On the Non-linear Dynamics of Financial Market Risk and Liquidity

A thesis submitted to the Department of Finance of the
London School of Economics and Political Science
for the degree of PhD in Finance

London School of Economics and Political Science,
September 2008

Christian Reusch

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

London, September 2008 Christian Reusch

To my family, the people I love and admire most -

my mother, my father, my sisters, my grandparents,
for always believing in me and helping me be who I am.

And to Jana.

Abstract

This thesis provides a novel empirical treatment of the dynamics of financial market risk and liquidity, two very important areas both for financial research as well as to practitioners in the financial markets: We devise empirical non-linear time series models of the two concepts that specifically take into account 'explosive', self-reinforcing dynamic patterns. While 'conventional' empirical models are often 'linear' and tend to neglect these effects, real-life evidence such as e.g. the 1987 crash, the large stock market drops on February 27^{th}, 2007 or the huge losses posted by investment banks and hedge funds during July and August 2007, suggest that such an approach is warranted:

In the first part of the thesis we extend a time series model of Value-at-Risk (VaR) with non-linear multiplicative features and endogenous regime thresholds. When estimated with a Markov Chain Monte Carlo (MCMC) method against real data, the resulting 'Self-Exciting Threshold CAViaR' (Conditional Autoregressive Value-at-Risk) model is able to detect trigger thresholds for explosive market risk as well as the scale of such a possible expansion in risk.

The second part of the thesis is dedicated to the 'Conditional Autoregressive Liquidity' (CARL) model, a multiplicative time series approach to the empirical modelling of market liquidity. The newly conceptualised model is capable of picking up self-reinforcing dynamics, i.e. autoregressive patterns in liquidity, which is in accordance with theoretical research on the topic. Moreover, by incorporating a multi-dimensional liquidity proxy, the model CARL is explicitly designed to take into account the fact that liquidity is a concept with many facets, unlike other empirical treatments that often view liquidity only along a single dimension (e.g. the bid-ask spread, volume, trade duration). In this thesis, we demonstrate the empirical versatility of the model using both fixed interval data (daily and weekly) as well as tick-by-tick intraday data, for which we propose a filtering technique in order to be able to use the model in such a data environment. We note that the model is able to pick up autocorrelation structures in liquidity rather well and find the forecast performance very encouraging for practical use.

Acknowledgements

This doctoral thesis would not be without the help and support of many. You know who you are! It shames me to admit that I will certainly forget to mention quite a few of you explicitly below. However, do rest assured that I am very much indebted to all of you and will always be grateful for your help and advice. In this limited space, I will merely try to highlight those that I owe most.

First and foremost, I would like to express my deepest gratitude to Dr. Jon Danielsson, my academic supervisor and mentor for many years. These few words cannot possibly capture how much I am indebted to him for his support. I met Jon already while being an undergraduate at the LSE and he has been a invaluable source of help and advice ever since. Many skills as a researcher, especially in scientific programming, I have learned from him. They will stay with me for the years to come.

I also wish to thank my second supervisor Dr. Andrew Patton, Prof. Greg Connor, Prof. Oliver Linton as well as Dr. Antonio Mele for their advice and help in showing me new avenues of progress and keeping me going when things got difficult. In this respect, I would also like to express my gratitude to the participants in the PhD Finance seminar and the Wednesday Econometrics workshop, especially Dr. Michaela

Verardo, Dr. Christian Julliard and Prof. Antoine Faure-Grimaud, for their helpful comments, advice and encouragement.

I also owe a special mention to Prof. David Webb to whom I am particularly grateful. I have known David since being an undergraduate at the LSE. His help, advice and understanding - also in areas other than academia - have kept me going, helped me become a successful researcher and make the most of my time at the Financial Markets Group.

My gratitude also extends to Dr. Sushil Wadhwani for his interest in my work and providing with an excellent professional perspective for the time after my research studies.

I would like to thank the London School of Economics and the Financial Markets Group, especially Mary Comben, Maria Komninou, Sooraya Mohabeer and Osmana Raie for providing me with excellent resources for my research and helping me find my way through LSE's administrative 'jungle'. The generous financial support by Concordia Advisors during the third year of my PhD studies is also gratefully acknowledged.

I have always been able to rely on my family, my mother, my father, my sisters and my grandparents for support. They have never stopped to encourage me and lifted me up when I was at loose ends. They are the people who have allowed me to pursue my dreams and become who I am today. To them I am indebted beyond words. It pains me especially that my grandparents are not with me anymore to witness what they have greatly supported me in doing. It is to them that I dedicate this work in particular. I will be forever grateful for their unconditional love and support.

I would also like to express my gratitude to Jana Klimecki who has been a tremendous source of inspiration and motivation during some of the difficult phases of my work. She kept me going in my efforts to produce research of the highest quality.

Erwin Bechtold used to be my English teacher at the Tilemannschule in Limburg. It is largely also because of him that I was able to start my studies at the LSE in the first place. I owe him many thanks for that.

Finally, I would like to thank Mohammed Fawaz, Anisha Gosh, Dr. Georg Kaltenbrunner, Dr. Manuel Klein, Ossi Lindström, Dr. Ryan Love, Ebrahim Rahbari, Dr. Miguel Segoviano and Bernhard Silli for being good friends and helpful colleagues. Anisha has often helped me grasp the odd difficult notion in (asymptotic) econometrics. My conversations with Georg, Manuel, Ossi, Ebrahim, Miguel and Bernhard have greatly shaped my understanding of financial economics, the academic world and the financial industry. And, without Mohammed and Ryan I would have probably never gotten around to learning C/C++.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

This thesis provides a novel empirical treatment of the dynamics of financial market risk and liquidity, two very important areas both for financial research as well as to practitioners in the financial markets: We devise empirical non-linear time series models of the two concepts that specifically take into account 'explosive', self-reinforcing dynamic patterns. While 'conventional' empirical models are often 'linear' and tend to neglect these effects, adverse events in the history of the financial markets suggest that such an approach is warranted:

On February 27^{th}, 2007 for example, the Dow Jones Stoxx 600 Index dropped by 3% in a manner that, according to the *The Wall Street Journal Europe, edn. February 28^{th}, 2007,* was "almost like (...) a cascade". Regarding these events, *The Economist, issue March 3^{rd}, 2007,* pointed out that investment banks often sell assets when volatility is rising in a move to cut capital allocated to trading, thereby creating self-reinforcing patterns in market risk: "...a sudden jump in volatility tends to generate further volatility", resulting in a large build-up of risk and potentially large losses. Another example for these mechanisms at work are the events in July and August 2007, when many investment banks and hedge funds posted major losses on their trading books: Goldman Sachs' main equity fund for example lost more than 30% of its value in a week caused by problems in its computer-driven trading strategies. The *Financial Times* on August 14^{th} 2007, analyses:

> "The issue of computer 'herding' appears to be a key factor behind this month's [August 2007] problems at the Goldman Sachs funds and others (...) computer models do not always take account of the way that their own behaviour is affecting markets (...) In practical terms, this means that when models evaluate markets, they often fail to recognise how their own behaviour is distorting prices (...) Computers have a nasty habit of all using similar strategies – partly because they are created by humans who have studied at the same institutions. Thus they can all dash for the exits at the same time."

Regarding the losses that materialised as a result of these issues, Goldman Sachs' chief financial officer, Mr. Davir Vaniar comments in the same paper:

> "We are seeing things that were 25-standard deviation events, several days in a row."

With a more explicit focus on liquidity, *The Economist, edn. April 28th, 2007* states:

> "Liquidity is a self-reinforcing process (...) if liquidity suddenly dries up, some investors might end up owning assets they neither want nor can get rid of. This might make a virtuous circle turn vicious."

Carlson (2007) further suggests that such liquidity problems were also hat the heart of the 1987 market crash. He argues that so-called margin calls on trading positions that suffered from losses, led to reduced market liquidity in the (futures) market through further selling, which in turn led to an even steeper decline.

The above evidence from the financial markets thus suggests that market risk and (il)liquidity are subject to feedback 'loops' as well as 'cross-over' effects, whereby market risk and (il)liquidity may feed on themselves and each other to create explosive patterns that can adverse affect trading conditions. This mechanism has been documented by Daníelsson and Shin (2003) and themed 'endogenous risk'. Morris and Shin (1998), Morris and Shin (2004) as well as Brunnermeier and Pedersen (2008) have provided theoretical models of these feedback mechanisms that can lead to liquidity 'black holes', excessive market risk and potentially large price drops in asset markets.

In this thesis we seek to develop new empirical time series models of market risk and liquidity that capture these 'unorthodox' dynamics: In the first part of the thesis we extend a time series model of Value-at-Risk (VaR) with non-linear multiplicative features and regime thresholds, to yield a non-linear time series model that is capable of estimating explosive patterns in market risk.

The second part of the thesis is dedicated to the 'Conditional Autoregressive Liquidity' (CARL) model, a multiplicative time series approach to the empirical modelling of market liquidity. The newly conceptualised model is capable of picking up self-reinforcing dynamics, i.e. autoregressive patterns in liquidity, in daily, weekly as well as intraday data. Moreover, in the setup of our model we explicitly take into account that liquidity is a difficult concept with multiple facets that are of varying importance both to different market participants and over time: Thus, as a liquidity proxy in the CARL model, we use the Hiu-Heubel liquidity ratio, a multi-dimensional measure of liquidity that proxies for the tightness as well as depth and breadth of a market.

Chapter 2 – Self-Exciting CAViaR Models with Endogenous Thresholds

This chapter introduces the empirical model class of 'Self-exciting Threshold' CAViaR (SET-CAViaR), in which threshold variables are determined endogenously. The proposed model is designed to incorporate reinforcing, explosive dynamics of market risk that are referred to as 'endogenous risk' and have recently received increased attention in the theoretical literature as well as among practitioners. As such, the SET-CAViaR constitutes a new empirical approach to risk modelling, which due to its semi-parametric and non-linear features calls for a Markov Chain Monte Carlo (MCMC) estimation routine. We illustrate the model in a Monte Carlo study and estimate it with CRSP IBM holding returns from 1982-2005. The estimation results generally indicate a good fit and are well in accordance with theoretical predictions: There is evidence for the SET-CAViaR model detecting sudden, explosive patterns in market risk, which 'conventional' models are not conceived to pick up.

Chapter 3 – CARL: An Empirical Conditional Autoregressive Model of Market Liquidity

In this chapter we present the Conditional Autoregressive Liquidity (CARL) model, an empirical time series model of market liquidity. We develop the model from two main building blocks: As an econometric framework, we use the multiplicative setup of Engle and Russell's (1998) ACD model; as a proxy for liquidity, we rely on the Hui-Heubel ratio, a multi-dimensional measure of (il)liquidity that captures many facets of market liquidity. For the resulting CARL model, we establish the econometric properties and demonstrate its empirical validity in the light of recent theoretical research in financial liquidity: Such theory predicts (il)liquidity to be self-reinforcing, thus strongly autocorrelated through time. In an empirical application using both daily as well as weekly CRSP Amazon data from 1997-2006 we find that the CARL model picks up such autocorrelation rather well and produces accurate forecasts.

Chapter 4 – Intraday Liquidity: A High-Frequency Application of the CARL Model

This chapter constitutes an application of the CARL model to an intraday context. In order to use the model in a high-frequency data environment characterised by irregular time intervals between subsequent tick-by-tick observations, we propose a volume technique: We sort data into volume durations during which a particular volume is both bought and sold and derive the maximum percentage range liquidity measure, a metric similar to the Hui-Heubel ratio, over the resulting intervals. The filtered series is then taken as an input into the CARL framework. We demonstrate the filtering technique as well as the versatility of the resulting intraday CARL model in an empirical application using TAQ intraday data on Amazon: Similar to previous empirical work in a daily and weekly data context and in line with theoretical research, we find that the model is able to pick up significant autocorrelation in the (il)liquidity proxy. Moreover, the model shows good forecasting performance.

Bibliography

Brunnermeier, M. and Pedersen, L. H. (2008), Market liquidity and funding liquidity. Working paper, Princeton University; forthcoming in the *Review of Financial Studies*; available from `http://www.princeton.edu/~markus/research/papers/liquidity`.

Carlson, M. (2007), A brief history of the 1987 stock market crash: with a dicussion of the Federal Reserve response, The finance and economics discussion series (feds) papers, The Federal Reserve Board; available from `http://www.federalreserve.gov/pubs/feds/2007/200713/200713pap.pdf`.

Daníelsson, J. and Shin, H. S. (2003), Endogenous risk, *in* P. Field, ed., 'Modern Risk Management', Risk Books, London, UK, pp. 297–316.

Engle, R. F. and Russell, J. R. (1998), 'Autoregressive conditional duration; a new model for irregularly spaced transaction data', *Econometrica* **66**, 1127–1162.

Morris, S. and Shin, H. S. (1998), 'Unique equilibrium in a model of self-fulfilling currency attacks', *American Economic Review* **88**, 587–597.

Morris, S. and Shin, H. S. (2004), 'Liquidity black holes', *Review of Finance* **8**, 1–18.

Chapter 2

Self-Exciting CAViaR Models with Endogenous Thresholds

2.1 Introduction

Most commonly used empirical models of market risk focus on the quantification of such risk in terms of 'Value-at-Risk' (VaR), i.e. the quantile of the returns distribution, through non-parametric empirical estimation or using time series models in linear single-regime setups. An example of an approach of the former kind is the popular method of 'Historical Simulation' (HS), whereas the latter includes quantile models based on conditional volatility or the 'Conditional Autoregressive Value-at-Risk' (CAViaR) model class. Recently developed by Engle and Manganelli (2004), CAViaR constitutes a semi-parametric autoregressive quantile regression approach, whereby the conditional quantile of the returns distribution at time t is typically directly modelled as a function of lagged values of the return and its lagged own conditional quantiles.

However, despite the wide-spread use, appeal and individual merits (in the case of CAViaR, cf. Engle and Manganelli, 2004), the above models are ill-equipped to incorporate self-exciting, reinforcing patterns in market risk, which recent theoretical research has dubbed 'endogenous risk': Daníelsson and Shin (2003) for example suggest that traders' selling decisions can lead to self-exciting downward spirals in asset prices, thus creating endogenous risk from within the financial system. More formally, Morris and Shin (1999) and Morris and Shin (2004) show how coordination effects and higher order beliefs along the lines of Morris and Shin (1998) can create endogenous risk and 'liquidity blackholes': Once certain thresholds are crossed the reinforcing actions of market participants can trigger to self-exciting patterns and endogenous downward spirals in asset returns.

Building on the CAViaR model class, in this chapter we provide a new framework in which to analyse endogenous risk empirically: We propose the 'Self-exciting Threshold' CAViaR (henceforth SET-CAViaR) model, an empirical time series model that explicitly allows for explosive, self-exciting risk dynamics over a part of the returns domain. When estimated with real data, the new model is able to detect 'trigger' thresholds for self-exciting endogenous risk and is therefore capable of identifying and predicting potential endogenously created crises. Further, the model also quantifies the scale of endogenous risk during such states of the world.

While the theoretical findings underlying the proposed SET-CAViaR model have been known for some time, an empirical attempt at providing a framework for the analysis of endogenous risk as in this chapter has to our knowledge so far not been attempted. Yet, the practical experience in financial markets also suggests that such an approach is warranted: Regarding the sizeable stock market fall on February 27^{th}, 2007, *The Economist, issue March 3^{rd}, 2007*, for example states on p.

82 that on this day "[the Dow Jones Industrial Average] plummeted [at] a rate of decline that traders said was unprecedented", suggesting mutually reinforcing sell decisions by traders leading to a cascading sell-off. The paper further concludes that "if conventional [VaR] models are correct, such an event should not have happened in the history of the known universe", indicating the need for better VaR modelling methodologies in the face of such events.

We would argue that the introduction of the SET-CAViaR model represents a first step into this direction: The proposed methodology builds on the CAViaR model as a univariate quantile 'baseline' process, which we combine with a multiplicative scaling factor that is allowed to differ across regimes. Regime shifts are triggered by lagged returns crossing endogenously estimated thresholds. In this setup, risk - as measured by the return quantile - can therefore increase exponentially once an explosive regime is entered.

