeated for the conventional Hill estimator. The conventional Hill estimator (16) requires an 'optimal' k . In this respect, we follow the procedure propagated by Jansen and de Vries (1991) and Koedijk and Kool (1994) among others.²⁴

From table 10, we conclude that the conventional Hill estimator yields on average unbiased tail-index estimates for α equal to 1. For larger α 's, the conventional Hill estimator is severely biased even if the assumption about the (unknown) underlying distribution happens to be correct. For example, when the true (and assumed) α equals 3, the average estimate of γ equals 0.4 for a sample size of 100; that is, the average estimate of α equals $1/0.4$ or 2.5. In general, therefore, the conventional Hill estimator overestimates the tail fatness of the underlying distribution. In order to obtain good estimates by applying the conventional Hill estimator, it is necessary to assume a much less fat-tailed underlying distribution than the true distribution actually is. From a comparison of tables 1 and 5, we conclude that the modified Hill estimator consistently outperforms the conventional one in small samples.

Overall, the simulation results in this section provide supporting evidence of the adequacy of the modified Hill estimator in small samples. In the following section, we apply the modified Hill estimators to real world data.

²⁴ First an assumption is made on the form of the underlying distribution (most commonly the Student-t distribution is used). For a given sample size n and assumed value of α of the underlying sample, a Monte Carlo experiment is performed in which the mean squared error (MSE) is calculated for different values of k . Then the value of k that minimizes the average MSE is chosen as the optimal number of observations to use in the Hill estimator.

10 Conventional Hill estimates

This table presents conventional Hill γ -estimates for simulated Student-t samples. The sample size n equals 100, 250, 500 or 1000. The true α (that equals the number of degrees of freedom) equals either 1,2,3,4 or 5. The number of tail observations k used in the Hill estimator is determined by a Monte Carlo simulation as described in footnote 5. 'assumed α ' refers to the degrees of freedom assumed in the Monte Carlo samples. The numbers presented are averages over 2000 replications. The cross sectional standard deviations over the estimates are presented in parentheses.

true α	assumed α	sample size n			
		100	250	500	1000
$\alpha = 1$	1	1.081 (0.177)	1.045 (0.123)	1.038 (0.087)	1.033 (0.067)
	2	1.021 (0.225)	1.006 (0.185)	1.006 (0.138)	1.008 (0.104)
	3	1.010 (0.296)	1.002 (0.221)	1.005 (0.186)	1.004 (0.159)
	4	1.007 (0.368)	0.996 (0.262)	1.003 (0.222)	1.003 (0.173)
	5	1.005 (0.343)	0.999 (0.320)	1.002 (0.240)	1.005 (0.208)
$\alpha = 2$	1	0.689 (0.106)	0.634 (0.070)	0.629 (0.052)	0.607 (0.037)
	2	0.597 (0.124)	0.553 (0.093)	0.545 (0.075)	0.540 (0.053)
	3	0.558 (0.152)	0.532 (0.108)	0.525 (0.096)	0.517 (0.078)
	4	0.538 (0.181)	0.522 (0.129)	0.519 (0.110)	0.514 (0.084)
	5	0.543 (0.173)	0.513 (0.162)	0.515 (0.120)	0.511 (0.100)
$\alpha = 3$	1	0.583 (0.088)	0.525 (0.057)	0.516 (0.040)	0.490 (0.029)
	2	0.476 (0.094)	0.427 (0.071)	0.413 (0.053)	0.404 (0.039)
	3	0.427 (0.111)	0.400 (0.080)	0.383 (0.067)	0.374 (0.057)
	4	0.400 (0.134)	0.386 (0.094)	0.371 (0.078)	0.370 (0.062)
	5	0.407 (0.126)	0.374 (0.112)	0.368 (0.084)	0.361 (0.074)

Conventional Hill estimates – Continued

true α	assumed α	sample size n				
		100	250	500	1000	
$\alpha = 4$	1	0.534 (0.082)	0.475 (0.050)	0.467 (0.036)	0.438 (0.026)	
	2	0.425 (0.085)	0.371 (0.059)	0.354 (0.044)	0.345 (0.032)	
	3	0.369 (0.097)	0.342 (0.070)	0.320 (0.055)	0.308 (0.045)	
	4	0.341 (0.112)	0.326 (0.079)	0.307 (0.063)	0.302 (0.049)	
	5	0.348 (0.105)	0.309 (0.091)	0.302 (0.068)	0.294 (0.058)	
	$\alpha = 5$	1	0.508 (0.075)	0.447 (0.048)	0.439 (0.033)	0.410 (0.024)
		2	0.394 (0.078)	0.338 (0.053)	0.322 (0.039)	0.310 (0.030)
		3	0.340 (0.088)	0.307 (0.059)	0.286 (0.048)	0.270 (0.040)
		4	0.312 (0.100)	0.288 (0.067)	0.273 (0.055)	0.264 (0.044)
		5	0.318 (0.096)	0.272 (0.078)	0.269 (0.059)	0.254 (0.051)

4.3 Results on stock and exchange rate returns

In the previous section we showed that both the OLS and WLS based modified Hill estimators produce much less biased tail-index estimates in small samples. In this section, we replicate the studies of Jansen and de Vries (1991) and Loretan and Phillips (1994). Jansen and de Vries examine returns on ten stocks selected from the S&P 100 list and the returns on the S&P 500. They define the risk on investing in those stocks in terms of probabilities on extreme negative returns and judge the probabilities on extreme events as the October 1987 stock market crash into a historical perspective. Loretan and Phillips test the stationarity of the unconditional covariance of time series on exchange rates and common stocks. Recall that the α tail-index equals the maximum number of finite

moments that exist for a particular sample. Applying extreme value theory, they note and observe that many financial data are fat-tailed, that mostly the first two moments, i.e. mean and variance, exist but that the third and fourth moment are possibly not finite. They also focus on the properties of tests on covariance stationarity when some of these moments are infinite.

Both Jansen and de Vries, and Loretan and Phillips use data covering a long period of time to have a large number of observations in order to obtain unbiased tail-index estimates. As discussed by Jansen and de Vries, one of the properties of the tail-index is that it is unaffected by aggregation as long as the underlying distribution varies regularly at infinity which is true for many fat-tailed distributions. Re-sampling from daily to weekly or monthly returns should therefore not affect the levels of the γ -estimates. Given this theorem, we apply the modified Hill estimator in small samples and test whether we can obtain estimates similar to the conventional Hill estimates obtained from large samples. To do so, we re-sample the data to a lower frequency. Then we use the re-sampled data having a considerably smaller number of observations, to produce almost similar tail-index estimates as Loretan and Phillips, and Jansen and de Vries. Re-sampling the data at a lower frequency not only reduces the size of the data to be processed, it is also likely to reduce or eliminate any dependency between observations. The independence assumption on which our and other approaches rely may be violated by high frequency financial return data. Kearns and Pagan (1997) indicate that conclusions on tail fatness of financial time series may be weakened by explicit consideration of the second moment dependence of returns on financial series. Their results do not contradict the point estimates of the cited studies but suggest that the reported confidence bands are too narrow.

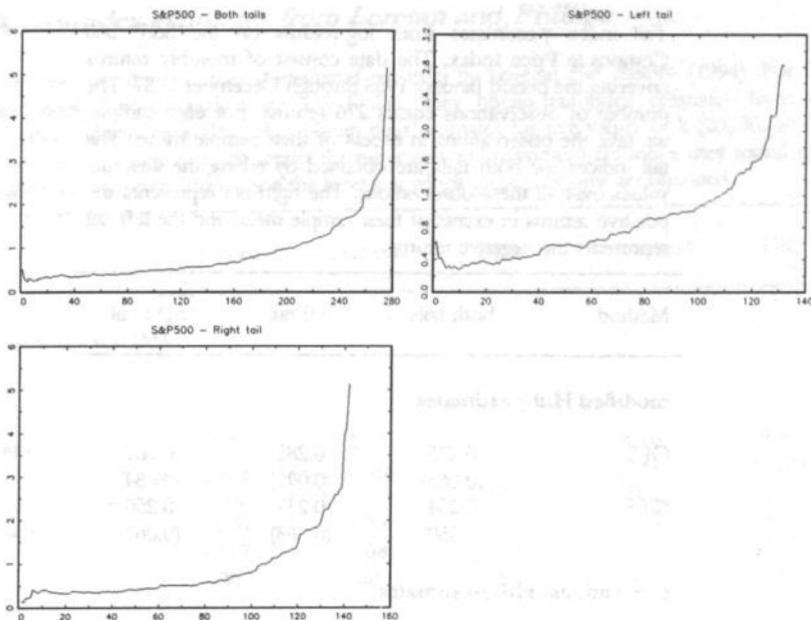
4.3.1 Returns on the S&P 500

Loretan and Phillips (1994) apply the conventional Hill estimator to obtain tail-index estimates for returns on the S&P 500 stock market index using more than 6400 daily observations from January 1962 through December 1987. Instead of selecting an optimal k , they consider a range of values for k . This way, no arbitrary assumptions about the underlying distribution of the stock returns are required. Loretan and Phillips report α -estimates between 3.44 (0.22) and 3.80 (0.54) for the left tail of the sample, and α -estimates between 3.08 (0.19) and 3.86 (0.39) for the right tail (their standard errors are in parentheses).

Jansen and de Vries (1991) examine the S&P 500 stock market index for almost the same period: January 1962 through December 1986. They use the conventional Hill estimator to produce α -estimates and determine the optimal k through Monte Carlo simulation; the resulting k equals 100. They obtain α -estimates equal to 4.30 (0.42) for the left tail and 3.96 (0.39) for the right tail. Their estimates are close to those of Loretan and Phillips.

We replicate these studies for the S&P 500 stock market index using the modified Hill estimator over approximately the same period at a lower frequency. Instead of a daily frequency, our data consist of the monthly returns on the S&P 500 Index from January 1965 through December 1987 (276 observations). The sample mean of the returns is subtracted prior to the estimation. Figure 6 presents the pattern of conventional Hill estimates (γ 's) for different values of k using the monthly S&P 500 returns. A similar pattern emerges as in Figure 5. Again, we observe the linearity for values of k smaller than some threshold level κ .

In table 11, we present γ -estimates obtained from the modified Hill estimator with $\kappa=n/2$. The OLS-based γ -estimates imply α 's of 3.51 (1.13) for the left tail and 3.75 (1.18) for the right tail. The WLS-based γ -estimates imply α 's of 4.22 (1.21) for the left tail and 4.00 (1.10) for the right tail. Apparently, the modified Hill estimator is able to produce tail-index estimates in a small sample (276 observations) almost similar to the estimates by Jansen and de Vries (1991) and Loretan and Phillips (1994) based on a sample of 6,000 observations. From the lower part of table 11, it seems that the modified Hill estimator overcomes the bias of the conventional Hill estimator. For the latter one, the γ -estimates imply α 's equal to 3.00 (0.86) for the left tail and 2.52 (0.72) for the right tail.



6 Hill estimates for the S&P 500 Composite Price Index

These graphs present the pattern of the Hill estimates as a function of k -- the number of tail observations used in the Hill estimator (16). The tail-indexes are obtained from monthly returns on the S&P 500 Composite Price Index covering January 1965 through December 1987. The upper-left graph shows the pattern for both tails simultaneously. The upper-right graphs presents the Hill estimates for the left tail of the sample and the bottom-left graph for the right tail.

4.3.2 Returns on foreign exchange rates

Loretan and Phillips (1994) also apply the conventional Hill estimator to obtain tail estimates for five major exchange rates against the U.S. dollar over the period 1978-1991. The currencies considered are the French Franc, the German Mark, the British Pound, the Swiss Franc, and the Japanese Yen. Their results are based on approximately 1550 daily observations and are summarized in table 12.

11 Tail estimates S&P 500

Tail index γ -estimates from log-returns on the S&P 500 Composite Price Index. The data consist of monthly returns covering the period January 1965 through December 1987. The number of observations equals 276 returns. For each sample, we take the observations in excess of their sample mean. The tail indices for both tails are obtained by taking the absolute values over all these observations. The right tail represents the positive returns in excess of their sample mean and the left tail represents the negative returns.

Method	both tails	left tail	right tail
modified Hill γ-estimates			
OLS	0.275 (0.065)	0.285 (0.092)	0.267 (0.084)
WLS	0.254 (0.050)	0.237 (0.068)	0.250 (0.069)
conventional Hill γ-estimates			
$k = 17 / 12$	0.335 (0.081)	0.334 (0.096)	0.397 (0.114)

OLS refers to the OLS based modified Hill estimator and WLS refers to the WLS based modified Hill estimator. The optimal number of tail observations k included in the conventional Hill estimator is determined by a Monte Carlo simulation. Standard errors are presented in parentheses. κ is set to the half of the sample size.

We apply both the conventional and the modified Hill estimator to the same five exchange rate series over approximately the same period (January 1979 - January 1991) using 620 weekly observations and compare our results with Loretan and Phillips. For illustrative purposes, we present the pattern of γ -estimates from the conventional Hill estimator on German Mark/US Dollar returns for different values of k in figure 7. Similar to figures 5 and 6, we again observe the linearity of the γ -estimates for values of k smaller than some threshold.

12 FX tail-index estimates from Loretan and Phillips

This table contains the tail-index α -estimates reported by Loretan and Phillips (1994). For five exchange rates denoted against the US Dollar, they obtain tail-index estimates from the conventional Hill estimator (16). They obtain their estimates for five values of k (20, 30, 50, 75, 100). For each exchange rate, we report the minimum and maximum tail index they found with the standard error in parentheses and the level k at which these estimates are obtained.

	France	Germany	Japan	Switzerland	UK
Daily returns - $n \approx 1550$					
left tail					
minimum	2.39 (0.24)	3.03 (0.35)	2.83 (0.28)	3.10 (0.31)	2.59 (0.26)
	100	75	100	100	100
maximum	3.00 (0.55)	3.70 (0.68)	3.09 (0.44)	5.13 (1.15)	3.25 (0.73)
	30	30	50	20	20
right tail					
minimum	2.29 (0.32)	2.70 (0.38)	3.30 (0.33)	2.71 (0.31)	3.27 (0.33)
	50	50	100	75	100
maximum	2.86 (0.64)	2.97 (0.66)	4.38 (0.80)	3.30 (0.74)	4.71 (1.05)
	20	20	30	20	20

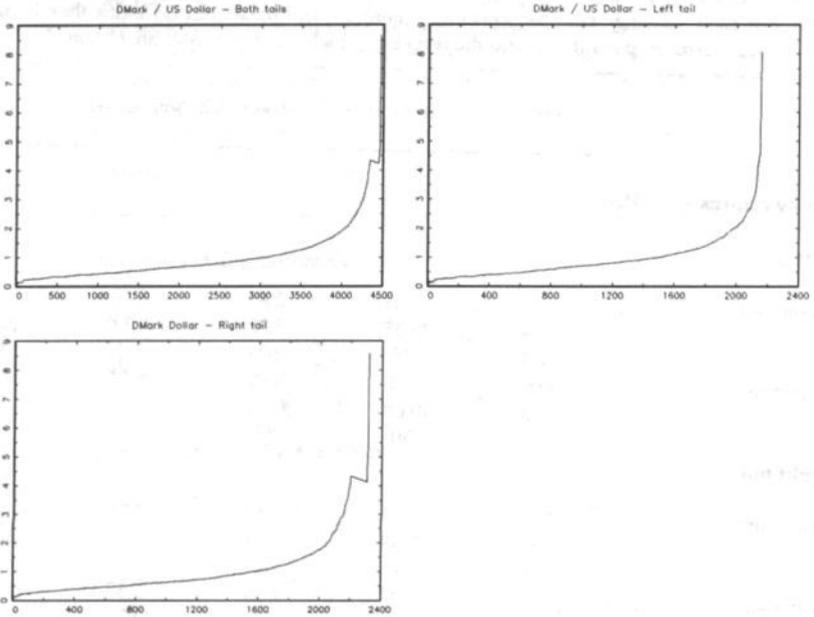
In table 13 we first present the modified Hill estimates. For all exchange rates under consideration our γ -estimates suggest higher α -estimates than found by Loretan and Phillips. That is, even the large number of observations used by Loretan and Phillips probably overestimates tail-fatness.²⁵

In the lower part of table 13, we present γ -estimates using the conventional Hill estimator.²⁶ Again, the conventional Hill procedure produces γ -estimates that are larger

²⁵ In further support of the unbiasedness of the modified Hill estimator in small samples, we note that Dacorogna, et al. (1995), for example, apply a bootstrap methodology in combination with high frequency data to obtain unbiased tail-index estimates and find α -estimates close to the values reported here for the same exchange rates.

²⁶ For the procedure that we used to select the optimal k , we refer to paragraph 4.2. Using $n = 620$, the optimal k is found to be 20 for the analysis of one tail, and 38 for the simultaneous analysis of both tails.

for all exchange rates than those for the modified Hill estimator. The results imply α -estimates between 3 and 4 for most exchange rates, whereas they are larger than 4 mostly if the tail-index is estimated by the modified Hill estimator.



7 Hill estimates for the German Mark vis-a-vis the US Dollar

These graphs present the pattern of the Hill estimates as a function of k - the number of tail observations used in the Hill estimator (16). The tail-indexes are obtained from daily returns on the German Mark-US Dollar exchange covering January 1979 through April 1996. The upper-left graph shows the pattern for both tails simultaneously. The upper-right graphs presents the Hill estimates for the left tail of the sample and the bottom-left graph for the right tail.

4.4 Concluding remarks

The tail structure of the distribution underlying the data-generating process of economic variables may provide important information. For instance, it may be used as a tool to assess the probability that some extreme event will occur without the need to make

further assumptions about the underlying distribution. Tail behavior can be summarized by one characteristic parameter α , the so-called tail-index. So far, all methods that have been developed to estimate α , suffer from small sample bias and are well behaved only asymptotically. Of these methods, the conventional Hill (1975) estimator is best known and most widely used.

For reliable applications, a large number of observations is required. That is, either high-frequency data or extremely long sample periods are necessary. Because of the fact that such methods have been used in the literature for relatively small samples, results must be interpreted cautiously and overestimation of tail fatness is likely to be a problem.

Small samples are bound to remain a fact of life in many applications. Moreover, splitting up long samples in shorter ones to analyze potential changes in tail behavior over time may be interesting as well. Consequently, we focus on the search for an unbiased tail estimator in small samples.

In this chapter, we modify the conventional Hill estimator. To this end, we exploit an important characteristic of its bias-function. In particular, it may be shown that the bias is an almost linear function of the number of tail observations (k) used in the estimation. Our estimator uses a number of conventional Hill estimates (for different k) as input which are different in the number of tail observations included, and calculates a weighted average to infer information about the tail from more than one single Hill estimate. In other words, the modified Hill estimator is a weighted average of a set of conventional Hill estimators, with weights obtained by using simple least squares techniques. We also take into account the issues of heteroskedasticity and autocorrelation due to overlapping data. A procedure is developed to correct the conventional least squares standard errors.

Subsequently, we test the adequacy of our estimator in a simulation study. Overall, we conclude that much less biased tail-estimates result even for samples as small as 100 observations for a range of underlying distributions: Pareto, Student-t, and Burr. Furthermore, our computed standard errors appear to be appropriate and are similar to cross-sectional standard errors that may be used as a benchmark. Overall, our modification provides an important gain over the conventional Hill estimator.

13 Tail estimates foreign exchange rates

This table contains conventional Hill and modified Hill tail-index estimates for daily and weekly returns on exchange rates expressed against the US Dollar for the period January 1979 through January 1991. For each sample, we take the observations in excess of their sample mean. The tail indices for both tails are obtained by taking the absolute values over all these observations. The right tail represents the positive returns in excess of their sample mean and the left tail represents the negative returns.

	France	Germany	Japan	Switzerland	UK
Weekly returns - modified Hill γ-estimates - $n \approx 620$					
Both tails					
OLS	0.203 (0.033)	0.171 (0.028)	0.203 (0.033)	0.170 (0.033)	0.207 (0.034)
WLS	0.202 (0.026)	0.195 (0.026)	0.173 (0.022)	0.189 (0.024)	0.189 (0.024)
Left tail					
OLS	0.162 (0.037)	0.152 (0.035)	0.185 (0.044)	0.135 (0.031)	0.217 (0.050)
WLS	0.166 (0.031)	0.168 (0.032)	0.121 (0.024)	0.137 (0.026)	0.202 (0.038)
Right tail					
OLS	0.248 (0.045)	0.171 (0.038)	0.183 (0.039)	0.175 (0.039)	0.216 (0.048)
WLS	0.231 (0.042)	0.194 (0.035)	0.166 (0.035)	0.195 (0.035)	0.201 (0.037)
Weekly returns - conventional Hill γ-estimates - $n \approx 620$					
Both tails					
k = 38	0.242 (0.039)	0.232 (0.038)	0.288 (0.047)	0.196 (0.032)	0.258 (0.042)
Left tail					
k = 20	0.204 (0.046)	0.213 (0.048)	0.277 (0.062)	0.234 (0.052)	0.274 (0.061)
Right tail					
k = 20	0.327 (0.073)	0.226 (0.051)	0.236 (0.053)	0.194 (0.043)	0.251 (0.056)

OLS refers to the OLS based modified Hill estimator and WLS refers to the WLS based modified Hill estimator. The optimal number of tail observations k included in the conventional Hill estimator is determined by a Monte Carlo simulation. Standard errors are presented in parentheses. κ is set to the half of the sample size.

To evaluate the performance of the modified Hill with respect to real world data, we apply it to obtain tail-index estimates for returns on the S&P 500 Composite Price Index and five main foreign exchange rates against the U.S. dollar. Our choice of economic series allows an explicit comparison with previously published work for the same series and (approximately) the same period, which uses extremely long samples (through higher frequency data). Our small-sample estimates are very close to the estimates that are obtained from the high-frequency datasets with many more observations. Our comparative work again shows that tail-fatness is easily exaggerated in small samples. Implicitly, this suggests that probabilities on extreme events may be over-estimated as well. In reality they may be much smaller than has been previously assumed. Our modified Hill estimator avoids this problem and provides better tail-index estimates even in small samples.

One attractive application of extreme value theory is that, if one knows the tail-index, exceedence probabilities or equivalently large quantiles can be estimated without making assumptions on the structure of the underlying sample. Since the various quantile estimators put forward in the literature depend on the tail-index estimate²⁷, our method is also likely to improve these quantile estimates. However, the issue of estimating quantiles in relation to our proposed tail index estimators goes beyond the scope of this dissertation and remains to be investigated.

One point of concern is that simulation evidence is presented only for the sample sizes used in this chapter. The large sample properties from the modified Hill estimator still need to be examined. Especially the rule of thumb for κ , always set κ equal to half of the sample size as a proxy for the number of informative tail observations, is likely to become less appropriate. In this case, one should always check, using graphs as presented in Figure 5, whether the choice of κ is such that the linearity assumption is not severely violated.

²⁷ See Dekkers and de Haan (1989) and the appendix of de Haan, Jansen, Koedijk and de Vries (1994) for two quantile estimators.

5 The Tail Fatness of FX Returns^a

In this chapter we shall focus on the tail characteristics of the distributions of returns on foreign exchange rates. It is a well-known stylized fact that foreign exchange returns, like the returns on many other financial assets, are non-normal and tend to have fat-tailed distributions. Although information about the precise magnitude of the tail-fatness is crucial for applications such as risk analysis, little consensus exists in this respect, due to estimation problems. Typically, empirical estimation of tail-fatness is conditional on the specific alternative (fat-tailed) distribution -- such as the Student-t or stable -- chosen.²⁹ Since the data-generating processes of these alternative distributions are non-nested, an incorrect choice with respect to the underlying distribution of returns may lead to significant estimation errors because of misspecification.

In recent years, extreme value analysis has been proposed to overcome this potential difficulty. It investigates the distribution of the maximum (minimum) in large samples, thereby determining the shape of the tails of a distribution. The limit law for the maximum is characterized by the tail-index introduced in the previous chapter, which corresponds one-to-one with the number of existing moments of the underlying distribution. Note that alternative distributions like the stable and Student-t now are nested within the limit law for extremes; for the stable distribution, it is the characteristic exponent (< 2) which equals the tail-index, while for the Student-t distribution it is the number of degrees of freedom. The gain from applying extreme value analysis is that one can nest and test for different tail sizes. However, information about the center

²⁸ This chapter is based on: R. Huisman, K. Koedijk, C. Kool, and F. Palm, 1998, 'The Tail fatness of FX Returns'.

²⁹ See Westerfield (1977), Rogalski and Vinso (1978) and Boothe and Glassman (1987), McCulloch (1997).

characteristics of the distribution is lost. Since it is information about the outlier observations that is crucial for studying issues like exchange rate volatility, exchange rate risk and value at risk, the use of extreme value analysis offers a positive net gain in our view.

Hill (1975) proposed the best-known and most often applied extreme value estimator for the tail-index. Although asymptotically unbiased, this maximum likelihood estimator suffers severely from small sample bias. In fact, this holds for most alternative estimators, see Pictet, Dacorogna, and Müller (1996). Consequently, the applicability of tail-index estimation is limited to cases of extremely large samples, either in the form of high frequency data or a very long sample period.³⁰ In real life, often neither of these two conditions is fulfilled and only a small sample analysis is feasible.

In this chapter, we re-examine the tail fatness of exchange rate returns using the tail-index estimator proposed in the previous chapter, referred to as the HKKP estimator (HKKP: Huisman, Koedijk, Kool and Palm), which is unbiased in small samples. We show that previous tail-estimates in the literature overestimate the amount of tail fatness in exchange rate returns. Additionally, goodness-of-fit statistics provide supportive evidence of the appropriateness of assuming that a Student-t distribution underlies the data-generating process of exchange rate returns. Both conclusions appear to hold more for floating than for fixed exchange rates.

5.1 Data

The data used is the same as in chapter 2 where we analyzed the validity of uncovered interest parity in relation with extreme exchange rate returns. Summarizing, the data consist of London closing mid-prices against the Pound Sterling for the period January 2, 1979 and April 4, 1996 obtained from DATASTREAM. The currencies of sixteen countries are considered: Austria, Belgium, Canada, Denmark, France, Germany, Ireland, Italy, Japan, the Netherlands, Norway, Spain, Sweden, Switzerland, the United Kingdom and the United States. We sample the data at the weekly frequency³¹ (using end-of-period data) to reduce the influences of ARCH effects and other types of dependency between subsequent observations that are present in many financial data, since the tail-index

³⁰ Recently, efficient nonparametric estimators have been proposed, see Beirlant et.al. (1996), for example. However, their implementation is complex and small-sample properties are as yet unknown.

³¹ Note that the true tail-index of the underlying return distribution is unaffected by time-aggregation; thus, the choice of sampling frequency should only affect the estimation results through the change in number of observations.

estimates are produced assuming independent observations³². All data are expressed as the number of foreign currencies to pay for one unit of the numeraire currency.

Table 14 contains summary statistics of the weekly data. With respect to skewedness and kurtosis the following observations can be made from the table. First, the EMS currencies denoted in German Marks are skewed to the right, whereas the amount of skewedness is much smaller and hovers around zero for the floating exchange rates denoted in US Dollars. Second, all exchange rate returns exhibit excess kurtosis³³, reflecting apparent tail fatness. The excess kurtosis is most apparent for the EMS exchange rates in general, and for the Spanish Peseta and Swedish Krone in particular.

