LIQUIDITY AND PORTFOLIO OPTIMISATION

BJÖRN HAGSTRÖMER

Doctor of Philosophy

Economics and Strategy Group

ASTON UNIVERSITY

July 2009

1

Aston University

Liquidity and Portfolio Optimisation

Björn Hagströmer

Ph.D. 2009

This thesis presents research within empirical financial economics with focus on liquidity and portfolio optimisation in the stock markets. The discussion on liquidity is focussed on measurement issues, including TAQ data processing and measurement of systematic liquidity factors. The portfolio optimisation section evolves around the properties of full-scale optimisation (FSO). Furthermore, a framework for treatment of the two topics in combination is provided.

The liquidity part of the thesis gives a conceptual background to liquidity and discusses several different approaches to liquidity measurement. It contributes to liquidity measurement by providing detailed guidelines on the data processing needed for applying TAQ data to liquidity research. The main focus, however, is the derivation of systematic liquidity factors. The principal component approach to systematic liquidity measurement is refined by the introduction of moving and expanding estimation windows, allowing for time-varying liquidity co-variances between stocks. Under several liquidity specifications this improves the ability to explain stock liquidity and returns, as compared to static window PCA and market average approximations of systematic liquidity. The highest ability to explain stock returns is obtained when using inventory cost as a liquidity measure and a moving window PCA as the systematic liquidity derivation technique. Systematic factors of this setting also have a strong ability in explaining cross-sectional liquidity variation.

Portfolio optimisation in the FSO framework is tested in two empirical studies. These contribute to the assessment of FSO by expanding the applicability to stock indexes and individual stocks, by considering a wide selection of utility function specifications, and by showing explicitly how the full-scale optimum can be identified using either grid search or the heuristic search algorithm of differential evolution. The studies show that relative to mean-variance portfolios, FSO performs well in these settings and that the computational expense can be mitigated dramatically by application of differential evolution.

Keywords: *stock market, differential evolution, systematic liquidity, commonality, high-frequency data*

To my father

Acknowledgements

Related Publications

Journal Articles

- Hagströmer, B., Anderson, R. G., Binner, J. M., Elger, T. and Nilsson, B.: 2008, Mean-Variance Vs. Full-Scale Optimization: Broad Evidence for the UK, *The Manchester School* **76**, 134–156. Previously published in:

  - *Aston Business School Working Paper Series* **RP0711**;
  - *Federal Reserve Bank of St Louis Working Paper* **2007-016**;
  - *Lund University Department of Economics Working Paper* **2008:1**.

- Hagströmer, B. and Binner, J. M.: 2009, Stock Portfolio Selection with Full-Scale Optimization and Differential Evolution, *Forthcoming in Applied Financial Economics*. Previously published in:

  - *Aston Business School Working Paper Series* **RP0808**;

Working Paper

- Hagströmer, B., Anderson, R. G., Binner, J. M. and Nilsson, B.: 2009, Dynamics in Systematic Liquidity. Published in:

  - *Federal Reserve Bank of St Louis Working Paper* **2009-025**;
  - *Lund University Department of Economics Working Paper* **2009:7**.

Conference Paper Presentations

- 14th International Conference on Computing in Economics and Finance 2008 (Paris);
- Arne Ryde Seminar in Financial Economics 2008, 2009 (Lund);
- Departmental Seminar, Lund University 2008;
- EFMA Annual Conference 2007 (Vienna);
- EFM "Merton H. Miller" Doctoral Student Seminar 2007 (Vienna);
- Financial Management Association European Conference 2007 (Barcelona), 2008 (Prague);
- FMA European Doctoral Student Seminar 2007 (Barcelona), 2008 (Prague);
- Forecasting Financial Markets 2009 (Luxembourg)
- Money, Macro, and Finance Conference 2007 (Birmingham);
- Nordic Finance Network Workshop 2008 (Bergen);
- PhD Colloquium 2007 (Birmingham).

External Support and Training

Visiting Scholar

- Federal Reserve Bank of St Louis, December 2007: *Using TAQ to Measure Liquidity*;
- Federal Reserve Bank of St Louis, May 2008: *Dynamics in Systematic Liquidity*.
- Lund University, Department of Economics, February–July 2008;

Extra-Curricular Training

- Computational Methods in Finance, 2007 (Summer School at CCFEA, Essex University);
- Econometrics, 2007 (PhD Course at Warwick University);
- Financial Market Microstructure and Contagion, 2009 (Summer School at University Aix-Marseille);
- Liquidity – Modelling, Recent Crisis and Challenges, 2009 (Oxford-Man Institute Conference);
- Principles of Finance, 2008 (PhD Course at Lund University).

Contents

List of Figures

List of Tables

List of Abbreviations

Δ	First difference operator
AMEX	American Stock Exchange
BBO	Best Bid and Offer
BID	TAQ: Bid price
BIDSIZ	TAQ: Order volume corresponding to BID
CAPM	Capital Asset Pricing Model
COND	TAQ: Variable containing information on sale conditions
CORR	TAQ: Variable containing information on corrected trades
CUSIP	TAQ: CUSIP stock identification number (Committee on Uniform Security Identification Procedures)
DENOM	TAQ: Tick size
E	Expected value operator
$Espread$	Effective relative bid-ask spread
EX	TAQ: Exchange variable indicating where the transaction took place
FSO	Full-scale optimization
ILLIQ	Amihud's (2002) illiquidity measure
IQR	Interquartile range
LTCM	Long Term Capital Management
MODE	TAQ: Variable containing information on specific conditions around quotes
NAME	TAQ: Full company name
MV	Mean-variance
NASDAQ	National Association of Securities Dealers Automated Quotations

OFR	TAQ: Ask price
OFRSIZ	TAQ: Order volume corresponding to OFR
OLS	Ordinary least squares
PCA	Principal component analysis
PDF	Probability distribution function
PRICE	TAQ: Transaction price
Qspread	Quoted relative bid-ask spread
QTIM	TAQ: Time stamp for quotes
S&P500	Standard & Poor's index of 500 large-cap stocks actively traded in the US
SHARESOUT	TAQ: Number of shares outstanding
SIZ	TAQ: Transaction volume
SYMBOL	TAQ: Ticker symbol of a stock
TAQ	Trades and Quotes database
TAQ3	Software for downloading data from TAQ
TTIM	TAQ: Time stamp for trades
TYPE	TAQ: Variable containing information on share type
UK	United Kingdom
US	United States
UT	TAQ: Minimum lot size

Part I

Introduction

Chapter 1

Introduction

This thesis is set in the field of empirical financial economics. The portfolio choice framework of mean-variance optimisation (MV; Markowitz 1952, 1959) and the capital asset pricing model (CAPM; Lintner 1965, Mossin 1966, Sharpe 1964) have had tremendous importance for both the development of financial economics as an academic discipline, and for the development of the financial industry to what it is today. Constituting the foundations of financial economics theory, these models have been subject to extensive assessments, improvements and extensions ever since they were founded. An important aspect of this work has been to clarify the implications of key assumptions in the models, and to find ways to relax these assumptions.

One clearly unrealistic assumption is that of frictionless financial markets — it is well-known that it takes time, money, and effort for a trade to happen. Another assumption states that investors' preferences can be described using quadratic utility functions, implying that the mean and variance of expected portfolio returns are enough to make portfolio allocation decisions. This differs considerably from most investors' reality, where risk of extreme events, illiquidity risk, and downside risk are important variables for the decision-making. The relaxation of these two assumptions set the stage for this thesis, whose contributions lie in the fields of liquidity and portfolio optimisation.

Liquidity is a concept describing the friction investors experience when trading securities. This friction matters to investors to the extent that it significantly affects asset prices (Amihud and Mendelson 1986a, Acharya and Pedersen 2005). Building on the finding that there are market-wide factors driving changes in the liquidity of stocks across financial markets (Chordia, Roll, and Subrahmanyam 2000, Huberman and Halka 2001, Hasbrouck and Seppi 2001), the aim of the liquidity study in Part II of this thesis is to establish how those factors are best measured.

Full-scale optimisation (FSO) is a recent method for portfolio choice that is fully flexible in the formulation of investor preferences (Cremers, Kritzman, and Page 2005, Adler and

Kritzman 2007), relaxing the classical assumption of quadratic utility functions. In the analysis of Part III of the thesis, this approach is assessed empirically in different portfolio choice settings, and different ways of finding the optimal portfolio are tested.

These parts have in common that they are both empirical in their nature, and that they deal with equity market applications with the aim to show facts of direct relevance to the financial industry. Both parts contain findings with implications for the investment decision. In the concluding part of the thesis (Part IV), some ideas on how the different concepts can be used in combination are offered.

1.1 Disposition

As indicated above, the thesis is divided into four parts: Introduction (this chapter), Liquidity (three chapters), Portfolio Optimisation (three chapters), and Conclusions (one chapter). To give the reader an overview of the work, I give here a brief introduction to each of the eight chapters.

The focus of Part II is measurement of liquidity and derivation of systematic liquidity factors. In Chapter 2, I give my understanding of the different dimensions of liquidity. In that context I present eight different liquidity measures that I consider for the liquidity analysis. Furthermore, I refer the literature on how liquidity affects stock prices. This gives the understanding necessary for the analysis in the subsequent two chapters on liquidity. To measure monthly stock liquidity I use high-frequency stock market data covering the constituent stocks of S&P500 from 1995–2007. The data for this are available in the Trades and Quotes database (TAQ). This is a vast database and a great resource for stock market research, but it takes careful processing to digest its billions of observations on transactions and orders. In Chapter 3, I present in full the processing I have done of this data set, referring to the rather limited literature available on the topic. Specifically, I discuss issues of raw data filtering, simultaneous observations on trades and quotes, matching of trades and quotes, and the detection of erroneous data points.

In Chapter 4, I turn to analysis of how individual stock liquidity co-varies across stocks — the main aim of the liquidity analysis in the thesis. The degree of co-variance is referred to as commonality in liquidity. Commonality is driven by systematic liquidity factors that have been identified as risk factors with influence on asset prices. My focus is the measurement of these risk factors. To account for the possibility of time-varying co-variance structures in liquidity, I introduce principal component analysis (PCA) with a dynamic estimation window as a way of estimating systematic liquidity factors. I evaluate different versions of this method along with traditional estimation techniques (full sample PCA and market average). My evaluation criteria are (1) the ability to explain cross-sectional stock liquidity, and (2) the ability to explain cross-sectional stock returns. For most of the common liquidity measures my

results suggest that an expanding window specification of PCA is appropriate for systematic liquidity estimation. For price impact liquidity measures (measuring inventory costs and adverse selection costs), however, I find support for a moving window specification. The market average proxy of systematic liquidity, which has been advocated in the literature, produces the same degree of commonality, but does not have the same ability to explain stock returns as the PCA-based estimates. Furthermore, I study dynamics in both liquidity commonality and the explanatory power liquidity on stock returns, showing that these can vary substantially over time. The findings of this chapter have also been published as a *Federal Reserve Bank of St Louis Working Paper* (Hagströmer, Anderson, Binner, and Nilsson 2009).

Part III of the thesis deals with portfolio optimisation. In Chapter 5, I present portfolio optimisation models in a utility maximisation setting. In particular, I focus on the FSO model, where the empirical return distribution is applied to a utility function. The advantage of this model is that preferences for skewness and kurtosis can be taken into account when formulating the utility function, as no analytical solution is pursued. In this context I also discuss the reason that return distributions are non-normal, why this matters to investors, and how it can be accounted for in utility functions.

The problem of models such as FSO is how to find the optimum when no analytical solution can be found. I demonstrate two different ways of solving this problem. In Chapter 6, I solve a 3 asset portfolio selection problem featuring UK equity indexes using a grid search. I identify several utility functions featuring loss aversion and prospect theory, under which FSO is a substantially better approach than the MV approach. As the equity indexes have return distributions with relatively small deviations from normality, the findings indicate much broader usefulness of FSO than has earlier been shown. The findings presented in this chapter have also been published in *The Manchester School* by Hagströmer, Anderson, Binner, Elger, and Nilsson (2008).

The grid search technique is computationally expensive and in large portfolio selection problems it becomes unfeasible. For problems with many assets more efficient algorithms are required to identify the optimal portfolio. In Chapter 7, I apply the heuristic technique differential evolution (DE) to solve FSO-type asset selection problems of 97 stocks under complex utility functions. I show that this problem is computationally feasible and that solutions retrieved with random starting values are converging to one optimum. The study constitutes the first FSO application to stock portfolio optimisation. The results indicate that when investors are loss averse, FSO improves stock portfolio performance compared to MV portfolios. The findings of this chapter have also been published in *Applied Financial Economics* by Hagströmer and Binner (2009).

Chapter 8 (Part IV) concludes the thesis by pointing out the main findings and policy implications. It also sketches a framework for how to combine the findings of the two preceding

parts into one model. Finally, future avenues of potential research are pointed out.

1.2 Contributions and Limitations

The contribution of this thesis is twofold. Firstly, it gives further insight in measurement of liquidity. A presentation of high frequency data processing is given at a high level of detail not seen elsewhere in the liquidity literature; it compares eight different liquidity measures; and it offers an extensive analysis of four different ways of deriving systematic liquidity measures, including dynamic estimation window PCA techniques that have not been applied elsewhere. This systematic liquidity study highlights many nuances on liquidity measurement that have not been known before. Accurate measurement of liquidity, both for individual stocks and on market-wide level is of high importance for investment decisions and for the handling of liquidity risk in a stock portfolio. Secondly, this thesis holds two studies on how FSO can be applied in different settings and how the optimum can be identified for these cases. The findings show that FSO is applicable to a much larger set of assets than previously shown, and that the portfolio optimum is feasible to identify also in a relatively large problem domain. This adds a useful tool for investment advice to investors whose view on risk goes beyond expected return variance.

In a context of presenting contributions of the research, it is appropriate to also point towards its limitations. All the work presented here is empirical in its nature, no new theoretical model is built. Throughout the thesis I rely on secondary data. Empirical investigations always come with important limitations with respect to data. Availability, quality, processing, and sample size are properties that always have to be questioned in a setting of empirical studies, and the implications of the findings in data sets with other properties can never be known beforehand. For the setting of FSO, I discuss the data limitations at some length in Section 7.6. For the liquidity studies, measurement quality is the subject matter itself. I do not investigate causes of liquidity or its co-variation, neither do I discuss corporate strategies for promoting stock liquidity.

Part II

Liquidity

Chapter 2

Liquidity: Definition, Measurement, and Impact on Asset Prices

In this chapter, I discuss the definition and economic intuition of liquidity, and connect that to the different liquidity measurement techniques that have been suggested in the literature. I present detailed data characteristics of eight different liquidity measures based on high-frequency data. As much of the interest in liquidity research is due to its impact on asset prices, I spend the final section of the chapter referring the literature on that topic. Overall, the purpose of this chapter is to present my understanding of liquidity and its role in financial economics, as well as introducing the data set that is the basis of my liquidity analysis in this thesis.

2.1 What is Liquidity?

Liquidity is a term spread around many fields of economics. The definition of liquidity is dependent on whether the setting is macroeconomics, corporate finance, or financial economics, and this is causing some confusion around its exact meaning.[1] According to Hicks' (1962) historical presentation of the term, it became popular with Keynes' work in the 1930's, and then was picked up in many different areas of economics. Researchers across the above-mentioned fields agree that liquidity is about the availability of liquid assets, but they have different settings in mind when defining it: the market as a whole, the balance sheet of a company, or a stock's tradability. However, in the context of this thesis the understanding of liquid-

---

[1] According to *The Economist* (February 8, 2007) liquidity is "one of the most mentioned, but least understood, concepts in the financial market debate".

ity is clear: liquidity is a concept describing the friction investors experience when trading securities.

This friction is understood as trading costs in terms of money, time, effort, and information. Broker fees and taxes constitute direct costs. If the investor wants to trade at a certain price, he may have to spend time and effort to find a counterpart for the trade, or accept a worse price to close the trade immediately with available market participants. The cost of immediate trading is typically described by the bid-ask spread. In an order-based market the bid-ask spread is the difference between the lowest available sale price and the highest available buy price in the order book; whereas in a market maker-based exchange the spread is set by the market maker in accordance with its cost structure.[2] To understand liquidity costs, it is useful to take the perspective of the market makers. Their business is to provide liquidity (immediacy) to the traders at an exchange. This intermediation activity carries, in excess of operating costs (such as fees, equipment, and staff), two types of costs:

1. *Inventory cost:* In order to provide immediacy the market maker needs to hold an inventory of the securities he is dealing with. After trading with sellers, he may also have to hold large quantities of stocks while searching for buyers. During the holding period, the market maker is facing a risk of fundamental value change. The cost of this inventory risk was first discussed by Stoll (1978) and has been modelled as part of the bid-ask spread by e.g. Amihud and Mendelson (1980) and Ho and Stoll (1981).

2. *Adverse selection cost:* When offering trading services the market maker exposes himself to a risk of trading with counterparties holding private information. Based on the reasoning of Akerlof (1970), traders holding private information will sell (buy) if they have bad (good) news, as their information is not incorporated in the market's pricing of the security. Non-informed traders have less incentive to trade, as they on average will see no arbitrage opportunity in trading at the market price. Hence, the market maker is likely to deal with traders with price driving information, which in the long term will cause him losses, called adverse selection costs. In order to stay in business, the market maker needs to balance their losses from informed trading by charging fees from uninformed traders (Bagehot 1971). How the adverse selection cost is affecting the pricing of securities has been modelled by e.g. Kyle (1985) and Glosten and Milgrom (1985).

In a market with perfect competition, where the profits of market makers are zero, these costs along with operating costs will constitute the bid-ask spread, making the market maker

---

[2]Of course, trading can also take place outside the stock exchanges (e.g. in over-the-counter markets). For such trading, liquidity costs are typically higher as it is harder to identify counter-parts and as there is less transparency in price dynamics. In this dissertation, I focus on stock exchanges, so these cases will not be discussed further. This also excludes derivatives and bond markets. Largely, similar discussions about liquidity apply to all these excluded markets, but I leave that discussion to others.

break even. As inventory risk should not affect the true value of the security, this part of the spread variation is regarded as transitory. Adverse selection cost, on the other hand, can cause permanent price changes to the extent that the order flow reveals private information that should be incorporated in security valuation. This transformation of information from private to public can be seen as an aspect of liquidity costs for an informed trader. The extent to which order flows affect the market price of a stock is often referred to as Kyle's λ, relating to his model where price changes depend on traded quantities of privately informed and uninformed traders together (1985).

Glosten and Harris (1988) summarise the modelling of how market maker costs motivate the bid-ask spread in what is called *the asymmetric information model*.[3] Following Kyle (1985), the order flow is divided between information-based trading and noise trading. The observed price at the time of trade k, p_k, is equal to the true price, m_k, adjusted for trade-specific costs that the market maker carries. They model this relation as

$$p_k = m_k + D_k C_k, \tag{2.1}$$

where C_k is a cost component containing fees and inventory costs incurred to the market maker (see Stoll 1978), but excluding costs of asymmetric information. D_k is a direction of trade dummy which is set to $+1$ when the trade categorised as buyer-initiated, and -1 when the trade is categorised as seller-initiated. $D_k C_k$ is called the *transitory spread component*, as future observed prices are not related to P_k.

The true price of the stock is assumed to be affected only by changes in information related to the company. Information reaches the market either trough a public information flow, y_k, or through information-based trading that reveals private information. Hence, according to the asymmetric information model the true price process is

$$m_k = m_{k-1} + D_k Z_k + y_k, \tag{2.2}$$

where $D_k Z_k$ is the cost component that is due to information-based trading. Z_k is called the adverse selection spread component, or the *permanent spread component*, as the dynamic specification of the true price process makes the impact of adverse selection costs on the stock price permanent.

The asymmetric information model offers a way of explaining the bid-ask spread in the perspective of a market maker who provides immediacy to investors. This perspective is useful as the immediacy isolates the temporal dimension of liquidity from the information and inventory costs. The reasoning is however not limited to market maker-based exchanges. At an order-based market, where the order book instead of the market maker constitutes the

---

[3]This model builds on the work by Glosten and Milgrom (1985), Glosten (1987), and Kyle (1985).

liquidity provision, a similar reasoning applies. The individual investors that as a collective form the order book should, on average, set their prices in the same way as a market maker.

As indicated in this discussion, the liquidity of a market has many aspects. Kyle (1985) summarise these aspects in three concepts:

1. *Tightness:* How much it costs to turn over a position in a short time. If there is no information content in the trade, this depends on C_k in the asymmetric information model.

2. *Depth:* The ability of the market to absorb quantities without large price impacts. This relates to Kyle's λ and to Z_k in the model by Glosten and Harris (1988).

3. *Resiliency:* The market's ability to quickly return to the underlying value of a security, e.g. after an uninformative price shock. This price discovery dimension of liquidity has received relatively little attention in the literature, but an analysis is available in Dong, Kempf, and Yadav (2007).

This multi-dimensionality of liquidity has triggered a voluminous literature on what is the most appropriate measurement methodology. Depending on the research problem, different researchers have found different measures more or less appropriate and no consensus on measurement has been reached. It is beyond the aim of this thesis to add to the individual liquidity measurement discussion, but as liquidity measures are needed for the subsequent analysis, I present different approaches in detail below.

2.2   Liquidity Measures

In this section, I discuss in detail eight liquidity measures that I use in the subsequent analysis. I refer the methodological discussion around each of them and look into data properties within my sample. I divide the measures in three categories; (1) spread-based measures; (2) volume-based measures; and (3) price impact measures. I estimate each liquidity measure ex post on monthly frequency. Ex post estimation allows more precise liquidity measurement than an ex ante approach. As no liquidity forecasting is done in this thesis, I choose the ex post approach, which conforms to most of the liquidity literature.

Most liquidity measures are based on the cost of immediate execution, setting the time cost to zero. This has the advantage that it yields comparable quantities of liquidity costs, but the obvious shortcoming of not quantifying the time costs. The temporal dimension is primarily important for resiliency and is not explicitly addressed by any measure referred here. Resiliency is, however, captured to some extent by the inventory cost measures. Before turning to the specific liquidity measurement discussion, I discuss the choice of data set and establish some notation.

2.2.1 Sample and Data

I study liquidity in a context of the S&P500 index constituent stocks. This index of large-cap companies traded on New York Stock Exchange (NYSE) and National Association of Securities Dealers Automated Quotations (NASDAQ), maintained by Standard & Poor's, is one of the most followed large-cap indexes in the world. Empirical liquidity literature has a strong bias towards US stock markets. This bias is likely to be due to the central role of US markets in the world financial markets, but also to the availability of high quality data. My choice of US stocks is due to both these reasons, as well as the ability to compare my results to previous studies in the US context. S&P500 stocks have their primary listing on NYSE (most), NASDAQ (about 100), and American Stock Exchange (AMEX; few). Accordingly, I restrict my data set to trades and quotes from these three markets. As NYSE and NASDAQ feature different market mechanisms, not many liquidity studies span all the S&P500 constituents. Being a continuous auction market, NYSE has a liquidity measurement problem in that trades frequently happen inside the bid-ask spread (Huang and Stoll 1996, Eleswarapu 1997), which makes spread measures less reliable as approximations of transaction cost. NASDAQ, which is a dealer market, has other liquidity measurement problems in the form of trade reporting delays (Vergote 2005) and artificially high volumes (Amihud 2002). Furthermore, rapid financial innovation in all stock markets creates time series inconsistencies in the market structures. For example, both NYSE and NASDAQ have in recent years added automated trading platforms to their systems (Arca and NASDAQ-ADF respectively). Rather than modelling different markets and periods separately, I deal with all these different trading mechanisms in one data set. This generalisation should be kept in mind when analysing the subsequent results.

I calculate stock liquidity measures on a monthly frequency. Liquidity measurement at the firm level can be divided into two schools of thought. One branch of literature is trying to minimise measurement error by looking in detail at behaviour of trade prices and volumes and order flows, usually utilising high-frequency data of order books and transactions. For assessing the impacts of liquidity over longer periods and for more markets, however, the low availability of high-frequency data has motivated research on measures that proxy liquidity using low-frequency data (daily/weekly/monthly), that are more readily available (Amihud 2002, Pastor and Stambaugh 2003, Hasbrouck 2009). For the analysis in this thesis, I use high frequency data, as this allows me to consider a wide variety of high quality liquidity measures. I use the high-frequency Trades and Quotes (TAQ) database, which holds data on all transactions and frequently updated quotes of best bid and best ask prices at major US stock markets, dating back to 1993. Data on individual trades and quotes are necessary for many of the liquidity measures considered, and TAQ is the largest accessible data source for such data. I limit my data set to stocks that were in the S&P500 index by the end of 2007.

The computational burden of dealing with recent TAQ data motivates the limitation with respect to number of stocks. The reason that I use the same stocks throughout the sample (rather than, e.g. updating the sample annually in accordance with S&P500 changes) is that I utilise co-variance matrices in the subsequent analysis that need consistent time series of the stocks considered. The maximum period that I consider is 1995–2007, but many stocks have shorter series due to name and/or ticker changes and company restructuring. The reason that I do not use data from 1993 and 1994 is that the number of order quote observations relative to trade observations is very low for these years, and many liquidity measures are dependent on matching quotes to trades.[4]

The raw data from TAQ need to be thoroughly processed before liquidity measurement can be made. The rules applied for this are described in detail in Chapter 3. In digested form, for each eligible trade k in month t I have data on transaction price ($p_{k,t}$), transaction volume ($V_{k,t}$), as well as bid and ask prices prevailing when the transaction occurred ($Bid_{k,t}$ and $Ask_{k,t}$). The exchange where the trade occurred is also recorded. For each stock i the data set also holds monthly information on lot size, shares outstanding and share types. All variables in the presentation below are firm specific (where the opposite is not noted), but the firm index i is dropped for brevity of exposition. The number of firms considered varies over time but is in general $N \leq 500$, implying $i = 1, 2, ..., N$.

2.2.2 Spread-Based Measures

The most straightforward measure of liquidity is the bid-ask spread.[5] This provides an ex ante measure of tightness in the market for a security. The spread is an intuitive measure of the cost of an immediately executed round-trip trade of a stock (for a single trade the half spread can be used). As stated in Section 2.1, it also summarises the costs of a market maker's exposure to inventory risk, adverse selection risk, and operating fees. According to Amihud and Mendelson (1989) the measure is related to characteristics associated with liquidity such as the number of investors, the transaction volume, the information availability, and the size of the firm. Using data on the best bid and offer (BBO) prices available, the monthly quoted spread ($Qspread$) is calculated for each transaction using the formula

$$Qspread_t = \frac{1}{K_t} \sum_{k=1}^{K_t} \frac{Ask_{k,t} - Bid_{k,t}}{Mid_{k,t}}, \qquad (2.3)$$

where months are indicated with the index t (which in my sample is $t = 1, 2, ..., 156$); trades in month t are indicated with the index k; and the number of trades in month t is K_t, implying

---

[4]In January 1993 the number of eligible quotes are 2.8% of the number of eligible trades. The corresponding number for January 1995 is 67.4%. Furthermore, large amounts of erroneous observations are recorded for these years, particularly 1993, see Section 3.6.

[5]The first study to relate the bid-ask spread to transaction costs was, according to Amihud and Mendelson (1986a), Demsetz (1968).

$k = 1, 2, ..., K_t$. In order to be cross-sectionally comparable it is common to calculate relative spreads by dividing the nominal spread by its midpoint ($Mid_{k,t}$), which is simply the mean of the bid and ask prices prevailing at each transaction. The monthly spread is the average spread across all trades that month.

Figure 2.1 shows the key statistics of a box plot of $Qspread$ for each month in my sample (the mean is the bold black line and the median is the bold grey line). It can clearly be seen in the figure that the $Qspread$ reacts to major events causing volatility in financial markets. Examples of events causing peaks in the spread time series are the Asian crisis (Fall 1997), Russian/Long Term Capital Management (LTCM) crisis (August–October 1998), the burst of the dot-com bubble (March 2000), the terrorist attacks on New York (September 2001), and the start of recognition of subprime loan losses by American banks (August 2007). The fall in spreads in July 1997 and January 2001 can be connected to changes in the rules of minimum tick size regulation at the NYSE. In July 1997, the minimum was changed from 1/8 to 1/16 and in January 2001 to 1/100 (referred to as decimalisation). Overall, a falling trend in average spread over the sample period is obvious, from around 0.75% in 1995 to around 0.10% by the end of 2007. All these observations are in line with previous findings in the literature. In addition, it can be noted that the mean is slightly higher than the median, which is not surprising considering that the spread by definition has a lower but no upper bound.

The bid-ask spread is a straightforward measure of transaction costs, but has a weakness in that not all trades actually occur at the spread. It is common that trades are executed within the spread (because of e.g. bettered quotes or hidden limit orders) both at the NYSE and NASDAQ (for discussion of this, see Huang and Stoll 1996, Petersen and Fialkowski 1994). Different approaches to account for this fact have been suggested in the literature. One method is the effective spread, which is the absolute difference between the actual trading price and the midpoint of the prevailing bid-ask spread. This measure is calculated as

$$Espread_t = \frac{1}{K_t} \sum_{k=1}^{K_t} \frac{|p_{k,t} - Mid_{k,t}|}{Mid_{k,t}}, \tag{2.4}$$

i.e. relative to the midpoint and averaged across all the trades in the month in the same way as done for the quoted spread.[6] Note that this is measuring the half spread, meaning that it will be half the size of $Qspread$ for trades occurring at the quoted price. It is argued by Huang and Stoll (1996), as for spreads in general, that the effective spread should cover the costs of market makers, otherwise the trades would not happen. On this basis, it is seen as more accurate than the quoted spread. A drawback with the effective spread is that it

---

[6]Other spread measures include the realised spread, which is the effective spread adjusted for adverse selection costs (Huang and Stoll 1996), and the amortised spread, which is the effective spread amortised across the holding period (Chalmers and Kadlec 1998).
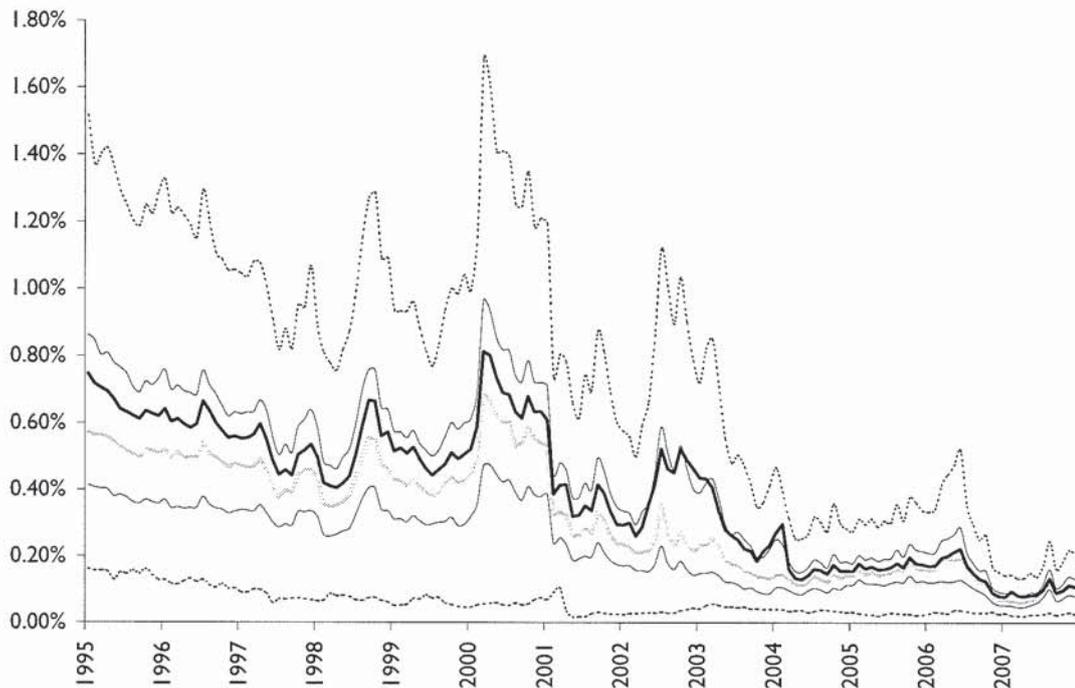
Figure 2.1: Quoted relative spread, box plot for each month

Quoted spread is calculated as in Equation 2.3, on monthly frequency for 1995–2007. Tick marks on the x-axis are placed at the January observation. The bold black line is the mean, and the bold grey line is the median. The thin solid black lines are the lower and upper quartiles, and the thin dashed black lines are the upper and lower whiskers associated with box plots. (The upper whisker is placed at the highest observation within 1.5 inter-quartile ranges (IQR) above the upper quartile. The same reasoning applies for the lower whisker, in opposite direction.)

demands data on both trades and quotes, and these are not usually reported together. For my analysis I consider both the quoted and the effective spreads. In Figure 2.2 it is seen that *Espread* follows the same general pattern as *Qspread*, with peaks at events causing financial volatility and lows at times of market-wide liquidity facilitating changes such as the decimalisation. The magnitude is, as expected, less than half of the corresponding *Qspread* observations, and there is a clear falling trend over time.

2.2.3 Volume-Based Measures

It is common (particularly in news media) to encounter the word liquidity in the meaning of traded volumes. This relates to market depth — the market's ability to absorb quantities without large price changes. Volume measures, either number of shares or the value of shares traded have also appeared in the liquidity literature, but mainly as a proxy of the bid-ask spread where data on that is not available (e.g. Brennan, Chordia, and Subrahmanyam 1998).[7] Good data availability make volume measures empirically appealing. Pure volume measures have little *theoretical* appeal as they do not measure trading friction explicitly. When divided

---

[7]According to Chordia, Roll, and Subrahmanyam (2002), "[s]tock trading volume is ... linked inextricably to liquidity" (they are referring this to Benston and Hagerman, 1974, and Stoll, 1978).
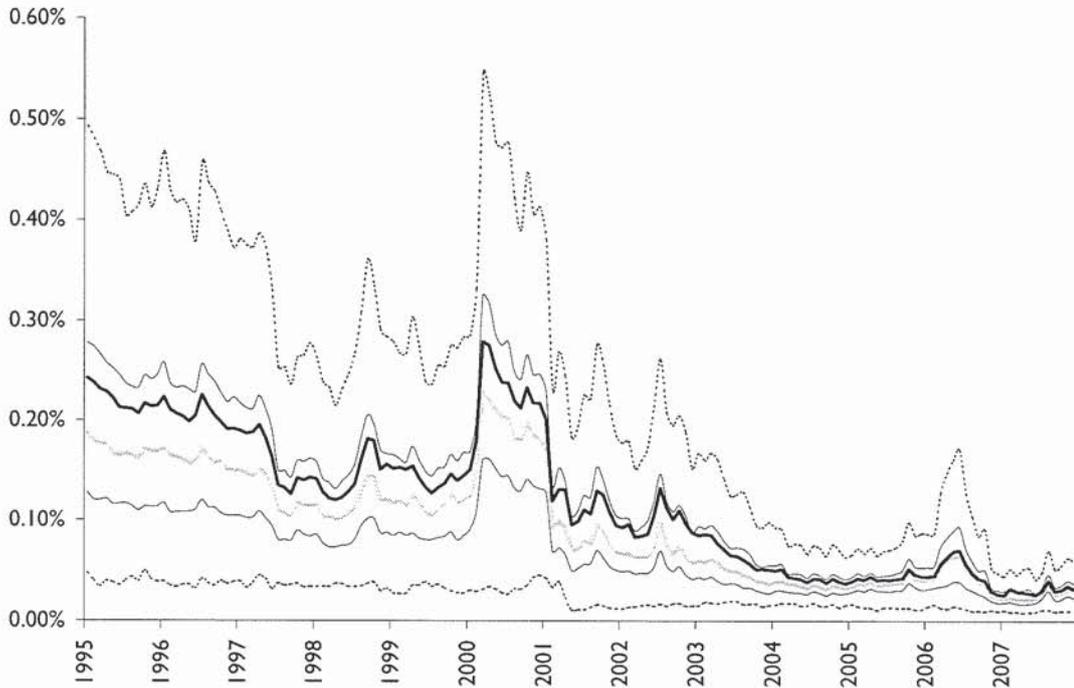
Figure 2.2: Effective relative spread, box plot for each month

Effective relative spread is calculated as in Equation 2.4, on monthly frequency for 1995–2007. The bold black line is the mean, and the bold grey line is the median. The thin solid black lines are the lower and upper quartiles, and the thin dashed black lines are the upper and lower whiskers associated with box plots. (The upper whisker is placed at the highest observation within 1.5 IQR above the upper quartile. The same reasoning applies for the lower whisker, in opposite direction).

by the number of shares outstanding, however, it describes the turnover of the firm; and the inverse of that ratio is a measure of the average holding period of the investors of a stock. This has theoretical appeal as a proxy of liquidity, as stocks with low liquidity are unattractive to trade frequently, and hence attract investors with long investment horizon (Amihud and Mendelson 1986a). *Turnover*, used by e.g. Datar, Naik, and Radcliffe (1998), can be calculated by the formula,

$$Turnover_t = \frac{1}{SO_t} \sum_{k=1}^{K_t} V_{k,t}, \tag{2.5}$$

where SO_t is the monthly number of shares outstanding for one stock. Unfortunately, SO_t is not reported for NASDAQ stocks in TAQ, so my investigation of *Turnover* is limited to NYSE and AMEX.

Turnover is bounded downwards (at zero) but has no limit upwards, making the mean higher than the median (see Figure 2.3). *Turnover* is increasing in liquidity, so the increasing trend seen over the sample considered corresponds to the falling trend in liquidity seen in spreads. However, *Turnover* has peaks at the major events in my sample, indicating rising rather than falling liquidity (that was seen for spreads). This highlights that *Turnover* measures a different aspect of liquidity — it is a measure of depth rather than tightness.
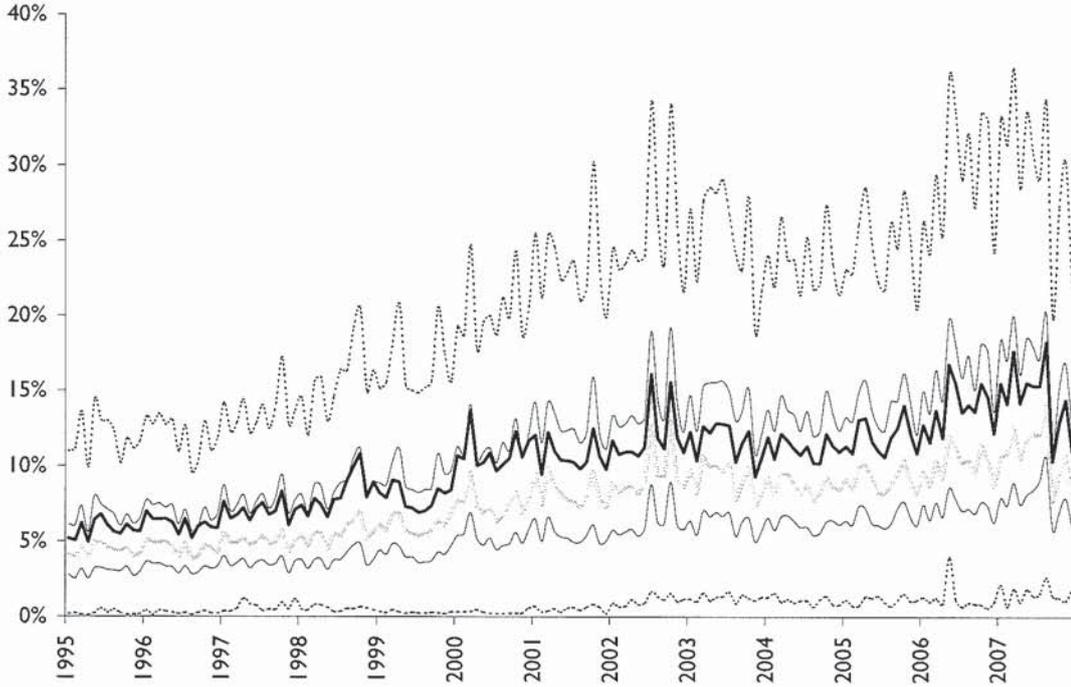
Figure 2.3: Turnover, box plot for each month

Turnover is calculated as in Equation 2.5, on monthly frequency for 1995–2007. The bold black line is the mean, and the bold grey line is the median. The thin solid black lines are the lower and upper quartiles, and the thin dashed black lines are the upper and lower whiskers associated with box plots. (The upper whisker is placed at the highest observation within 1.5 IQR above the upper quartile. The same reasoning applies for the lower whisker, in opposite direction).

It differs from many other depth measures though, in that it does not distinguish between buyer- and seller-initiated volumes. This is done in the price impact measures presented next.

2.2.4 Price Impact Measures

Building on the models by Kyle (1985), Glosten and Milgrom (1985), and Glosten (1987) several ways of measuring liquidity in terms of price impact exist. This relates to depth of a market — how well a market can absorb trade quantities without major price changes. As will be seen below, however, these measures capture elements of market tightness as well. The asymmetric information model by Glosten and Harris (1988), referred in Section 2.1, is one way to estimate the price impact of trading.

By taking the first difference of Equation 2.1, i.e. $\Delta p_k = p_k - p_{k-1}$, and inserting Equation 2.2, Glosten and Harris (1988) retrieve

$$\Delta p_k = D_k Z_k + D_k C_k - D_{k-1} C_{k-1} + y_k, \tag{2.6}$$

i.e. the observed price change is related to the change in signed inventory costs ($D_k C_k - D_{k-1} C_{k-1}$) and the signed information cost ($D_k Z_k$). Both spread components are modelled to be linear functions of volume of trade k, V_k, so the transitory (inventory) and permanent

(information-based) spread components are expressed

$$C_k = \tilde{\Psi} + \tilde{\lambda} V_k, \qquad (2.7)$$

and

$$Z_k = \Psi + \lambda V_k, \qquad (2.8)$$

where Ψ and λ are coefficients for fixed and variable costs and "$\sim$" distinguishes transitory from permanent costs. Inserting these expressions in Equation 2.2 yields

$$\Delta p_k = \Psi D_k + \lambda D_k V_k + \tilde{\Psi} \Delta D_k + \tilde{\lambda} \Delta(D_k V_k) + y_k, \qquad (2.9)$$

where y_k is again the public information flow at the time of trade k, and Δ is the first difference operator.