By relying on the recently developed CAViaR model, we benefit from its semi–parametric properties: Specifying the law of motion of the return distribution quantiles directly, CAViaR models the evolvement of *all* moments of return distribution at different quantile levels – without the (ad hoc) assumption of a particular error term distribution. In this respect, CAViaR contrasts with more 'traditional' empirical risk modelling methods based on models for conditional volatility: Particular examples of the latter include models in the ARMA-GARCH class[1], which explicitly lay out the law of motion of the first and second moments of the returns distribution, but do *not* address higher moment dynamics or quantile-specific behaviour. As a consequence, models based on conditional volatility have been shown to

[1] cf. e.g. Engle (1982) and Bollerslev (1986).

produce poor VaR forecasts in situations in which returns exhibit excess kurtosis and strong skewness (cf. e.g. McNeil et al., 2005, p. 49), thus creating a case for the more general CAViaR approach:

Indeed, as also shown in this chapter, the CAViaR model class constitutes a superset that includes a variety of ARMA-GARCH models. The qualitative logic of our self-exciting threshold methodology therefore also carries over to an ARMA-GARCH model setting.

Yet, while the semi-parametric nature of our SET-CAViaR model is advantageous in many ways, it comes at a specific cost that is particularly relevant in practice: Contrary to e.g. ARMA-GARCH-type models, CAViaR makes *no* explicit parametric assumptions about an error term distribution[2], rendering conventional estimation according to '(Quasi) Maximum Likelihood' ((Q)ML) infeasible. Rather, CAViaR models are estimated by means of minimising Bassett and Koenker's (1978) quantile regression objective function, which is not everywhere differentiable and, for common (non-linear) CAViaR specifications, often exhibits multiple local extrema. Commonly used gradient-based optimisation routines such as for example the 'Quasi-Newton' method are thus not applicable in this environment. While methods such as linear programming or interior point algorithms have been proposed for specific (mostly linear) CAViaR processes, they have been documented to be problem-laiden (cf. Koenker and Park, 1996, p. 277) and are not feasible in the case of more complex, non-linear CAViaR representations including this chapter's endogenous structural breaks setting (cf. Komunjer, 2005, p. 151). Estimation in this paper is therefore carried out by means of an algorithm based on Chernozhukov and Hong's (2003) 'Markov Chain Monte Carlo' (MCMC) 'LaPlace-type

[2]In the case of ARMA-GARCH models usually taken to be a Gaussian or Student t distribution.

Estimation' (LTE) routine, a simulation-based 'global' optimisation method that does not rely on gradients and sufficiently explores the parameter space, avoiding the problem of getting trapped in local extrema.

We also benefit from the circumstance that while the need for more elaborate estimation routines already in the case of 'simple' CAViaR models is evidently costly in terms of computing effort and estimation time, once implemented it also affords the opportunity to estimate the richer, more complex quantile regression specifications such as the SET-CAViaR version in this chapter. Moreover, the proposed MCMC LTE routine also facilitates the novel approach of estimating the regime thresholds endogenously as variables of the model. Drawing on the above, the identification of such thresholds is further aided by the semi-parametric properties of the CAViaR model in the sense that shifts in the law of motion of the CAViaR process affecting the moments of the returns distribution subject to regimes should be easier to detect by taking into account all moments (as in the case of CAViaR) as opposed to e.g. just the mean and variance in the case of ARMA-GARCH.

When put to use in the computationally challenging setting of the proposed model, the estimation routine is capable of producing meaningful results:

An estimation of two concrete examples of SET-CAViaR models against a time series of 1982-2005 CRSP IBM holding returns indicates indeed that beyond loss thresholds of c. 3.5% and 10.5% for lagged returns, market risk as measured by VaR can suddenly increase by a factor of c. 2.0 and 3.5 on VaR levels of 95% and 99% respectively. Further, the results suggest that beyond yet an even larger loss threshold for lagged returns, these explosive VaR dynamics are dampened by a factor below unity allowing for a eventual return to more normal market risk levels.

While feasible from an econometric point of view, these results are in line with the before-mentioned theories on endogenous risk and co-ordination effects as well as empirically observed patterns in financial markets: As documented by Morris and Shin (2004) and Daníelsson and Peñaranda (2007) there cannot only be sharp increases in risk and associated large negative returns, but prices also are expected to eventually 'bounce back' following sudden, large negative returns, producing 'v-shaped' price patterns, which can be seen as a reaction to market risk returning to more normal levels.

The rest of the chapter proceeds as follows: Section 2.2 provides an overview of quantile regression as well as CAViaR models. In Section 2.3, the links between the popular ARMA-GARCH class and CAViaR models are established. Section 2.4 argues the case for self-exciting non-linear dynamics in risk modelling and introduces the SET-CAViaR model class. Section 2.5 is dedicated to the variant of MCMC LTE methodology used in this chapter. In Section 2.6 the SET-CAViaR model is subjected to a Monte Carlo sensitivity study. Section 2.7 presents an application of the methodology to the financial markets: SET-CAViaR models are estimated with CRSP IBM holding returns from 1980-2005. The results, computational issues surrounding the MCMC LTE routine as well as possible areas of application are discussed here before section 2.8 concludes.

2.2 CAViaR and Quantile (Auto-)Regression

CAViaR models describe autoregressive conditional quantiles of financial returns over time: $Q_\tau(y_t|\theta)$, the conditional τ quantile of the return y at time t is modelled as a function of a parameter vector θ, of variables included in the information set $\mathcal{F}_{t-1}$ available at time t, typically lagged values of the return, and its lagged conditional τ quantiles, such that $\mathbb{P}(y_t \leq Q_\tau(y_t|\theta)|\mathcal{F}_{t-1}) = \tau$. As such, CAViaR is rooted in the literature on the estimation of quantiles, of which a short account is given in the following section.

2.2.1 Quantile Estimation

Following initial work on the analysis of quantiles by Fox and Rubin (1964), Koenker and Bassett (1978) provided the first rigorous and comprehensive analysis of quantile regression: Most notably, they established the quantile regression objective function, which we also employ in the estimation of the SET-CAViaR model:

Quantiles are estimated by means of a standard decision theoretic approach (cf. e.g. Ferguson, 1967), whereby one is to minimise the expectation $\mathbb{E}\left[\rho_\tau\left(u_i(\phi)\right)\right]$ of the following piecewise linear loss function

$$\rho_\tau\left(u_i(\phi)\right) = u_i(\phi) \cdot [\tau - \mathbb{I}(u_i(\phi) < 0)], \tag{2.1}$$

i.e., in a sample setting,

$$\min_{\phi\in\Theta} \mathbb{E}[\rho_\tau(u_i(\phi))] \Rightarrow \min_{\phi\in\Theta} \mathcal{L}[u_i(\phi);\tau;N] = \min_{\phi\in\Theta} N^{-1}\sum_{i=1}^{N}\rho_\tau[u_i(\phi)] \tag{2.2}$$

in which $\mathbb{I}(\cdot)$ denotes the indicator function, $\tau \in (0,1)$ the cumulative probability of quantile to be analysed, N the sample size and $u_i(\phi)$ a criterion function with the p-dimensional parameter vector $\phi \in \Theta \subset \mathbb{R}^p$. The compact set Θ denotes the parameter space.

2.2.2 The General CAViaR Model

In the time series context of CAViaR models one usually specifies the time t quantile of y to depend on past information, i.e. lagged values of y and its quantiles, e.g.

$$\min_{\theta\in\Theta} T^{-1}\sum_{t=1}^{T}\rho_\tau[y_t - Q_\tau(y_t|\mathcal{F}_{t-1};\theta)]. \tag{2.3}$$

In general, the autoregressive CAViaR structures, $Q_\tau(y_t|\mathcal{F}_{t-1};\theta)$ are typically modelled in a 'semi-linear' way, such as

$$Q_\tau(y_t) = \theta_0 + \sum_{i=1}^{p}\theta_i Q_\tau(y_{t-i}) + l(\mathcal{F}_{t-1};\theta_{p+1},...,\theta_{p+q}), \tag{2.4}$$

in which $\mathcal{F}_{t-1}$ is the information set up to and including time $t-1$ and $l(\cdot)$ is a possibly non-linear function (cf. Engle and Manganelli, 2004).

2.2.3 Popular CAViaR Models

Building on the general model above, Engle and Manganelli (2004) propose a range of possible specifications for CAViaR models: A simple CAViaR model setup is 'Symmetric-Absolute-Value' CAViaR (SAV-CAViaR), which may generally be written as

$$Q_\tau(y_t) = \theta_0 + \sum_{i=1}^{p}\theta_i Q_\tau(y_{t-i}) + \sum_{j=1}^{q}\theta_{p+j}\left|y_{t-j}\right|. \tag{2.5}$$

Allowing for asymmetry in the effect of lagged variables on the quantile evolvement depending on their signs, the SAV-CAViaR model extends to the 'Asymmetric-Slope' CAViaR (AS-CAViaR) model:

$$Q_\tau(y_t) = \theta_0 + \sum_{i=1}^{p} \theta_i Q_\tau(y_{t-i}) + \sum_{j=1}^{q} \theta_{p+j}(y_{t-j})^+ + \sum_{l=1}^{q} \theta_{p+q+l}(y_{t-l})^- . \quad (2.6)$$

This specification may be rewritten as

$$Q_\tau(y_t) = \theta_0 + \sum_{i=1}^{p} \theta_i Q_\tau(y_{t-i}) + \sum_{j=1}^{q} \left(\theta_{p+j} + \theta'_{p+q+j}\mathbb{I}[y_{t-j} < 0]\right) |y_{t-j}| , \quad (2.7)$$

where $\theta'_{p+q+j} = -(\theta_{p+q+j} + \theta_{p+j})$ and which is the form used in this chapter. Setting $\theta'_{p+q+j} = 0$, the AS-CAViaR collapses to the SAV-CAViaR model.

Another model adopted in this chapter is the so-called 'Indirect-GARCH' CAViaR (IGARCH-CAViaR) model, specified as

$$Q_\tau(y_t) = (1 - 2 \cdot \mathbb{I}[\tau < 0.5]) \cdot \sqrt{\left(\theta_0 + \sum_{i=1}^{p} \theta_i (Q_\tau(y_{t-i}))^2 + \sum_{j=1}^{q} \theta_{p+j} y_{t-j}^2\right)} . \quad (2.8)$$

Including an asymmetric component similar to the AS-CAViaR model above is obviously also possible (although not mentioned in Engle and Manganelli, 2004) and might be done in the following way (for reasons explained in the next section, this model is dubbed GJR-IGARCH-

CAViaR):

$$Q_\tau(y_t) = (1 - 2 \cdot \mathbb{I}[\tau < 0.5]) \cdot$$
$$\sqrt{\left(\theta_0 + \sum_{i=1}^{p} \theta_i (Q_\tau(y_{t-i}))^2 + \sum_{j=1}^{q} (\theta_{p+j} + \theta_{p+q+j} \mathbb{I}[y_{t-j} < 0]) y_{t-j}^2 \right)};$$
$$\theta_k > 0 \text{ for } \forall k. \quad (2.9)$$

All these CAViaR processes above exhibit similarities to GARCH-type models and indeed the resemblance is not a coincidental: The next section examines the link between the ARMA-GARCH and CAViaR model classes.

2.3 CAViaR in Relation to ARMA-GARCH

CAViaR models describe the evolution of quantiles of the distribution of financial returns over time: As such they implicitly model the evolvement of *all* moments of this distribution, contrary e.g. to ARMA-GARCH-type models, which only describe the first two. Furthermore, contrary to ARMA-GARCH, CAViaR models in their general form are semi-parametric in the sense that they do not assume a particular error distribution or other properties of this distribution such as '*i.i.d.*-ness'[3].

Therefore, CAViaR models nest many other popular model choices in financial econometrics and risk modelling, including notably quantile models based on the ARMA-GARCH class. This also means that there is a direct link between ARMA-GARCH and CAViaR:

[3]'*i.i.d.*-ness' here denotes the property of *i*ndependently and *i*dentically *d*istributed errors.

Proposition 1 *Let y_t follow an ARMA-GARCH-type process of the sort $y_t = \mu_t(\phi)+\sigma_t(\omega)\varepsilon_t$, with $\varepsilon_t \sim$ i.d.d.,$\forall t \in \{0, ..., T\}$ and $\mu_t(\phi), \sigma_t(\omega)$ are $\mathcal{F}_{t-1}$ – adapted with parameter vectors ϕ, ω respectively. The corresponding CAViaR models is then given by $Q_\tau(y_t|\mathcal{F}_{t-1};\theta) = \mu_t(\phi) + \sigma_t(\omega)Q(\varepsilon)$.*

Proof. Proposition 1 can be derived as follows. Given the assumptions, one has

$$
\begin{aligned}
Q_\tau(y_t|\mathcal{F}_{t-1};\theta) &= Q_\tau[\mu_t(\phi) + \sigma_t(\omega)\varepsilon_t|\mathcal{F}_{t-1};\theta] \\
&= Q_\tau[\mu_t(\phi)|\mathcal{F}_{t-1};\theta] + Q_\tau[\sigma_t(\omega)\varepsilon_t|\mathcal{F}_{t-1};\theta] \\
&= \mu_t(\phi) + \sigma_t(\omega) \cdot Q(\varepsilon_t) \\
&= \mu_t(\phi) + \sigma_t(\omega) \cdot Q(\varepsilon).
\end{aligned}
$$

The second last equality follows from that $\mu_t(\phi), \sigma_t(\omega)$ are both $\mathcal{F}_{t-1}$-adapted. The last equality is a consequence of $\varepsilon_t \sim$ i.d.d. $\forall t \in \{0, ..., T\}$. ∎

Mirroring that CAViaR models do not rely on a specific (parametric) assumptions for an error term, the existence of a ARMA-GARCH-type model as in proposition 1 is a sufficient condition for the derivation of a corresponding CAViaR model, however it not necessary one: Thus, a ARMA-GARCH (under the assumptions of proposition 1) has a unique corresponding CAViaR model, yet the converse does not hold.

Based on this result it is now possible to establish a one-to-one correspondence of popular ARMA-GARCH models with the specific CAViaR models above. Since AS-CAViaR nests SAV-CAViaR, the following result is presented only for AS-CAViaR:

Corollary 2 *If y_t follows Zakoan's (1994) TGARCH process of the sort*

$$y_t = \mu_t + \sigma_t \varepsilon_t \tag{2.10a}$$
$$\text{with } \varepsilon_t \sim i.i.d. \ \forall t \in \{0, ..., T\} \tag{2.10b}$$
$$\text{and } \sigma_t = \omega + \sum_{i=1}^{p} \alpha_i \sigma_{t-i} + \sum_{j=1}^{q} (\beta_j + \gamma_j \mathbb{I}[y_{t-j} < 0]) \left| y_{t-j} \right|, \tag{2.10c}$$
$$\text{for } p, q \in \{1, ..., t\}; \omega > 0; \gamma_j \geq 0; j = 1, ..., q; i = 1, ..., p,$$

with $\alpha_i, \beta_j, \forall i, j$ adhering to the conditions for the positivity of σ_t as in Nelson and Cao (1992) and a constant mean component $\mu_t = \bar{\mu}; \forall \{t\}$, the conditional τ quantile of y_t is correctly specified by an AS-CAViaR process

$$Q_\tau(y_t) = \theta_0 + \sum_{i=1}^{p} \theta_i Q_\tau(y_{t-i}) + \sum_{j=1}^{q} (\theta_{p+j} + \theta_{p+q+j} \mathbb{I}[y_{t-j} < 0]) \left| y_{t-j} \right| \tag{2.11}$$

with

$$\begin{aligned}
\theta_0 &= \bar{\mu} + \omega Q_\tau(\varepsilon); \\
\theta_i &= \alpha_i, \forall i \in \{1, ..., p\}; \\
\theta_{p+j} &= \beta_j Q_\tau(\varepsilon), \forall j \in \{1, ..., q\}; \\
\theta_{p+q+j} &= \gamma_j Q_\tau(\varepsilon), \forall j \in \{1, ..., q\}.
\end{aligned}$$

For SAV-CAViaR, the result is the same, except that the asymmetric terms $\theta_{p+q+j} = \gamma_j = 0$, therefore linking SAV-CAViaR with Taylor's (1986) and Schwert's (1998) 'Absolute-Value GARCH' (AV-

GARCH) model, in which

$$\sigma_t = \omega + \sum_{i=1}^{p} \alpha_i \sigma_{t-i} + \sum_{j=1}^{q} \beta_j \left| y_{t-j} \right| . \tag{2.12}$$

An analogous result can obviously be established for the IGARCH-CAViaR model above, linking it in the same way as above directly to Bollerslev's (1986) well-established GARCH model. When an asymmetric term is introduced as in (2.9), again, a correspondence can be established in the same way as above to the 'GJR-GARCH' model by Glosten et al. (1993), thus prompting the name GJR-IGARCH-CAViaR[4].

Incidentally, for the benefit of this chapter, having established the relationship between the model classes these theoretical parametric links allow for straight-forward Monte Carlo simulation studies: Data may be simulated via a parametric ARMA-GARCH-type model with e.g. a Gaussian error term, the corresponding CAViaR model can then be 'estimated back' and the estimated parameters checked against the assumed correct ones. We carry out such a study in section 2.6.