14 Summary statistics

These summary statistics cover the sample period from January 2, 1979 through April 4, 1996. The data is sampled at the weekly frequency. The mean and standard deviations of the (std.) returns are in annualized percentages. In the other columns, the skewedness and excess kurtosis are provided.

Currency	numeraire: German Mark				numeraire: US Dollar			
	mean	std.	skew	kurt.	mean	std.	skew	kurt.
Austria	-0.005	1.172	-0.116	3.551	-0.634	5.219	-0.101	1.469
Belgium	0.350	1.542	1.661	20.718	-0.279	5.173	-0.093	1.518
Canada	0.957	5.454	0.020	1.500	0.328	2.177	0.092	2.669
Denmark	0.561	1.666	1.534	11.758	-0.068	5.093	-0.099	1.150
France	0.780	1.848	2.864	28.210	0.151	5.139	0.059	1.857
Germany	---	---	---	---	-0.629	5.248	-0.065	1.457
Ireland	0.925	2.518	1.760	16.855	0.296	5.171	0.207	2.656
Italy	1.574	3.250	2.284	18.090	0.945	5.044	0.676	5.855
Japan	-0.542	4.471	-0.252	1.307	-1.171	4.993	-0.537	2.774
Netherlands	0.098	0.869	0.255	13.376	-0.531	5.166	-0.094	1.269
Norway	0.803	2.669	2.089	15.334	0.174	4.475	0.454	3.218
Spain	1.472	3.919	8.687	148.89	0.843	5.573	2.958	33.490
Sweden	1.109	3.936	6.319	84.030	0.480	4.948	3.145	34.096
Switzerland	-0.234	2.501	0.149	3.111	-0.863	6.002	-0.123	1.305
U. Kingdom	0.961	3.960	0.592	3.026	0.332	5.158	0.235	3.379
U. States	0.629	5.248	0.065	3.457	---	---	---	---

³² Kearns and Pagan (1997) argue that the dependency between the observations results in serious biases for GARCH and IGARCH type models, especially when those estimates are obtained from small samples. Lucas (1997) shows that the tail-index estimator that we use in this study produces rather unbiased tail-index estimates for a GARCH(1,1) process in a sample of 250 observations. This, combined with the fact that we use weekly data to eliminate as much dependency as possible, makes that we feel comfortable about the accuracy of the estimates.

³³ Excess with respect to the normal distribution which has a kurtosis equal to 3.

5.2 Results

In this section, we present tail-index estimates for the 16 currencies described in the previous section with the US dollar and the DM alternating as benchmark currency. We explicitly compare the tail-index estimates obtained from the HKKP estimator with results from the literature that are mostly obtained using maximum likelihood techniques. Subsequently, goodness-of-fit tests are done to investigate whether a hypothesized Student-t distribution with degrees of freedom equal to the obtained HKKP tail-index estimates is able to adequately reproduce the actual distribution of exchange rate returns.

5.2.1 Tail-index estimates

The observed excess kurtosis in table 14 already supports the finding of many other studies that the distribution of foreign exchange rate returns is fat-tailed. Note though that the kurtosis estimate is conditional on existence of the fourth moment of the underlying distribution. Especially for cases where quite large excess kurtosis values are computed, such as, for example, the Spanish Peseta and the Swedish Krone against the DM, this assumption might be unwarranted. A direct way to test for the existence of the fourth moments is to evaluate the tail-index estimates for each exchange rate. For this purpose, HKKP-tail estimates with appropriate standard errors are reported in table 15. Estimates are shown for left-tail and right-tail observations separately and for the combination of both tails.³⁴

According to table 15, the fourth moments are likely to exist for all exchange rates vis-à-vis the US dollar. Considering both tails simultaneously, we find tail-index estimates vary between 4.0 for the Spanish Peseta and 8.1 for the British Pound. The tail-index estimates are smaller for currencies vis-à-vis the German Mark, indicating a higher amount of tail fatness. The estimates vary between 2.2 for the French Franc and 5.4 for the Canadian Dollar. For many of these exchange rates, the standard errors are such that a test of whether the fourth moment is finite can be rejected.

³⁴ All observations are taken in excess of their sample mean. The left tail is examined using the absolute value of all negative returns, the right tail is examined using all positive returns and to examine both tails simultaneously we use the absolute values of all returns.

15 Tail-index α estimates

The table provides tail-index estimates for the weekly returns on the sixteen currencies over the sample period from January 2, 1979 through April 4, 1996. The estimates presented are for both tails simultaneously, the left, and the right tail respectively. T-test statistics are for tests of the hypothesis that the tail-index estimate for the left tail equals the estimate for the right tail.

Currency	numeraire: German Mark				numeraire: US Dollar			
	both	left	right	t	both	left	right	t
Austria	3.419 (0.333)	3.461 (0.476)	3.185 (0.452)	0.421	5.753 (0.561)	5.621 (0.798)	5.986 (0.822)	-0.318
Belgium	3.248 (0.317)	3.953 (0.541)	2.718 (0.388)	1.856	5.382 (0.524)	5.120 (0.721)	5.701 (0.789)	-0.544
Canada	5.454 (0.531)	4.371 (0.605)	7.259 (1.024)	-2.428	4.609 (0.449)	4.006 (0.551)	5.660 (0.803)	-1.698
Denmark	3.290 (0.321)	4.157 (0.562)	2.695 (0.389)	2.139	5.280 (0.515)	4.844 (0.679)	5.724 (0.795)	-0.842
France	2.205 (0.215)	2.707 (0.344)	2.489 (0.389)	0.421	5.236 (0.510)	5.520 (0.762)	4.772 (0.673)	0.736
Germany	---	---	---	---	5.323 (0.519)	5.937 (0.834)	4.828 (0.670)	1.037
Ireland	2.616 (0.255)	2.890 (0.376)	2.761 (0.418)	0.230	5.757 (0.561)	5.441 (0.744)	6.010 (0.858)	-0.501
Italy	2.281 (0.222)	2.364 (0.304)	2.426 (0.373)	-0.129	4.714 (0.459)	5.283 (0.720)	4.228 (0.605)	1.122
Japan	5.250 (0.512)	5.546 (0.782)	5.303 (0.732)	0.217	6.099 (0.594)	9.037 (1.318)	5.801 (0.777)	2.115
Netherlands	2.790 (0.272)	2.707 (0.376)	2.857 (0.401)	-0.271	5.365 (0.523)	6.278 (0.880)	4.705 (0.653)	1.436
Norway	3.314 (0.323)	4.945 (0.664)	2.512 (0.366)	3.209	4.442 (0.433)	4.886 (0.672)	4.035 (0.571)	0.965
Spain	3.187 (0.311)	4.596 (0.610)	2.525 (0.372)	2.897	4.034 (0.393)	5.084 (0.690)	3.283 (0.472)	2.155
Sweden	3.034 (0.296)	4.072 (0.546)	2.434 (0.356)	2.515	4.264 (0.416)	5.079 (0.696)	3.594 (0.512)	1.719
Switzerland	4.786 (0.466)	5.613 (0.793)	4.179 (0.577)	1.462	5.734 (0.599)	8.314 (1.157)	4.552 (0.637)	2.849
U. Kingdom	3.847 (0.375)	4.162 (0.555)	4.019 (0.590)	0.177	8.111 (0.790)	8.027 (1.092)	7.787 (1.117)	0.154
U. States	5.232 (0.519)	4.828 (0.670)	5.937 (0.834)	-1.037	---	---	---	---

The overlapping data corrected standard errors are presented in parentheses. The test on equality is a t-test for H_0 : left tail-index - right-tail-index = 0 vs. H_1 : left tail-index - right tail index \neq 0.

From table 15 we also conclude that the tail-index of the left and the right tail often differ considerably. For example, we obtain a tail-index estimate equal to 4.3 for the left

tails of Canadian vis-à-vis the German Mark and an estimate equal to 7.3 for the right tail. If these estimates differ significantly over both tails, one should take the tail-specific estimate into account when inferences are drawn from the tails; i.e. it is inappropriate to use the estimate obtained from the combined information in both tails. The columns headed 't' contain the t-statistics from testing the hypothesis that the tail-index estimate of the left tail equals the tail-index estimate of the right tail³⁵. The results show that the difference between the left tail-index and the right tail-index is insignificant for the majority of exchange rates. It is significant at the 5% level for the Canadian Dollar, the Danish Krone, the Norwegian Krone, the Spanish Peseta and the Swedish Krone all vis-à-vis the German Mark and for the Japanese Yen, the Spanish Peseta, and the Swiss Franc vis-à-vis the US Dollar.

5.2.2 *The tail-index of EMS rates versus floating rates*

Koedijk, Stork and de Vries (1992) and Koedijk and Kool (1994) show that EMS exchange rate returns exhibit fatter tails than do floating rate returns. The information in table 15 provides clear evidence in support of this conclusion. For all EMS currencies, tail-estimates are lower when measured vis-à-vis the German Mark than vis-à-vis the US dollar, regardless of whether the left-tail, right-tail or both are considered. Table 16 contains the results of formally testing the null hypothesis that the tail-index of a currency vis-à-vis the German Mark equals the tail-index of the same currency but now against the US Dollar. The test statistic is asymptotically distributed as a Student-t. For negative values the tail fatness against the German Mark exceeds that against the US Dollar.

For most of the countries that have floated against both the US dollar and the DM, the equality of tail-fatness across numeraire currencies can not be rejected. This is the case for Canada, Japan, Spain, and Switzerland. Note that Spain has participated in the ERM only from 1989 onward, and has devalued a few times against the DM in 1992 and 1993. For all ERM countries equality of tail-estimates across numeraire currencies is firmly rejected with the tails against the DM showing significantly more fatness than the tail-estimates against the US dollar. The same holds (marginally) for Norway and Sweden, even though they did not formally participate in the EMS over the sample. Consequently, the HKKP estimates support earlier conclusions that the return distribution of non-

³⁵ See the previous chapter for more details on this t-test. We assumed here that the observations are mutually independent.

floating exchange rates has substantially fatter tails than does the distribution under floating.

16 EMS versus floating

The t-statistics are the result of testing the hypothesis that the tails of a currency vis-à-vis the German Mark equals the tails of the same currency vis-à-vis the US Dollar. For many European currencies, this can be seen as a test on tail equality of EMS and floating rates. 'Both' represents the test on both tails simultaneously. 'Left' and 'right' refer to the left and the right tail respectively.

currency	both	left	right
Austria	-3.578	-2.325	-2.986
Belgium	-3.485	-1.295	-3.393
Canada	1.215	0.446	1.229
Denmark	-3.279	-0.779	-3.422
France	-5.476	-3.365	-2.935
Germany	---	---	---
Ireland	-5.097	-3.065	-3.404
Italy	-4.772	-3.735	-2.535
Japan	-1.083	-2.278	-0.467
Netherlands	-4.368	-3.732	-2.412
Norway	-2.088	0.065	-2.246
Spain	-1.690	-0.530	-1.261
Sweden	-2.409	-1.138	-1.860
Switzerland	-1.249	-1.926	-0.434
U. Kingdom	-4.876	-3.155	-2.983
U. States	---	---	---

The tests on equality is a t-test for H_0 : tail-index German Mark -/- tail-index US Dollar = 0 vs. H_0 : tail-index German Mark -/- tail-index US Dollar $\neq 0$.

5.2.3 A comparison: HKKP estimates versus the literature

The HKKP tail-index estimates that are reported in table 15 are much higher than values that were found in previous studies for floating rates. For example, Loretan and Phillips (1994) report 's for the German Mark / US Dollar around 3.4 where we find 5.3 for both tails instead. Boothe and Glassman (1987) report a tail-index estimate equal to 3.2 for the German Mark / US Dollar.

These level differences can be explained twofold. Firstly, they might be due to the different sample period, where Loretan and Phillips examine the period 1978 through 1991 and Boothe and Glassman examine the period 1973 through 1984. Our sample is partly overlapping with theirs as it contains data over the period 1979 through 1996. However, we presented tail-index estimates obtained over roughly the same sample as Loretan and Phillips and find systematically higher tail-index estimates for all currencies examined than reported in Loretan and Phillips in the previous chapter.

The alternative explanation focuses on the differences in estimation procedures. Loretan and Phillips (1994) obtain their estimates from the conventional Hill estimator (16). This estimator is biased in small samples, as we already discussed, while the HKKP-estimates in table 15 are unbiased. Boothe and Glassman (1987) obtain tail-index estimates use a maximum likelihood procedure to estimate the tail index conditional on the true distribution being Student-t. As argued before, this distributional assumption may be incorrect in general and may lead to estimation errors. In this case, an additional concern is the inequality of left and right tails for many currencies, as illustrated in table 15. The assumption of a symmetric underlying distribution definitely leads to misspecification like a Student-t distribution, a sum-stable distribution and a mixture of normals. Finally, the conventional Hill-estimator is a maximum likelihood estimator as is Boothe and Glassman's method, raising the issue of small sample bias again. Overall, we conclude that exchange rate returns may be fatter than normal, but that they are considerably less fat than has been argued in the literature before. This holds especially true for floating exchange rate returns.

5.3 Distributional characteristics

The fact that the tail-fatness of exchange rate returns has been considerably overestimated in the literature, has important implications for the issue of which statistical distribution captures the observed distribution of exchange rate returns best. Our results, for instance, indicate that the stable distribution is unlikely to be an appropriate choice to model exchange rate returns. The characteristic exponent of the stable distribution which equals the tail-index is by definition below 2, while we find tail-index values around 5 for most floating exchange rate returns.

Westerfield (1977), Rogalski and Vinso (1978), and Boothe and Glassman (1987) examined several distributions to capture tail fatness. The Student-t distribution, a mixture of normals, and the sum-stable distribution are shown to provide the best fit

among various alternative candidates, although their goodness-of-fit tests only provide statistically weak evidence. This may be due to the poor quality of previous tail-estimates.

Here, we investigate whether a (potentially asymmetric) Student-t distribution with the number of degrees of freedom equal to the estimated tail-index, is able to adequately reproduce the observed empirical distribution of exchange rate returns.

Let x_i be the exchange rate returns under consideration. To fit the Student-t distribution, we assume that $y_i = \theta(x_i - m)$ is Student-t distributed with α degrees of freedom. Here m is a location parameter that equals the sample mean. The scale parameter θ is a function of the variance and can be derived as follows. For a student-t distribution with α degrees of freedom, the variance of y_i equals $\alpha/(\alpha-2)$ where $\alpha > 2$. We then obtain:

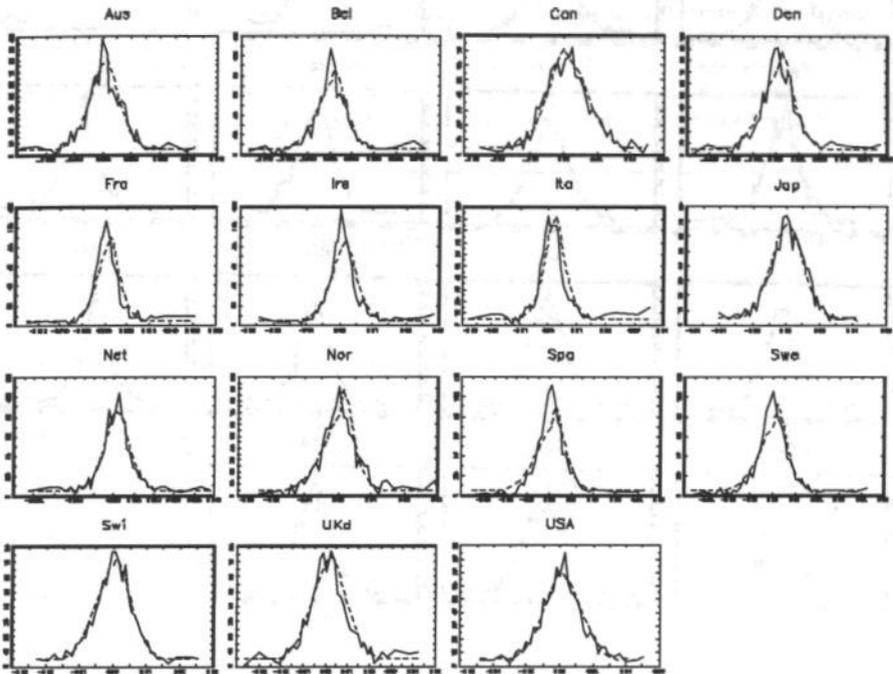
$$(36) \quad \frac{\alpha}{\alpha - 2} = \text{var}(y_i) = \theta^2 \text{var}(x_i)$$

Consequently, the scale parameter θ can be computed as the following function of the sample variance and the number of degrees of freedom α :

$$(37) \quad \theta = \sqrt{\frac{\alpha}{(\alpha - 2) \text{var}(x_i)}}$$

Now, we set the number of degrees of freedom equal to the (unbiased) HKKP tail-index estimate for each currency instead of estimating this number using the maximum likelihood procedure as in Boothe and Glassman. This way, we circumvent the usual small sample bias in other methods. Because the tail-index estimates for the left and right tail differ significantly for some exchange rates, we model and reproduce each tail separately. The above procedure allows for the reproduction of data-generating distribution for exchange rate returns conditional on the HKKP- tail estimate and the observed sample variance. In figures 1 and 2 both these fitted (hypothetical) Student-t distribution and the observed empirical distribution of each exchange rate return series is given in the form of histograms. Note that the whole distribution is fitted, not just the tail, although it is actually only the information in the tail that has been used for the fit. For example, consider the US Dollar vis-à-vis the German Mark. We fit the left tail with a Student-t distribution with α equal to 5.2 and the right tail with a Student-t distribution with α equal to 5.9. Figure 8 shows the fitted distribution for the exchange rates in terms

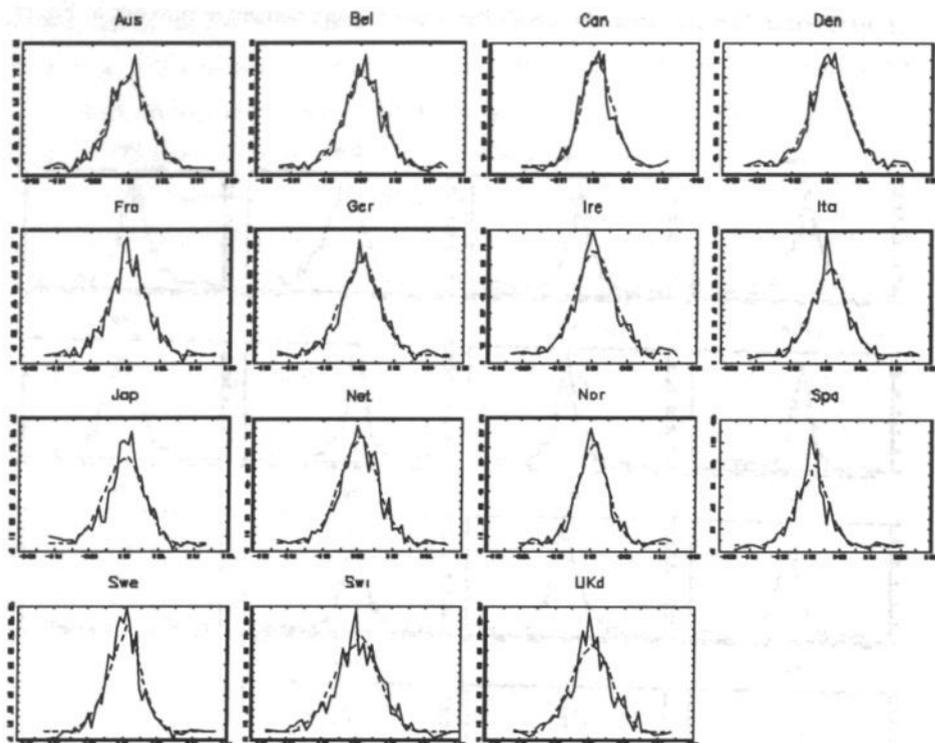
of the German Mark, whereas the US Dollar is taken as the numeraire currency in Figure 9.



8 Student-t fits of exchange rate returns in German Marks

All graphs show the observed frequencies (straight lines) and the expected frequencies as a result of the Student-t fits (dashed lines).

At first glance, the Student-t distribution seems to fit the complete empirical distribution of exchange rate returns surprisingly well for many exchange rates. Some important deviations from the symmetric Student-t exist for the French Franc, the Italian Lira, the Spanish Peseta and the Swedish Krone against the German Mark due to non-zero skewedness in EMS rates as is reflected in table 14. In figure 9, skewedness problems appear less severe. Substantial differences between actual and fitted distribution only seem present for the Japanese Yen and the Spanish Peseta.



9 Student-t fits of exchange rate returns in US Dollars

All graphs show the observed frequencies (straight lines) and the expected frequencies as a result of the Student-t fits (dashed lines).

To test statistically the hypothesis that the fitted Student-t distribution is a good approximation for the unconditional empirical distribution of exchange rate returns, we apply goodness-of-fit tests similar to the one used by Boothe and Glassman (1987). They rejected the Student-t distribution for returns on the British Pound, the German Mark and the Japanese Yen vis-à-vis the US Dollar and statistically accept the Student-t only for the Canadian Dollar. Table 17 contains the results both for the whole distribution and for the tails.³⁶ With respect to the latter, we distinguish between fitting the 5% largest and 5% smallest observations and fitting the 10% largest and 10% smallest observations.

³⁶ More details on the goodness-of-fit test can be found in appendix.

17 Goodness-of-fit statistics for Student-t

The table provides chi-squared goodness-of-fit statistics to test the null hypothesis that the Student-t distribution with the number of degrees of freedom equal to the tail-index estimate α provides a good fit of the exchange rate returns. The tests are performed over all observations (all), over the 10% largest and the 10% smallest observations (10%), over the 5% largest and the 5% smallest observations (5%) and over the 5% largest and smallest obtained from the sample from which all realignments are excluded. The expected and observed frequencies are calculated over 20 intervals for all and 10%; they are calculated over 10 intervals in case of 5%.

Currency	numeraire: German Mark				numeraire: US Dollar		
	5%	5% (R)	10%	all	5%	10%	all
Austria	9.487	10.822	18.567	22.952	4.869	22.874	20.815
Belgium	18.132	19.385	31.061	44.003	5.485	13.455	13.879
Canada	12.655	14.963	24.293	23.268	6.363	12.058	18.200
Denmark	18.292	11.170	24.936	46.153	7.753	15.144	12.390
France	21.792	18.474	37.924	100.795	5.302	16.572	20.582
Germany	---	---	---	---	11.350	16.812	21.391
Ireland	17.892	16.011	25.640	77.832	4.888	15.027	16.169
Italy	36.178	44.268	48.642	92.668	9.811	7.750	14.866
Japan	10.591	10.383	13.166	18.843	13.291	20.651	45.325
Netherlands	12.263	15.799	26.580	28.459	7.573	16.317	12.752
Norway	46.615	23.638	59.060	80.006	10.280	19.000	26.829
Spain	46.507	47.795	51.876	118.308	10.573	12.285	48.066
Sweden	21.269	22.692	40.355	74.582	15.816	16.925	32.819
Switzerland	3.462	5.651	13.990	10.741	16.277	23.588	33.518
U. Kingdom	27.569	32.558	33.325	41.453	7.566	14.793	27.681
U. States	11.207	13.093	17.279	21.502	---	---	---

Chi-squared critical value for 19 degrees of freedom at the 5% level is 30.140; for 9 degrees of freedom it equals 16.920. Reject the null that the Student-t is a good approximation for the exchange rate returns if the reported goodness-of-fit statistic exceeds the corresponding critical value. The fitted Student-t distributions are not tail symmetric; i.e. we used the separate tail-index estimates for the left and right tail.

From table 17, we observe that the Student-t is accepted to be a good approximation for the complete distribution of returns on many of the floating exchange rates (see the 'all' column). For the tails, the fit is even better. For all currencies against the US dollar, the test fails to reject the Student-t fit both for the 5% and 10% categories. Note that this result also holds for the exchange rates for which the Student-t fit over all observations was rejected: the Japanese Yen, the Spanish Peseta, the Swedish Krone and the Swiss Franc. This result implies that one could safely use the Student-t distribution to obtain tail information for these floating rates in practice.

Especially the fit of the tails is relevant here. Although the Student-t fits the overall distribution rather well too, we stress that if the goal is to fit the complete distribution, empirical distribution kernel estimates and splines are likely to produce even better fits. However, if the Student-t distribution could provide a very good local fit for the tails it would bring the opportunity to calculate exceedence probabilities or quantiles easily and accurately. Especially for events that occur with a probability smaller than $1/n$, a good fit of the tails through a Student-t distribution is potentially worthwhile. Extrapolation out-of-sample using the functional form of the Student-t can therefore be done with great ease.

The results are not that strong for the exchange rates vis-à-vis the German Mark. The Student-t is rejected to fit the tails for the Belgian Franc, the Danish Krone, the French Franc, the Irish Punt, the Italian Lira, the Norwegian Krone, the Spanish Peseta, the Swedish Krone and the British Pound. Still, the Student-t is accepted to fit the tail for all other exchange rates. One reason why the Student-t distribution is rejected in the tails for the exchange rates mentioned above might be the realignments that took place in the EMS. To test for the impact of these realignments we computed HKKP tail-index estimates from a sample where all realignment observations were excluded. Fitting a Student-t with the new tail-estimate for the 5% smallest and 5% largest observations leads to goodness-of-fit statistics in the 5% (R) column. The results improve somewhat and show that we cannot reject the Student-t distribution for the Danish Krone and the Irish Punt. Note that the Student-t is rejected to fit the whole distribution but accepted to fit the tails for the Belgian Franc, the Danish Krone and the Irish Punt due to the fact that all of these exhibit highly skewed distributions. Clearly, realignments do not provide the complete story about the difference between fixed and floating exchange rate return distributions.

5.4 Conclusions and practical implications

In this chapter, we focus on the tail characteristics of the unconditional distribution of exchange rate returns. It is a stylized fact that exchange rate returns are fat-tailed. However, methods used to compute the degree of tail fatness generally suffer from small sample bias. We present the unbiased HKKP tail-index estimate, which essentially is a modified Hill-estimate corrected for small sample bias. The correction exploits the fact that the bias is an approximately linear function of the number of tail observations used.

We apply this method to currencies of 16 countries, measured against the US dollar and the German Mark for the period 1979-1996. We conclude first, that the amount of tail fatness in exchange rate returns has been seriously over-estimated in the past, especially for many floating rates. Second, the tail-indexes for the left tail (negative returns) differ significantly from the right tail for many exchange rates; the assumption of an underlying symmetric distribution has to be rejected, therefore. Third, the tails of EMS exchange rate returns are significantly fatter than the tails of floating rate returns. Fourth, the Student-t distribution is statistically accepted to fit the tails of all floating exchange when the degrees of freedom parameter is set equal to our HKKP tail estimate. For some of the exchange rate returns in terms of the German Mark, apparent skewedness and realignments in the EMS result in somewhat poorer fits even when realignments are excluded. More research is needed to better understand the difference between fixed and floating exchange rate return distributions. Nevertheless, the Student-t is still accepted to fit the tails well for almost all exchange rates under consideration.

The results put forward in this chapter have important implications for managing exchange rate risk. Risk on extreme events on floating exchange rate positions has seriously been over-estimated in the past. Furthermore, one needs to analyze positive and negative returns separately and cannot directly assume that the probability on extreme events is the same for short and long foreign exchange positions. Also, the probability on extreme events is much higher for EMS rates. Finally, the Student-t distribution can be used as an accurate approximation for tails of the unconditional distribution of foreign exchange rate returns if one sets the number of degrees of freedom equal to the unbiased HKKP tail-index estimate. This holds for all floating rates and for most of the non-floating (EMS) exchange rates that are expressed against the German Mark. Exceedence probabilities and quantile estimates used as inputs for Value at Risk formulae can easily be obtained by extrapolation of the fitted Student-t distribution. This is especially attractive for the analysis of the risk on out-of-sample events which occur with probability smaller than $1/n$.