In estimating this relationship, Glosten and Harris (1988) find that $\tilde{\lambda} = \Psi = 0$, implying that inventory cost is the effect of changes in trade direction, $D_k - D_{k-1}$, and that the adverse selection cost is the effect of signed trading volume, $D_k V_k$. Hasbrouck (1988, 1991) points out that adverse selection costs can only appear in response to *unexpected* order flows, which has been implemented in the model framework by Foster and Viswanathan (1993) and Brennan and Subrahmanyam (1996) (by using the trade innovations from a model with lagged order flows and lagged price changes). Sadka (2006) incorporates this in the Glosten and Harris's (1988) model by writing the adverse selection cost function as

$$Z_k = \Psi \left(1 - \frac{E_{k-1}[D_k]}{D_k} \right) + \lambda \left(V_k - \frac{E_{k-1}[D_k V_k]}{D_k} \right), \qquad (2.10)$$

where E_{k-1} denotes conditional expectations. Inserting this in the framework of Glosten and Harris (1988) yields an expression of changes in observed stock prices showing that adverse selection costs are incurred from unexpected trade flow changes:

$$\Delta p_k = \Psi(D_k - E_{k-1}[D_k]) + \lambda(D_k V_k - E_{k-1}[D_k V_k]) + \overline{\Psi} \Delta D_k + \overline{\lambda} \Delta(D_k V_k) + y_k. \quad (2.11)$$

Sadka (2006) estimates the unexpected signed order flow as the residuals of an AR(5) model of observed signed order flows and denotes it $\epsilon_{\lambda,k}$. The direction of trade dummy D_k is set +1 (-1) when the transaction price is higher (lower) than the prevailing spread midpoint, implying that the trade is buyer-initiated (seller-initiated) (following the algorithm by Lee and Ready 1991). By assuming a normal distribution of $\epsilon_{\lambda,k}$ with constant variance, Sadka (2006) is able to derive the unexpected order flow *direction*, which he denotes $\epsilon_{\Psi,k}$, using the cumulative density function. Letting the estimated time series of $\epsilon_{\lambda,k}$ and $\epsilon_{\Psi,k}$ replace $D_k V_k - E_{k-1}[D_k V_k]$ and $D_k - E_{k-1}[D_k]$ respectively, he is able to analyse the following

relationship using ordinary least squares (OLS):

$$\Delta p_{k,t} = \alpha_t + \Psi_t \epsilon_{\Psi,k,t} + \lambda_t \epsilon_{\lambda,k,t} + \overline{\Psi}_t \Delta D_{k,t} + \overline{\lambda}_t \Delta(D_{k,t}V_{k,t}) + y_{k,t}, \qquad (2.12)$$

where α_t is an intercept and the index t indicates that the relationship is estimated on a monthly basis. It is important to note that trades typically take place on irregular time intervals, e.g. around the release of news intense trading can take place. Following Sadka (2006), I treat the observations as if they happened on regular time intervals. An alternative approach using durations is presented by Engle and Russell (1998). This could be an interesting way to improve measurement in general, and possibly the resiliency dimension in particular. As it is not the aim of this chapter to improve individual stock liquidity measurement techniques, I do not pursue this here, but I regard it as an interesting future research topic. In the exposition below, I refer to the four coefficients estimated here as *price impact coefficients*.

I estimate Equation 2.12 for each firm and each month, retrieving estimates of the coefficients in the inventory cost (and other market making cost) function $[\tilde{\Psi}$ and $\tilde{\lambda}]$; and of the coefficients in the adverse selection cost function $[\Psi$ and $\lambda]$. Following Brennan and Subrahmanyam (1996) and Sadka (2006), I scale the coefficients by the end of month stock price, in order to reflect relative costs rather than absolute. The distributional properties of each of the coefficients are presented in Figure 2.4. These coefficients are falling in liquidity, so falling trends over the sample periods are expected, and this holds for every coefficient except $\tilde{\lambda}$. This coefficient is also negative on average throughout the sample, which may seem peculiar, but this is in line with the findings of Sadka (2006). Sadka interprets this as that total costs are increasing with trade quantity, but that the information-based costs take a larger proportion of the costs as volumes increase. This hypothesis is supported in my results as the two variable cost series have a clear negative correlation, indicating that they counteract each other. Sadka (2006) also emphasises that the total inventory cost is positive (up to a certain quantity).

The transitory non-information fixed cost ($\tilde{\Psi}$; Panel C), as many other liquidity measures, features peaks of low liquidity in the months previously identified as volatile in the stock markets (October 1997, August 1998, February 2000, September 2001, August 2007), and it is falling sharply in the two months of market-wide decreasing tick sizes (July 1997 and February 2001). The permanent information variable cost (λ) features peaks in roughly the same months, but does not show much reaction to the tick size changes, which is in line with theory as these have nothing to do with informed trading. The permanent fixed cost (Ψ) and the transitory variable costs ($\tilde{\lambda}$) have less emphasised peaks.

In Figure 2.5 some diagnostics from the regression analysis are given. Overall, it shows increasing significance of each coefficient over the period covered. λ and $\tilde{\Psi}$ have high rejection rates of the t test throughout the sample, meaning that they are consistently significantly

Panel A: Permanent Fixed Costs (Ψ)

Panel B: Permanent Variable Costs ($\lambda \times 10^6$)

Figure 2.4: Price impact coefficients, box plots for each month (see caption on p.35)

Panel C: Transitory Fixed Costs ($\tilde{\Psi}$)

Panel D: Transitory Variable Costs ($\tilde{\lambda} \times 10^6$)
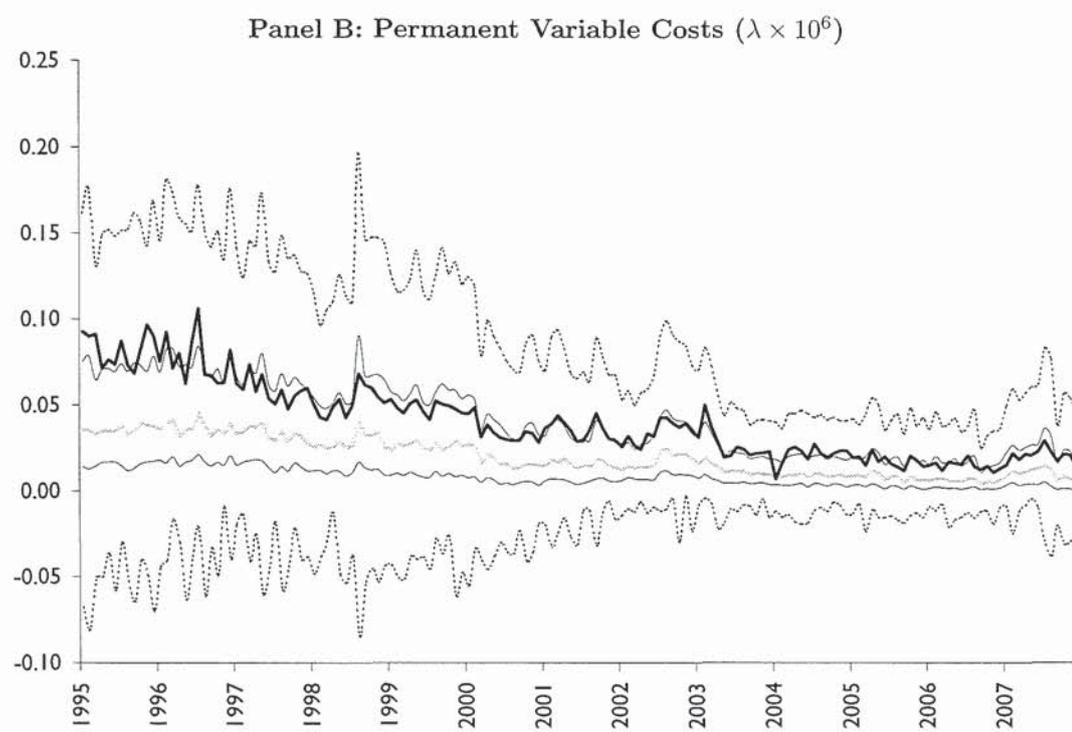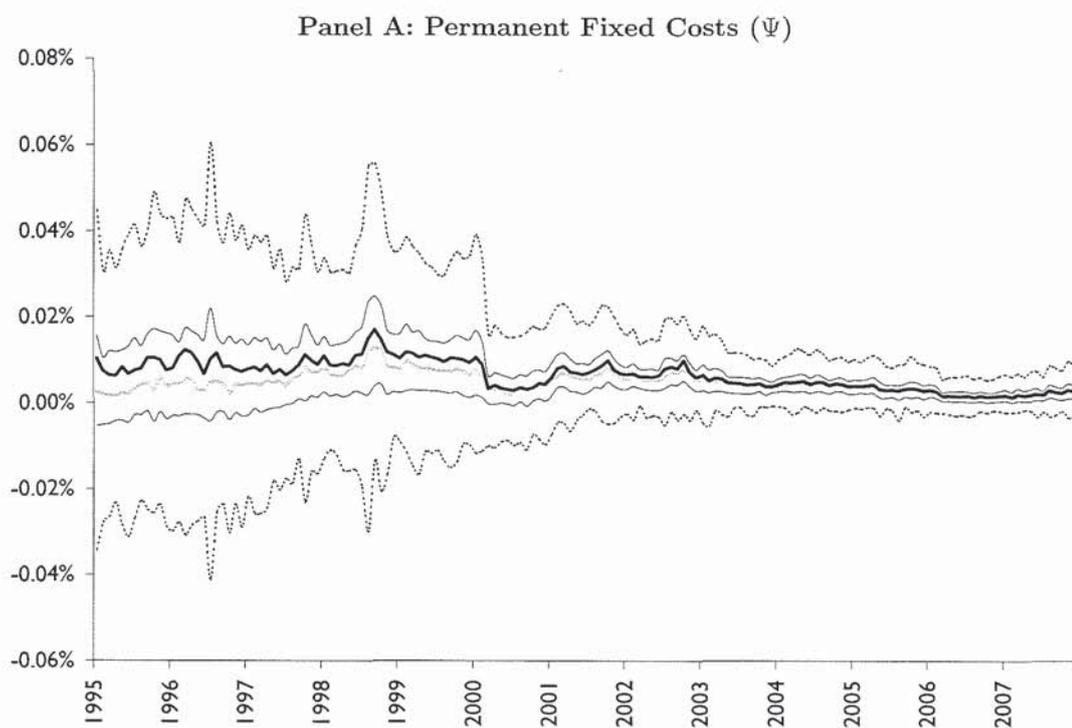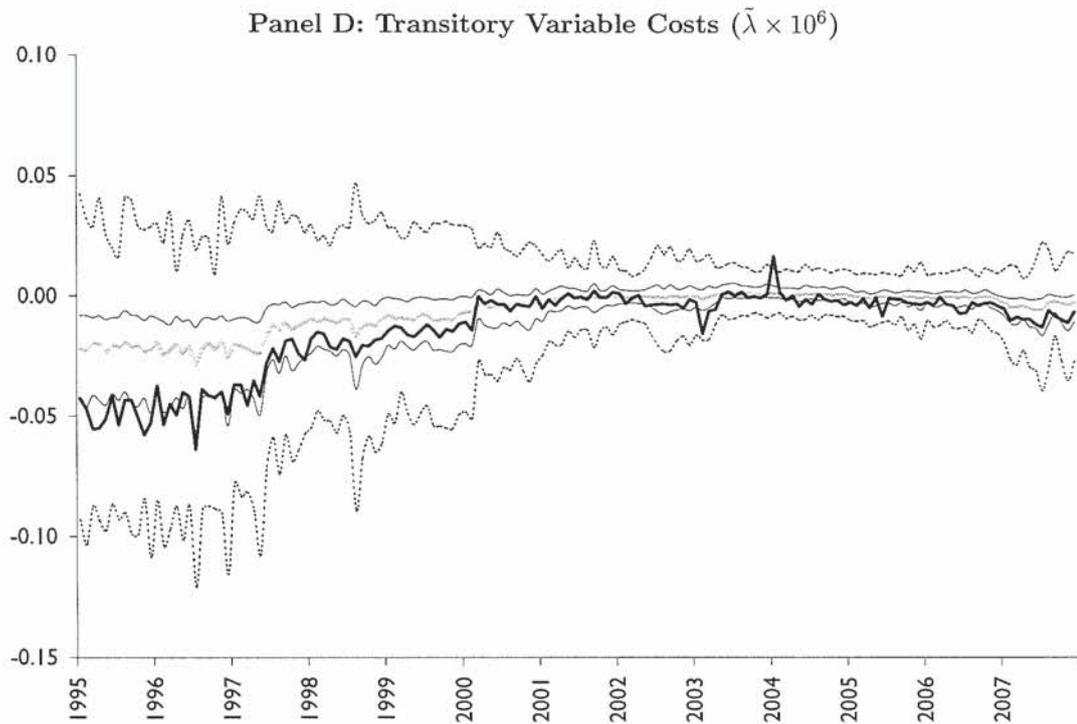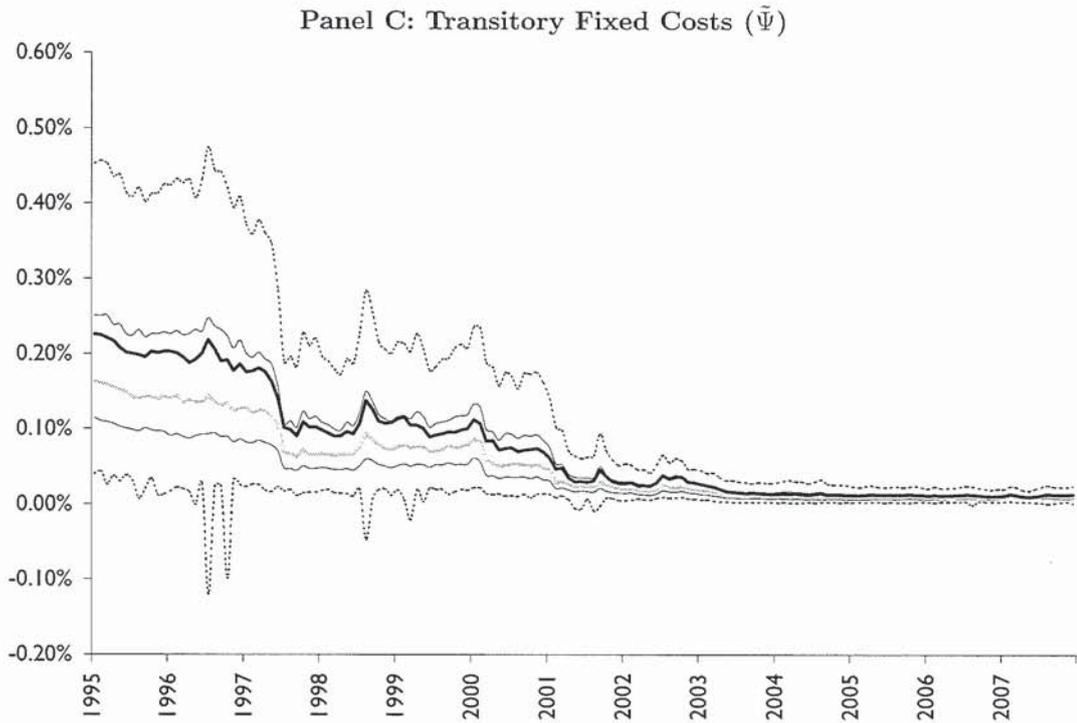
Figure 2.4: Price impact coefficients, box plots for each month (see caption on p.35)

Figure 2.4: Price impact coefficients, box plots for each month

Properties of the coefficients from the regression analysis of Equation 2.12 are presented in Panels A–D. The estimation was run for each month over the years 1995–2007. The bold black line is the mean, and the bold grey line is the median. The thin solid black lines are the lower and upper quartiles, and the thin dashed black lines are the upper and lower whiskers associated with box plots. (The upper whisker is placed at the highest observation within 1.5 IQR above the upper quartile. The same reasoning applies for the lower whisker, in opposite direction).
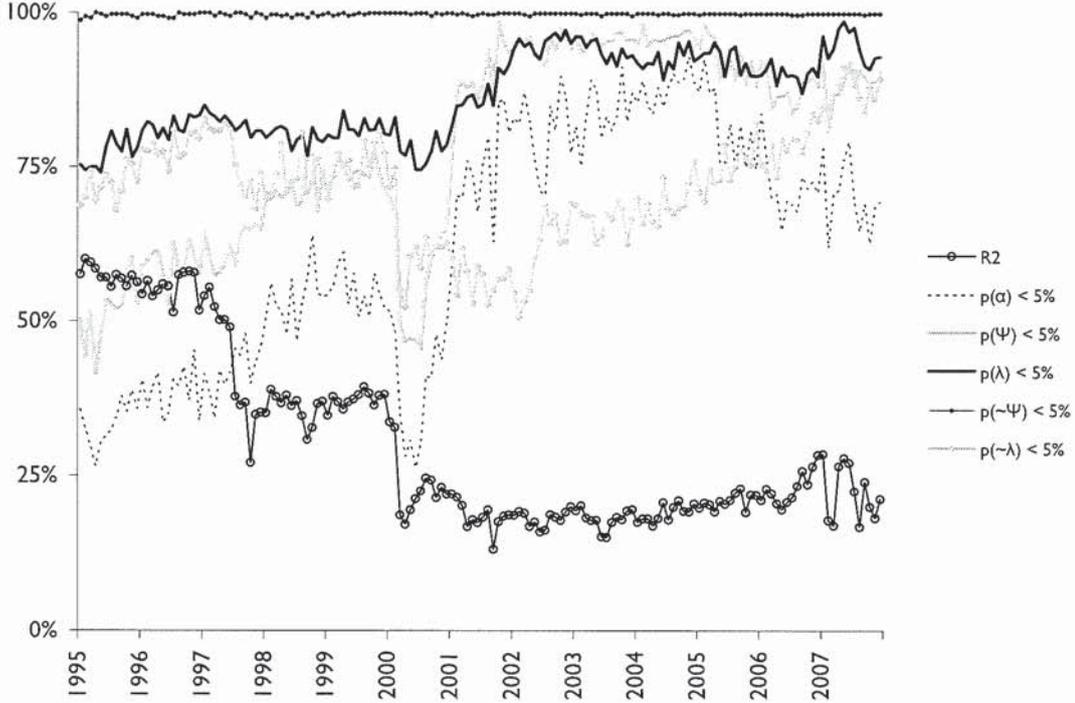
Figure 2.5: Price impact regression estimation diagnostics

The figure shows diagnostics from the regression analysis of Equation 2.12. The number of rejections of the null hypothesis of the t test (stating that the coefficient in question is equal to zero) divided by the number stocks investigated is given for each coefficient and each month. A confidence level of 95% is applied. In addition, the coefficient of determination, R^2, of the regression is given for each month, averaged across stocks. Note that in the legend $\bar{\lambda}$ and $\bar{\Psi}$ are written $\sim \lambda$ and $\sim \Psi$.

different from zero. The other two coefficients ($\tilde{\lambda}$ and Ψ), as well as the regression intercept α, show increasing rejection rates of the t test null hypothesis over time. The increasing significance of the coefficients over the period considered can be explained in several ways. Firstly, the amount of data available on bid-ask spreads has increased immensely, which is due to that reporting has been made automatic rather than changes in the market conditions. This makes the measurement of price impact more exact, as it is possible to match trades to market conditions prevailing when the trade happened. Secondly, the number of trades has increased substantially. This increases the sample size and hence the accuracy of the price impact regression. Thirdly, the decreased tick size that has been discussed above removes noise from the observed price process. This is likely to make the deviations from the true price process be more precisely described by inventory costs (see Equation 2.2). In fact, a dramatic increase in significance for the inventory cost coefficients (Ψ and λ) as well as the

35

intercept (α) is seen in January 2001, when NYSE moved to decimalisation. The intercept is positive on average and significant in around 75% of the cases after decimalisation. This can be due to the remaining discreteness of price movements.

The finding of Glosten and Harris (1988) that $\Psi = \tilde{\lambda} = 0$ is not supported. The t test null hypothesis of each of the coefficients is rejected in more than 50% of the cases for almost all months throughout the sample. Towards the end of the sample the rejection rate is higher than 75%. Figure 2.5 also shows the explanatory power of the model, and this is falling during the period 1995–2000, and then stabilises around 20% on average. This was not observed by Sadka (2006). The stepwise decline indicates that this can be related to tick size reforms. I do not address this issue further here, but regard it as interesting problem for future research.

For many applications, in particular asset pricing studies where long time series are needed, the transaction data needed to estimate price impact measures are unavailable. To circumvent this problem, Amihud (2002) suggests a liquidity measure based on low-frequency data with the intention to proxy the λ of Kyle's model (1985). He derives the price impact by dividing absolute daily returns with the daily dollar volume trade. Aggregated to a monthly frequency and adjusted for market capitalisation[8], this measure takes the following form:

$$ILLIQ_t = \frac{mcap_{ref}}{mcap_t} \sum_{j=1}^{J_t} \frac{|r_{j,t}|}{dvol_{j,t}}, \tag{2.13}$$

where $r_{j,t}$ and $dvol_{j,t}$ are the return and dollar volume on day j in month t, $j = 1, 2, ..., J_t$, and J_t is the number of trading days in month t. The Amihud measure is denoted $ILLIQ$, which demonstrates the fact that it is falling in liquidity and hence rather measures *illiquidity*. Key features of $ILLIQ$ are that it is an economically intuitive proxy of liquidity and that it is utilising data available for most markets and long time series. This has made Amihud's (2002) measure a popular choice in the liquidity asset pricing literature. Figure 2.6 illustrates the distributional properties of $ILLIQ$ within the sample considered. The figure shows the price impact in relative terms (%) in response of average daily trading volume of $1 million. What is striking here is that the mean series is clearly influenced by outliers (as it is higher than the upper quartile for most of the sample). Focussing on the median, there is a falling tendency over time and some of the major volatility periods are marked by peaks (October 1997, October 1998, March 2000, August 2007), which complies with other liquidity measures. Furthermore, $ILLIQ$ falls steeply at the time around the decimalisation of the NYSE in January 2001.

---

[8]As suggested by Acharya and Pedersen (2005), on the basis that dollar volume tends to grow over time, indicating inflation rather than increasing liquidity. Note that *market* capitalisation (*mcap* is used, so the measure is corrected by the same factor throughout the cross-section of firms. I use January 1993 as reference date.
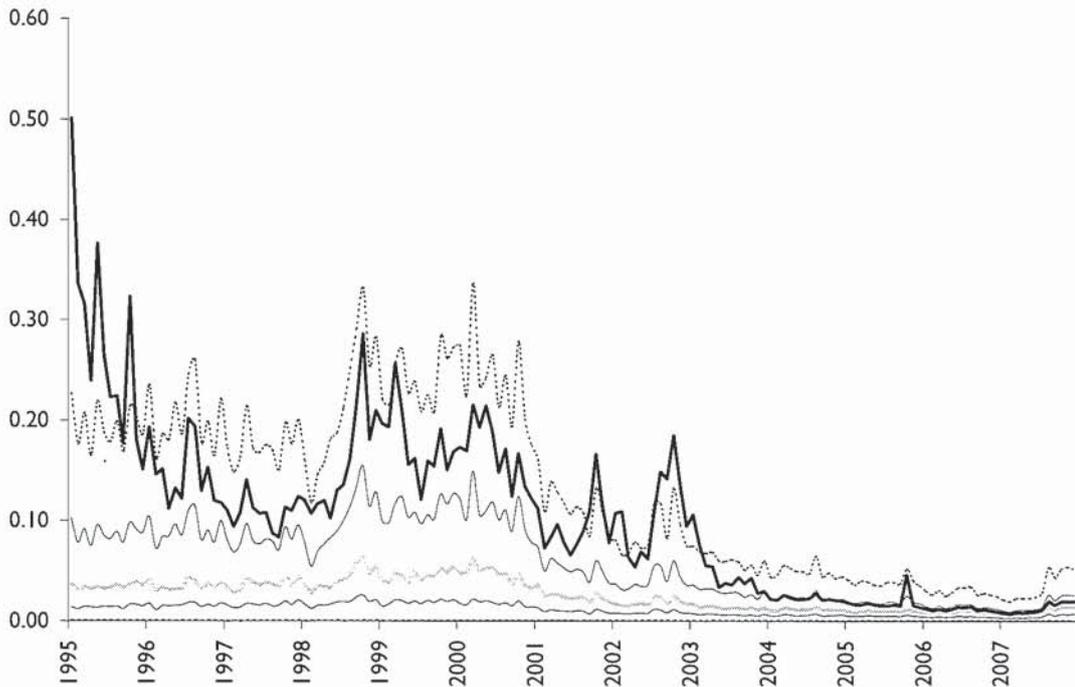
Figure 2.6: Amihud's illiquidity, box plot for each month

ILLIQ is calculated as in Equation 2.13, on monthly frequency for 1995–2007. Displayed values give the price impact in percentage terms of a \$1 million average daily volume. The bold black line is the mean, and the bold grey line is the median. The thin solid black lines are the lower and upper quartiles, and the thin dashed black lines are the upper and lower whiskers associated with box plots. (The upper whisker is placed at the highest observation within 1.5 IQR above the upper quartile. The same reasoning applies for the lower whisker, in opposite direction).

2.3 Descriptive Statistics of Liquidity Measures

In the exposition above, I have presented eight different liquidity measures that I will consider in the subsequent analysis. Henceforth, I denote liquidity L, and the liquidity measures are distinguished by superscript $l = 1, 2, ..., 8$ referring to order of liquidity measures as shown in Table 2.1, where basic statistical properties of the monthly liquidity measures are also presented. I also provide correlations between liquidity measures, see Table 2.2.

The first four liquidity measures are by definition positive. Liquidity measures 5–8 are based on the price impact regression, and the distributions of these regression coefficient series all contain negative values. This is to be expected when a large number of regressions are run. Overall, the values of the price impact coefficients are smaller than what Sadka (2006) found, which reflects the fact that liquidity in general is known to be higher in my sample than in his, both with respect to firms and with respect to time (he investigated 1983-2001, with smaller firms on average). My results correspond well with those of Sadka (2006) in that the transitory fixed costs are much higher than the permanent fixed costs (by a factor of 10); and that the variable costs are dominated by the permanent adverse selection component (which is 0.37 on average, as compared to -0.01 for the transitory variable cost).

		Mean	Median	Std.Dev	Max	Min
L_t^1	$QSpread$ (%)	0.37	0.26	0.41	9.99	0.02
L_t^2	$ESpread$ (%)	0.11	0.07	0.14	3.50	0.01
L_t^3	$Turnover$ (%)	10.49	7.70	11.41	556.58	0.122
L_t^4	$ILLIQ \times 10^6$	0.10	0.02	0.70	47.77	0.00
L_t^5	Ψ (%)	0.006	0.004	0.014	0.779	-0.591
L_t^6	$\lambda \times 10^6$	0.37	0.16	1.38	63.98	-53.42
L_t^7	$\tilde{\Psi}$ (%)	0.06	0.03	0.12	3.96	-0.58
L_t^8	$\tilde{\lambda} \times 10^6$	-0.12	-0.02	0.84	71.89	-46.58

Table 2.1: Descriptive statistics of liquidity measures

The liquidity measures are calculated monthly for 1995–2007. Ψ and λ are coefficients in the adverse selection cost function, with permanent influence on stock prices. $\tilde{\Psi}$ and $\tilde{\lambda}$ are coefficient in the inventory cost function and have a transitory impact on stock prices. See Equation 2.11 for details.

As discussed above, the two variable costs show a very high (negative) correlation coefficient. *Turnover* and the transitory variable cost are in general negatively correlated with other liquidity measures, showing that they are growing in liquidity. All the other measures are measuring illiquidity. As expected, a very high correlation is seen between the two bid-ask spread measures. The spreads are also highly correlated with *ILLIQ* and the transitory fixed costs. *Turnover* has in general low correlation to other measures, which can be related to the observation that it shows high liquidity in times of market turmoil or that it lacks observations on NASDAQ stocks.

		$QSpread$	$ESpread$	$Turnover$	$ILLIQ$	Ψ	λ	$\tilde{\Psi}$	$\tilde{\lambda}$
L_t^1	$QSpread$	1							
L_t^2	$ESpread$	0.74	1						
L_t^3	$Turnover$	-0.24	-0.16	1					
L_t^4	$ILLIQ$	0.58	0.45	-0.36	1				
L_t^5	Ψ	0.06	0.04	-0.1	0.19	1			
L_t^6	λ	0.19	0.13	-0.26	0.28	-0.13	1		
L_t^7	$\tilde{\Psi}$	0.52	0.53	-0.21	0.44	0.01	0.23	1	
L_t^8	$\tilde{\lambda}$	-0.18	-0.14	0.22	-0.22	0.11	-0.72	-0.25	1

Table 2.2: Average correlation between liquidity measures

This table shows the correlation between liquidity measures, averaged across stocks. The sample covers the stocks that constituted the S&P500 index on December 31, 2007, over the period 1995–2007. Monthly observations are used. Correlations are calculated on basis of those months in the sample where there are observations on both liquidity measures. Ψ and λ are coefficients in the adverse selection cost function, with permanent influence on stock prices. $\tilde{\Psi}$ and $\tilde{\lambda}$ are coefficient in the inventory cost function and have a transitory impact on stock prices. See Equation 2.11 for details.

2.4 Liquidity Impact on Asset Prices

One of the main reasons for research on stock liquidity is that it influences stock prices. In this section I first refer to the theory and evidence of individual stock liquidity impact on individual stock prices briefly, as this is an important foundation to understand liquidity effects on asset pricing. I then turn to the literature on market-wide liquidity and its impact

on individual stock prices. As this is the focus of my analysis in Chapter 4, I cover this literature in greater detail.

2.4.1 Individual Stock Liquidity and Stock Prices

The literature on liquidity impacts on asset pricing builds on a seminal paper by Amihud and Mendelson (1986a). They formulate a model describing the relation between illiquidity, in the form of relative bid-ask spreads, and gross stock returns. This model builds on two hypotheses:

1. Investors will in general prefer a stock with low trading costs to an identical stock with high trading costs.

2. As investors amortise the trading costs over their holding period, stocks with high (low) trading costs will in equilibrium be allocated to investors with long (short) investment horizons.

These two intuitive hypotheses together imply a positive and concave relationship between gross returns and the relative bid-ask spread. The relation is positive because the spread is discounted in the price, making returns higher when spreads are high, as the payoffs are unaffected by the spread. Concavity follows from the clientele effect described in the second hypothesis. As illiquid stocks are bought by long-term investors, the spread is amortised over a longer period and will hence be discounted less in relative terms. Liquid stocks are bought by short-term investors that discount trading costs over a shorter period. Amihud and Mendelson (1986a) find empirical support for this model in a setting of annual observations on NYSE stocks.

A related reasoning is given in a model by Constantinides (1986) dealing with portfolio choice in a multi-period setting. He argues that there is a trade-off between the utility of rebalancing the portfolio to its optimal allocations and the transaction cost, which has the implication that assets with low trading costs will be traded more frequently than assets with high trading costs.

A substantial empirical literature has focussed on assessing these hypotheses, using different return adjustment approaches, different liquidity measures, different market mechanisms, and different data frequencies. This includes Amihud and Mendelson (1989) who find support for their model while accounting for return volatility; Brennan, Chordia, and Subrahmanyam (1998) who use dollar trading volume and Datar, Naik, and Radcliffe (1998) who use *Turnover* as liquidity measure, both finding support for the illiquidity relationship to returns. Brennan and Subrahmanyam (1996) use the framework by Glosten and Harris (1988) and find support for the positive return relationship for both fixed and variable costs, but they do not find support for concavity in this relationship. Jacoby, Fowler, and Gottesman (2000) argue

that concavity of the relationship is likely in a highly liquid market (such as the one studied by Amihud and Mendelson 1986a), but in the limit (when the expected relative spread approaches 100%), a level effect will dominate and cause convexity.

Eleswarapu and Reinganum (1993) find that the liquidity effect is only present in January, giving rise to a seasonality discussion followed by Eleswarapu (1997) and Amihud (2002) who both find support for a liquidity effect also in non-January months. Hasbrouck (2009), however, using a liquidity measure based on a Gibbs sampler, does not find support for the liquidity effect when accounting for the January effect. A detailed survey of the literature on liquidity's influence on asset pricing in general is available in Amihud, Pedersen, and Mendelson (2005). As my main interests are the measurement and effects of market-wide liquidity, I now turn the focus to that topic.

2.4.2 Systematic Liquidity, Liquidity Risk, and Asset Pricing

As illustrated in Section 2.2, liquidity varies substantially over time. In accordance with the findings on liquidity impacts on prices, this time variation should be important for asset pricing too, forming the concept liquidity risk. Furthermore, liquidity co-varies in the cross-section, implying that it is driven by some underlying market factor structure. Several reasons for cross-sectional co-variance have been suggested. These include the following categories (similar arguments can be used to explain time-series variations in liquidity):

- Trading activity, market volatility, and funding conditions can affect the inventory cost of market makers on a market-wide scale (Brunnermeier and Pedersen 2009, Chordia, Roll, and Subrahmanyam 2000);

- Market makers and institutional investors that deal with many stocks simultaneously can cause market-wide liquidity movements (Chordia, Roll, and Subrahmanyam 2000, Coughenour and Saad 2004, Kamara, Lou, and Sadka 2008);

- Seasonality effects, such as the January effect discussed above (Eleswarapu 1997), are a source of co-variation;

- Various macroeconomic indicators can lead market movements in liquidity (Chordia, Roll, and Subrahmanyam 2001, Watanabe 2004).

The co-variation of stock liquidity across firms is called commonality in liquidity, and the underlying market factors causing this are referred to as systematic liquidity factors. To the extent that commonality in liquidity is strong and persistent over time, it should be treated as a risk factor in asset pricing. Pioneering articles that established the existence of systematic liquidity movements were Chordia, Roll, and Subrahmanyam (2000) and Huberman and Halka (2001). Their results were reinforced by Chordia, Roll, and Subrahmanyam (2001)

who used a larger set of firms over a much longer time span. In a contemporary article, Hasbrouck and Seppi (2001) studied commonality in liquidity using PCA without finding evidence thereof. Numerous studies have followed, studying commonality in liquidity with respect to different markets, time periods, and liquidity measures. Systematic liquidity factors are typically derived either as a market- or value-weighted average (following Chordia, Roll, and Subrahmanyam 2000), or through a factor structure investigated using PCA (used by Hasbrouck and Seppi 2001). Details on these methods are given in Chapter 4 — I here focus on the theory and findings on systematic liquidity factors.

While the above studies investigate the US stock exchanges, commonality has also been found in the London Stock Exchange (Galariotis, Giouvris, and Lane 2007), Athens Stock Exchange (Giouvris and Galariotis 2008), the Frankfurt Stock Exchange (Xetra) (Kempf and Mayston 2008, Kopp, Hütl, Loistl, and Prix 2008), the Swiss Stock Exchange (Bauer 2004), and Oslo Stock Exchange (Chollete, Næs, and Skjeltorp 2007). Karolyi, Lee, and van Dijk (2007) provide a study of commonality in liquidity in 40 different countries. Several studies have looked further into commonality on NYSE, AMEX, and NASDAQ (Amihud 2002, Chen 2007, Chollete, Næs, and Skjeltorp 2008, Chordia, Huh, and Subrahmanyam 2009, Goyenko, Holden, and Trzcinka 2009, Kamara, Lou, and Sadka 2008, Korajczyk and Sadka 2008, Sadka 2006), reinforcing the evidence of systematic liquidity factors.

In order to study asset-specific liquidity shocks, liquidity data can be adjusted for the effects that are known to cause market-wide liquidity changes. Applying the adjustment methodology of Gallant, Rossi, and Tauchen (1992), Chordia, Sarkar, and Subrahmanyam (2005) show how liquidity can be detrended and adjusted (using dummy variables) for financial crisis, tick size changes, seasonality (with respect to weekday and month), holidays, and macroeconomic shocks. Such adjustment puts the emphasis on the idiosyncratic liquidity shocks rather than the common shocks. I do not adjust the liquidity data presented in this chapter, as this would remove the commonality that I seek to analyse.

The degree of commonality depends on the liquidity measure applied. In the extensive study by Korajczyk and Sadka (2008), commonality is compared across the same liquidity measures as those presented in Section 2.2, showing strong commonality in quoted and effective spreads, and fixed inventory costs ($\tilde{\Psi}$). Commonality is also detected in *Turnover* and in fixed information-based costs (Ψ), whereas the variable costs of inventory and asymmetric information show little commonality. They also find persistence in the systematic factors for all the eight liquidity measures, which is in line with other liquidity commonality studies and a prerequisite for systematic liquidity factors' importance in asset pricing.

The price implications of liquidity risk and systematic liquidity has been modelled by Acharya and Pedersen (2005). They adjust CAPM (Sharpe 1964, Lintner 1965, Mossin 1966) to study returns net of liquidity costs ($r_i - L_i$). This amendment implies that expected gross

41

return is a function of the market risk as well as three different types of liquidity risk. These can be described as follows:

- The first type of liquidity risk is the co-variance between stock liquidity and market liquidity, which relates to the commonality in liquidity discussion. Stocks that become more illiquid when markets in general are illiquid are unattractive to investors, and a positive co-variance is hence compensated by a return premium. This type of liquidity risk has been discussed by Chordia, Roll, and Subrahmanyam (2000) and Kamara, Lou, and Sadka (2008). The latter shows that this type of liquidity risk has increased over the years 1963–2005.

- The second type of liquidity risk is the co-variance between returns on individual stocks and market liquidity cost. Stocks that show higher returns when markets are illiquid are attractive to investors motivating the negative sign on this co-variance. This relationship has been shown to hold empirically in several market liquidity studies, including Chen (2007), Gibson and Mougeot (2004), Korajczyk and Sadka (2008), Pastor and Stambaugh (2003), and Sadka (2006).

- The third type of liquidity risk describes how liquidity costs of individual stocks react to movements in market returns. In market downturns investors often need to liquidate positions. Stocks whose liquidity costs are low during downturns are hence relatively highly valued, implying a negative sign on this kind of risk. The observations in Section 2.2 indicate a tendency of negative co-variance between market return and individual stock liquidity, as many measures demonstrate spikes in liquidity costs during sharp falls in market returns, e.g. the Russian/LTCM crisis and the burst of the dot-com bubble. This type of risk has not received much attention in the literature, but the empirical investigation of Acharya and Pedersen (2005) find it to be the most important of the three liquidity risk types, and they find support in the work of Lynch and Tan (2004). Other empirical work on this type of risk includes Chordia, Roll, and Subrahmanyam (2001) and Coughenour and Saad (2004) that find an asymmetry in spread responses to up- and down-markets.

Many of the risk assessment studies referred above are performed in factor models as specified by Fama and MacBeth (1973) and Fama and French (1993), in some cases including the momentum factor of Carhart (1997). In this context the momentum anomaly (Jegadeesh and Titman 1993, the possibility to make money by selling stocks in a falling trend, and vice versa) has showed none or very little significance (Chen 2007, Chordia, Goyal, Sadka, Sadka, and Shivakumar 2009, Sadka 2006), indicating that this anomaly may be explained by liquidity costs.

2.5 Way Forward: Liquidity Measurement Issues

Clearly, liquidity has some influence on stock pricing. Measurement of this pricing effect and of liquidity risk is however complicated by the measurement problem of liquidity itself. The purpose of this chapter has been to give some insight in that measurement problem and different ways of dealing with it. It is, however, not straightforward to calculate liquidity directly from the empirical data even when the measures are well defined. Several details on data processing remain to be addressed to achieve good measures. In the next chapter, I discuss in full detail how the TAQ database can be processed to a form suitable for liquidity measurement in accordance with Section 2.2.

Another measurement issue, not discussed above, is how to best derive the systematic liquidity factor(s), i.e. the underlying mechanisms driving market-wide changes in liquidity. This is the focus of Chapter 4. There I present different approaches to systematic liquidity measurement in detail, and suggest improvements on these. I then evaluate the different methods' ability to capture commonality in liquidity as well as ability to explain stock returns.

Chapter 3

Using TAQ to Measure Liquidity

This chapter shows in full detail the data processing preceding liquidity measurement when using high-frequency data, with specific focus on the TAQ data set. This is a process necessary for liquidity research that has not previously been fully covered in the literature. In recent years the size of the TAQ database, in particular the quote database, has increased immensely, creating a new set of problems that have not been covered in previous processing guides. The purpose of this chapter is to fill that gap in the literature by providing a transparent exposition of the decisions necessary before liquidity measurement. In particular, data downloading and filtering, the matching of trade and quote observations, and remedies for simultaneous observations and erroneous observations are discussed.

3.1 Introduction

In the previous section I present monthly observations of liquidity measured in eight different ways. These are based on the stocks that were in S&P500 on 31 December 2007, covering the years 1995–2007. For each liquidity measure, this implies an unbalanced panel of 156 months and 500 stocks. The data used for calculating this liquidity data set comes from the TAQ database. The processing preceding any calculations is extensive, but it is rarely documented in detail. As this is a tedious but necessary part of liquidity research today, the lack of guidance constitutes a clear gap in the literature. Several choices, often arbitrary, have to be made in the processing of data, and transparency of these is important for analysis of the end results. The purpose of this chapter is to provide a detailed exposition of the different steps necessary to take prior to liquidity measurement when using TAQ data.