2.4 The Self-exciting Threshold CAViaR Model

In basic applications, time series models are usually applied to data in affine, stationary, single-regime setups usually in order to facilitate estimation and meaningful statistical analysis. Empirical and theoretical research, however, suggests that in certain settings this (often implicit) modelling assumption might not be ideal to capture the highly

[4]As there are many other GARCH-type models (often including asymmetric 'leverage terms'; popular choices are inter alia Ding et al., 1993; Engle and Ng, 1993; Hentshel, 1995; Nelson, 1991), more corresponding CAViaR processes are conceivable.

varying dynamics observed in financial time series data. Especially in the case of (market) risk, recent theoretical findings suggest that endogenous self-exciting behavioural patterns play a significant role and should be incorporated in empirical models.

2.4.1 Endogenous Self-Reinforcing Risk

Stylised behavioural facts as well as theoretical research show that there are explosive, i.e. non-stationary patterns to be observed in financial returns data and associated (market) risk: From a theoretical perspective, one avenue through which such endogenous risk that has been documented by Morris and Shin (1999), Daníelsson and Shin (2003) and Morris and Shin (2004), can arise within the financial system are higher order coordination effects in the sense of Morris and Shin (1998). Loosely speaking, the mechanism at work at the creation of endogenous risk consists of a big enough initial 'white noise' shock creating amplification effects that "gather momentum from the endogenous responses of the market participants themselves" (cf. Morris and Shin, 2004, p. 2): When asset prices fall below a certain "trigger point", market participants' selling orders create selling pressure among other market participants, sparking further selling rounds and so on - leading to a downward spiral in asset prices without the need of any fundamental shocks driving the price development. However, as in Morris and Shin (2004), the 'fallacy' is eventually corrected, asset prices rebound and revert to fundamentals, resulting in v-shaped price paths. From a risk perspective, these patterns constitute a sudden, explosive build-up of risk and returns volatility feeding on themselves, followed by a collapse back to a 'normal' state of the world once critical benchmarks are passed.

2.4.2 The Case for Self-exciting Non-linear Dynamics

The above suggests that affine, stationary single regime setups in empirical risk modelling cannot be expected to capture the entirety of risk and return dynamics in financial markets. While this issue has been known for some time and is documented in theoretical research, the translation into an empirical model is far from trivial: As suggested by the variety of modelling approaches in empirical research in different fields of finance and economics, there is no unique 'correct' empirical methodology to address issues of non-linearity. In the following, we provide an (non-exhaustive) overview of such empirical modelling approaches and highlight this chapter's modelling choice for the issue of endogenous risk:

Generally speaking, while a stationary, single regime model setup might be sufficient in small samples over which the 'state of the world' can be assumed to be stable, there appears to be considerable evidence to suspect that (financial) time series data, especially when analysed over longer horizons, exhibit different regimes or structural breaks and therefore cannot be described by a single, stable process: This point is mirrored in early work by Chow (1960), Quandt (1960) and subsequently Brown et al. (1975) who devise tests for structural breaks in simple linear regression models. Since then numerous studies have been undertaken on the estimation as well as on the testing for structural breaks in all kinds of econometric models and data, including time series, cross-sectional and multivariate models. While early work has focused on the testing of exogenously determined break points (often only in simple linear regressions), more recent research has picked up on the issue of the determination and testing of unknown change points in the model setup. A non-exhaustive list of papers in this area includes Hansen (1992), Andrews (1993), Bai (1994), Andrews et al. (1996), Ghysels et al. (1997), Bai and Perron (1998), Bai et al. (1998), Bai (1999) as well as Bai and Perron (2003).

However, despite the added model and estimation complexity through the 'endogenisation' of break points, econometric modelling with structural breaks has an important shortcoming, in particular in the context of time series: By definition, the *ex-post* estimation and analysis of structural break dates in a given data set is backward-looking and holds little informative value for forecasting and prediction. For example, detecting two structural break dates in the autoregressive process governing the UK post-war consumer price index (CPI) inflation rate from 1947-1987 as in Bai and Perron (2003) does not help in forecasting *future* structural breaks in the process[5]. Since a desired feature of the proposed SET-CAViaR model is forecastability based on past information, structural breaks do not appear to be an appropriate setup for this chapter.

A more suitable non-linear modelling choice in the light of forecastability might be to opt for an Markov-switching approach as pioneered by Goldfeld and Quandt (1973) and further developed by Hamilton (1989), who introduced an AR(1) unit root model in levels with a Markov-switching trend, motivated by the idea of "formalising the statistical identification of turning points of a time series". In the context of volatility modelling, Hamilton and Susmel (1994) consider a three-regime Markov-switching ARCH (dubbed SWARCH) model, again motivated by the fact that GARCH-type models imply very

[5]Strictly speaking, this only holds if one rules out that structural breaks occur on a cyclical basis.

high persistence in volatility over time, which they find is empirically contradicted by their poor forecasting performance for weekly NYSE return volatility from 1962-1987. They further argue that an ARCH model with a four-regime Markov-switching scale is better suited to explain the empirically documented low persistence and provides a better fit to the data as well as better volatility forecasts. Given the documented close ties of CAViaR with GARCH models in the previous section one might therefore be led to believe that a Markov-switching approach might be the natural choice for a non-linear CAViaR model. Here, however, we opt *against* this approach for two specific reasons:

(i) As documented by Hamilton and Susmel (1994, cf. p. 317), autoregressive components cannot be modelled easily in a Markov-switching context, thus ruling out a wide range of interesting CAViaR models, including the ones presented above[6].

(ii) In order to render estimation and statistical inference tractable, Markov-switching models are usually assumed to exhibit stationary behaviour both across and within regimes. While this is a standard assumption, however, it may not be suitable for all situations, especially in the context of risk modelling. According to the above-mentioned theoretical research, at least over some domain, the behaviour of returns and the corresponding risk cannot be expected to be stationary as it becomes endogenised and therefore explosive.

Furthermore, theoretical findings suggest that the evolvement of risk (volatility) and returns is governed by different regimes which are entered into once past returns or risk proxies pass certain *thresholds*:

From an econometric modelling point of view, this calls for a type

[6]Incorporating regime switching into an *autoregressive* structure results in a non-Markovian switching process for which estimation becomes intractable.

of model similar to the 'Self-exciting Threshold Autoregressive' (SETAR) model class by Tong (1983, 1990). The SETAR model allows for explosive, non-stationary behaviour over a part the domain of the autoregressive variable while still maintaining desirable statistical tractability, including overall geometric ergodicity, as long as the so-called deterministic 'skeleton' of the model is 'stable' (cf. Tong, 1990, and s. the following section). The following outlines our take on the SETAR model and presents the Self-exciting CAViaR model and its properties.

2.4.3 SET-CAViaR

The SET-CAViaR model used in this chapter has the following general form:

Definition 3 *$Q_\tau(y_t)$, the τ-quantile at time t of a variable y_t, follows an SET-CAViaR model if*

$$Q_\tau(y_t) = \sum_{j=1}^{J} \mathbb{I}_{y_{t-d}\in R_j} \cdot \left(\theta_0^{(j)} + \sum_{i=1}^{p} \theta_i^{(j)} Q_\tau(y_{t-i}) + l(\mathcal{F}_{t-1}; \theta_{p+1}^{(j)}, ..., \theta_{p+q}^{(j)}) \right),$$

where $\mathbb{I}_{y_{t-d}\in R_j} = \begin{cases} 1 & \text{if } y_{t-d} \in R_j \\ 0 & \text{otherwise} \end{cases}$; $d \in \{1, ..., t-1\}; R_j = [r_{-1}, r_j), j = 1, 2, ..., J$ *($J \in \mathbb{N}^+$ indicating the number of regimes) and* $-\infty = r_0 < r_1 < ... < r_J = \infty$.

In analogy to the literature on SETAR models, d is called the delay parameter and the thresholds are denoted by r_j. The above model is specified for J regimes and thus requires the estimation of $J \cdot (1 + p + q)$ model parameters. If the thresholds and the delay parameter are to be estimated as well, this adds further J parameters to be determined. The AS-CAViaR model in (2.7) in the simple specification with $p = q = 1$ for example would therefore require the estimation of ten parameters in a basic two-regime ($J = 2$) setup. Clearly, with more regimes and more complex base model specifications, the number of parameters to be estimated can increase very quickly, which in the case of CAViaR models poses a problem since the estimation and identification of parameters is not as straight-forward as in the case of parametric (Q)ML estimation.

As the above suggests that a setup with two thresholds and three regimes, i.e. a setting with a normal, an explosive and a 'calming' regime, is apposite to model the dynamics of endogenous risk, there is clearly a need for a more parsimonious model specification.

Scaled SET-CAViaR with Endogenous Thresholds

In this chapter, the dynamics of endogenous risk are modelled by means of a three-regime scaled SET-CAViaR model: This amounts to employing either AS-CAViaR or IGARCH-CAViaR as a constant 'base' quantile model, which is then 'scaled' according to the reigning regime. This way, one avoids having to estimate all base model parameters differently for all regimes. On the basis of AS-CAViaR the specification, dubbed SET-AS-CAViaR, looks as follows:

$$Q_\tau(y_t) = \sum_{j=1}^{3} \mathbb{I}_{|y_{t-1}|\in R_j} \cdot \kappa^{(j)} \cdot Q_\tau^{BASE}(y_t), \text{with}$$

$$Q_\tau^{BASE}(y_t) = \begin{pmatrix} \theta_0 + \sum_{i=1}^{p} \theta_i Q_\tau(y_{t-i}) + \\ \sum_{j=1}^{q} (\theta_{p+j} + \theta_{p+q+j}\mathbb{I}[y_{t-j} < 0])\,|y_{t-j}| \end{pmatrix} \tag{2.13}$$

and

$$\mathbb{I}_{|y_{t-1}|\in R_j} = \begin{cases} 1 & \text{if } |y_{t-1}| \in R_j \\ 0 & \text{otherwise} \end{cases}, \; R_j = \begin{cases} [0, r_1) & \text{for } j = 1 \\ [r_1, r_2) & \text{for } j = 2 \\ [r_2, \infty) & \text{for } j = 3 \end{cases},$$

$$r_1, r_2 \in \mathbb{R}^+; \kappa^{(j)} \in \mathbb{R}^+; j = 1, 2, 3.$$

One notes that the regimes and thresholds are positioned symmetrically around 0, further facilitating parsimonious estimation. The asymmetry in the reaction of quantiles to past returns will be 'picked up' by the asymmetry terms with parameters $\theta_{p+q+j}, j = 1, .., q$ in the base model. Furthermore, the delay parameter is set to 1 to reflect the influence of the immediate returns history on risk as suggested by the literature on endogenous risk.

In order to facilitate the identification of the parameters $\kappa^{(1)}$ is normalised to 1. Obviously, in the light of endogenous risk, one would expect $\kappa^{(2)} > \kappa^{(1)}$ and $\kappa^{(3)} \leq 1$. This way, the quantile, sc. risk, evolvement would be 'explosive' for absolute lagged returns between r_1 and r_2 and enter into a 'calming' regime for absolute lagged returns passing the larger (in absolute terms) threshold r_2.

At the estimation stage, one may choose to estimate the scales freely in an unrestricted model and then check the results against the theoretically desirable restrictions. Alternatively, the restrictions may be enforced during the MCMC estimation routine outlined in the next section. Evidently, an unrestricted estimation producing 'sensible' results is to be preferred to a restricted routine.

In this chapter, as mentioned above, we also estimate the thresholds r_1 and r_2, instead of determining them exogenously as is for example commonplace in a lot of the literature on regime breaks. The MCMC

LTE routine used in this chapter allows for the endogenous determination of the thresholds 'in one go' and can be set up, as with other restrictions, to ensure their staggered order. Contrary to the 'desirable' restrictions on $\kappa^{(1)}$,$\kappa^{(2)}$ and $\kappa^{(3)}$ above, it is necessary to enforce a restriction on the thresholds, i.e. $r_1 < r_2$, at the estimation stage to avoid 'cycling' of the estimation algorithm and to be able to identify the model[7].

Due to the semi-parametric nature of CAViaR models, the analysis of statistical properties such as stationarity is not very meaningful for the above non-linear threshold model. It is however instructive (also given the analysis in section 2.3) to study a 'Self-exciting Threshold' GARCH specification (SET-GARCH) from which (however not exclusively) the above SET-AS-CAViaR model obtains:

$$\begin{aligned}
y_t &= \sigma_t \varepsilon_t, \text{with } \varepsilon_t \sim i.i.d., \forall t \in \{0, ..., T\} \text{ and} \\
\sigma_t &= \sum_{j=1}^{3} \mathbb{I}_{|y_{t-1}| \in R_j} \cdot \kappa^{(j)} \cdot \sigma_t^{BASE}, \text{whereby} \qquad (2.14) \\
\sigma_t^{BASE} &= \left(\omega + \sum_{i=1}^{p} \alpha_i \sigma_{t-i} + \sum_{j=1}^{q} (\beta_j + \gamma_j \mathbb{I}[y_{t-j} < 0]) \, |y_{t-j}| \right),
\end{aligned}$$

and

$$\mathbb{I}_{|y_{t-1}| \in R_j} = \begin{cases} 1 & \text{if } |y_{t-1}| \in R_j \\ 0 & \text{otherwise} \end{cases}, \; R_j = \begin{cases} [0, r_1) & \text{for } j = 1 \\ [r_1, r_2) & \text{for } j = 2 \\ [r_2, \infty) & \text{for } j = 3 \end{cases},$$

$$r_1, r_2 \in \mathbb{R}^+; \kappa^{(j)} \in \mathbb{R}^+, j = 1, 2, 3; \omega, \alpha_k, \beta_l, \gamma_m \text{ for } \forall k, l, m$$

[7]For details on the estimation of the thresholds please consult appendix 2.5.1 below. Further information on the dynamic enforcement of parameter restrictions can be found in section 2.7.2.

The parameters of the base volatility model σ_t^{BASE} map into the base model parameters of the SET-CAViaR model above according to Corollary 2. Following Lee and Shin (2005, p. 28), the base volatility model is strictly stationary if

$$\left[\sum_{i=1}^{p}|\alpha_i| + \sum_{j=1}^{q}(|\beta_j| + \frac{|\gamma_j|}{2})\mathbb{E}[|\varepsilon_{t-1}|]\right] < 1. \tag{2.15}$$

The scale parameters $\kappa^{(1)}$,$\kappa^{(2)}$ and $\kappa^{(3)}$ are the same in both the GARCH and CAViaR representation. Thus, as long as the above theoretical restrictions on the scales, i.e. $\kappa^{(1)} = 1, \kappa^{(2)} > \kappa^{(1)}$,$\kappa^{(3)} \leq 1$, hold and the base volatility model is stationary in the sense of Lee and Shin (2005), the above SET-GARCH specification has got a 'stable' deterministic skeleton and is geometrically ergodic. The proof of this claim is almost identical to Zhang et al. (2001, p. 203) and is therefore omitted here. Intuitively, however, it can be seen that the volatility process in (2.14) can only become wholly explosive through the symmetrical 'top' regime R_3. However, with a stationary base volatility model σ_t^{BASE} and $\kappa^{(3)} \leq 1$, this is prevented from happening.