5.5 Appendix: Goodness-of-fit test

The goodness-of-fit test compares the observed and expected number of observations in c intervals over which the data is divided as follows:

$$(38) \quad G = \sum_{j=1}^c \frac{(o_j - e_j)^2}{e_j},$$

where o_j and e_j are the observed and expected number of observations in interval j . The test statistics G is chi-squared distributed with $(c-1)$ degrees of freedom. The null hypothesis of the Student-t being a good fit is rejected at the γ significance level if G is larger than the $(1-\gamma)^{\text{th}}$ percentile of the chi-squared distribution. We set the number of intervals c equal to 20.³⁷ The intervals (except the first in the left tail and the last in the right tail that range to minus or plus infinity respectively) are chosen such that they are of equal length and that they have at least 5 expected observations (which is the minimum for the test statistic to be chi-squared distributed).

³⁷ As Boothe and Glassman (1987), who use 14 intervals, we find that our results are very robust with respect to the number of intervals chosen. See Bain and Engelhardt (1987) for more details on goodness-of-fit tests.

ADVENTURES

II

THE MICROSTRUCTURE OF MULTIPLE DEALER FINANCIAL MARKETS

6 Introduction to the 2nd part

Since 1986, we have witnessed a strong increase in competition between the European exchanges. In that year, the London Stock Exchange introduced a new electronic system of quote display (SEAQ: Stock Exchange Automated Quotation) that replaced the trading floor. Standards for trade reporting were relaxed. This radical change of trading structure is since then referred to as the 'Big Bang'. Big Bang resulted in a much more liquid market with significantly lower spreads, especially for large transactions. For this reason, SEAQ I (SEAQ International) became the focus of an important market in international equities in London in the decade following Big Bang. In particular, wholesale trading in European blue chip companies gravitated to London. Pagano (1997) overviews this period and gives the following description of the competition between European exchanges: 'In the late 1980's, London appeared on its way toward becoming the main marketplace for all the blue-chip European stocks. But in those same years, the main Continental exchanges underwent a radical restructuring of their trading mechanisms and regulations, introducing continuous, electronic order-driven systems, liberalizing access to its membership reducing transaction taxes etc. This strategy allowed them to stage a formidable comeback in the early 1990's'. The trading process thus played an essential role in the competition between financial markets. In response to the Continental developments, London brought its trading mechanisms more in line with the Continental exchanges.

The introduction of the Euro and the start of the European Monetary Union will further enhance this process of competition. As capital mobility increases within an integrating Europe, investors increasingly compare the several exchanges where the stock of their focus can be traded on. They will trade there where the costs of trading are lowest. Furthermore, companies that issue new stock are increasingly willing to list their

stock on the exchange where the transaction costs of their shareholders are the lowest. With the focus on trading costs, the importance of the trading process design increases strongly since the outcome of the trading process directly influences the performance and trading costs of exchanges.

The direct influence of trading process design on market performance is the central issue in the following chapters. The market microstructure literature has concentrated on this issue in the last decade and a half. Although many theoretical papers have been published on the issue, most of them have focused on markets in which only one market maker or dealer makes the market for a single stock. Markets like the New York Stock Exchange (NYSE) for example. However, many financial markets worldwide are organized in such a way that more than one dealer makes the market in a particular stock. These are called multiple dealer markets and examples are the London Stock Exchange, NASDAQ, and the interbank foreign exchange market. Moreover, many markets have installed 'upstairs' markets in which large orders can be traded under different exchange trading rules. These markets can be found on the NYSE and on many Continental European exchanges and are all organized as multiple dealer markets. Despite the importance of multiple dealer markets they have gained much less attention than the single dealer markets in the literature for reasons explained below. Until now basically, as we shall focus directly on the performance of multiple dealer markets in the following chapters.

6.1 Market microstructure

The market microstructure literature focuses on the question how the particular trading mechanism installed on a financial market influences the performance of that market and the price-path of the securities that are being traded on that market. The literature has grown extensively since the theoretical papers by Kyle (1985) and Glosten and Milgrom (1985) and is excellently reviewed by O'Hara (1995). Typical in the microstructure literature is the following setup. The financial market is the place in which potential buyers and sellers for a specific asset meet. The asset is being traded for a finite period of time. At the end of this period, the asset liquidates and the liquidation value equals V (think of a single trading day in a stock where an analogy of the liquidation value is the closing price). The liquidation value is not common knowledge at the beginning of the trading period. Transactions in the asset occur due to the intervention of dealers or market makers. These are obliged to make a market in the asset in the sense

that they set a bid price that sellers can hit and an ask price that the buyers can hit. This combination of prices is called the quote and the quality of the quote influences how much of the asset is being traded. In the Kyle and Glosten and Milgrom model, a single dealer (the specialist as she is called on the New York Stock Exchange) is confronted with the actions of traders that represent the potential buyers and sellers. Two types of traders are distinguished. The first type represents the informed traders. They have perfect information about the asset that is being traded in the sense that they have a perfect signal about V . The informed traders exploit their knowledge strategically. They buy when the ask price of the dealer is below V or sell when the dealers bid is above V . Note that the informed traders make money on any trade they set as they exploit their excess information. In turn, the dealer loses on those trades. The second type of traders are the liquidity or noise traders. Their trades are assumed to be not information driven. In other words, these traders have no information about V and trade in the market for many possible reasons other than exploiting excess information (in order to rebalance investor portfolios for example). The trades initiated by liquidity traders are important for the dealers as the latter can offset the losses from their trades with the informed traders with the trades set by the liquidity traders.

The following extremely stylized example gives some insight in the implications of the above setting. Suppose that there is one dealer that is confronted with the trades of one informed trader and two liquidity traders. The market is quote driven in the sense of the Glosten and Milgrom model. This implies that the dealer has a quote outstanding that the traders can hit. Furthermore, assume that one liquidity trader has to buy one asset and that the other has to sell one and that the informed trader trades only once maximally. The liquidation value V of the asset is 120 and is known by the informed trader only. The following table shows the effects of different quotes on the profits of all agents.

bid	quote		total profits / losses		
	ask	spread	liquidity traders	informed trader	dealer
90	110	20	-20	10	10
95	105	10	-10	15	-5
115	125	10	-10	0	10
140	150	10	-10	20	-10

Consider the first row in the table. The dealer sets a bid price of 90 and an ask of 110. The difference between the ask and the bid is called the spread and its size is 20. The informed trader hits the ask of the dealer and buys one share for 110. Knowing V this implies a sure profit of 10. The dealer thus makes a loss equal to 10. One liquidity trader buys her asset for 110 and the other sells for 90. The first wins 10 whereas the second loses 30. In total the liquidity traders thus lose 20 on their trades and, being the counterparty, the dealer wins 20 on the trades of the liquidity traders. Note that the dealer compensates her loss of 10 of the trade with the informed with the 20 earned at the expense of the liquidity traders. It is important to see that the gain of 20 from the liquidity traders is equal to the spread size. The spread is the tool of the dealer to make money. The wider the spread the more she earns from the liquidity trader, but then the latter has less incentives to trade in real life markets as she perceives the spread she faces as transaction costs. If the spread becomes narrower as for example in the second row of the table, the dealer receives too little compensation to offset the loss incurred from the informed trader. The trick of the dealer is thus to set the 'right' spread.

There are two factors that define the quality of a quote. One is the spread size that has been illustrated above. The second is the level of the bid and ask prices. The third row in the table shows a case in which the liquidation value ($V = 120$) lies between the bid and ask of the dealer. The informed trader cannot initiate a profitable strategy as she can buy for 125 or sell for 115. The dealer does not lose any money and earns the total spread from the liquidity traders. The informed trader thus only gains if the bid and ask do not bracket V (see rows 1,2, and 4) and makes more money as the prices deviate more from V . So the quality of a quote is affected by the level of the bid and ask.

Although the above example was highly simplified it reflects the nature of financial markets and that is the basis of all microstructure models. Dealers set quotes to attract order flow from the liquidity traders to make money in order to offset the losses incurred by the transactions from the informed traders. The dealer in the above example could have made the most money by setting her spread as wide as possible to get rid of the informed trader but this is not realistic in real life financial markets. Liquid supply and demand is a function of the spread; the wider the spread the less liquidity traders are attracted as they face high transaction costs. In multiple dealer markets where more than one dealer sets a quote for one particular asset, the price competition between dealers puts a competitive pressure on the spread. The dealer with the most competitive price

gets the liquidity trades, as well as the informed trades if the level of the quote is not in line with the assets liquidation value. The spreads, the quote levels and the profits of the agents in the market all say something about the quality of the market and are important market performance variables for this reason. The market microstructure literature therefore focuses on these variables to assess the effects of trading mechanism design on the performance of a financial market.

Although the research on dealer markets has grown extensively since the theoretical models of Kyle and Glosten and Milgrom, it concentrated mainly on single dealer or specialist markets as for example the New York Stock Exchange. This focus on single dealer markets can be explained by the difficulties that a market with multiple heterogeneous dealers introduces relative to a single dealer market. For example, imagine that in the above example of a financial market two dealers instead of one compete for order flow. They compete in terms of the quotes they set. These quotes reflect their willingness to trade, the trader with the best price gets the order flow. To cope with these heterogeneous dealers in theoretical models is tedious unless one restricts the price behavior of dealers. However, the trading process on many important exchanges is organized as a multiple dealer market. Examples of these are for example NASDAQ (National Association of Securities Dealers Automated Quotation), the London Stock Exchange SEAQ International and the interbank foreign exchange market. In addition, European equity markets have converged to a dual trading system in the late 1980's and 1990's, combining single dealer markets and limit order books for small transactions and dealership markets for large transactions.

Many have studied the effects that trading mechanism characteristics have on the performance of multiple dealer markets using an empirical methodology³⁸. Two important caveats arise when an empirical methodology is used. The first caveat lies in the fact that it is very hard to disentangle the effects of trading mechanisms alone if a comparison is made over different time periods with different states of the world or over exchanges in different countries that have a different economical environment. The second caveat lies in the fact that many data sources do not record exactly the informational process like what type information enters the market and at what point in time; important examples of these missing 'information observations' are streettalk and rumors. One way to overcome these problems is to record exactly what market makers

have done and what happens in their environment by observing one or more dealers closely for a specific period in time as is done by Lyons (1995) regarding the foreign exchange market.

In order to overcome many of these difficulties we shall apply experimental economics. The basic idea is to bring the setup as given in the example to the laboratory; dealers make the market for an imaginary stock and are confronted with the actions of liquidity and informed traders. As some of the roles are played by human subjects, we do not need to restrict behavior and can confront them with various settings without changing the economic environment. Although experiments offer many opportunities, one could possibly introduce problems of another kind. The comment heard most commonly is that experiments lack reality, i.e. it is hard to get the natural feeling and sensation of the trading process in a laboratory. Although this is true one could argue that this is an advantage of experiments over real life data, as many disturbing effects do not take place in an experimental environment. When the research goal is to compare two trading regimes, one is able to do this in exactly the same economic environment.

Perhaps the greatest difficulty lies in bringing the real life behavior of economic agents in the laboratory. The experimental economist wants his or her subjects to act as they would do in reality. The tool most experimentalist use is the payoff scheme that defines how the subjects are rewarded for their actions. Although the payoff scheme works satisfactorily in many experiments, its relevance is flawed when the focus of the research is on the trading process within financial markets. The feeling of investing millions of dollars in a risky asset is hard to bring alive in the laboratory. Furthermore, trading shares is different from buying a car or a house in the sense that the price process on financial markets consists of two prices instead of only one price (the bid and the ask) which also introduces the spread size as an important variable. For this reason, one needs to be careful when selecting subjects. Although many experiments in psychology rely on 'people from the street', it is hard to get them used to trading even after a few days of training. Therefore, many rely on MBA or PhD students in finance as subjects³⁹. We have used professional traders of Dutch commercial banks and market makers from the European Option Exchange in Amsterdam as subjects. The advantage of using these professionals is that they are almost directly familiar with the experimental trading

³⁸ Some examples of these empirical studies are Porter and Weaver (1997) for NASDAQ, Naik, Neuberger, and Viswanathan (1994), de Jong, Nijman, and Röell (1995), Gemmill (1996) for the London Stock Exchange and Evans (1998), Lyons (1995, 1998) for the foreign exchange market.

process. We thus do not need to simplify the experimental trading process for the subjects, which offers more opportunities to build in many real-life aspects. An important caveat exists in the fact that these professionals are less influenced by the payoff scheme. However, an important experience we learnt from the experiments is that 'professional pride' takes over some of the functions of the payoff scheme, although this does not solve all the problems.

The subjects in our experiment deserve further comment. We emphasize at the outset that our subjects are professional traders and that the experimental software we employ is designed to mimic trading systems that traders actually employ. This makes us less vulnerable to the criticism that they lack professional expertise as might be the case with students. Finally, the financial reward for success within the experiments is not large. However, this does not imply that our subjects do not take the trading seriously. In fact, these traders know each other well; the rivalry this implies generates a certain 'poker-table competitiveness'.

6.2 Objectives

The experimental methodology is applied to examine the effects of trading mechanism characteristics on the performance of financial markets. One of the fundamental characteristics of trading mechanisms that has received much attention in the last decade is transparency⁴⁰. The transparency of a particular financial market reflects the way information that is a part of the trading process is transmitted among the market participants. Two information flows are a part of the trading process: information before a transaction is cleared and information that is revealed after a trade is settled. Transparency with respect to the first information flow is called pre-trade transparency whereas the latter information flow is referred to as post-trade transparency. In chapters 8 and 9, we shall focus on pre-trade and post-trade transparency respectively. The effects of the amount of pre-trade transparency on market performance is examined in terms of quote disclosure, to study the differences between markets where all quotes are disclosed immediately versus markets where no quotes are disclosed and where market participants obtain price information by calling one another. The effects of post-trade transparency are examined in terms of the disclosure of transaction information, to study the

³⁹ See for example Bloomfield and O'Hara (1996), Schnitzlein (1996) and Lamoreux and Schnitzlein (1997a, 1997b).

differences between markets where transaction information of all trades on the exchange is disclosed immediately and markets where market participants only have transaction information about the trades they participated in. Chapter 10 analyses the costs of trading in a financial market in terms of the spreads market participants face. The experimental methodology is used to study a possible spread component that could otherwise hardly be obtained from transaction data: search costs. These are the costs market participants should make to obtain price information. In chapter 11, we review the impacts of transparency and search costs, and discuss their role in the competition process between exchanges.

The design of the trading mechanism influences the performance of financial markets importantly since the trading mechanism controls the routing of information and the actual process of trading. Chapters 8 and 9 focus on the information routing function whereas a variable related to the actual trading process is discussed in 10. Chapter 11 summarizes the results and concentrates on the importance of trading mechanisms in the competition process between exchanges given their relation to the market's performance.

⁴⁰ Pagano and Röell (1996) review the impacts of both types of transparency for various market structures.

7 The Experimental Environment

The following chapters concentrate on trading mechanism characteristics in the performance of multiple dealer financial markets and their influences on overall market performance. The analyses are based on an experimental methodology from which the general settings will be presented in this chapter; the exact details for the experimental environment studied in a specific chapter are discussed there.

7.1 Overall

The experimental methodology involves a computerized experimental securities market in which a specific number of professional dealers act as market makers. They trade with each other and with a number of computerized traders operating as non-market-making customers of the exchange.⁴¹ Each experiment consists of a number of mutually independent trading rounds. Each round is independent, in the sense that the parameter settings are unique for each round; information about the security's value derived from one round is not relevant in another.

7.2 The microstructure

The microstructure of our experimental environment can be seen as a continuous, multiple-dealer version of the pure dealership market introduced by Glosten and Milgrom (1985). In this model, a specialist sets quotes for investors, who view the quoted bid and ask prices and decide whether to trade one unit at a time. The specialist is free to

reset the bid and ask at any time. The investors represent both informed and liquidity motivated traders. Informed traders have received a perfect signal about the asset being traded and trade to exploit their excess information. Liquidity motivated traders trade for reasons that have nothing to do with information. They trade for portfolio rebalancing needs for example. The informed and liquidity traders do not price compete with the specialist, since they do not set limit orders in the market. Note that the Glosten and Milgrom market is quote driven; in this sense, our experimental design differs from the framework introduced by Kyle (1985). In the Kyle model, the quantities traded by both informed and liquidity traders are batched, and the market maker determines the market-clearing price.⁴² Our quote-driven experimental design is to a certain extent similar to the quote-driven experimental markets used in Bloomfield (1996), Bloomfield and O'Hara (1996). Compared with the experimental design in these studies, there are three important differences. First, the market makers in our market both set quotes (this is obligatory) and initiate trades (voluntary). We thus allow for interdealer trading, which is a salient characteristic of multiple-dealer markets.⁴³ Second, our market is continuous instead of sequential. A sequential framework is organized by several stages of trading where a certain number of actions - like the transactions of all market makers and investors - are undertaken in each stage. Our continuous market has the feature that market makers may trade and revise their quotes at any time; continuous trading has the notable benefit that it yields large amounts of data. Third, we use professional security traders as subjects rather than students.

The order driven Kyle model is the underlying framework of the experimental markets used by Schnitzlein (1996) and Lamoreux and Schnitzlein (1997a, 1997b). The choice of modeling either a quote-driven or order-driven market is not a choice made on performance characteristics or importance. In real life, many market structures co-exist. Furthermore, many financial markets like the NYSE, the Amsterdam, Frankfurt, London and Paris stock exchanges combine a multiple of trading systems, mostly a combination of (limit) order books for small transactions and quote-driven dealer markets with relaxed standards regarding information disclosure for large transactions.

⁴¹ We use the terms "market makers", "traders" and "dealers" interchangeably to refer to the human subjects. Similarly, we use the terms "robots" and "customers" interchangeably to refer to the computerized non-market-making customers.

⁴² See Madhavan (1992) for an overview of the differences between quote-driven and order-driven markets.

⁴³ For example, in the foreign exchange market, Perraudin and Vitale (1996) report that the magnitude of interdealer trading is roughly 80 percent of total volume.

7.3 Dealer objectives and the liquidation value of the asset

At the start of each trading round, each market maker is given an initial endowment of 1000 esquires (a fictional numeraire currency used in the laboratory). All traders are instructed to maximize their end-of-round wealth by trading for a single security. End-of-round wealth is expressed in esquires; it equals the cash balance at the end-of-round plus the end-of-round inventory valued against the liquidation value or true price of the asset. The true price of the asset can be seen as an *ex post* liquidation value as used in, among many others, the Glosten and Milgrom (1985) and Kyle (1985) models. This liquidation value differs across rounds to achieve independence of the rounds.

Before the start of the experiment, the market makers are told that the true price is uniformly distributed between a minimum of 1 esquire and a maximum of 200 esquires; this is the only *a priori* information traders have about the true price at the start of every round. The dealers can obtain information about the true price through transactions with the informed investors, as we shall discuss below. Dealers are not instructed about possible trading strategies. Each dealer is allowed to trade according to his or her own expectations and predictions. Traders obtain their end-of-round profit by buying and selling (i.e., jobbing) and/or by building a long or short inventory of the security that is to be converted to cash at the true price at the end of trading (i.e., speculating on the true price). During each round, traders can trade at any time they desire in unit quantities.

The true prices are chosen as follows. First, recall that we tell the market makers that the true price lies between 1 and 200 before the start of the experiment. Therefore, a uniform prior expectation of the true price would be 100 esquires; we refer to this as their initial guess. In the experiments, we want to have instances of both upward and downward price discovery as both positive and negative trends exist in real markets. Therefore, we draw the true prices for half of all rounds randomly from the [1,100] interval and for the other rounds from the (100,200] interval. Within these intervals, we would like to have prices both far away and close to the initial guess of 100. Therefore, we chose an equal number of prices uniformly from each of the following four intervals: [1,50), [50,100), (100,150] and (150,200]. The true price settings for each study are reported in the summary statistics tables presented in the corresponding chapters.

7.4 The trading process

Each market maker is obliged to enter a quote - both bid and ask prices - within 10 seconds of the start of each round (otherwise, a penalty of 10 esquires per second accrues). The dealers can change their quote at any time during the round, but cannot withdraw their quotes. Therefore, at any time all traders have a quote outstanding in the market. The maximum spread is limited to 30 esquires.

In the following chapters, we study two pre-trade transparent trading systems: one with and one without public disclosure of quotes. In the first market, all quotes are continuously visible on the trading screen. The bids (asks) are ranked from the highest (lowest) on top through the lowest (highest) at the bottom, so that the market spread is always obtainable at the top of the queues; the highest bid and the lowest ask outstanding are directly obtainable. To buy securities, the trader simply presses the "B" key; the trader automatically buys against the lowest ask available. To sell, the trader presses the "S" key; the trader automatically sells against the highest bid. After this the trader that called enters the number of shares he or she wants to trade. In the market without public quote disclosure, information about the quotes of other market makers is, by definition, not publicly available. Instead, in order to obtain direct price information and to trade, traders have to call one another bilaterally. This is done by pressing the "C" key and entering the identity of the trader whose quote the caller wishes to observe.⁴⁴ The quote of the dealer called then appears on the caller's trading screen for a maximum of 7 seconds, during which time the caller can initiate a trade. To buy assets, the trader presses the "B" key; to sell the "S" key. As in the Glosten and Milgrom (1985) model, we normalize the trade size to exactly 1 share per transaction; there is, however, no limit on the number of transactions other than the length of each round.

After the transaction is cleared, the trade details are presented in the private transaction history window of both counter parties. We shall study two post-trade transparency regime, i.e. the disclosure of transaction information, in the following chapters. In the opaque regime, a market maker only views the information about the transactions in which he or she directly participated. In the transparent regime, a market makers also views all transactions of the investors, including those he or she did not participate in.

⁴⁴ The identity of traders is denoted with letters while "R" refers to the investors or the robot. Robots do not quote prices, however.

7.5 The behavior of investors

As discussed previously, two types of agents are involved in the experiment. In addition to the (human) market makers, who set quotes and trade, there are two computerized investors (robot) that trade in the market. The robot initiate trades by themselves do not set quotes and cannot be called by the market makers. The robots represent external customers of the exchange and proxy for both informed traders and liquidity or noise traders. Whether a certain robot transaction was based on information or noise is not distinguishable to the market makers. The robot is programmed to check every 3.5 seconds whether it can trade profitably in the market. If so it initiates the trade against the best price in the market.

Whether a robot trade is to be informed or uninformed is determined at random, just prior to each robot transaction; the probability that the trade is informed equals one half.⁴⁵ The market makers are told this probability; they therefore know that, on average, half of the robot trades are informed. If the robot initiates an informed trade (with probability $\frac{1}{2}$), it buys (sells) if the lowest ask price (highest bid) outstanding at that time is below (above) the true price of the security. An informed robot does not trade if the best price equals the true price. Note that the informed robot tries to maximize its profit only at the trade level; over the whole round it is restricted to trading only at multiples of 3.5 seconds. Since the informed robot trades are a function of the true price, these transactions convey information about the true price of the asset. If the robot initiates an uninformed trade, it is determined with probability $\frac{1}{2}$ whether the robot sells or buys; if it sells (buys) it does so against the (highest bid price) lowest ask price available in both market structures.

When a market maker is involved in a robot transaction, he or she receives a special notification. That is, as for the interdealer trades, the transaction information appears in his or her private transaction history window with identity "R" for robot. Note that the trader does not know whether the robot trade was informed. Given their knowledge of the probability that a robot trade is informed, traders must filter the relevant price information from these robot transactions. If a robot trade is informed and the robot sells (buys), it is a clear signal that the bid price (ask price) of the trader is above (below) the true price of the asset.

⁴⁵ This setup of the noise structure parallels the setup in Glosten and Milgrom (1985) and Röell (1996).

7.6 The trading mechanism

Each market maker has his or her own computer trading screen and a keyboard. A sample computer screen appears in figure 10.

Your Position		Price Quotes		Transactions	
Money	: 790.00	ID Sell	Buy ID	#	buy sel units price
LIFE Shares:	2			1	* B 1 100.00
Your quote :				2	E * 1 110.00
bid:	90.00 ask: 110.00			3	* J 1 110.00
Approximate wealth				4	* R 1 110.00
based on last price:					
1010.00 esquires					
Time remaining: 158 sec					

PageUp/PageDown = scroll

10 The screen of the trading mechanism

This is an example of a trading screen in the market without quote disclosure that is used in the following chapters.

The window on the upper left presents the cash balance (790 esquires), the inventory position of this dealer (2 LIFE shares long), the current outstanding quote of this trader (90 - 110) and the approximate profit of this trader based on the price of the last transaction in which he or she was involved (1010 esquires). The middle window on the left side shows the time remaining in this round. The black window in the center of the screen is where a quote appears when this trader calls another trader. Under the heading « ID Sell Buy ID », ID denotes the identity of the dealer presenting the quote and Sell denotes the quoted bid (at which this trader can sell); the Buy column contains the

quoted ask (at which this trader can buy the share). In the markets where all quotes are disclosed publicly, all bids are presented below Sell ordered from high to low and all asks are presented below the heading Buy ordered from low to high. At the right of the trading screen, information of past transactions is presented. By default, it displays the details of the last 20 transactions. The trader can scroll through the list with the PageUp and PageDown keys. For all transactions, the identities of the buying (under heading buy) and the selling (sel) trader is displayed, along with the number of shares involved in the trade and the price at which the trade cleared. For example, the first row indicates that this trader (his identity is shown as an asterisk *) bought one LIFE share for 100 esquires from trader B. The identities of the market makers are denoted by letters starting with 'A'. The robot transactions are denoted with the letter 'R'. The fourth transaction is thus an example of a trade in which the robot sold to this trader one share for 110 esquires.

8 Quote Disclosure

This chapter concentrates on pre-trade transparency as a characteristic of the institutional design of financial markets. We are most interested in transactable prices (i.e., live quotes), rather than indicative prices or post-trade transaction prices. Specifically, we ask how the degree of public quote disclosure of transactable prices affects important market characteristics, such as transaction costs and the dynamic adjustment to strong-form efficient prices (price discovery).

Worldwide, many financial markets coexist while differing crucially in the way market participants receive price information. For example, electronic limit order books, such as those on the Paris Bourse, are very price-transparent, since brokers can obtain information on quotes and order-sizes from all limit orders available in the book. On the other hand, the London SEAQ display system (and, to a lesser extent, NASDAQ) provide participants with quotes, but these are largely indicative, since most market makers will improve upon their quoted bids and asks in telephone negotiations, offering within-the-quotes prices for most transactions. Even close substitutes, such as foreign currency futures and forward contracts, trade in markets with very different degrees of quote disclosure. One explanation for these coexisting quote disclosure regimes in financial markets is clearly found in the type of securities being traded. For example, futures trading is typically centralized (while forwards trading usually is not) at least in part because futures contracts are standardized. However, it is frequently the case that the same security trades in different microstructures, often simultaneously, and sometimes even in the same room. For example, the NYSE runs a call market at the morning opening, and continuous specialist trading throughout the day, along with a separate upstairs market for block trades. The foreign exchange market operates

⁴⁶ This chapter is based on: M. Flood, R. Huisman, K. Koedijk, and R. Mahieu, 'Quote Disclosure and Price Discovery in Multiple Dealer Financial Markets', *Review of Financial Studies*, Spring, 1999.

simultaneously as a decentralized interbank direct market and a more centralized brokered market.