The reason that arbitrary decisions have to be made in the data processing is that the size of the database prohibits manual evaluation of each data point. For the years 1995–2007 and the stocks and exchanges that I cover, the TAQ database holds in total 5.2 billion observations on trades and 24.3 billion observations on quotes. Figure 3.1 illustrates how the

44
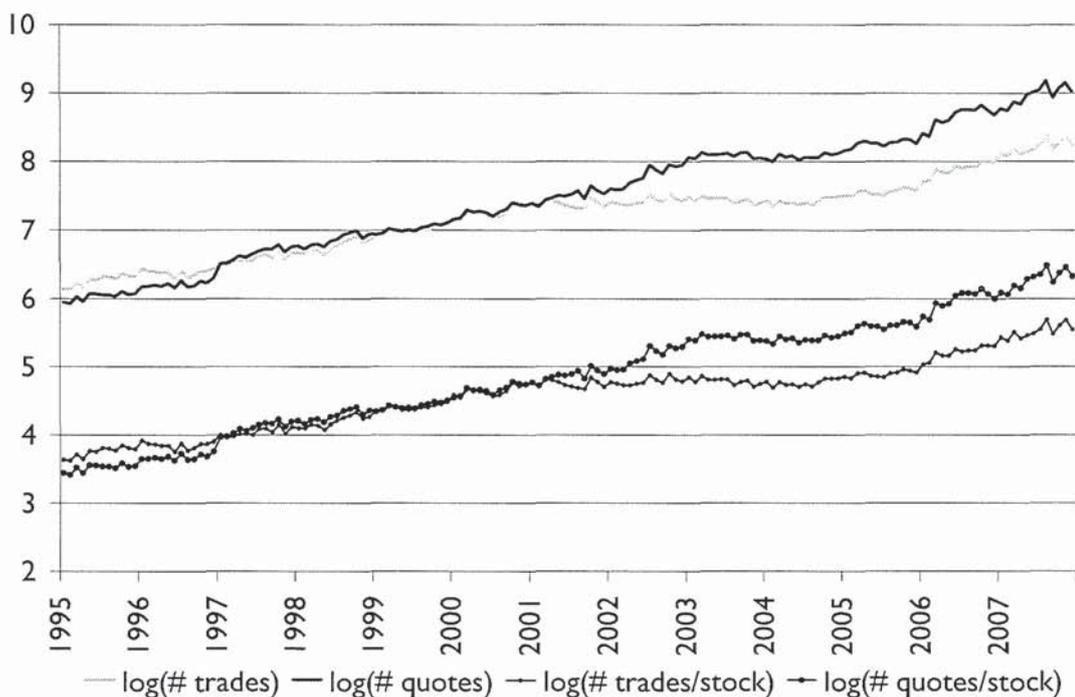
Figure 3.1: Number of observations from TAQ database 1995–2007

The number of trades and quotes observations are shown on a logarithmic scale monthly for the period 1995–2007, both in total and divided with number of stocks considered each month.

size of the database has developed over the years covered. As seen in the figure, which has a logarithmic scale, there were more trades than quotes observation in 1995 (1.3 and 0.9 million per month respectively). The number of quote observations have increased much faster than the number of trade observations, and in 2007 there were almost 6 times more quotes than trades (1 billion quotes and 0.17 billion trades per month).

An earlier description of TAQ data processing is available in Bauwens and Giot (2001, Chapter 2), which describes the database structure in more detail than I do, and has a focus on financial durations and intraday seasonalities — two topics that I do not cover here. Furthermore, Boehmer, Broussard, and Kallunki (2002, Chapter 10), give an overview of how to read TAQ data in the SAS software, and have some general advice on liquidity measurement in terms of bid-ask spreads. Both of these guides overlap with the current chapter in that they show how matching of trades and quotes can be done, but they do not incorporate the latest research on the appropriate time stamp delay of trades (see Section 3.5). In addition, they do not address the increasing problem of simultaneous observations (see Section 3.4), which has appeared after the publication of their work, and they do not address erroneous observations.

The processing of data preceding liquidity calculations can be described in five steps: (1) data downloading; (2) data filtering (3) aggregation of simultaneous data; (4) matching of trades and quotes; and (5) data cleaning. I present each of these steps below, with a summary

of the workflow provided in Table 3.1, which can be used as a reference card throughout this chapter.

3.2 Data Downloading

TAQ holds transaction data (trades) and order data (quotes) in separate files that can be translated to trade and quote files using the software *TAQ3*. Trade records hold data on price (PRICE), volume (SIZ), time stamp (TTIM), and an indicator of the exchange where the transaction took place (EX). All major exchanges in the United States are covered by the database. Furthermore, information on special conditions of the trade and on whether the trade was marked as corrected, cancelled or erroneous is given in two indicator variables (COND and CORR). I am interested primarily in the first four variables, but all variables need to be downloaded for data filtering purposes. Quote records contain information on best bid price (BID) and volume (BIDSIZ) and best offer price (OFR) and volume (OFRSIZ) at a certain time (QTIM) and exchange (EX). Specific conditions around quotes are given in an indicator variable (MODE).

For each stock and each month, I download all trades recorded between 09.30 and 16.00 (the opening hours on AMEX, NYSE, and NASDAQ). It is also possible to specify what exchanges should be included, but an indicator change of NASDAQ records on June 28 2006 that is not accounted for by TAQ3 makes it necessary to download data from all exchanges in order to access NASDAQ trades (see TAQ3 documentation, read me file). Table 3.2 shows an example of all the input parameters for TAQ3.[1] It is important to note that trade and quote records are downloaded to separate files, they are not matched to each other at this stage. TAQ also contains some monthly stock characteristics data stored in *master* files. I read these separately, not using the function in TAQ3.[2] These files give information for each ticker (SYMBOL) on full company name (NAME), CUSIP number (CUSIP), number of shares outstanding (SHARESOUT), lot size (UT), tick size (DENOM) and share type (TYPE).

3.3 Data Filtering

The stocks considered in this thesis are traded at many different exchanges, but have their primary listing on NYSE, NASDAQ, or AMEX. Liquidity provided in smaller exchanges can be interesting, but can also disturb the data set as trading rules differ across exchanges. To minimise such effects, I limit the data set to the exchanges where stocks have their primary listing. NASDAQ and NYSE have introduced automated trading platforms (NASDAQ-ADF

---

[1]For this project I used monthly job files for each of the 500 stocks considered. These job files were generated by a C program written by Yu Man at the Federal Reserve Bank of St Louis. This program was very helpful in automating the downloading process.

[2]Statistical software used throughout this data processing is R 2.8, Windows version.

1. Data downloading

Data files downloaded	*Variables*
Trade files (156 months, 500 stocks)	PRICE, SIZ, TTIM, COND, CORR, EX
Quote files (156 months, 500 stocks)	BID, BIDSIZ, OFR, OFRSIZ, QTIM, MODE, EX
Master files (156 months)	SYMBOL, NAME, CUSIP, SHARESOUT, DENOM, TYPE

2. Data filtering

a) Filters applied to both trades and quotes

EX_i =1,N,T,A,D,Q,P*	NYSE, AMEX, NASDAQ, and NASDAQ-ADF data are retained. Records from other exchanges are discarded. *After February 2006, when NYSE took over Arca, records from that exchange (P) are also retained.
Number of observations > 10	Stocks that have less than 10 observation on either trades or quotes in one month are excluded.

b) Filters applied to trades

$0.5 < PRICE_i < 15000$	Trades with extreme transaction prices are excluded.
$CORR = 0,1$	Records containing correct trades are retained
$COND$ = [blank],*,E,0	Trades without conditions, and trades marked as NYSE Direct+ records are retained.

c) Filters applied to quotes

$OFR_i > BID_i$	Only positive spreads are retained.
$OFR_i - BID_i < \$5$	Large spreads are excluded.
$2(OFR_i - BID_i)/(OFR_i + BID_i) < 0.25$	Spreads larger than 25% of their midpoint are excluded.

3. Aggregation of simultaneous observations

a)

if $TTIM_i = TTIM_j$ $(i < j)$ $\Rightarrow PRICE_i = \frac{PRICE_i SIZ_i + PRICE_j SIZ_j}{SIZ_i + SIZ_j}$ $\Rightarrow SIZ_i = SIZ_i + SIZ_j$ $\Rightarrow$ Discard trade j	If two or more trades are recorded at the same second, their volumes are summed and the aggregated trade price is calculated as the volume-weighted average trade price.

b)

if $QTIM_i = QTIM_j$ $(i < j)$ $\Rightarrow$ Discard trade i	If two or more quotes are recorded at the same second, the last quote record of that second is used.

4. Matching of trades and quotes

The quote prevailing at each trade is taken to be the last quote at least one second before the trade. If no quote exists before the trade at that day, no quote is assigned. Quotes that are not assigned to trades are discarded.

Table 3.1: Workflow for data processing (see caption on p.48)

5. Data cleaning

a) Trades:

if $|\text{PRICE}_i - p^-_{\kappa,i}| > 3\sigma_{\kappa,i}(\text{PRICE}_i) + \zeta_i$	Trades where the price is more than 3 standard
$\Rightarrow$ Discard trade i	deviations $(\sigma_{\kappa,i})$ away from its δ-trimmed mean
$\kappa = min(10, \# \text{ trades in the current day}/270)$	$(p^-_{\kappa,i})$ over the κ surrounding observations are
$\delta = 10\%; \zeta_i = 2 \times \text{DENOM}_i$	discarded. ζ_i is a granularity term set to be
	twice the tick size.

b) Spreads:

if $\text{PRICE}_i > 1.1 \times \text{OFR}_i$	When the trade price is more than 10% higher
or $\text{PRICE}_i < 0.9 \times \text{BID}_i$	(lower) than the ask (bid) price, the quote
$\Rightarrow$ Discard quote i	observation is discarded. The trade is not
	discarded.

a) Shares outstanding:

if $\Delta\text{SHARESOUT}_i > 25\%$	Changes in number of shares outstanding that
and $\text{SHARESOUT}_{i+j} = \text{SHARESOUT}_{i-1} \pm 10\%$	are reversed within 10 months are regarded as
$(j = 1, ..., 10)$	erroneous. All observations from the change
$\Rightarrow$ Discard $\text{SHARESOUT}_i{:}\text{SHARESOUT}_j$	until the reversal are discarded.

Table 3.1: Workflow for data processing

The table shows the workflow of data processing prior to liquidity calculation. The indexes i and j are discrete positive numbers referring to records of trades and quotes. After the matching in step 4, trades and quotes are in the same records.

and Arca respectively) parallel to their traditional exchanges. I include data from these platforms as well. Records from all other exchanges are discarded, see Table 3.1:2a for exact filter specifications.

In order to avoid noise in the liquidity measurement it is common to apply a number of filters on trade and quote records. In this regard I follow the conventions of previous literature. When a stock does not have at least 10 observations on both trades and quotes in a given month, I exclude it from liquidity measurement that month, as the risk of measurement error would be too large.

Trades that are cancelled, corrected, out of sequence, or have special conditions attached to them are excluded. Such information is covered by the CORR and COND indicators in TAQ. By applying the filters specified in Table 3.1:2b, I exclude trades with the conditions: Cash-only basis, Bunched, Cash Sale, Next Day Settlement Only, Bunched Sold, Rule 127, Rule 155, Sold Last, Next Day, Opened Last, Prior Reference Price, Seller, Split Trade, Pre-/Post-Market Trades, Average Price Trades, Opened After Trading Halt, Sold Sale, and Crossing Section. In addition, I exclude trades at very high ($15000) or very low ($0.5) prices.

I follow Korajczyk and Sadka (2008) and classify quotes with negative spreads, spreads larger than $5, or spreads larger than 25% when the spread midpoint is less than $20, as implausible. The implausible quotes are discarded, as described in Table 3.1:2c.

```
[JOBSETTINGS]                              [QUOTEFIELDS]
JOBDESCRIPTION=Example                      QTIM=1
INPUTSYMBOLFILE=                            BID=1
QUOTESFILENAME=D:\...\200702NYXq            OFR=1
TRADESFILENAME=D:\...\200702NYXt            QSEQ=0
STATSFILENAME=                             BIDSIZ=1
TIMEPERIOD=0                                OFRSIZ=1
INPUTOPTIONS=0                              MODE=0
INPUTSYMBOLLISTTYPE=0                       EX=1
NUMBERFORMAT=1                              MMID=0
TIMEFORMAT=2                                [TRADEFIELDS]
DATEFORMAT=1                                TTIM=1
OUTPUTOPTIONS=1                             PRICE=1
OUTPUTFORMAT=1                              SIZ=1
SELECTEDTIMERANGE=1                         TSEQ=0
MONTHTOPROCESS=0                            G127=1
TIMERANGEINTERVAL=1                         CORR=1
INCLUDEQUOTES=1                             COND=1
INCLUDETRADES=1                             EX=1
INCLUDESTATS=0                              [EXCHANGES]
INCLUDEHEADER=1                             AMEX=1
HEADERFIRSTRECORDONLY=1                     BOSTON=1
INCLUDEMAST=0                               CINCINNATI=1
INCLUDEDIV=0                                MIDWEST=1
OVERWRITEOUTPUTFILES=1                      NYSE=1
AFTERHOURSSTATISTICS=0                      PACIFIC=1
DATABASEIMPORT=0                            NASD=1
INCLUDECORRECTIONS=1                        PHILADELPHIA=1
JOBACTIVE=1                                 INSTINET=1
STARTDATE=2/1/2007                          CBOE=1
ENDDATE=2/28/2007                           NASDADF=1
STARTTIME=09:30:00
ENDTIME=16:05:00
SYMBOL/CUSIPLIST=NYX
```

Table 3.2: Example job file for TAQ3

The table shows the inputs given to the trade and quote data downloading program TAQ3 through job files. Job files were created for each stock and each month. The example covers ticker NYX in February 2007.

3.4 Aggregation of Simultaneous Observations

As the number of quote records in TAQ has increased dramatically lately it is common to find many quote observations of one stock that are reported in the same second. As TAQ does not distinguish time units smaller than one second, all these quotes can not be considered. There is not much guidance in the literature on how to deal with this problem. Brownlees and Gallo (2006) recommend using the median quote but also say that the first or last observation could be used (the latter is the choice of e.g. Korajczyk and Sadka 2008). There does not appear to be any theoretical foundation for either, so I make an arbitrary choice and use the last observation each second (see Table 3.1:3b).

Trades in the same stock that are recorded within one second are less common but happen throughout the sample considered. As these will all be matched to the same quote, several researchers merge the simultaneous trades (Hansch 2003, Engle and Patton 2004, Brownlees and Gallo 2006). These researchers all agree that volumes can straightforwardly be aggregated by summation, but there are different approaches to find a merged trade price. Engle and Patton (2004) use the price of the first trade in each second, assuming that all trades in the second are executed at the same price. Brownlees and Gallo (2006) use the median price in each second. As price impact of trading is an important aspect of liquidity, I think that the volumes of the simultaneous trades are important to consider. Hence, I use a volume-weighted average as measure of aggregated trade price, and sum of volumes as measure of aggregated trade volume (see Table 3.1:3a). It is noteworthy that if the assumption of Engle and Patton (2004) is true, all these approaches will yield the same aggregated trade price.

3.5 Matching of Trades and Quotes

For *Espread* measurement and price impact regression analysis, trades need to be matched with quotes prevailing at the time just before the trade was executed. This is needed for determining whether the trade is buyer- or seller-initiated. Most of the liquidity literature has used the findings of Lee and Ready (1991) as guidance on how to match trades and quotes for this approximation. Based on delays in trade reporting, their recommendation is to use the latest quote recorded at least five seconds before the trade. Henker and Wang (2006) argue that the rule by Lee and Ready (1991) is based on manual trade reporting that is no longer in use in the stock exchanges. According to Henker and Wang (2006), the reporting process was made automatic during the period 1994–2001, leading to much shorter reporting delays, or even simultaneous reporting. A history of trade reporting procedures is available in Vergote (2005). Both Henker and Wang (2006) and Vergote (2005) have investigated what time lag is appropriate after these reforms, finding one and two seconds respectively. Both studies are limited to NYSE.

50

Reporting delays in NASDAQ trades have been investigated by Ellis, Michaely, and O'Hara (2000). In their sample from 1996–1997, they find that even though about 12.4% of the trades are reported within 1 second after execution, a more typical delay is 15–16 seconds (58.6%), and 6.4% of the trades have a longer delay than that. Still, they conclude that a zero delay matching rule is appropriate, as they see no improvement in matching accuracy when applying delays.[3]

I am not aware of any investigation of trade reporting delays on AMEX.

In order to match aggregate trades consisting of trades from several exchanges to appropriate quotes, I need to use the same matching rule for all trades in my sample. I choose to follow the one second rule recommended by Henker and Wang (2006), as that study was performed on the post–1993 S&P500 index stocks with primary listing on NYSE, which is matching the majority of the trades in my sample (Vergote, 2005, uses a much smaller set of NYSE stocks). The same rule has been used for NASDAQ trades in a recent study by Stoll and Schenzler (2006), and this is close to the recommendation by Ellis, Michaely, and O'Hara (2000). I do not investigate how sensitive my results are to the chosen trade delay rule. As the rule has been derived on the basis of a sample matching the majority of the trades in my data set, it is likely that a different trade reporting delay rule would yield less accurate matching of trades and quotes, and accordingly induce more noise in the subsequent estimates of systematic factors.

3.6   Data Cleaning

In a huge data set as the one at hand, there is an obvious risk of erroneous entries that can have slipped through the above processing. Many studies deal with this process by Winzorising the liquidity measures once they are calculated, i.e. they trim the tails of the liquidity probability distributions (Amihud 2002, Sadka 2006, Korajczyk and Sadka 2008). As this is arguably a crude method, I am instead dealing with unreasonable data points *before* calculating liquidity measures. To distinguish erroneous data points from true observations of extreme values is sometimes difficult. The data cleaning described in this section is done with the purpose of removing errors but not true outliers. Unfortunately, the size of the data set prohibits double-checking of each removed observation, making the correct design of detection algorithms important.

---

[3]Ellis, Michaely, and O'Hara (2000), and later also Chakrabarty, Li, Nguyen, and Van Ness (2007), develop algorithms for how to classify NASDAQ trades as buyer- or seller-initiated. These algorithms differ from the quote rule that I apply. I stick to the simpler rule for computational reasons. Furthermore, simultaneous trades (see section 3.3) can contain trades from both NYSE and NASDAQ, so applying different rules for different exchanges would cause a classification problem for these trades.

3.6.1 Trades

Trades in my relatively liquid sample typically occur in smooth price sequences. When inspecting the data, it appears that in rare cases trades appear that are far off the price sequence for the stock. This may be due to erroneous recording of transaction price, time, or trade conditions. Brownlees and Gallo (2006) provide an algorithm for dealing with such trade price inconsistencies. They suggest that for each day, trades that deviate more than three standard deviations from the delta-trimmed mean should be regarded erroneous. Details of this algorithm are given in Table 3.1:5a.

Brownlees and Gallo (2006) suggest that the number of observations (κ) considered when calculating the mean and the standard deviation should be chosen in accordance with the trading activity in the stock. This allows for some time-variation in the standard deviation. They choose $\kappa = 60$ for their sample, which corresponds to a trading intensity of $270 \times \kappa$ per day. I use this factor (270) to calculate κ for each day for each stock, setting 10 as a minimum value for κ. For days with less than κ observations, all observations are used. The granularity coefficient (ζ), which is used to avoid zero variances, should be chosen as a multiple of the minimum tick size. Brownlees and Gallo (2006) choose $\zeta = 0.02$, which is a multiple of 2 of the tick size used in their sample (0.01). As the minimum tick size decreases substantially over the years of my data set, I let the parameter be set by the double monthly tick size variable (DENOM) available in TAQ.[4]

A potential development of the algorithm would be to introduce an autoregressive conditional heteroscedasticity (ARCH) specification when calculating mean and standard deviation. This could improve the precision in the filter, but would also add computational burden to the filtering. I consider this an interesting venue for future research, but for the application in this thesis I follow the algorithm as specified by Brownlees and Gallo (2006).

In 1995–2007, this filter captures 3.6 million trades, which is 0.07% of the total number of observations on trades. The intensity of erroneous trades peaks at the burst of the dot-com bubble in March 2000. Closer inspection shows that these observations primarily come from NASDAQ companies, a tendency also reported by Brownlees and Gallo (2006). In Figure 3.2, I provide a snapshot of a trading day with many trades identified as erroneous in one firm. It shows that the algorithm captures trades that are clearly outside the core main price sequence. It is known that the trade reporting sometimes has delays at NASDAQ, which can make a trade look like it is out of sequence. If this is the case, it is good to remove those trades, as they otherwise will have a time mismatch between trade and quote. Another theory could be that informed traders would have reason to transact out of the price sequence. However, deviations from the price sequence of this magnitude would be inefficient behaviour by those traders, as they would reveal the private information to the market. Hence, I see it as

---

[4]Tick size data is not available in TAQ for all stocks. For these cases, the first 1000 non-aggregated trades in the month are scanned to identify the minimum non-zero price movement.
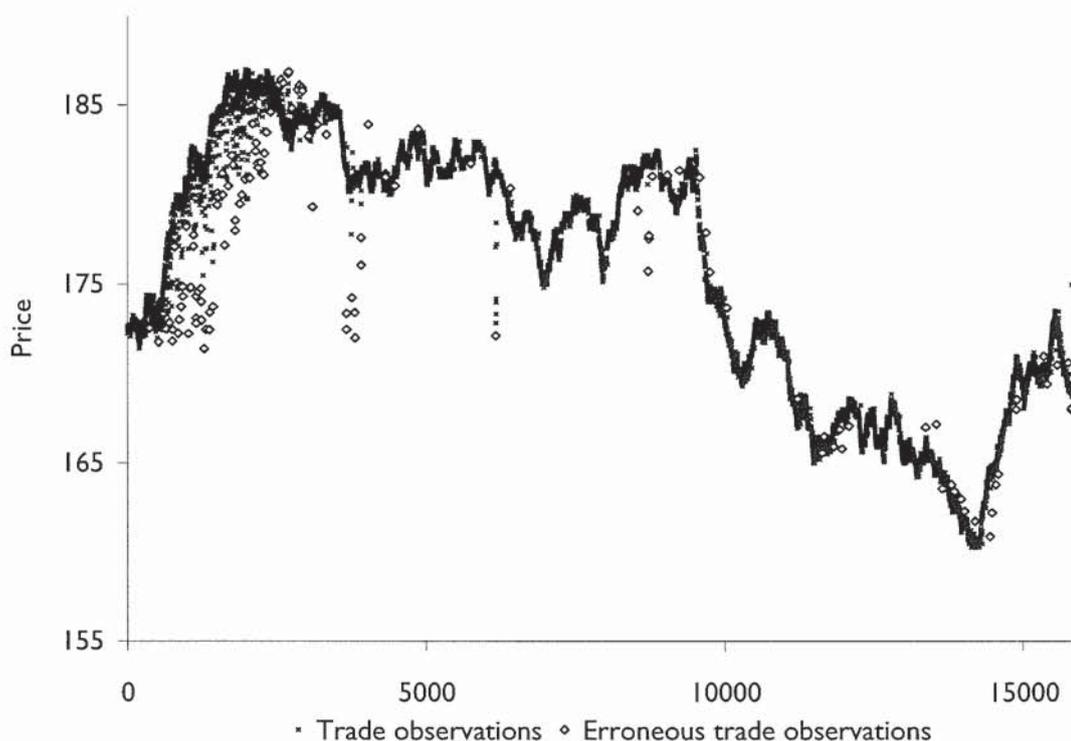
Figure 3.2: Illustration of erroneous trade observations, YHOO, March 30, 2000

Each black data point represents one trade in YHOO in March 30, 2000. Trades are sorted chronologically on the horizontal axis. In total there were 15826 trades on the day, out of which 127 were identified as erroneous by the Brownlees and Gallo (2006) algorithm (0.8% of total).

unlikely that the detected trades mirror true deviations from the trade sequence. The finding of typical trade reporting delays of more than 15 seconds at NASDAQ (Ellis, Michaely, and O'Hara 2000) implies that on days with high volatility many trades can appear to be out of sequence.

3.6.2 Quotes

Another problem found in the TAQ database is quotes that are not on the same level as the trade that they have been matched to, i.e. the trade price is outside the spread. I exclude quotes observations if the trade is more than 10% outside the bid-ask spread, as described in Table 3.1:5b. The decision level, 10%, is chosen arbitrarily, as no guidance has been found in the literature. These trade and quote mis-matches are likely due to a lack of recent quotes in the data set. If the trade price changes with more than 10% in one day, without any updated quote observations, this type of erroneous outliers appear. This is not mirroring economic events, and the observations should hence be removed. It does happen that trades are actually executed outside the bid-ask spread, but in cases where this happens at more than 10% price premium or discount, it is likely due to trade conditions not covered in this data set. This filter excludes less than 0.001% of all quotes that are matched to trades in the sample 1995–2007,

but has a much larger impact for 1993–1994 (16%), where quote observations are scarce. For those years, there exists whole days where quote prices are unrelated to trade prices. This is due to unknown problems in the TAQ database. The problem serves as further motivation for excluding 1993–1994 data from the analysis.

3.6.3 Shares Outstanding

Finally, a measure in TAQ that appears to have many erroneous entries is shares outstanding (SHARESOUT). This measure is used when calculating *Turnover*, see Equation 2.5. An example of erroneous data is the ticker TMK (Torchmark Corp.), in 1997. This stock had a split in August 1997 that doubled its number of shares, which is correctly reported in TAQ. SHARESOUT is typically not a volatile measure, and in the two subsequent months its value is constant. In November 1997, however, TAQ reports the pre-August level of SHARESOUT, and then in December it is back at the newer value. This is illustrated in Figure 3.3. Such jumps in SHARESOUT are clearly erroneous, and they induce outliers in the turnover measure. As this problem has not been discussed in the literature before (to my knowledge), I design an algorithm for removing data periods with dubious reporting. If a change in SHARESOUT of $> 25\%$ is reversed within 10 months, the period of the temporary level is discarded. A reversal is defined as a move back to the previous level $(\pm 10\%)$.[5] In the given example, illustrated in Figure 3.3, this means that the observations for August-November 1997 are discarded. The algorithm filters out 0.47% of all observations on SHARESOUT. The number of erroneous observations is decreasing over time, indicating increasing quality of the data.

3.7 Concluding Remarks

The purpose of this chapter was to describe the data processing preceding actual liquidity measurement that is necessary when utilising TAQ data. I described this process in five steps with the objective to produce a transparent and useful guidance for liquidity research. Some of the problems of TAQ processing have appeared over the last few years (post-2004), in particular the problem of simultaneous observations. My summary on treatment on this problem is a novel contribution to this small literature. In addition, the detection of errors in the data on shares outstanding has not been dealt with elsewhere, and the approach described in Section 3.6 is a first attempt to remedy that problem.

---

[5]The threshold levels used in this algorithm are arbitrarily selected.
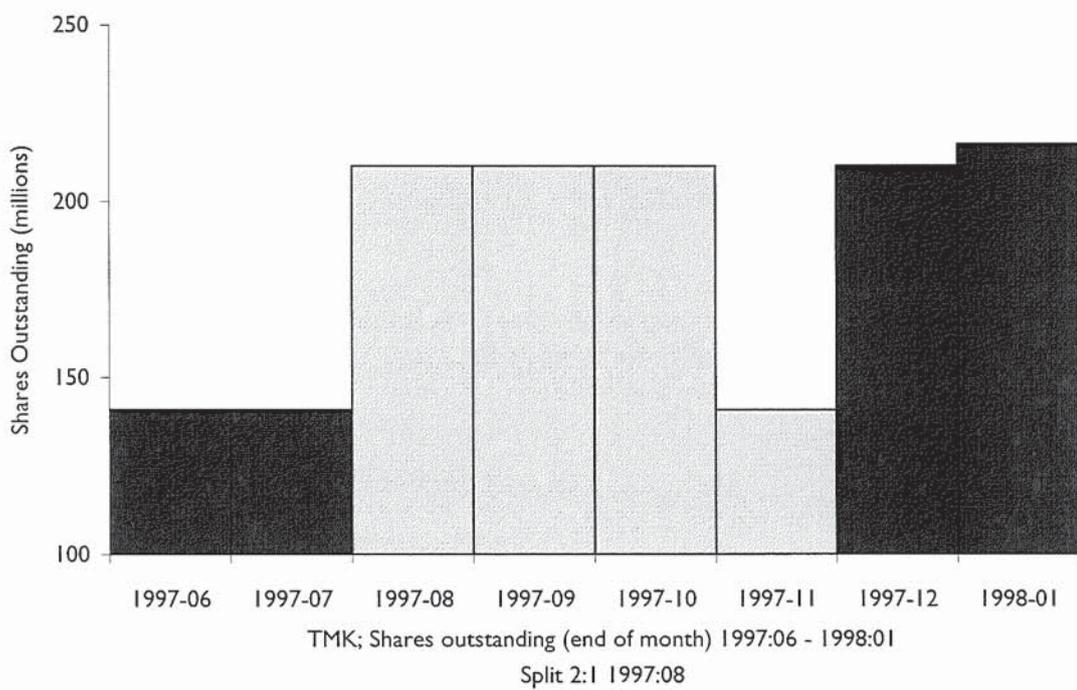
Figure 3.3: Illustration of erroneous SHARESOUT observations

Each bar represents number of shares outstanding (end of month) in ticker TMK, covering June 1997–January 1998. Light grey columns are excluded from the data set as they feature a reversal in November 1997. Dark grey columns are retained.

Chapter 4

Dynamics of Systematic Liquidity

I develop the principal component analysis (PCA) approach to systematic liquidity measurement by introducing moving and expanding estimation windows.[1] I evaluate these methods along with traditional estimation techniques (full sample PCA and market average) in terms of ability to explain (1) cross-sectional stock liquidity and (2) cross-sectional stock returns. For several traditional liquidity measures my results suggest an expanding window specification for systematic liquidity estimation. However, for price impact liquidity measures I find support for a moving window specification. The market average proxy of systematic liquidity produces the same degree of commonality, but does not have the same ability to explain stock returns as the PCA-based estimates.

4.1 Introduction

As referred in Section 2.4, it is well-established that liquidity affects equity prices and returns. Furthermore, co-movements in stocks' liquidity, driven by unobservable market forces called systematic liquidity factors (Chordia, Roll, and Subrahmanyam 2000, Huberman and Halka 2001, Hasbrouck and Seppi 2001), have also been shown in many studies to affect individual stock returns (Amihud 2002, Pastor and Stambaugh 2003, Acharya and Pedersen 2005, Korajczyk and Sadka 2008). Hence, systematic liquidity factors are relevant as risk factors to any investor.

To manage liquidity risk, accurate measurement of systematic liquidity is crucial. In this

---

chapter I use the familiar principal components estimation technique to derive the factors. I extend previous research by introducing dynamic estimation windows in this setting, allowing an investigation of the temporal stability (robustness) of the systematic liquidity measure.

PCA is a popular method for explaining the co-variation between many variables in terms of a small number of common factors. Most investigations use a static specification of PCA, based on analysis of the co-variance matrix of variables of interest, which is assumed to be time-invariant. In the case of liquidity, where both systematic liquidity and commonality in liquidity are known to vary over time (Chordia, Roll, and Subrahmanyam 2001, Kamara, Lou, and Sadka 2008), this assumption may be counterfactual. Still, all studies using PCA to derive systematic liquidity choose this static approach.[2] I test the appropriateness of this assumption by introducing moving and expanding (recursive) estimation windows for PCA.[3] I compare my results to those obtained from PCA with static window (i.e. where the full sample is used); and those from a cross-sectional (equal weighted) liquidity average, another common approximation of systematic liquidity (Chordia, Roll, and Subrahmanyam 2000); giving me four methods to approximate systematic liquidity.

For each of the eight liquidity measures described in Section 2.2, I evaluate the four different methods for deriving systematic liquidity factors, using two different evaluation criteria: (A) ability to explain cross-sectional variation in stock liquidity (i.e. commonality in liquidity) and (B) ability to explain cross-sectional variation in stock returns. I argue that high commonality is an indicator of systematic liquidity measurement accuracy. PCA is based on the estimated co-variance matrix. If the estimated co-variance matrix differs over different estimation horizons it can be due to either sampling error or actual time-variation in the underlying co-variance matrix. If it is due to noise, a longer estimation window should yield higher commonality, as the sampling error is smaller with more data. If there is time-variation in the underlying co-variance matrix, the PCA with moving estimation window should be able to capture this time-variation and therefore produce a higher degree of commonality. Hence, by running commonality tests I investigate which estimation method yields factors that best summarise cross-sectional liquidity variation.

As explaining stock returns is central to the application of systematic liquidity factors, I make this ability my second evaluation criterion in my assessment of estimation methods for systematic liquidity. I run an extended market model including systematic liquidity factors, and study the improvement in return variation explanation relative to the standard market model.

My first finding is that time-series properties of the liquidity co-variance matrix differ

---

[2] Hasbrouck and Seppi (2001); Beltran-Lopez, Giot, and Grammig (2005); Chen (2007); Chollete, Næs, and Skjeltorp (2007); Chollete, Næs, and Skjeltorp (2008); Kempf and Mayston (2008); Kopp, Hütl, Loistl, and Prix (2008); Korajczyk and Sadka (2008).

[3] The only related liquidity study that does not stay with the static window (full sample) PCA is Chen (2007) who uses an expanding window specification; however she does not discuss the implications of this and the constant co-variance assumption remains present.

across liquidity measures. When considering liquidity measured by the *Qspread*, *Espread*, *Turnover* or *ILLIQ*, the accuracy of estimated systematic liquidity is not improved by allowing the co-variance matrix to vary over time. The static, expanding and moving window approaches produce roughly the same average commonality. This suggests that the co-variance matrices of these measures are quite stable over time. For the price impact measures based on Glosten and Harris (1988) and Sadka (2006), however, I find that the moving window approach improves the accuracy of estimated systematic liquidity, i.e. the average commonality is consistently higher when applying the moving window PCA. Hence, there appears to be time-variation in the underlying co-variance matrix for these measures. My investigation of ability to explain stock returns point in the same direction as that of stock liquidity. For most liquidity measures, different estimation window specifications for PCA do not make any difference for explaining returns. Transitory fixed market maker costs, which is Sadka's (2006) measure of fixed inventory cost, is an important exception to this. The systematic factors of this liquidity measure, relative to the other measures considered, have a good ability in both explaining stock liquidity and stock returns (highest of all measures), and it is best estimated using the moving window PCA.

A second finding is that the market average proxy of systematic liquidity, suggested by Chordia, Roll, and Subrahmanyam (2000), produces a commonality that on average is in line with the various PCA factors for all measures considered (with the exception of *ILLIQ*, which is plagued by outliers). For *Turnover* and spread measures of liquidity, the cross-sectional average can be used as proxy for systematic liquidity. For the other measures, time series properties in commonality produced by the market average differ substantially from those of PCA measures. Furthermore, when applied to stock return variation, the PCA-based measures are superior to the market average regardless of the choice of liquidity measure.

Thirdly, I find that the properties of static window (full sample) PCA and the expanding window (recursive) PCA in terms of ability to explain both stock liquidity and stock returns are converging over time. The latter is computationally more expensive, but the former is suffering from a forward looking bias that undermines its applicability in practice. My findings indicate that the static window PCA can be replaced by expanding window PCA without loss of accuracy.

In the next section I present the framework applied to estimate systematic liquidity factors with PCA, along with a discussion on how to deal with missing data, outliers, and computational burden for this kind of problems. In Section 4.3, I perform commonality tests and analyse results thereof, and in Section 4.4 I evaluate systematic liquidity in terms of explanatory power with respect to stock returns. Finally, in Section 4.5, I summarise the main implications of the chapter.

4.2   Systematic liquidity derivation

I assume that the data-generating process of liquidity for a given stock is driven by some underlying market liquidity factor (F) and an idiosyncratic liquidity variable (Γ^*). In a market of N firms for which τ periods are considered, yielding a stock liquidity matrix L^* of dimension $(N \times \tau)$, this process can be described as

$$L^* = \mu(L^*)\iota' + \Xi F + \Gamma^*, \tag{4.1}$$

where the underlying market liquidity factor is described by H row vectors $(h = 1, 2, ..., H)$ in the matrix F $(H \times \tau)$ and each stock's sensitivity to these vectors is given in the matrix Ξ $(N \times H)$. The column vectors of F are defined to be orthogonal to each other. The matrix of idiosyncratic liquidity shocks, Γ^*, has dimension $(N \times \tau)$. The average liquidity across time for each stock is given in the vector $\mu(L^*)$ of length N and ι is a vector of ones of length τ.

In order to estimate F with equal influence of each stock's variance, I standardise liquidity of each stock to have unit variance and zero mean over the τ periods of time.[4] Denoting the vector of stock liquidity standard deviation for each firm $\sigma(L_i^*)$, the elements of the standardised liquidity matrix takes the form

$$L_{i,t} = (L_{i,t}^* - \mu(L_i^*))/\sigma(L_i^*). \tag{4.2}$$

Using this expression, the factor model described in Equation 4.1 can be expressed as

$$L = \xi F + \Gamma, \tag{4.3}$$

where $\xi_{i,h} = \Xi_{i,h}/\sigma(L_i^*)$ and $\Gamma_{i,t} = \Gamma_{i,t}^*/\sigma(L_i^*)$. Allowing the idiosyncratic liquidity shocks Γ to be weakly cross-sectionally correlated, I estimate this as an approximate factor model (Chamberlain and Rothschild 1983). The cross-sectional correlation is decreasing with the numbers of factors used (H). The idiosyncratic liquidity shocks may also feature autocorrelation. When τ is large enough, the model can still be consistently estimated using PCA (see Stock and Watson, 2002a, 2002b). The principal components are the eigenvectors of the co-variance matrix of L (Σ_L) multiplied by L, and sorted into row $h = 1, 2, ..., H$ of $\hat{F}$ by its corresponding eigenvalues, starting with the highest.

---

[4]Standardisation of the data is customary for PCA, but the exact methodology to do this has varied in the liquidity literature (and has often not been disclosed). For example, Korajczyk and Sadka (2008) are using an expanding estimation window to calculate market-wide mean and standard deviation. I standardise on a stock-by-stock basis using means and standard deviations calculated for the τ periods considered, as market-wide standardisation may give disproportionate weight to stocks with high liquidity volatility.

4.2.1 PCA with Dynamic Estimation Window

In the systematic liquidity literature, τ has typically been set to the sample size T, yielding a static $\hat{F}$. This implies an assumption of constant co-variances in the cross-section over the sample size T. As mentioned in Chapter 2, due to tick size changes, financial crises, and macroeconomic events the data entering this analysis is likely to feature structural breaks. Furthermore, there is evidence of seasonality patterns in liquidity time series (Chordia, Sarkar, and Subrahmanyam 2005). This makes the assumption of constant co-variances questionable.

I challenge the validity of constant co-variances by running a moving window PCA, utilising only the τ latest observations. This excludes old observations that may otherwise yield biased estimates, e.g. in the presence of structural breaks. With a moving estimation window, time-varying co-variances are allowed. Formally, this implies a structure with time index t on each of the variables in Equation 4.3 and a time span $(t - \tau + 1) : t$ considered for estimation in each period t. The estimation is repeated for $t = \tau, \tau + 1, \tau + 2, ..., T$.

I also consider expanding window PCA (or recursive PCA), where the problem dimension is growing with t, implying a time span $1 : t$. This specification is appropriate if the co-variances are believed to be time-invariant, and where the usage of future data should be avoided. Such a specification has earlier been considered in a systematic liquidity study by Chen (2007). Henceforth, I shall refer to the the moving and expanding window PCA methods as *PCA with dynamic estimation window*.

Earlier economics applications of PCA with dynamic estimation window have appeared in literature on integration of equity markets (Volosovych 2005, Gilmore, Lucey, and McManus 2008) and macroeconomic forecasting (Heij, van Dijk, and Groenen 2008), and more commonly outside economics, in particular process monitoring (Li, Yue, Valle-Cervantes, and Qin 2000, Wang and Xia 2002).

For both versions of dynamic estimation window PCA, I start the estimation at $\tau = 36$ and for the moving window specification I keep the sample size constant at that size throughout the sample. This time horizon, three years, should be long enough to identify commonality due to monthly seasonality in the liquidity time series. It is also short enough to get a useful comparison between expanding and moving estimation windows. As a comparison to previous studies, I also provide the traditional version of PCA, estimated over the whole sample (I refer to this as *static window PCA*).

The three specifications of PCA can all be nested into the moving window notation of time span given above, setting the window size to $\tau = [36, t, T]$. For further comparison to previous literature, I evaluate these methods together with an equal-weighted cross-sectional average of liquidity, used by Chordia, Roll, and Subrahmanyam (2000).

When using liquidity data on a higher frequency, such as intra-day or daily, it is possible that high-liquidity stocks are affected by market-wide liquidity changes before low-liquidity

stocks see any effects (as discussed by Kempf and Mayston, 2008; and Hallin, Mathias, Pirotte, and Veredas, 2009). Such intertemporal co-variations are ignored in PCA, as the analysis is based on the static co-variance matrix. Dynamic PCA models (or generalised dynamic factor models), introduced by Forni, Hallin, Lippi, and Reichlin (2000, 2005), consider leads and lags in the estimation and derive common factors that are orthogonal at all leads and lags of the series. In my application, with liquidity data on monthly frequency, lagged market liquidity effects appear unlikely, which is why I do not use the dynamic PCA model. An application of dynamic PCA on systematics in daily liquidity data is available in Hallin, Mathias, Pirotte, and Veredas (2009). They also investigate commonality across liquidity measures. In their setting, with daily data and multiple liquidity measures, the ability to investigate leads and lags is important, and the dynamic PCA is an appropriate modelling choice.