Obviously, an SET-CAViaR model with endogenous thresholds can also be defined on the basis of GJR-IGARCH-CAViaR, which would look as follows:

$$Q_\tau(y_t) = \sum_{j=1}^{3} \mathbb{I}_{y_{t-1}^2 \in R_j} \cdot \kappa^{(j)} \cdot Q_\tau^{BASE}(y_t), \text{with}$$
$$Q_\tau^{BASE}(y_t) = (1 - 2 \cdot \mathbb{I}[\tau < 0.5]) \cdot$$
$$\sqrt{\left(\theta_0 + \sum_{i=1}^{p} \theta_i (Q_\tau(y_{t-i}))^2 + \sum_{j=1}^{q} (\theta_{p+j} + \theta_{p+q+j} \mathbb{I}[y_{t-j} < 0]) y_{t-j}^2\right)} \tag{2.16}$$

and

$$\mathbb{I}_{y_{t-1}^2 \in R_j} = \begin{cases} 1 & \text{if } y_{t-1}^2 \in R_j \\ 0 & \text{otherwise} \end{cases}, \; R_j = \begin{cases} [0, r_1) & \text{for } j = 1 \\ [r_1, r_2) & \text{for } j = 2 \\ [r_{2,}\infty) & \text{for } j = 3 \end{cases},$$
$$r_{1,} r_2 \in \mathbb{R}^+; \kappa^{(j)} \in \mathbb{R}^+, j = 1, 2, 3; \theta_k \text{ for } \forall k,$$

Again, this SET-CAViaR model, which from hereon will be dubbed SET-GJR-IGARCH-CAViaR can be linked in the simple, straightforward way presented in section 2.3 to the following SET-GARCH representation:

$$y_t = \sigma_t \varepsilon_t, \text{with } \varepsilon_t \sim i.i.d. \; \forall t \in \{0, ..., T\} \text{ and}$$
$$\sigma_t = \sum_{j=1}^{3} \mathbb{I}_{y_{t-1}^2 \in R_j} \cdot \kappa^{(j)} \cdot \sigma_t^{BASE}, \text{whereby} \tag{2.17}$$
$$(\sigma_t^{BASE})^2 = \left(\omega + \sum_{i=1}^{p} \alpha_i \sigma_{t-i}^2 + \sum_{j=1}^{q} (\beta_j + \gamma_j \mathbb{I}[y_{t-j} < 0]) y_{t-j}^2\right),$$

and

$$\mathbb{I}_{y_{t-1}^2 \in R_j} = \begin{cases} 1 & \text{if } y_{t-1}^2 \in R_j \\ 0 & \text{otherwise} \end{cases}, \quad R_j = \begin{cases} [0, r_1) & \text{for } j = 1 \\ [r_1, r_2) & \text{for } j = 2 \\ [r_2, \infty) & \text{for } j = 3 \end{cases},$$

$$r_1, r_2 \in \mathbb{R}^+; \kappa^{(j)} \in \mathbb{R}^+, j = 1,2,3; \omega, \alpha_k, \beta_l, \gamma_m \text{ for } \forall k, l, m$$

The same properties as above apply, except that the volatility base model for σ_t^{BASE} in this case is stationary if

$$\left[\sum_{i=1}^{p} |\alpha_i| + \sum_{j=1}^{q} (|\beta_j| + \frac{|\gamma_j|}{2}) \mathbb{E}[|\varepsilon_{t-1}|^2] \right] < 1. \tag{2.18}$$

The SET-GARCH model is geometrically ergodic by the same virtues as above if the same parameter restrictions on $\kappa^{(1)}$, $\kappa^{(2)}$ and $\kappa^{(3)}$ apply and the base volatility model is stationary.

2.5 MCMC LTE

As mentioned above, autoregressive quantile models like CAViaR are difficult to estimate. This stems from the fact that the objective function $\mathcal{L}[u_i(\phi); \tau; N]$ in the minimisation problem (2.2) can be (i) not linear, is (ii) not everywhere convex, thus involving several local minima and maxima and (iii) is not everywhere differentiable. Therefore, standard gradient-based optimisation routines do not apply in a straight-forward way; a way to recover the applicability of gradient-based methods has been proposed by Komunjer (2005), who rewrites equation (2.2) as a QML estimation characterisation with a density from the 'tick-exponential' family and transforms it into the well-known 'minimax' problem. While recovering the differentiability property, her approach still does not mitigate the non-convexity problem, thus running the risk of becoming 'trapped' in local extrema[8]. Furthermore, despite employing the linear CAViaR processes as base models, incorporating endogenous thresholds as parameters in SET-CAViaR 'injects' non-linearity and further non-differentiability into the objective function. A reconciliation of our approach with the characterisation proposed by Komunjer (2005) therefore does not appear feasible.

Even non-gradient-based methods such as simplex routines need to be considered problematic, because, also in their case, the algorithm might get stuck in a local extremum and has been shown to exhibit slow convergence, especially in the case of higher argument dimensions[9]. The same reservations, as well as documented problems surrounding stopping criteria and the rank deficiency of the non-linear quantile regression Jacobians also cloud the feasibility of the interior point algorithm proposed by Koenker and Park (1996) for practical use.

The most promising class of methods for complicated optimisation problems such as the one at hand are simulation-based routines such as simulated annealing, MCMC methods or the 'Differential Evolutionary Genetic Algorithm' (DEGA) by Price and Storn (1997) - a method based on the biological principles of reproduction and mutation. These routines have the advantage that they stochastically explore the parameter space, thereby avoiding the 'local extremum

[8]This might of course be alleviated by starting the routine off with several starting values, thereby however increasing processing time considerably.

[9]A good treatment of the relative merits of various optimization methods can be found e.g. in (Judd, 1998).

trap' and are not reliant on gradients, which allows for the accommodation of non-linear and not everywhere differentiable objective functions. This feature also makes it feasible to introduce thresholds as parameters of the model as in this chapter.

Generally speaking, the DEGA and simulated annealing work by simulating sequences of parameter vector estimates ('parameter candidate vectors') that ultimately converge to a parameter estimate $\hat{\theta} = \theta^*$, which minimises the objective function. MCMC methods on the other hand take an 'indirect route' by forming a Markov chain of parameter vector estimates $\hat{\theta}^{(0)}, ..., \hat{\theta}^{(C)}$ such that the sequence of these parameter vector candidates converges to the posterior distribution of the true θ, i.e. to $f(\theta|\mathbf{y})$, where $\mathbf{y}$ denotes the data. One then proceeds by then taking the mean, median or mode of this empirical distribution, depending on the loss function used[10].

A fairly recent development, the DEGA seems to be a promising algorithm in many practical applications as documented by Price and Storn (1997), but has not (yet) reached much prominence in econometrics. This might be due to the fact that statistical convergence and asymptotic properties have not yet been established (cf. Price and Storn, 1997, p. 11).

In this chapter, we therefore implement an estimation procedure that closely follows the MCMC LTE method developed by Chernozhukov and Hong (2003) and combines it with features taken from simulated annealing. Extending Chernozhukov and Hong (2003), the proposed algorithm implements features for the 'cooling' (lowering) of a 'temperature' parameter and therefore allows for the tuning of the Markov Chain to achieve faster convergence in finite time The reason for tak-

[10]An extensive overview of MCMC methodology can be found e.g. in Robert and Casella (2004).

ing this route is that this implementation of the MCMC LTE has close theoretical ties with Bayesian econometrics, representing a more intuitive approach for econometricians and allows for convergence of the chain of parameter candidate vectors to the global optimum.

2.5.1 MCMC LTE Details

The following MCMC LTE scheme used in this chapter is based on Chernozhukov and Hong (2003): Given a broadly specified objective function $\mathcal{L}_n(\theta; \mathbf{y})$ with $\theta \in \Theta \subset \mathbb{R}^p$ that is to be minimised, the *'quasi' posterior density* is defined as

$$\mathfrak{p}_n(\theta|\mathbf{y}) = \left\{ \frac{e^{-\left[\frac{\mathcal{L}_n(\theta;\mathbf{y})}{\mathcal{T}}\right]}\pi(\theta)}{\int_\Theta e^{-\left[\frac{\mathcal{L}_n(\theta;\mathbf{y})}{\mathcal{T}}\right]}\pi(\theta)d\theta} \right\} \propto e^{-\left[\frac{\mathcal{L}_n(\theta;\mathbf{y})}{\mathcal{T}}\right]}\pi(\theta), \tag{2.19}$$

where $\pi(\theta) > 0$ is the continuous prior density function of θ and n is the sample size. $\mathcal{T} \in \mathbb{R}^+$is called the 'temperature' parameter, which Chernozhukov and Hong (2003) and most MCMC applications set to unity. This parameter simply scales the objective function without harming the results.

Obviously, the quasi posterior density is a true probability distribution density in the probability theoretical sense; however, it is in general not a proper Bayesian posterior as it may not incorporate conditional data densities (likelihood functions).

Given a loss function $\lambda_n[\mathbf{u}]$, we can now construct an estimator $\hat{\theta}$ by deriving

$$\hat{\theta} = \arg\min_{\xi\in\Theta}[\mathcal{Q}_n(\xi|\mathbf{y})], \tag{2.20}$$

where

$$\mathcal{Q}_n(\xi|\mathbf{y}) = \int_{\Theta} \lambda[\theta - \xi] \left\{ \frac{e^{-\left[\frac{\mathcal{L}_n(\theta;\mathbf{y})}{T}\right]}\pi(\theta)}{\int_{\Theta} e^{-\left[\frac{\mathcal{L}_n(\theta;\mathbf{y})}{T}\right]}\pi(\theta)d\theta} \right\} d\theta. \tag{2.21}$$

Depending on the choice of the loss function $\lambda_n[\mathbf{u}]$, the programme in (2.20) will e.g. yield the quasi posterior mean, median or marginal quantiles if $\lambda_n[\mathbf{u}]$ is the squared loss, absolute deviation or check loss function respectively. Clearly, the posterior mode is also an interesting metric.

In the case of the posterior mean for example, with $\lambda_n[\mathbf{u}] = (\sqrt{n}\mathbf{u})^2$, problem (2.20) yields

$$\hat{\theta} = \int_{\Theta} \theta \mathfrak{p}_n(\theta|\mathbf{y})d\theta. \tag{2.22}$$

$\hat{\theta}$ can now readily be obtained via MCMC methods, i.e. by sampling a sequence $\hat{\theta}^{(0)}, ..., \hat{\theta}^{(C)}$ from $\mathfrak{p}_n(\theta|\mathbf{y})$ and then taking the sample average $\hat{\theta} = C^{-1} \sum_{i=1}^{C} \hat{\theta}^{(i)}$. Analogous procedures also hold for the posterior median and marginal quantiles, always involving drawing a sample from the quasi posterior density. In practice, the sampling is accomplished be means of the Metropolis-Hastings (MH) algorithm, first devised by Metropolis et al. (1953) and generalised by Hastings (1970). The algorithm works by creating a Markov chain of candidate vectors $\hat{\theta}^{(i)} \xrightarrow{d} f(\theta)$, by sampling from instrumental density $q(\theta|\xi)$, from which it is easier to sample from than from the true $f(\theta)$. Is is especially useful in those cases, in which sampling from $f(\theta)$ would in practice be infeasible. The requirement for this procedure to work is that the ratio $\frac{f(\theta)}{q(\theta|\xi)}$ needs to be known up to a constant M independent of ξ.

In the given case, the density to simulate from is the quasi posterior in (2.19) which is known up to a constant $\int_{\Theta} e^{-\left[\frac{\mathcal{L}n(\theta;\mathbf{y})}{T}\right]}\pi(\theta)d\theta$. We thus set $f(\theta) \propto e^{-\left[\frac{\mathcal{L}n(\theta;\mathbf{y})}{T}\right]}\pi(\theta)$. The Metropolis-Hastings algorithm in the given case now looks as follows:

I. Choose a starting value $\hat{\theta}^{(0)}$.

II. Generate ξ from $q(\xi\ |\hat{\theta}^{(j)})$, where $j = 0, .., C$.

III. Update $\hat{\theta}^{(j+1)} = \begin{cases} \xi & \text{with probability } \mathbb{P}(\xi) = p(\hat{\theta}^{(j)}, \xi) \\ \hat{\theta}^{(j)} & \text{with probability } \mathbb{P}(\hat{\theta}^{(j)}) = 1 - p(\hat{\theta}^{(j)}, \xi) \end{cases}$,

where the 'acceptance probability' $p(\mathbf{x}, \mathbf{y})$ is

$$p(\mathbf{x}, \mathbf{y}) = \inf \left\{ \frac{e^{-\left[\frac{\mathcal{L}n(\mathbf{y})}{T}\right]}\pi(\mathbf{y})q(\mathbf{x}|\mathbf{y})}{e^{-\left[\frac{\mathcal{L}n(\mathbf{x})}{T}\right]}\pi(\mathbf{x})q(\mathbf{y}|\mathbf{x})}, 1 \right\}. \tag{2.23}$$

Typically, as in the simpler original random walk version of the Metropolis algorithm, the instrumental distribution used is symmetric, i.e. $q(\mathbf{x}|\mathbf{y}) = q(\mathbf{y}|\mathbf{x}) = \psi(|\mathbf{x} - \mathbf{y}|)$, with ψ taken to be a symmetric density around 0, typically Gaussian or Cauchy. In this case, one has

$$p(\mathbf{x}, \mathbf{y}) = \inf \left\{ \frac{e^{-\left[\frac{\mathcal{L}n(\mathbf{y})}{T}\right]}\pi(\mathbf{y})}{e^{-\left[\frac{\mathcal{L}n(\mathbf{x})}{T}\right]}\pi(\mathbf{x})}, 1 \right\} = \inf \left\{ \frac{\pi(\mathbf{y})}{\pi(\mathbf{x})} e^{\left[\frac{\mathcal{L}n(\mathbf{x}) - \mathcal{L}n(\mathbf{y})}{T}\right]}, 1 \right\}.$$

Furthermore, assuming a non-dogmatic 'flat' prior of the sort $\pi(\mathbf{y}) = \pi(\mathbf{x}) = c$, where c is a positive constant, one obtains

$$p(\mathbf{x}, \mathbf{y}) = \inf \left\{ e^{\left[\frac{\mathcal{L}n(\mathbf{x}) - \mathcal{L}n(\mathbf{y})}{T}\right]}, 1 \right\}, \tag{2.24}$$

which is also the setup used in this chapter.

The theoretical convergence properties of the MH algorithm and the

MCMC LT estimator are well established, even for non-convex and not everywhere differentiable objective functions (cf. Chernozhukov and Hong, 2003; Robert and Casella, 2004). However, in practice, with strongly non-linear models and objective functions as in this chapter, convergence of the above algorithm in finite time is problematic. Luckily, in such a situation, 'cooling' the temperature parameter is a promising measure to achieve faster convergence. The next section lays out the details of this procedure.

2.5.2 Cooling the 'Temperature'

In their paper, Metropolis et al. (1953) also provide a way to use the temperature parameter introduced above. The method builds on the following general result:

Proposition 4 *If $g(\theta)$ is a real-valued function with $\theta \in \Theta \subset \mathbb{R}^p$ and $\exists$ a unique θ^* such that*

$$\theta^* = \arg\max_{\xi}[g(\theta)], \text{ it follows for } \mathcal{T} \in \mathbb{R}^+$$

$$\Longrightarrow \lim_{\mathcal{T}\to 0} \frac{\int_\Theta \theta e^{\frac{g(\theta)}{\mathcal{T}}} d\theta}{\int_\Theta e^{\frac{g(\theta)}{\mathcal{T}}} d\theta} = \lim_{\mathcal{T}\to 0} \int_\Theta \theta \frac{e^{\frac{g(\theta)}{\mathcal{T}}}}{\int_\Theta e^{\frac{g(\theta)}{\mathcal{T}}} d\theta} d\theta = \theta^*,$$

if $g(\theta)$ is continuous at θ^*.

Proof. Proposition 4 is based on the 'Laplace' approximation of $g(\theta)$. See Robert and Casella (2004), pp. 188-189 for a sketch of the proof. More rigorous treatments include Pincus (1968), Tierney and Kadane (1986) and Tierney et al. (1989). ∎

The above Proposition mirrors a result by Hwang (1980), whereby the 'Gibbs measure' $e^{\frac{g(\theta)}{\mathcal{T}}}$ exhibits $\mathcal{T}$-convergence to a uniform distribution on the set of global maxima of $g(\theta)$. One now notes the resemblance of the 'pseudo measure'

$$\mathfrak{f}_{\mathcal{T}}(\theta) = \frac{e^{\frac{g(\theta)}{\mathcal{T}}}}{\int_{\Theta} e^{\frac{g(\theta)}{\mathcal{T}}} d\theta} \tag{2.25}$$

to the quasi posterior in (2.19), prompting the following procedure:

Firstly, without loss of generality, one may set $g(\theta) = -\mathcal{L}_n(\theta; \mathbf{y})$[11]. Then, 'cooling the temperature'[12] $\mathcal{T}$ throughout the run of the MCMC chain above, i.e. gradually lowering $\mathcal{T}^(i)$ for moves $i \in \{1, ...b\}$ such that $\mathcal{T}^(1) > \mathcal{T}^(2) > ... > \mathcal{T}^(b)$, $\lim_{m} \to \infty \mathcal{T}^(b) \to 0$, and using

$$\mathfrak{p}_{n,\mathcal{T}^{(i)}}(\theta|\mathbf{y}) = \left\{ \frac{e^{-\left[\frac{\mathcal{L}_n(\theta;\mathbf{y})}{\mathcal{T}^{(i)}}\right]}}{\int_{\Theta} e^{-\left[\frac{\mathcal{L}_n(\theta;\mathbf{y})}{\mathcal{T}^{(i)}}\right]} d\theta} \right\} \tag{2.26}$$

as the now $\mathcal{T}^(i)$-varying quasi posterior[13] in the algorithm, yields a sequence of parameter candidate vectors $\overset{(}{_{\mathcal{T}}}i)^(i)$ that converge to

Obviously, the 'appropriate' cooling of the temperature plays an important role for the convergence of the algorithm: If $\mathcal{T}^{(i)}$ is lowered too slowly, convergence will be slow. If it is lowered to quickly, one runs the risk of becoming trapped in a local maximum or worse, ending up with no extremum point at all.