It is generally assumed that price-transparent microstructures better allow traders to extract information from outstanding quotes, leading (rather intuitively) to prices impounding a maximum of available information. For example, O'Hara (1995, p. 270) notes that, "the transparency of prices allows traders better ability to extract price information from the market price, a process that surely abets the goal of equilibrium price discovery."⁴⁷ From this point of view, it is not obvious why markets with different levels of pre-trade transparency coexist. In this chapter, we measure the effects of pre-trade transparency on the pricing and inventory strategies of market makers, in terms of liquidity (spreads and volume) and the price discovery process.

Despite the numerous studies of post-trade transparency (i.e., the amount of transaction information available to market makers which will be discussed in the following chapter), the effects of (pre-trade) quote disclosure on market performance have received much less attention.⁴⁸ Madhavan (1992) distinguishes between "order-driven" and "quote-driven" markets, comparing their relative price efficiency. In quote-driven markets, investors trade against outstanding prices, while in order-driven-markets participants must submit orders first, after which prices are determined. The quote-driven markets are necessarily more pre-trade transparent than order driven markets. Madhavan finds that a quote-driven market is more price efficient than an order-driven market, but that the equilibria coincide. Bollerslev and Domowitz (1993) compare different "book lengths" for an electronic limit order book (patterned on the CME/Reuters Globex trading system) in a computer simulation. Their general conclusion is that increasing book length - the number of limit orders held on the book and broadcast to participants - improves price discovery (with the exception of trading in unit lots, for which a technical artifact of the order-posting algorithm slows price discovery).

Using the experimental environment that is discussed in the previous chapter, we compare directly the effects of two market structures on the pricing and trading strategy

⁴⁷ There is a vast literature on market transparency, including empirical and experimental studies of price discovery. Surveys of the relevant microstructure literature are available in O'Hara (1995, esp. ch. 9), Yu (1993), and Admati (1991). Stoll (1992), and Harris (1993) survey issues related to institutional design. Goodhart and O'Hara (1994), and Guillaume, et al. (1994) cover recent empirical results. Davis and Holt (1993, esp. chs. 3 and 5) and Duxbury (1995) survey the experimental literature.

⁴⁸ Recent examples of the former include Bloomfield and O'Hara (1996), Gemmill (1996), Pagano and Röell (1996), Röell (1996).

of professional market makers, and therefore on the price efficiency of those markets. The markets differ only in the way quotes are disclosed.

8.1 Experimental Design

The effects of quote disclosure on market performance are examined by combining the data obtained from two market types. In the transparent market, all quotes are disclosed publicly and immediately, which means that all market makers have all outstanding quotes presented on their private trading screens. In the opaque market, no quotes are publicly disclosed, and the market makers must "call" one another for price quotes. See chapter 7 for more details about the trading mechanism. The market is low post-trade transparent in the sense that each market maker only sees the history of transactions in which she directly participated.

8.1.1 *Payoff to the subjects*

The payoff scheme is most commonly used to have the subjects act properly regarding their economic behavior. At the end of each experimental round, the market makers are informed about the true price of the asset and their end-of-round wealth. Esquire amounts are transformed into guilders as follows. The dealer with the greatest end-of-round wealth receives 7 guilders; the second best receives 6 guilders etc.⁴⁹ The traders are informed about this payment arrangement before the start of the experiment.

The guilder payoffs are not an affine function of the dealers' esquire results. Instead, the guilder payoffs are, in effect, a simplistic ordinal function of esquire results. We implemented this payoff function for practical reasons. However, we argue that any bias induced herewith must be small. First, it is our experience that the professional traders who participated as subjects were strongly motivated by professional pride (in addition to any direct interest in the cash payoffs, which are ultimately relatively small by their non-student standards). The role of professional pride as compensation is essentially untouched by the literature on experimental rewards. It is reasonable to suppose that such pride is approximately linear in esquire payoffs, since the dealer compensation schemes that they face in their professional lives are typically linear in trading profits (and explicitly so).

⁴⁹ At the time of the experiments, the guilder-dollar exchange rate was roughly: 1 USD = 1.90 NLG.

Moreover, a caveat arises because of the ex-ante truncation of the guildler payoffs at a maximum of 7 guilders per round. The problem alleged here is that this truncation may induce risk-aversion (when a dealer is convinced in mid-round that her profits are already unbeatable). This induced risk-aversion is by design, however, as it resolves a potential moral hazard problem that arises in the absence of truncation. This problem cannot be alleviated by simply re-calibrating the payoffs (although it is of secondary importance when working with professional subjects, as noted above). To exploit the moral hazard, a subject reasons (correctly) that the ex-ante optimal inventory is infinite, either long or short (it doesn't matter which). Thus, if the true price is revealed to be higher than the cost of inventory, the subject collects a large profit; if the true price is lower, the subject cannot – by law – leave the lab with less cash than zero. (Precisely this moral hazard situation arose in very early pretests of the software prototype using student subjects.) In financial practice, it is analogous to the role of deadweight bankruptcy costs in inducing an optimal interior debt level for firm financing, or, more pertinently, it is analogous to the role of dealer position constraints (enforced by job termination and other sanctions) as a means of controlling dealer risk-taking.

8.2 Data

The data in this chapter were collected from an experiment that was held at the Center for Research in Experimental Economics and Political Decision Making (CREED) at the University of Amsterdam on 29 April 1997. The subjects were seven professional traders from de Generale Bank, Optiver, and Rabobank. All subjects traded as market makers in eight independent rounds. Each round lasts five minutes.

Table 18 presents the settings and summary statistics. We started with four rounds with markets without public quote disclosure, followed by four rounds where all quotes were publicly disclosed. This order of microstructures might have caused learning effects over the rounds; i.e., skills formed in the first market might have enhanced trading behavior in the last four rounds. However, this chapter mimics the experiments from Flood, Huisman, Koedijk and Mahieu (1996), with qualitatively very similar results. We reverse the sequence of the market structures relative to Flood, et al. who started with public price disclosure and ended with opaque markets. Since the results we present in this chapter are qualitatively similar to the results they obtained, we feel comfortable about the robustness of the results.

18 Summary Statistics

In this table, we provide information about the experimental data that we obtained.

	No quote disclosure (opaque market)				Public quote disclosure (transparent market)			
	1	2	3	4	5	6	7	8
Round settings								
True price	69	163	134	15	22	75	118	185
Results								
# quotes set	80	72	64	66	76	83	80	88
# trades								
- total	461	473	533	549	1300	1557	1841	1423
- interdealer	387	399	459	475	1226	1483	1769	1349
- robot	74	74	74	74	74	74	72	74
% uninformed	57	38	47	46	50	46	51	51

Average end-of-round capital in esquires^a (initial capital of 1000 esquires is excluded)

market maker								
1	-486	1881	834	-5071	6127	3656	4633	301
2	29	-6869	-816	-3409	2459	1536	-319	-2745
3	-35	2686	-5894	6855	-5070	-21166	-13041	2436
4	-282	-1794	487	-2265	15074	6012	119	-14364
5	-376	466	169	882	-75664	-8302	787	13651
6	234	-1410	424	-2954	12436	6311	5565	-13683
7	571	3481	3224	3757	41689	10656	1386	12495
avg.	-49	-223	-225	-315	-421	-185	-124	-273
robots								
- informed	155	752	745	972	1549	550	328	707
- uninformed	18	28	42	131	-74	99	108	248

^a None of the average profits of the dealers and uninformed robots are significant. The average profits of the informed robots are all significant.

In table 18 we find, on average, 76 quote settings and 430 interdealer transactions in the opaque markets, and 82 quote settings and 1457 interdealer transactions in the

transparent markets.⁵⁰ There is an average of 74 robot transactions initiated per round in both market types. Table 18 also shows that, on average, dealers made small losses (although the zero-sum nature of the experiment means that if some market makers make large gains, others generally make large losses). On average, the market makers make slightly negative profits, largely because they must trade with informed robots, to which they always lose. The upper bound of thirty esquires for the bid-ask spread means that they cannot recover these adverse selection losses by trading gains from the uninformed robots. Note that market makers' average losses do not differ systematically across the transparent and opaque market settings - these are determined mostly by informed robots' profitability; however, the dispersion of profits is higher in the transparent market, reflecting the more intensive interdealer trading there.

8.3 Results

In this section, we test two hypotheses concerning the effects of quote disclosure on market performance. The tests are based on the experimental data drawn from seven professional traders over eight rounds, as discussed in the preceding section. In brief, we find a marked trade-off between liquidity and price efficiency. We conclude that both the higher liquidity and slower price discovery of the more transparent market are the result of speculating dealers exploiting the lower search costs in this market. It is worth noting that the results reported here correspond qualitatively with those found in two pilot experiments.

8.3.1 Volume

Regarding liquidity, we first observe the sharp discrepancy in trading volume between the transparent and opaque markets. The average trading volume in the markets with no public quote disclosure was 504, while the average for the transparent markets was roughly three times higher at 1530. We attribute this difference directly to transparency: dealers spent relatively less time sampling prices (and correspondingly more time engaging in transactions) due to reduced search costs in the transparent market. Note that there is both an informational component (knowing more prices sooner) and a logistical component (less time is expended canvassing the market and identifying the

⁵⁰ This higher number of transactions in the market where all quotes are publicly disclosed is largely due to the trading mechanism, since prices are more difficult to obtain in the markets without quote disclosure, which increases the time needed to initiate a transaction.

counterparty for a transaction) to this reduction in search costs. Both components attribute to the increased volume.

8.3.2 Spreads

In financial markets, the costs of market making and information asymmetry among market participants are seen as the primary motivations for the bid-ask spreads. In the literature on post-trade transparency, the general consensus is that opening spreads in transparent markets are wider, but that the difference in spreads between transparent and opaque markets disappears over time.⁵¹ Madhavan (1995) rationalizes this by stating that market makers compete for order flow to obtain information; they therefore quote more competitive spreads in (post-trade) opaque markets. This competition is reduced in transparent markets, since information is disclosed publicly, and quoted spreads therefore widen.

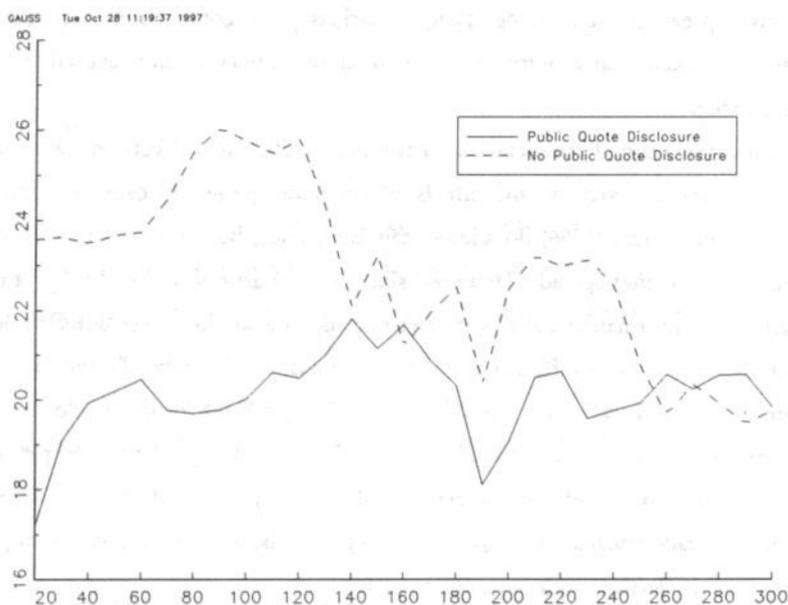
Thus, the disclosure of transaction information has substantial effects on spread sizes, but there is less evidence on the effects of pre-trade quote disclosure on spreads. Bloomfield and O'Hara (1996) do address this latter issue, but find no effect of quote disclosure on (opening) spreads. However, since spreads arise to cover the operational and asymmetric information costs of market making, one would expect different quote disclosure regimes to have different impacts on spreads. For instance, the higher search costs imposed in a market with a low level of public price disclosure should result in higher spreads. (It is worth emphasizing that the literature predicts that pre-trade transparency and post-trade transparency will have opposite effects on spreads. Increasing pre-trade transparency narrows opening spreads, whereas an increase in post-trade transparency leads to wider opening spreads.)

In addition, disclosing quotes publicly reduces the asymmetric information problem, and should therefore lead to smaller spreads. This creates a second-order effect. As trading evolves over the trading period, more information is shared by the market makers and the informational differences between market structures diminish. This would imply that the differences in spreads, due to information asymmetry, would decline over the trading period. Combining both effects, we test the following hypothesis:

⁵¹ See Bloomfield and O'Hara (1996) and the following chapter for experimental results. Röell (1996) analyzes theoretically the effects of transparency on spreads. Gemmill (1996) presents an empirical investigation of the London Stock Exchange, but finds mixed results.

H1₀: Dealer spreads are on average wider in a market without public price information than in a market in without public disclosure of quotes; furthermore, the difference between the spread sizes of both markets converges to zero over time.

To examine the spread differences between microstructures, we split each trading round in 30 intervals of 10 seconds each. In figure 11, we present the average spread size over all market makers and all rounds for all intervals in both the market with public quote disclosure and without public information on quotes.



11 Average dealer spread size over time

This figure presents the time-series patterns of the average dealer spreads, in esquires, over 30 intervals of 10 seconds each. Each line represents the average of the spreads over rounds with the same market structure, defined as the difference between the quoted ask and bid price, over all quotes outstanding during each interval.

As Madhavan (1992) predicts for quote-driven markets, the average dealer spread shrinks over the trading period in both market types, which we attribute to declining information asymmetries as more transactions are settled. Furthermore, figure 11 presents clear differences in spreads between both market types, as stated under the null hypothesis. The opening spreads in the rounds without quote disclosure are significantly wider than in the rounds with public information on quotes.⁵² Over the trading period, the market maker spreads narrow in the markets without quote disclosure such that eventually no differences in spread size are apparent. Both aspects of the null hypothesis are supported by the data.

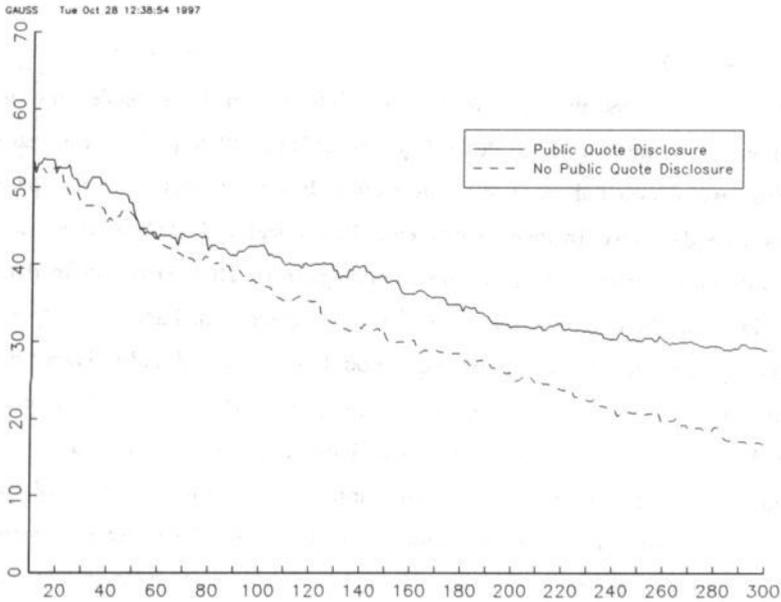
8.3.3 Price efficiency

Intuitively, one expects that the more price information is available, the more information will be impounded in dealers' quotes. O'Hara (1995, p. 270), for example, makes this case, as noted above in the introduction. In other words, price transparency improves price discovery. In their experiments, Bloomfield and O'Hara (1996) do not find a significant effect of quote disclosure on pricing errors. These errors, defined as the absolute difference between the true price of the asset being traded and the midpoint of the market spread, decline over the trading period, but Bloomfield and O'Hara find no significant evidence that they decline more rapidly in markets where all quotes are publicly disclosed. On the other hand, Flood, Huisman, Koedijk and Mahieu (1996) show that the price discovery process is significantly slower in markets where all quotes are publicly disclosed than in markets without quote disclosure. The experimental market they used is similar to the market discussed here, but without noise traders; their market makers were confronted instead only with informed customers. Nonetheless, we state a null hypothesis to conform to the intuitive consensus:

H₂₀: *Public disclosure of dealer quotes enhances price efficiency.*

⁵² Significance is at a 10% confidence level in approximately the first 120 seconds. We compared the average spreads of each market maker over both market structures. At each interval we therefore have two samples consisting of 7 matching spreads, i.e. one for each market structure. Then we conducted a two-sample sign test (see Bain and Engelhardt (1987) for more details). This test compares the medians of both samples, with the advantage that we do not need to assume normality of the spreads. We do, however, need to assume independence, but this is likely given that we are comparing two samples from different market structures.

To examine the effects of quote disclosure on prices, we first consider the time path of dealer pricing errors, defined as the absolute difference between the midpoint of a quote and the true price of the asset. In figure 12, we plot the average price error over all dealers and rounds with the same market structure for each second. We discard the first ten seconds, since this is the time-span we allowed the market makers to enter their initial quotes. The price errors decline over time and we observe that these price errors decline more rapidly in the market without public quote disclosure.



12 Average price errors over time

This figure presents the time-series patterns of the price error, averaged over all dealers and rounds with the same market structure at each second during the trading round.

Figure 12 clearly suggests differences in price efficiency between the two market structures. To examine this more formally, let $P(j,t,r)$ be the price error obtained from the quote of market maker j in round r at time t (measured in seconds). We test whether the pricing errors decline more rapidly in a transparent market by regressing the pricing

errors on a time trend variable, t , whereby the effect of quote disclosure is split over two slope coefficients.

$$(39) \quad P(j, t, r) = \alpha + \alpha_j + \beta_1 I(r \in \text{QD}) + \beta_2 t I(r \in \text{QD}) + \beta_3 t I(r \in \text{NQD}) + \varepsilon(j, t, r),$$

where $I(\cdot)$ is an indicator function that equals 1 if the condition in parentheses is true and equal to 0 elsewhere, t is the time trend, QD is the set of all rounds in which quotes are publicly disclosed, NQD is the set of rounds in which quotes are not disclosed, and ε is an error term that is assumed to be iid. Note that we pooled the data over all eight rounds and the 7 market makers in the regression.⁵³ We added specific constants α_j (individual effects) to control for the specific characteristics of trader j ; we assume that the α_j 's sum to zero. The effect of quote disclosure on the price errors is captured both by the initial level effect - measured by the market structure dummy (β_1) - and by the difference between the slope coefficients for the markets where all quotes are disclosed (β_2) and where no quotes are disclosed (β_3). Table 19 contains the estimates of these slope coefficients.

⁵³ Note that, if price discovery (defined here as quoted prices within three esquires of the true price) had been achieved in only some of the rounds, then a test with pooled data would be biased, since market makers' pricing strategy would likely change in the vicinity of the true price. However, the dealers never achieved full price discovery. Furthermore, note that robot trades can only indicate the direction of the true price (higher or lower) to the dealer, so that variation in absolute levels of the true prices across rounds should not bias the test either.

19 Price efficiency

This table contains the estimates of the slope coefficients in the fixed individual effects model (39). The time path of price errors of individual market makers is regressed on a constant, on a market structure dummy (β_1) that equals 1 if the market has public quote disclosure and 0 elsewhere, and on two slope coefficients (β_2 and β_3) for the time trend in the rounds with and without public disclosure of quotes respectively. The individuals in the panel are the market makers, so that we correct for cross-sectional differences between all market makers.

	Intercept (α)	Market Structure (β_1)	Public Quote Disclosure (β_2)	No Public Disclosure (β_3)
estimated slope coefficient	53.169 (1.926)	0.528 (3.426)	-0.098 (0.010)	-0.134 (0.007)
difference in slopes ($\beta_2 - \beta_3$)			0.037 (0.012)	

The data are pooled over 7 market makers and 8 rounds; $n = 16,800$.

For details on estimating fixed individual effects panel models, see Baltagi (1995).

Robust standard errors for a general variance-covariance matrix are presented in parentheses (see Arellano (1987)).

$R^2=0.149$

From Table 19, we first observe that all parameter estimates are significant except for the market structure dummy. The constant term equals 53, which equals the average absolute true price deviation from 100 which is the initial market makers expectation of the true price when no price information has entered the market. The market structure dummy suggests that the initial level of the price errors is not significantly different between both market structures. Since these initial price errors are equal for both market structures, we use the differences in the speed with which the price errors are corrected, captured by the slope coefficients of the linear trends in (39), as a measure of relative price efficiency. The slope coefficients β_2 and β_3 are negative and significant, indicating that the pricing errors decline over time reflecting the accumulation of information over time as trading proceeds in both markets. The difference between the two time trend coefficients equals 0.04, indicating that pricing errors decline more rapidly (the slope is more negative) in markets where quotes are not publicly disclosed. Using price errors as a measure of efficiency, we conclude that markets without pre-trade information are more efficient than markets with public prices, thus rejecting H_2 . On the other hand, this

result is in line with Flood, et al. (1996), and not inconsistent with Bloomfield and O'Hara (1996) who find no significant relationship between quote disclosure and pricing errors.

To confirm this rather striking result, we examine average dealer price responses to new information. If prices indeed impound new information, then this fact should be evident in dealer quote revisions following information arrivals. We examine the magnitude of such quote revisions using the most clear cut informational innovations for a dealer namely robot trades. For each dealer j in round r , we collected all quote revisions, $P(t)$, where P is the new midpoint of the dealer spread and t ranges from 1 (the initial quote posting) up to the total number quote revisions by trader j . From this, we calculate all log-price changes $\Delta P(t) = \ln P(t) - \ln P(t-1)$ ($t = 2, \dots$). Using the number of robot trades in the period $[t, t-1]$, we construct a dummy variable, $R(j, t, r)$, that equals one if a robot bought from trader j in period $[t, t-1]$, minus one if a robot sold to j , and zero otherwise. Furthermore, we define indicator variables - $I(r \in \text{QD})$ and $I(r \in \text{NQD})$ - as before to represent the transparency level in the market. Finally, we pool all rounds and dealers in a panel data set, thereby controlling for dealer-specific (e.g., skill) effects by allowing for an individual specific intercept, α_j (assuming that they sum to zero), and testing for the differences in price responses between the two market structures in a fashion similar to equation (39):

$$(40) \quad \Delta P(j, t, r) = \alpha + \alpha_j + \beta_1 I(r \in \text{QD}) + \beta_2 I(r \in \text{QD}) R(j, t, r) + \beta_3 I(r \in \text{NQD}) R(j, t, r) + \varepsilon(j, t, r)$$

The results appear in table 20. All parameters estimates are significant. The intercept α is negative, whereas the market structure dummy (β_1) is positive. The sum of both is a function of the level of the true prices relative to the initial price expectation of 100, since these parameters reflect the average price adjustment. Regarding the price response coefficients (β_2 and β_3), they are both positive and significant as expected, indicating that dealers do impound robot trade information appropriately in both microstructures, raising prices when robots buy, and lowering prices when robots sell. More importantly, the difference between the transparent and opaque markets, $\beta_2 - \beta_3$, is negative and significant, implying that dealers in the opaque market are more aggressive in responding to new information. This is entirely consistent with the result in table 19, again leading to a rejection of the null hypothesis.

20 Price responses

This table contains the estimates of the slope coefficients in the price response fixed individual effects panel model (40). Here, the individual price adjustments (log-price returns) following robot trades is regressed on a constant (α), on a market structure dummy (β_1) that equals 1 if the market has public quote disclosure and 0 elsewhere, a robot trade dummy, that equals +1 for a robot purchase and -1 for a robot sale, times an indicator variable for the transparent markets (β_2), or times an indicator for the opaque markets (β_3). β_2 and β_3 thus represent the average price increase (decrease) - measured as log price relatives - following a robot purchase (sale). The individuals in the panel are the market makers, so that we correct for cross-sectional differences between all market makers.

	Intercept (α)	Market Structure (β_1)	Public Quote Disclosure (β_2)	No Public Disclosure (β_3)
estimate	-0.015 (0.004)	0.032 (0.007)	0.063 (0.009)	0.119 (0.013)
difference in slopes ($\beta_2 - \beta_3$)				-0.056 (0.016)

The data are pooled over 7 market makers and 8 rounds; $n = 553$.

For details on estimating fixed individual effects panel models, see Baltagi (1995).

Robust standard errors for a general variance-covariance matrix are presented in parentheses (see Arellano (1987)).

$R^2 = 0.227$

8.4 Summary and concluding remarks

While the issue of post-trade transparency has received considerable attention in the literature, the degree of pre-trade transparency is also of central importance. Asset markets worldwide employ a bewildering variety of degrees of public quote disclosure, with implications that are not well understood. We consider the impact of quote disclosure in a continuous multiple-dealer market in an experimental setting. Seven professional market makers set quotes, and trade with other market makers and customers that represent both informed and liquidity motivated traders. Over multiple rounds, we vary the amount of quote information that is publicly available between two extremes: a market where all quotes are publicly disclosed and a market with no public quote disclosure. Our study represents an advance, in that we provide a direct comparison of alternative transparency arrangements, with a focus on multiple-dealer interactions. An experimental methodology is used to reduce the need for restrictive

behavioral and *ceteris paribus* assumptions. A continuous experimental market also generates large amounts of data, and permits us to examine more closely the relationship between the price and inventory components of a dealer's strategy.

We can summarize our results as follows. Pre-trade transparency significantly reduces search costs, thus alleviating some uncertainty and facilitating trade (search costs will be examined in chapter 10). As a result, market liquidity, measured by spreads and volume, is greater in the transparent market: opening spreads are smaller, and interdealer-trading volume is much higher. Dealers learn about the underlying price both directly, by trading with (potentially) informed robots, and indirectly, by observing changes in the prices quoted by other dealers. Dealers actively attempt to attract robot trades by quoting relatively tight spreads until a robot trade is achieved. Negligible search costs in the transparent market imply that speculating dealers see clearly how to shade quotes to avoid being picked off in a trade on the "wrong" side of the market, and how to keep their price competitive on the "right" side. The important and counterintuitive upshot is that rational speculating dealers use less aggressive price adjustments in the transparent market, thus slowing price discovery. Conversely, in the opaque market, search costs prevent accurate observation of competing prices, so that finely tuned quote shading is discouraged, and price adjustments are significantly more aggressive. The overall result is a distinct trade-off between liquidity and price efficiency.

Our explanation for this trade-off revolves around search costs. Pre-trade transparency reduces search costs, thus alleviating some uncertainty and facilitating trade. Market liquidity, measured by spreads and volume, is therefore greater. At the same time, however, these reduced search costs imply that dealers can (optimally) more narrowly fine-tune their price improvements. The upshot is that rational speculating dealers use less aggressive price adjustments, thus slowing price discovery. In light of these results, it seems plausible that some combination of both microstructures could offset the disadvantages of each individually. We note with interest that curious combinations of microstructures do exist in practice (direct versus brokered foreign exchange trading or upstairs versus floor trading of common stocks). Further research is warranted to gain insights into the performance of such combined market structures.