The dynamic PCA model also has a clearer distinction between common and idiosyncratic components than what is seen in PCA. Common components are orthogonal not only to each other but also to the idiosyncratic component. Still, in all versions of PCA (static, dynamic estimation window, and dynamic) some commonality remains in the idiosyncratic component. Hence, the benefit of this property is limited for my application, though it can be useful for determining how many common factors to use (traditional PCA does not have an explicit rule for this). Also, in a recent Monte Carlo experiment with a finite sample setting, Kapetanios, Marcellino, Schifanoia, and Salasco (2009) show that the correlation between true and estimated factors is consistently higher when using the PCA estimator than when using the dynamic PCA estimator.

In my analysis, I set the number of factors $H = 3$. This choice is based on the choice in previous studies (Hasbrouck and Seppi 2001, Korajczyk and Sadka 2008, Chollete, Næs, and Skjeltorp 2008), as well as visual inspection of eigenvalues, that shows the proportion of total variance explained by each factor.

4.2.2 Estimation Methodology: Robust Asymptotic PCA

PCA has some technical complications that have rarely been discussed in the applications to liquidity. When applying PCA with dynamic estimation windows, where PCA is run hundreds of times, these complications become important for the robustness of the results. As PCA is based on finding eigenvectors in the co-variance matrix of the underlying variables, it is important that the co-variances are robustly estimated. In this subsection, I discuss problems associated with that estimation in some detail, including issues with missing values, outliers, negative eigenvalues, and computational burden. For a more general discussion of PCA and technical considerations thereof, see e.g. Jolliffe (2002).

Firstly, the estimation of the co-variance matrix needs to be adapted for missing values. Typically, when dealing with stock market data there is an unbalanced panel of observations

as companies rise and fall and go through mergers and acquisitions. As exclusion of companies that do not exist throughout the sample would cause material information losses, a method for dealing with long series of missing values is needed. With a moving window specification, stocks with long series of missing values can be excluded until they have enough observations. This partially sidesteps the missing value problem, but temporary cases of missing values remain. To estimate co-variances in the presence of missing values, I use element-wise estimation, using all pairs of observations available for the two variables (referred by Jolliffe 2002, Ch. 13). For a stock to be included in the analysis, I demand that more than 12 observations are available in the time span considered.[5]

Secondly, outliers can inflate or deflate variances and co-variances to an extent that they mislead the PCA substantially (see Jolliffe 2002, Ch.10). In my setting, large idiosyncratic stock liquidity shocks are easily imaginable and these should not be allowed to influence the estimation of systematic liquidity. To deal with this, I use the robust co-variance matrix estimation technique by Mehrotra (1995). This technique is based on using medians rather than means when calculating slopes needed in variance and co-variance estimation. This is a computationally expensive technique, but it yields high stability in the principal components.

Thirdly, a problem appearing when using element-wise estimation of the co-variance matrix is that it may no longer be positive semi-definite. The robust estimation limits this problem, but I still detect several negative eigenvalues, showing that the co-variance is not positive semi-definite. To limit this problem, I use asymptotic PCA as suggested by Connor and Korajczyk (1986). With this methodology, the co-variances between time periods rather than the co-variances between stocks are applied as the basis for PCA. Connor and Korajczyk (1986) show that components derived using this method are asymptotically equivalent to those of PCA, and my investigation shows that in my setting of 500 stocks, the methods are equivalent. When using asymptotic PCA, the problem of negative eigenvalues is much smaller than for PCA, though it is still present. Furthermore, the computational burden of PCA is in this way limited, as the co-variance matrix has dimension τ rather than N.

4.3 Commonality in Liquidity

The degree of commonality is the proportion of stock liquidity variation that is due to systematic liquidity variation. This can be measured by estimating the relationship

$$\Delta_{AR} L^*_{i,l} = \phi_{l,\tau} + \Phi_{l,\tau} \Delta_{AR} \hat{F}_{1:h,l,\tau} + u_{i,l,\tau}, \tag{4.4}$$

---

[5]For two alternative methods dealing with missing values in a static setting, see Korajczyk and Sadka (2008).

which is closely related to Equation 4.1. Here, $\phi_{l,\tau}$ is an intercept and $\Phi_{l,\tau}$ is a vector of systematic liquidity factor loadings. Stock liquidity is regressed on the first h rows of the estimated systematic liquidity factor (i.e. the h first principal components). The residuals, $u_{i,l,\tau}$ are taken to be idiosyncratic liquidity shocks. In accordance with Chordia, Roll, and Subrahmanyam (2000), the proportion of variation in stock liquidity explained by systematic liquidity, R^2, averaged across stocks, is taken as the degree of liquidity commonality. As indicated by Δ_{AR}, I follow Chordia, Roll, and Subrahmanyam (2000) and run the commonality test on innovations in liquidity. These are retrieved by fitting autoregressive models to liquidity measures and systematic liquidity factors.[6] In some previous studies, the commonality test has been run in levels, but I experience substantial non-stationarity problems with such a specification.[7] Note that in addition to the firm index i, I add the subscripts l and τ to distinguish the different liquidity measures and PCA specifications that I evaluate in this framework. Note also that I use non-standardised liquidity data, as denoted by L^*.[8]

I investigate commonality in liquidity for $h = 1, 2, 3$. This means that the three first systematic liquidity factors are included in the $h = 3$ case, implying that R^2 is growing with h. As I am interested in the time-variation of commonality, I run this regression using a moving estimation window of 36 observations. To capture potential short-term trends in the underlying co-variance structure, I use a three year window for my regressions, the same as I use for the moving window specification of systematic liquidity. I have experimented with different window sizes, finding that a two year window yields a less stable systematic liquidity measure (with lower commonality), whereas slightly longer horizons yield results similar to the three year window. A larger estimation window has been considered for the commonality regressions as well, but that naturally makes the differences between moving and expanding windows less significant. I require more than 12 stock liquidity observations in the specified time frame for a firm to be considered in this analysis. To investigate how commonality depends on the estimation of the underlying liquidity factor estimation, the procedure is repeated for my different settings of τ.[9]

Before analyzing these differences between systematic liquidity factors, it is useful to

---

[6] Autoregressive order is chosen using the Schwarz Bayesian information criterion, with up to two lags allowed. I allow the lag length to differ across liquidity measures. The model is run using a 36 months moving window.

[7] Other ways to overcome this stationarity problem would be either to run the PCA on shocks in liquidity rather than levels, or to use a dynamic PCA specification.

[8] An alternative way of measuring commonality when using PCA is to study the magnitude of the co-variance matrix eigenvalues. The variation explained by each component vector is equal to the ratio of its corresponding eigenvalue to the sum of all eigenvalues. This method for liquidity commonality measurement was introduced by Hasbrouck and Seppi (2001) and has also been applied by e.g. Chen (2007), Kempf and Mayston (2008) and Kopp, Hütl, Loistl, and Prix (2008). I do not apply it here, as I study shocks of principal components, not the principal components directly.

[9] E.g. Kamara, Lou, and Sadka (2008) have looked at commonality in terms of the coefficient $\Phi_{l,\tau}$. When using dynamic estimation window PCA, I have a sign indeterminacy in the factors, making that regression coefficient unstable. It is also difficult to interpret the absolute size of $\Phi_{l,\tau}$. Hence, in my commonality test, I focus on R^2, which is unaffected by these problems. The problem of sign indeterminacy across liquidity measures was discussed for the static window PCA case by Korajczyk and Sadka (2008). In the dynamic estimation window setting the same problem is added in the time dimension.

τ		36			t			T			
h		1	2	3	1	2	3	1	2	3	$\bar{L}$
L_t^1	$Qspread$	0.11	0.19	0.26	0.16	0.22	0.26	0.15	0.21	0.27	0.23
L_t^2	$Espread$	0.12	0.21	0.27	0.20	0.24	0.27	0.20	0.25	0.28	0.29
L_t^3	$Turnover$	0.06	0.13	0.20	0.10	0.15	0.20	0.08	0.17	0.21	0.20
L_t^4	$ILLIQ$	0.08	0.16	0.21	0.15	0.21	0.25	0.16	0.22	0.25	0.14
L_t^5	Ψ	0.05	0.10	0.14	0.04	0.08	0.12	0.04	0.08	0.12	0.11
L_t^6	λ	0.05	0.10	0.15	0.04	0.08	0.12	0.04	0.08	0.12	0.07
L_t^7	$\tilde{\Psi}$	0.11	0.20	0.26	0.14	0.18	0.22	0.13	0.16	0.20	0.26
L_t^8	$\tilde{\lambda}$	0.05	0.10	0.15	0.04	0.08	0.11	0.04	0.08	0.11	0.07
r_t^m	Returns	0.11	0.18	0.24	0.10	0.17	0.23	0.10	0.16	0.23	0.22

Table 4.1: Average degree of commonality in liquidity

The table shows the degree of commonality for eight different measures of liquidity. Degree of commonality is defined as R^2 of regression run on Equation 4.4 run with a moving estimation window of 36 monthly observations. Numbers in the table are averaged across stocks and time (December 1997–December 2007). The estimation window size used for deriving systematic liquidity factors with PCA is given by τ, and the number of systematic liquidity factors considered is given by h. The same procedure is run using cross-sectional equal-weighted average liquidity, denoted $\bar{L}$. Ψ and λ are coefficients in the adverse selection cost function, with permanent influence on stock prices. $\tilde{\Psi}$ and $\tilde{\lambda}$ are coefficient in the inventory cost function and have a transitory impact on stock prices. See Equation 2.11 for details. The routine is repeated for returns, measuring to what degree common factors in returns can explain variation in individual stock returns.

ask whether a high degree of commonality is good. Does high commonality imply that the systematic liquidity measure is more accurate? In the regression commonality test I use systematic liquidity factors that have been estimated using different window specifications τ, but I evaluate them over the same time span. PCA is based on extraction of eigenvalues and eigenvectors of the co-variance matrix. Estimated co-variance matrix differences between short and long estimation windows can be due to either estimation error (sampling error) or actual variation over time in the underlying co-variance matrix. If the latter is true, the systematic liquidity factor utilising a short estimation window should be better suited to capture commonality. If the variation is due to noise; however, a long estimation horizon should yield higher accuracy. Based on this discussion, I argue that high commonality is desirable; and more specifically, that high commonality is a sign of accuracy in the systematic liquidity measurement.

The commonality regression as specified above allows me to investigate time dynamics of commonality for different liquidity measures and different measurement methods. I look at the dynamics over time below, but first I study overall differences between systematic liquidity measures by looking at commonality averaged both cross-sectionally and over time. These results are presented in Table 4.1.

To get an intuition of what different magnitudes of commonality imply I include the return measure in the table. It is well known that the market as an aggregate to a large extent drives returns in individual stocks. Using exactly the same methodology as for liquidity measures, I find that the degree of commonality in returns at $h = 3$ is 20 to 25% (i.e. variation in common factors of stock prices can explain more than 20% of the variation in individual stock price

variations). For the bid-ask spread measures, $Turnover$, $ILLIQ$, and the fixed (transitory) inventory cost ($\tilde{\Psi}$), I find commonality of about the same degree as for returns. The other measures register lower degrees of commonality. Before turning to differences across τ, which is my main interest, it is interesting to compare my results to some previous studies. In the case of returns, a commonality of the same level has been found by both Hasbrouck and Seppi (2001) and Korajczyk and Sadka (2008). Commonality tests of previous literature are not directly comparable to my results, as specification of commonality tests, liquidity variables, as well as data frequencies differ. Regardless, a brief review of their findings is useful. Chordia, Roll, and Subrahmanyam (2000) find commonality in quoted and effective spreads. Kamara, Lou, and Sadka (2008) identify commonality in the illiquidity measure by Amihud. Korajczyk and Sadka (2008) perform level commonality tests on the same measures as I use. They find high commonality in the bid-ask spread (quoted and effective), $ILLIQ$, and $\tilde{\Psi}$; and low commonality in Ψ, λ and $\tilde{\lambda}$. This corresponds well to my results, although the magnitudes of commonality they find are higher due to their commonality specification. Their finding of weak commonality in $Turnover$ is not reflected in my results.

The differences across my specifications of PCA with regard to commonality are on average small, but some tendencies can be seen. The degree of commonality found in the first four liquidity measures is slightly higher when using an expanding window rather than a moving window. For the four price impact coefficients the tendency is the opposite. In accordance with my interpretation above, the lower degree of commonality registered for the moving window factors for the first four measures means that the underlying market liquidity is disturbed by sampling error when a short estimation window is applied. If the co-variance matrix is time-varying that sampling error should be counter-acted by an ability to capture that variation. For the four price impact coefficients, such short-term effects appear to be present, as the moving window consistently generates higher commonality than the expanding window specification.

I also run the commonality test using the cross-sectional equal-weighted average liquidity ($\bar{L}$), which is another popular proxy for systematic liquidity (Chordia, Roll, and Subrahmanyam 2000, Kamara, Lou, and Sadka 2008). This can be expressed as

$$\Delta_{AR} L_{i,l}^* = \phi_{l,L} + \Phi_{l,L} \Delta_{AR} \bar{L}_l + u_{i,l,L}, \tag{4.5}$$

where L replaces τ as subscript to distinguish the notation from Equation 4.4 and $\Phi_{l,L}$ is a scalar showing the sensitivity to mean liquidity. Results of this regression are given in the rightmost column of Table 4.1. Performance of this average liquidity approximation of the systematic liquidity factor varies widely across liquidity measures. Comparing it to PCA with the first three factors considered ($h = 3$), it appears unsuited to capture underlying market liquidity in terms of $ILLIQ$ and $Qspread$ as well as some of the price impact measures. In

the case of $ILLIQ$, I believe that the low commonality is due to that the mean is driven by some large outlier observations, particularly in the first half of the sample.

The static window PCA ($\tau = T$) has often been used in previous systematic liquidity literature (e.g. Hasbrouck and Seppi 2001, Korajczyk and Sadka 2008). With this method the full time sample is considered in the estimation, implying that future information is involved in the systematic factors that I evaluate in the commonality test. When looking at liquidity in retrospect this is not a problem, but if these factors are applied to explain (or even forecast) asset prices, there is a forward-looking bias.[10] I note here that the static window PCA factor registers degrees of commonality similar to those of the expanding window. If the methods can be regarded as equivalent in terms of outcome, it would be appropriate to use expanding window PCA rather than static, but before concluding this I have to see if the findings are consistent over time.

4.3.1 Dynamics in Liquidity Commonality

I now turn to the commonality dynamics over time. Figure 4.1 displays the time dynamics of commonality for different τ (within each panel) and different liquidity measures (one in each panel). The only difference between the curves within each panel is the horizon used when estimating systematic liquidity factors.[11] Commonality depicted is based on the three first systematic liquidity factors and regressions run on 36 observations.

The graphs reveal that commonality varies over time, and for some liquidity measures it varies substantially. The largest variations are seen in the spread measures and the transitory fixed cost measure ($\tilde{\Psi}$). For these measures the period chosen for measuring commonality will matter substantially for what answer is retrieved. In many previous studies of commonality, measurement data have been chosen by availability, making the point of measurement arbitrary. Furthermore, the graphs uncover some cases of substantial differences between measurement methodologies. The difference between the most commonly used proxies of systematic liquidity, static window PCA and average liquidity, can be large (differences of > 0.1 in degree of commonality can be seen temporarily for most liquidity measures).

I now turn to analysis of the dynamics of the individual liquidity measures. For $ILLIQ$ I see that the expanding window consistently yields higher commonality than the moving window. This implies that there are long term co-variances between asset liquidities that are not revealed when using a short estimation window, and that these are more important than potential short term changes in the co-variance matrix. For $ILLIQ$, average liquidity appears to be a consistently worse predictor of firm liquidity than PCA proxies. This is notable as $\bar{L}$ was the proxy used in a recent study of commonality in $ILLIQ$ (Kamara, Lou,

---

[10]Several studies have used the static window PCA for such applications.

[11]This implies that the moving and the expanding window PCA factors are by definition the same in the first time period at $t = 36$ (Dec. 1997, not shown), and that the expanding and static window factors are by definition the same in the last period (Dec. 2007).

Panel A: *Qspread*

Panel B: *Espread*

Figure 4.1: Dynamics in systematic liquidity pricing power (see caption on p.71)

Panel C: *Turnover*

Panel D: *ILLIQ*

Figure 4.1: Dynamics in systematic liquidity pricing power (see caption on p.71)

Panel E: Permanent Fixed Cost (Ψ)

Panel F: Permanent Variable Cost (λ)

Figure 4.1: Dynamics in systematic liquidity pricing power (see caption on p.71)

Panel G: Transitory Fixed Cost ($\tilde{\Psi}$)

— Mov. PCA — Exp. PCA — Static PCA — Mean

Panel H: Transitory Variable Cost ($\tilde{\lambda}$)

— Mov. PCA — Exp. PCA — Static PCA — Mean
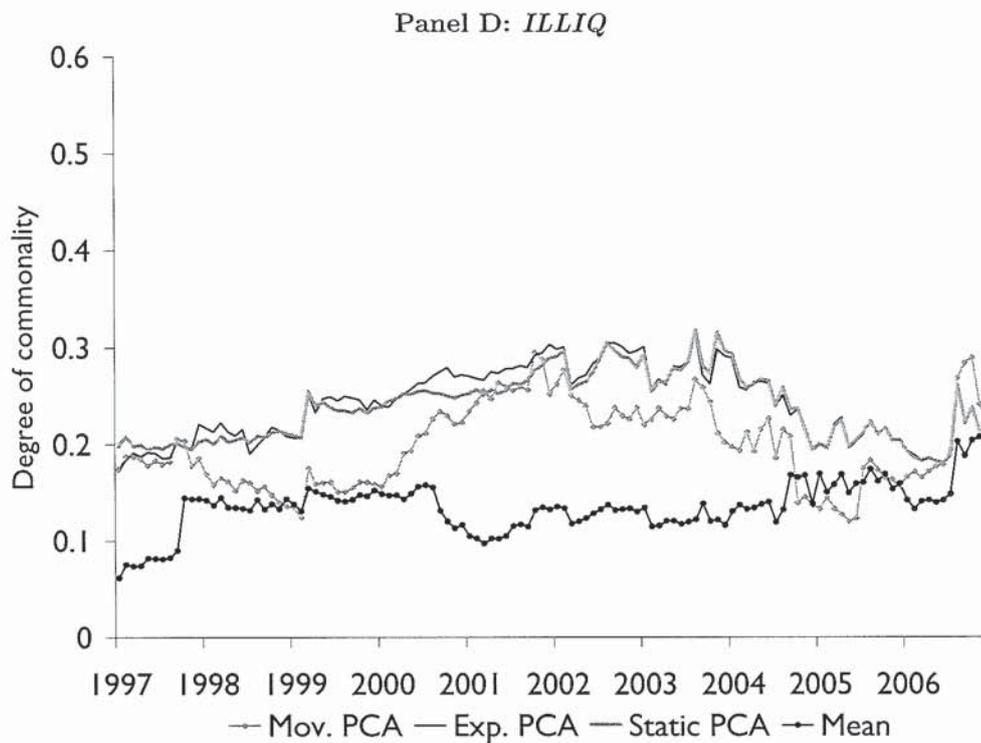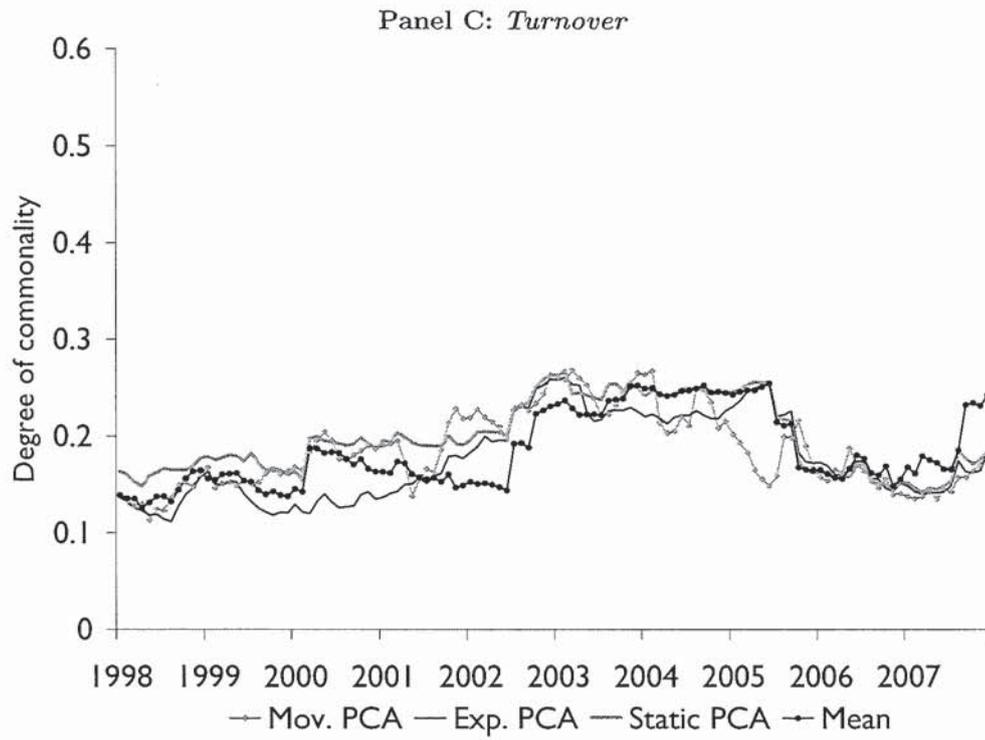
Figure 4.1: Dynamics in systematic liquidity pricing power (see caption on p.71)

Figure 4.1: Commonality dynamics

In each panel, time dynamics of the degree of commonality is graphed for 1998–2007. Degree of commonality displayed is the R^2 value of Equation 4.4 estimated with $h = 3$. Commonality measures are averages across stocks, estimated for the period (t-35):t. In each panel, the curves represent different settings of $\tau = [36, t, T]$, denoted *Mov. PCA, Exp. PCA, Static PCA* in the legend for each panels. Also, commonality as predicted by the cross-sectional average is given, estimated as in Equation 4.5. This curve is denoted *Mean*. Each panel represents commonality in different liquidity measures: A) Quoted spread (*Qspread*); B) Effective spread (*Espread*); C) *Turnover*; D) Amihud's (2002) illiquidity measure (*ILLIQ*); E) Permanent fixed cost (Ψ); F) Permanent variable cost (λ); G) Transitory fixed cost ($\bar{\Psi}$); H) Transitory variable cost ($\bar{\lambda}$). The last four measures are based on Sadka's (2006) implementation of the price impact regression from Glosten and Harris (1988).

and Sadka 2008). As noted above, a reason for the poor performance of the average liquidity is that it can be driven by large outliers. For *Turnover*, the degree of commonality is relatively stable over time and no consistent differences between systematic liquidity measures are seen. The average liquidity proxy can hence be used without loss of precision. For the two measures of the bid-ask spread, on the other hand, I find significant variation over time. The spike in commonality in March 2000 coincides with the burst of the dot-com bubble. It is interesting to see that this causes a sharp increase in commonality of spreads. This can be related to discussion by e.g. Amihud, Mendelson, and Wood (1990) about strong liquidity co-movement in times of crisis. For the spread measures, no recommendation on which systematic liquidity measurement technique that should be used can be given. There appear to be important short-term trends captured by the moving window PCA during parts of the sample, but during other periods this method is less efficient in capturing trends appearing in the long term sample.

For the price impact liquidity measures, the highest commonality is consistently recorded by the moving window PCA. This implies time-varying co-variance matrices — revealing significant short term trends. In general I find low commonality in the price impact measures, except for the transitory fixed cost $\bar{\Psi}$, which is exactly the same finding as in Korajczyk and Sadka (2008). This is the inventory cost that the market maker has to carry. When it is large the market maker has to charge wider spreads in order to avoid losses. Systematic liquidity in terms of $\bar{\Psi}$ is also well proxied by its cross-sectional average. Sadka (2006) and Glosten and Harris (1988) find that this transitory fixed cost ($\bar{\Psi}$) along with the permanent variable cost (λ) (i.e. the volume related adverse selection cost) are more important than the other two price impact measures, and the magnitudes shown in Section 3 point in the same direction for my study. My commonality regressions; however, which are more about variation than magnitudes in the underlying liquidity measures, show that the other price impact measures are as efficient as the permanent variable cost in explaining liquidity variation. This is another indication of that these (Ψ and $\bar{\lambda}$) are non-zero.

In general, the expanding window PCA and the static window PCA yield very similar degrees of commonality. In most cases they have converged around 2003, implying that an

eight year window is enough to capture long term trends in systematic liquidity. If such a time frame is available, the forward-looking bias of static window PCA can be avoided by applying the expanding window specification, as was done by Chen (2007), without loss of accuracy.

4.3.2 Correlation Between Estimated Commonality Series

Although similar on average, commonality series for one liquidity measure often have quite different time-series properties, i.e. are only weakly correlated, depending on the method used for systematic liquidity estimation. This is especially so for the cross-sectional mean when compared to the PCA-based estimates of systematic liquidity. Table 4.2 shows correlations for the estimated time-series of commonality for each liquidity measure (i.e. correlations between the series graphed in Figure 4.1).

	Mov–Exp	Mov–Stat	Mov–Mean	Exp–Stat	Exp–Mean	Stat–Mean
$Qspread$	0.52	0.72	0.43	0.82	0.91	0.77
$Espread$	0.43	0.62	0.37	0.81	0.90	0.82
$Turnover$	0.75	0.80	0.61	0.86	0.81	0.82
$ILLIQ$	0.70	0.68	-0.08	0.96	-0.01	-0.09
Ψ	0.68	0.33	0.33	0.31	0.07	0.46
λ	0.38	0.52	0.29	0.90	0.00	0.30
$\bar{\Psi}$	0.48	0.58	0.30	0.91	-0.03	-0.05
$\bar{\lambda}$	0.09	0.15	0.33	0.84	0.45	0.21

Table 4.2: Correlation between estimated commonality

For each liquidity measure, Pearson correlation coefficients between pairs of commonality series in terms of R^2 of the models described by Equations 4.4 and 4.5 are calculated. Each series contains monthly observations from December 1997 to December 2007. *Mov* is based on Equation 4.4 and $\tau = 36$; *Exp* is based on Equation 4.4 and $\tau = t$; *Stat* is based on Equation 4.4 and $\tau = T$; *Mean* is based on Equation 4.5. Ψ and λ are coefficients in the adverse selection cost function, with permanent influence on stock prices. $\bar{\Psi}$ and $\bar{\lambda}$ are coefficient in the inventory cost function and have a transitory impact on stock prices. See Equation 2.11 for details.

Apparently, the liquidity measures can be divided in two groups. The first group contains quoted spread, effective spread, and *Turnover*. For this group, correlations are relatively high for all estimation methods. The lowest correlation, 0.37, is between the cross-sectional mean and the moving window (for *Espread*) and the highest correlation 0.91 between the expanding and static window (for *Qspread*). For this group of measures, time series properties of systematic liquidity are in general similar across estimation methods, implying that average liquidity is a good proxy of systematic liquidity in terms of these measures.

The second group contains the price impact measures: $ILLIQ$, permanent market maker costs (adverse selection costs; Ψ and λ), and transitory market maker costs (inventory costs; $\bar{\Psi}$ and $\bar{\lambda}$). For this group, correlations vary considerably. Correlations between PCA-based estimates are in most cases high, whereas correlations between the cross-sectional average and the PCA-based estimates are much lower. For $ILLIQ$ estimated by the cross-sectional

mean, point estimates of correlation with the static, expanding and moving window are in all cases negative: -0.09, -0.01 and -0.08, respectively. In other words, the cross-sectional average estimate of systematic liquidity produces a commonality with very different time-series properties than the PCA-based estimates of systematic liquidity. Given the fact that the cross-sectional mean yields an average commonality that is much lower than PCA, this suggests that the time-series properties of systematic liquidity are not fully captured by the cross-sectional mean for $ILLIQ$. This is interesting as many studies use the cross-sectional average to estimate systematic liquidity in terms of $ILLIQ$ (see Kamara, Lou, and Sadka 2008).

Turning to the price impact measures of Sadka (2006), similar results hold, i.e. commonality estimated using the cross-sectional mean is often in principle uncorrelated with commonality estimated using different specification of PCA. For example, for the fixed inventory costs measure, which produces the highest commonality among Sadka's (2006) four measures, correlations between commonality for the cross-sectional average and the PCA-based estimates of systematic liquidity are -0.05, -0.02 and 0.30. This is in spite of average commonality being at the same level for PCA- and mean-based measurement. Again, this indicates that time-series properties of estimated systematic liquidity are very different depending on the method used for estimation.

4.4   Systematic Liquidity Factors and Stock Prices

Above I argue that commonality is an indicator of liquidity factor measurement accuracy. A different matter is whether the liquidity factor is useful for explaining stock prices. As it is well-known that individual stock liquidity affects prices (Amihud, Pedersen, and Mendelson 2005) it is sensible to believe that if systematic liquidity explains stock liquidity, it should to some extent be priced in the cross-section of stocks. This hypothesis has been tested and confirmed by Pastor and Stambaugh (2003) and Korajczyk and Sadka (2008), and is related to the second type of liquidity risk in the theoretical framework by Acharya and Pedersen (2005).

In the preceding section I study correlations between the commonality time series retrieved with different systematic liquidity derivation techniques. High correlations between these series would indicate that they also have similar abilities in explaining stock returns. In my investigation above I find many correlations being low, implying fundamentally different time series properties of the underlying market factors. If that is the case, the different systematic liquidity estimation techniques may lead to differences in the ability to explain stock returns too. This is the topic of investigation in the current section.

I adopt an extended market return model to evaluate the explanatory power of my sys-

tematic liquidity measures. I let

$$r_{i,t} = \pi_{i,l,\tau,1} + \pi_{i,l,\tau,2} r_{M,t} + \Pi_{i,l,\tau} \Delta_{AR} \hat{F}_{1:h,l,\tau} + v_{i,l,\tau,t}, \tag{4.6}$$

where $r_{i,t}$ is the return of stock i in time t, and $r_{M,t}$ is the market return in time t. Intercept and market return sensitivity are denoted $\pi_{i,l,\tau,1}$ and $\pi_{i,l,\tau,2}$ respectively, and return sensitivities to systematic liquidity are given by the coefficient vector $\Pi_{i,l,\tau}$ (which is of length h). Returns unexplained by the market returns and systematic liquidity factors are put in the residual term $v_{i,l,\tau,t}$. As indicated by the indexes l and τ, the regressions are repeated for the eight liquidity measures and the three different PCA-based systematic liquidity measurement techniques. I set $h = 3$. In the same fashion as above, I run the regression using a 36 months moving estimation window. From these regressions I am primarily interested in the explanatory power of the systematic liquidity factors with respect to returns. Hence, I record the R^2 of each run of the regression and take the cross-sectional average, which gives me time series of the explanatory power of liquidity. To isolate the effect of liquidity, I run three versions of this model: (I) full regression; (II) setting all elements of $\Pi_{i,l,\tau}$ to zero; (III) setting $\pi_{i,l,\tau,2} = 0$. In Table 4.3 I present averages across time of the full model (I) in Panel A; the excess explanatory power of the liquidity factor [(I)–(II)] in Panel B; and finally the explanatory power of the liquidity factor alone (III) in Panel C.[12] Also, as with the commonality regressions I consider the alternative relation

$$r_{i,t} = \pi_{i,l,L,1} + \pi_{i,l,L,2} r_{M,t} + \Pi_{i,l,L} \Delta_{AR} \bar{L}_l + \epsilon_{i,l,L,t}, \tag{4.7}$$

simply replacing the systematic liquidity factor with the cross-sectional mean liquidity and changing the subscripts to L. As L only has one dimension, $\Pi_{i,l,L}$ is here a scalar.

The results in Table 4.3 show that the systematic liquidity factor estimated using PCA can have a substantial explanatory power for stock returns. When studied in isolation (Panel C), it ranges from 10.9% up to 16.4%. When accounting for the market return factor (Panel B), the excess explanatory power of systematic liquidity ranges from 7.9% to 10.9%. Cross-sectional mean liquidity has a substantially lower explanatory power regardless of what liquidity measure I look at.

The results for the static window PCA systematic liquidity measure are, as was found for commonality, similar to those of the expanding window PCA. The difference between moving window and expanding window estimation for systematic liquidity is in general small, but some cases where they differ can be noted. For $ILLIQ$, the expanding window is slightly better than the moving, which is in accordance with my findings on commonality in Section 4.3. Among the price impact regression coefficients, an expanding estimation window specification

---

[12]As mentioned in conjunction to the commonality tests, I am unable to interpret the size and the sign of $\Pi_{i,l,\tau}$ due to the sign indeterminacy of PCA factors. This is the reason that I focus on R^2 values.

Panel A: Average R^2 of market return and liquidity factor

l		$\tau = 36$	$\tau = t$	$\tau = T$	L
1	$Qspread$	0.286	0.286	0.284	0.235
2	$Espread$	0.287	0.287	0.289	0.235
3	$Turnover$	0.291	0.290	0.293	0.234
4	$ILLIQ$	0.292	0.296	0.293	0.236
5	Ψ	0.290	0.290	0.289	0.239
6	λ	0.283	0.290	0.292	0.238
7	$\tilde{\Psi}$	0.313	0.306	0.306	0.241
8	$\tilde{\lambda}$	0.285	0.291	0.284	0.234

Panel B: Average R^2 of liquidity factor in excess of market return

l		$\tau = 36$	$\tau = t$	$\tau = T$	L
1	$Qspread$	0.081	0.082	0.080	0.031
2	$Espread$	0.083	0.083	0.085	0.031
3	$Turnover$	0.087	0.086	0.088	0.030
4	$ILLIQ$	0.087	0.091	0.089	0.032
5	Ψ	0.086	0.085	0.085	0.035
6	λ	0.079	0.086	0.088	0.034
7	$\tilde{\Psi}$	0.109	0.102	0.102	0.037
8	$\tilde{\lambda}$	0.081	0.087	0.080	0.029

Panel C: Average R^2 of liquidity factor alone

l		$\tau = 36$	$\tau = t$	$\tau = T$	L
1	$Qspread$	0.120	0.129	0.127	0.054
2	$Espread$	0.128	0.128	0.124	0.059
3	$Turnover$	0.123	0.109	0.119	0.036
4	$ILLIQ$	0.112	0.117	0.114	0.038
5	Ψ	0.131	0.129	0.120	0.044
6	λ	0.111	0.122	0.125	0.063
7	$\tilde{\Psi}$	0.164	0.143	0.135	0.107
8	$\tilde{\lambda}$	0.127	0.140	0.154	0.038

Table 4.3: Systematic liquidity factor ability to explain returns

Panel A shows, for each liquidity measure l, R^2 averaged across stocks and across time (December 1997–December 2007) for the model described in Equation 4.6 and 4.7; Panel B shows R^2 averaged across stocks and across time for the model described in Equation 4.6 and 4.7 minus the corresponding number from the same model but with all elements of $\Pi_{i,l,\tau}$ equal to 0; Panel C shows R^2 averaged across stocks and across time for the model described in Equation 4.6 and 4.7 with $\pi_{i,l,\tau,2} = 0$. τ indicates how the estimation window is specified in the systematic liquidity derivation in Equation 4.4. Ψ and λ are coefficients in the adverse selection cost function, with permanent influence on stock prices. $\tilde{\Psi}$ and $\tilde{\lambda}$ are coefficient in the inventory cost function and have a transitory impact on stock prices.

appears to be better for variable costs, whereas a moving window is better for fixed costs.

The spread measures yielded highest commonality of all measures, but do not stand out as strong measures in terms of explanatory variables for returns. Hence, high commonality does not necessarily imply high explanatory power of returns. The measure that recorded the second highest commonality, transitory fixed costs ($\tilde{\Psi}$), has much higher pricing power than all the other liquidity measures (10.9% in Panel B and 16.4% in Panel C, when using moving window specification). Sadka (2006) argues that out of the four coefficients, $\tilde{\Psi}$ and λ should be most important for pricing, but my results show that they are all on par with the other liquidity measures in this regard. Again, the co-variation rather than the magnitude of the underlying liquidity measure is what matters for the investigation.

4.4.1 Dynamics in Systematic Liquidity Pricing Power

Following the same disposition as in the previous section, I now turn to time series properties of the systematic liquidity pricing regression results. Figure 4.2 shows the dynamics of systematic liquidity pricing power over time for different τ (within each panel) and for different liquidity measures (one in each panel). The statistic displayed is the excess R^2 achieved when adding systematic liquidity to the market model, i.e. average R^2 from model (I) minus average R^2 from model (II) (the series covered in Panel B of Table 4.3).

It is clear from the figure that the PCA-based methods for deriving systematic liquidity factors yield factors that are better than the equal-weighted average in explaining stock returns. The difference between PCA-based methods and the average is consistent over time and across all liquidity measures tested. This is the dominant difference between approaches to systematic liquidity estimation — no consistency in differences between PCA-based methods is seen.

The pricing effects of the spread measures and for $ILLIQ$ (Panels A, B and D) are virtually constant over time. This is an important difference from the pattern seen in commonality (see Figure 4.1), where commonality is much higher from the beginning of 2000 to the beginning of 2003, compared to the rest of the sample. Interestingly, the pricing power of the transitory fixed cost ($\bar{\Psi}$) is increasing sharply in the beginning of 2000 and stays high until the beginning of 2003 (see Panel G). This measure captures the part of the bid-ask spread that is associated with inventory cost.

In my analysis above I established that the two different dynamic estimation window PCA specifications on average resulted in systematic liquidity factors with rather similar explanatory power on returns. If the co-variance structure is time-varying, the gap between moving PCA and other methods should get wider over time. For most of the liquidity measures, there is no such pattern, and no other consistent difference between PCA methods either. There is, however, one exception. For $\bar{\Psi}$, the graph in Panel G shows clearly that the difference

Panel A: *Qspread*

— Mov. PCA — Exp. PCA — Static PCA — Mean

Panel B: *Espread*

— Mov. PCA — Exp. PCA — Static PCA — Mean

Figure 4.2: Dynamics in systematic liquidity pricing power (see caption on p.81)

Panel C: *Turnover*

Panel D: *ILLIQ*

Figure 4.2: Dynamics in systematic liquidity pricing power (see caption on p.81)

Panel E: Permanent Fixed Cost (Ψ)

Panel F: Permanent Variable Cost (λ)

Figure 4.2: Dynamics in systematic liquidity pricing power (see caption on p.81)

Panel G: Transitory Fixed Cost ($\tilde{\Psi}$)

Panel H: Transitory Variable Cost ($\tilde{\lambda}$)
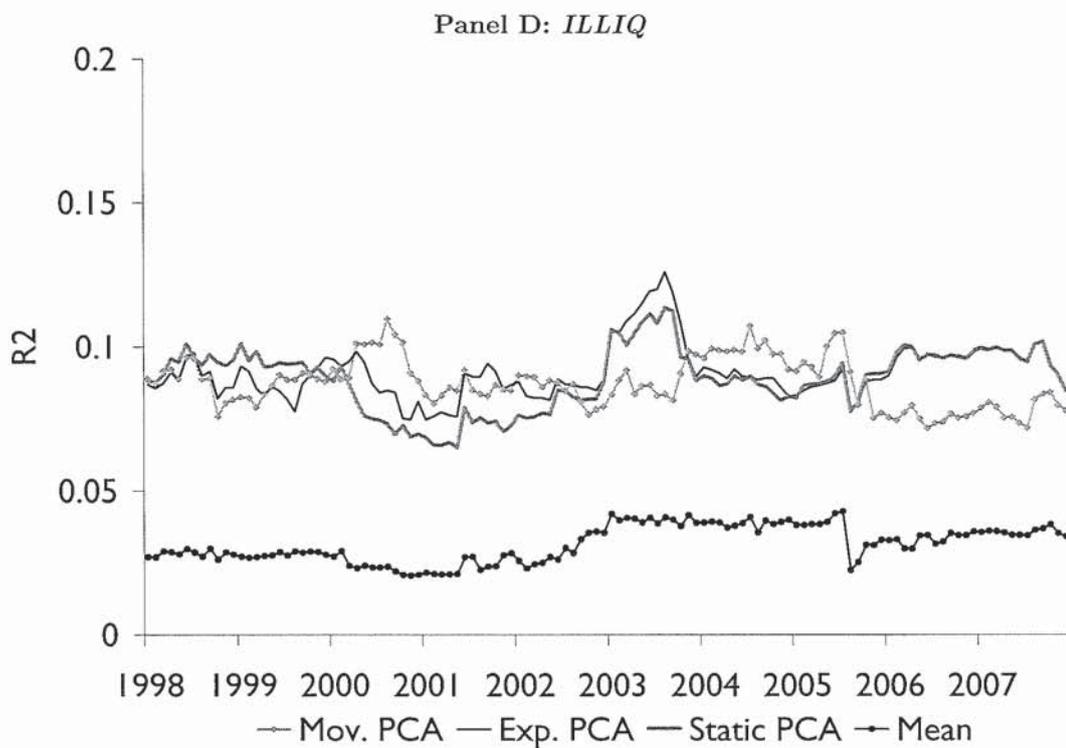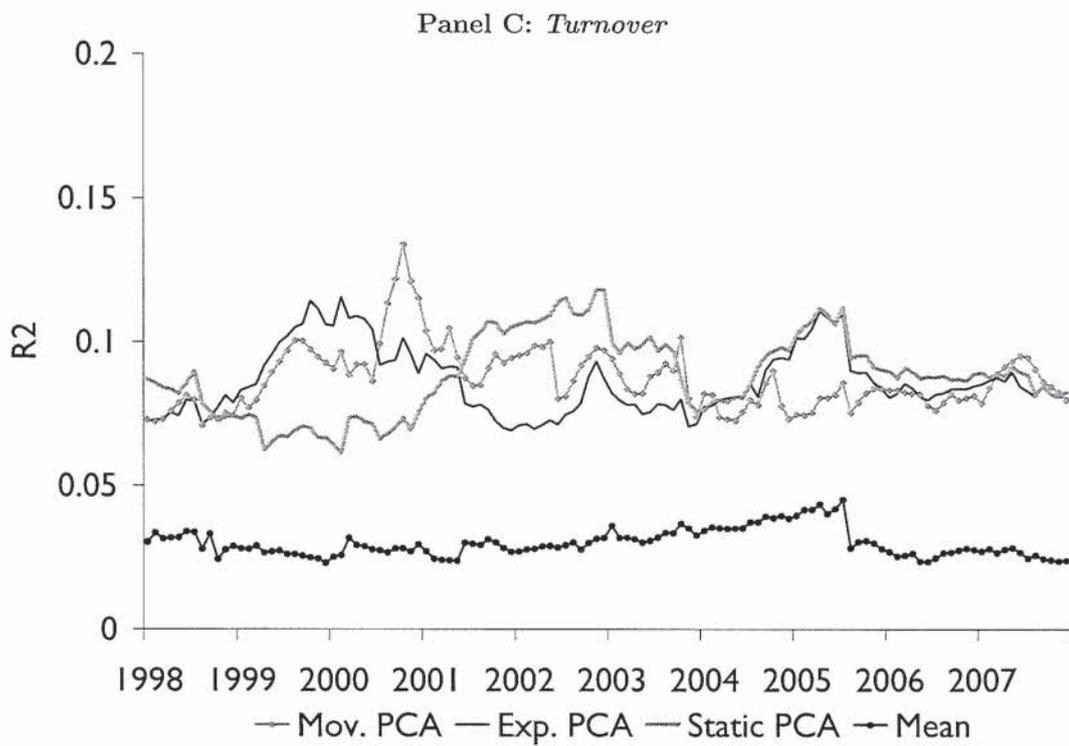
Figure 4.2: Dynamics in systematic liquidity pricing power (see caption on p.81)

Figure 4.2: Dynamics in systematic liquidity pricing power

In each panel, time dynamics of the liquidity pricing power is graphed for 1998–2007. The curves represent different settings of $\tau = [36, t, T]$, denoted *Mov. PCA, Exp. PCA, Static PCA* in the legend, referring to the estimation method used for derivation of systematic liquidity factors. Also, systematic pricing power as predicted by the cross-sectional liquidity average is given, estimated as in Equation 4.7. This curve is denoted *Mean*. Liquidity pricing power displayed is in excess of market returns. That is calculated as R^2 averaged across stocks for the model described in Equation 4.6 and 4.7 minus the corresponding number from the same model but with all elements of $\Pi_{i,l,\tau}$ and $\Pi_{i,l,L}$ equal to 0. The equations are estimated with $h = 3$, meaning that the first three systematic liquidity factors are used. The systematic liquidity pricing power for time t is estimated using the estimation window $(t - 35) : t$. Each panel presents different liquidity measures: A) Quoted spread (*Qspread*); B) Effective spread (*Espread*); C) *Turnover*; D) Amihud's (2002) illiquidity measure (*ILLIQ*); E) Permanent fixed cost (Ψ); F) Permanent variable cost (λ); G) Transitory fixed cost ($\tilde{\Psi}$); H) Transitory variable cost ($\tilde{\lambda}$). The last four measures are based on Sadka's (2006) implementation of the price impact regression from Glosten and Harris (1988).

on average between expanding and moving windows is due to the latter half of the sample. Hence, moving window PCA appears to be the appropriate choice for this measure, whereas a longer estimation window may be more appropriate for the other measures.