---

[11] A maximisation problem $\max_{\phi \in \mathbb{R}} \{\mathcal{L}(\phi)\}$ is equivalent to a mininisation problem $\min_{\phi \in \mathbb{R}} \{-\mathcal{L}(\phi)\}$.

[12] This procedure resembles simulated annealing, which has its roots in the engineering sciences: In metallurgy, a metal is hardened by slowly decreasing (annealing) the temperature in which it is forged. The negative of the function $b(\xi)$ to be maximized, i.e. $w(\xi) = -b(\xi)$, which is then to be minimised, is called *energy*.

[13] A result in analogy to proposition 4 can also be derived for the 'full' log likelihood case, whereby $\lim_{\mathcal{T} \to 0} \frac{\int_{\Theta} \theta e^{\mathcal{T}^{-1} l(\theta|\mathbf{y})} \pi(\theta) d\theta}{\int_{\Theta} e^{\mathcal{T}^{-1} l(\theta|\mathbf{y})} \pi(\theta) d\theta} = \theta^*$ for a continuous log likelihood $l(\theta|\mathbf{y})$ with data $\mathbf{y}$, a positive prior density $\pi(\theta)$ and θ^* being the unique maximun likelihood estimator (cf. Tierney et al., 1989). This paper opts for the simpler implemenation of a flat, unformative prior in order to express on agnostic view about prior distributions.

There have been numerous attempts to determine the optimal speed of decrease, notably one by Hjek (1988) who establishes that the optimal $\mathcal{T}^{(i)} = \Im/\log i$, for $i \in \{1, ...b\}$ and $\Im \geq \zeta$, with ζ being a purely theoretical quantity that depends on the sets of global and local maxima and cannot be determined in practice. This result is thus infeasible to be implemented, prompting the use of more 'ad hoc' rules such as a logarithmic decrease with $\mathcal{T}^{(i)} = \gamma/\log i$ with $\gamma > 0$, or an exponential decrease in α, i.e. $\mathcal{T}^{(i)} = \alpha_0^i \mathcal{T}_0$, where $0 < \alpha_0 < 1$ and $\mathcal{T}_0 > 0$ are usually calibrated to ensure convergence.

We implement the latter cooling rule for the initial b moves, the 'tuning phase', during which the temperature is lowered according to a cooling schedule after every $\mathfrak{NB}$ moves until the chain has stabilised. After the tuning phase the chain then runs freely with the cooled temperature value for another m moves, from which parameter estimates are obtained using the posterior mean and mode; details regarding the choice of α_0, $\mathcal{T}_0$, the cooling schedule $\mathfrak{NB}$ during the tuning phase as well as the construction of parameter estimates are laid out in section 2.7.2.

Obviously, it is quite crucial to determine when to end the tuning phase of the chain in order to avoid cooling the temperature too much and thus producing a degenerate chain. The next section contains an heuristic rule to determine b, i.e. when to stop cooling the temperature.

2.5.3 Monitoring Chain Performance

The MH algorithm is known to be a powerful instrument for global optimisation. However, there exist a number of drawbacks that one needs to be aware of when using it (the following statements pertain to the random walk form):

- If the target distribution (in this case the quasi posterior) and the instrumental distribution (which is typically Gaussian) are strongly misaligned, many new draws ξ will be rejected, thus resulting in a lot of 'waste'.

- If the instrumental distribution is too 'large' (in terms of support or, say, variance), draws will often fall in the tail and be rejected, wasting many draws ξ without improving the probability of visiting all the modes (if there are several) of the objective function. In this case the acceptance rate is low.

- If the instrumental distribution is too 'tight', draws will be too close too each other and convergence is bound to be slow, albeit the acceptance rate is high. In the situation of a multimodal objective function in which modes that are separated by areas of very low probability, the identification of global extrema might be problematic in this case, as 'jumps' from one mode to another are very unlikely, thus leaving the risk of becoming trapped in local extrema.

The acceptance rate is thus a paramount metric in assessing the performance of the algorithm. Based on the above, it needs to be noted that for the random walk MH algorithm a high acceptance rate is, counter-intuitively, not desirable. As a heuristic rule Roberts et al. (1997) recommend to use instrumental distributions with an acceptance rate of 0.5 for models with a parameter dimension of up to 2 and 0.25 for models of higher dimension. They also show that a volatility of 2.4 is the optimal choice for a Gaussian instrumental distribution in case of a Gaussian $\mathcal{N}(0, \sigma)$ target distribution, corresponding to a acceptance rate of 0.44.

In the given case the target is not Gaussian, yet monitoring the acceptance rate is still important: This chapter takes the route of algorithmic calibration, introduced by Müller (1991), albeit in a simplified fashion: The acceptance rate is monitored throughout the tuning phase of the chain (b moves) so to adjust the volatility of the Gaussian instrumental distribution whenever the acceptance rate falls out of bounds around 0.25 in accordance with the argument above. Thereafter, once the chain has stabilised, it is allowed to run freely for m moves in order to produce the parameter estimates.

There are appears to be no definitive hard decision criterion in the literature by which to determine when to stop tuning the chain, i.e. cooling the temperature and calibrating the instrumental distribution. However, the implementation of the MCMC LTE algorithm in this chapter will produce highly autocorrelated parameter candidate vectors as $\mathcal{T}^{(b)} \to 0$ and in the limit always the same one as no draws will ever be accepted. Given this property, we propose to use the following heuristic rule: Following an argument analogous to Gallant and Tauchen (2008, in particular pp. 31-45), a good point at which to stop tuning the chain appears to be one at which the autocorrelation functions of the parameter candidate draws start to flatten and rise significantly. It might be argued that at this point the chain has gained traction and needs no further cooling and tuning.

The concrete implementation of this rule in this chapter looks as follows: At various points during the tuning phase of the chain, non-tuned posterior sequences with the latest temperature and instrumental distribution parameterisation are 'branched out' and an average autocorrelation function across all the model parameters is constructed. When this average autocorrelation function starts to list itself up and flatten out, the tuning phase is ended. Obviously, this heuristic criterion requires individual judgement and a 'feel' for the model at hand and is therefore far from perfect. However, in the context of this chapter and given the lack of a hard decision rule, it appears to perform quite well.

2.5.4 Statistical Inference

As with the estimation of CAViaR models, statistical inference on the estimated parameters is difficult. Due to the analytical non-differentiability of the objective function, the construction of 'robust' standard errors based on the 'tick-exponential' density QML estimation interpretation of quantile (auto-)regression as suggested by Komunjer (2005) is not straight-forward in practice.

Moreover, the inclusion of thresholds in the proposed SET-CAViaR model 'injects' further non-linearity and thus complicates analysis even more: The experience gathered in statistical inference on CAViaR models for our purposes shows that in particular the estimation of the inverse Hessian matrix needed for the construction of the 'sandwich' robust covariance matrix produces ambiguous, unstable results already in a simple CAViaR setup without thresholds. This is due to the need for numerical procedures in order to approximate the Hessian, which, in combination with the 'kinked' nature of the objective function, results in high-varying estimates of the second derivatives depending on the choice of numerical differentiation routine[14] and its

[14]A comprehensive overview of methods that can be used for numerical differentiation can be found in Press et al. (1992).

parameterisation.

Furthermore, treating thresholds as model parameters and estimating them endogenously does not permit standard statistical inference: As pointed out by Hansen (1997, 2000) for the case of threshold estimation in a linear regression context (including threshold autoregressive (TAR) models), the asymptotic distribution of the threshold estimates is non-standard; for more complicated threshold model structures the statistical theory of threshold estimation has not even been fully developed yet. In the case of the SET-CAViaR model and its semi-parametric quantile estimation setting, it is thus to be expected that statistical inference on the threshold estimates is equally if not actually more difficult. Obviously, this circumstance also aggravates standard joint statistical inference on the other model parameters, as for example in the case of the estimation of a covariance matrix.

This chapter therefore adopts the route of statistical inference via the construction of 'Laplace-type' (LT) confidence intervals as proposed by Chernozhukov and Hong (2003): After tuning the chain for b moves and achieving convergence for the parameter vector, the $\frac{\alpha}{2}$ and $(1-\frac{\alpha}{2})$ empirical quantiles of the last m parameter vector candidates $\hat{\theta}^{(b+1)}, ..., \hat{\theta}^{(b+m)}$ from the MCMC sequence are used to form the $(1-\alpha)$ confidence interval for the parameter estimate vector $\hat{\theta}$. Naturally, statistical inference via confidence intervals does not allow for the testing of hypotheses or parameter restrictions. In the given circumstances however, they do provide a good empirical measure by which to judge the meaningfulness of the parameter estimates.

Judging the goodness of model fit across specifications is aided by the AIC and BIC information criteria, which in this chapter are constructed from the objective function evaluated at the final parameter vector estimate. This procedure, again, builds on Komunjer's (2005) finding that quantile (auto)regression via the minimisation of Koenker and Bassett's (1978) objective function in (2.1) is equivalent to the QML estimation of a non-standard 'tick-exponential' likelihood function. Therefore, the objective function evaluated at the final parameter vector can be seen as the likelihood estimate, from which the above information criteria may be calculated.

2.6 A Monte Carlo Study

As indicated above, this section provides a Monte Carlo study to establish the properties of the SET-CAViaR model and the MCMC LTE routine when put to use. The aim of this section is in particular to establish the precision and usefulness of the estimation method for different data series containing various degrees of endogenous risk-type, explosive data elements: Loosely speaking and *ceteris paribus*, one should for example expect that more 'extreme' explosive behaviour - i.e. larger scales κ - should be easier and more clearly identifiable than only mild non-linear patterns. Equally, thresholds r_1 and r_2 that lie 'further out' should be harder to estimate simply due to the fact that there are less data points in the extreme ends of the returns spectrum available in order to 'tie down' the threshold estimates precisely.

The route taken in this section is as follows: First, financial return series are simulated via the following self-exciting threshold GARCH specification, which builds on Zakoan's (1994) TGARCH model in its (1,1) form as the volatility base model σ_t^{BASE} and uses a standard

Gaussian error term:

$$y_t = \sigma_t \varepsilon_t, \text{with } \varepsilon_t \sim \text{i.i.d. } \mathcal{N}(0,1), \forall t \in \{0, ..., T\} \text{ and} \tag{2.27a}$$

$$\sigma_t = \mathbb{I}_{|y_{t-1}| \in (R_1 \cup R_3)} \cdot \sigma_t^{BASE} + \mathbb{I}_{|y_{t-1}| \in R_2} \cdot \kappa \cdot \sigma_t^{BASE}, \text{whereby} \tag{2.27b}$$

$$\sigma_t^{BASE} = (\omega + \alpha\sigma_t + (\beta + \gamma\mathbb{I}[y_{t-1} < 0]) |y_{t-1}|), \tag{2.27c}$$

and

$$R_j = \begin{cases} [0, r_1) & \text{for } j = 1 \\ [r_1, r_2) & \text{for } j = 2 \\ [r_2, \infty) & \text{for } j = 3 \end{cases}, \; r_1, r_2 \in \mathbb{R}^+; \kappa \in \mathbb{R}^+.$$

To add realism, the parameters of the volatility base model in (2.27c) are calibrated to values obtained from an estimation of the TGARCH(1,1) version

$$\sigma_t = \omega' + \alpha' \sigma_t + \beta' (|y_{t-1}| - \gamma' y_{t-1}) \tag{2.28}$$

against the 1980-1992 CRSP tape IBM holding returns. The results of this estimation can be found in table 2.1, which for comparison also includes the estimation results for the 1980-2005 CRSP IBM holding returns.

As shown in section 1, the estimates for ω', α', β' and γ' for the 1980-1992 period correspond to the parameters in (2.27c) in the following

Table 2.1: **TGARCH(1,1) parameter estimates**

This table reports parameter estimates of a TGARCH(1,1) model of the form $\sigma_t^{BASE} = \omega' + \alpha' \sigma_t + \beta'(|y_{t-1}| - \gamma' y_{t-1})$ against daily CRSP IBM holding returns from 1980-1992 and 1980-2005. The estimation is carried out using respectively 3287 and 6563 observations in total. P-values are based on robust standard error estimates.

Parameter	**Period** (Beginning-End)			
	1980-1992		**1980-2005**	
	Estimate	**P-value**	**Estimate**	**P-value**
ω'	0.038	0.007	0.015	0.003
α'	0.923	0.017	0.939	0.000
β'	0.067	0.001	0.071	0.000
γ'	0.462	0.000	0.456	0.000

way:

$$\begin{aligned}
\omega &= \omega' = 0.038 \\
\alpha &= \alpha' = 0.923 \\
\beta &= \beta' - \beta'\gamma' = 0.036 \\
\gamma &= \beta'\gamma' = 0.062.
\end{aligned}$$

Given (2.15), a single-regime TGARCH(1,1) of the sort in (2.27c) with the above parameter values is strictly stationary, which would not hold for a model parameterised with the estimates for the 1980-2005 period[15].

The calibrated (stationary) volatility base model is then used to simulate return series for various combinations of R_2 and κ, with

$$\begin{aligned}
&R_2 = [r_1, r_2) \in \{[1.0, 2.0), [2.0, 3.0), ..., [6.0, 7.0)\}, \text{ and} \\
&\kappa \in \{1.5, 2.0, 2.5, ..., 6.0\}.
\end{aligned}$$

[15]Since a stationary base volatility process is needed to allow for 'controlled explosive' behaviour through the scaling with κ between thresholds r_1 and r_2 while still preserving overall 'skeleton' stability, the parameters estimates obtained with the 1980-2005 period are not deemed feasible in this context.

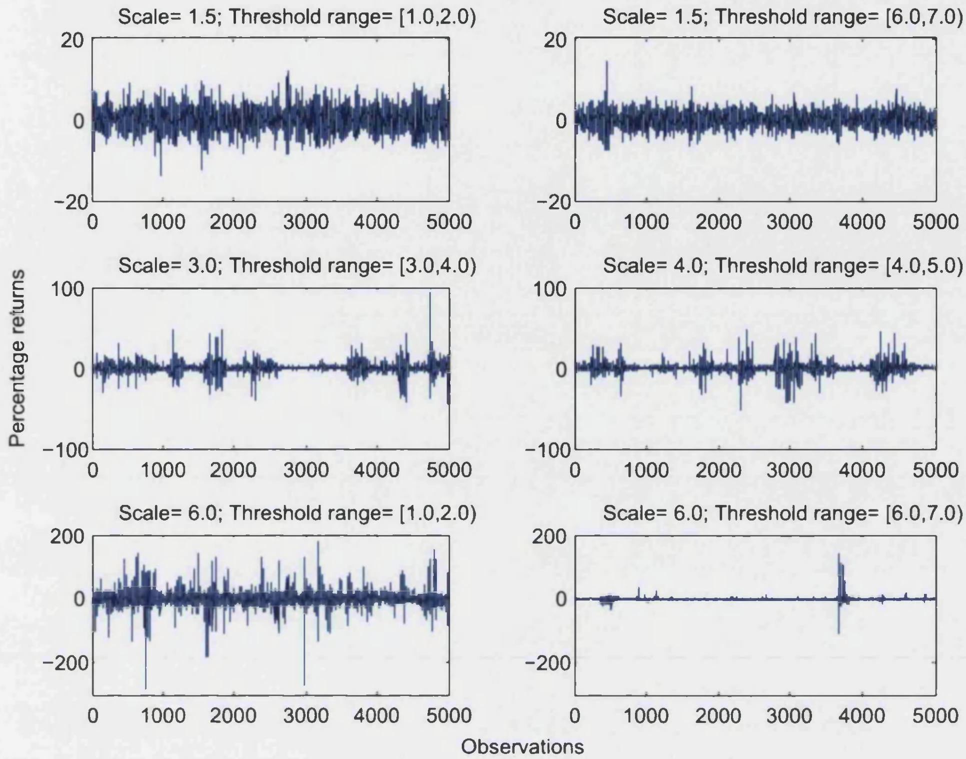

Figure 2.1: **Simulated returns using an SET-GARCH model**

This figure shows various simulated return series of length 5000 using an SET-GARCH model on with a TGARCH(1,1) base volatility process. The figure shows 6 different parameter combinations for threshold ranges and scale.