9 Post-Trade Transparency⁵⁴

Post-trade transparency is a very contentious issue among stock exchange regulators around the world. In Europe, the issue first came to attention in the wake of the U.K. 'Big Bang' in 1986, when a new electronic system of quote display (SEAQ: Stock Exchange Automated Quotation) was introduced, replacing the trading floor. Over the last decade there has been considerable public debate about the appropriate delay in the reporting and publication of trade prices and volumes on the domestic market for UK equities; a debate accompanied by periodic shifts in the transparency regime set by stock exchange authorities. Market makers argue in favor of delayed publication of the details of trades, because they feel that they cannot provide immediate liquidity for large transactions unless they are given time to unwind the resulting inventory positions before their competitors become aware of their predicament. Their position was challenged by the Director General of Fair Trading (1990), who argued that delays in publication are unnecessarily anti-competitive. Until very recently, an uneasy compromise (S.I.B. 1995a and 1995b) was in force, with immediate publication for smaller transactions, and publication delays of up to 60 minutes for all but the very largest transactions. This regime is currently being replaced by a tighter one.

Meanwhile, in the decade following Big Bang, SEAQ I (SEAQ International) became the focus of an important market in international equities in London. In particular, wholesale trading in European blue chip companies gravitated to London. Continental European authorities reacted to this migration of trading volume by lowering or abolishing turnover taxes and modernizing and reforming their domestic stock

⁵⁴ This chapter is based on: M. Flood, R. Huisman, K. Koedijk, and A. Röell, 1998, 'Post-Trade Transparency in Multiple Dealer Financial Markets'.

exchanges, thus repatriating wholesale trading volume with considerable success (see Pagano and Steil, 1996, for an overview of recent developments). Regarding post-trade transparency, some have argued that London's lax standards (before 1996 there was no trade publication on SEAQ I) might be the source of its success: market makers who can keep the news of their large trades confidential can outcompete those who have to publish details immediately, even though such delays in publication have a detrimental effect on the liquidity of the market at large because other market participants defensively widen their spreads for normal-sized trades because there are more asymmetries of information. In the early 1990s the imposition of tight post-trade transparency standards at the European Union level was debated intensively in the context of drafting the Investment Services Directive. But the imposition of stringent standards, advocated in particular by France and Italy, was abandoned in the face of strong opposition from the UK and Germany. The political agreement reached in 1992 sets very minimal standards for post-trade transparency.

In the USA, fast reporting of recent trades has been a policy objective since the adoption of the National Market System (NMS) in 1975; and the Securities and Exchange Commission (SEC) actively monitors exchanges' reporting and publication speeds, currently holding them to a standard of no less than 90% of trades reported within 90 seconds under normal conditions. The NYSE (New York Stock Exchange) has complied with extremely fast trade publication: a median delay in printing transactions on the consolidated tape system of well below 20 seconds, as reported by Hasbrouck et al. (1993). Regarding NASDAQ (the National Association of Securities Dealers Automated Quotation system), it is more difficult to effectively impose such standards because of the endemic practice of "protected trading", discussed by Franks and Schaefer (1995). With a protected trade, a market maker provisionally agrees to a price for a large deal (with a possible price improvement if market conditions warrant it), but the deal is not finalized until he has had time to 'work' the order without the market knowing about it. In effect, this type of trading circumvents publication rules and at the same time, gives reduced immediacy and hinders regulatory surveillance.

This points to a general problem with the enforcement of fast trade publication on decentralized markets such as NASDAQ or SEAQ. By their very nature, it is hard to impose strict post-trade transparency on dealer markets because large deals are conducted bilaterally, by telephone negotiation, and reported ex post to exchange authorities, who cannot really verify the exact time at which the deal is struck. Board and

Sutcliffe (1995) provide compelling evidence of the importance of protected trading in London: for deals of over roughly 10 times the median transaction size, they find that on average 26% of the transaction is pre-positioned on the same day, in advance of the transaction; and only 37% of trades have no pre-positioning at all (in the sense that less than 10% of the transaction is positioned beforehand on the same day). Markets such as the NYSE (with its very visible floor-based trading combined with electronic trading on a centralized system) and the Paris Bourse (where transactions on the electronic limit order book are again centralized, and the very act of transacting involves a report to the exchange authorities) are much more amenable to fast trade publication. Thus a debate about the pros and cons of post-trade transparency is, at the same time, a debate about the centralization of order flow processing.

The academic literature on the issue of market transparency has taken three main approaches: theoretical modeling, empirical work and experimentation. We review each in turn. On the theoretical front, Pagano and Röell (1992, 1996) point out that the more transparent the market, the more order flow information market makers have at their disposal, on which to condition their quote setting behavior. In their static model, transparency enables market makers to distinguish better between information-driven and liquidity-driven order flow. It is shown that this generally improves market liquidity for all trade sizes. A dynamic approach is taken by Kyle (1987), Röell (1988, 1996), Madhavan (1995) and Naik et al. (1994). These papers argue that market makers are willing to pay for order flow if they can keep information about their transactions confidential, because they can make trading profits in subsequent periods using the information. Thus, in a two-period setting, the first-period market for, in particular, large and informative orders is more liquid under an opaque publication regime. In the second period, the market as a whole is less liquid because market makers must defend themselves against the possibility that their competitors are better informed. Broadly speaking, these models predict that adoption of a more transparent regulatory regime harms informed traders, benefits small liquidity traders, and may have an ambiguous effect on large liquidity trades. Regarding competition between transparent and opaque markets (including off-exchange dealing), it is interesting to note that, irrespective of the relative comparative static merits of the two transparency regimes, dealers on opaque markets can always outbid for order flow because they have the advantage of keeping information confidential. Both Naik et al. (1994) and Lyons (1996) incorporate inventory control considerations into their models as well as the adverse selection problem stressed

by the other models. Lyons (1996) takes a positive approach, showing that from dealers' viewpoint there is an optimal degree of transparency in the market, which weighs its benefits in terms of enhancing investor confidence and participation against its cost in terms of prematurely conveying information, thus limiting risk sharing opportunities.

Empirical evidence on the impact of post-trade transparency is rather sparse and inconclusive. An early investigation (NASD, 1983) of the impact of subjecting NASDAQ securities to immediate trade reporting found a small decrease in average bid-ask spreads. This evidence is matched by a UK study (DG of Fair Trading, 1990) which finds that the introduction of a 24-hour delay in trade publication was accompanied by a small increase in spreads (for both small and large trade sizes) and a fall in the size for which market makers are willing to provide firm quotes. The most thorough study of the impact of trade reporting on liquidity is that of Gemmill (1996), who compares price movements associated with large transactions across different regulatory regimes for trade reporting. Gemmill finds mixed evidence concerning the hypothesis that block spreads are narrower relative to those for small trades when trade publication is delayed. Such price improvements for large trades seem primarily determined by market volatility, which varied substantially across the regimes; and Gemmill is unable to disentangle the effects of publication speed and market volatility. Gemmill also finds no evidence that delayed publication slows down price adjustment, or that returns volatility is affected. The problem with empirical work is that historically, other factors influencing market liquidity varied considerably across the periods in which different trade publication regimes were in force. This severely hampers empirical work attempting to identify the impact of transparency alone. Empirical work on cross-listed securities suffers from the same problem: different trading systems vary across so many dimensions that it would be rash indeed to attribute differences in liquidity to differences in transparency alone. Thus the large body of work that shows that very transparent auction markets give better terms for small- and medium-sized deals, but do not offer immediacy for large deals (see, for example, De Jong, Nijman and Röell, 1995) offers at best tangential evidence. These problems point to the advantages of experimental work. For in such work a controlled trading environment can be designed, in which all factors other than market transparency are held constant.

Related experimental work considers both pre- and post-trade transparency. In the previous chapter, we examined the concept of pre-trade transparency experimentally. The focus of Bloomfield and O'Hara (1996) is closer to what we deal with in this

chapter, in that they focus on the effect of trade publication. They find that post-trade transparency speeds up the price discovery process but increases opening bid-ask spreads, apparently reducing market makers' incentive to compete for order flow; and that informed traders are willing to pay a premium to conceal their trades by trading in a less transparent setting. This chapter also focuses on post-trade transparency; but it differs from Bloomfield and O'Hara's (1996) chapter in a number of important ways. Firstly, our experimental subjects are professional security traders; in preliminary runs using financial economics undergraduate students, the participants were noticeably less quick to grasp the setting in which they were playing, and much less driven by a wish to 'beat' the other members of their cohort. Secondly, our experiment takes place in continuous time; and the subjects are constantly forced to allocate scarce time between a number of different activities, such as modifying their quotes, initiating trades, and collecting information by finding out about the other participants' quotes and studying the trading records. Such a setting where traders need to make imperfect, rule-of-thumb decisions about where to focus their energies is designed to mimic the rough-and-tumble of real-world security markets. Furthermore, this continuous setup makes that the length of each experimental round in time is the single environmental variable that limits trading. We therefore obtain much more observations, about 500 transactions and 65 quote changes per run, than in for example sequential trade experiments. Thirdly, our market participants, realistically, not only post quotes but also initiate trades on their own account; and given that they are quite numerous (nine agents post quotes rather than two), the poor pre-trade transparency setting means that agents not only know little about other agents' quotes when deciding about their own quotes, they also do not know the best bid and ask prices at which they can initiate their trades, unless they waste valuable time finding out many other market makers' quotes. Thus post-trade transparency - the publication of recent trades - plays a dual role. Not only does it reveal how potentially better informed traders have recently been trading, it also tells traders about the best quotes that are currently on offer.

Some U.K. market participants stressed this role of post-trade transparency in their responses to a Securities and Investments Board discussion chapter regarding regulation of the U.K. equity markets: 'a significant number of institutional investors indicated that while they attached importance to liquidity, they saw transparency too as important. Several expressed concern that on occasion they might be dealing at the 'wrong' price. ... A range of 'foreign owned' market participants and some derivative market interests ...

suggested that a trade-off between transparency and liquidity should not be an issue. They argued that immediate publication of all trades would enhance investors' confidence that they were dealing at the 'right' price, and that it was investor confidence, rather than market maker commitment, that maximized market liquidity.' (S.I.B. 1995a). In our experiment, post-trade transparency reveals not only general market sentiment but also the current level of available price quotes.

9.1 Experimental design

In order to examine the effects of the post-trade transparency regime on market performance, we vary the market microstructure in terms of the amount of transaction information that is provided to the market makers over the trading rounds. As can be seen in table 21, we experiment in 6 rounds with an opaque market. In this market structure, the transaction history window of a trader only contains the transactions in which that particular trader is involved. In the transparent market structure, the transaction history window contains all trades in which the particular trader is involved (as in the opaque market) plus all other transactions in the market in which the external customers of the exchange are involved (the robot trades). The transaction information in the opaque market structure is therefore a subset of the transaction information that would be available in a transparent market. Trades are disclosed as soon as they occur. There is no delay in disclosure; only the amount of transaction information is varied.

9.1.1 Payoff to the subjects

At the end of each round, each trader is informed of the true price of the asset and his or her end-of-round profit. Esquire amounts are translated into Dutch guilders as follows. The dealer with the greatest end-of-round wealth receives 5 guilders; the dealer with the smallest end-of-round wealth receives 1 guilder; payments to other dealers are interpolated between the minimum and maximum result.⁵⁵ The traders are informed about this payment arrangement before the start of the experiment.

⁵⁵ At the time of the experiments, the guilder-dollar exchange rate was roughly 1 USD = 1.70 NLG. See section 8.1.1 for a discussion about the truncation of the payoff.

9.2 Data

The data are collected from experiments held at the Center for Research in Experimental Economics and Political Decision Making (CREED) at the University of Amsterdam on September, 16 1996. The subjects were nine professional traders from ABN Amro Bank, de Generale Bank, Optiver and Oudhoff Effecten. All subjects traded as market makers in 12 independent rounds; i.e., information about asset values obtained in one round is irrelevant in other rounds. To diminish learning effects we varied the amount of noise, the true price and the level of transparency randomly over all rounds.

21 Summary statistics

The experiment consisted of 12 rounds (in the following order) which lasted 300 seconds each. Nine professional traders participated as market maker in each round.

settings	1	2	3	4	5	6	7	8	9	10	11	12
market type ^a	O	T	O	T	O	T	O	T	O	T	O	T
α^b	0.5	0.5	0.5	0.5	0.3	0.3	0.3	0.3	1	1	1	1
true price	26	147	122	33	125	190	12	40	178	17	59	135
results												
# quotes set	54	82	70	56	71	63	72	53	70	78	95	62
# trades												
- total	287	264	406	504	419	603	444	636	625	648	526	537
- interdealer	203	180	322	418	334	518	360	551	541	569	455	490
- robot total	84	84	84	86	85	85	84	85	84	79	71	47
- uninformed	42	49	44	36	61	63	62	53	--	--	--	--
average end-of-round capital in esquires; initial capital is excluded.												
dealers	-289	-130	-180	-307	-110	-293	-264	-221	-519	-355	-109	-157
robot												
- informed	1315	588	648	1334	334	1048	987	938	2336	1598	491	708
- uninformed	-15	-5	163	50	162	273	202	57	--	--	--	--

^a O = Opaque market; T = Transparent market.

^b α equals the probability that a robot trade is informed.

Table 21 presents the settings and summary statistics. On average we obtained 69 quote settings, 412 transactions initiated by the market makers and 80 transactions initiated by the robots per round. Note that these values do not differ systematically in size across the two market types. Only in rounds 10, 11 and 12 is the number of robot trades smaller than in the other rounds. This is due to fact that the traders discovered the true price before the end of the round (note that in these rounds, there was no noise, i.e. α equals 1); the average dealer price converged and when the best price equals the true price, informed robot transactions do not take place. Table 21 also shows that, on average, dealers made losses, although not significantly, in each round. Here we should comment that the zero-sum nature of the experiment means that if some market makers make large gains, others generally make large losses. On average the market makers make slightly negative profits; for they are obliged to maintain a bid-ask spread of at most thirty esquire in the face of a mix of an informed trading robot (who always gains at their expense) and, in 8 out of 12 rounds, a noise trading robot (who loses on average provided the market spread is nonnegative), even when a wider spread than 30 esquires would be needed to ensure nonnegative expected profits on quoted prices. Observe again that market makers' average profits do not differ systematically across the transparent and opaque market settings.

9.3 Results

In this section, we compare the data obtained from the 6 opaque rounds with the results derived from 6 transparent rounds. In the opaque markets, traders only had transaction information about interdealer trades and robot trades in which they were personally involved. For the transparent markets, dealers also had information on all robot transactions. We formulate and test three hypotheses on the effects of market transparency. The first is in terms of the price efficiency, the second emphasizes the pre-trade profits and the last relates spreads and transparency.

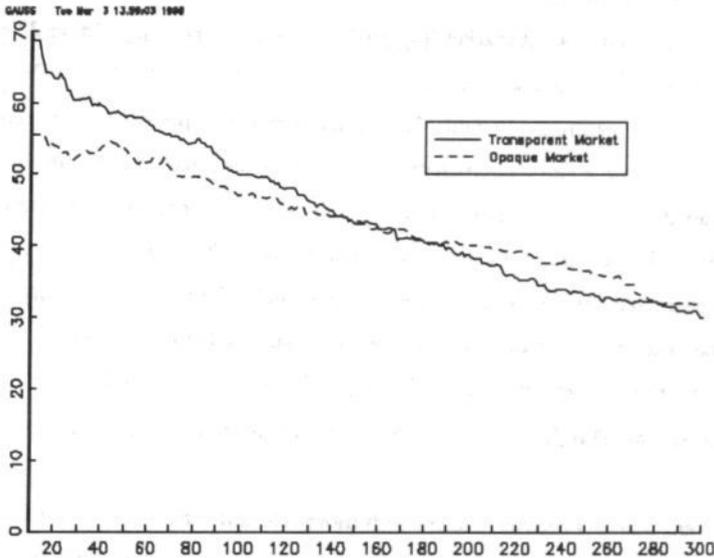
9.3.1 *Price efficiency*

In the absence of public price information, a high level of post-trade transparency provides the traders with (lagged) information, not only about the robot order flow but also about the prices quoted by other traders and their identities. Price information is therefore transmitted more quickly in transparent markets, and this is expected to improve the price discovery process. Bloomfield and O'Hara (1996) indeed observe that

price errors decrease more rapidly in transparent markets. We therefore test the following hypothesis:

$H1_0$: *Post-trade transparency improves the informational efficiency of prices.*

In order to examine the informational efficiency of the quotes of the market makers, we concentrate on the price changes of market makers to robot trades that occur anywhere in the market. Since these robot trades are disclosed to all traders in the transparent market, traders can observe them and incorporate them into their price quotes more quickly. Then, price errors, defined as the absolute difference between the true price of the asset and the midpoint of the market spread, decline faster over time in transparent markets. In the figure 13, we show the time path of the price errors for each market as an average obtained over all rounds in which a particular market type was installed.



13 Average price errors over time

Figure 13 reveals the difference in price efficiency between the two market structures. To examine this more formally, let $P(j,t,r)$ be the price error obtained from the quote market maker j in round r at time t (measured in seconds). We test whether the price errors decline more rapidly in a transparent market by regressing the price errors on a time trend variable t , whereby the effect of disclosing transaction information is split over two slope coefficients.

$$(41) \quad P(j, t, r) = \alpha + \alpha_j + \beta_1 I(r \in Tr) + \beta_2 t I(r \in Tr) + \beta_3 t I(r \in Op) + \varepsilon(j, t, r),$$

where $I(\cdot)$ is an indicator function that equals 1 if the condition in parentheses is true and equal to 0 elsewhere, t is the time trend, Tr is the set of all rounds in which the market is transparent, Op is the set of all rounds in which the market is opaque, and ε is assumed to be an iid error term. We pooled the data over all twelve rounds and the nine market makers. We added specific constants α_j in (41) to control for the specific characteristics of market maker j , whereby we assume that they sum to zero. The effect of disclosing transaction information or transparency is captured both by the initial level effect - measured by the market structure dummy β_1 - and by the difference between the slope coefficient for the transparent market (β_2) and the opaque market (β_3). Table 22 contains the estimates of these slope coefficients.

Evaluating the estimates in table 22, we learn that the price errors decline more rapidly over time as the slope coefficient for the transparent market is significantly more negative than the slope coefficient for the opaque market. Note that the average price errors differ in the beginning. In the transparent markets, the average initial price error equals approximately 70, whereas it equals approximately 55 in the opaque markets. This is due to the differences in true prices selected for both market types as the average price error reflects the average difference between the true prices and the initial price expectation of 100. The β_1 estimate corrects for this difference as can be seen from the table..

Transparent markets provide traders with more relevant information contained in the robot trades, and we find that this is reflected in their quotes since the price errors decline more rapidly over time. Therefore, post-trade transparency does indeed improve the informational efficiency of price quotes, as hypothesized. At first glance, this result may seem contradictory to what we found in the previous chapter. There we observed that a market that is transparent in terms of quote disclosure is less price efficient than an opaque market. The main difference lies in the focus of either pre-trade or post-trade

transparency. Although both pre- and post-trade transparency are frequently mentioned together, their effects on market performance is the opposite.

22 Price efficiency

This table contains the estimates of the slope coefficients in the fixed individual effects model (41). Here, the time path of price errors of individual market makers is regressed on a constant, on a market structure dummy (β_1) that equals 1 if the market has public quote disclosure and 0 elsewhere and on two slope coefficients (β_2 and β_3) for the time trend in the rounds with and without public disclosure of quotes respectively. The individuals in the panel are the market makers, so that we correct for cross-sectional differences between all market makers.

	Intercept (α)	Market Structure (β_1)	Public Quote Disclosure (β_2)	No Public Disclosure (β_3)
estimated slope coefficient	57.86 (3.05)	8.14 (5.17)	-0.13 (0.01)	-0.09 (0.02)
difference in slopes ($\beta_2 - \beta_3$)				-0.044 (0.021)

The data are pooled over 9 market makers and 12 rounds; $n = 32,400$.

For details on estimating fixed individual effects panel models, see Baltagi (1995).

Robust standard errors for a general variance-covariance matrix are presented in parentheses (see Arellano (1987)).

$R^2=0.117$

Regarding the volatility of price changes, we have split each round in 20 intervals of 15 seconds each and analyzed the differences in the logs of the average over the midpoints of all outstanding quotes over two succeeding intervals. In both regimes, the standard deviation of the log price changes, being our measure of market volatility, equals 0.03 on average. Thus we find no effects of transparency on market volatility, which is to some degree in line with both Gemmill (1996) and Bloomfield and O'Hara

(1996) who find mixed results concerning the effects of transparency on market volatility.⁵⁶

9.3.2 Profits

Since in opaque markets only private transaction information is available to the traders, informed traders are better able to hide their information. In transparent markets however, the trades of informed traders are publicly observable. Transparent markets therefore enable other market participants to learn faster about the true price of the asset. An increase in the level of post-trade transparency should increase the profits or decrease the losses of the uninformed market makers at the expense of the informed traders, who have less opportunities to hide their trades. In Röell's (1996) model the gains to uninformed traders from an increase in the level of transparency are exactly offset by a loss in the profits of informed traders. In Bloomfield and O'Hara's (1996) experiment both liquidity investors and informed traders lose from transparency while market makers gain. Given these results, we shall focus on testing the following hypothesis in this section:

H2₀: Uninformed traders gain from transparency at the expense of the informed traders.

The summary statistics in Table 21 present the end-of-round average profits of the market makers and the robots. The market makers make losses on average but because there is enormous variation among individual dealers these losses are not significant as reflected by their large standard errors. The informed robots make significant profits. The uninformed robots generally make a small but insignificant profit, as explained by the pattern of spreads to be discussed in the next section.

To examine the effects of transparency on the profits in more detail, we compare the profit made on each transaction by the parties involved across opaque and transparent markets. Per trade we identify two parties: the party who initiates the trade and the one who has provided the quote that is being hit (to be referred to as the 'quoter'). The initiators are either market makers who hit the quote of another trader (we refer to them

⁵⁶ We also examined autocorrelations of the price changes in a similar way, but did not find any effects of transparency on the autocorrelation of price changes. However, the autocorrelation function does differ of both transparency regimes if we analyze changes in transaction prices. In both regimes, the first order autocorrelation is negative and significant, but it is less negative in the transparent rounds. We attribute this to the fact that in transparent markets, good prices are easier to find (which will be

as 'hitters'), informed customers or uninformed customers ('informed' and 'uninformed' respectively).

Define $\pi(r, \tau, j)$ as the profit for trader j involved in the τ^{th} transaction in round r . For a buyer j this is equal to the difference between the true price $P^*(r)$ in round r , and the transaction price, $P(r, \tau)$:

$$(42) \quad \pi(r, \tau, j) = P^*(r) - P(r, \tau).$$

Similarly, for seller j :

$$(43) \quad \pi(r, \tau, j) = P(r, \tau) - P^*(r).$$

For each transaction, we calculate the profits given in (42) and (43), for both parties involved. For each of the four trader types, we collect the profits over all transactions in which the particular trader type was involved. For each trader type (hitter, quoter and the robots), we regress the profit obtained from transaction τ in round r on the following variables. First, a dummy variable $T(r)$ that takes the value one if the market in round r is transparent and zero if the market is opaque. Second, the variable $\alpha(r)$ which inversely reflects the amount of noise used in round r . Third, a variable $D(r)$ equal to the absolute value of the difference between the true price and the 'best initial estimate' of 100 esquires⁵⁷. If $D(r)$ is large, the true value lies 'far away' which makes it easier to obtain large profits. Finally, although we use professional traders as subjects, their trading skills can differ substantially, *inter alia* because they trade in different securities markets in their professional lives. To correct for these skill differences, we introduce a trader specific constant term λ_j (summing to zero) in addition to the conventional constant β_0 ⁵⁸. Thus, we pool the profits over all rounds for each type of trader and perform the following regression, in which $\varepsilon(r, \tau, j)$ represents an iid error term:

$$(44) \quad \pi(r, \tau, j) = \beta_0 + \lambda_j + \beta_1 T(r) + \beta_2 \alpha(r) + \beta_3 D(r) + \varepsilon(r, \tau, j),$$

analyzed in the profits section) such that the amount of price dispersion (which can be substantial in our poor pre-trade setting) is of less importance in transparent markets.

⁵⁷ The traders are informed that the true price is uniformly distributed between 1 and 200. Their best initial estimate should therefore be 100 esquires.

⁵⁸ In effect, model (44) is a fixed individual effects panel model, in which the trader dummy is the fixed effect. This model can be estimated by first subtracting the average over time from the variables on the left hand side and the right hand side of (44); the result is a regression equation in which the individual constants drop out. Then OLS is applied to obtain estimates for the slope parameters in (44); see Baltagi (1995). A regression without these individual effects did not yield substantially different results.

where $j = 1 \dots 9$ for the regression with the hitters and quoters, while the λ_j terms are omitted in the regressions for the informed and uninformed. Estimates of the parameters of regression (44) are presented in table 23 for the hitters, quoters, informed customers, and uninformed customers.⁵⁹

The estimates for β_1 in table 23 indicate that the hitters gain significantly from an increase in the level of transparency. For quoters this effect is negative but not significant; the same result holds for the uninformed robots. Note that the uninformed robot trades can be regarded as transactions from liquidity or noise traders. The gain to the hitters clearly comes from the informed robots, who make significantly less profit per trade in the more transparent setting.

23 Profits

Panel estimates of the parameters in the pooled regression equation (44) for the following profits definitions: (a) the per-transaction profit of the market makers that hit the quote of another market maker (hitters); (b) of traders that are being hit by another market maker or a robot (quoters); (c) of the informed robot trades; and (d) of the uninformed robot trades. Each of these variables is regressed on a constant term (β_0), a transparency dummy (β_1), the level of noise (measured as the probability that a robot trade is informed) (β_2) and a variable that equals the absolute value of the difference between 100 esquires and the true price in the round (β_3). The column avg. presents the average per trade profits for each definition over all rounds.

Profits definition	avg.	β_0	β_1	β_2	β_3
Traders					
- Hitters ^a	22.093 (48.652)	-7.583 (8.514)	9.579* (4.842)	-8.146 (8.777)	0.482* (0.095)
- Quoters ^b	-22.983 (48.685)	7.177 (9.216)	-7.177 (8.907)	4.972 (9.508)	-0.480* (0.080)
Robots					
- Informed ^c	44.891 (26.506)	29.886* (1.273)	-3.633* (0.083)	-37.998* (1.079)	0.730* (0.015)
- Uninformed ^d	4.330 (60.462)	11.384 (14.395)	-1.066 (1.808)	-17.723 (36.826)	0.004 (0.045)

The panel estimates are obtained from a within regression as described in Baltagi (1995). Robust standard errors for a general variance-covariance matrix are presented in parentheses (see Arellano (1987)).

^a n = 4941

^b n = 5899

^c n = 549

^d n = 409

* significant at the 5% level or smaller

⁵⁹ For reasons of space, we do not present the estimates of the individual constants.

The estimates for the noise parameter β_2 are only significant for the informed. The negative sign of the estimate indicates that an increase in the level of noise (a decrease in $\alpha(\tau)$) leads to higher profits for the informed customers, as they gain from the uncertainty in the market.

Lastly, we observe from the estimates of β_3 that a true price that is further away from the initial estimate of 100 equires increases the per trade profits of the hitters and informed robots significantly at the expense of the quoters. For uninformed robots, the effect is not significant since they buy and sell randomly with probability $1/2$, so that they gain and lose roughly equally from a large deviation of the true price.