The small differences that are seen between expanding window PCA and full sample PCA, all seem to be due to pre-2003 observations, reinforcing the view that the two methods converge after approximately 8 years. For long data sets, an 8 year moving window specification may be preferred, as the computational burden of expanding window PCA is increasing with time.

4.4.2 Correlation Between Systematic Liquidity Pricing Power Series

As I did for commonality time series in the previous section, I now look at correlation between time series on ability to explain returns. These correlations are presented in Table 4.4. In Panel A, correlations between R^2 in excess of the market returns are given; and in Panel B correlations between R^2 series of models estimated using liquidity factors as the only explanatory variable are presented.

Similar to the case of commonality, the correlation between the moving window and expanding window PCA specifications is in general high and positive. I interpret this as a sign that the two methods explain the same pricing patterns. For *ILLIQ*, *Turnover* and permanent variable costs (λ) that correlation is lower, which can be interpreted as meaning that the two methods capture different aspects of return variation. That implies that there are both short and long term trends that contribute to the return data generating process, captured by moving and expanding estimation windows respectively. When comparing PCA specifications to the mean liquidity, correlations are in many cases very low or even negative. This shows that mean liquidity explains other pricing patterns than the PCA measures.

The correlation between the static window PCA and expanding window PCA in terms of pricing power is again found to be high. In the previous section I found that the two methods

Panel A: Systematic liquidity pricing effect in excess of market returns

l		Mov–Exp	Mov–Stat	Mov–Mean	Exp–Stat	Exp–Mean	Stat–Mean
1	$Qspread$	0.67	0.48	0.20	0.73	0.42	0.55
2	$Espread$	0.57	0.48	0.19	0.67	0.32	0.35
3	$Turnover$	0.29	-0.12	-0.31	-0.31	0.10	0.41
4	$ILLIQ$	-0.19	-0.30	0.04	0.77	0.57	0.63
5	Ψ	0.76	0.73	0.56	0.85	0.39	0.58
6	λ	0.31	0.34	0.06	0.80	0.80	0.80
7	$\tilde{\Psi}$	0.85	0.89	0.85	0.95	0.93	0.93
8	$\tilde{\lambda}$	0.55	0.54	0.50	0.82	0.77	0.55

Panel B: Systematic liquidity pricing effect alone

l		Mov–Exp	Mov–Stat	Mov–Mean	Exp–Stat	Exp–Mean	Stat–Mean
1	$Qspread$	0.86	0.86	0.34	0.95	0.50	0.54
2	$Espread$	0.53	0.46	0.49	0.20	0.71	-0.20
3	$Turnover$	0.24	0.56	-0.08	-0.06	0.12	0.03
4	$ILLIQ$	0.44	0.24	0.27	0.75	0.38	0.17
5	Ψ	0.88	0.81	-0.18	0.89	-0.21	0.00
6	λ	0.41	0.41	0.48	0.96	0.48	0.48
7	$\tilde{\Psi}$	0.14	0.48	0.34	0.83	0.86	0.78
8	$\tilde{\lambda}$	0.74	0.77	-0.09	0.89	0.01	-0.18

Table 4.4: Systematic liquidity factor pricing power: correlation between estimation techniques

For each liquidity measure l, Pearson correlation coefficients between pairs of R^2 series of the models described by Equations 4.6 and 4.7 are calculated. Panel A shows correlations between R^2 averaged across stocks for the model described in Equation 4.6 and 4.7 minus the corresponding number from the same model but with all elements of $\Pi_{i,l,\tau}$ equal to zero; Panel B shows correlations between R^2 averaged across stocks for the model described in Equation 4.6 and 4.7 with $\pi_{i,l,\tau,2} = 0$. Each series contains monthly observations from December 1997 to December 2007. Mov is based on Equation 4.6 and $\tau = 36$; Exp is based on Equation 4.6 and $\tau = t$; $Stat$ is based on Equation 4.6 and $\tau = T$; $Mean$ is based on Equation 4.7. Ψ and λ are coefficients in the adverse selection cost function, with permanent influence on stock prices. $\tilde{\Psi}$ and $\tilde{\lambda}$ are coefficient in the inventory cost function and have a transitory impact on stock prices. See Equation 2.11 for details.

also achieved equivalent levels of explanatory power. Hence, I conclude that the expanding window PCA is a good substitute for the static window PCA, which has the important drawback that it uses future information for its estimation.

4.5   Concluding Remarks

I improve the measurement of systematic liquidity, a factor describing market variation of liquidity that has been shown to be an important risk factor in asset pricing. Traditionally systematic liquidity is either proxied by the market average liquidity or derived using PCA. By running PCA it is implicitly assumed that the co-variance matrix of cross-sectional liquidity is constant. As this assumption may be unrealistic, I introduce PCA with dynamic estimation window, which re-estimates the co-variance matrix for each period of time. This can be done either with a moving or an expanding estimation window. Which one is appropriate depends on the properties of the co-variance matrix over time. I assess these two specifications of dynamic estimation window PCA together with the two traditional systematic liquidity estimation techniques.

The evaluation of the four estimation techniques is run in terms of (A) ability to explain cross-sectional stock liquidity and (B) ability to explain cross-sectional stock returns. The first criterion is the explicit purpose of systematic liquidity factors, whereas the second is important for its application in asset pricing. For both cases, I find support for using the dynamic estimation window specification of PCA.

Which measurement technique that is most appropriate for systematic liquidity is dependent on what liquidity measure is considered. For measuring illiquidity as defined by Amihud (2002), a long time frame is beneficial both in terms of liquidity and return variation, so an expanding window PCA estimation is appropriate. For *Turnover* and bid-ask spread measures of liquidity, no single measurement technique dominate the other when measuring commonality in liquidity. The simplest possible proxy, such as a cross-sectional average can therefore be used. For explaining stock returns; however, all the PCA-based measures are substantially better. For price impact coefficients as estimated by Sadka (2006), the highest commonality is retrieved when using a moving window to estimate systematic liquidity, implying a time-varying co-variance matrix. The most important factor from Sadka's price impact regression, both in terms of liquidity and return variability, is the transitory fixed cost variable, which is associated with the inventory cost that a liquidity provider carries. Systematic liquidity in this variable is best estimated using a moving window PCA. Using that methodology, 26% of the cross-sectional variation in liquidity can be explained, and 16% of the cross-sectional variation in asset returns. This is higher than what I find for all the other liquidity measures tested.

My investigations also show that the expanding window PCA yield results equivalent

to those of the traditional static window PCA, both for explaining asset liquidity and for explaining asset returns. Hence, where possible, expanding window PCA should replace static window PCA, as the latter utilises future information for its estimation. Preliminary evidence shows that the two methods converge after 8 years (96 observations). Hence, a moving window PCA with 8 years of observations may be an appropriate method when no time-variation in co-variances is expected. This would reduce the computational burden of PCA in long time series.

Part III

Portfolio optimisation

Chapter 5

Introduction to Full-Scale
Optimisation

Portfolio selection problems can in general be described as utility maximisation problems where utility is dependent on expected portfolio return distributions. A portfolio selection problem with N assets is described in Equation 5.1. The utility function U is the objective function, and the vector θ with dimensions $(N \times 1)$ contains the decision variables — the portfolio weights. The decision variables are subject to constraints given by the matrix Ω. Typically, Ω contains a budget constraint ($\theta' \iota = 1$, where ι is a vector of ones). Other constraints may include e.g. upper and lower bounds on allocations, borrowing restrictions, and Value at Risk restrictions. The optimal allocation vector $\theta^\star$ is the vector that maximises the expected utility function while adhering to these constraints.

$$\theta^\star = \arg \max_\theta U$$
$$\theta \in \Omega \tag{5.1}$$

The utility function U describes the investor's preferences with respect to the trade-off between returns and risk of investments. Typically, the investor defines risk as some function of the probability distribution function (PDF) of portfolio returns. When maximising expected utility, the investor may consider a matrix R of different outcomes (returns) for each admissible asset ($n = 1, ..., N$) in S different scenarios (giving R dimension $N \times S$). As portfolio returns are portfolio allocations multiplied by asset returns, $\theta'R$ will be a vector of length S that forms a probability distribution of portfolio returns for each θ.

The portfolio choice problem can hence be summarised in three sub-problems: (a) formulating investor preferences in an utility function; (b) determining the PDF of portfolio returns; and (c) finding the optimal portfolio allocation vector θ. In this chapter I first introduce two portfolio choice frameworks, mean-variance (MV) and full-scale optimisation (FSO),

and compare them in terms of the three sub-problems (Sections 5.1–5.2). Then, I go into further detail on each of the three sub-problems (Sections 5.3–5.5). Overall, the purpose of this chapter is to lay the foundations for the subsequent empirical chapters on FSO.

5.1 Mean-Variance Optimisation

The portfolio choice literature was founded with the seminal work by Markowitz (1952, 1959) that introduced the MV model. Based on the assumption that the utility function is concave, i.e. that investors are risk averse, Markowitz found that for a wide range of returns, a quadratic utility function can serve as an approximation of investor preferences. Using that approximation the expected utility maximisation can be solved analytically for a set level of risk aversion, η. The utility function is then expressed as

$$U_q = r_p - \eta \sigma_p^2, \tag{5.2}$$

where r_p is the portfolio return and σ_p^2 its variance, which can be expressed in terms of the stock return co-variance matrix Σ,

$$\sigma_p^2 = \theta' \Sigma \theta. \tag{5.3}$$

In terms of the three sub-problems sketched above, by specifying the utility function as quadratic Markowitz addresses sub-problem (a) and reduces (b) to estimation of expected returns ($E[R]$) and the stock return co-variance matrix (Σ). Application of dynamic programming gives the solution (c),

$$\theta_{MV} = \arg \max_{\theta \in \Omega} E[U_q(r_p, \sigma_p)]. \tag{5.4}$$

The portfolio return and variance can be derived from R, considering all possible future scenarios. In practice the parameters are typically based on historical data and fundamental analysis that in combination forms the future scenarios ((see Sharpe 2007).

The attractiveness of MV is that it is computationally economical to find the solution vector. In addition, the trade-off between level and variance of returns is intuitive, which has made the method popular in the investment industry. The usage of return variance as measure of risk is, however, also the main drawback of the method. For the mean and variance of returns to be the only parameters of importance for an investor's risk analysis, either the PDF must be spherical symmetric[1], or the investor must be indifferent to higher moments of the distribution, such as skewness and kurtosis. It is well-known that return distributions in general do not conform to the normal PDF.[2] It is also straightforward to understand that

---

[1]This was pointed out by Chamberlain (1983) Below, I refer to such distributions as *normal*, even though e.g. the Joint Normal, the Uniform and the Binomial distributions also feature spherical symmetry.

[2]First recognised by Mandelbrot (1963), there is now overwhelming evidence available on the non-normality

investors want to protect themselves against extreme events and are more averse to downside than to upward volatility, which would imply that the second condition for MV efficiency is violated as well. These limitations of the MV approach were established at an early point, not least by Markowitz (1959). Arrow (1965) and Pratt (1964) showed that the quadratic utility function implies that risk aversion is increasing in wealth, to the extent that it at some point turns negative, which is an unrealistic property. However, it has been shown by many studies, including Levy and Markowitz (1979), that the difference between the MV approximation and the true utility maximum under exponential and logarithmic utility functions is small. Also, they argue that the risk aversion problem argued by Arrow (1965) and Pratt (1964) does not apply if the quadratic approximation is allowed to vary between portfolios.

As computational power has increased, the MV advantage of computational economies has decreased in importance, making true utility maximisation models more popular. One of those models is FSO, which is the portfolio choice model that I focus on in this thesis.

5.2   Full-Scale Optimisation

In FSO, empirical returns over a time period $t = 1, ..., T$ for N assets constitute the scenarios and can hence be expressed as a return matrix R of dimensions $(N \times T)$. The elements of R are defined

$$R_{n,t} = \frac{p_{n,t}}{p_{n,t-1}} - 1, \tag{5.5}$$

where $p_{n,t}$ is the price of asset n at time t. Once these returns are calculated, each column R_t of the matrix is treated as a future scenario with probability T^{-1}. Utility is evaluated for each possible θ vector and each scenario in the sample. The θ vector with highest average utility across scenarios will be the optimal allocation combination, which is given by

$$\theta_{FSO} = \arg\max_{\theta \in \Omega} \left(T^{-1} \sum_{t=1}^{T} U(\theta' R_t) \right). \tag{5.6}$$

To compare FSO to MV it is useful to look at how the three sub-problems of portfolio optimisation are addressed. In FSO the utility function (a) is predetermined, just as it is in MV, but no restrictions are imposed on its functional form. This allows the optimisation over more complex investor preferences than otherwise assumed. The PDF of the returns (b) is given by the empirical distribution. This means that each historical observation of returns in the cross-section of admissible stocks is treated as a possible future scenario. Each scenario is assigned the same probability (T^{-1}), which implies that the chronological order of these scenarios is ignored, but the sample co-variance pattern between assets persists. In order to find the solution vector of portfolio allocations (c), numerical/computational methods can

of returns. I discuss this further in Section 5.3.

be used — no analytical solution is pursued. As the optimisation problem is not convex in general, a search algorithm has to be used.

The term FSO was introduced by Cremers, Kritzman, and Page (2005), but the idea is not unique. Gourieroux and Monfort (1998), Brandt (1999) and Aït-Sahalia and Brandt (2001) all work with historical return distributions to approximate expected portfolio utility but do not use the term FSO. Properties of the estimated portfolios have been explored by Gourieroux and Monfort (2005). That each observation in history in general can be viewed as a future scenario relates the studies based on empirical return distributions to scenario-based approaches (using hypothesised outcomes with different probabilities attached), which have been dealt with by Grinold (1999) and Sharpe (2007). The idea of utility maximisation as a methodology for portfolio optimisation problems, based on the utility theory founded by Von Neumann and Morgenstern (1947), can be traced back at least to Tobin (1958), and also appears in several assessments of the MV approach (e.g. Levy and Markowitz 1979, Markowitz 1987).

The strength of FSO is the lack of restrictions on the utility function. It is straightforward to implement risk preferences considering higher moments of the portfolio return PDF. As empirical returns are used directly, performance of FSO is dependent on how well sample scenarios mirror actual future scenarios. I discuss this further in Section 7.6, providing suggestions on how to deal with the problem. It has been shown that FSO is particularly useful when the utility function features different types of loss aversion, or when many constraints are imposed (Cremers, Kritzman, and Page 2005, Adler and Kritzman 2007, Hagströmer, Anderson, Binner, Elger, and Nilsson 2008, Hagströmer and Binner 2009). Accordingly, the utility surface is often non-convex, which makes analytical solutions unfeasible. Instead search techniques must be used, testing different solution vectors, θ.

In the following two chapters of this thesis I solve different empirical FSO applications. Before doing that, it is useful to improve the understanding of the three sub-problems of portfolio optimisation. The remainder of this chapter is hence dedicated to understanding why return distributions are non-normal; why this matters to investors and how that can be accounted for in utility functions; and finally how a solution vector can be found once a utility function has been specified.

5.3   Return Distributions

Understanding the deviations from normality in financial return distributions is important in forming an efficient portfolio selection model. We need to understand both what is causing the deviations and what preferences investors have for different distribution characteristics. With that knowledge, appropriate utility functions can be formed and used for FSO portfolio selection.

As discussed by Mandelbrot (1963), financial asset price changes form distributions that are more peaked than samples drawn from Gaussian distributions.[3] After Mandelbrot's article, an extensive literature on the proper mathematical formulation of financial return distributions emerged, and eventually the problem was addressed with techniques such as *conditional heteroscedasticity models* (Engle 1982, Bollerslev 1986), *stochastic volatility models* (Clark 1973, Taylor 1982, Taylor 1986), and *regime switching models* using a mixture of normal distributions (Goldfeld and Quandt 1973, Hamilton 1989). As the return distribution parameterisation is super-ceded by the FSO model, I do not refer to such problems further here, but e.g. Aparicio and Estrada (2001) give a brief summary of the literature.

To illustrate the phenomenon of non-normal returns, I randomly generate a price series of 251 observations (that can be regarded as daily observations of one year), shown in Figure 5.1. Prices and returns are displayed in Panel A and B respectively. Panel C displays how the returns are distributed and how they differ from the normal distribution. As discussed in Chapter 2, prices are driven by information flows. Information reaches the market in a non-linear fashion, causing many calm days when no significant news appear and clusters of price volatility trading on days of price-driving news (as seen in Panel B). This is the reason that financial assets typically display leptokurtic return distributions (i.e. distributions with higher kurtosis than the normal distribution) (Clark 1973). Days with relatively few news and small returns form the high peak of the return probability distribution, and the large price changes due to important news cause fat tails. Non-linear behaviour of investors has also been pointed out as a reason for this pattern (Aparicio and Estrada 2001), which may be due to uncertainty of information, or insider trading (Clark 1973). A high level of kurtosis means that the asset has a high probability of extreme events (Lai, Yu, and Wang 2006). This uncertainty typically makes risk averse investors dislike kurtosis. Such kurtosis aversion is shown theoretically by Scott and Horvath (1980). Aparicio and Estrada (2001) point out that investors that mistakenly assume normality when a return distribution is leptokurtic will underestimate the risk of the asset substantially.

A second non-normality feature that is important for investors to follow is skewness (first discussed by Beedles 1979). Skewness appears when the mean and the median do not coincide, describing the degree of asymmetry in the PDF. A mean higher (lower) than the median characterise positive (negative) skewness, which implies a higher chance (risk) of extreme positive (negative) events. The example in Figure 5.1 features negative skewness (-0.5), likely due to the two occurrences of returns around -7% that are not matched by positive returns of the same magnitude. Negative skewness means that the risk for a substantial loss is bigger than the chance of a substantial gain. Typically, investors are averse to negative skewness.

Skewness is created if a company's value falls lower when negative news arrives than it rises

---

[3]Mandelbrot credits Wesley C. Mitchell as being the first one to point out this fact in 1915. See Mandelbrot (1963, footnote 3) for full reference.
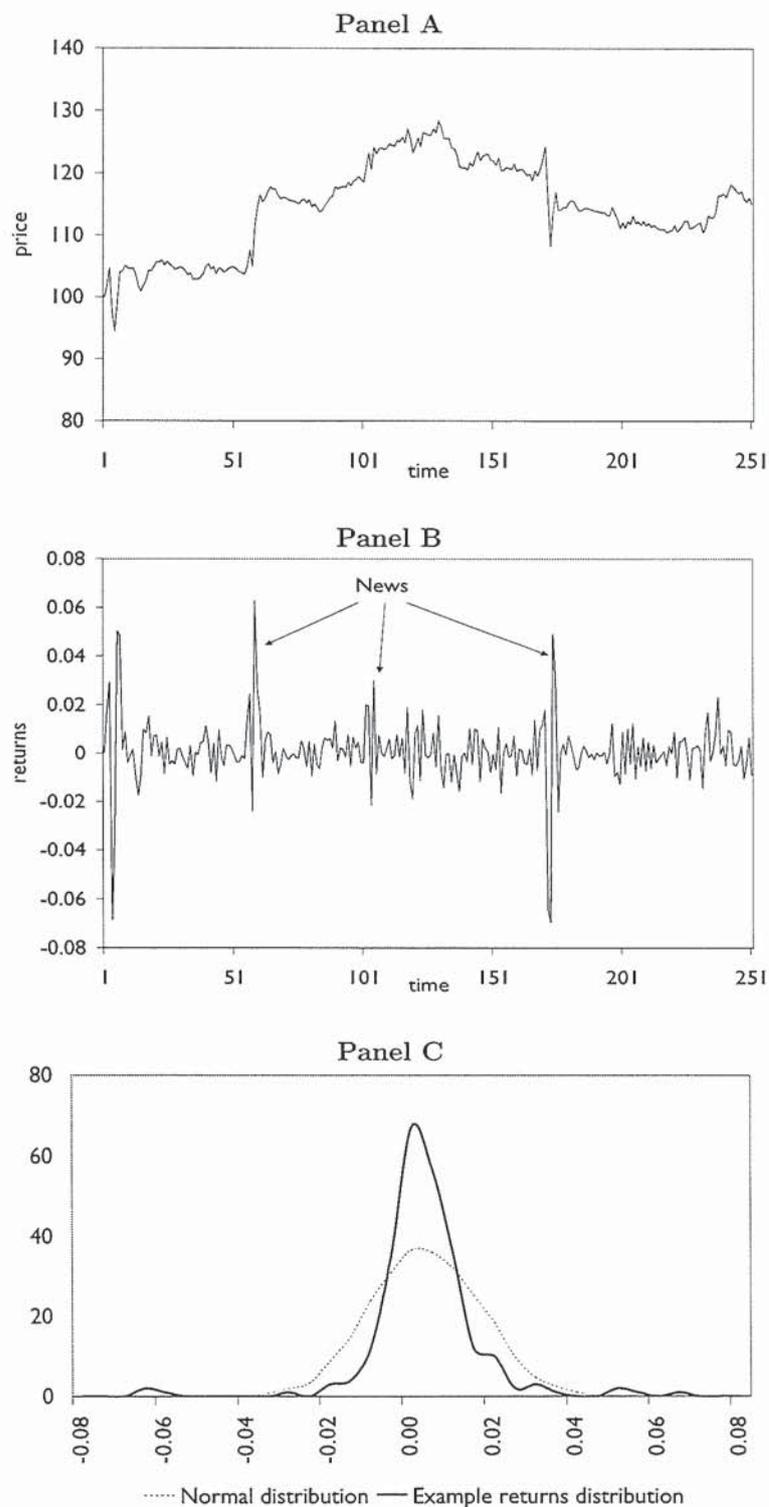
Figure 5.1: Example price and return series

Panel A shows an example price series containing 251 observations of prices. The series was generated in R 2.8 using the function `garch.sim` (TSA package). Panel B shows the same series in terms of returns, pointing out that clusters of volatility are caused by news information flows. Panel C shows the empirical distribution of the return series, showing that the returns are leptokurtic and negatively skewed in relation to a normal distribution. The series has excess kurtosis of 10.0, and skewness of -0.5.

when good news arrive, or vice versa. Damodaran (1985) argue that this can be related to a company's propensity to release positive or negative news. Guidolin and Timmermann (2005) show in a regime switching model that large negative skewness appears in the switch from a "bull" to a "bear" state of the market. Preferences for positive skewness have been shown theoretically by Arditti (1967) and Scott and Horvath (1980), and empirically by e.g. Sortino and Price (1994), Levy and Sarnat (1984), and Sortino and Forsey (1996). This implies that an investor would be willing to trade some average return for a lower risk of high negative returns (as argued by Harvey, Liechty, Liechty, and Muller 2004). This reasoning links to the literature on downside risk, including Value at Risk models. Opposite to typical investor preferences stock return distributions usually feature some degree of negative skewness.

Using variance (the 2nd moment of the return distribution) as a stand-alone measure of risk is the main drawback of using the quadratic utility function, applied in the MV approximation of optimal portfolios, as it ignores preferences for skewness and kurtosis (the 3rd and 4th moments). The focus of the next section is how other utility functions account for such preferences.

5.4   Utility Functions

Investor preferences are in economics usually described in terms of utility functions. Utility can be a function of any variable that investors value. In the stock market setting the main variable is wealth, which changes with stock returns. Investor utility can also be derived from e.g. ethical values or liquidity preferences, but these are not common in the literature and I do not consider such preferences in this part of the thesis. I define utility in terms of the portfolio return (r_p), rather than over wealth. This approach may be interpreted as a normalisation of initial wealth to one, i.e. $W_0 = 1$. A motivation for defining utility directly in terms of returns can be found in Kahnemann and Tversky (1979). They argue that investors focus more on the return on an investment than on the level of wealth.

Investor preferences implied by utility functions can be investigated by expanding the utility function in a Taylor series around the mean (μ_1, the expected return on investment) and taking expectations on both sides. This yields measures of the investors' preferences in terms of the distribution's moments. Set up in the same fashion as in Scott and Horvath (1980), the expected utility takes the following form:

$$E(U) = U(\mu_1) + \frac{U^2(\mu_1)}{2}\mu_2 + \Sigma_{i=3}^{\infty} \frac{U^i(\mu_1)}{i!}\mu_i, \tag{5.7}$$

where U^k denotes the kth derivative of the utility function and μ_j the expected j^{th} moment of the portfolio return. The expression shows that the expected utility equals the utility of the expected returns, plus the impact on utility of deviations from the expected return.

The influence of each moment on expected utility is weighted by the corresponding order derivative of the utility function. Typically, $U^2(\mu_1)$, for variance is negative; $U^3(\mu_1)$ for skewness is positive; and $U^4(\mu_1)$ for kurtosis is negative. As stated above, the MV approach is based on either assuming quadratic utility or a normal return distribution. The quadratic utility function, given in Equation 5.2, implies $U^1(\mu_1) = 1$, $U^2(\mu_1) < 0$ (usually referred to as the risk aversion parameter, η), and $U^k(\mu_1) = 0$ for all $k > 2$. If normally distributed returns are assumed, all odd moments ($k = 3, 5...$) will be zero, and all even moments will be functions of the variance (see Appendix to Chapter 1 in Cuthbertson and Nitzsche 2004).

In the MV model the *co-variance* of the assets plays an important role. Utility functions with higher moment preferences different from zero also take co-moments, such as *co-skewness* and *co-kurtosis*, into account.[4] In principle there is no limit to the number of moments to consider, but higher moments than kurtosis ($k > 4$) have not been considered in the finance literature, and will not be discussed here. However, according to Scott and Horvath (1980), most investors have utility functions where moments of odd order (i.e. $k = 1, 3, 5...$) have positive signs on its respective derivative, and moments of even order have negative derivatives.

In the subsequent analysis I consider five families of utility functions. Their general mathematical forms are presented in Table 5.1. Figure 5.2 displays how the utility varies with returns for certain specifications of the four utility function families.

The parametric, closed form utility functions that are most common in the finance literature are the families of *exponential* and *power* utility functions. The former is characterised by constant *absolute* risk aversion (CARA), and the latter by constant *relative* risk aversion (CRRA), meaning that risk aversion varies with wealth level. In Figure 5.2 examples of exponential and power utility function graphs are given in Panels A and B respectively.

For investors that put more emphasis on the higher moments skewness and kurtosis it may be appropriate to specify utility functions that feature a critical level of return, under which returns are given disproportionally bad utility. Examples of such utility functions include the *bilinear* (Panel C in Figure 5.2), the *kinked power* (Panel D), and the *S-shaped* utility functions (Panel E). The kinked functions capture a phenomenon that is central in investment management today: loss aversion. The objective of limiting losses is motivated by monetary as well as legal purposes. The issue is traditionally treated with Value-at-Risk models, and can also be incorporated in FSO theory through a constraint on the maximisation problem (as shown by Gourieroux and Monfort 2005). The bilinear utility functions have a kink at the critical return level. Above the critical point, it follows a power utility function with $\gamma = 1$, which implies a very low risk aversion for returns above that point. Below the critical point it is a straight line, implying risk neutrality. The main incentive for the investor is to achieve

---

[4]Co-skewness is a phenomenon extensively discussed by Harvey, Liechty, Liechty, and Muller (2004).

Panel A: Exponential Utility
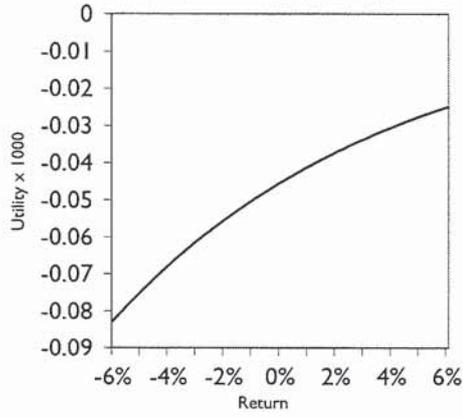
Panel B: Power Utility
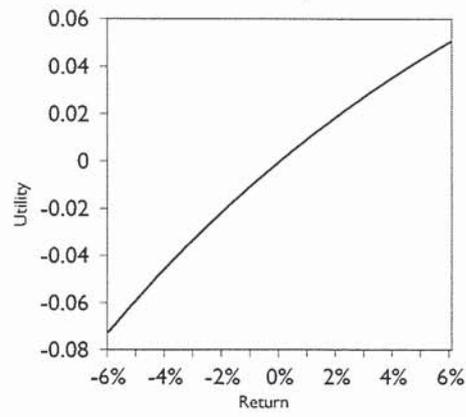
Panel C: Bilinear Utility
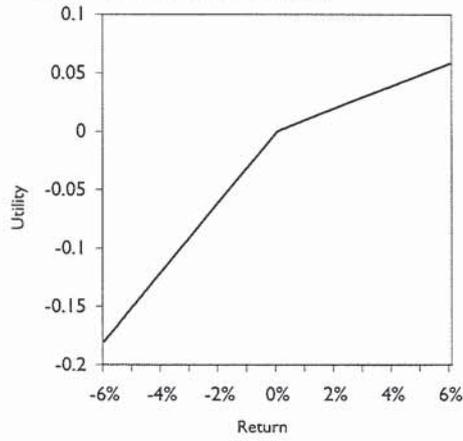
Panel D: Kinked Power Utility
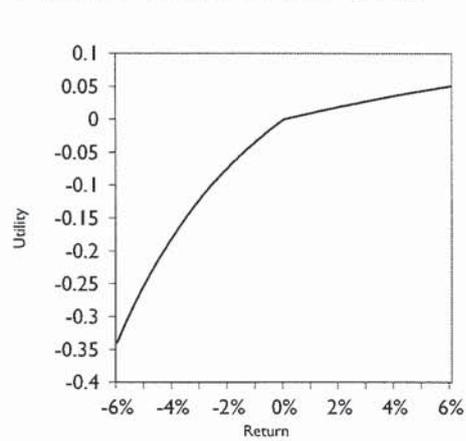
Panel E: S-Shaped Utility
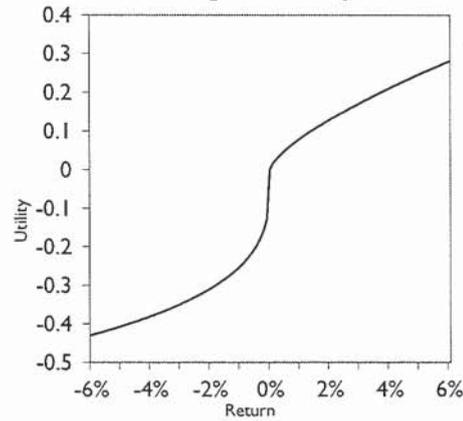
Figure 5.2: Utility function graphs

The figures show how utility varies with portfolio return under different specifications of utility functions. Panel A shows exponential utility where the risk aversion parameter is set to $A = 10$. Panel B shows power utility with risk aversion set to $\gamma = 6$. All the utility functions featuring threshold in form of a kink or an inflection point has that set to $r_p^* = 0\%$. Panel C shows the bilinear utility function with $\varphi = 5$. Panel D is the kinked power utility function, with parameters $\gamma = 6$ and $\chi = 3$. In Panel E, the S-shaped utility function depicted has parameters $A = 1$, $B = 2$, $\gamma_1 = 0.3$, and $\gamma_2 = 0.7$.

Exponential Utility:	$U_{exp} = -exp(-A(1 + r_p))$
Power Utility:	$U_{power} = \begin{cases} ((1 + r_p)^{1-\gamma} - 1)/(1 - \gamma) & \text{for} \quad \gamma > 0 \\ ln(1 + r_p) & \text{for} \quad \gamma = 1 \end{cases}$
Bilinear Utility:	$U_{bilinear} = \begin{cases} ln(1 + r_p^\star) + \varphi(r_p - r_p^\star) & \text{for} \quad r_p < r_p^\star \\ ln(1 + r_p) & \text{for} \quad r_p \geq r_p^\star \end{cases}$
Kinked power Utility:	$r_p^{adj} = \begin{cases} r_p^\star - \chi(r_p^\star - r_p) & \text{for} \quad r_p < r_p^\star \\ r_p & \text{for} \quad r_p \geq r_p^\star \end{cases}$
	$U_{KP} = U_{power}(r_p^{adj})$
S-shaped Utility:	$U_{SSHA} \begin{cases} -A(r_p^\star - r_p)^{\gamma_1} & \text{for} \quad r_p \leq r_p^\star \\ B(r_p - r_p^\star)^{\gamma_2} & \text{for} \quad r_p > r_p^\star \end{cases}$

Table 5.1: Utility function equations

In all the utility functions, r_p represents portfolio return. A represents the degree of (absolute) risk aversion in the exponential utility functions. In power utility functions, γ is the degree of (relative) risk aversion. The special case when $\gamma = 1$ is also called logarithmic utility. In the bilinear, kinked power, and S-shaped utility functions, $r_p^\star$ is the critical return level, referred to as the *kink* in the former two, and as the inflection point in the S-shaped. φ is the penalty level for returns lower than the kink in the bilinear utility function. In the kinked power utility function, γ is the degree of (relative) risk aversion, and χ the degree of loss aversion. The latter is used to adjust sub-kink portfolio returns to get r_p^{adj}, which is evaluated in the power utility function. In the S-shaped utility functions, A and B determine the disproportion between returns on different sides of the critical level, and γ_1 and γ_2 are parameters determining the curvature of the S-shape.

returns above the critical level. The function has a discontinuous first derivative. This type of functions has previously been applied by Cremers, Kritzman, and Page (2005) and Adler and Kritzman (2007). The kinked power utility function works in the same way, but utilises the relative risk aversion featured in power utility. The function has been used in portfolio optimisation by Maringer (2008b).

The S-shaped utility function is motivated by the fact that it has been shown in behaviour studies that investors prefer a certain *gains* to uncertain gains with higher expected value, but also prefer uncertain *losses* to certain losses with higher expected value (see Kahnemann and Tversky 1979). The utility function features an inflection point where these certainty preferences change. This implies high absolute values of marginal utility close to the inflection point, but low absolute marginal utility for higher absolute returns. The first derivatives are continuous, but second derivatives are not.

In the subsequent empirical chapters I apply each of these utility functions to portfolio choice problems. This creates utility maximisation problems that are non-trivial to solve. Before turning to these chapters, in the next section I discuss the last sub-problem of portfolio choice problems: how to find the optimum?

5.5 Finding the Optimal Portfolio Allocation Vector

Once a decision is made on how to characterise the expected return distribution and the investor preferences, the remaining problem is to solve for the optimal portfolio weights. The more involved the specifications of the return PDF and the utility functions are, the more expensive it is to find the optimum. When Markowitz (1952, 1959) formulated his MV theory, he was aware that both quadratic utility and normality of returns were unrealistic assumptions to make (see Markowitz 1987, Chapter 3). He emphasised that the quadratic specification was merely an approximation of the true utility. The advantages of that he pointed out were the model's economical properties — that the MV optimisation is much less expensive than true utility maximisation, and that the specification of the utility function is down to choice of the level of return variance aversion. The MV efficient portfolio can be found by application of dynamic programming, using e.g. the programme in Markowitz (1987).

The low computational cost was very important when the MV model was founded in the 1950's, and it certainly was a factor that earned it its popularity, not least in the investment industry. With the technical progress over the years since, this advantage of the MV model has decreased. As the cost of computations has fallen immensely, extensions of the MV model and alternative portfolio choice models have appeared, utilising more computational power. With the computing speed available today, true utility maximisation is becoming a viable alternative. FSO constitutes hard evidence of this, as it has its origin in the investment industry, where it is used for portfolio optimisation of e.g. hedge funds.[5]

FSO does not have an analytical solution and there is no quick way of identifying the global optimum of the solution space. The number of candidate solutions (W) grows fast with the number of assets (N) considered, as each asset added imposes another dimension on the problem. As the return PDF is not parameterised, no analytical solution can be defined, and any solution is bound to be an approximation of the true optimum. Hence, the number of candidate solutions also depends on the precision (ρ) pursued in the solution vector. Using the formula for combinations of discrete numbers with repetition, derived by Leonhard Euler (1707–1783) [see e.g. Epp (2003)], the number of candidate solutions will be approximately

$$ W = \frac{(N + (1/\rho) - 1)!}{(1/\rho)!(N-1)!} = \frac{((1/\rho)+1) * ((1/\rho)+2) * \ldots * ((1/\rho)+N-1)}{1 * 2 * \ldots * (N-1)}, \qquad (5.8) $$

provided that no short-selling is allowed and that portfolio weights add up to one. To give an understanding of the scope of this problem, I demonstrate in Table 5.2 how the number of solutions is growing with the number of assets and the precision of the problem. In spite of this dimensionality problem, Gourieroux and Monfort (2005), Cremers, Kritzman, and Page (2003, 2005) and Adler and Kritzman (2007) all argue that the computational burden of the

---

[5]The authors of the original articles on FSO work at *Windham Capital Management, LLC*, and *State Street Associates, LLC*, both in Cambridge, MA, USA.

N	$\rho = 10\%$ W	$\rho = 5\%$ W	$\rho = 1\%$ W	$\rho = 0.1\%$ W
1	1	1	1	1
2	11	21	101	1001
3	66	231	5151	501501
4	286	1771	176851	1.7E+8
5	1001	10626	4.6E+6	4.2E+10
10	92378	1.0E+7	4.3E+12	2.9E+21
20	2.0E+7	6.9E+10	4.9E+21	9.9E+39
50	6.3E+10	1.16E+17	6.7E+39	5.5E+84

Table 5.2: Number of possible FSO solutions

This table shows the number of possible discrete solutions (W) of a portfolio selection problem with N assets, where the smallest change of the discrete numbers are given by the precision parameter ρ, and the portfolio weights are constrained to be larger than or equal to zero and to sum to one. The numbers are calculated using the formula given in Equation 5.8.

FSO technique has become obsolete with the ample computational power on hand. There is however, to my knowledge, no study verifying this. In Cremers, Kritzman, and Page (2005) and Adler and Kritzman (2007), a search algorithm is applied to find the FSO optimum. Such algorithms are necessary when using larger number of assets. Cremers, Kritzman, and Page (2005) consider 61 assets in their application, and use a precision of 0.1% and do not allow for short selling. This implies $W = 7.23 \times 10^{98}$, which is the number of vectors to be evaluated over their 10 annual observations. They do not disclose their search algorithm.

5.5.1 Finding the Full-Scale Optimum: Two Empirical Studies

There are many alternatives on how a search algorithm can be applied to find the solution to this problem. The purpose of the next two chapters of this thesis is to explore different ways of doing this. Perhaps the most intuitive approach is the grid search, i.e. to evaluate the utility function for each possible solution vector. In Chapter 6, I apply a grid search to a portfolio optimisation problem featuring three different equity indexes. I evaluate the FSO solutions identified under different utility function specifications in relation to the MV solution, comparing utility achieved both in-sample and out-of-sample.