The set from which to choose R_2 contains threshold intervals reflecting ranges in percentage points within which lagged absolute returns need to lie in order to trigger the explosive regime. In this state of the world, the volatility process is self-exciting, which is modelled by an upscaled base process, with scales ranging from 1.5 to 6.0.

For each combination of R_2 and κ, 20 return series of length $T = 5000$ are simulated. Figure 2.1 shows typical simulated series for various combinations of the scale parameters κ and explosive regime threshold intervals R_2.

The simulated series are then taken as data input to estimate the

Table 2.2: **Parameter linkages SET-CAViaR**

This table shows the correspondence of parameters of a Gaussian SET-GARCH with those of an SET-CAViaR model for the 95% VaR (i.e. for 5% quantile with $\tau = 0.05$). $Q^{\mathcal{N}}_{0.05}(\varepsilon)$ denotes the 5% quantile of a standard Normal distribution.

SET-GARCH	**SET-CAViaR**	
Coefficient	Coefficient	Corresponding Value
ω	θ_0	$Q^{\mathcal{N}}_{0.05}(\varepsilon) \cdot \omega = -1.645 \cdot \omega$
α	θ_1	$Q^{\mathcal{N}}_{0.05}(\varepsilon) \cdot \alpha = -1.645 \cdot \alpha$
β	θ_2	β
γ	θ_3	$Q^{\mathcal{N}}_{0.05}(\varepsilon) \cdot \gamma = -1.645 \cdot \omega$
r_1	r_1	r_1
r_2	r_2	r_2
κ	κ	κ

following SET-CAViaR model on the basis an AS-CAViaR(1,1) base model for the 5% quantile, i.e. for $\tau = 0.05$, thus corresponding to the dynamics of the 95% VaR, a metric widely used in the financial industry:

$$\begin{aligned} Q_{0.05}(y_t) =& \mathbb{I}_{|y_{t-1}|\in(R_1\cup R_3)} \cdot Q^{BASE}_{0.05}(y_t) + \\ & \mathbb{I}_{|y_{t-1}|\in R_2} \cdot \kappa \cdot Q^{BASE}_{0.05}(y_t), \text{with} \end{aligned} \tag{2.29}$$

$$Q^{BASE}_{0.05}(y_t) = \left(\theta_0 + \theta_1 Q_{0.05}(y_{t-1}) + (\theta_2 + \theta_3 \mathbb{I}[y_{t-1} < 0]) \, |y_{t-1}|\right) \tag{2.30}$$

and

$$R_j = \begin{cases} [0, r_1) & \text{for } j = 1 \\ [r_1, r_2) & \text{for } j = 2 \\ [r_2, \infty) & \text{for } j = 3 \end{cases}, \; r_1, r_2 \in \mathbb{R}^+; \kappa \in \mathbb{R}^+.$$

In the light of the analysis of section 2.3, the parameters of the SET-CAViaR model and the Gaussian SET-GARCH model above are linked as displayed in table 2.2.

The aim of the Monte Carlo study is to exploit the theoretical relationship between the model parameters and check the obtained values against the theoretically correct ones. To ease calculation time required for the estimation of the model for $20 \cdot 10 \cdot 6 = 1200$ data series[16] of length 5000 each in the Monte Carlo study only the asymmetry parameter γ, the scale parameter κ and the thresholds r_1 and r_2 are estimated, whereas ω, α and β are fixed to their theoretically correct values during the estimation. This route is motivated by our auxiliary empirical finding (not documented here) that the estimation of SET-CAViaR models produces fairly accurate estimates of the latter model parameters whereas the 'non-linearity' parameters γ, κ and the thresholds tend to be estimated less precisely and are thus of greater interest.

Once the parameter estimates have been obtained for all data series and across all threshold-scale combinations[17], the average of the 20 estimates for every parameter in each category is calculated. This average estimate is then compared to the theoretically correct value according to table 2 and, based on the deviation from the correct value, each threshold-scale category receives an aggregate error score calculated in the following way:

$$ErrScore_{\kappa}^{R_2} = \frac{100}{4} \cdot \begin{pmatrix} |\hat{\gamma}_{av} - \gamma| \\ |\hat{\kappa}_{av} - \kappa| \\ |\hat{r}_{1,av} - r_1| \\ |\hat{r}_{2,av} - r_2| \end{pmatrix}^{\top} \begin{pmatrix} |\gamma^{-1}| \\ |\kappa^{-1}| \\ |r_1^{-1}| \\ |r_2^{-1}| \end{pmatrix}, \tag{2.31}$$

where $\hat{\gamma}_{av}$, $\hat{\kappa}_{av}$, $\hat{r}_{1,av}$ and $\hat{r}_{2,av}$ are the mean parameter estimates for each R_2, κ-category. The proposed error score metric thus constitutes

[16]This number stems from 20 return series for all combinations of 10 threshold intervals and 5 scale parameters values.

[17]For details on the concrete implementation of the estimation algorithm refer to section 2.7.2.

Table 2.3: $ErrScore_{\kappa}^{R_2}$ **for different scale/threshold combinations**

The table shows the $ErrScore_{\kappa}^{R_2}$ metric generated from estimating an SET-CAViaR model on the basis of a simple AS-CAViaR process against several simulated dataseries generated with an associated SET-GARCH model. The $ErrScore_{\kappa}^{R_2}$ is computed from 20 dataseries estimation runs with $\tau = 0.05$ in each scale/threshold range (κ/R_2) combination for $\kappa \in \{1.5, 2.0, 2.5, ..., 6.0\}$ and $R_2 \in \{[1.0, 2.0), [2.0, 3.0), ..., [6.0, 7.0)\}$

Scale	**Threshold range**					
κ	**[1.0-2.0)**	**[2.0-3.0)**	**[3.0-4.0)**	**[4.0-5.0)**	**[5.0-6.0)**	**[6.0-7.0)**
1.5	14.243	3.360	3.409	6.245	12.214	19.718
2.0	0.840	15.248	9.345	3.042	3.575	37.448
2.5	0.713	3.762	4.321	7.409	3.719	10.002
3.0	3.767	7.231	4.056	15.591	2.849	7.811
3.5	3.175	1.850	3.397	12.897	1.885	9.581
4.0	0.814	3.222	3.453	17.922	15.065	2.301
4.5	3.119	3.669	9.787	7.788	3.061	14.914
5.0	3.697	2.608	3.622	21.095	6.084	13.984
5.5	1.230	0.729	10.244	14.425	4.151	11.197
6.0	1.085	2.652	4.328	2.902	8.636	6.762

the average absolute percentage deviation of the estimates from the correct values.

The results of the Monte Carlo study are shown in table 2.3 and graphically in figure 2.2:

For each R_2, κ-category the $ErrScore_{\kappa}^{R_2}$ is plotted against R_2 and κ. It can be seen that, in general, the average percentage deviation is small for low interval thresholds with high scale values and vice versa, as expected. Moreover, the $ErrScore_{\kappa}^{R_2}$ generally increases with magnitude of the thresholds and is also quite high for low scale values. The latter can be interpreted as a consequence of the fact the low scale values κ do not result in very 'marked' return patterns, i.e. not many spikes and poignant explosive patterns, and are thus harder to pin down exactly.

For the increase in the average percentage deviation along with

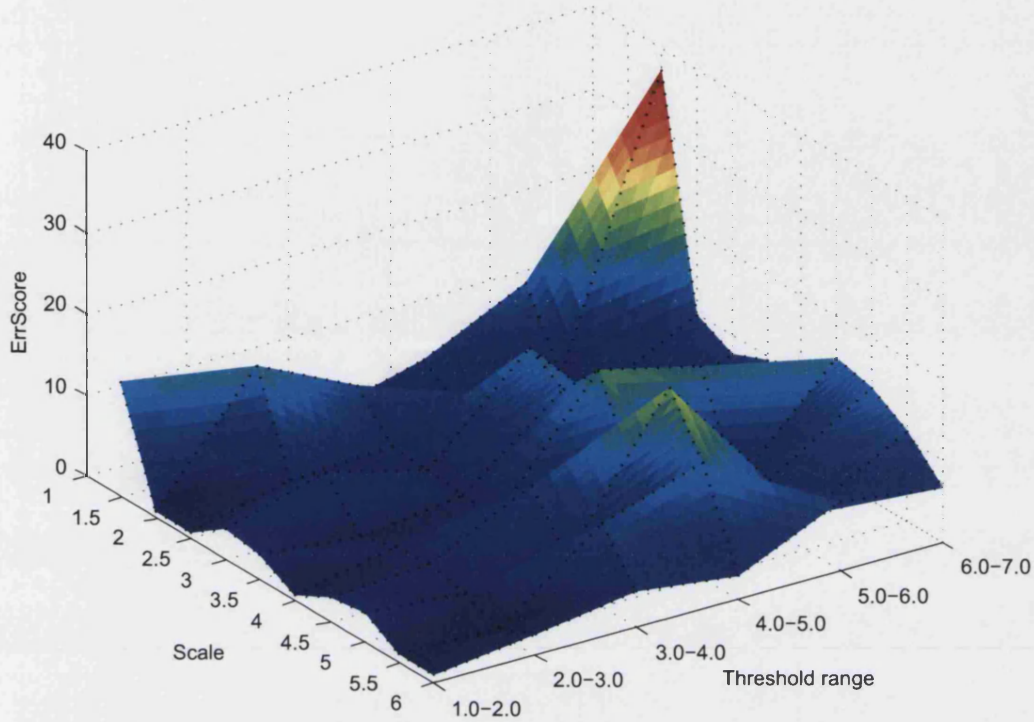

Figure 2.2: $\boldsymbol{ErrScore_{\kappa}^{R_2}}$ **surface plot**

This plot displays the $ErrScore_{\kappa}^{R_2}$ statistic computed from 20 estimates of the SET-CAViaR model in (2.30) for each scale/threshold range (κ/R_2) combination. The scale κ starts at 1.5 and is increased in steps of 0.5 to 6.0, i.e. $\kappa \in \{1.5, 2.0, 2.5, ..., 6.0\}$, and threshold ranges are $R_2 \in \{[1.0, 2.0), [2.0, 3.0), ..., [6.0, 7.0)\}$.

higher threshold magnitudes a different argument is apposite: Since higher thresholds in this setting mean that observations in the explosive regime are relatively scarce compared to data generated with low threshold levels, the precise estimation of scales and thresholds is comparatively harder. In the extreme, as evident from $ErrScore_{\kappa}^{R_2}$ being the highest for the combination of the highest threshold interval $R_2 = [6.0, 7.0)$ with the lowest scale $\kappa = 1.5$, the thresholds can be 'too high', which in combination with a low scale results in the explosive regime being visited only very scarcely and possibly not at all.

Indeed, although not reported here, the high $ErrScore_{\kappa}^{R_2}$ in this combination category is mainly due to a gross misestimation of the threshold levels. Along the same lines as above, as a consequence, an increase in the scale on high thresholds levels eases the identification of the threshold and scale parameters, thereby reducing $ErrScore_{\kappa}^{R_2}$, as can be seen in figure 2.2.

2.7 Empirical Application

In this section, SET-CAViaR models are estimated with 1980-2005 CRSP IBM holding returns. We carry out the estimation for realistic and sensible Value-at-Risk significance levels of 95% and 99%, corresponding to CAViaR models with $\tau = 0.01$ and $\tau = 0.05$ respectively. First, we provide an account of that data used in the estimation.

2.7.1 Data Description

Given the insights from the Monte Carlo section above, to facilitate 'clean' estimation one ideally needs a dataset that besides 'normal' trading activity returns also includes sequences of extreme (negative) returns in order to cleanly identify the model parameters for realistic and sensible Value-at-Risk significance levels.

The financial returns data used in this chapter are the CRSP tape IBM holding returns from the beginning of January 1980 to the end of December 2005, with a total of 6563 observations.

The IBM share constitutes a very liquid and actively traded security for which a long history of holding returns data is available without undesirable data characteristics such share class changes or frequent merger effects. Further, given the high level of trading activity and liquidity in the stock, endogenous risk or herding effects should be identifiable more easily and cleanly than in securities with less trading activity, which usually exhibit more volatility and risk and thus less marked build-ups of extreme returns to start with. It is also to be expected that the proportion of 'informed' traders that move markets solely based on fundamentals and therefore limit the potential for endogenous risk or herding effects as proposed in the theoretical literature on market microstructure[18] is much less significant in a widely traded security such as IBM[19].

Compared to a liquid index such as the e.g. the S&P500 using returns of a single stock such as IBM also has the benefit that in addition to factors influencing the overall market sentiment, company-specific information also plays a role, therefore further facilitating collective effects such herding or endogenous risk. Moreover, the CRSP database allows for the extraction of holding returns only in the case of individual securities. Returns on indices would thus be clouded by frequent price drops due to dividend payments.

The returns data is displayed in figure 2.3, which also details the subperiod from January 1980 to December 1992, used for obtaining the parameters estimates for the TGARCH(1,1) simulation base model in the Monte Carlo section above.

Comparing the two periods, it becomes apparent that the shorter subperiod contains relatively less extreme observations and generally exhibits 'well-behaved' trading conditions - with the notable exception of the 1987 crash clearly standing out. This relative 'calmness' of the market during the 1980-1992 subperiod is also the likely reason for

[18]A comprehensive overview can be found e.g. in (O'Hara, 1995).

[19]Many mutual funds for example are required to hold the stock simply for the very reason of a requirement to track market indices such as the S&P500, in which a large stock such as IBM is a major constituent.

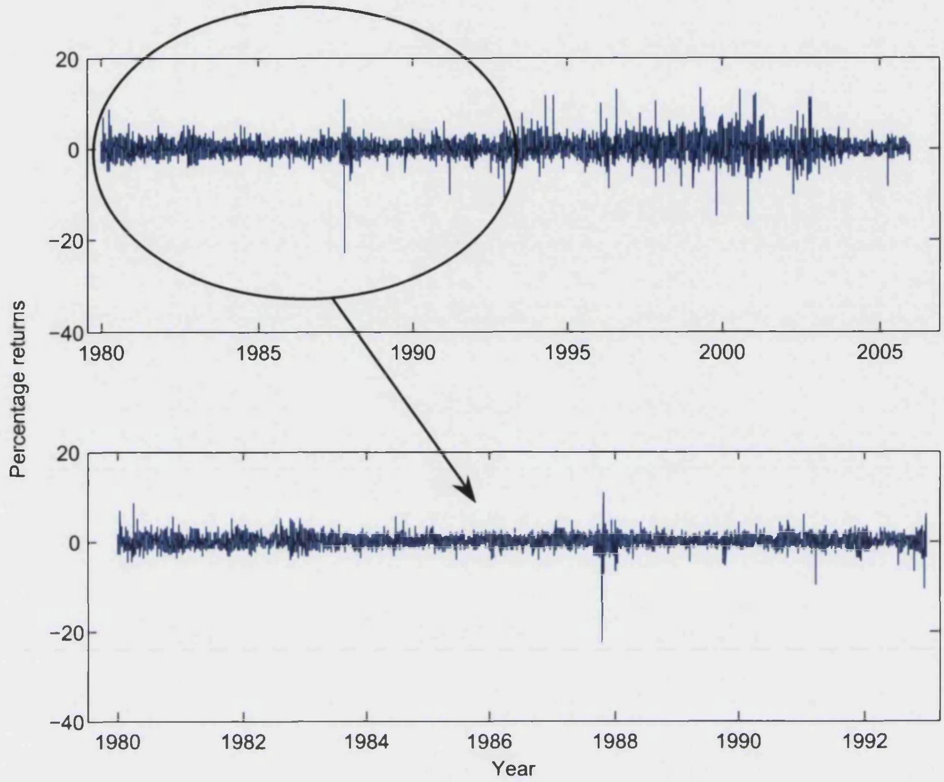

Figure 2.3: **CRSP IBM holding returns**

This figure displays the CRSP IBM holding returns for two horizons: From the beginning of 1980 to the end of 2005 and to the end of 1992 respectively.

obtaining stationary parameter estimates for a simple single-regime TGARCH(1,1) model when applied to the data. Estimating the model with the whole returns dataseries, however, results in non-stationary parameter estimates, leading to the conclusion that a simple, one-regime setup may not be the appropriate modelling choice (cf. table 2.1).

Moreover, the data is strongly non-Gaussian and very leptokurtic with a mean close to zero, albeit these characteristics are more pronounced for 1980-1992 than for the whole time period, as confirmed by table 2.4 and the histograms in figure 2.4.

Table 2.4: **CRSP IBM holding return data statistics**

This table shows basic statistics of the CRSP IBM holding returns for two horizons. The data is strongly non-Gaussian.