The main result from these tests is that (uninformed) dealers who initiate trades gain from transparency at the expense of the informed customers, supporting the hypothesis $H2_0$. However, this result does not hold for all dealers; the quoters, i.e. traders whose quotes are hit, do not gain from an increase of transparency. A possible explanation could be that the hitters use the transaction information to find the best quotes outstanding in the market, since subjects in the transparent microstructure are provided with information about market-wide flows of transaction prices and trader identities. Transparent markets give them opportunities to find good prices faster than in opaque markets, increasing the profits made per transaction. Note that only the hitters are able to gain from this information directly.

9.3.3 Spreads faced by the different agents

The above test derived from (44) does not provide insight into whether hitters are able to obtain better transaction prices in transparent markets. To test this formally, we should relate transaction prices to other quotes outstanding at the time the trade is cleared, to see whether the hitters obtain the best price. We therefore calculate the effective spreads for each transaction. As defined in de Jong, Nijman and Röell (1995), the effective spread measures the difference between the transaction price and a proxy for the ruling market price at the time the transaction clears. In transparent markets traders can obtain (lagged) price information from the displayed transaction history, which offers them an opportunity to trade against better prices than in opaque markets, i.e. face lower effective spreads. Therefore, we expect *a priori* that transparency leads to lower effective spreads.

Following de Jong, et al., we define the effective spread facing the initiator j of the τ^{th} transaction in the r^{th} round transaction as $S_{\text{E}}(r, \tau, j)$, when trader j is the buyer, as:

$$(45) \quad S_E(r, \tau, j) = P(r, \tau) - M(r, \tau).$$

If trader j is the seller, the effective price he or she faces is:

$$(46) \quad S_E(r, \tau, j) = M(r, \tau) - P(r, \tau).$$

In both formulas, $P(r, \tau)$ denotes the price at which transaction τ in round r is settled and $M(r, \tau)$ is the average midpoint over all quotes outstanding at the moment transaction τ is initiated. The variable $M(r, \tau)$ can be seen as a proxy for the market price of the asset⁶⁰. Note the similarity between the profit definitions (42) and (43) and the effective spreads (45) and (46). In the latter, the true prices are replaced by proxies for market prices. The effective spread therefore measures the instantaneous loss made by the initiator of transaction τ , relative to the market's current estimate of the security's value.

The effective spread facing the hitters, informed robots and uninformed robots is calculated using (45) and (46). The average effective spread obtained by the quoters is calculated obversely. Both types of robots are programmed to hit the best price available in the market when they initiate a trade, so that robots always pay the minimum possible effective spread. This is not the case for the hitters, who may not know the best price available. We regress the effective spreads on a transparency dummy $I(r \in Tr)$ as define in (44), on the level of noise in the round r , represented by $\alpha(r)$, and on the variable $D(r)$ which equals the absolute value of the difference between the true price and the best initial guess of 100 esquires. Again we added an individual constant term for each trader to correct for differences in skills. Given the above notation and definitions, the regression equation becomes:

$$(47) \quad S_E(r, \tau, j) = \beta_0 + \lambda_j + \beta_1 I(r \in Tr) + \beta_2 \alpha(r) + \beta_3 D(r) + \varepsilon(r, \tau, j),$$

in which $j = 1 \dots 9$ in both regressions for the hitters and quoters and $\varepsilon(\tau, r, j)$ is assumed to be iid. Table 24 presents the parameter estimates of the above regression model.

From table 24, we observe that increasing the level of transparency in the market significantly decreases the average effective spread for the hitters. For them, it is easier to obtain better price quotes due to the extra amount of (lagged) transaction information in transparent markets. They gain at the expense of the quoters, who obtain a significantly smaller effective spread.

⁶⁰ Recall that, due to the poor pre-trade transparency in our setup, the market spread is not readily available in the market, since no price information is presented on the trading screens of the traders.

A decreasing amount of noise in the markets (an increase in $\alpha(r)$) increases the average level of the effective spread for the hitters, informed and uninformed customers significantly. More certainty hastens the price discovery process and the consensus of the traders in their pricing strategy. Note that a decrease in the level of noise does not affect the quoters. Lastly, we observe that the further away the true price is from 100 esquires, the smaller the effective spreads that are obtained by the quoters and faced by the robots. This variable has no significant effect on the hitters.

24 Effective spread

Panel estimates of the parameters in the pooled regression equation (47) for the following definitions of the effective spread (transaction price - average midpoint over all outstanding quotes): (a) the per-transaction profit of the market makers that hit the quote of another market maker; (b) of traders that are being hit by another market maker or a robot; (c) of the informed robot trades; and (d) of the uninformed robot trades. Each of these variables is regressed on a constant term (β_0), a transparency dummy (β_1), a noise variable (β_2) and a variable that corrects for differences in the true prices (β_3). The column avg. presents the average per trade effective spread faced or obtained over all rounds.

Spreads faced by	avg.	β_0	β_1	β_2	β_3
- Hitters ^a	-3.111 (12.484)	-4.181* (1.607)	-3.147* (0.752)	6.230* (1.642)	-0.019 (0.018)
- Informed ^c	-3.937 (9.058)	-6.682* (0.476)	-1.519* (0.072)	0.116* (0.055)	0.056* (0.008)
- Uninformed ^d	-3.963 (8.484)	-4.446* (0.254)	-2.387* (0.014)	2.261* (0.480)	0.013* (0.002)
Spreads obtained by					
- Quoters ^b	3.247 (11.968)	-4.142 (4.062)	-2.489* (0.872)	4.317 (4.026)	-0.008* (0.051)

The panel estimates are obtained from a within regression as described in Baltagi (1995). Robust standard errors for a general variance-covariance matrix are presented in parentheses (see Arellano (1987)).

^a n = 4941

^b n = 5899

^c n = 549

^d n = 409^e significant at the 5% level or smaller

Since in both market types information about the true price arrives over time, the traders become better informed, which should be reflected in the profits they make on average per transaction. The more they know about the true price, the larger the profits per transaction. We test the time path of profits as follows. Each round is split into 10 intervals of 30 seconds each. For each interval we calculate the average profits per

transaction of the hitters, quoters, informed and uninformed. These average profits are then regressed on a time trend t corresponding with the interval. More formally, define $\pi(r,t)$ as the average profits made in interval t where $t = 1..10$. Then the regression equation becomes:

$$(48) \quad \pi(r, t) = \beta_0 + \lambda_r + \beta_1 t + \varepsilon(r, t),$$

where λ_r is an individual constant for round r , to correct for differences in the true price and noise and $\varepsilon(r,t)$ is an error term. We perform the regression for the four trader types and both market types. Table 25 presents the estimates for the slope parameter β_1 .

25 Profits over time

Estimates of the slope parameter in the trend regression (48). The average profits over 10 intervals are regressed on a trend variable that denotes the interval number.

Profits definition	Opaque	Transparent
Traders		
- Hitters	0.296 (1.536)	-3.802 ^a (1.909)
- Quoters	0.213 (1.121)	2.887 ^b (1.491)
Robots		
- Informed	-2.616 ^a (0.840)	-3.807 ^a (1.449)
- Uninformed ^c	0.788 (0.994)	-0.478 (1.012)

Number of observations equals 6 rounds. Robust standard errors for a general variance-covariance matrix are presented in parentheses (see Arellano (1987)).

^a Significantly different from zero at the 5% confidence level.

^b Significantly different from zero at the 10% confidence level.

^c Number of observations equals 4 rounds.

Table 25 shows that in the opaque market the informed customers make less profits over time, due to the price discovery process. The estimates for the hitters, quoters and uninformed are positive but not significant, indicating that their profits do not significantly increase over time. For the transparent market we observe a different result. Here, the hitters' per trade profit decreases over time, the same holds for the informed, whereas it increases for the quoters. More information becomes available over time and

is incorporated into price quotes, thus reducing the losses of the quoters. In the opaque market, the profits do not change over time for the quoters and the hitters, reflecting the difference in the amount of information that becomes available over time to traders compared with transparent markets.

Concluding, we accept $H2_0$, in the sense that uninformed market makers gain from transparency in the transactions they initiate. For quoting market makers who are hit by other participants, transparency has no significant effect. Combining the hitters and quoters, transparency improves the profit position of the uninformed market makers. This gain is offset by losses to the informed robots.

9.3.4 Spreads

Opening spreads in transparent markets are wider than in opaque markets. The difference disappears over time. During the trading day the spreads in transparent markets narrow, so that differences in spread between opaque and transparent markets disappear over time. This is exactly what Bloomfield and O'Hara (1996) find in their experiments. If, as argued by Madhavan (1995) and Röell (1988, 1996), trading information is not made publicly available, market makers compete for order flow to obtain information by offering small spreads. Since more transaction information is publicly available in the transparent market, the market makers have no incentive to offer small spreads in the beginning. They might even want to wait and see how the market evolves throughout the opening, and therefore set wide spreads to minimize the probability of being hit in early trading. During the trading period, the willingness to trade increases and, when trades have occurred, the information differential between the two market types decreases, narrowing the spreads in the transparent market. In financial markets, the determinants of spreads typically include inventory costs, the costs of market making, and an adverse selection component. Note that our setup abstracts from inventory costs and costs of market making. As in Glosten and Milgrom (1985), the main source of non-zero spreads in our experiment is the adverse selection problem; i.e. quote setters' protection against better informed counterparties.

To examine the difference in spreads between the opaque and transparent markets, we split each round into 20 intervals of 15 seconds each.⁶¹ For each interval t in round r ,

⁶¹ The results are robust to the choice of the number of intervals; 10 intervals or 30 intervals do not change the qualitative results.

we calculated the average spread across all outstanding quotes, denoted as $s(r,t)$.⁶² With respect to the measurement of the spreads we note that Bloomfield and O'Hara (1996) perform their analysis in terms of the market spread (i.e. lowest ask minus highest bid). Due to the absence of centralized price information in our experiments, the market spread is not observable for our traders. For instance, a negative market spread induces an arbitrage opportunity if this is readily observable and exploitable in the market. This is only the case in markets with a high level of pre trade transparency (e.g., the setting in Bloomfield and O'Hara (1996)). To overcome this problem we analyze the average dealer spread instead of the market spread. We regress the average spread on a trend variable equal to the interval index t ($t=2..20$). To obtain a direct comparison between the opaque market and the transparent market, we restrict the slope parameters β_1 to be equal in all rounds that are from the same market type. For each round r from the same market type, the regression equation becomes:

$$(49) \quad s(r, t) = \beta_0 + \lambda_r + \beta_1 t + \varepsilon(r, t),$$

where $\varepsilon(r,t)$ denotes an IID error term. Observe that the slope parameter is equal for each round r within the same market type, but that each round is allowed to have a individual constant term λ_r . The estimated slope coefficients of the model (49) are presented in table 26.

From table 26 we observe the downward trend of the spread over time in transparent markets in contrast with no significant trend in the opaque market. The spreads in the beginning of the round are wider in the transparent market, indicating the absence of an incentive to trade, as discussed in Madhavan (1995). It should be noted that, because trade initiators in a transparent market can more easily find the best quote, the wider initial spread in the transparent market does not necessarily imply higher trading costs. We also present, in figure 14, the averages of $s(r,t)$ over each round in which we used the same market type.

⁶² We discarded the values of the first interval, because the traders set their initial quotes in this time period. We therefore have no quotes for part of this interval.

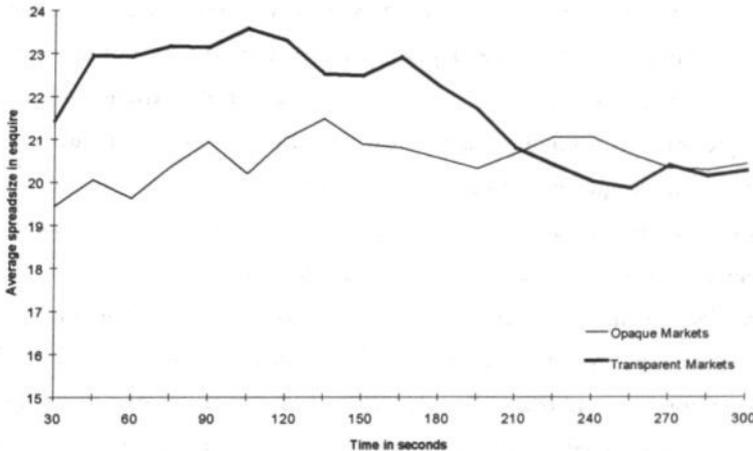
26 Spreads

Slope estimates of the regression model (49). The average spread size across time intervals is regressed on a trend variable, indicating the interval index. The slope coefficients are restricted to be equal for all rounds of the same market type. Each round is allowed to have a unique constant term.

	Opaque market	Transparent market
Estimated slope coefficient	0.033 (0.072)	-0.189 ^a (0.085)

The panel estimates are obtained from a within regression, as described in Baltagi (1995). Number of observations equals 114 intervals. Robust standard errors for a general variance-covariance matrix are presented in parentheses (see Arellano (1987)).

^a Significantly different from zero at the 5% confidence level.



14 Spreads over time

The lines in the figure represent the average spread over all quotes outstanding averaged over all rounds with the same transparency regime at intervals of 15 seconds.

Again, we see that the difference in spread between opaque markets and transparent markets is substantial in the first part of the round and that the difference diminishes towards the end of the round, which is in line with the results obtained by Bloomfield and O'Hara (1996).⁶³

9.4 Conclusions

In this chapter, we examine the effects of different levels of post-trade transparency on an experimental financial market. The question of transparency is important, since there is an unresolved policy debate about its desirability. In the U.S., the SEC has consistently (since 1975) held that rapid reporting yields in effect a public good requiring exchanges to provide rapid post-trade disclosure. The key benefits are seen to come through price efficiency, competition, and fairness. In the U.K. on the other hand, the S.I.B. (1995b) has recognized that there may be a trade-off between transparency and liquidity, acknowledging "the need to strike a balance between the desirability of promoting transparency and the risk of reducing liquidity".

In our continuous experimental financial market trading a single fictional security, nine subjects act as market makers, who both set quotes and trade continuously. In addition to the market makers, computerized external customers trade on the exchange, representing both informed and liquidity (i.e. noise) traders. The level of transparency is defined by the amount of transaction information shown on the private trading screens of the market makers. Over the experimental runs the post-trade transparency regime varies between an opaque market structure, where only those transactions appear on the dealer's screen where that particular dealer was involved, and a transparent market, where also all trades of the external customers are publicly disclosed.

The results presented here are obtained from twelve experimental rounds where professional securities traders of Dutch banks participated as market makers. Having over 5800 transactions and 800 quote settings, we find that an increase in transparency enhances the price efficiency of the market and improves the profit positions of uninformed traders since crucial information derived from transactions with informed external customers is transmitted faster to all market participants. The uninformed traders gain significantly from transparency, at the expense of the informed traders, because uninformed traders have additional opportunities to discover relevant price

⁶³ The difference is significant at the 10% confidence level in the first part of the trading period.

information, facing a significantly smaller effective spread per transaction. On the other hand we find that opening spreads are significantly wider in transparent markets, since traders have less incentive to compete for order flow to obtain information, resulting in initially better prices for customers on opaque markets.

Concluding, apart from wider opening spreads, greater transparency clearly improves the performance of financial markets in terms of price efficiency and the profit positions of uninformed traders. Worldwide these benefits are seen by many exchanges and explains the trend towards higher transparency standards. In this respect the introduction of the new order-driven trading service SETS on the London Stock Exchange is a very recent example. This will probably temper the transparency discussion on the LSE but will raise the issue of protected-trading, i.e. reporting a large trade at the moment they are worked out instead of reporting at the moment it enters the market, which is essentially closely related to the issue of post-trade transparency since it will give traders opportunities to hide trade information from the market.

The bid-ask spread and its components are of considerable practical significance. In Europe, for example, the push to lower spreads is driving major redesign of exchanges in preparation for monetary union. In the U.S., recent controversy over odd-eighth quotes on NASDAQ is also driving reforms. How reforms lower the spread, however, depends on how reforms affect particular spread components. These links are not yet well understood.

This chapter offers evidence of an overlooked component of the spread — search costs. Earlier work on spreads in competitive markets emphasizes three other components: order-processing costs, inventory holding costs, and adverse selection costs (arising from asymmetric information).⁶⁵ Search costs are overlooked in this earlier work because markets are assumed to be consolidated — all participants trade at the best bid and offer. But this is not the case in many key financial markets, for example, foreign exchange and multiple-dealer equity markets. In this chapter, search costs arise in our experimental market because obtaining quotes requires dealers to call one-another; it is not possible to observe the best bid and ask in the market directly. The resulting shopping for quotes is time-consuming and adds uncertainty to dealers' price information (since quotes lapse when revised).⁶⁶ This induces trades at prices that frequently depart from the best bid/ask, implying a wider effective spread. We show that these costs contribute substantially to the spread, accounting in many cases for more than 40 percent.

⁶⁴ This chapter is based on: M. Flood, R. Huisman, K. Koedijk, and R. Lyons, 1998, 'Search Costs: The Neglected Spread Component'.

⁶⁵ The literature on spread components is extensive. See Campbell, Lo, and MacKinlay (1997, ch. 3) and Huang and Stoll (1997) for recent surveys. Huang and Stoll (1997) is a recent addition to a long line of papers that disentangle these three components from market data.

We also evaluate leading models of how the spread affects transaction prices. On the whole the theoretical implications of “bid-ask bounce” are borne out with remarkable accuracy. For example, the first model we consider is that of Roll (1984). Roll’s insight is that liquidity traders who trade randomly against the bid and ask should produce negative auto-correlation in transaction price changes. Specifically, assuming equally likely buy and sell orders and otherwise frictionless markets, his model predicts an auto-correlation in transaction price changes of $-1/2$. This is precisely what we find from the trades between our subjects.

10.1 Effective Spreads: Four Leading Models

Consider first Roll’s (1984) model of effective spreads.⁶⁷ Roll shows that the first-order return auto-correlation in the data can be used to measure the effective spread faced by investors. Specifically, if S is the constant effective spread and P_t is the price of transaction at time t , then Roll’s model implies the following estimator for the effective spread, S_{Roll} :

$$(50) \quad S_{\text{Roll}} = 2\sqrt{-\text{cov}(\Delta P_t, \Delta P_{t-1})}$$

A crucial assumption in Roll’s model is that the conditional probability of a trade flow reversal equals $1/2$; i.e., $P\{t \text{ is a sell} \mid t-1 \text{ was a buy}\} = P\{t \text{ is a buy} \mid t-1 \text{ was a sell}\} = 0.5$. That is, there is no serial dependence in transaction type. Choi, Salandro and Shastri (1988), henceforth CSS, present an extension of Roll’s model which loosens this restriction. In particular, the conditional probability of a trade reversal, π , can depart from $1/2$ due to the serial dependence.⁶⁸ In the CSS model the effective spread becomes:

$$(51) \quad S_{\text{CSS}} = \frac{1}{\pi} \sqrt{-\text{cov}(\Delta P_t, \Delta P_{t-1})}$$

Both the Roll model (50) and the CSS model (51) assume that the fundamental value of the asset being traded is constant over time, i.e. price volatility and serial correlation in

⁶⁶ That dealers shop for quotes is shown by Lyons (1998) for the foreign exchange market. In his sample more than 70 percent of bilateral quote requests do not generate a trade.

⁶⁷ As in the rest of the literature, our focus here is on transaction prices and effective spreads. Effective spreads can differ from quoted spreads for many reasons. The reason emphasized in the current literature is that dealers negotiate prices within their quoted spreads (see, for example, Fialkowski and Petersen (1994)). The reason emphasized in this chapter — search costs — is quite different.

⁶⁸ In their original formulation, Choi, Salandro and Shastri (1988), state the model in terms of the conditional probability, δ , of no change in trade direction, so that $\pi = (1-\delta)$.

transaction price changes are due to bid-ask bounce. In financial markets, the equilibrium value of a security can change for a number of reasons, including especially the arrival of news. In our experimental market, the fundamental value is fixed over time – i.e., there is no news in the traditional sense. However, dealers learn over time about fundamental value by filtering information from computer-generated customer trades, as described below. In other words, as price discovery occurs, the fundamental value as perceived by the dealers changes over time. To capture the effect of a changing fundamental value (or, as here, a perceived fundamental value) de Jong, Nijman and Röell (1995) (JNR) propose the following estimator for the effective spread, S_{JNR} :

$$(52) \quad S_{JNR} = \sum_{t=1}^T |P_t - M_t|,$$

where M_t is the mid-spread price at the time of the t^{th} transaction, defined as the average of the best bid and the best ask outstanding. This “mid-price” is a proxy for the fundamental value as perceived by the market.

Finally, and most recently, Huang and Stoll (1997) develop a general model of the spread, in which they nest three other models, including the Roll and CSS equations introduced above. They begin their generalization by defining a trade-direction indicator, Q_t , which equals +1 when the ask price is hit (by a buyer), and -1 when the bid price is hit (by a seller).⁶⁹ We use their regression equation (7):

$$(53) \quad \Delta P_t = \frac{S}{2} \Delta Q_t + e_t,$$

and obtain spread estimates using ordinary least squares (for details see Huang and Stoll (1997), p. 1001). Huang and Stoll argue theoretically that $\text{cov}(\Delta Q_t, \Delta Q_{t-1})$ should equal -1 when the conditional probability of a trade flow reversal, π , equals 1/2, implying that our equation (53) should be identical to Roll’s estimator in equation (50) above. However, we face a different institutional arrangement and a modified definition of Q_t from that presumed by Huang and Stoll, allowing our estimators (53) and (50) to diverge from one another.

Below we use each of the four alternatives in equations (50) - (53) to estimate the effective spread for trading in an experimental securities market.

10.2 Experimental Design

The design is similar to that used in the previous chapters, though our parameter settings are different. Here we outline the basic structure of the experiments.

We conducted the experiments on Nov. 17, 1997 in the computer laboratory at the University of Amsterdam's Center for Research in Experimental Economics and Political Decision Making (CREED). The seven subjects are professional options traders from Optiver, a Dutch options trading house. Each subject had his/her own trading screen and keyboard. The subjects trade as dealers in 16 four-minute rounds, with a single security in each round. The rounds are independent in the sense that information about the security's value in one round is unrelated to values in other rounds.

10.2.1 Information Structure

At the end of each trading round each dealer's position is converted into a fictitious laboratory currency, esquires, at the security's terminal value. The terminal value has both a learnable and unlearnable component. The learnable component is distributed uniformly between 0 and 200 (esquires per share). The unlearnable component is distributed uniformly over three possible values, $\{-50, 0, \text{ and } +50\}$. Dealers are informed *ex-ante* about the distributions of the two components, but are not told the realized terminal values until their inventories are converted at the end of trading. The learnable component is learnable in the sense that it is a key factor in the trading behavior of the informed customers (described below), so that dealers can filter this information from observing customer trades. For the unlearnable component, in contrast, there is no conditioning information revealed that is useful for predicting it. Rather, the unlearnable component is an end-of-round price shock that ensures some residual risk, even if the price discovery process has succeeded in identifying the learnable component.

⁶⁹ More specifically, Huang and Stoll (1997), p. 999, define Q_i relative to the spread midpoint, M_i , so that $Q_i = +1$ (-1, respectively) if a trade is buyer-initiated (seller-initiated) and occurs above (below) M_i ; $Q_i = 0$ if the trade occurs at M_i . However, our experimental market involves bilateral quotations, implying informational imperfections that allow negative market spreads to arise. Defining Q_i as do Huang and Stoll - i.e., relative to M_i - would thus leave the direction of trade undefined for some transactions. We can observe the direction of trade directly in our data set, however, and therefore we define $Q_i = +1$ (-1) if the initiator of the trade bought (sold).

10.2.2 Trade Transparency, Payoffs and Summary Statistics

The market is low pre-trade and low post-trade transparent. This implies that market makers do not observe quotes publicly but have to call other market makers instead. Furthermore, the market makers can only observe their private transaction history on their computer screens.

Payoffs for each of the 16 rounds are converted from esquires into Dutch guilders according to the following scheme. The dealer with the highest profits in each round receives 7 guilders, and the dealer with the lowest receives 1 guilder. Guilder payoffs for the remaining 5 dealers are interpolated linearly between these two extremes.⁷⁰ Thus, the conversion into guilders is normalized to ensure that all participants receive a positive guilder payoff for their participation (and to constrain the absolute guilder cost to the experimenters).

Table 27 contains summary statistics describing the experiment. The parameter α (the probability that a customer trade is informed) alternates between 0.3 and 0.7 each round. Except for the two cases in which the learnable component is fully discovered (rounds 6 and 15), customer volume fluctuates between 92 and 96 shares; interdealer volume is anywhere from 3 to 10 times larger than dealer-customer volume. For the most part, average dealer profits are negative, while the informed customers have the highest profits. Of course, the unlearnable component of the terminal value can shock profits at the end of the trading round. For example, in round 8 average dealer profits are positive, while the informed customers lose money.

⁷⁰ Specifically, if dealer i 's esquire profits are denoted E_i , and the best and worst trader's earnings are E_{\max} and E_{\min} , then i 's guilder payoff is: $G_i = 1 + 6(E_i - E_{\min}) / (E_{\max} - E_{\min})$. Thus, G_i is an affine transformation of E_i , albeit a different transformation in each round.

27 Summary Statistics

This table summarizes the data from each of the 16 rounds of the experiment.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Parameters																
α	0.3	0.7	0.3	0.7	0.3	0.7	0.3	0.7	0.7	0.3	0.7	0.3	0.7	0.3	0.7	0.3
Terminal Value	200	142	82	102	138	121	81	102	171	55	2	114	205	67	27	23
- learnable part	150	142	82	52	188	121	31	152	171	55	52	64	155	117	77	23
- unlearn. part	50	0	0	50	-50	0	50	-50	0	0	-50	50	50	-50	-50	0
Data																
No. quotes set	59	40	55	61	54	58	60	64	63	60	59	47	52	57	59	66
No. of trades	438	1038	480	505	720	566	370	690	434	626	781	1080	976	780	762	676
- interdealer	343	942	388	410	625	486	274	594	339	531	686	985	880	685	698	581
- customer	95	96	92	95	95	80	96	96	95	95	95	95	96	95	64	95
Dealer profit	-734	-538	-73	-32	-292	-90	14	264	-370	-171	-567	-119	-861	-25	-53	-303
Profit uninformed	1087	446	78	65	459	41	58	-65	-37	248	46	520	831	19	172	65
Profit informed	1482	1436	180	47	562	274	-108	-858	1331	349	1939	-103	2184	70	15	995
% informed	27	68	29	71	28	73	31	75	67	27	67	26	74	27	59	27
Price discovery at	--	--	--	--	--	178	--	--	--	--	--	--	--	--	--	169

The parameter α is the probability that a given customer trade is informed. The row "Price discovery at" is the time within the four-minute round (in seconds) at which the best bid and ask first bracket the learnable component of the terminal value.

10.3 Results

10.3.1 *Serial correlation in changes in transaction prices*

Tables 28-31 present return autocorrelations and effective spreads from each of the four estimators — ROLL, CSS, JNR, and HS. The estimates in Table 28 reflect all transactions in the market, whereas Tables 29, 30 and 31 present the estimates for specific agent types (informed customers, uninformed customers, and dealers, respectively). Particularly striking is the extent to which these results confirm Roll's model: the first-order autocorrelation in Table 28 is consistently near Roll's predicted value of -0.5 , and, indeed, the average of these autocorrelations is exactly -0.5 . Note that this occurs despite the price discovery in these experiments. Price discovery implies a gradual movement in quoted prices toward the fundamental value, inducing positive serial correlation in quoted prices, and presumably transaction prices as well. Clearly this has little effect on the aggregate results in Table 28. Instead, interdealer trades are so frequent that movements in quoted spreads over time only have a negligible effect. This preponderance of interdealer trading is confirmed in Table 28: an overwhelming majority of trades are interdealer. Note that this is consistent with the result in Table 31 that the effective spreads for interdealer trades are close to the effective spreads for all trades.