The grid search quickly becomes computationally unfeasible when more assets are added to the problem. In Chapter 7 I address a stock portfolio optimisation problem of 97 stocks. This is a largest data set in an FSO application to date and demands a more sophisticated search algorithm than that used in Chapter 6. To avoid searching the whole set of possible solutions, different heuristic search algorithms exist. Heuristic methods map the solution set and are self-learning — the search is based on previous iterations of the algorithm. I apply a heuristic method called differential evolution to solve my stock portfolio optimisation problem. My results show that the problem at hand can be treated in a FSO framework. I provide several performance metrics of the algorithm I use, as well as a comparison to MV in terms of utility. I conclude with a word on precautions to take when using FSO in practice.

Chapter 6

Full-Scale Optimisation using Grid Search: Application to UK Equity Indexes

In this chapter I apply FSO to a portfolio optimisation problem featuring three equity indexes.[1] In this setting I identify several utility functions featuring loss aversion and prospect theory under which FSO is a substantially better approach than MV. As the equity indexes have return distributions with relatively small deviations from normality, the findings indicate much broader usefulness of FSO than has earlier been shown. The results hold in-sample and out-of-sample, and the performance improvements are given in terms of utility as well as certainty equivalents.

In the first two sections of this chapter I present the problem setting for this empirical application of FSO and the four different types of utility functions that I consider. Next, in Section 6.3, I discuss the optimisation method used to solve the problem. In Section 6.4, I present the methodology for evaluating the results, and in Section 6.5 I present the outcome of that evaluation. In Section 6.6, I summarise the main findings of the chapter.

6.1 Setting: Equity Index Portfolios

There exist two earlier studies that evaluate the FSO methodology. Cremers, Kritzman, and Page (2005) show in a hedge fund selection problem that the performance of FSO (in terms

---

of utility) is substantially better than Markowitz's (1952, 1959) MV approach when investor preferences are modelled to include loss aversion or prospect theory (based on Kahnemann and Tversky 1979). The results are confirmed in an out-of-sample application by Adler and Kritzman (2007). Both studies deal with hedge funds, an asset class that is well-known to have return distributions that deviate much more from the normal distribution than e.g. equity index returns or stock returns. When strong non-normalities exist, the potential for FSO performance being superior to MV is higher, as investor preferences for such will matter more.

In the empirical application in this chapter, I use an approach that strongly resembles the two preceding evaluations of FSO, with the important difference that I use equity indexes instead of hedge funds as admissible assets. This is a more challenging setting for the FSO approach, as the returns from equity indexes typically deviate much less from the normal distribution than hedge fund returns do. Specifically, I use three indexes that are published by the Financial Times: FTSE 100, FTSE 250, and FTSE All-World Emerging Market Index (EMI). FTSE 100 includes the 100 largest firms on the London Stock Exchange (LSE) and FTSE 250 include mid-sized firms, i.e. the 250 firms following the hundred largest. FTSE EMI reflects the performance of mid- and large-sized stocks in emerging markets.[2] The data are downloaded from Datastream.[3] I calculate return series for eight years of monthly observations (Jan 1999 - Dec 2006), yielding 96 observations. As shown in Figure 6.1, the data feature two expansionary periods and one downward trend. The return distribution properties given in Table 6.1 show that the three indexes display positive means over the sample period. FTSE100 is the least volatile of the three, followed by FTSE250 and FTSE EMI. All indexes feature negative skewness. Excess kurtosis is observed for FTSE 100 and FTSE 250. It is shown with a Jarque-Bera test of normality (Jarque and Bera 1980) that normality can be rejected for the UK indexes, but not for the EMI. For comparison, the hedge funds considered by Cremers, Kritzman, and Page (2005) had skewness and kurtosis averaging -0.12 and 6.44 respectively, and they rejected normality in 85% of the return distributions considered.

6.2 Utility Functions

I apply my portfolio selection problem to exponential, power, bilinear, and S-shaped utility functions. The same utility function types were investigated by Cremers, Kritzman, and Page (2005) and Adler and Kritzman (2007), but they considered only a few cases of each utility function type. I perform my investigation under several different utility function parameter values, chosen with the intention to cover all reasonable levels.

The bilinear and S-shaped utility functions are my main focus, as it has been shown

---

[2] For exact definition, see http://www.ftse.com/Indices/FTSE_Emerging_Markets/Downloads /FTSE_Emerging_Market_Indices.pdf

[3] The Datastream codes for the indexes are FT100GR(PI), FT250GR(PI), and AWALEG£(PI).

	Mean	Variance	Skewness	Kurtosis	J–B stat.	p(J–B)
FTSE 100	0.0010	0.0015	-0.93	3.91	17.07	0.00
FTSE 250	0.0096	0.0025	-0.83	4.36	18.32	0.00
FTSE EMI	0.0122	0.0043	-0.24	2.91	0.93	0.63

Table 6.1: Summary statistics

The table shows the first four moments of the monthly data for the three equity indexes FTSE100, FTSE250 and FTSE EMI. Kurtosis is the estimator of Pearson's kurtosis using the function kurtosis in R (this is not excess kurtosis). J-B stat. is the Jarque-Bera test statistic, and p(J–B) is the probability that the series is following a normal distribution according to this test.

in a hedge fund setting that FSO using these functions yields portfolio weights that differ substantially from those of MV optimisation (Cremers, Kritzman, and Page 2005, Adler and Kritzman 2007). I seek to test whether this difference holds in a portfolio selection problem of equity indexes. I also include the traditional investor preferences of exponential and power utility. Utility maximisation using these functions have repeatedly been shown to differ only marginally to quadratic utility (Levy and Markowitz 1979, Markowitz 1987, Cremers, Kritzman, and Page 2005), and I include them for the purpose of illustration. Gourieroux and Monfort (2005) apply the FSO model to the exponential and power utility functions, which allows them to derive the asymptotic properties of the FSO estimator. They establish in this context that the utility maximising estimator yields greater robustness than the MV counterpart, as the full empirical return distribution is considered.

The range of utility parameters tested is given in Table 6.2. For the exponential utility function, the only parameter to vary is the level of risk aversion (A), which I vary between 1 and 10. The γ parameter in the power utility function determines level of risk aversion and how risk aversion decreases with wealth. As I let it vary between 1 and 5, I include the special case when the power utility function is logarithmic ($\gamma = 1$). The higher γ and A are, the higher is the risk aversion. For the bilinear utility function, I vary the critical point (the kink, $r_p^\star$, varied from -4% to +0.5%) under which returns are given a disproportionate bad utility. I also vary the magnitude, φ, of this disproportion from 1 to 10. In the S-shaped utility function there are five parameters to vary. I test three levels for the inflection point, $r_p^\star$: 0%, −2.5% and −5%. The parameters γ_1 and A respectively determine the shape and magnitude of the downside of the function, whereas γ_2 and B determine the upside characteristics in the same way. The disproportion between gains and losses can be determined either by the γ parameters or the A and B parameters, or both. I perform one set of tests where the γ's vary ($\gamma_2 \geq \gamma_1$) and the magnitude parameters are held constant and equal, and one set of tests where the γ_1 and γ_2 are constant and equal, but where A and B vary ($A \geq B$).

Figure 6.1: Development of the three indexes over the time period considered

Monthly prices for FTSE100, FTSE250, and FTSE EMI, indexed for January 1999, are plotted for 1999-2006.

6.3 Optimisation Method: Grid Search

To identify the FSO optimum in this three-asset case I use grid search. In the grid search, all possible solution vectors are evaluated, and the one that maximises utility is taken as the FSO optimum. If the precision of the grid is high enough, the identified solution should be near the global optimum of the utility surface. I let the precision parameter of the grid be $\rho = 0.5\%$. As I do not allow for short-selling, this three-asset problem has 20301 potential solutions that I evaluate over 96 monthly observations. The choice of grid precision can be made on the basis of the trade-off between the marginal utility of increasing ρ and the additional computational cost of doing so. Setting the grid precision to 0.1% instead of 0.5% increases the grid size by a factor of almost 25 (from 20301 to 501501). For most portfolio choice problems such increased computational cost can not be motivated by the increased utility achieved.

An alternative to increasing the precision of the complete grid is to perform a second grid search around the optimum found in the first search. I perform a second step grid search using $\rho = 0.1\%$ around the $\rho = 0.5\%$ optima, covering all possible allocation within the range of $\theta_i \pm 1\%$. This is the range needed to cover all solutions not covered in the previous grid search, as if two assets change by 0.5% in the same direction, the third has to change by 1%. In general the range required for the second grid search is $\rho(N-1)$, where N is the number of assets in the problem and ρ is the precision of the first grid search. The second grid dimension is in

101

Utility function	Parameter values
Exponential:	$0.5 \leq A \leq 6$
Power:	$1 \leq \gamma \leq 5$
Bilinear:	$-4\% \leq r_p^* \leq 0.5\%$; $1 \leq \varphi \leq 10$
S-shaped:	$-5\% \leq r_p^* \leq 0\%$; $0.05 \leq \gamma_1 \leq 0.5$; $0.95 \leq \gamma_2 \leq 0.5$; $A = 1.5$; $B = 1.5$
	$-5\% \leq r_p^* \leq 0\%$; $\gamma_1 = 0.5$; $\gamma_2 = 0.5$; $1.5 \leq A \leq 2.9$ $1.5 \leq B \leq 0.1$;

Table 6.2: Utility function parameters

The table describes the intervals for parameter values applied to the utility functions considered in this chapter. The utility function definitions are given in Table 5.1. Exact parameter combinations applied are given in the results tables for each utility function (see Tables 6.5–6.9).

my case $67 \leq W \leq 331$, depending on whether the first stage optimum contains allocations close to the allowed limits 0 and 1 or not. I found that both the computational cost and the utility improvement of the second grid search are minute. None of the utility improvements exceeded 1%, and only 8 out of 132 utility functions had improvements exceeding 0.1%. This utility improvement is what I sacrifice when choosing $\rho = 0.5\%$ rather than $\rho = 0.1\%$. As a grid search never can yield an exact solution, I also applied Simulated Annealing on the area around the optimum (for a description of this technique, see e.g. Goffe, Ferrier, and Rogers 1994). Again, the utility improvement is very small. I hence concluded that $\rho = 0.5\%$ is enough precision for this application. In portfolio choice problems with large grids resulting from a larger number of assets, the two-stage technique presented here may be a viable option to decrease computational cost. Caution is needed however, as non-convexity may lead an imprecise grid to an area not including the global optimum. The computational cost of the second step grid search, measured in computation time, is also relatively small. Times needed for performing FSO on each utility function type used in this article, for first and second step grid searches respectively, are given in Table 6.3.

The study is performed in a one-period setting – no rebalancing of the portfolio is considered. In order to mitigate the probability of corner solutions, I scale all returns to conform to implied returns of an equal-weighted portfolio ($\theta_n = N^{-1}$).[4] This does not change the shape of the probability distribution, which is what matters for the optimisation and hence the comparison between FSO and MV. Corner solutions need to be avoided, as solutions that are bounded by the allocation constraints can give a false impression of equality of the two methods.

---

[4]The difference between the average return of all three indexes and the return corresponding to the variance of the equally weighted portfolio is added to each observation. In this case, the difference added amounts to 0.086%.

Utility function	Time 1st step	Time 2nd step
Exponential:	1.25	0.00
Power:	1.47	0.01
Bilinear:	35.78	0.13
S-shaped:	44.75	0.14

Table 6.3: Computational cost of FSO

Time is specified in seconds. Optimisation was run over 96 time periods. The platform used was an Intel Core 2, 2.16 GHz processor with 3GB RAM. Software used was R v2.6.1, applying the function `system.time`.

6.4 Model Evaluation

The methodology for comparing FSO to the MV approach is to a large extent inspired by that applied by Cremers, Kritzman, and Page (2005) and Adler and Kritzman (2007), where the performance of the different approaches is measured in utility.

In order to compare the FSO and MV optima, the expected FSO portfolio return is calculated $(E[r_{p,FSO}] = E[\theta'_{FSO}R])$. The corresponding MV optimal portfolio, θ_{MV}, is taken to be the variance-minimising allocation that yields the same expected portfolio return, a point on the MV efficient frontier.[5]

6.4.1 Utility Differences and Certainty Equivalents

The performance of the two solutions (θ_{FSO} and θ_{MV}) is calculated in terms of the utility function applied for the FSO method. The relative expected utility difference between the two portfolios can be calculated as

$$\epsilon_{MV} = \frac{E[U(\theta'_{FSO}R)] - E[U(\theta'_{MV}R)]}{|E[U(\theta'_{MV}R)]|}, \tag{6.1}$$

where ϵ_{MV} is a measure of the MV approximation error.[6]

Measuring the difference in terms of improved utility has the drawback that utility is hard to interpret in economic terms. Also, different utility functions yield different magnitudes of utility variation for a set of returns. An alternative measure, which is more straightforward to interpret, is the certainty equivalent.[7]

The certainty equivalent of a risky investment is the certain return yielding the same

---

[5]I use the original MV model as benchmark in this study. In this way, it can be established whether the FSO model is superior to that model. The rich supply of MV extensions, however, is yet to be compared to the FSO model.

[6]The MV solution can be calculated with much higher detail than the FSO portfolio, which is limited to 0.5% precision. Accordingly, a minor source of utility difference will be due to this limitation, which in the comparison is to the MV method's advantage.

[7]Certainty equivalents have earlier been used for FSO-MV comparisons in a working paper by Cremers, Kritzman, and Page (2003).

utility as the expected return of the risky investment. Denoting the certainty equivalent return r_{CE}, this relationship can be described as

$$U(1 + r_{CE}) = E[U(1 + \theta' R)]. \tag{6.2}$$

The difference between the expected return on the risky investment and the certainty equivalent is usually defined as the risk premium, i.e. the additional expected return the investor requires to take on the risky investment instead of the risk-free investment (Meucci 2005). This number has an obvious economic interpretation and can easily be restated in monetary terms. For example, if the risk-premium is 0.5%, the investor requires an extra $500 of expected return per $100,000 invested.

Using the return series calculated for each optimum above, I can derive measures of expected average utility for FSO and MV respectively over time, $\overline{U}_{FSO}$ and $\overline{U}_{MV}$. This simplifies Equation 6.2 to

$$U(1 + r_{CE}) = \overline{U}, \tag{6.3}$$

which can be solved for each of the utility functions presented above, see Table 6.4. Having calculated the certainty equivalent for the FSO and MV optima, the two can straightforwardly be compared by letting

$$\Delta r_{CE} = r_{CE}^{FSO} - r_{CE}^{MV}, \tag{6.4}$$

where Δr_{CE} is the improvement in certain return on investment corresponding to the utility improvement of using FSO rather than MV. This difference can be interpreted in the same way as the risk premium. Note that Δr_{CE} by definition always (in-sample) is positive or equal to zero. As the corresponding MV solution is evaluated in the same utility function, it can never (in-sample) achieve higher utility than the utility maximum identified by FSO. Under quadratic utility however, it will be equivalent.

6.4.2 Success Rates

For the bilinear and S-shaped utility functions, I also calculate success rates of the same type as in Cremers, Kritzman, and Page (2005). The success rate describes the relative frequency of portfolio returns superior to the investor's specified critical level (kink and inflection point respectively). Denoted SR, the success rate is

$$SR = \frac{\sum_{t=1}^{T} I(r_p > r_p^\star)}{T}, \tag{6.5}$$

where $I(\cdot)$ is an indicator variable taking value 1 when the condition holds and zero otherwise.

Utility function	CE derivation			
Exponential:	$-exp(-A(1+r_p))$		$\Leftrightarrow$	$r_{CE} = -\frac{1}{A}ln(-\overline{U}) - 1$
Power:	$\frac{(1+r_p)^{1-\gamma}-1}{1-\gamma}$	for $\gamma > 0$	$\Leftrightarrow$	$r_{CE} = [1 + (1-\gamma)\overline{U}]^{\frac{1}{1-\gamma}} - 1$
	$ln(1+r_p)$	for $\gamma = 1$	$\Leftrightarrow$	$r_{CE} = exp(\overline{U}) - 1$
Bilinear:	$ln(1+r_p)$	for $r_p \geq r_p^\star$	$\Leftrightarrow$	$r_{CE} = exp(\overline{U}) - 1$
	$\varphi(r_p - r_p^\star) + ln(1+r_p^\star)$	for $r_p < r_p^\star$	$\Leftrightarrow$	$\star$
S-shaped:	$-A(r_p^\star - r_p)^{\gamma_1}$	for $r_p \leq r_p^\star$	$\Leftrightarrow$	$r_{CE} = r_p^\star - \left(\frac{\overline{U}}{-A}\right)^{\frac{1}{\gamma_1}}$
	$+B(r_p - r_p^\star)^{\gamma_2}$	for $r_p > r_p^\star$	$\Leftrightarrow$	$r_{CE} = r_p^\star + \left(\frac{\overline{U}}{B}\right)^{\frac{1}{\gamma_2}}$

Table 6.4: Certainty equivalent equations

The utility function definitions are given in Table 5.1. $\star$Under Bilinear utility when $\overline{U} < U(r_p^\star)$, no explicit solution exists. This case is handled by a standard search algorithm.

6.4.3 Out-of-Sample Testing

In order to further examine the robustness of the FSO methodology, I repeat the procedure described above in an out-of-sample setting. This is done using essentially the same methodology as in Adler and Kritzman (2007). For this purpose, the sample is split in two halves. The first half is used for estimating optimal portfolio allocations, and the performance of the optimal portfolio retrieved is measured on the second half.

I generate 10,000 samples of cross-sectional monthly returns by drawing from the second half of my sample (with replacement), which was not employed for the portfolio optimisation. This bootstrapping procedure allows me to study performance of the optimised portfolio out-of-sample. The exercise is repeated using the second half of the sample for estimation and the first half for the diagnostics. Finally, the average utility difference between FSO and MV portfolios over the 20,000 samples is calculated.

6.5 Results

I run the portfolio selection problem described under 103 different utility function specifications: 12 with exponential utility; 9 with power utility; 31 with bilinear utility; and 51 with S-shaped utility. Below the results are presented and interpreted for each utility function type. The section is concluded with some general observations.

6.5.1 Exponential and Power Utility Portfolios

Running the exponential utility maximisation yields, as expected, portfolios with high weights to risky assets when A is low and less so when the risk aversion parameter A is set higher

	FSO approach			MV approach			ϵ_{MV}		Δr_{CE}	
A	θ_1	θ_2	θ_3	θ_1	θ_2	θ_3	IS	OOS	IS	OOS
0.5	0.000	0.000	1.000	0.000	0.001	0.999	0.0%	0.0%	0.1%	0.0%
1	0.000	0.000	1.000	0.000	0.001	0.999	0.0%	0.0%	0.1%	0.0%
1.5	0.000	0.000	1.000	0.000	0.001	0.999	0.0%	0.0%	0.1%	0.0%
2	0.000	0.230	0.770	0.000	0.230	0.770	0.0%	0.0%	0.0%	0.0%
2.5	0.000	0.400	0.600	0.000	0.400	0.600	0.0%	0.0%	0.0%	0.0%
3	0.000	0.515	0.485	0.000	0.515	0.485	0.0%	0.0%	0.0%	0.0%
3.5	0.000	0.595	0.405	0.000	0.595	0.405	0.0%	0.0%	0.0%	0.0%
4	0.000	0.655	0.345	0.000	0.655	0.345	0.0%	0.0%	0.0%	0.0%
4.5	0.000	0.705	0.295	0.000	0.705	0.295	0.0%	0.0%	0.0%	0.0%
5	0.000	0.740	0.260	0.000	0.740	0.260	0.0%	0.0%	0.0%	0.0%
5.5	0.000	0.770	0.230	0.000	0.770	0.230	0.0%	0.0%	0.0%	0.0%
6	0.000	0.795	0.205	0.000	0.795	0.205	0.0%	0.0%	0.0%	0.0%

Table 6.5: Exponential utility results

This table shows results for FSO and MV portfolios using exponential utility. Each row represents a different setting of the absolute risk aversion A, given in the first column. The following six columns present portfolio allocations for FSO and MV portfolios. Portfolio allocations notation are as follows: θ_1 is allocation to FTSE100, θ_2 is to FTSE250, and θ_3 is to FTSE EMI. The given allocations are those estimated on the full data set. In the four rightmost columns the utility improvement of using FSO rather than MV is given in terms of utility and in terms of certainty equivalents. For each of these, both in-sample results (IS) and out-of-sample results (OOS) are given. OOS results given are the averages of the two OOS applications. Certainty equivalent improvements are given in annual terms.

(see Table 6.5). The portfolios identified are all very close to the MV efficiency frontier, resulting in extremely small utility improvements, if any, when using FSO instead of MV. As the FSO and MV solutions are close to identical, the differences in performance out-of-sample are close to zero. Portfolio optimisation based on the power utility function is run with 9 different levels of the relative risk parameter γ. As shown in Table 6.6, the power utility function yields portfolios allocating the whole investment to the most risky asset (in terms of variance), FTSE EMI, for $\gamma \leq 1.5$, and then gradually moves towards the medium-risky asset (FTSE 250) as γ grows. The deviations from the MV frontier are slightly larger than in the exponential utility cases, but differences in utility outcome both in-sample and out-of-sample are still minute. In-sample utility improvement never exceed 0.02%, and out-of-sample differences display neither substantial magnitude, nor consistent directions on deviations from zero.

The improvement in terms of certainty equivalents is zero or close to zero under both exponential and power utility functions. Hence, an investor following these utility functions is not prepared to pay anything to move from MV to FSO. These results conform well to those of earlier assessments (Levy and Markowitz 1979, Markowitz 1987), showing that portfolio allocations chosen by the utility maximising approaches yield very small improvements relative to the MV approach, when the investor's utility is well described by the exponential or the power utility functions.

	FSO approach			MV approach			ϵ_{MV}		Δr_{CE}	
γ	θ_1	θ_2	θ_3	θ_1	θ_2	θ_3	IS	OOS	IS	OOS
1	0.000	0.000	1.000	0.000	0.001	0.999	0.0%	0.1%	0.1%	0.0%
1.5	0.000	0.000	1.000	0.000	0.001	0.999	0.0%	-0.1%	0.1%	0.0%
2	0.000	0.220	0.780	0.000	0.220	0.780	0.0%	-0.2%	0.0%	0.0%
2.5	0.000	0.395	0.605	0.000	0.395	0.605	0.0%	0.1%	0.0%	0.0%
3	0.000	0.510	0.490	0.000	0.510	0.490	0.0%	0.1%	0.0%	0.0%
3.5	0.000	0.590	0.410	0.000	0.590	0.410	0.0%	0.0%	0.0%	0.0%
4	0.000	0.650	0.350	0.000	0.650	0.350	0.0%	0.2%	0.0%	0.0%
4.5	0.000	0.700	0.300	0.000	0.700	0.300	0.0%	0.0%	0.0%	0.0%
5	0.000	0.735	0.265	0.000	0.735	0.265	0.0%	-0.1%	0.0%	0.0%

Table 6.6: Power utility results

This table shows results for FSO and MV portfolios using power utility. Each row represents a different setting of the relative risk aversion γ, given in the first column. The following six columns present portfolio allocations for FSO and MV portfolios. Portfolio allocations notation are as follows: θ_1 is allocation to *FTSE100*, θ_2 is to *FTSE250*, and θ_3 is to *FTSE EMI*. The given allocations are those estimated on the full data set. In the four rightmost columns the utility improvement of using FSO rather than MV is given in terms of utility and in terms of certainty equivalents. For each of these, both in-sample results (IS) and out-of-sample results (OOS) are given. OOS results given are the averages of the two OOS applications. Certainty equivalent improvements are given in annual terms.

6.5.2 Bilinear Utility Portfolios

The tests on bilinear utility functions are performed with the kink $(r_p^{\star})$ at different levels and with various penalties (φ) on sub-kink returns. Each kink value is tested for 3 different penalty levels. As is shown in Table 5.1 above, the disutility of sub-kink returns is amplified by the factor φ. When $\varphi = 1$, there is no kink, and the utility function features no loss aversion.

The results retrieved for bilinear utility functions are shown in Table 6.7. As expected, the risk level of the optimal portfolios retrieved with FSO decreases as the kink and penalty parameters increase. When the penalty parameter is set to 1 all wealth is allocated to the most risky asset (FTSE EMI), which is due to the lack of risk aversion in this case. For higher penalty levels, the portfolios are increasingly weighted towards the less risky assets (FTSE 100 and FTSE 250). At $\varphi = 10$, no allocations are made to the most risky asset. This occurs when the incentive to avoid returns less than those associated with the kink dominates other investor incentives, such as maximising returns or minimising risk by diversification. Portfolio diversification is highest in portfolios optimised with $\varphi = 5$.

Whereas many of the portfolios optimised under bilinear utility are close to the MV frontier, there are cases where the utility improvement is substantial (up to 4% in-sample and 13% out-of-sample). Looking at certainty equivalents, there is a positive, fairly consistent, but small difference in favour of the FSO portfolios. On average, r_{CE} improvement amounted to 0.02% in-sample and 0.04% out-of-sample (see bottom of Table 6.7). This means that usage of FSO rather than MV when investor preferences are correctly described with a bilinear utility function, yields a utility improvement equivalent to that of a certain 0.02% annual return. Of the 20,000 draws made for the out-of-sample test, 14% yield r_{CE} improvements higher than

1%, and 8% yield less than -1% (in annual terms).

Looking at the distribution of returns out-of-sample, there are on average (of the 20,000 bootstrapped samples) no differences between FSO and MV in mean and variance. The solutions under bilinear utility, however, yield higher skewness and higher kurtosis than what MV solutions yield on average (as showed in Table 6.8). Bilinear utility does not penalise kurtosis, as there is no decrease in marginal utility with returns (i.e. no risk aversion) except for the jump at the kink. Accordingly, it can be seen that the corresponding MV portfolios are more diversified than the FSO portfolios.

The success rates (ratio of portfolio returns exceeding the kink) are slightly higher on average using FSO to implement the bilinear preferences than when using MV portfolios, which can be related to the higher skewness resulting under bilinear utility (see bottom of Table 6.7). Differences appeared only in 10 cases, out of which two were in favour of MV. Success rates implemented by Cremers, Kritzman, and Page (2005) yielded similarly small differences under the bilinear utility specification.

6.5.3 S-Shaped Utility Portfolios

As discussed in Section 5.4, the S-shaped utility function features risk loving behaviour when returns are below a critical value $r_p^\star$, and risk aversion when returns are above that value. I perform my tests with the inflection point set to 0%, -2.5% and -5%, with varying settings of either the gamma values (γ_1 and γ_2) or the magnitude parameters (A and B). These parameters regulate the curvature of the S-shape.

As seen in Table 6.9, the utility from the FSO approach under S-shaped utility is considerably better than the utility obtained with S-shaped preferences when evaluating allocations chosen via the MV approach. The average utility difference is 10% in-sample and 15% out-of-sample. Certainty equivalent improvements are on average 0.2%, both in-sample and out-of-sample (10 times higher than under bilinear utility). That is, using FSO causes an increase in utility corresponding to a certain annual return of 0.2% on average – a substantial gain. This Δr_{CE} is consistent throughout the 20,000 out-of-sample draws. It exceeds 1% in 35% of the cases, whereas it is lower than -1% in only 15% of the tests.

Parameters		FSO approach			MV approach			ϵ_{MV}		Δr_{CE}		Success rates	
φ	$r_p^\star$	θ_1	θ_2	θ_3	θ_1	θ_2	θ_3	IS	OOS	IS	OOS	SR_{FSO}	SR_{MV}
1	-	0.000	0.000	1.000	0.000	0.001	0.999	0.0%	-0.1%	0.1%	0.0%	-	-
2.5	-0.04	0.000	0.745	0.255	0.000	0.745	0.255	0.0%	0.0%	0.0%	0.0%	87.4%	87.4%
2.5	-0.035	0.000	0.760	0.240	0.000	0.760	0.240	0.0%	0.0%	0.0%	0.0%	84.2%	84.2%
2.5	-0.03	0.000	0.860	0.140	0.000	0.860	0.140	0.0%	-0.4%	0.0%	0.0%	80.0%	80.0%
2.5	-0.025	0.000	0.955	0.045	0.010	0.912	0.078	2.2%	-2.5%	0.0%	0.0%	76.8%	75.8%
2.5	-0.02	0.000	1.000	0.000	0.023	0.902	0.075	4.1%	-0.2%	0.1%	0.0%	75.8%	74.7%
2.5	-0.015	0.000	1.000	0.000	0.023	0.902	0.075	3.1%	-0.3%	0.1%	0.0%	73.7%	73.7%
2.5	-0.01	0.000	1.000	0.000	0.023	0.902	0.075	2.5%	10.2%	0.1%	0.4%	69.5%	69.5%
2.5	-0.005	0.000	1.000	0.000	0.023	0.902	0.075	1.7%	5.7%	0.1%	0.2%	67.4%	67.4%
2.5	0	0.000	1.000	0.000	0.023	0.902	0.075	1.3%	2.6%	0.1%	0.1%	63.2%	64.2%
2.5	0.005	0.000	1.000	0.000	0.023	0.902	0.075	0.7%	12.9%	0.1%	0.2%	60.0%	59.0%
5	-0.04	0.045	0.805	0.150	0.023	0.902	0.075	1.5%	0.1%	0.1%	-0.2%	88.4%	86.3%
5	-0.035	0.030	0.805	0.165	0.005	0.916	0.080	0.6%	1.0%	0.1%	-0.1%	84.2%	83.2%
5	-0.03	0.160	0.755	0.085	0.148	0.807	0.045	0.4%	-1.5%	0.0%	-0.1%	84.2%	83.2%
5	-0.025	0.225	0.775	0.000	0.233	0.742	0.025	0.4%	2.3%	0.0%	0.2%	81.1%	79.0%
5	-0.02	0.320	0.680	0.000	0.321	0.675	0.004	0.1%	7.2%	0.0%	0.2%	76.8%	76.8%
5	-0.015	0.400	0.600	0.000	0.400	0.600	0.000	0.0%	4.4%	0.0%	0.1%	73.7%	73.7%
5	-0.01	0.335	0.665	0.000	0.335	0.664	0.001	0.0%	6.1%	0.0%	0.1%	68.4%	68.4%
5	-0.005	0.305	0.695	0.000	0.307	0.686	0.007	0.1%	8.6%	0.0%	0.0%	67.4%	67.4%
5	0	0.210	0.790	0.000	0.219	0.753	0.028	0.3%	8.5%	0.0%	0.1%	65.3%	66.3%
5	0.005	0.190	0.810	0.000	0.200	0.767	0.033	0.6%	12.9%	0.0%	0.1%	59.0%	57.9%
10	-0.04	0.595	0.405	0.000	0.595	0.405	0.000	0.0%	0.0%	0.0%	0.0%	88.4%	88.4%
10	-0.035	0.450	0.550	0.000	0.450	0.550	0.000	0.0%	-0.1%	0.0%	0.0%	85.3%	85.3%
10	-0.03	0.520	0.480	0.000	0.520	0.480	0.000	0.0%	-0.1%	0.0%	0.0%	83.2%	83.2%
10	-0.025	0.500	0.500	0.000	0.500	0.500	0.000	0.0%	-0.1%	0.0%	0.0%	80.0%	80.0%
10	-0.02	0.565	0.435	0.000	0.565	0.435	0.000	0.0%	5.1%	0.0%	0.1%	77.9%	77.9%
10	-0.015	0.510	0.490	0.000	0.510	0.490	0.000	0.0%	1.1%	0.0%	0.0%	72.6%	72.6%
10	-0.01	0.510	0.490	0.000	0.510	0.490	0.000	0.0%	0.0%	0.0%	0.0%	70.5%	70.5%
10	-0.005	0.600	0.400	0.000	0.600	0.400	0.000	0.0%	0.0%	0.0%	0.0%	65.3%	65.3%
10	0	0.590	0.410	0.000	0.590	0.410	0.000	0.0%	0.0%	0.0%	0.0%	63.2%	63.2%
10	0.005	0.485	0.515	0.000	0.485	0.515	0.000	0.0%	1.5%	0.0%	0.1%	59.0%	59.0%
Averages								0.6%	2.7%	0.02%	0.04%	74.4%	74.1%

Table 6.7: Bilinear utility results (see caption on p.110)

Table 6.7: Bilinear utility results

This table shows results for FSO and MV portfolios using bilinear utility functions. Each row represents a different parameter setting of the utility function. The values of the parameters φ and r_p^* are given in the first two columns. The following six columns present portfolio allocations for FSO and MV portfolios. Portfolio allocations notation are as follows: θ_1 is allocation to *FTSE100*, θ_2 is to *FTSE250*, and θ_3 is to *FTSE EMI*. The given allocations are those estimated on the full data set. In the next four columns the utility improvement of using FSO rather than MV is given in terms of utility and in terms of certainty equivalents. For each of these, both in-sample results (IS) and out-of-sample results (OOS) are given. OOS results given are the averages of the two OOS applications. Certainty equivalent improvements are given in annual terms. In the two rightmost columns success rates are given for FSO and MV portfolios respectively (both in-sample for the full data set).

Utility function	Mean	Variance	Skewness	Kurtosis
Bilinear utility	0.0000	0.0000	0.0029	0.0180
S-shaped utility	0.0004	0.0001	0.0687	0.0150

Table 6.8: Average out-of-sample differences in return distribution properties between FSO and MV portfolios

The table gives the difference in portfolio return distribution properties between FSO and MV portfolios, averaged over the 20,000 bootstrapped samples and across utility function specifications.

Return distributions on average of the 20,000 bootstrapped samples have higher means and variances when optimising with FSO than with MV (see Table 6.8). As for the bilinear utility case, skewness and kurtosis are also higher, the former with much higher magnitude than under the bilinear preferences (0.069). In other words, the portfolios based on S-shaped preferences have a higher risk level but a lower probability of very low returns in a single month. The success rates are also clearly in favour of the FSO approach in cases where S-shaped utility is believed to describe the investor's preferences well, which was also found by Cremers, Kritzman, and Page (2005). In 60% of the utility specifications the FSO success rate is higher than the MV counterpart – another reflection of the increased skewness of the FSO return distribution.

Variation in the ratio γ_1/γ_2 causes only minor changes in the allocations. These parameters primarily determine the bends of the S-shape, and the influence on allocations is marginal. The variations of A and B, on the other hand, are more influential on allocations. The higher the ratio A/B gets, the less risk (in terms of variance) is chosen for the portfolio. Similar findings on the influence of S-shaped utility function parameters have been reported by Benartzi and Thaler (1995). My results also indicate that, as expected, higher inflection points correspond to higher loss aversion.

Parameters					FSO approach			MV approach			ϵ_{MV}		Δr_{CE}		Success rates	
γ_1	γ_2	A	B	r_p^*	θ_1	θ_2	θ_3	θ_1	θ_2	θ_3	IS	OOS	IS	OOS	SR_{FSO}	SR_{MV}
0.05	0.95	1.5	1.5	0	0.090	0.905	0.005	0.105	0.840	0.056	5.3%	0.2%	0.1%	0.0%	66.3%	64.2%
0.1	0.9	1.5	1.5	0	0.090	0.905	0.005	0.105	0.840	0.056	4.6%	1.8%	0.1%	0.0%	66.3%	64.2%
0.15	0.85	1.5	1.5	0	0.090	0.910	0.000	0.107	0.838	0.055	4.0%	1.8%	0.1%	0.0%	66.3%	64.2%
0.2	0.8	1.5	1.5	0	0.090	0.910	0.000	0.107	0.838	0.055	3.6%	3.2%	0.1%	0.0%	66.3%	64.2%
0.25	0.75	1.5	1.5	0	0.090	0.910	0.000	0.107	0.838	0.055	3.4%	5.4%	0.1%	0.0%	66.3%	64.2%
0.3	0.7	1.5	1.5	0	0.090	0.905	0.005	0.105	0.840	0.056	3.6%	20.6%	0.1%	0.1%	66.3%	64.2%
0.35	0.65	1.5	1.5	0	0.090	0.905	0.005	0.105	0.840	0.056	4.8%	11.2%	0.1%	0.1%	66.3%	64.2%
0.4	0.6	1.5	1.5	0	0.000	0.735	0.265	0.000	0.735	0.265	0.0%	5.9%	0.0%	0.1%	65.3%	65.3%
0.45	0.55	1.5	1.5	0	0.000	0.735	0.265	0.000	0.735	0.265	0.0%	0.0%	0.0%	0.0%	65.3%	65.3%
0.5	0.5	1.5	1.5	0	0.000	0.730	0.270	0.000	0.730	0.270	0.0%	0.0%	0.0%	0.0%	65.3%	65.3%
0.5	0.5	1.7	1.3	0	0.000	0.735	0.265	0.000	0.735	0.265	0.0%	0.5%	0.0%	0.1%	65.3%	65.3%
0.5	0.5	1.9	1.1	0	0.090	0.910	0.000	0.107	0.838	0.055	30.2%	-3.4%	0.1%	0.1%	66.3%	64.2%
0.5	0.5	2.1	0.9	0	0.090	0.910	0.000	0.107	0.838	0.055	4.7%	11.7%	0.1%	0.1%	66.3%	64.2%
0.5	0.5	2.3	0.7	0	0.095	0.905	0.000	0.111	0.835	0.054	2.2%	12.2%	0.1%	0.0%	66.3%	64.2%
0.5	0.5	2.5	0.5	0	0.130	0.870	0.000	0.144	0.810	0.046	0.5%	5.4%	0.0%	0.0%	66.3%	66.3%
0.5	0.5	2.7	0.3	0	0.205	0.795	0.000	0.214	0.757	0.030	0.3%	0.7%	0.0%	0.0%	66.3%	66.3%
0.5	0.5	2.9	0.1	0	0.205	0.795	0.000	0.214	0.757	0.030	0.2%	0.3%	0.0%	0.0%	66.3%	66.3%
0.05	0.95	1.5	1.5	-0.025	0.295	0.700	0.005	0.296	0.694	0.010	6.4%	35.8%	0.0%	0.8%	83.2%	82.1%
0.1	0.9	1.5	1.5	-0.025	0.295	0.705	0.000	0.298	0.693	0.010	5.8%	194.6%	0.0%	1.4%	83.2%	82.1%
0.15	0.85	1.5	1.5	-0.025	0.295	0.705	0.000	0.298	0.693	0.010	6.0%	228.3%	0.0%	1.8%	83.2%	82.1%
0.2	0.8	1.5	1.5	-0.025	0.295	0.705	0.000	0.298	0.693	0.010	10.1%	65.4%	0.0%	1.7%	83.2%	82.1%
0.25	0.75	1.5	1.5	-0.025	0.295	0.705	0.000	0.298	0.693	0.010	11.5%	20.7%	0.0%	1.4%	83.2%	82.1%
0.3	0.7	1.5	1.5	-0.025	0.295	0.705	0.000	0.298	0.693	0.010	2.1%	14.3%	0.0%	1.1%	83.2%	82.1%
0.35	0.65	1.5	1.5	-0.025	0.295	0.705	0.000	0.298	0.693	0.010	0.9%	24.8%	0.0%	0.8%	83.2%	82.1%
0.4	0.6	1.5	1.5	-0.025	0.295	0.705	0.000	0.298	0.693	0.010	0.6%	10.4%	0.0%	0.7%	83.2%	82.1%
0.45	0.55	1.5	1.5	-0.025	0.025	0.975	0.000	0.046	0.885	0.070	1.9%	2.0%	0.1%	0.7%	79.0%	76.8%

Table 6.9: S-shaped utility results (see caption on p.113)

111

	Parameters					FSO approach			MV approach			ϵ_{MV}		Δr_{CE}		Success rates	
γ_1	γ_2	A	B	$r_p^\star$		θ_1	θ_2	θ_3	θ_1	θ_2	θ_3	IS	OOS	IS	OOS	SR_{FSO}	SR_{MV}
0.5	0.5	1.5	1.5	-0.025		0.025	0.975	0.000	0.046	0.885	0.070	1.4%	1.1%	0.1%	0.6%	79.0%	76.8%
0.5	0.5	1.7	1.3	-0.025		0.295	0.705	0.000	0.298	0.693	0.010	0.5%	1.5%	0.0%	0.6%	83.2%	82.1%
0.5	0.5	1.9	1.1	-0.025		0.295	0.705	0.000	0.298	0.693	0.010	0.6%	3.3%	0.0%	0.6%	83.2%	82.1%
0.5	0.5	2.1	0.9	-0.025		0.295	0.705	0.000	0.298	0.693	0.010	0.7%	10.2%	0.0%	0.6%	83.2%	82.1%
0.5	0.5	2.3	0.7	-0.025		0.295	0.705	0.000	0.298	0.693	0.010	1.2%	21.8%	0.0%	0.7%	83.2%	82.1%
0.5	0.5	2.5	0.5	-0.025		0.295	0.705	0.000	0.298	0.693	0.010	10.8%	6.9%	0.0%	0.7%	83.2%	82.1%
0.5	0.5	2.7	0.3	-0.025		0.300	0.700	0.000	0.303	0.689	0.008	0.9%	18.6%	0.0%	0.7%	83.2%	82.1%
0.5	0.5	2.9	0.1	-0.025		0.365	0.635	0.000	0.365	0.635	0.000	0.0%	20.8%	0.0%	0.0%	83.2%	83.2%
0.05	0.95	1.5	1.5	-0.05		0.115	0.500	0.385	0.022	0.902	0.075	68.9%	0.0%	1.3%	0.0%	90.5%	88.4%
0.1	0.9	1.5	1.5	-0.05		0.115	0.500	0.385	0.022	0.902	0.075	270.5%	-4.0%	1.3%	-0.2%	90.5%	88.4%
0.15	0.85	1.5	1.5	-0.05		0.115	0.500	0.385	0.022	0.902	0.075	26.8%	-1.3%	1.3%	-0.4%	90.5%	88.4%
0.2	0.8	1.5	1.5	-0.05		0.115	0.500	0.385	0.022	0.902	0.075	10.0%	-0.7%	1.3%	-0.3%	90.5%	88.4%
0.25	0.75	1.5	1.5	-0.05		0.135	0.520	0.345	0.052	0.880	0.068	4.8%	-0.4%	1.0%	-0.3%	90.5%	88.4%
0.3	0.7	1.5	1.5	-0.05		0.150	0.535	0.315	0.075	0.863	0.063	2.4%	-0.2%	0.9%	-0.1%	90.5%	88.4%
0.35	0.65	1.5	1.5	-0.05		0.000	0.850	0.150	0.000	0.850	0.150	0.0%	-8.0%	0.0%	-2.0%	89.5%	89.5%
0.4	0.6	1.5	1.5	-0.05		0.000	0.835	0.165	0.000	0.835	0.165	0.0%	-3.3%	0.0%	-1.1%	89.5%	89.5%
0.45	0.55	1.5	1.5	-0.05		0.000	0.825	0.175	0.000	0.825	0.175	0.0%	-2.1%	0.0%	-1.1%	89.5%	89.5%
0.5	0.5	1.5	1.5	-0.05		0.000	0.815	0.185	0.000	0.815	0.185	0.0%	-1.4%	0.0%	-0.9%	89.5%	89.5%
0.5	0.5	1.7	1.3	-0.05		0.000	0.830	0.170	0.000	0.830	0.170	0.0%	-0.9%	0.0%	-0.4%	89.5%	89.5%
0.5	0.5	1.9	1.1	-0.05		0.000	0.840	0.160	0.000	0.840	0.160	0.0%	-0.7%	0.0%	-0.2%	89.5%	89.5%
0.5	0.5	2.1	0.9	-0.05		0.215	0.595	0.190	0.170	0.790	0.040	0.9%	-0.1%	0.3%	-0.1%	90.5%	88.4%
0.5	0.5	2.3	0.7	-0.05		0.230	0.595	0.175	0.188	0.776	0.036	2.0%	2.6%	0.2%	-0.1%	90.5%	87.4%
0.5	0.5	2.5	0.5	-0.05		0.875	0.125	0.000	0.875	0.125	0.000	0.0%	3.1%	0.0%	0.1%	90.5%	90.5%
0.5	0.5	2.7	0.3	-0.05		0.880	0.120	0.000	0.880	0.120	0.000	0.0%	0.0%	0.0%	0.0%	90.5%	90.5%
0.5	0.5	2.9	0.1	-0.05		0.880	0.120	0.000	0.880	0.120	0.000	0.0%	0.0%	0.0%	0.0%	90.5%	90.5%
Averages												10.1%	14.5%	0.2%	0.2%	79.6%	78.5%

Table 6.9: S-shaped utility results (see caption on p.113)

Table 6.9: S-Shaped utility results

This table shows results for FSO and MV portfolios using S-shaped utility functions. Each row represents a different parameter setting of the utility function. The values of the parameters γ_1, γ_2, A, B, and $r_p^\star$ are given in the five leftmost columns. The following six columns present portfolio allocations for FSO and MV portfolios. Portfolio allocations notation are as follows: θ_1 is allocation to *FTSE100*, θ_2 is to *FTSE250*, and θ_3 is to *FTSE EMI*. The given allocations are those estimated on the full data set. In the next four columns the utility improvement of using FSO rather than MV is given in terms of utility and in terms of certainty equivalent. For each of these, both in-sample results (IS) and out-of-sample results (OOS) are given. OOS results given are the averages of the two OOS applications. Certainty equivalent improvements are given in annual terms. In the two rightmost columns success rates are given for FSO and MV portfolios respectively (both in-sample for the full data set).