Period (Beginning-End)	**1980-1992**	**1980-2005**
Observations	3287	6563
Mean	0.021	0.052
Median	0.000	0.000
Maximum	10.580	13.160
Minimum	−22.960	−22.960
Standard deviation	1.490	1.808
Skewness	−1.111	−0.012
Kurtosis	23.556	13.502
Jarque-Bera (P-value)	58545.93 (0.000)	301656.80 (0.000)

The whole data series therefore constitutes a mix of a strongly non-Gaussian initial subperiod with few very extreme returns and a more volatile second period from 1993-2005 which exhibits more frequent large returns, thus constituting a good basis for the identification of a CAViaR base model as well as for threshold and scale parameters as indicated by the Monte Carlo study.

2.7.2 Implementation and Computing

As the estimation of quantile models is rather elaborate and relies on numerical techniques that are more time consuming than standard gradient-based methods, computation speed is of key importance when implementing estimation routines: For this reason, the code for the simulation and estimation routines in this chapter is written in *C/C++* using standard and widely available free libraries such as the *GNU Scientific Library (GSL)*. Code compilation was carried out using the *gcc compiler v3.3* under Linux[20].

[20]The code is available from the author upon request.

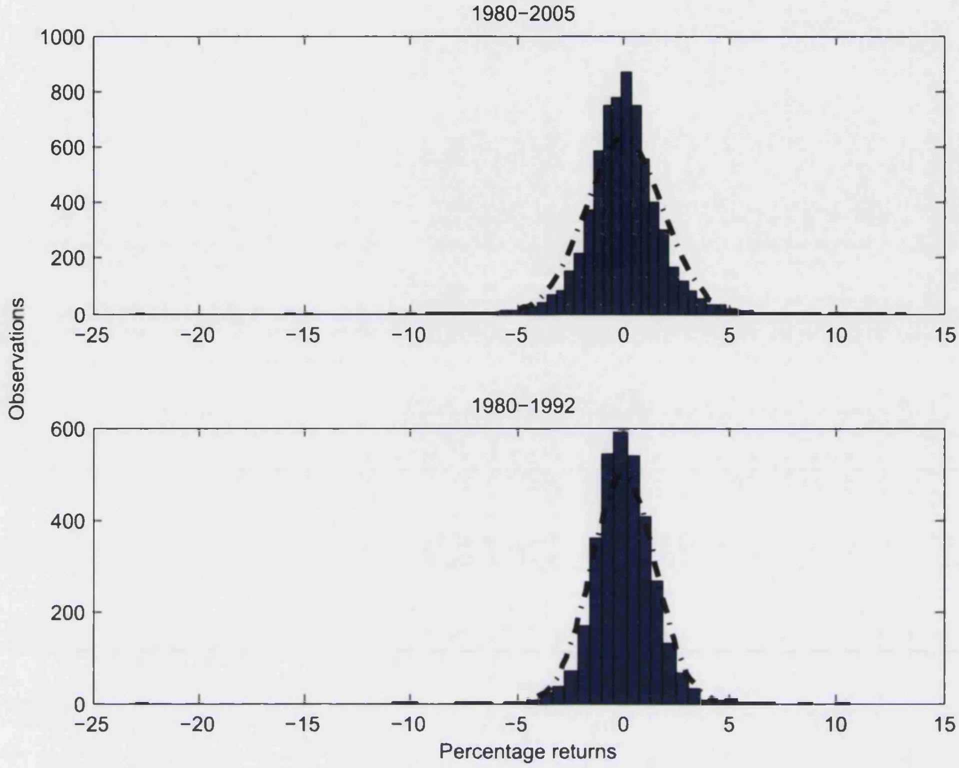

Figure 2.4: **CRSP IBM holding returns histogram**

This figure shows histograms of CRSP IBM holding returns for the 1980-2005 and 1980-1992 time periods respectively. Also displayed as dash-dotted lines are fitted normal distributions with the same mean and variance as the respecitve data series.

The implementation of the MCMC LTE routine consists of $C = (b + m)$ estimation 'moves' for the parameter candidate vector: The first b moves are used for the tuning of the chain (s. also appendices 2.5.2 and 2.5.3) and are divided into estimation 'batch rounds': Each batch round is made up of $\mathfrak{NG}$ 'outer cycles', of which each in turn contains $\mathfrak{NT}$ parameter candidate vector moves, forming the 'inner cycle'. The outer and inner cycle length have both been calibrated to $\mathfrak{NG} = \mathfrak{NT} = 50$. A batch rounds thus comprises $\mathfrak{NB} = \mathfrak{NG} \cdot \mathfrak{NT} = 2500$ individual moves, whereby the inner cycle length $\mathfrak{NT}$ is used to control the frequency at which acceptance rate is monitored and the

Gaussian instrumental distribution adjusted accordingly:

After each inner cycle (thus every $\mathfrak{NT}$ moves), the acceptance ratio is calculated using the previous $\mathfrak{NT}$ moves and, if need be, the volatilities $\sigma_{j,i}$, $j \in \{1, ..., p\}, i \in \{1, ..., b\}$ of the Gaussian instrumental distribution $\phi_i(|\mathbf{x} - \mathbf{y}|) \implies \mathbf{x} \sim \mathcal{N}(\mathbf{y},\Sigma_i) \iff \mathbf{y} \sim \mathcal{N}(\mathbf{x},\Sigma_i)$, with $\mathbf{x}, \mathbf{y} \in \Theta \subset \mathbb{R}^p$, $\Sigma_i = diag(\sigma^2_{1,i}, ..., \sigma^2_{p,i})$[21] are adjusted to keep the acceptance ratio for each parameter in a band of 0.1 around 0.25. This procedure is carried out up to the last m moves, for which the previously 'calibrated' volatilities are kept for each parameter to ensure convergence.

The outer cycle length $\mathfrak{NS}$ us used in combination with $\mathfrak{NT}$ to steer the cooling schedule for the temperature: The initial temperature of the Markov chain is calibrated to $\mathcal{T}^{(0)} = 1.3$, which is lowered after each every batch round, i.e. after every $\mathfrak{NB}$ moves, by a calibrated factor of 0.985. Again, for the last m moves, after the chain has stabilised, the temperature is kept steady to ensure convergence.

It is worth noting that once the chain has stabilised, the 'posterior move length' m is not of crucial importance but should not be too small in order to obtain meaningful estimates from the chain; m is taken to be either $250,000$ for estimation purposes or $10,000$ in order to construct the autocorrelation functions also described in section 2.5.3.

Parameter estimates are obtained by choosing the parameter candidate vector of the last m moves from the chain that yields the mode of the objective function or by taking the sample mean of the last m

[21]In the context of the Markov chain, σ_{ij} in this implementation corresponds to the stepsize with which parameter candidate values are 'proposed' randomly from the instrumental distribution. The routine uses individual stepsizes for each parameter depending on the relative magnitude of the parameters and for some parameters does not need to be adjusted throughout the tuning phase.

candidate vectors. The latter corresponds to estimating the mean of the quasi posterior in equation (2.26) as an estimator for θ^*.

2.7.3 Model Setup

The SET-CAViaR models estimated in this section are sub-cases of the models presented in section 2.4.3. In particular, the (1,1) versions (i.e. $p = q = 1$) of the SET-CAViaR specifications in (2.13) and (2.16) are chosen, resulting in the following respective base quantile processes:

$$Q_\tau^{BASE}(y_t) = \theta_0 + \theta_1 Q_\tau(y_{t-i}) + (\theta_2 + \theta_3 \mathbb{I}[y_{t-1} < 0])\,|y_{t-1}| \qquad (2.32)$$

for SET-AS-CAViaR, and

$$\begin{aligned} Q_\tau^{BASE}(y_t) &= (1 - 2 \cdot \mathbb{I}[\tau < 0.5]) \cdot \\ &\quad \sqrt{\left(\theta_0 + \theta_1 (Q_\tau(y_{t-i}))^2 + (\theta_2 + \theta_3 \mathbb{I}[y_{t-1} < 0]) y_{t-1}^2\right)} \qquad (2.33) \end{aligned}$$

for SET-GJR-IGARCH-CAViaR.

The models are parameterised in their (1,1) form to have a parsimonious enough base quantile process as this simple SET-CAViaR model setup already requires the estimation of 8 parameters against more than 6000 data points in the sample[22]. Moreover, in a GARCH context, the GARCH(1,1) setup has proved itself to be a very resilient and appropriate model in many applications as for example established by Hansen and Lunde (2005): Given that their findings are

[22] As a rough guide, with 100 batch rounds (of 2500 moves each) used for burning in and tuning of the chain and a further 250,000 moves to establish the estimates (cf. section 2.7.2), the computational effort involved in the estimation of the models takes in excess of 5 hours on a standard Intel Pentium® IV processor with 3.0 GHz.

more poignant and in favour of the simple GARCH(1,1) setup in the case FX data in comparison to equity returns, where the leverage effect seems to play a significant role, we take the route of using the (1,1) setup and augmenting it with an asymmetry parameter θ_3. This way, the asymmetry in the reaction of quantiles to past returns, which is analogous to the leverage effect in GARCH models and also expected in the case of CAViaR models when applied to equity data, is 'picked up', while still allowing for a parsimonious model setup (cf. section 2.4.3).

The estimation of the models in this section is carried out for the 95% and 99% levels of VaR, two widely used metrics in the financial industry, which have also been suggested as appropriate measures of market risk by the Basel Committee on Banking Supervision (1996). In the econometric context of CAViaR, 95% and 99% VaR correspond to the $\tau = 0.05$ and $\tau = 0.01$ quantiles of the returns distribution respectively. Below, the results of the estimation with $\tau = 0.05$ are shown first.

2.7.4 Estimation Results for 95% VaR ($\tau = 0.05$)

Applying the above models to the 1980-2005 CRSP IBM holding returns[23] on a 95% VaR level, results in parameter estimates displayed in table 2.5. The table shows the results both for the SET-AS-CAViaR and the SET-GJR-IGARCH-CAViaR models.

Immediately catching the eye is the fact that the magnitude as well as the signs of the parameter estimates 'make sense' in the light of the literature laid out in section 2.4.1. This outcome is quite significant and can by no means be taken for granted given the parameterisation,

[23]The returns are on a percentage scale.

Table 2.5: **SET-CAViaR parameter estimates ($\tau = 0.05$)**

This table shows parameter estimates as well as 95% confidence intervals for an MCMC LTE of an SET-AS-CAViaR and an SET-GJR-IGARCH-CAViaR model with $\tau = 0.05$ and base quantile processes given by (2.32) and (2.33) respectively against 1980-2005 CRSP IBM holding returns. Parameter estimates are displayed as the posterior mean and mode constructed from the Markov chain. Also shown are the loglikelihood as well as the AIC and BIC at the estimated parameter values.

$\mathbf{Q}_\tau^{BASE}(\mathbf{y}_t) = (\theta_0 + \theta_1 Q_\tau(y_{t-i}) + (\theta_2 + \theta_3 \mathbb{I}[y_{t-1} < 0])\,\lvert y_{t-1}\rvert)$			
Metric	**Estimates**		**95% Conf interval**
($\tau = \mathbf{0.05}$)	**Posterior mean**	**Posterior mode**	
$\hat{\theta}_0$	-0.027	-0.021	$(-0.040, -0.015)$
$\hat{\theta}_1$	-0.053	-0.047	$(-0.069, -0.037)$
$\hat{\theta}_2$	0.941	0.948	$(0.926, 0.954)$
$\hat{\theta}_3$	-0.088	-0.079	$(-0.112, -0.067)$
$\hat{r}_1$	10.512	10.671	$(9.944, 10.730)$
$\hat{r}_2$	10.990	10.753	$(10.757, 11.248)$
$\hat{\kappa}^{(2)}$	2.570	1.790	$(1.632, 4.804)$
$\hat{\kappa}^{(3)}$	0.525	0.495	$(0.384, 0.719)$
Loglikelihood	-1189.646	-1188.764	
AIC	2395.292	2393.529	
BIC	2449.604	2447.841	
$\mathbf{Q}_\tau^{BASE}(\mathbf{y}_t) = (1 - 2\cdot\mathbb{I}[\tau < 0.5])\sqrt{(\theta_0 + \theta_1 (Q_\tau(y_{t-i}))^2 + (\theta_2 + \theta_3 \mathbb{I}[y_{t-1} < 0]) y_{t-1}^2)}$			
Metric	**Estimates**		**95% Conf interval**
($\tau = \mathbf{0.05}$)	**Posterior mean**	**Posterior mode**	
$\hat{\theta}_0$	0.059	0.053	$(0.039, 0.091)$
$\hat{\theta}_1$	0.026	0.024	$(0.016, 0.037)$
$\hat{\theta}_2$	0.954	0.956	$(0.936, 0.964)$
$\hat{\theta}_3$	0.116	0.111	$(0.089, 0.162)$
$\hat{r}_1$	113.712	112.609	$(104.508, 115.314)$
$\hat{r}_2$	125.352	118.839	$(115.739, 133.322)$
$\hat{\kappa}^{(2)}$	2.114	1.557	$(1.415, 3.471)$
$\hat{\kappa}^{(3)}$	0.511	0.510	$(0.378, 0.643)$
Loglikelihood	-1200.415	-1199.825	
AIC	2416.833	2415.650	
BIC	2471.146	2469.962	

non-linear structure and semi-parametric nature of the SET-CAViaR models used:

The parameters estimates pertaining to the self-exciting features of the models, i.e. $\hat{r}_1$, $\hat{r}_2$, $\hat{\kappa}^{(2)}$ and $\hat{\kappa}^{(3)}$, suggest that market risk (as measured by 95% VaR, i.e. the 5% return quantile) becomes explosive by a scale factor of roughly 2 (indicated by $\hat{\kappa}^{(2)}$) once yesterday's absolute returns or squared returns pass a threshold of c. 10.5% or 113% (given by $\hat{r}_1$) for SET-AS-CAViaR and SET-GJR-IGARCH-CAViaR respectively. However, once $\hat{r}_2$, the second threshold of c. 11% (in the case of SET-AS-CAViaR) or c. 120% (SET-GJR-IGARCH-CAViaR), is crossed by yesterday's absolute and squared returns respectively, the explosive behaviour changes into a 'calming' regime, in which market risk 'cools down' and shrinks by a factor of roughly 0.5 (indicated by $\hat{\kappa}^{(3)}$). Moreover, the estimates for the asymmetry parameter $\hat{\theta}_3$ also indicate in the case of both SET-CAViaR models that a negative return yesterday prompts a higher increase in market risk (i.e. results in a more negative 5% return quantile) then a positive return, which is in line with similar asymmetry and leverage effect arguments in the GARCH model literature, whereby negative returns are followed by higher increases in volatility than positive returns (cf. e.g. Nelson, 1991).

It is also noteworthy that the scale parameter estimates $\hat{\kappa}^{(2)}$ and $\hat{\kappa}^{(3)}$ have been obtained *without* restrictions to their domain such as $\hat{\kappa}^{(2)} > 1$, $\hat{\kappa}^{(2)} < 1$, which would enforce the explosive and calming model behaviour. Rather, they have been estimated freely on the positive domain.

The table further indicates that even though parameter estimates show close correspondence between the two SET-CAViaR models (in particular the scale estimates are very similar; the threshold estimates are also very close to each other once the square root is applied to those for SET-GJR-IGARCH-CAViaR), the AIC and BIC indicate a better model fit for SET-AS-CAViaR. This is despite the 95% confidence interval being slightly narrower for the scales in the case of SET-GJR-IGARCH-CAViaR.

The Relation of the Base Model to GARCH and Stationarity Issues

For both SET-CAViaR specifications, the base model parameters estimates $\hat{\theta}_i, i = \{0, 1, 2, 3\}$ are in the region of magnitude of that one would expect if the underlying return generator was a SET-GARCH-type model with a constant mean component and an *i.i.d.* error term. According to proposition 1 and the findings in section 2.3, a model such as the one in (2.14) with $p = q = 1$ and a standard Gaussian *i.i.d.* error term for example would map into an SET-AS-CAViaR model in which the base model parameter θ_2 is the same as β_1, and θ_1 for instance would be given by $Q^{\mathcal{N}}_{0.05}(\varepsilon) \cdot \alpha_1 = -1.645 \cdot \alpha_1$, where $Q^{\mathcal{N}}_{0.05}(\varepsilon)$ is the 5% quantile of the standard Gaussian error term. Equally, θ_0 would equal $Q^{\mathcal{N}}_{0.05}(\varepsilon) \cdot \omega = -1.645 \cdot \omega$ and θ_3 correspond to γ_1 via $Q^{\mathcal{N}}_{0.05}(\varepsilon) \cdot \gamma_1 = -1.645 \cdot \gamma_1$.