Further confirming evidence appears in Tables 29 and 30 covering the customer trades. Since customers trade only every 3.5 seconds, more time elapses on average between customer trades, and therefore more price discovery occurs between them. This causes first-order autocorrelation to be more pervasive (i.e., less negative). For example, the average autocorrelation for price changes between customer trades (informed and uninformed together) is just -0.29 .⁷¹

Frequency of trade is only one force at work on the measured autocorrelations. We expect the negative autocorrelation to be smaller for informed customers since they tend to enter the market repeatedly on the same side. That is, price discovery in these experiments is typically monotonic in that prices start out either too high or too low and move gradually toward the learnable component of terminal value (with noisy fluctuations around this trend); once prices converge on the learnable component, however, the trend ceases, noise is dampened, and spreads narrow. Through the price

⁷¹ Statistics combining all customer trades do not appear in the tables here, but are available upon request.

discovery period there should be not be much bid-ask bounce in prices from the informed customer trades. Negative autocorrelation from informed trades would have to come from noise in dealer prices, much of which is idiosyncratic.⁷²

Table 29 reveals that the return autocorrelation from informed customer trades is in fact significantly less negative than in the sample of all transactions. Though this is consistent with the absence of bid-ask bounce, part of this higher (i.e., less negative) autocorrelation might be attributable to the lower average frequency of customer trades, as discussed above. To disentangle these we can compare informed customers to uninformed customers. Uninformed customers are liquidity traders who by design generate a bid-ask bounce when they trade. Comparing the autocorrelations in Tables 29 and 30 we find that, on average, the autocorrelation for informed customers is only half the magnitude of that for the uninformed (-0.14 versus -0.28), confirming the impact of bid-ask bounce.⁷³

10.3.2 Consistency across effective spreads models

Tables 28-31 also report effective spreads calculated from each of the four models described in section 10.1. Although there is occasionally considerable dispersion across the four measures in a given round, consistency across the measures is the rule. Table 32 measures the consistency of the effective spread estimators by comparing correlations among the estimators across all 16 rounds (i.e., correlating the rows of Table 28). The Roll and CSS measures, in particular, are closely related, both theoretically and empirically; the JNR measure is also highly correlated with these two. The HS measure, while showing the lowest correlation with the other measures, is nonetheless, strongly positive correlated.

⁷² Recall that the full queue of price quotes is not public information at any time, so that dealers do not have perfect knowledge of the inside spread or price history. Thus, rational dealers can generate seemingly anomalous price phenomena, such as unintended price reversals, or even negative inside (market) spreads, due solely to limited information.

⁷³ This difference remains qualitatively the same, even when one distinguishes between high and low alpha values. Recall that a high alpha value means that roughly 70% of customer trades are informed (and 30% uninformed), meaning that informed customer trades are relatively high frequency; a low alpha value reverses these percentages. The average autocorrelation over all 8 high-alpha rounds is -0.12 for informed customers, and -0.28 for uninformed customers; the numbers in low-alpha rounds are -0.16 for informed customers and -0.29 for the uninformed. The numbers are nearly identical to the overall average (i.e., ignoring the distinction between alpha values), arguing strongly against trade frequency as a possible factor in the difference between autocorrelations for informed and uninformed customers.

10.3.3 Effective spreads and the role of search costs

A more novel result from the experiments is the difference in effective spreads faced by customers and dealers. Customers here enjoy the (perhaps unrealistic) ability to trade against the best dealer price in the market. Dealers have to search the available prices on a bilateral basis. As a result, the customers face smaller effective spreads. This is noteworthy because the role of interdealer search as a component of the effective spread has been neglected in the literature.

The effective spread for customer trades thus provides a benchmark for evaluating the effective spread on interdealer trades. Dealers would also trade at the best available price if they knew it. The cost of their imperfect information corresponds to the difference between effective interdealer spreads and effective customer spreads. That is, at the margin a dealer ceases searching and executes a trade when the cost of continuing search just equals the benefit. The benefit of eliminating the search costs is the gap between the two effective spreads. This is our rationale for assigning the term “search costs” to this gap.

A comparison of the effective spreads in Table 31 with those in Tables 29 and 30 shows that these search costs are a significant component of the effective spread in a bilateral dealer market. For example, we find that the average effective spread for interdealer trades is 27.4 esquires using the Roll measure. The same measure produces an average effective spread for customer trades of 14.7 esquires. The gap between the two is 46 percent of the effective interdealer spread. Though this percentage varies across the other three modeling approaches, the message is the same: the cost of bilateral price search is a significant component of the spread.

10.4 Concluding remarks

We examine trading costs in markets where dealers search for quotes. Earlier work shows spreads are affected by three costs—order-processing, adverse-selection, and inventory-holding. Our results indicate a fourth cost component is present in bilateral dealer markets: search costs.

A central message of this chapter is that the effects of search are much larger and more pervasive than is currently recognized. Work to date on search within microstructure focuses narrowly on block trading. There is ample room to expand application. Indeed, as microstructure branches beyond the NYSE into some of the world’s largest markets, like spot FX, ignoring search costs becomes increasingly

untenable. In this chapter we examine trading costs in markets where dealers search for price quotes bilaterally. Using an experimental approach, we compare four popular models for estimating effective spreads faced by different trader types. Earlier papers show that spreads are affected by the costs of order processing, adverse selection, and inventory. Our results indicate the presence of a fourth factor in bilateral dealer markets: search costs. Search costs are a distinct component of the spread that has not been considered in the literature.

We also find that the theoretical implications of bid-ask bounce are borne out with remarkable accuracy in the time series of transaction prices. Roll's (1984) hypothesis of a negative auto-correlation in transaction price changes equal to -0.5 is matched nearly exactly in our interdealer trades. In contrast, the lower trading frequency of customers means that gradual price discovery raises the return autocorrelation from customer transactions toward zero, contaminating the relatively pure bid-ask bounce effect from interdealer trading.

28 All transactions

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Ave.
ROLL	38.29	31.74	27.44	6.50	13.63	50.15	27.44	20.10	5.18	33.61	15.10	57.02	106.31	29.29	13.79	14.97	30.66
CSS	57.30	34.72	43.23	14.49	15.07	67.15	34.44	23.39	7.14	83.37	18.46	72.04	159.95	45.63	34.75	23.40	45.91
JNR	25.15	28.64	19.18	8.87	13.53	44.69	26.07	10.63	4.52	19.00	12.80	69.23	197.13	25.44	9.27	11.95	32.88
HS	(1.39)	(1.44)	(1.43)	(0.39)	(0.45)	(31.79)	(1.59)	(0.74)	(0.29)	(1.75)	(0.57)	(20.03)	(58.56)	(2.97)	(1.22)	(0.69)	(17.44)
π	0.33	0.46	0.32	0.22	0.45	0.37	0.40	0.43	0.36	0.20	0.41	0.40	0.33	0.32	0.20	0.32	0.35
ACF1	-0.59*	-0.55*	-0.46*	-0.39*	-0.51*	-0.58*	-0.50*	-0.48*	-0.41*	-0.65*	-0.46*	-0.58*	-0.67*	-0.43*	-0.33*	-0.46*	-0.50*
ACF2	0.20*	0.11*	-0.11*	0.07	0.06	0.09	-0.05	-0.04	-0.08	0.31*	-0.05	0.18*	0.31*	0.07	-0.14*	0.10*	0.06
ACF3	-0.12*	-0.04	0.21*	-0.02	-0.02	0.04	0.18*	0.03	0.08	-0.13*	0.05	-0.05	-0.16*	-0.16*	0.02	-0.02	-0.01
ACF4	0.10	0.04	-0.06	-0.07	0.06	-0.08	-0.12	-0.00	-0.12	-0.02	-0.02	-0.01	0.13*	0.13*	-0.06	0.00	-0.01
ACF5	-0.22*	-0.02	-0.05	0.03	-0.08	0.14*	0.07	0.10*	0.05	0.01	0.00	0.11*	-0.08	-0.12*	-0.01	-0.07	-0.01

Standard errors are in parentheses. π is the estimated conditional probability of a trade reversal for the CSS model. * denotes significantly different from zero at the 5% level. ACF# denotes the estimated autocorrelation at order #.

29 Transactions of Informed Customers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Ave.
ROLL	11.23	30.30	16.81	8.20	1.71	2.75	3.23	1.63	1.35	7.75	5.24	9.69	35.39	18.70	12.83	6.26	10.82
CSS		386.34	96.67			23.82					136.27	96.93		107.50	64.17		113.10
JNR	13.11 (1.95)	66.87 (25.04)	10.49 (3.46)	7.23 (0.75)	14.13 (3.82)	9.26 (2.73)	12.68 (1.92)	6.61 (1.79)	3.47 (0.69)	11.76 (2.74)	12.13 (1.50)	14.79 (2.78)	15.52 (5.21)	13.25 (2.77)	3.29 (1.52)	8.79 (1.80)	13.96 (6.76)
HS		22.53 (11.60)	19.44 (6.77)			24.82 (5.17)					11.46 (2.81)	11.98 (5.95)		22.88 (5.98)	11.58 (5.17)		17.81 (2.53)
π	0.00	0.04	0.09	0.00	0.00	0.06	0.00	0.00	0.00	0.00	0.02	0.05	0.00	0.09	0.10	0.00	0.03
ACF1	-0.24	-0.23	0.18	-0.28	-0.00	-0.02	-0.04	-0.01	0.03	-0.25	-0.07	-0.19	-0.33*	-0.11	-0.44*	-0.18	-0.14
ACF2	0.20	-0.09	-0.07	-0.17	-0.36	-0.44*	-0.20	-0.51*	-0.03	-0.28	0.03	0.11	-0.05	0.04	-0.04	-0.14	-0.12
ACF3	-0.01	0.15	-0.18	-0.05	0.01	0.02	0.12	-0.00	-0.05	-0.04	-0.07	-0.47	-0.00	0.04	0.14	0.06	-0.02
ACF4	0.30	-0.01	-0.06	-0.05	0.20	-0.12	-0.09	0.04	-0.09	-0.11	0.14	0.22	-0.00	-0.46*	-0.15	0.09	-0.01
ACF5	-0.02	-0.02	0.01	-0.17	-0.07	0.03	-0.04	-0.01	-0.41*	0.23	-0.03	0.04	0.00	0.11	-0.00	-0.26	-0.04

Standard errors are in parentheses. π is the estimated conditional probability of a trade reversal for the CSS model. * denotes significantly different from zero at the 5% level. ACF# denotes the estimated autocorrelation at order #.

30 Transactions of Uninformed Customers

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Ave.
ROLL	25.92	30.90	28.07	4.48	12.80	28.24	18.72	9.42	4.49	50.07	8.59	39.41	2.45	19.67	3.76	11.54	18.66
CSS	29.68	31.63	25.42	6.73	11.52	36.72	18.18	12.16	2.83	42.92	7.34	43.59	3.78	20.86	3.34	11.22	19.24
JNR	17.51	27.52	13.75	7.39	9.71	10.92	15.57	7.33	3.24	18.02	9.50	24.10	10.73	12.56	3.72	10.15	12.61
HS	(3.17)	(4.25)	(3.43)	(1.22)	(1.15)	(3.45)	(2.04)	(3.14)	(0.58)	(8.99)	(1.49)	(4.69)	(3.09)	(1.61)	(1.21)	(1.60)	(3.45)
HS	16.16	37.47	16.65	3.94	9.26	19.36	16.21	3.14	4.02	18.91	11.42	31.50	5.53	13.81	0.34	11.66	13.71
	(2.79)	(3.42)	(3.47)	(1.73)	(1.06)	(4.55)	(2.16)	(2.58)	(0.75)	(5.64)	(1.35)	(4.15)	(1.88)	(2.22)	(1.11)	(1.81)	(2.87)
π	0.44	0.49	0.55	0.33	0.56	0.38	0.51	0.39	0.79	0.58	0.59	0.45	0.33	0.47	0.56	0.51	0.50
ACF1	-0.40*	-0.35*	-0.47*	-0.13	-0.45*	-0.55*	-0.29*	0.23	-0.26	-0.47*	-0.19	-0.43*	-0.02	-0.38*	-0.17	-0.20	-0.28*
ACF2	-0.02	-0.03	-0.04	-0.01	0.05	-0.01	-0.24	-0.34	-0.18	-0.07	-0.34*	-0.12	0.16	-0.12	-0.29	-0.20	-0.11
ACF3	-0.13	0.03	0.11	0.04	-0.15	0.03	-0.07	-0.38	0.08	0.05	-0.17	0.08	-0.50*	0.04	0.13	-0.21	-0.06
ACF4	0.06	-0.15	-0.16	-0.29	0.12	0.16	0.17	0.03	0.03	-0.04	0.28	-0.02	0.15	-0.04	0.20	0.06	0.03
ACF5	0.02	0.03	0.11	0.21	0.10	-0.14	0.15	0.01	0.00	0.05	0.07	-0.11	0.08	0.09	-0.26	0.14	0.03

Standard errors are in parentheses. π is the estimated conditional probability of a trade reversal for the CSS model. * denotes significantly different from zero at the 5% level. ACF# denotes the estimated autocorrelation at order #.

31 Interdealer transactions

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Ave.
ROLL	37.33	31.39	30.44	7.55	13.00	45.25	26.23	21.53	5.05	24.05	12.72	40.96	84.08	29.70	13.51	15.45	27.39
CSS	76.90	36.37	66.19	18.16	15.78	67.73	40.23	31.30	8.71	96.57	16.69	56.77	152.07	54.04	47.08	28.35	50.81
JNR	27.56	26.58	20.67	9.24	13.95	50.43	30.03	11.24	4.84	19.43	13.05	73.78	217.17	27.21	9.79	12.30	35.49
	(1.61)	(0.73)	(1.64)	(0.46)	(0.48)	(37.02)	(2.02)	(0.82)	(0.34)	(1.65)	(0.64)	(21.95)	(64.92)	(3.38)	(1.33)	(0.77)	(19.52)
HS	40.81	29.70	30.96	9.05	11.98	23.12	27.68	16.44	8.94	22.87	15.03	34.82	27.01	21.32	13.14	13.61	21.65
	(1.87)	(0.49)	(1.60)	(0.38)	(0.33)	(3.06)	(1.63)	(0.79)	(0.44)	(1.57)	(0.55)	(1.66)	(3.48)	(1.38)	(0.92)	(0.80)	(1.59)
π	0.24	0.43	0.23	0.21	0.41	0.33	0.33	0.34	0.29	0.12	0.38	0.36	0.28	0.27	0.14	0.27	0.29
ACF1	-0.57*	-0.59*	-0.56*	-0.50*	-0.49*	-0.50*	-0.45*	-0.50*	-0.47*	-0.51*	-0.39*	-0.34*	-0.52*	-0.42*	-0.30*	-0.49*	-0.48*
ACF2	0.18*	0.13*	0.19*	0.20*	0.05	0.01	-0.14*	-0.04	-0.00	-0.02	-0.11*	-0.30*	0.04	0.08*	-0.16*	0.32*	0.03
ACF3	-0.14	-0.07	-0.15*	-0.14*	-0.03	0.03	0.22*	0.07	0.09	0.15*	0.04	0.36*	0.11*	-0.17*	-0.02	-0.27*	0.01
ACF4	0.11*	0.06	0.18*	0.06	0.03	-0.03	-0.04	-0.03	-0.16*	-0.15*	0.04	-0.03	-0.10*	0.12*	-0.03	0.12*	0.01
ACF5	-0.19*	0.00	-0.22*	-0.15*	-0.05	0.02	-0.04	0.03	0.10	0.04	-0.01	-0.14*	0.12*	-0.07	0.03	-0.13*	-0.04

Standard errors are in parentheses. π is the estimated conditional probability of a trade reversal for the CSS model. * denotes significantly different from zero at the 5% level. ACF# denotes the estimated autocorrelation at order #

32 Correlation between spreads estimators

Correlations across all 16 rounds among the estimated effective spread measures, using all transactions (i.e., correlations among the first four rows of Table 28).

	Roll	CSS	JNR	HS
Roll	1.00	0.96	0.95	0.71
CSS		1.00	0.90	0.65
JNR			1.00	0.48
HS				1.00

11 Financial Market Competition

Big Bang - the radical reforms regarding the trading process on the London Stock Exchange in 1986 – resulted in an intense competition process between the exchanges in Europe. In that year, London installed the Stock Exchange Automatic Quotation International (SEAQ I) trading system, a dealership market with relaxed transparency regime and on which (indicative) quotes of market makers were publically disclosed on trading screens. The combined result was that investors faced smaller spreads and trading international stocks on SEAQ I was cheaper than on other European exchanges for this reason. This resulted in a significant migration of blue chip stock order flow from Continental Exchanges to London. The Continental exchanges underwent radical reforms in response which led to a successful comeback. In the decade following Big Bang, the microstructure of the European exchanges became increasingly similar⁷⁵. Pagano (1997) notes on the issue: 'This underscores an impressive convergence among European exchanges towards a common dualistic structure, formed by an automated auction system specializing in small and medium-sized trades and a market-making segment devoted to large trades.'

This competition process between the European exchanges will be further enhanced by the introduction of the Euro and the start of the European Monetary Union. As investors are willing to trade shares where trading costs are lowest and as capital mobility increases within an integrating Europe, investors will increasingly compare the trading costs over the different exchanges where they are able to trade shares of their interest.

⁷⁴ This chapter is based on: R. Huisman and K. Koedijk, 'Financial Market Competition: The Effects of Transparency', *De Economist*, 1998, 146, pp. 463-473.

⁷⁵ See Pagano (1997) for an extensive review of the reforms on the European exchanges.

With the focus on trading costs, the importance of financial market organization and trading mechanisms increases strongly.

In the following 'transparency grid' we summarize the effects of trading mechanism design on market performance and trading costs found in the previous chapters. In detail, the grid presents the impact of the trading mechanism characteristics pre-trade and post-trade transparency on liquidity and price efficiency. When choosing the suitable transparency of the trading process on a particular market, one faces a clear trade-off between liquidity and price efficiency. Clearly, a general optimum is unlikely to exist. The choice for a certain transparency regime must then be based on the needs of the investors that trade on the market and on the characteristics of the stocks that are being traded. This explains partly the convergence towards dual trading structures on European exchanges as discussed by Pagano (1997). Large transactions are better off in multiple dealer markets with a relaxed post-trade transparency regime as this improves the liquidity and thereby the cost of executing these block trades. This is exactly what happened after Big Bang in 1986, as most block trades in continental European stock were executed in the London SEAQ I dealer market.

Transparency Grid

This grid summarizes the effects of a ceteris paribus increase in either pre-trade or post-trade transparency on market variables as liquidity and price efficiency.

a ceteris paribus increase in		
	pre-trade transparency (quote disclosure)	post-trade transparency (transaction information)
liquidity	+/+	-/ (during the opening)
price efficiency	-/-	+/+

The grid also explains why different market structures co-exist, like the NASDAQ dealer market versus the NYSE specialist market and the new markets such as NMAX in Amsterdam, AIM in London and Le Nouveau Marché in Paris and EASDAQ in addition to the traditional exchanges. The Economist (1998) notes for example that 'NASDAQ has also managed to hold on to its best-known stocks as Intel and Microsoft, despite the efforts of the NYSE to poach them'. One of the reasons for this success lies in the relative advantages of the NASDAQ trading structures for Intel and Microsoft stock.

After the merger between NASDAQ and Amex, the US stock market is dominated by two exchanges. Since all eleven European countries that will participate in the Euro have their own financial markets, at least eleven stock exchanges exist that compete for order flow in the European capital market. It is expected that not all of them can survive in the same way as they operate now⁷⁶. This stresses the importance of the installed trading processes as these influence trading costs, liquidity and price efficiency; fundamental variables that investors will compare between the various exchanges before they will decide on where to trade. Given the fact that the European Stock Exchanges all converge to a similar trading process and knowing that no single optimal transparency regime exists, some European exchanges should restructure their trading process to gain from the trade-off between liquidity and efficiency in order to survive in the increased competitive capital market in the European Monetary Union. In order to survive, exploiting the differences between the transparency regimes will become increasingly important. From this point of view, it is unlikely that similar market designs among the European exchanges will prevail.

⁷⁶ Schinasi and Prati (1997) for a discussion on the effects of the European Monetary Union on the European capital market. Boot (1997) discusses the effects of order flow fragmentation on the competition process between European exchanges.

12 Overall concluding remarks

Although many pages have been devoted to the issue, the single real conclusion we can draw is that we gained a piece of extra insight into the puzzles in international financial markets. However, new results and ideas always lead to new questions. A researcher who is able to solve a problem does not keep his or her associate researchers with empty hands, but provides work instead. One could then make the claim that trying to solve all puzzles is the best way to fight unemployment, perhaps. Therefore, the main conclusions and ideas for future research will be presented in this chapter.

In chapter 2 we find extreme support for the uncovered interest parity. Research done in this area has rejected the null of uncovered interest parity frequently in the past decades. Many of these tests were conducted on a single exchange rate only and we argued that this leads to potential biases in the test results if exchange rates depend on worldwide economic factors in addition to idiosyncratic shocks. A change in the value of the dollar is likely to change several exchange rates that are denoted in terms of the dollar simultaneously. We propose a panel model to test for the validity of uncovered interest parity in order to be able to account for these common factors influencing exchange rate returns. The results show clearly that the rejection of uncovered interest parity is not as severe as many thought, although the null is still rejected. The results improve if we identify periods with high volatility. Then, it is shown that uncovered interest parity almost perfectly holds in periods with high volatility whereas it needs to be rejected in other periods; extreme support for uncovered interest parity. A sufficient explanation for this apparent difference between both periods has not been presented yet. It could be the case that the existence of transaction costs blocks FX traders from exploiting arbitrage

opportunities in low volatile periods, although much more research will be needed to make this case strong.

In chapter 3 and 4 we focus on extreme returns on financial assets. Extreme negative returns directly reflect the downside risk an investor faces on his or her investment. Extreme is defined as a tail event; we concentrate us on the tails of the distribution of returns on financial assets and on exchange rates in particular. The fatter a tail is, the more probability mass is in the tail, and the higher risk is that an investor faces. Measuring the tail fatness is thus essential to have a good insight in the exposure to such a risk factor. However, it is shown that common estimators to measure the tail fatness suffer from relatively big biases, especially for small samples that are commonly used in practice. We present a new estimator in chapter 3 that basically corrects for this small sample bias. Using this estimator we show that we have overstated the amount of tail fatness for, and thus the amount of downside risk faced from, many financial assets including stocks and exchange rates. This is shown in more detail for returns on exchange rates in chapter 4. The estimator makes that we could measure the amount of tail fatness correctly and we use this information to simply fit the tails of the empirical return distribution using a simple statistical distribution in which others usually failed. The simplicity makes that it can easily be incorporated to obtain better Value at Risk estimates, for example, to measure downside risk. This is demonstrated successfully by Huisman, Koedijk, and Pownall (1998). This gives many opportunities for future research in downside risk management for many assets or portfolios. Although the estimator presented in chapter 3 is shown to produce good results for many financial assets, we only know little about its (statistical) properties and how it behaves under different return generating processes, although Lucas (1997) showed that it provides good estimates for GARCH(1,1) processes that are apparent in financial data.

From chapter 5 we follow an alternative route to learn more about international financial markets. The experimental methodology is adopted to be able to study the performance of multiple dealer financial markets (exchanges). Although many important markets worldwide are organized as a multiple dealer market (examples are the interbank foreign exchange market, NASDAQ and the London Stock Exchange) their performance and the price formation process has gained only little attention in the literature. A multiple of competing dealers in a market is hard to implement in a theoretical framework without restricting behavior. Empirical studies have the disadvantage that it is very hard to disentangle the effects of the market structure alone

due to noise that might show up if one compares markets either over different time periods or over different countries. Experiments then offer a good alternative. This type of experimental literature is still in its infancy but it is promising according to Glosten (1998). The details of the experimental markets that we use in the following chapters are presented in chapter 6.

In chapter 7 we show the effects of disclosing price information to dealers in a financial market. We compare the results from markets in which market makers observe the quotes of other market makers continuously (as is the case in NASDAQ for example) with the results from markets in which market makers do not observe quotes from other market makers (as is the case in the foreign exchange market). Although one could expect that more information (thus quote disclosure) would lead to a higher efficiency, we show that this is not the case. The results indicate that market makers use the information from the quotes they see to fine-tune their price changes. They cannot do so in markets without quote disclosure. Therefore price discovery is blocked in markets with quote disclosure as market maker put more attention to the others than to their own thoughts. On the other hand, it is relatively time consuming to obtain a good overview of prices in markets without quote disclosure. It is therefore costly to find the good prices and makes that the implicit costs of trading in such markets are higher as is shown in detail in chapter 9.

Whereas chapter 7, and chapter 9 also basically, focuses on *pre-trade* transparency (i.e. quote disclosure), chapter 8 examines the effects of *post-trade* transparency on the price formation process in financial markets. We compare the results from markets in which market makers observe the immediate transaction information (the most transparent market) with the results from markets in which market makers do not observe transaction information. Interestingly, the results here are different from those found in chapter 7. Price efficiency is better in post-trade transparent markets whereas the implicit trading costs are higher. The latter result is due to the fact that market makers have less incentive to compete for order flow in post trade transparent markets because they can also retrieve information from observing the order flow.

The small amount of research projects that has been conducted so far in multiple dealer markets brings the opportunities that many things can still be done. One could think of applying the same methodology to option markets, examine the effects of the number of market makers on the market performance, having differently informed

traders etc. Another important extension could lie in the link between the experiments and theoretical literature. One could write a second thesis about this ...

13 References

- Admati, A.R. (1991) The informational role of prices: A review essay. *Journal of Monetary Economics*, 28, pp. 347-60.
- Arrelano, M. (1987) Computing Robust Standard Errors for Within-Groups estimators. *Oxford Bulletin of Economics and Statistics*, 49, pp. 431-434.
- Baillie, R.T. and T. Bollerslev (1997) The Forward Premium Anomaly is not as Bad as You Think. *Working Paper*.
- Bain, L.J. and M. Engelhardt (1987) *Introduction to Probability Theory and Mathematical Statistics*. Duxbury Press Boston.
- Baldwin, R.E. (1990) Re-interpreting the Failure of Foreign Exchange Market Efficiency Tests: Small Transaction Costs, Big Hysteresis Bands. *CEPR Discussion Paper*, # 407.
- Baltagi, B.H. (1995) *Econometric Analysis of Panel Data*. John Wiley & Sons.
- Beirlant, J., P. Vynckier and J. Teugels (1996) Tail-index Estimation, Pareto Quantile Plots and Regression Diagnostics. *Journal of the American Statistical Association*, 91, pp. 1659-1667.
- Bekaert, G. and R.J. Hodrick (1992) Characterizing Predictable Components in Excess Returns on Equity and Foreign Exchange Markets. *Journal of Finance*, 47, pp. 467-510.