6.6 Conclusions

The empirical application of this chapter constitutes the widest FSO-MV comparison to date with respect to utility functions. The results extend earlier findings by Cremers, Kritzman, and Page (2005) and Adler and Kritzman (2007), establishing the robustness of those studies' results. The FSO methodology is useful when investor utility function features a threshold, such as in the bilinear and the S-shaped utility functions (especially the latter). For traditional utility functions (exponential and power utility) there is no clear performance difference between MV and FSO.

The fact that these results appear in an application of equity indexes increases the scope of the FSO applicability considerably. It has earlier only been shown that FSO is useful in allocation problems involving hedge funds, which have very non-normal return distributions. Using assets with return distributions closer to normality yields smaller gains in utility, but the improvements are still substantial.

FSO has great theoretical appeal in that it does not build on assumptions simplifying the world of the investor. Return distributions are used in their entirety and utility functions can be chosen with complete flexibility, without mathematical convenience considerations. Challenges remaining for the investment advisor include correctly specifying the investor's preferences in a utility function, and to overcome the computational burden of the technique that appears when more assets are added. I look closer at the latter issue in the next chapter.

Chapter 7

Full-Scale Optimisation using Differential Evolution: Application to FTSE100 Stocks

In this chapter I apply the heuristic technique differential evolution (DE) to solve relatively large portfolio selection problems under complex utility functions using FSO.[1] I show that this is computationally feasible and that solutions retrieved with random starting values are converging to one optimum.

Furthermore, the study constitutes the first FSO application to stock portfolio optimisation. The assets considered are the constituent stocks of the FTSE100 index. My results indicate that when investors are loss averse, FSO improves stock portfolio performance compared with MV portfolios. This finding widens the scope of applicability of FSO, but it is also stressed that out-of-sample success will always be dependent on the forecasting ability of the input return distributions.

This chapter follows the same structure as the previous chapter. I begin with presenting the problem setting and the utility functions applied. Then, in Section 7.3 I present the DE algorithm in some detail. The model evaluation methodology used is similar to that of the previous chapter, but some differences due to optimisation method and choice of utility functions are presented in Section 7.4. In the results section, 7.5, I first present findings on the reliability of DE optimisation, then I present results on FSO performance in this large stock portfolio setting. Before concluding the chapter, in Section 7.6, I provide a discussion

of FSO weaknesses.

7.1 Setting: Large Stock Portfolios

In the previous chapter I showed that FSO is useful for the choice between equity indexes and that this is robust for several different utility function specifications. That chapter extends the findings of Cremers, Kritzman, and Page (2005) and Adler and Kritzman (2007), who introduce FSO and demonstrate that it is useful for hedge fund selection. When investor preferences are complex FSO has hence repeatedly been shown to be performing better than traditional portfolio choice models. What has not been addressed by these studies is how large portfolios can be optimised in a time-efficient manner. For larger problems (with many assets), the grid search quickly becomes too burdensome computationally; and with investors interested in higher moments of returns, gradient searches are likely to get stuck in a local optimum. In the setting of FSO, there is currently no study on how to overcome the massive computational burden associated with problems of several assets. Gourieroux and Monfort (2005) state that the computational burden problem is obsolete due to modern computing capacity. Cremers, Kritzman, and Page (2003) and Cremers, Kritzman, and Page (2005) solve hedge fund selection problems of up to 61 assets, but they do not disclose what search algorithm they apply to do that. The study presented in this chapter is the first to show explicitly how the FSO optimum can be found in problems of that scale.

Large stock selection problems are common in portfolio optimisation practice. They are encountered in actively as well as passively managed portfolios (such as index tracking portfolios). For the latter it is common to choose among hundreds of stocks. In extreme cases, portfolio choice problems can include thousands of assets. For an example, see Perold (1984). By making assumptions on the shape of return distributions and on the utility functional form, optima of such large scale problems are straightforward to find using MV. In the FSO framework, however, the solution comes at great computational cost. To deal with large stock selection problems within the FSO framework, I demonstrate how a heuristic search algorithm such as DE can overcome the dimensionality problem.

Admissible assets in my empirical application are the constituents of the FTSE 100 index (by 1 Jan 2005). I use daily price data for two years (3 Jan 2005–29 Dec 2006), yielding 502 observations, retrieved from Datastream. One observation is lost when prices are transformed to returns. Due to missing data, five companies are excluded.[2] Properties of each asset return series are given in Table A.1 in the appendix. The Jarque-Bera normality test (Jarque and Bera 1980) shows that the null hypothesis of a normal probability distribution can be rejected for all assets except one (BP). The fact that return distributions deviate from normality implies scope for asset allocation superior to that of MV evaluation.

---

[2]Drax Group, Experian Group, Kazakhmys, Royal Dutch Shell A, and Standard Life.

7.2 Utility Functions

In this chapter I consider two different families of utility functions: kinked power utility and S-shaped utility (see Section 5.4 for definitions and discussion). Both of these feature loss aversion of different types.

The kinked power utility function has earlier been implemented by Maringer (2008b). To calculate this utility, returns are adjusted before they are evaluated in the utility function. As shown in Table 5.1, returns lower than a set level $r_p^\star$, are given disproportionate weight in accordance with a loss aversion parameter χ. This yields a kink on the utility function located at the critical value. A graphical example of this function can be seen in Figure 5.2, Panel D).

The second group of utility functions applied is the prospect theory S-shaped utility functions, which I also used in the previous chapter. For a graphical example, see Figure 5.2, Panel E). The utility curve is concave as long as returns are positive, but for negative returns it turns convex, showing risk seeking behaviour. Hence, under such preferences, the investor wants to take on more risk when losses are made. In spite of this irrational feature, S-shaped utility is applied in portfolio optimisation as a way of incorporating loss aversion, see e.g. Aït-Sahalia and Brandt (2001) and Cremers, Kritzman, and Page (2005).

The definitions of the two utility functions considered here are reprinted in Table 7.1. I vary the critical point $r_p^\star$ between the values -1.5%, -1%, -0.5%, and 0%. I let the loss aversion parameter for the power utility vary between $\chi = 1, 2$ and 3 (hence including a traditional power utility function without kink when $\chi = 1$), and the risk aversion parameter between $\gamma = 0, 2, 4$ and 6. This yields 36 different specifications of the power utility function.

For the S-shaped utility I use the same variations of $r_p^\star$ as above. The gamma values here regulate the curvature of the S-shape and are assumed to be equal, $\gamma_1 = \gamma_2 = 0.5$ (as they were found in the previous chapter to have little influence on portfolio allocations). The disproportion between losses and gains (relative to $r_p^\star$) is instead regulated by the magnitude parameters A and B. I vary the ratio between these parameters to $\frac{B}{A} = 1, 2$ and 3. This yields 12 different specifications of the S-shaped utility function, so in total I have 48 different functions to evaluate.

7.3 Optimisation Method: Differential Evolution

When testing all possible solutions is not practicable, heuristic optimisation techniques constitute one alternative. These are self-learning algorithms able to find optima in rough solution surfaces. Local optima are avoided by use of stochastic elements. Clearly sub-optimal solution areas are quickly by-passed, whereas areas around candidate optima are searched in detail. Maringer and Oyewumi ((2007); dealing with index tracking applications) and Maringer and

116

$$\text{Kinked power utility:} \quad r_p^{adj} = \begin{cases} r_p^\star - \chi(r_p^\star - r_p) & \text{for} \quad r_p < r_p^\star \\ r_p & \text{for} \quad r_p \geq r_p^\star \end{cases}$$

$$U_{KP} = U_{power}(r_p^{adj})$$

$$\text{S-shaped utility:} \quad U_{SSHA} = \begin{cases} -A(r_p^\star - r_p)^{\gamma_1} & \text{for} \quad r_p \leq r_p^\star \\ B(r_p - r_p^\star)^{\gamma_2} & \text{for} \quad r_p > r_p^\star \end{cases}$$

Table 7.1: Utility function equations

In both the utility functions, r_p represents portfolio return and $r_p^\star$ is the critical return level. The latter is referred to as the kink in the kinked power utility function, and as the inflection point in the S-shaped utility function. In the kinked power utility function, γ is the degree of (relative) risk aversion, and χ the degree of loss aversion. The latter is used to adjust sub-kink portfolio returns to get r_p^{adj}, which is evaluated in the power utility function. In the S-shaped utility functions, A and B determine the disproportion between returns on different sides of the critical level, and γ_1 and γ_2 are parameters determining the curvature of the S-shape (that are held constant and equal).

Meyer (2008; choosing between STAR models) have explored the efficiency and reliability of different heuristic optimisation techniques in portfolio choice problems. In the competition of threshold accepting, simulated annealing and stochastic differential equations, they find DE to be well suited for non-convex portfolio choice problems. None of these studies has assessed a FSO framework. A study closely related to this one is Maringer (2008b). It applies a FSO-like framework (not using the term FSO), but focuses on higher moment outcomes rather than utility performance differences.

DE has been shown to be efficient in converging to one (presumably global) optimum. The algorithm utilises a low number of parameters and these need little calibration, which makes it user friendly and time efficient. As independent restarts are part of the DE routine, the computation time of DE can straightforwardly be cut by running many machines parallel to each other. DE was introduced by Storn and Price (1997), and is described in detail in Price, Storn, and Lampinen (2005). A description of the DE algorithm in a portfolio choice setting is available in Maringer (2008b). Below, I give a brief summary of the six steps of the algorithm, establishing the notation and understanding necessary for my analysis. This summary is based on Maringer (2008b) and Price, Storn, and Lampinen (2005). To facilitate understanding of DE, I also provide an illustration of the algorithm, see Figure 7.1.

(1) *Initialise population*: Randomly generate a set of Q starting value vectors, θ_{ig}^P, of length N, where $i = 1, 2, ..., Q$ and Q is the population size, such that $\theta_{ig}^P \in \Omega$ (recall that Ω is the permissible set of allocation combinations and that N is the number of admissible assets). The index g denotes the vector generation which in the initialisation case is 1.

(2) *Mutation*: Generate a second set of Q vectors of length N by the equation

$\theta_{ig}^M = \theta_{k_1g}^P + (F + z_1)(\theta_{k_2g}^P - \theta_{k_3g}^P + z_2)$, where F is a difference vector scale factor, k_1, k_2 and k_3 are randomly drawn discrete numbers from the set $1, 2, ..., Q$, and z_1 and z_2 are

117

noise elements (see below for further definition of z_1 and z_2).

(3) *Offspring*: Generate a third set of Q vectors θ_{ig}^{O*} of length N with crossover probability π of equalling θ_{ig}^{P} and probability $(1-\pi)$ of equalling θ_{ig}^{M}. Adjust these vectors to satisfy the problem constraints using some function f_Ω such that $f_\Omega(\theta_{ig}^{O*}) = \theta_{ig}^{O} \in \Omega$.

(4) *Selection*: Generate a fourth set of Q vectors of length N consisting of the best solution vectors in sets θ_{ig}^{P} and θ_{ig}^{O} by using the equation $\theta_{ig}^{S} = \arg\max\{\theta_{ig}^{P}, \theta_{ig}^{O}\}(U(\theta'R))$.

(5) *Iteration*: Create a new vector generation by setting $\theta_{g+1}^{P} = \theta_{g}^{S}$ and repeat steps (2)–(4) until $g = G$, where G is the halting criterion setting how many generations should be used.

(6) *Optimum*: The DE optimum is given by $\theta^{DE} = \arg\max_{\theta_{iG}^{S}}(U(\theta'R))$. Typically, to ensure that the random starting values do not affect the DE optimum, the whole procedure is repeated several times.

Specific to the application in this chapter is that I use noise terms in the mutation (both in the scale factor and the difference vector), and a limit on number of generations as halting criterion.

The stock selection problem I assess contains 97 assets ($N = 97$). I apply a budget constraint and a short selling constraint, yielding the following constraint set (where ι is a vector of ones):

$$\Omega = \begin{cases} 0 \leq \theta_{ig} \leq 1 \\ \theta_{ig}'\iota = 1 \end{cases}. \tag{7.1}$$

These constraints are to be applied when forming θ_{ig}^{P} and θ_{ig}^{O} in Steps (1) and (3) respectively. The two constraints imply a function $f_\Omega(\cdot)$ involving two steps. First, all negative elements of θ_{ig} are set to zero, and second, all elements are divided by the solution vector sum $\theta_{ig}'\iota$. It should be noted that other constraints would make these adjustments insufficient. Price, Storn, and Lampinen (2005) discusses different approaches on how to deal with constraints.

The mutation equation scale parameter chosen is $F = 0.6$ and I add noise elements following the choices made by Maringer and Parpas (2009), generating $z_1 \sim N(0, 0.01)$ with 50% probability and zero otherwise; and $z_2 \sim Uni(-0.005, 0.005)$ with 1% probability and zero otherwise (noise parameters are redrawn in every generation). The noise elements add randomness to the algorithm, with the purpose of avoiding getting stuck in a local optimum distant from the global optimum. Population size is set to about $3 \times N$, $Q = 300$, as recommended by Maringer (2008a). The halting criterion is set to $G = 400$ for kinked power utility applications, and $G = 700$ for S-shaped utility, which was found suitable in optimisation monitoring. Crossover probability is set to $\pi = 0.6$, which is within the limits recommended in the literature. For each utility function specification I perform 5 restarts. As will be seen below, these parameter settings make the restarts converge to close to the same optimum.
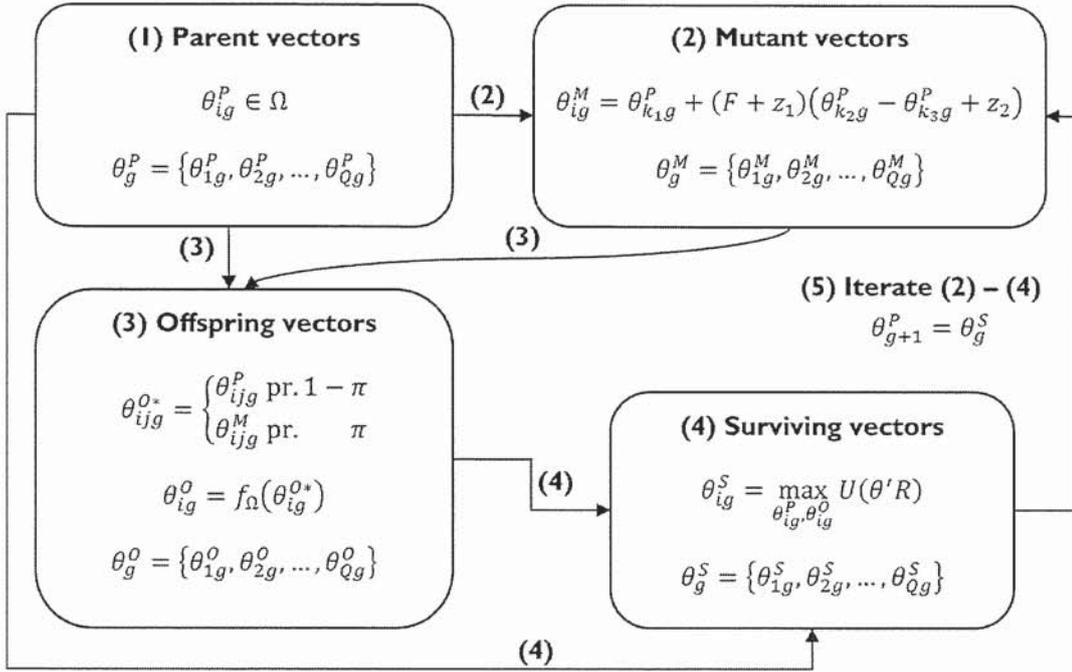
Figure 7.1: Flow chart for the DE algorithm

The figure describes the search algorithm for DE optimisation in five steps (1)–(5). The variable to be optimised is θ, which has dimension N, i.e. the number of assets. In step (1) a set of Q solution vectors are generated, denoted using index $i = 1, 2, ..., Q$. This step is done at initialisation only, implying that the generation index is $g = 1$. The superscript P indicates that these vectors are parent vectors. In step (2) a set of Q mutant vectors are generated, each being a function of three randomly chosen parent vectors and a difference vector scale factor F. Two noise terms, z_1 and z_2, introduce randomness in the vector scale factor and the difference vector. The set of mutant vectors are denoted with superscript M. In step (3) offspring vectors are generated. Each element j ($j = 1, 2, ..., N$) in each vector i have a probability $1 - \pi$ to equal element j of parent vector i, and a probability π to equal element j of mutant vector i. Once all vectors are determined, they are transformed to satisfy the constraints given by Ω using the function f_Ω. The set of Q offspring vectors are denoted with superscript O. In step (4) the surviving vectors are chosen. The surviving vector θ_{ig}^S is equal to the one of θ_{ig}^P and θ_{ig}^O that yields the highest utility according to utility function $U(\theta'R)$. In step (5) the surviving vector set θ_g^S forms the next parent vector generation θ_{g+1}^P, which is entered into step (2) as long its subscript g is smaller than G. When $g = G$, the final step (6, not illustrated) involves choosing the vector θ_{iG}^S out of the set θ_G^S, that maximises $U(\theta'R)$. This vector gives the DE optimum.

7.4 Model Evaluation

The empirical application performed in this study is assessing the performance of FSO using DE for solving a stock selection problem. The performance is evaluated towards MV, using the same evaluation procedure as applied in Cremers, Kritzman, and Page (2005) and the previous chapter of this thesis.

To compare the FSO and MV optima, the expected return of the FSO portfolio over the considered return history is calculated ($E[r_p] = E[\theta'_{FSO}R]$). A corresponding MV portfolio can then be found where the MV efficiency frontier intersects that return level. I take this to be the corresponding MV optimum, θ_{MV}. Now, by calculating portfolio return distributions for each optimum θ_{FSO} and θ_{MV}, average utility can be calculated in accordance with investor preferences. Comparing these two utilities yields a measure of the performance improvement achieved by FSO. As discussed in the previous chapter, utility differences are difficult to

119

Kinked power utility:	$r_{CE} = [1 + (1 - \gamma)\overline{U}]^{\frac{1}{1-\gamma}} - 1$	for $r_p \geq r_p^\star$
	$r_{CE} = \left([1 + (1 - \gamma)\overline{U}]^{\frac{1}{1-\gamma}} - (1 + r_p^\star)\right)\chi^{-1} + r_p^\star$	for $r_p < r_p^\star$
S-shaped utility:	$r_{CE} = r_p^\star - \left(\frac{\overline{U}}{-A}\right)^{\frac{1}{\gamma_1}}$	for $r_p \leq r_p^\star$
	$r_{CE} = r_p^\star + \left(\frac{\overline{U}}{B}\right)^{\frac{1}{\gamma_2}}$	for $r_p > r_p^\star$

Table 7.2: Certainty equivalent equations

This table shows equations for calculating certainty equivalents of utility under kinked power utility and S-shaped utility. For brevity, the certainty equivalent of the logarithmic utility function $\gamma = 1$ is not given. This case is also not applied in the empirical application of this article.

compare in an intuitive way. To get around this problem I use certainty equivalents. Following Equation 6.3, certainty equivalent expressions for each utility function can be derived. The expressions for the utility functions considered in this chapter are given in Table 7.2.

When the full utility space is searched, so that the true global optimum is located, Δr_{CE} is by definition never negative. When using a heuristic algorithm I can never be certain that the optimum located is the global optimum, I only get an approximation. Hence, under this optimisation scheme it is possible that the difference in certainty equivalents (Δr_{CE}) takes on a negative value.

As in the previous chapter, when comparing portfolios resulting from FSO and MV respectively, I study differences in the portfolio return distribution's moments. As the MV portfolio choice is based on the FSO portfolio return sample mean, the mean will be the same for the two portfolios. Variance, skewness and kurtosis differences are, however, interesting to interpret behaviour resulting from different utility function specifications.

No out-of-sample evaluation is performed in this chapter. I discuss the reasoning for this in Section 7.6.

7.5 Results

In this section I first discuss the performance of the search algorithm used to find the full-scale optimum. After that, I turn to evaluate the performance of FSO in the large stock portfolio optimisation problem applied.

7.5.1 Differential Evolution Performance

The true full-scale optimum is by definition always superior or equal to the MV portfolio when compared in-sample, as the MV solution is on the utility surface searched by FSO. The optimum found using DE is an approximation of the full-scale optimum, and its precision is dependent on the parameter settings in the algorithm. There is an obvious trade-off between

solution quality and computational cost when using DE. If a high cost can be afforded, the speed of convergence can be set low and the dispersion of candidate solutions can be set high. In practice, there are few managers who would find it worthwhile to act on portfolio weight differences smaller than a tenth of a percentage point, as administrative costs of such trades would be too high. In this section I assess the performance of DE solutions to FSO in terms of optimum stability over restarts (precision) and computational cost (in terms of time).

To analyse the precision of DE, it is interesting to look at the variation in results over restarts. As the starting point of this heuristic optimisation is random, consistency in outcomes would indicate that the number of restarts can be set low. I perform five restarts for each utility function specification and rank the solutions by utility. To study the solution stability, I measure the difference between the best and the worst solution for each utility specification. Differences across restarts in terms of portfolio weights and annual certainty equivalents are given in Tables 7.3 (kinked power utility functions) and 7.4 (S-shaped utility functions). For kinked power utility functions, the allocations to one stock never differ by more than 3 percentage points, and the performance difference in terms of certainty equivalents is less than 0.08% in all cases. This indicates that the solutions are close to the same in all restarts of the algorithm. Under the S-shaped utility specification, larger differences are observed. In one case the allocations to one stock differ by 27 percentage points between the DE optimum and the other candidate portfolios. Under such circumstances further restarts can be useful. The performance differences between restarts are in general lower when using kinked power utility than when using S-shaped utility (ranging up to 0.59%). The average certainty equivalent differences are 0.02% and 0.15% respectively. I conclude from these results that under the current DE parameter settings, it is useful with some restarts (e.g. 5), but that there is little reason to increase the number, except for a couple of cases.

An alternative to random starting points is to start with an educated guess (e.g. the MV portfolio), which may speed up the search process, but at the risk of getting stuck in a local optimum. Another approach would be to use a vector of an optimally dispersed portfolio as starting values, to limit the risk of missing corners of the search surface. Fixed starting values limit the benefit of restarting the algorithm and are primarily useful when restarts are computationally expensive. For my purposes, restarts do not come at great cost. However, choosing fixed starting values is an area of potential future research.

In order to illustrate the computational requirements for DE in an application like mine I provide measures of the time needed for the algorithm under each type of utility function. I relate this to the number of function evaluations (FE) needed, which is the product of population size (Q), number of generations in halting criterion (G), number of restarts (R), and number of scenarios considered (T). In Table 7.5, I present the settings used and the resulting computational cost expressed in number of functional evaluations and in compu-

Parameters			Maximum absolute weight differences				
$r_p^\star$	γ	χ	$\lvert\theta_{n,1}-\theta_{n,2}\rvert$	$\lvert\theta_{n,1}-\theta_{n,3}\rvert$	$\lvert\theta_{n,1}-\theta_{n,4}\rvert$	$\lvert\theta_{n,1}-\theta_{n,5}\rvert$	$\Delta r_{CE(1,5)}$
-	0	1	0.00	0.00	0.00	0.00	0.01%
-	2	1	0.00	0.00	0.01	0.01	0.01%
-	4	1	0.01	0.01	0.01	0.01	0.00%
-	6	1	0.01	0.00	0.01	0.01	0.00%
0	0	2	0.01	0.02	0.01	0.01	0.01%
0	2	2	0.01	0.01	0.01	0.01	0.01%
0	4	2	0.01	0.01	0.01	0.01	0.02%
0	6	2	0.01	0.01	0.01	0.01	0.01%
0	0	3	0.01	0.02	0.01	0.01	0.01%
0	2	3	0.01	0.01	0.01	0.02	0.02%
0	4	3	0.01	0.02	0.01	0.02	0.01%
0	6	3	0.02	0.01.	0.01	0.01	0.01%
-0.005	0	2	0.01	0.02	0.01	0.01	0.03%
-0.005	2	2	0.02	0.02	0.01	0.02	0.03%
-0.005	4	2	0.01	0.01	0.01	0.01	0.02%
-0.005	6	2	0.01	0.01	0.02	0.02	0.03%
-0.005	0	3	0.01	0.01	0.01	0.01	0.05%
-0.005	2	3	0.01	0.02	0.01	0.01	0.07%
-0.005	4	3	0.01	0.01	0.01	0.02	0.05%
-0.005	6	3	0.02	0.01	0.02	0.02	0.02%
-0.01	0	2	0.01	0.03	0.03	0.01	0.01%
-0.01	2	2	0.01	0.01	0.01	0.01	0.01%
-0.01	4	2	0.01	0.02	0.02	0.01	0.01%
-0.01	6	2	0.01	0.01	0.01	0.01	0.01%
-0.01	0	3	0.01	0.01	0.01	0.01	0.04%
-0.01	2	3	0.01	0.03	0.01	0.01	0.02%
-0.01	4	3	0.02	0.02	0.01	0.01	0.02%
-0.01	6	3	0.02	0.02	0.02	0.03	0.02%
-0.015	0	2	0.00	0.00	0.00	0.00	0.00%
-0.015	2	2	0.00	0.01	0.00	0.00	0.00%
-0.015	4	2	0.01	0.01	0.02	0.01	0.00%
-0.015	6	2	0.01	0.01	0.01	0.01	0.00%
-0.015	0	3	0.00	0.01	0.01	0.01	0.01%
-0.015	2	3	0.01	0.02	0.00	0.01	0.02%
-0.015	4	3	0.02	0.01	0.02	0.01	0.03%
-0.015	6	3	0.01	0.01	0.01	0.01	0.01%

Table 7.3: Stability of DE solutions under kinked power utility functions

In this table the DE solution for each utility specification is compared to each of the other four candidate solutions. The comparison is performed in terms of portfolio weights and in terms of annual certainty equivalent differences. The three leftmost columns are parameter settings for the kinked power utility function. $r_p^\star$ and χ are used for sub-kink return adjustment, $r_p^{adj} = r_p^\star - \chi(r_p^\star - r_p)$ for $r_p < r_p^\star$; $r_p^{adj} = r_p$ for $r_p \geq r_p^\star$. Adjusted returns are inserted in the utility function $U = \left((1+r_p^{adj})^{1-\gamma} - 1\right)/(1-\gamma)$. The next four columns are showing the largest absolute differences in portfolio weights between the DE solution and each of the other candidate solutions, ordered by utility from 1 to 5. Finally, the rightmost column shows the performance difference between the DE solution and the worst candidate solution in terms of annual certainty equivalents.

Parameters					Maximum absolute weight differences				
$r_p^\star$	γ_1	γ_2	A	B	$\lvert\theta_{n,1}-\theta_{n,2}\rvert$	$\lvert\theta_{n,1}-\theta_{n,3}\rvert$	$\lvert\theta_{n,1}-\theta_{n4}\rvert$	$\lvert\theta_{n,1}-\theta_{n,5}\rvert$	$\Delta r_{CE(1,5)}$
0	0.5	0.5	1.5	1.5	0.04	0.04	0.04	0.04	0.06%
0	0.5	0.5	1	2	0.27	0.27	0.27	0.27	0.27%
0	0.5	0.5	0.7	2.1	0.00	0.00	0.00	0.00	0.00%
-0.005	0.5	0.5	1.5	1.5	0.02	0.07	0.07	0.03	0.59%
-0.005	0.5	0.5	1	2	0.03	0.02	0.03	0.04	0.23%
-0.005	0.5	0.5	0.7	2.1	0.02	0.02	0.04	0.05	0.16%
-0.01	0.5	0.5	1.5	1.5	0.00	0.01	0.01	0.02	0.26%
-0.01	0.5	0.5	1	2	0.00	0.01	0.01	0.02	0.16%
-0.01	0.5	0.5	0.7	2.1	0.03	0.03	0.03	0.02	0.12%
-0.015	0.5	0.5	1.5	1.5	0.00	0.00	0.00	0.00	0.00%
-0.015	0.5	0.5	1	2	0.00	0.00	0.00	0.00	0.00%
-0.015	0.5	0.5	0.7	2.1	0.01	0.00	0.01	0.00	0.00%

Table 7.4: Stability of DE solutions under S-shaped utility functions

In this table the DE solution for each utility specification is compared to each of the other four candidate solutions. The comparison is performed in terms of portfolio weights and in terms of annual certainty equivalent differences. The five leftmost columns are parameter settings for the S-shaped utility function, given by $U = -A(r_p^\star - r_p)^{\gamma_1}$ for $r_p \leq r_p^\star$ and $U = B(r_p - r_p^\star)^{\gamma_2}$ for $r_p > r_p^\star$. The next four columns are showing the largest absolute differences in portfolio weights between the DE solution and each of the other candidate solutions, ordered by utility from 1 to 5. Finally, the rightmost column shows the performance difference between the DE solution and the worst candidate solution in terms of annual certainty equivalents.

Utility function	Q	G	R	T	FE	Time (seconds)
Kinked power utility:	300	400	5	501	3.01×10^8	310
S-shaped utility:	300	700	5	501	5.26×10^8	592

Table 7.5: Computational cost of FSO using DE

Time is specified in seconds. P is population size in DE; G is number of generations in DE; R is number of restarts performed of the DE algorithm; T is the number of scenarios considered. FE is number of function evaluations, defined as the product population size, number of generations, number of restarts, and number of scenarios, given by the equation $FE = QGRT$. The platform used was an Intel Core 2, 2.16 GHz processor with 3GB RAM. Software used was R v2.6.1, applying the function system.time. The time given in the table is the average of two different parameter specifications of each utility function.

tation time. This measure of computational cost in terms of time can also be compared to those in the previous chapter. In that three asset portfolio choice problem, which I solve using a grid search with 0.5% precision over 96 scenarios, the computations took 45 seconds when applied to S-shaped utility. The time cost for my application in this chapter, with 97 assets, no restrictions on precision and 501 scenarios was about 10 minutes.[3] A cluster of 5 processors of the same speed could bring this computational time down to 2 minutes. This shows that the DE algorithm is immensely more efficient than the grid search.

Having established that DE produces stable optima in a time efficient manner, I now turn to looking at the second issue of this chapter: how good the FSO solutions are in stock selection problems.

---
[3]The time measures are directly comparable, as the same computer facility was used for the two studies.

7.5.2 FSO Performance

Results of model comparisons for the 48 chosen utility function specification are given in Table 7.6 (kinked power utility) and Table 7.7 (S-shaped utility). Here, only the best (elitist) solution for each specification is considered. I find no case of MV portfolios yielding higher utility than FSO using DE, which is a first indication that the DE approximation is close to the FSO optimum (as the FSO optimum by definition yields a higher or equal utility than the MV optimum).

I evaluate kinked power utility functions at different levels of the critical value (the kink, $r_p^\star$) and different penalties (χ) on the sub-kink returns, with different levels of risk aversion (γ). When the penalty parameter is set to one, I get a power utility function without kink. It has earlier been found that FSO under power utility yields negligible improvements compared to MV performance (see discussion in Sections 5.1 and 6.5.1). This is confirmed here, as no utility improvement is seen, regardless of level of risk aversion. When the power utility functions are kinked, however, FSO consistently shows substantially better performance than MV. The difference is consistently larger the stronger the loss aversion χ is. For example, with $r_p^\star = -1\%$, the utility difference between the FSO and MV portfolios correspond to an annual certain return of 0.9% on average when $\chi = 2$, and 2.8% on average when $\chi = 3$. The CE difference is the risk free annual return that would yield the same utility improvement as the improvement achieved when using FSO instead of MV.

Looking at the portfolio properties resulting from FSO and MV respectively (Table 7.6 show how FSO portfolio moments compare to those of MV portfolios), it can be seen that investors with kinked power preferences in general take on portfolios with higher variance (and hence lower Sharpe ratio, as the mean return is fixed) than MV investors do. This is a consequence of taking higher moments into account, and I observe the same pattern under S-shaped preferences. Investors whose preferences correspond to the kinked power utility function are averse to risk in general, and in particular to downside risk — losses are heavily penalised (especially large losses). This behaviour implies that the investor pursues positive skewness, trying to avoid assets with large risk of extreme negative returns. Such skewness preference is well documented in the literature, see discussion in Section 5.4. The risk aversion implies aversion towards variance and kurtosis (as shown in the Taylor expansion of the utility function above). In my application, I see that the incentive to achieve positive skewness is stronger than the incentive to limit variance and kurtosis (loss aversion dominates risk aversion). This holds regardless of critical level, risk aversion, and penalty level on sub-kink returns. Naturally, the effect is stronger (skewness is higher) when the loss aversion parameter χ is higher. Furthermore, when the kink $r_p^\star$ is lower, the difference in portfolio return skewness between FSO and MV portfolios tend to be higher (with a few exceptions). The number of stocks with allocations higher than 0.1% is also in general higher in the

MV portfolios than those optimised under kinked power utility functions, indicating that diversification is more important for MV investors. This relates to the MV focus on variance as sole measure of risk.

Results for portfolios optimised for S-shaped utility functions are given in Table 7.7. The FSO utility improvement as compared to the corresponding MV portfolio is highest when the inflection point $r_p^\star$ is set at zero. For lower inflection points, the annual certainty equivalent improvement is decreasing but is still substantial. Certainty equivalent improvements are on average higher than 10% when the inflection point is at zero, 6.4% when $r_p^\star = -0.5\%$, 2.6% when $r_p^\star = -1\%$, and 1.5% when $r_p^\star = -2\%$. The reason that the difference is decreasing as the inflection point gets lower, is probably that there are fewer observations in daily return series that fall short of the lower inflection points. This decreases the scope for avoiding such returns, thus making loss aversion at this level less influential.

Investors adhering to S-shaped utility are primarily concerned with getting returns exceeding the critical value (the inflection point). As opposed to kinked power utility, large losses are not penalised much more than small losses (below the inflection). Hence, a return distribution with high median is pursued at the cost of relatively high risk of large losses. In terms of moments, this implies acceptance of negative skewness, and little emphasis on variance and kurtosis. The next three columns of Table 7.7 show how FSO portfolio moments compare to those of MV portfolios. It turns out that skewness is indeed more negative in FSO portfolios when the inflection point is 0% and −0.5%. This is not the case for lower inflection points, which can again be explained with the limited scope for optimisation as there are fewer daily returns falling short of these critical values. Diversification, measured as number of stocks with weights exceeding 0.1%, does not differ consistently between S-shaped and MV optimised portfolios.

7.6 FSO in Practice

There are two main points of critique towards methodologies such as FSO. Firstly, its computational cost makes it expensive for many applications. I have shown in this chapter how this cost can be substantially alleviated using heuristic optimisation techniques. Secondly, FSO can be criticised for its dependence on the empirical return distribution as predictor of the future return distribution. Previous studies assessing the performance of FSO have performed out-of-sample studies to confirm the performance achieved in-sample (see Chapter 6 and Adler and Kritzman, 2007). These studies found that the performance improvement of FSO persists out-of-sample, but that it is not as large as in-sample. However, as was pointed out by Adler and Kritzman (2007), FSO performance out-of-sample is dependent on how persistent the properties of the return distributions are over time. Skewness and kurtosis in the historical sample of returns may not forecast skewness and kurtosis in the future. For

Parameters				Distribution characteristics			Stocks used	
$r_p^\star$	γ	χ	Δr_{CE}	ΔVariance	ΔSkewness	ΔKurtosis	FSO	MV
-	0	1	0.0%	0.000000	0.00	0.06	1	2
-	2	1	0.0%	0.000001	0.10	1.00	4	3
-	4	1	0.0%	0.000000	0.08	0.57	6	6
-	6	1	0.0%	0.000000	0.06	0.28	7	7
0	0	2	0.5%	0.000001	0.04	0.50	20	20
0	2	2	0.4%	0.000001	0.07	0.42	18	20
0	4	2	0.4%	0.000001	0.06	0.35	19	20
0	6	2	0.4%	0.000001	0.09	0.40	19	23
0	0	3	0.6%	0.000002	0.06	0.56	25	26
0	2	3	0.6%	0.000001	0.06	0.53	28	27
0	4	3	0.5%	0.000001	0.08	0.52	26	27
0	6	3	0.5%	0.000001	0.09	0.51	25	28
-0.005	0	2	1.1%	0.000003	0.30	0.84	16	17
-0.005	2	2	0.9%	0.000002	0.27	0.65	18	18
-0.005	4	2	0.7%	0.000002	0.19	0.12	19	20
-0.005	6	2	0.8%	0.000001	0.20	0.15	19	20
-0.005	0	3	1.9%	0.000002	0.22	-0.11	20	26
-0.005	2	3	2.1%	0.000002	0.23	-0.05	22	27
-0.005	4	3	2.1%	0.000002	0.24	-0.15	20	27
-0.005	6	3	2.2%	0.000002	0.24	-0.09	21	27
-0.01	0	2	0.8%	0.000003	0.28	0.61	9	11
-0.01	2	2	0.8%	0.000004	0.32	0.45	10	13
-0.01	4	2	0.9%	0.000003	0.31	0.29	11	16
-0.01	6	2	1.2%	0.000003	0.32	0.16	11	17
-0.01	0	3	2.7%	0.000005	0.40	-0.02	11	18
-0.01	2	3	2.8%	0.000005	0.39	-0.06	11	18
-0.01	4	3	2.9%	0.000005	0.34	-0.30	12	20
-0.01	6	3	2.8%	0.000004	0.33	-0.31	13	20
-0.015	0	2	0.9%	0.000007	0.30	0.60	6	7
-0.015	2	2	0.9%	0.000007	0.35	0.89	6	9
-0.015	4	2	0.7%	0.000006	0.38	1.02	7	10
-0.015	6	2	0.8%	0.000006	0.43	1.15	8	12
-0.015	0	3	2.6%	0.000012	0.69	2.03	6	13
-0.015	2	3	2.6%	0.000011	0.72	2.39	10	13
-0.015	4	3	2.6%	0.000011	0.77	2.40	9	16
-0.015	6	3	2.5%	0.000011	0.82	2.63	8	17

Table 7.6: Results for kinked power utility

This table shows results of the empirical application performed under kinked power utility. The three leftmost columns are parameter settings for the kinked power utility function. $r_p^\star$ and χ are used for sub-kink return adjustment, $r_p^{adj} = r_p^\star - \chi(r_p^\star - r_p)$ for $r_p < r_p^\star$; $r_p^{adj} = r_p$ for $r_p \geq r_p^\star$. Adjusted returns are inserted in the utility function $U = \left((1 + r_p^{adj})^{1-\gamma} - 1)\right)/(1 - \gamma)$. The column denoted Δr_{CE} describes differences in utility between the FSO and the MV solutions, expressed in certainty equivalent differences. Positive numbers indicate that FSO is better than MV. The next three columns are differences in variance, skewness and kurtosis between the FSO and MV portfolio return distributions. Positive numbers indicate that the FSO portfolio measure is larger. Note that mean returns are by definition the same for FSO and MV portfolios. The two rightmost columns shows the number of stocks with allocations $> 0.1\%$ for FSO and MV portfolios respectively.

Parameters						Distribution characteristics			Stocks used	
$r_p^\star$	γ_1	γ_2	A	B	Δr_{CE}	ΔVariance	ΔSkewness	ΔKurtosis	FSO	MV
0	0.5	0.5	1.5	1.5	3.3%	0.000030	-0.35	1.36	20	13
0	0.5	0.5	1	2	10.8%	0.000323	-0.85	-7.27	2	4
0	0.5	0.5	0.7	2.1	18.2%	0.000333	-0.85	-7.31	2	4
-0.005	0.5	0.5	1.5	1.5	4.4%	0.000005	-0.19	0.74	32	27
-0.005	0.5	0.5	1	2	7.2%	0.000012	-0.28	1.00	30	17
-0.005	0.5	0.5	0.7	2.1	7.5%	0.000020	-0.26	1.32	27	17
-0.01	0.5	0.5	1.5	1.5	2.4%	0.000002	0.08	0.19	21	24
-0.01	0.5	0.5	1	2	2.8%	0.000003	-0.03	0.36	22	20
-0.01	0.5	0.5	0.7	2.1	2.6%	0.000004	0.02	0.93	22	20
-0.015	0.5	0.5	1.5	1.5	2.8%	0.000002	0.26	-0.05	20	23
-0.015	0.5	0.5	1	2	2.0%	0.000002	0.25	0.06	22	23
-0.015	0.5	0.5	0.7	2.1	1.8%	0.000002	0.17	0.36	22	23

Table 7.7: Results for S-shaped utility

This table shows results of the empirical application performed under S-shaped utility. The five leftmost columns are parameter settings for the S-shaped utility function, given by $U = -A(r_p^\star - r_p)^{\gamma_1}$ for $r_p \leq r_p^\star$ and $U = B(r_p - r_p^\star)^{\gamma_2}$ for $r_p > r_p^\star$. The column denoted Δr_{CE} describes differences in utility between the FSO and the MV solutions, expressed in certainty equivalent differences. Positive number indicates that FSO is better than MV. The next three columns are differences in variance, skewness and kurtosis between the FSO and MV portfolio return distributions. Positive numbers indicate that the FSO portfolio measure is larger. Note that mean returns are by definition the same for FSO and MV portfolios. The two rightmost columns shows the number of stocks with allocations $> 0.1\%$ for FSO and MV portfolios respectively.

example, skewness may be generated by one very large outlier that is unlikely to be repeated. The historical observations may be viewed as estimators of future scenarios. The extent to which in-sample performance does not persist out-of-sample is called *estimation error* by Cremers, Kritzman, and Page (2005). In fairness, however, this line of critique applies to all portfolio choice models based on historical return information. No matter if the portfolio choice is based on estimates of a theoretical return distribution (such as Markowitz's, 1952, assumption of normally distributed returns or Harvey, Liechty, Liechty, and Muller, 2004, fitting of skew normal distributions) or empirical return distributions, it will be dependent on historical observations ability to predict the future. The quality of a model's output will never be higher than its input data. On the basis of this discussion, I do not perform an out-of-sample study in this chapter.

In the MV framework the co-variance matrix plays an important role, and Harvey, Liechty, Liechty, and Muller (2004) emphasises the role of co-skewness. The FSO framework implicitly takes all these co-movements into account when treating historical observations as potential future scenarios. If the investor expects a return different from the historical average, he can adjust the historical returns to reflect this. Similar adjustments can be done if the investor expects higher moments to differ from their historical levels, without affecting the information on co-movements of asset prices. This discussion highlights the similarity between FSO and Sharpe's (2007) scenario-based utility maximisation, where mean, variance and correlation adjustments are described as common in practice. Another possible adjustment at hand for a practitioner would be to impose a weighting scheme on observations, so that more recent

information is given higher influence in the optimisation (i.e. relaxing the assumption of equal probability T^{-1} of all scenarios).

FSO is a technique that has been developed for the financial industry and it is applied in different forms by several practitioners. A substantial amount of research has focussed on how to forecast financial returns and their higher moments. If this knowledge can be implemented in the FSO inputs, the benefits of the technique should become even clearer. I encourage future research to look at how this can be done to limit the estimation error. The results presented above should be viewed as indications that there are considerable gains to be made by optimising stock portfolios using FSO, when higher moments are important to the investor. These gains can be realised if the investor is able to construct good estimators of future returns.

7.7 Conclusions

This chapter contributes to the assessment of the FSO technique in two respects. It shows that:

1. DE yields approximations that appear to converge to the FSO optimum;

2. using DE to solve FSO problems the computational burden inherent in the technique can be handled, and hence its applicability can be expanded to stock portfolio optimisation (of up to 100 stocks).

FSO has great theoretical appeal as simplifying assumptions on expected return distributions and investor preferences are avoided. This yields however a non-convex problem of great dimension, in particular when many assets are considered and when investor preferences are complex. The computational burden inherent in the FSO framework is its main drawback.

By introducing DE in this framework, I show that the computational burden of FSO can be shortcut and broken down to achieve short computing times. This is important for transferring the technique to the financial industry, and constitutes the main contribution of this article. DE is easy to calibrate and can be divided between computers. I show that the approximations it gives of the FSO optimum is stable, and that for most problems no more than 5 restarts are necessary.

This chapter also constitutes the first application of FSO in a pure stock selection problem. Previous assessments have dealt with hedge fund selection and equity index selection problems. I show that the in-sample utility improvement identified in those studies persist in the stock selection setting. I also provide a discussion about out-of-sample performance, claiming that it is dependent on how persistent the historical returns' properties are over time, and encouraging practitioners to adjust the empirical distribution to reflect their analyst's expectation of the future.

Utility functions considered in this chapter feature loss aversion and prospect theory. Loss aversion has gained increasing attention recently, and different types of this concept are captured with the chosen utility functions. My application shows how each type of loss aversion influences the preferred portfolio properties: investors following kinked power utility strive for positive skewness, whereas S-shaped utility investors are indifferent towards skewness.

The computational problem of FSO has been called the curse of dimensionality. With heuristic algorithms such as DE this problem can be substantially alleviated.

Part IV

Conclusions

Chapter 8

Concluding Remarks and Policy Implications

In this final chapter of the thesis, I summarise the main findings on liquidity (Section 8.1) and portfolio optimisation (Section 8.2). In Section 8.3 I then sketch a framework for future analysis of how liquidity costs influence the portfolio choice. This highlights how the two topics in this thesis do intersect and influence each other. Finally, in Section 8.4, I point out further directions for future research.

8.1 Liquidity

Liquidity is a concept describing the friction investors experience when trading securities. This definition does not specify what is meant by friction, thereby opening for discussion on how liquidity should be measured. As discussed in Chapter 2, liquidity has many dimensions and can be measured in many different ways (each way capturing different dimensions of the concept). Many pages in the liquidity literature have been dedicated to the choice of liquidity measure. This thesis presents that discussion, but its contribution lies in improved insights in (a) the data processing preceding liquidity measurement; and once the individual liquidity measures have been derived (b) the measurement of systematic liquidity factors.

The first of these two contributions is minor, but nevertheless important. Everyone studying liquidity using high-frequency data needs to go through extensive data processing. Until now, little guidance is available in the literature on how to do this. Chapter 3 can be used as a step by step guide on how to process TAQ data. The chapter aim is to provide a transparent exposition of the necessary steps towards a data set that can be used for reliable liquidity measurement. The growth of the TAQ database in recent years has created new problems of data processing that have been scarcely covered elsewhere. My presentation contains novel

discussions on how to treat such problems, including handling of simultaneous observations and detection of erroneous observations.

It is well-known that liquidity is co-varying in the cross-section of stocks, implying some underlying liquidity market force. Systematic liquidity shocks, i.e. changes in that market force, are recognised as an important risk factor in asset pricing (Acharya and Pedersen 2005, Korajczyk and Sadka 2008, Pastor and Stambaugh 2003). In order to handle this liquidity risk it is crucial to have accurate approximations of the systematic liquidity factors. Herein lies my second contribution on liquidity measurement (Chapter 7). The purpose of systematic liquidity factors is to capture underlying forces that induce co-variance between stocks' individual liquidity costs. Hence, it is natural to base the analysis on the co-variance matrix, which I do by using PCA. This approach is however only appropriate as long as the sample co-variance forms a good approximation of the true co-variance structure. Previous literature has assumed that this is the case, i.e. that co-variances are constant over time. In my analysis, I allow for time-varying co-variances by introducing dynamic estimation window PCA.

Analysing different techniques for derivation of systematic liquidity factors, my findings differ across liquidity measures. For many liquidity measures, including quoted and effective spreads and turnover, the cross-sectional average is a good approximation of systematic liquidity. For price impact measures, including proxies for inventory cost and adverse selection cost, I find that moving window PCA is a better methodology. In general, PCA-based measures are better than the cross-sectional average for explaining return variations. I find general support for using expanding window PCA rather than static window PCA, as these methods converge after approximately 8 years, and the latter has a forward-looking bias. One measure stands out in the investigation: fixed inventory costs. When systematic liquidity in this measure is derived using moving window PCA, strong explanatory power is found for both liquidity variation and return variation.

The contributions made here on liquidity measurement are important for the understanding of liquidity in itself and its impact on asset prices. Extensive processing and computation is necessary to achieve high quality measures, and there is reason for a practitioner to ask whether all the work is worthwhile. My investigation shows that a computationally expensive technique such as moving window PCA applied to a liquidity measure derived from high frequency data explains stock returns better than any other approach. This should be relevant for any fund manager and investment adviser that aim to handle liquidity risk.

8.2 Portfolio Optimisation

FSO is a portfolio optimisation framework where the empirical return distribution is used as a set of future scenarios. These scenarios are the basis for expected utility maximisation.

This method is attractive in that the utility function and portfolio constraints can be specified without the restriction of mathematical convenience often seen in traditional portfolio optimisation frameworks. The solution is approximated by application of a search algorithm to the set of candidate portfolio allocations. Previous literature has shown empirically that when investors are loss averse, FSO portfolios differ substantially from MV portfolios, and that the former yield higher utility (Cremers, Kritzman, and Page 2005, Adler and Kritzman 2007). The existing empirical evidence is however limited to hedge fund selection problems and a small selection of utility function specifications. Hedge fund returns are typically featuring large deviations from the normal distribution, larger than what is typical for e.g. individual stock returns. Furthermore, these previous articles do not show explicitly how the full-scale optimum is found and to what extent the computational burden is a drawback of FSO.

In two empirical studies I address these gaps in the literature (Chapters 6 and 7). Firstly, I extend the test ground by performing one application on equity index portfolio selection (a three asset portfolio) and one on stock portfolio selection (the FTSE100 constituent stocks). In comparison to MV, I find that FSO portfolios yield higher expected utility in-sample in the setting of stock indexes as well as in the setting of individual stocks. This widens the applicability of FSO to equity indexes and stocks. Secondly, I consider a wide range of utility function parameters, which adds to the robustness of the results as well as the understanding of the utility function parameters. Thirdly, I show explicitly how the optimum can be found using either a grid search or a heuristic algorithm such as DE, and I give measures of the computational burden associated with each method. The application of DE shows that FSO can be solved in relatively large portfolio problems with good computational economy and high solution stability. This is an important contribution to the FSO literature, as no previous study has shown explicitly how the computational burden of the technique can be handled.

Out-of-sample evidence showing further robustness is provided in the first of the two empirical studies. That application shows that the scenarios used in the estimation of portfolio allocations are good approximations of the future. However, I also argue that the portfolio estimated in FSO will be optimal only to the extent that the empirical return distribution is a good approximation of future events. This limitation applies to any portfolio choice framework that utilises historical data.

FSO is a technique with roots in the investment industry, and the findings here are hence of clear interest to practitioners, but for successful management its limitations also need to be kept in mind.

8.3 A Framework for Liquidity-Adjusted Portfolio Optimisation

In this thesis I address problems of liquidity and portfolio optimisation in isolation, but the two problems are actually intimately related in the investment decision. For instance, how can a portfolio be said to be optimised if trading costs are ignored? Including assets with different levels of liquidity in a portfolio choice problem without adjusting for trading costs will, ceteris paribus, due to the illiquidity return premium result in a portfolio with high weight to illiquid assets. This is misleading, as the trading costs of illiquid assets in practice undermine the returns. In this section I aim to sketch a framework that can be used as a starting point for combining the findings of this thesis into one model.

Several asset pricing studies provide frameworks for calculating liquidity risk premiums. Using either the liquidity-adjusted CAPM (Acharya and Pedersen 2005) or a Fama and French (1993) style factor model with systematic liquidity (as in Korajczyk and Sadka 2008), liquidity risk premiums can be derived. By adjusting expected returns for these risk premiums, traditional portfolio optimisation models can straightforwardly be applied to retrieve liquidity-adjusted portfolios. The weakness of augmenting the asset prices for liquidity characteristics is that liquidity is then handled in a risk management framework. As was pointed out by Amihud and Mendelson (1986b), diversification does not mitigate liquidity costs. On the contrary, there is a trade-off between diversification (risk management) and liquidity costs.

According to the clientele effect hypothesised by Amihud and Mendelson (1986a), the excess returns from illiquid assets should be relatively attractive to long-term investors. The intuition of this is that the high liquidity costs of illiquid assets can be spread over a long time when the holding period is longer. Then the trade-off between excess returns and liquidity costs will be favourable. For investors with a short holding period, on the other hand, the liquidity costs will dominate the excess returns. Swensen (2000) discusses how long term investors can benefit from the illiquidity return premium. His investment model for very long term investors (such as university endowments) utilises illiquid investments to achieve returns superior to bond and public stock investments.

In order to account for this feature of liquidity cost sensitivity, a natural starting point for the portfolio optimisation problem is to formulate a utility function describing the trade-off between risk management and liquidity costs, as well as the trade-off between illiquidity premiums and the investment horizon. That function could in general be formulated as

$$U_L = f\left(r_p, \frac{L_p}{IH}\right),\tag{8.1}$$

where IH is the investment horizon, L_p the portfolio liquidity cost, and r_p the portfolio return. The ratio $\frac{L_p}{IH}$ is the portfolio trading cost amortised over the investment horizon.

How to determine the functional form and the parameters of the utility function would be the core problem of research in this direction. The utility function would typically be strictly increasing in portfolio returns and decreasing in the amortised liquidity costs. The latter implies a growing preference for illiquid stocks as the investment horizon grows.

Once the utility function is specified, one way to optimise the portfolio is to maximise expected utility by applying FSO. Using empirical distributions of both returns and liquidity costs as inputs gives a model that accounts for risk with respect to both returns and liquidity, including co-variance structures and higher moments. Other utility-maximising methods could also be considered.

Portfolio optimisation with respect to liquidity has earlier been approached by Longstaff (2001), who analyses the effect of market restrictions on trade (such as dry-ups); and Ghysels and Pereira (2008), who investigate the impact of liquidity on portfolio allocations under different investment horizons. Further insights are given by Domowitz, Hansch, and Wang (2005), demonstrating that liquidity volatility and its relation to return co-variance are important factors for effective diversification of a portfolio. The model proposed above should be evaluated in relation to these earlier approaches. If it can be shown to handle liquidity risk successfully, it can be a useful tool for fund managers as well as for investment advice to institutional investors.

8.4   Recommendations for Future Research

Liquidity has received considerable attention since Amihud and Mendelson's (1986a) seminal paper, and the systematic liquidity literature has also grown large. In spite of this, several directions of further research on systematic liquidity are possible. Analysis of the suggested systematic liquidity factors in the factor model framework of Fama and French (1993) or the liquidity adjusted CAPM by Acharya and Pedersen (2005) would be interesting for further understanding of its impact on returns. How this differs across time and markets would be an interesting research question. If liquidity behaves similarly across markets the systematic factors of S&P500 liquidity may be possible to use to compose a liquidity index with applicability to a wider array of stocks, as is often done with S&P500 returns. Other interesting research venues include the causes of systematic liquidity shocks, which can help increasing its measurement accuracy; and the relevance of Sadka's (2006) price impact regressions, which has declined in latter years (as discussed in Chapter 2), even though its coefficients show good ability in explaining both liquidity and returns. Furthermore, it would be interesting to implement durations in the liquidity measurement context, following the modelling by Engle and Russell (1998).

Future research in the area of FSO should look further into how the limitations of the technique can be handled. Also, solution stability over time is a topic that has not been

covered in this literature. This could be studied using a moving or expanding window in the estimation — similar to the systematic liquidity study in this thesis. Another interesting avenue of research would be to look at how to determine what utility function is appropriate for an investor. Possibly, experimental economics can contribute in this area. This relates directly to the model on liquidity-adjusted portfolio optimisation presented in the previous section, where utility function determination is crucial. Liquidity influence on portfolio choice is an important area, in particular to handle financial crisis. What properties of stock liquidity are important to monitor when managing a portfolio? How can liquidity factors be used to optimise the trade-off between illiquidity return premiums and the ability to liquidate positions on short notice?

Liquidity crisis behaviour also forms a link to how macroeconomics can be better understood. Measures of money supply are typically aggregates of different asset classes weighted together by their respective degree of liquidity (see e.g. Elger and Binner 2004). If stocks were included in such aggregates, would the stock illiquidity related to financial crisis be a possible explanation for how a financial crisis can cause a recession in the real economy?

Part V

References

Bibliography

Acharya, V., and L. Pedersen, 2005, "Asset Pricing with Liquidity Risk," *Journal of Financial Economics*, 77(2), 375–410.

Adler, T., and M. Kritzman, 2007, "Mean-Variance versus Full-Scale Optimization: In and Out of Sample," *Journal of Asset Management*, 7(5), 302–311.

Aït-Sahalia, Y., and M. Brandt, 2001, "Variable Selection for Portfolio Choice," *The Journal of Finance*, 56(4), 1297–1351.

Akerlof, G., 1970, "The Market for Lemons: Quality Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 84(3), 488–500.

Amihud, Y., 2002, "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects," *Journal of Financial Markets*, 5(1), 31–56.

Amihud, Y., and H. Mendelson, 1980, "Dealership Markets: Market-Making with Inventory," *Journal of Financial Economics*, 8, 31–53.

——— , 1986a, "Asset Pricing and the Bid-Ask Spread," *Journal of Financial Economics*, 17(2), 223–250.

——— , 1986b, "Liquidity and Stock Returns," *Financial Analysts Journal*, 42(3), 43–48.

——— , 1989, "The Effects of Beta, Bid-Ask Spread, Residual Risk, and Size on Stock Returns," *The Journal of Finance*, 44(2), 479–486.

Amihud, Y., H. Mendelson, and R. Wood, 1990, "Liquidity and the 1987 Stock Market Crash," *Journal of Portfolio Management*.

Amihud, Y., L. Pedersen, and H. Mendelson, 2005, "Liquidity and Asset Prices," *Foundations and Trends in Finance*, 1(4).

Aparicio, F., and J. Estrada, 2001, "Empirical Distributions of Stock Returns: European Securities Markets, 1990-95," *The European Journal of Finance*, 7(1), 1–21.

Arditti, F., 1967, "Risk and the Required Return on Equity," *The Journal of Finance*, 22(1), 19–36.

Arrow, K., 1965, "Aspects of a Theory of Risk Bearing.," Yrjo Jahnsson Lectures, Helsinki. Reprinted in Essays in the Theory of Risk Bearing 1971.

Bagehot, W., 1971, "The Only Game in Town," *Financial Analysts Journal*, 27(2), 12–14.

Bauer, W., 2004, "Commonality in Liquidity in Pure Order-Driven Markets," *National Centre of Competence in Research Working Paper*.

Bauwens, L., and P. Giot, 2001, *Econometric Modelling of Stock Market Intraday Activity.* Kluwer, Dordrecht, The Netherlands.

Beedles, W., 1979, "Return, Dispersion and Skewness: Synthesis and Investment Strategy," *Journal of Financial Research*, 2(1), 71–80.

Beltran-Lopez, H., P. Giot, and J. Grammig, 2005, "Commonalities in the Order Book," *Universite Catholique de Louvain Working Paper*, 14.

Benartzi, S., and R. Thaler, 1995, "Myopic Loss Aversion and the Equity Premium Puzzle," *Quarterly Journal of Economics*, 110(1), 73–92.

Benston, G., and R. Hagerman, 1974, "Determinants of Bid-Ask Spreads in the Over-the-Counter Market," *Journal of Financial Economics*, 1(4), 353–364.

Boehmer, E., J. Broussard, and J. Kallunki, 2002, *Using SAS in Financial Research.* SAS Publishing, Cary, NC.

Bollerslev, T., 1986, "Generalized Autoregressive Conditional Heterocedasticity," *Journal of Econometrics*, 31, 307–327.

Brandt, M., 1999, "Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach," *The Journal of Finance*, 54(5), 1609–1645.

Brennan, M., T. Chordia, and A. Subrahmanyam, 1998, "Alternative Factor Specifications, Security Characteristics, and the Cross-Section of Expected Stock Returns," *Journal of Financial Economics*, 49(3), 345–373.

Brennan, M., and A. Subrahmanyam, 1996, "Market Microstructure and Asset Pricing: On the Compensation for Illiquidity in Stock Returns," *Journal of Financial Economics*, 41, 441–464.

Brownlees, C., and G. Gallo, 2006, "Financial Econometric Analysis at Ultra-High Frequency: Data Handling Concerns," *Computational Statistics and Data Analysis*, 51(4), 2232–2245.

Brunnermeier, M., and L. Pedersen, 2009, "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, 22(6), 2201–2238.

Carhart, M., 1997, "On Persistence in Mutual Fund Performance," *The Journal of Finance*, 52(1), 57–82.

Chakrabarty, B., B. Li, V. Nguyen, and R. Van Ness, 2007, "Trade Classification Algorithms for Electronic Communications Network Trades," *Journal of Banking and Finance*, 31(12), 3806–3821.

Chalmers, J., and G. Kadlec, 1998, "An Empirical Examination of the Amortized Spread," *Journal of Financial Economics*, 48(2), 159–188.

Chamberlain, G., 1983, "Characterization of the Distributions that Imply Mean-Variance Utility Functions.," *Journal of Economic Theory*, 29(1), 185–201.

Chamberlain, G., and M. Rothschild, 1983, "Arbitrage, Factor Structure, and Mean-Variance Analysis on Large Asset Markets," *Econometrica*, 51(5), 1281–1304.

Chen, J., 2007, "Understanding Equity Returns," Ph.D. thesis, Columbia University, New York, NY.

Chollete, L., R. Næs, and J. Skjeltorp, 2007, "What Captures Liquidity Risk? A Comparison of Trade and Order Based Liquidity Factors," *Norges Bank Working Paper*, 3.

——— , 2008, "The Risk Components of Liquidity," *Norges Bank Working Paper*, 3.

Chordia, T., A. Goyal, R. Sadka, G. Sadka, and L. Shivakumar, 2009, "Liquidity and the Post-Earnings-Announcement Drift," *Financial Analysts Journal*, 65(4), 31–56.

Chordia, T., S. Huh, and A. Subrahmanyam, 2009, "Theory-Based Illiquidity and Asset Pricing," *Forthcoming in Review of Financial Studies*.

Chordia, T., R. Roll, and A. Subrahmanyam, 2000, "Commonality in Liquidity," *Journal of Financial Economics*, 56(1), 3–28.

——— , 2001, "Market Liquidity and Trading Activity," *The Journal of Finance*, 56(2), 501–530.

——— , 2002, "Order Imbalance, Liquidity, and Market Returns," *Journal of Financial Economics*, 65(1), 111–130.

Chordia, T., A. Sarkar, and A. Subrahmanyam, 2005, "An Empirical Analysis of Stock and Bond Market Liquidity," *Review of Financial Studies*, 18(1), 85–129.

Clark, P., 1973, "A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica*, 41(1), 135–155.

Connor, G., and R. Korajczyk, 1986, "Estimating Pervasive Economic Factors with the Arbitrage Pricing Theory: A New Framework for Analysis," *Journal of Financial Economics*, 15, 323–346.

Constantinides, G., 1986, "Capital Market Equilibrium with Transaction Costs," *Journal of Political Economy*, 94(4), 842–862.

Coughenour, J., and M. Saad, 2004, "Common Market Makers and Commonality in Liquidity," *Journal of Financial Economics*, 73(1), 37–69.

Cremers, J. H., M. Kritzman, and S. Page, 2003, "Portfolio Formation with Higher Moments and Plausible Utility," *Revere Street Working Paper Series*, 272-12.

———, 2005, "Optimal Hedge Fund Allocations," *Journal of Portfolio Management*, 31(3), 70–81.

Cuthbertson, K., and D. Nitzsche, 2004, *Quantitative Financial Economics: Stocks, Bonds and Foreign Exchange*. Wiley, New York, NY.

Damodaran, A., 1985, "Economic Events, Information Structure, and the Return-Generating Process," *Journal of Financial and Quantitative Analysis*, pp. 423–434.

Datar, V., N. Naik, and R. Radcliffe, 1998, "Liquidity and Stock Returns: An Alternative Test," *Journal of Financial Markets*, 1(2), 203–219.

Demsetz, H., 1968, "The Cost of Transacting," *The Quarterly Journal of Economics*, pp. 33–53.

Domowitz, I., O. Hansch, and X. Wang, 2005, "Liquidity Commonality and Return Co-Movement," *Journal of Financial Markets*, 8(4), 351–376.

Dong, J., A. Kempf, and P. Yadav, 2007, "Resiliency, the Neglected Dimension of Market Liquidity: Empirical Evidence from the New York Stock Exchange," *University of Lancaster Working paper*.

Eleswarapu, V., 1997, "Cost of Transacting and Expected Returns in the NASDAQ Market," *The Journal of Finance*, pp. 2113–2127.

Eleswarapu, V., and M. Reinganum, 1993, "The Seasonal Behavior of the Liquidity Premium in Asset Pricing," *Journal of Financial Economics*, 34(3), 373–386.

Elger, T., and J. Binner, 2004, "The UK Household Sector Demand for Risky Money," *Topics in Macroeconomics*, 4(1), 1136–1136.

Ellis, K., R. Michaely, and M. O'Hara, 2000, "The Accuracy of Trade Classification Rules: Evidence from NASDAQ," *Journal of Financial and Quantitative Analysis*, pp. 529–551.

The Economist, February 8, 2007, "Economics Focus: A Fluid Concept," 382(8515), 90.

Engle, R., 1982, "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50(4), 987–1008.

Engle, R., and A. Patton, 2004, "Impacts of Trades in an Error-Correction Model of Quote Prices," *Journal of Financial Markets*, 7(1), 1–25.

Engle, R., and J. Russell, 1998, "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data," *Econometrica*, 66(5), 1127–1162.

Epp, S. S., 2003, *Discrete Mathematics with Applications*. Brooks Cole, Pacific Grove, CA, 3rd edn.

Fama, E., and K. French, 1993, "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33(1), 3–56.

Fama, E., and J. MacBeth, 1973, "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 81(3), 607–636.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin, 2000, "The Generalized Dynamic Factor Model: Identification and Estimation," *Review of Economics and Statistics*, 82(4), 540–554.

———, 2005, "The Generalized Dynamic Factor Model," *Journal of the American Statistical Association*, 100(471), 830–840.

Foster, F., and S. Viswanathan, 1993, "Variations in Trading Volume, Return Volatility, and Trading Costs: Evidence on Recent Price Formation Models," *The Journal of Finance*, 48, 187–211.

Galariotis, E., E. Giouvris, and M. Lane, 2007, "Liquidity Commonality in the London Stock Exchange," *Journal of Business, Finance & Accounting*, 34(1-2), 374–388.

Gallant, A., P. Rossi, and G. Tauchen, 1992, "Stock Prices and Volume," *The Review of Financial Studies*, 5(2), 199–242.

Ghysels, E., and J. Pereira, 2008, "Liquidity and Conditional Portfolio Choice: A Nonparametric Investigation," *Journal of Empirical Finance*, 15(4), 679–699.

Gibson, R., and N. Mougeot, 2004, "The Pricing of Systematic Liquidity Risk: Empirical Evidence from the US Stock Market," *Journal of Banking and Finance*, 28(1), 157–178.

Gilmore, C., B. Lucey, and G. McManus, 2008, "The Dynamics of Central European Equity Market Comovements," *Quarterly Review of Economics and Finance*, 48(3), 605–622.

Giouvris, E., and E. Galariotis, 2008, "Systematic Liquidity and Excess Returns: Evidence From The Athens Stock Exchange," *Journal Of Money, Investment and Banking*, 1(2), 81–96.

Glosten, L., 1987, "Components of the Bid-Ask Spread and the Statistical Properties of Transaction Prices," *The Journal of Finance*, pp. 1293–1307.

Glosten, L., and L. Harris, 1988, "Estimating the Components of the Bid/Ask Spread," *Journal of Financial Economics*, 21(1), 123–142.

Glosten, L., and P. Milgrom, 1985, "Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Agents," *Journal of Financial Economics*, 14(71), 71–100.

Goffe, W., G. Ferrier, and J. Rogers, 1994, "Global Optimization of Statistical Functions with Simulated Annealing," *Journal of Econometrics*, 60(1/2), 65–100.

Goldfeld, S., and R. Quandt, 1973, "A Markov Model for Switching Regressions," *Journal of Econometrics*, 1(1), 3–15.

Gourieroux, C., and A. Monfort, 1998, "The Econometrics of Efficient Frontiers," *CREST Working Paper*, 4.

———, 2005, "The Econometrics of Efficient Portfolios," *Journal of Empirical Finance*, 12, 1–41.

Goyenko, R., C. Holden, and C. Trzcinka, 2009, "Do Liquidity Measures Measure Liquidity?," *Journal of Financial Economics*, 92(2), 153–181.

Grinold, R., 1999, "Mean-Variance and Scenario-Based Approaches to Portfolio Selection," *Journal of Portfolio Management*, 25(2), 10–22.

Guidolin, M., and A. Timmermann, 2005, "International Asset Allocation under Regime Switching, Skew and Kurtosis Preferences," *Federal Reserve Bank of St. Louis Working Paper*, 2005–034B(2005–034B).

Hagströmer, B., R. G. Anderson, J. M. Binner, T. Elger, and B. Nilsson, 2008, "Mean-Variance vs. Full-Scale Optimization: Broad Evidence for the UK," *The Manchester School*, 76, 134–156.

Hagströmer, B., R. G. Anderson, J. M. Binner, and B. Nilsson, 2009, "Dynamics in Systematic Liquidity," *Federal Reserve Bank of St Louis Working Paper*, 2009-025A.

Hagströmer, B., and J. M. Binner, 2009, "Stock Portfolio Selection woth Full-Scale Optimization and Differential Evolution," *Forthcoming in Applied Financial Economics*.

Hallin, M., C. Mathias, H. Pirotte, and D. Veredas, 2009, "Market Liquidity as Dynamic Factors," *ECORE Discussion Paper*, 004.

Hamilton, J., 1989, "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57(2), 357–384.

Hansch, O., 2003, "Island Tides: Exploring ECN Liquidity," *Pennsylvania State University Working Paper*.

Harvey, C., J. Liechty, M. Liechty, and P. Muller, 2004, "Portfolio Selection with Higher Moments," *Duke University Working Paper*.

Hasbrouck, J., 1988, "Trades, Quotes, Inventories, and Information," *Journal of Financial Economics*, 22(2), 229–52.

———, 1991, "Measuring the Information Content of Stock Trades," *The Journal of Finance*, 46, 179–207.

———, 2009, "Trading Costs and Returns for US Equities: The Evidence from Daily Data," *Forthcoming in The Journal of Finance*.

Hasbrouck, J., and D. Seppi, 2001, "Common Factors in Prices, Order Flows, and Liquidity," *Journal of Financial Economics*, 59(3), 383–411.

Heij, C., D. van Dijk, and P. Groenen, 2008, "Macroeconomic Forecasting with Matched Principal Components," *International Journal of Forecasting*, 24(1), 87–100.

Henker, T., and J. Wang, 2006, "On the Importance of Timing Specifications in Market Microstructure Research," *Journal of Financial Markets*, 9(2), 162–179.

Hicks, J., 1962, "Liquidity," *The Economic Journal*, 72(288), 787–802.

Ho, T., and H. Stoll, 1981, "Optimal Dealer Pricing under Transactions and Return Uncertainty," *Journal of Financial Economics*, 9, 47–73.

Huang, R., and H. Stoll, 1996, "Dealer Versus Auction Markets: A Paired Comparison of Execution Costs on NASDAQ and the NYSE," *Journal of Financial Economics*, 41(3), 313–357.

Huberman, G., and D. Halka, 2001, "Systematic Liquidity," *Journal of Financial Research*, 24(2), 161–178.

Jacoby, G., D. Fowler, and A. Gottesman, 2000, "The Capital Asset Pricing Model and the Liquidity Effect: A Theoretical Approach," *Journal of Financial Markets*, 3(1), 69–81.

Jarque, C., and A. Bera, 1980, "Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals," *Economics Letters*, 6(3), 255–259.

Jegadeesh, N., and S. Titman, 1993, "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency," *The Journal of Finance*, 48, 65–91.

Jolliffe, I., 2002, *Principal Component Analysis*. Springer, New York, NY.

Kahnemann, D., and A. Tversky, 1979, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47(2), 263–291.

Kamara, A., X. Lou, and R. Sadka, 2008, "The Divergence of Liquidity Commonality in the Cross-Section," *Journal of Financial Economics*, 89, 444–466.

Kapetanios, G., M. Marcellino, V. Schifanoia, and V. Salasco, 2009, "A Parametric Estimation Method for Dynamic Factor Models of Large Dimensions," *Journal of Time Series Analysis*, 30(2), 208–238.

Karolyi, G. A., K.-H. Lee, and M. A. van Dijk, 2007, "Commonality in Returns, Liquidity, and Turnover Around the World," *Fisher College of Business Working Paper Series*, 2007-16.

Kempf, A., and D. Mayston, 2008, "Liquidity Commonality Beyond Best Prices," *Journal of Financial Research*, 31(1), 25–40.

Kopp, E., M. Hütl, O. Loistl, and J. Prix, 2008, "Systematic Liquidity in the Xetra Order Book: A Multi-Stage Approach," *Vienna Univeristy of Economics and Business Administration, Working paper*.

Korajczyk, R., and R. Sadka, 2008, "Pricing the Commonality Across Alternative Measures of Liquidity," *Journal of Financial Economics*, 87, 45–72.

Kyle, A., 1985, "Continuous Auctions and Insider Trading," *Econometrica*, 53(6), 1315–1336.

Lai, K., L. Yu, and S. Wang, 2006, "Mean-Variance-Skewness-Kurtosis-Based Portfolio Optimization," in *Proceedings of the First International Multi-Symposiums on Computer and Computational Sciences (IMSCCS'06)*.

Lee, C., and M. Ready, 1991, "Inferring Trade Direction from Intraday Data," *The Journal of Finance*, 46(2), 733–746.

Levy, H., and H. Markowitz, 1979, "Approximating Expected Utility by a Function of Mean and Variance," *The American Economic Review*, 69(3), 308–317.

Levy, H., and M. Sarnat, 1984, *Portfolio and Investment Selection: Theory and Practice*. Prentice/Hall International, Englewood Cliffs, NJ.

Li, W., H. Yue, S. Valle-Cervantes, and S. Qin, 2000, "Recursive PCA for Adaptive Process Monitoring," *Journal of Process Control*, 10(5), 471–486.

Lintner, J., 1965, "Security Prices, Risk, and Maximal Gains From Diversification," *The Journal of Finance*, 20(4), 587–615.

Longstaff, F., 2001, "Optimal Portfolio Choice and the Valuation of Illiquid Securities," *Review of Financial Studies*, 14(2), 407–431.

Lynch, A., and S. Tan, 2004, "Explaining the Magnitude of Liquidity Premia: The Roles of Return Predictability, Wealth Shocks and State-Dependent Transaction Costs," *NBER Working Paper*, 10994.

Mandelbrot, B., 1963, "The Variation of Certain Speculative Prices," *The Journal of Business*, 36(4), 394–419.

Maringer, D., 2008a, *Constrained Index Tracking under Loss Aversion Using Differential Evolution*, vol. 100. Springer, Berlin / Heidelberg, Germany.

——— , 2008b, "Risk Preferences and Loss Aversion in Portfolio Optimization," in *Computational Methods in Financial Engineering*, ed. by E. Kontoghiorghes, B. Rustem, and P. Winker, Berlin / Heidelberg, Germany. Springer.

Maringer, D., and M. Meyer, 2008, "Smooth Transition Autoregressive Models — New Approaches to the Model Selection Problem," *Studies in Nonlinear Dynamics & Econometrics*, 12(1), 5.

Maringer, D., and O. Oyewumi, 2007, "Index Tracking with Constrained Portfolios," *International Journal of Intelligent Systems in Accounting and Finance Management*, 15(1), 57–71.

Maringer, D., and P. Parpas, 2009, "Global Optimization of Higher Order Moments in Portfolio Selection," *Journal of Global Optimization*, 43(2), 219–230.

Markowitz, H., 1952, "Portfolio Selection," *The Journal of Finance*, 7(1), 77–91.

——— , 1959, *Portfolio Selection: Efficient Diversification of Investments*. Yale University Press, New Haven, MA.

Markowitz, H., 1987, *Mean-Variance Analysis in Portfolio Choice and Capital Markets*. Basil Blackwell, Oxford, UK.

Mehrotra, D., 1995, "Robust Elementwise Estimation of a Dispersion Matrix," *Biometrics*, 51, 1344–1351.

Meucci, A., 2005, *Risk and Asset Allocation*. Springer, Berlin / Heidelberg, Germany.

Mossin, J., 1966, "Equilibrium in a Capital Asset Market," *Econometrica*, 34(4), 768–783.

Pastor, L., and R. Stambaugh, 2003, "Liquidity Risk and Stock Returns," *Journal of Political Economy*, 11, 642–685.

Perold, A., 1984, "Large-Scale Portfolio Optimization," *Management Science*, 30(10), 1143–1160.

Petersen, M., and D. Fialkowski, 1994, "Posted Versus Effective Spreads: Good Prices or Bad Quotes?," *Journal of Financial Economics*, 35(3), 269–292.

Pratt, J., 1964, "Risk Aversion in the Small and in the Large," *Econometrica*, 32(1/2), 122–136.

Price, K., R. Storn, and J. Lampinen, 2005, *Differential Evolution: A Practical Approach to Global Optimization*. Springer, Berlin / Heidelberg, Germany.

Sadka, R., 2006, "Momentum and Post-Earnings Announcement-Drift Anomalies: The Role of Liquidity Risk," *Journal of Financial Economics*, 80(2), 309–349.

Scott, R., and P. Horvath, 1980, "On the Direction of Preference for Moments of Higher Order than the Variance," *The Journal of Finance*, 35(4), 915–919.

Sharpe, W., 1964, "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *The Journal of Finance*, 19(3), 425–442.

Sharpe, W. F., 2007, "Expected Utility Asset Allocation," *Financial Analysts Journal*, 63(5), 18–30.

Sortino, F., and H. Forsey, 1996, "On the Use and Misuse of Downside Risk," *Journal of Portfolio Management*, 22(2), 35–42.

Sortino, F., and L. Price, 1994, "Performance Measurement in a Downside Risk Framework," *Journal of Investing*, 3(3), 50–58.

Stock, J., and M. Watson, 2002, "Forecasting Using Principal Components from a Large Number of Predictors," *Journal of the American Statistical Association*, pp. 1167–1179.

Stoll, H., 1978, "The Pricing of Security Dealer Services: An Empirical Study of NASDAQ Stocks," *The Journal of Finance*, 33(4), 1153–1172.

Stoll, H., and C. Schenzler, 2006, "Trades Outside the Quotes: Reporting Delay, Trading Option, or Trade Size?," *Journal of Financial Economics*, 79(3), 615–653.

Storn, R., and K. Price, 1997, "Differential Evolution — A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces," *Journal of Global Optimization*, 11(4), 341–359.

Swensen, D., 2000, *Pioneering Portfolio Management*. The Free Press, New York, NY.

Taylor, S., 1982, "Financial Returns Modeled by the Product of Two Stochastic Processes — A Study of the Daily Sugar Prices 1961–75," *Time Series Analysis: Theory and Practice*, 1, 203–226.

Taylor, S., 1986, *Modelling Financial Time Series*. Wiley, New York, NY.

Tobin, J., 1958, "Liquidity Preference as Behavior Towards Risk," *The Review of Economic Studies*, 25(2), 65–86.

Vergote, O., 2005, "How to Match Trades and Quotes for NYSE Stocks?," *Center for Economic Studies Discussion Paper Series*, 10.

Volosovych, V., 2005, "Measuring Financial Market Integration Over the Long Run: Is There a U-Shape?," *Florida Atlantic University Department of Economics, Working Paper*.

Von Neumann, J., and O. Morgenstern, 1947, *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ.

Wang, S., and Y. Xia, 2002, *Portfolio Selection and Asset Pricing*. Springer, New York, NY.

Watanabe, A., 2004, "Macroeconomic Sources of Systematic Liquidity," *Yale University Working Paper*.

Part VI

Appendix