- Bilson, J.F.O. (1981) The Speculative Efficiency Hypothesis. *Journal of Business*, 54, pp. 435-451.
- Bloomfield, R. (1996) Quotes, Prices, and Estimates in a Laboratory Market. *Journal of Finance*, 5, pp. 1791-1808.
- Bloomfield, R. and M. O'Hara (1996) Market Transparency: Who wins and who loses?. *Cornell University Working Paper*.
- Board, J. and C. Sutcliffe (1995) The Effects of Trade Transparency in the London Stock Exchange: A Summary. *LSE Financial Markets Group Special Paper*, # 67.
- Bollerslev, T., and I. Domowitz (1993) Some Effects of Restricting the Electronic Order Book in an Automated Trade Execution System. In: *The Double Auction Market: Institutions, Theories, and Evidence, Santa Fe Institute Studies in the Sciences of Complexity, Proceedings Vol. XIV*, eds. D. Friedman and J. Rust, Addison-Wesley, pp. 221-252.
- Boot, A. (1997) Financiële Markten en Concurrentie, *Pre-adviezen voor de Koninklijke Vereniging voor de Staathuishoudkunde*.
- Boothe, P. and D. Glassmann (1987) The Statistical Distribution of Exchange Rates: Empirical Evidence and Economic Implications. *Journal of International Economics*, 22, pp. 297-320.
- Campbell, J.Y., A.W. Lo and A.C. MacKinlay (1997) *The Econometrics of Financial Markets*. Princeton University Press.
- Choi, J.Y., D. Salandro and K. Shastri (1988) On the Estimation of Bid-Ask Spreads: Theory and Evidence. *Journal of Financial and Quantitative Analysis*, 23, pp. 219-230.
- Copeland, L.S. (1994) *Exchange Rates and International Finance*, 2nd edition. Addison-Wesley.
- Cox, D.R. and D.V. Hinkley (1974) *Theoretical Statistics*. Chapman and Hall.

- Cumby, R.E. and M. Obstfeld (1981) A Note on Exchange-Rate Expectations and Nominal Interest Differentials: A Test of the Fisher Hypothesis. *Journal of Finance*, 36, pp. 697-703.
- Cumby, R.E. and M. Obstfeld (1984) International Interest Rate and Price Level Linkages under Flexible Exchange Rates: A Review of Recent Evidence. In: *Exchange Rate Theory and Practice*, eds. J.F.O. Bilson and R.C. Marston, University of Chicago Press.
- Dacorogna, M.M., U.A. Müller, O.V. Pictet and C.G. de Vries (1995) The Distribution of Extremal Foreign Exchange Rate Returns in Extremely Large Data Sets. *Tinbergen Institute Discussion Paper*, #95-70.
- Davis, D.D., and C.A. Holt (1993) *Experimental Economics*, Princeton University Press.
- De Haan, L., D.W. Jansen, C.G. Koedijk and C.G. de Vries (1994) Safety First Portfolio Selection, Extreme Value Theory and Long Run Asset Risk. In *Extreme Value Theory and Applications*, eds. Galambos, pp. 471-487, Kluwer.
- Dekkers, A.L.M. and L. de Haan (1989) On the Estimation of the Extreme-Value Index and Large Quantile Estimation. *The Annals of Statistics*, 17, pp. 1795-1832.
- Director General of Fair Trading (1990) Trade Publication and Price Transparency on the International Stock Exchange: A Report to the Secretary of State for Industry. *Office of Fair Trading London Report*, April.
- Duxbury, D. (1995) Experimental Asset Markets within Finance. *Journal of Economic Surveys*, 9, pp. 331-71.
- Engel, C.M. and J.D. Hamilton (1990) Long Swings in the Dollar: Are They in the Data and Do Markets Know It? *American Economic Review*, 80, pp. 689-713.

- Evans, M.D.D. (1998) *The Microstructure of Foreign Exchange Dynamics*, working paper Georgetown University.
- Fama, E.F. (1984) Forward and Spot Exchange Rates. *Journal of Monetary Economics*, 14, pp. 319-338.
- Fialkowski, D. and M. Petersen (1994) Posted Versus Effective Spreads: Good Prices or Bad Quotes? *Journal of Financial Economics*, 35, pp. 269-292.
- Fisher, I. (1930) *The Theory of Interest*. Macmillan New York.
- Flood, M.D., R. Huisman, K. Koedijk and R. Mahieu (1996) Institutional Limits on Price Discovery in Multiple Dealer Financial Markets. *LIFE Working Paper*, 96-37.
- Flood, R.P. and A.K. Rose (1994) Fixes: of the Forward Discount Puzzle. *NBER Working Paper*, # 4928.
- Flood, R.P. and M.P. Taylor (1996) Exchange Rate Economics: What's Wrong with the Conventional Macro Approach. In *The Microstructure of Foreign Exchange Markets*, eds. J.A. Frankel, G. Galli and A. Giovannini, pp. 261-294, NBER.
- Franks, J. and S. Schaefer (1995) Equity Market Transparency on the London Stock Exchange. *Journal of Applied Corporate Finance*, 8, pp. 70-77.
- Gemmill, G. (1996) Transparency and Liquidity: A Study of Block Trades on the London Stock Exchange under Different Publication Rules. *Journal of Finance*, December, pp. 1765-1790.
- Glosten, L.R. (1998) Introductory Comments: Bloomfield and O'Hara, and Flood, Huisman, Koedijk, and Mahieu. Forthcoming in: *The Review of Financial Studies*.
- Glosten, L.R. and P.R. Milgrom (1985) Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders. *Journal of Financial Economics*, 14, pp. 71-100.

- Goodhart, C.A.E. and M. O'Hara (1994) High Frequency Data in Financial Markets: Issues and Applications. Paper presented at the First International Conference on High-Frequency Data in Finance, Olsen and Associates, Zurich.
- Guillaume, D.M., M.M. Dacoragna, R.R. Dav, U.A. Müller, R.B. Olsen, and O.V. Pictet (1994) From the Bird's Eye to the Microscope: A survey of new stylized facts of the intra-daily foreign exchange markets. *Olsen and Associates Working Paper*, Zurich.
- Harris, L. (1993) Consolidation, fragmentation, segmentation and regulation. *Financial Markets, Institutions, and Instruments*, pp. 1-28.
- Hasbrouck, J., G. Sofianos and D. Sosebee (1993) New York Stock Exchange Systems and Trading Procedures. *NYSE Working Paper*, #93-1.
- Hill, B. (1975) A Simple General Approach to Inference about the Tail of a Distribution. *Annals of Mathematical Statistics*, 3, pp. 1163-1174.
- Ho, T.S.Y. and H.R. Stoll (1981) Optimal Dealer Pricing under Transactions and Return Uncertainty. *Journal of Financial Economics*, 9, pp. 47-73.
- Huang, R.D., and H.R. Stoll (1997) The Components of the Bid-Ask Spread: A General Approach. *Review of Financial Studies*, 10, pp. 995-1034.
- Huisman, R., K.G. Koedijk, R.A.J. Pownall (1998) VaR+: Fat Tails in Financial Risk Management. *LIFE Working Paper*.
- Jansen, D.W. and C.G. de Vries (1991) On the Frequency of Large Stock Returns: Putting Booms and Busts into Perspective. *The Review of Economics and Statistics*, 73, pp. 18-24.
- Jong, F. de, T. Nijman and A.A. Röell (1995) A Comparison of the Cost of Trading French Shares on the Paris Bourse and on SEAQ International. *European Economic Review*, 39, pp. 1277-1301.

- Jorion, P. (1996) *Value at Risk: The New Benchmark for Controlling Derivatives Risk*. Irwin.
- Kearns, P. and A. Pagan (1997) Estimating the Density Tail Index for Financial Time Series. *The Review of Economics and Statistics*, 79, pp. 171-175.
- Koedijk, K.G. and P. Schotman (1990) How to Beat the Random Walk: An Empirical Model of Real Exchange Rates. *Journal of International Economics*, 29, pp. 311-332.
- Koedijk, C.G. and C. Kool (1994) Tail Estimates and the EMS Target Zone. *Review of International Economics*, 2, pp. 153-165.
- Koedijk, C.G., M.M.A. Schafgans and C.G. de Vries (1990) The Tail Index of Exchange Rate Returns. *Journal of International Economics*, 29, pp. 93-108.
- Koedijk, C.G., P.A. Stork and C.G. de Vries (1992) Differences between Foreign Exchange Rate Regimes: the View from Tails. *Journal of International Money and Finance*, pp. 462-473.
- Kyle, A.S. (1985) Continuous Auctions and Insider Trading. *Econometrica*, November, 6, pp. 1315-1335.
- Lamoureux, C.G. and C.R. Schnitzlein (1997a) When It's not the Only Game in Town: The Effect of Bilateral Search on the Quality of a Dealer Market. *Journal of Finance*, June, 2, pp. 683-712.
- Lamoureux, C.G. and C.R. Schnitzlein (1997b) Herd Through the Grapevine: Winner's Curse in a Fragmented Asset Market. *University of Arizona Working Paper*.
- Lewis, K.K. (1995) Puzzles in International Financial Markets. In *Handbook of International Economics*, eds. K. Rogoff and G. Grossman, 3, North Holland.
- Loretan, M. and P.C.B. Phillips (1994) Testing the Covariance Structure of Heavy-Tailed Time Series. *Journal of Empirical Finance*, 1, pp. 211-248.

- Lucas, A. (1997) *Discussion on: Fat Tails in Small Samples*. Presented at the 5th Workshop on Finance and Econometrics in Brussels, December.
- Lyons, R. (1995) Tests of Microstructural Hypotheses in the Foreign Exchange Markets. *Journal of Financial Economics*, 39, pp. 321-351.
- Lyons, R. (1996) Optimal Transparency in a Dealer Market with an Application to Foreign Exchange. *Journal of Financial Intermediation*, 5, pp. 225-254.
- Lyons, R. (1998) Profits and Position Control: A Week of FX Dealing. *Journal of International Money and Finance*, forthcoming.
- Madhavan, A. (1992) Trading Mechanisms in Securities Markets. *Journal of Finance*, June, 2, pp. 607-641.
- Madhavan, A. (1995) Consolidation, Fragmentation, and the Disclosure of Trading Information *Review of Financial Studies*, 8, pp. 579-603.
- Mahieu, R. and P. Schotman (1994) Neglected Common Factors in Exchange Rate Volatility. *Journal of Empirical Finance*, 1, pp. 279-311.
- McCulloch, J.H. (1997) Measuring Tail Thickness to Estimate the Stable Index α : A Critique. *Journal of Business Economics and Statistics*, 15, 11, pp. 74-81.
- Meese, R.A. and K. Rogoff (1983a) Empirical Exchange Rate Models for the Seventies: Do They Fit Out of Sample?, *Journal of International Economics*, 14, pp. 3-24.
- Meese, R.A. and K. Rogoff (1983b) The Out-of-Sample Failure of Empirical Exchange Rate Models: Sampling Error or Misspecification? In: *Exchange Rates and International Macroeconomics*, ed. J.A. Frenkel, University of Chicago Press.
- Meese, R.A. and K. Rogoff (1988) Was it Real? The Exchange Rate - Interest Differential Relation over the Modern Floating Rate Period. *Journal of Finance*, 43, pp. 933-948.

- Naik, N., A. Neuberger and S. Viswanathan (1994) Disclosure Regulation in Competitive Dealer Markets: Analysis of the London Stock Exchange. *London Business School IFA Working Paper*, # 193.
- National Association of Securities Dealers (1983) Analysis of the Impact of Last Sale Reporting on Market Characteristics of Tier 2 NASDAQ Securities. *Unpublished*, NASD, Washington D.C.
- O'Hara, M. (1995) *Microstructure Theory*. Basil Blackwell.
- Pagano, M. (1997) The Changing Microstructure of European Equity Markets. Forthcoming in: *The European Securities Markets: Implementing the Investment Services Directive and Beyond*, ed. G. Ferrarini, Kluwer Law International.
- Pagano, M. and A.A. Röell (1992) Auction and Dealership Markets: What is the Difference? *European Economic Review*, 36, pp. 613-623.
- Pagano, M. and A.A. Röell (1996) Transparency and Liquidity: A Comparison of Auction and Dealer Markets with Informed Trading. *Journal of Finance*, June, pp. 579-611.
- Pagano, M. and B. Steil (1996) Equity Trading I: The Evolution of European Systems. In *The European Equity Markets*, ed. B. Steil, European Capital Markets Institute, pp. 553-611.
- Perraudin, W. and P. Vitale (1996) Interdealer Trade and Information Flows in a Decentralized Foreign Exchange Market. In *The Microstructure of the Foreign Exchange Market*, eds. J. Frankel, G. Gallie and A. Giovannini, University of Chicago Press and NBER, pp. 183-201.
- Pictet, O., M. Dacorogna and U. Müller (1996) Hill, Bootstrap and Jackknife Estimators for Heavy Tails. *Olsen & Associates Working Paper*.
- Porter, D.C. and D.G. Weaver (1987) Post-Trade Transparency on NASDAQ's National Market System. *University of Wisconsin Working Paper*.

- Reiss, P. and I. Werner (1997) Interdealer Trading: Evidence from London. Stanford Graduate School of Business Research Paper, # 1430, February.
- Röell, A.A. (1988) Regulating Information Disclosure among Stock Exchange Market Makers. *LSE Financial Markets Group Discussion Paper*, # 51.
- Röell, A.A. (1996) Stock Market Transparency. *Working Paper*.
- Rogalski, R.J. and J.D. Vinso (1978) Empirical Properties of Foreign Exchange Rates. *Journal of International Business Studies*, 9, pp. 69-79.
- Roll, R. (1984) A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market. *Journal of Finance*, 39, pp. 1127-1139.
- Schnitzlein, C.R. (1996) Call and Continuous Trading Mechanisms under Asymmetric Information: An Experimental Investigation. *Journal of Finance*, June, 2, pp. 613-636.
- Schotman, P., S. Straetmans and C.G. de Vries (1997) Big News in Small Samples. *Maastricht University Working Paper*.
- Securities and Investments Board (1995a) Regulation of the United Kingdom Equity Markets: Market Views. A Digest of Responses to S.I.B.'s Discussion Paper. London, June.
- Securities and Investments Board (1995b) Regulation of the United Kingdom Equity Markets: Report, London, June.
- Schinasi, G.J. and A. Prati (1997) European Monetary Union and International Capital Markets: Structural Implications and Risks, *paper presented at the IMF seminar on EMU and the International Monetary System*, March 17-18, Washington D.C.
- Stockmarket Mergers: Marry in Haste, *The Economist*, 1998, March 21st, 92-93

- Stoll, H.R. (1992) Principles of Trading Market Structure. *Journal of Financial Services Research*, May, pp. 75-107.
- Westerfield, J.M. (1977) An Examination of Foreign Exchange Risk under Fixed and Floating Rate Regimes. *Journal of International Economics*, 7, pp. 181-200.
- Yu, G.G. (1993) Information Revelation and Aggregation in Financial Markets, *Financial Markets, Institutions and Instruments*, December, pp. 29-54.

14 Nederlandse Samenvatting

Aan het einde van de 20^e eeuw zijn de internationale financiële markten onderhevig aan enorme structurele veranderingen die naar verwachting zullen leiden tot een flinke groei maar wel gepaard gaan met enorme turbulentie in die markten. Technologische ontwikkelingen, het wegvallen van allerlei (handels) barrières (denk bijvoorbeeld aan de Europese Monetaire Unie), maar ook onverwachtse gebeurtenissen zoals een wel-of-niet seksuele relatie tussen president Clinton en de stagiaire Lewinsky (het eerste nieuwsbericht hierover leidde tot een verlies in de waarde van de Amerikaanse Dollar gelijk aan \$120 miljard gemeten aan de geldhoeveelheid van de VS aldus het NRC Handelsblad) hebben een significant effect. Inzicht in de werking van deze markten is en blijft dus van belang en vormt de motivatie voor de studies waarvan de resultaten worden gepresenteerd in dit proefschrift. *Adventures in international financial markets* bestaat uit twee onderdelen welke verbonden zijn door de rode draad - de werking van internationale financiële markten - maar qua natuur en methodologie sterk van elkaar verschillen. Het eerste deel is geconcentreerd rond het gedrag van wisselkoersen en het modelleren van de koersbewegingen. In deel 2 wordt onderzocht hoe de structuur van het handelsproces zoals dat is geïnstalleerd op een effectenbeurs van invloed is op de karakteristieken van het verhandelde financiële instrument en op de kwaliteit van die financiële markt.

14.1 Deel 1: Wisselkoersen

Hoofdstuk 2 handelt om een inmiddels klassiek probleem: de ongedekte rentepariteit blijkt in de praktijk nauwelijks op te gaan. De ongedekte rentepariteit relateert wisselkoersveranderingen aan renteversillen tussen landen en zou theoretisch een zuivere schatting van toekomstige wisselkoersveranderingen moeten geven. Traditioneel

wordt deze pariteit getoetst door te concentreren op de rentestanden in slechts twee landen en de tussenliggende wisselkoers. In hoofdstuk 2 wordt op een niet bilaterale manier naar deze pariteit gekeken. Immers, een economische schok in de VS die een effect heeft op de Euro/Dollar wisselkoers zal naar verwachting ook een effect hebben op de Yen/Dollar wisselkoers. Er wordt een panel-model gepresenteerd waarmee de hypothese dat de ongedekte rentepariteit opgaat getoetst wordt op basis van meer dan één valuta uitgedrukt in dezelfde onderliggende valuta. De resultaten tonen dat de hypothese nog steeds verworpen dient te worden, maar dat de afwijkingen niet zo groot zijn als die in de traditionele toetsen naar voren kwamen. De panel methodologie wordt vervolgens gebruikt om het Bilson (1981) effect te onderzoeken. Bilson toonde aan dat de empirische afwijkingen van de ongedekte rentepariteit afhangen van de mate van volatiliteit (koersschommelingen) in de markt. De resultaten in hoofdstuk 2 laten zien dat het effect inderdaad opgaat en, sterker, dat de ongedekte rentepariteit terdege opgaat in zeer volatiele perioden.

Hoofdstukken 3 en 4 concentreren beide op extreme bewegingen van wisselkoersen. Om de verdeling van rendementen op financiële waarden als aandelen of wisselkoersen te modelleren wordt er vaak van uitgegaan dat de rendementen normaal verdeeld zijn. Het is daarentegen welbekend dat de verdelingen van rendementen op financiële waarden dikstaartig zijn, wat betekent dat de kans waarmee grote opwaartse en/of neerwaartse koersschommelingen zullen plaatsvinden groter is dan wat de normale verdeling veronderstelt. Een goed beeld van de mate van dikstaartigheid is dus noodzakelijk om bijvoorbeeld een goede inschatting te maken van de hoeveelheid neerwaarts risico dat men loopt. In hoofdstuk 4 wordt getoond hoe een expliciete maatstaf voor dikstaartigheid, de tail-index, gebruikt kan worden om de verdeling van wisselkoersrendementen succesvol en gemakkelijk te modelleren. Het bepalen van de waarde van de tail-index is echter moeilijk omdat van schatters die deze waarde bepalen bekend is dat zij leiden aan niet triviale afwijkingen in kleine steekproeven. Alhoewel tegenwoordig grote bestanden aanwezig zijn met hoge frequentie data, moeten voor vele toepassingen binnen de financiering nog steeds relatief kleine steekproeven gebruikt worden. In hoofdstuk 3 wordt een schatter voor de tail-index gepresenteerd die corrigeert voor de afwijkingen in kleine steekproeven. Met behulp van deze schatter wordt aangetoond dat men de hoeveelheid dikstaartigheid van rendementen op aandelen en wisselkoersen veelal overdreef in het verleden. Uit hoofdstuk 4 blijkt tevens dat deze

schatte gebruikt kan worden om de staarten van wisselkoersrendement verdelingen te modelleren.

14.2 Deel 2: De microstructuur van financiële markten

Een van de belangrijkste vraagstukken in het kader van de Europese eenwording is wat er in de toekomst zal gebeuren met de effectenbeurzen. De VS kennen slechts twee grote dominante beurzen, de New York Stock Exchange (NYSE) en de combinatie Nasdaq/Amex, naast een aantal regionale beurzen. In het Europa van voor de eenwording heeft ieder land tenminste één effectenbeurs en de vraag is in hoeverre deze situatie kan blijven bestaan op de lange termijn. Gezien de strategische allianties (bijvoorbeeld London & Frankfurt) die worden besproken en uitgevoerd, vrezen een aantal beurzen voor hun bestaan. De hamvraag is dan ook welke beurzen wel en welke niet blijven bestaan. Veel factoren zijn hierop van invloed. In deel 2 wordt gekeken naar hoe de structuur van het handelsproces op een beurs van invloed is op de prijsvorming van financiële waarden die er verhandeld worden en op de impliciete transactiekosten. Deze zijn dus bepalend in de perceptie van beleggers aangaande de markt en spelen dus een belangrijke rol in de veranderingen van het Europese effectenverkeer.

De focus in het tweede deel is op zogenaamde multi-dealer financiële markten. Op deze markten concurreren meerdere dealers (market makers) om transacties binnen te halen. Voorbeelden van dit soort markten zijn Nasdaq, de London Stock Exchange, de wholesale markt in Amsterdam en de interbancaire wisselkoersmarkt. Een tegenhanger is bijvoorbeeld de NYSE welke een single-dealer ofwel een specialist markt is. De voorkeur voor multi-dealer financiële markten in dit proefschrift komt door het feit dat de literatuur zich met name gericht heeft op single-dealer markten. Deels kan dit worden verklaard door de dominante rol van de NYSE. Een andere belangrijke oorzaak is dat het in theoretische modellen zeer moeilijk is om het gedrag van meer dan één dealer te modelleren zonder de heterogeniteit te beperken. Empirische onderzoeken hebben als nadeel dat indien men de voordelen van een multi-dealer markt wilt vergelijken met een single-dealer markt de testen altijd verstoord worden doordat de vergelijking niet kan worden gedaan op basis van dezelfde aandelen of vanuit hetzelfde land. Het is dus moeilijk om het multi-dealer effect direct te destilleren. Om deze problemen te omzeilen is experimentele economie gebruikt in dit proefschrift. Professionele handelaren van Nederlandse banken en de EOE participeerden in een fictieve markt waarin het effect van fundamentele karakteristieken op de markt onderzocht werd.

De hoofdstukken 7 en 8 onderzoeken het effect van transparantie op de performance van de markt. Transparantie geeft aan hoeveel informatie aan de dealers gegeven wordt, en is dus bepalend voor welke informatie zij gebruiken in hun prijsstrategie. Twee soorten transparantie worden onderscheiden: pre-trade en post-trade transparantie. De eerste staat centraal in hoofdstuk 7. Pre-trade transparantie bepaalt hoeveel prijsinformatie handelaren krijgen in de markt. Ziet men alleen de eigen bied- en laatprijs of ook die van de anderen in de markt? Het blijkt dat een markt die weinig pre-trade transparant is (i.e. handelaren zien alleen hun eigen prijzen) een veel snellere prijsontwikkeling doormaakt. Prijzen passen zich sneller aan nieuwe informatie. Deze verklaring lijkt in tegenstelling met wat men zou verwachten in efficiënte markten: hoe meer informatie hoe beter. Echter, als handelaren de prijzen van anderen zien gebruiken zij deze informatie ook in hun prijsstrategie. Prijsaanpassingen zijn dus deels gebaseerd op wat er elders in de markt gebeurt. Indien handelaren niet de prijzen van anderen zien, reflecteren hun prijzen veel meer hun eigen informatie en inzichten en worden dus niet beïnvloed door wat anderen doen. Een voorbeeld verduidelijkt dit. Stel voor dat een handelaar een informatie signaal krijgt dat een aandeel eigenlijk 100 gulden waard moet zijn, maar dat het aandeel nu gemiddeld 90 gulden doet in de markt. In een markt waar deze handelaar alle prijzen ziet zou zij deze informatie niet direct in de prijzen weergeven; anders weet de rest het ook direct. Een betere strategie is juist om de informatie niet in de prijs te reflecteren maar direct het aandeel te kopen in de markt voor 90 waardoor de informatie relatief langzaam in de prijzen komt. Als de handelaar nu de prijzen van anderen niet gezien zou hebben, is deze strategie moeilijker uit te voeren omdat het relatief veel tijd kost om uit te zoeken wat de prijzen van de anderen zijn en of anderen deze informatie bijvoorbeeld ook allang in hun prijzen verwerkt hebben (deze zoekkosten worden behandeld in hoofdstuk 9). Een handelaar heeft dus een prikkel om de informatie snel in haar bied- en laatprijs te verwerken. Tegenover het nadeel dat een pre-trade transparante markt minder efficiënte prijzen heeft, staat het voordeel dat de competitie tussen handelaren heviger is wat resulteert in weidere spreads en hogere transactiekosten.

In hoofdstuk 8 staat post-trade transparantie centraal. In een post-trade transparante markt zien handelaren direct alle transacties die waar dan ook in de markt plaatshebben. In een niet post-trade transparante markt gebeurt deze informatieverstrekking of met een vertraging of helemaal niet waardoor handelaren slechts aangewezen zijn op hun eigen transactiehistorie. De effecten van post-trade transparantie liggen anders dan bij pre-trade

transparantie. Een post-trade transparante markt (i.e. transactieinformatie wordt direct aan iedere handelaar vertoond) leidt tot betere prijzen maar tot weidere spreads en dus hogere transactiekosten dan in een niet post-trade transparante markt. De trade-offs tussen prijsefficiëntie en spreads werken dus verschillend voor pre-trade en post-trade transparantie. In hoofdstuk 10 wordt dan ook geconcludeerd dat én het bestaan van de trade-offs én de verschillen tussen pre-trade en post-trade transparantie als gevolg hebben dat niet één handelssysteem optimaal is. Sterker, een economie waarin verschillende financiële markten bestaan waarbij de handelsprocessen karakteristiek verschillen profiteert hier dus het meeste van. Het streven naar één enkele effectenbeurs in Europa waar alle aandelen op een zelfde wijze verhandeld worden is dan ook niet optimaal.

CURRICULUM VITAE

Ronald Huisman, born on January 17 1972 in The Hague, received his Master's degree in Business Econometrics at the Erasmus University Rotterdam in August 1995. He immediately joined the finance department at Maastricht University and worked as a Ph.D. student on this thesis. In March 1997 he became an Assistant Professor in Finance at Maastricht University. He moved to the Erasmus University Rotterdam in October 1998, where he is still employed as an Assistant Professor in Finance. Ronald Huisman published in international journals such as *Review of Financial Studies*, *Journal of International Money and Finance*, *Journal of Risk*, *The Economist*, and *Real Estate Economics*. In the Netherlands, he published in the *ESB*, *VBA Journaal en Bedrijfskunde*. His current research interests are Financial Risk Management.

Ronald Huisman werd op 17 januari 1972 geboren te 's Gravenhage. Hij studeerde bedrijfseconometrie aan de Erasmus Universiteit Rotterdam en slaagde voor het doctoraal examen in Augustus 1995. Hij startte daarna als AIO bij de sectie financiering aan de Universiteit van Maastricht. Vanaf maart 1997 werkte hij daar als universitair docent in de financiering. Sinds oktober 1998 werkt hij aan de Erasmus Universiteit Rotterdam als universitair docent in de financiering. Ronald Huisman publiceerde in internationale tijdschriften als *Review of Financial Studies*, *Journal of International Money and Finance*, *Journal of Risk*, *The Economist*, and *Real Estate Economics*. Verder publiceerde hij in Nederlandse tijdschriften als de *ESB*, het *VBA Journaal en Bedrijfskunde*. Zijn huidige onderzoeksinteresse is financieel risicomanagement.

THELA • THESIS