Similarly, in the case of SET-GJR-IGARCH-CAViaR, the model parameters can be correctly associated with a Gaussian SET-GJR-

GARCH model as in (2.17) with $p = q = 1$, in which

$$\omega = \frac{\theta_0}{\left[Q^{\mathcal{N}}_{0.05}(\varepsilon)\right]^2} = \frac{\theta_0}{-1.645^2}$$
$$\alpha_1 = \frac{\theta_1}{\left[Q^{\mathcal{N}}_{0.05}(\varepsilon)\right]^2} = \frac{\theta_1}{[-1.645^2]^2}$$
$$\gamma_1 = \frac{\theta_3}{\left[Q^{\mathcal{N}}_{0.05}(\varepsilon)\right]^2} = \frac{\theta_3}{[-1.645^2]^2},$$

whereas the GARCH model parameter β_1 would map identically into θ_2.

Again, parameter estimates for the quantile base model in both SET-CAViaR setups have been obtained *without* the need to impose binding constraints on their domain, such as e.g. positiveness in the case of SET-GJR-IGARCH-CAViaR. In fact, the only enforced restriction in the entire estimation procedure for both models with $\tau = 0.05$ is the need to have $\hat{r}_1 < \hat{r}_2$, without which the models would not be identified.

On the basis of the above relationships it is also easily verifiable that, with the exception of the posterior mode estimates for SET-AS-CAViaR, all other sets of parameters estimates may be traced back to corresponding geometrically ergodic SET-GARCH-type specifications: Using (2.15) and (2.18), the quantile base model parameter estimates may be interpreted as corresponding to Gaussian stationary GARCH base volatility processes. In conjunction with the parameter estimates for the scales and thresholds, which would map identically from an SET-CAViaR to a corresponding SET-GARCH model, it is apparent in accordance with the argument in section 2.4.3 that such an SET-GARCH model would possess a stable deterministic skeleton and would thus be geometrically ergodic.

The Performance of the MCMC LTE Routine for 95% Value-at-Risk ($\tau = 0.05$)

The variation of the chain for the different parameter values and thus the width of the confidence intervals decrease with the temperature and depend on the length of tuning phase of the chain[24]. Therefore, finding a good stopping point for temperature cooling and adjustment of the instrumental density variance (cf. section 2.5.3) is important both from a computation time point of view as well as to facilitate meaningful statistical analysis.

Figure 2.5 illustrates the use of the heuristic decision criterion as to when to stop the tuning phase in the MCMC LTE of the two SET-CAViaR models. The figure displays average autocorrelation functions (ACFs), which are constructed as follows: At different points during the tuning phase, which is measured in batch rounds (cf. section 2.7.2 for details), the chain is branched out and allowed to run without interference for 10,000 runs. From these posterior runs the ACFs are computed for the first 1000 lags. Using the ACFs, the chain is tuned until a point is reached where the ACF starts to flatten out and rise (cf. section 2.5.3). In the given estimation, this point is reached after 110 and 140 batch rounds for SET-AS-CAViaR and SET-GJR-IGARCH-CAViaR respectively. It can be seen from the plot that the ACFs after 110 and 140 batch rounds are very similar to the ACF after the relatively low number of 40 batch rounds. Yet, in the case of SET-AS-CAViaR, after 110 batch rounds, the ACF starts to rise as is the case for SET-GJR-IGARCH-CAViaR after 140 batch rounds.

[24] In the limit, with $\mathcal{T}^{(b)} \rightarrow 0$, the width of any confidence interval is 0.

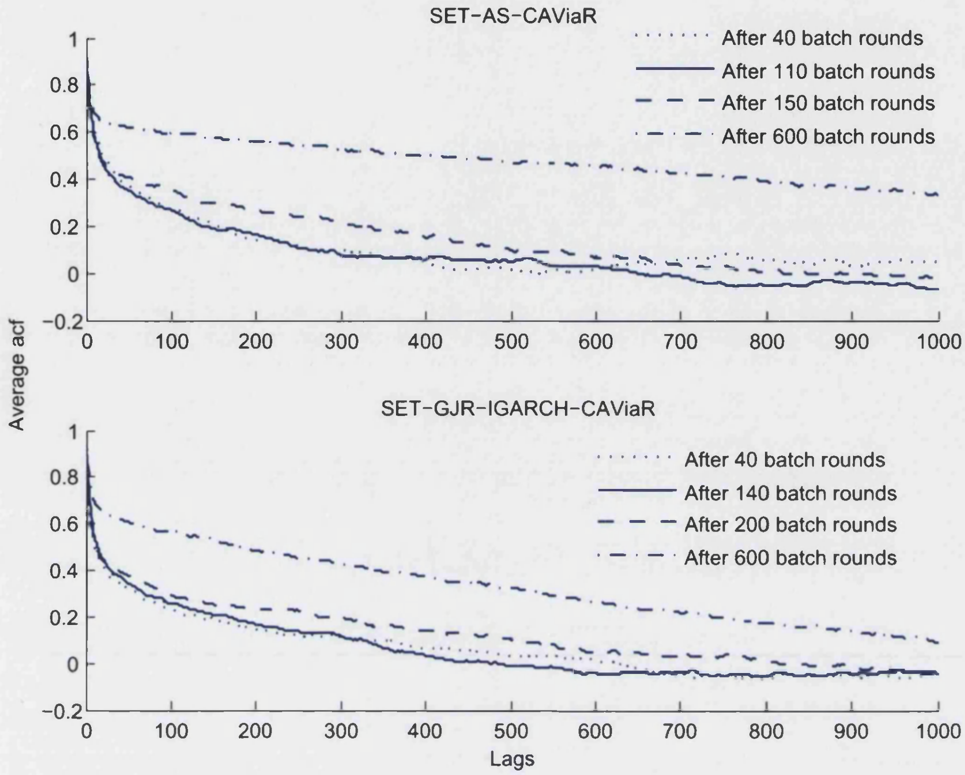

Figure 2.5: **Average autocorrelation functions ($\tau = 0.05$)**

This figure displays the average autocorrelation functions (ACFs) for parameter candidate value in an MCMC LT estimation of the SET-AS-CAViaR and SET-GJR-IGARCH-CAViaR models with $\tau = 0.05$ and base quantile specifications given by (2.32) and (2.33) respectively against 1980-2005 CRSP IBM holding returns. Following tuning phases of different lengths, the average ACFs are computed for the first 1000 lags.

In the extreme case, after a long tuning phase of 600 batch rounds, corresponding to a temperature $\mathcal{T}^{(600)} \approx 0$, the ACFs are very flat and indicate highly autocorrelated parameter candidate draws and thus a degenerate chain.

Figures 2.6 and 2.7 show the evolution of the parameter candidate draws during the tuning phase of the estimation for the SET-AS-CAViaR and the SET-GJR-IGARCH-CAViaR model respectively. For all parameters it can be observed that the variation over the re-

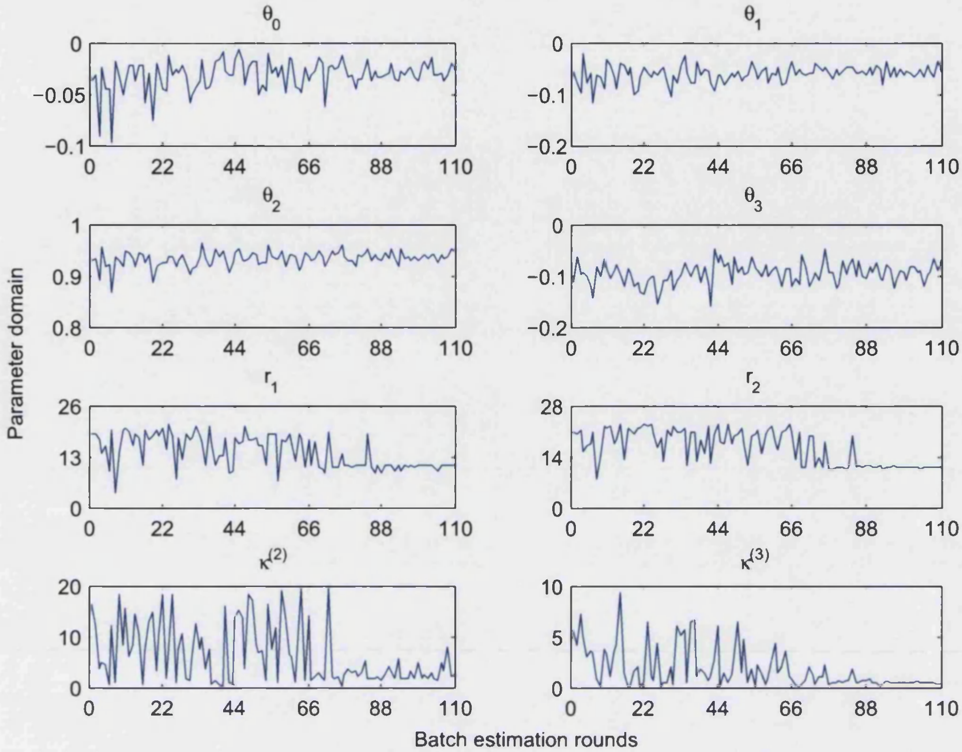

Figure 2.6: **SET-AS-CAViaR para's ($\tau = 0.05$)**

This figure shows the evolution of the parameter candidate values in the MCMC LTE of the SET-AS-CAViaR model with $\tau = 0.05$ and a base quantile process given by (2.32) against 1980-2005 CRSP IBM holding returns. The parameter candidate value draws are plotted after each batch estimation round during the tuning phase of the Markov chain.

spective parameter domains decreases the longer the tuning phase and therefore the more the temperature is cooled as a lower temperature in general makes it more difficult for new draws to be accepted (cf. section 2.5.2). From the figures it is also apparent that the evolution of the base quantile model parameter candidate values is fairly standard, i.e. a fairly steady but not too smooth progression towards the final parameter estimates without extreme changes in variation, which is line with other practical applications of MCMC LTE, e.g. in Gallant and Tauchen (2008). The evolution of the parameters candidate values

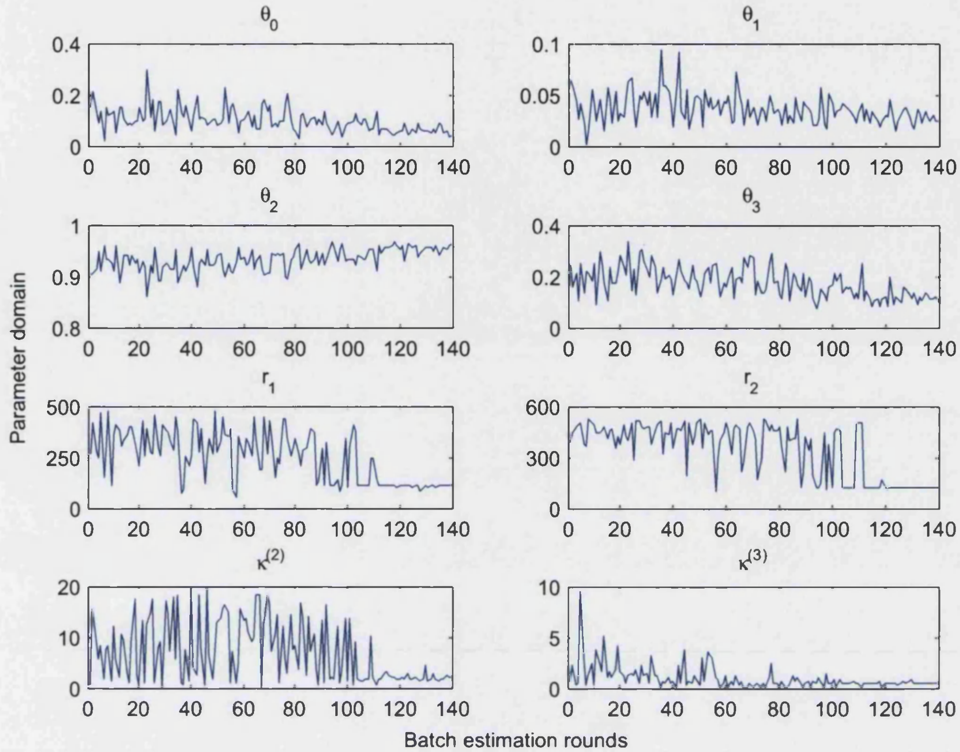

Figure 2.7: **SET-GJR-IGARCH-CAViaR para's ($\tau = 0.05$)**

This figure shows the evolution of the parameter candidate values in the MCMC LTE of the SET-GJR-IGARCH-CAViaR model with $\tau = 0.05$ and a base quantile process given by (2.33) against 1980-2005 CRSP IBM holding returns. The parameter candidate value draws are plotted after each batch estimation round during the tuning phase of the Markov chain.

pertaining to the self-exciting, non-linear features of the models, i.e. $\hat{r}_1$, $\hat{r}_2$, $\hat{\kappa}^{(2)}$ and $\hat{\kappa}^{(3)}$, however, is less standard for both SET-CAViaR models: Initially the variation is very large, especially in the case of $\hat{r}_1$ and $\hat{r}_2$ for SET-GJR-IGARCH-CAViaR, where the thresholds relate to lagged squared percentage returns. Yet, at some point during the tuning phase the large variation is more damped (after c. 80 and 110 batch rounds for SET-AS-CAViaR and SET-GJR-IGARCH-CAViaR respectively) and gives way to a more subdued parameter evolution. A possible explanation for this circumstance is that after sufficient

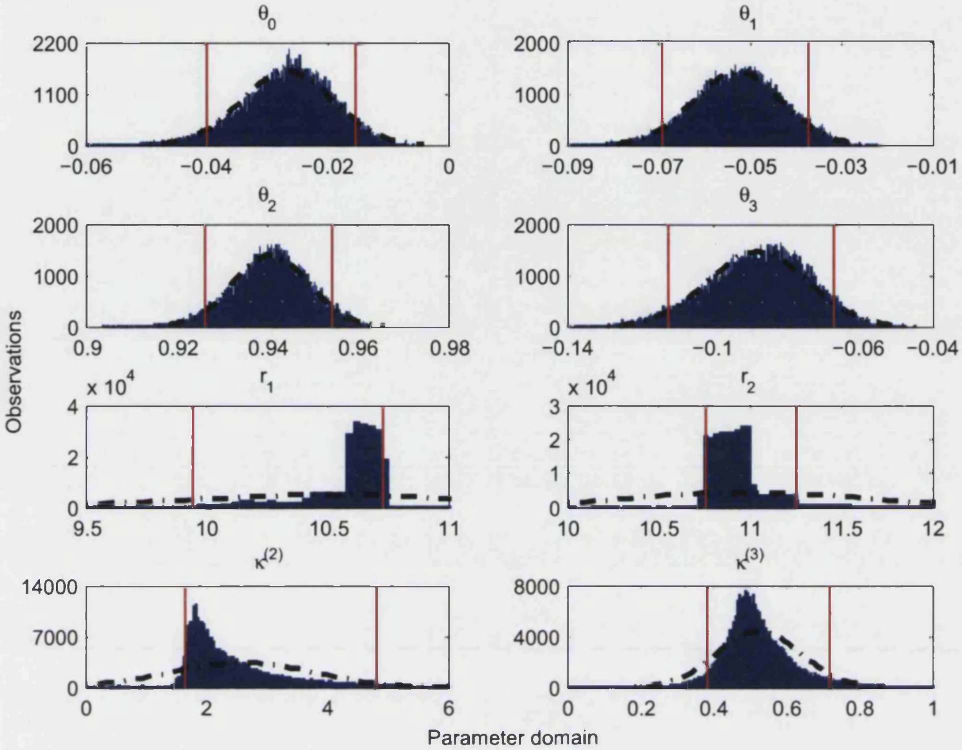

Figure 2.8: **SET-AS-CAViaR posteriors ($\tau = 0.05$)**

This figure displays the posterior distribution histograms for parameter candidate values of the MCMC LTE of an SET-AS-CAViaR model with $\tau = 0.05$ and a base quantile process given by (2.32) against 1980-2005 CRSP IBM holding returns. Also displayed as dash-dotted lines are fitted normal distributions with the same mean and variance as the respective parameter posterior. Additionally, the 95% confidence intervals bounds are shown as vertical lines.

random exploration of the parameter space at these points during the tuning phase, the chain has progressed enough towards the area of the global maximum of the likelihood. In conjunction with a fairly low temperature at these points, the chain now does not 'wander' out of this area again, but produces narrower parameter draws.

Given the strong non-linearity of both SET-CAViaR models in the objective function as well as in the model specification itself, such non-standard behaviour is to be expected[25]. This holds for the non-

[25]The loglikelihood function in such a case is very 'kinked' with a pronounced global maximum,

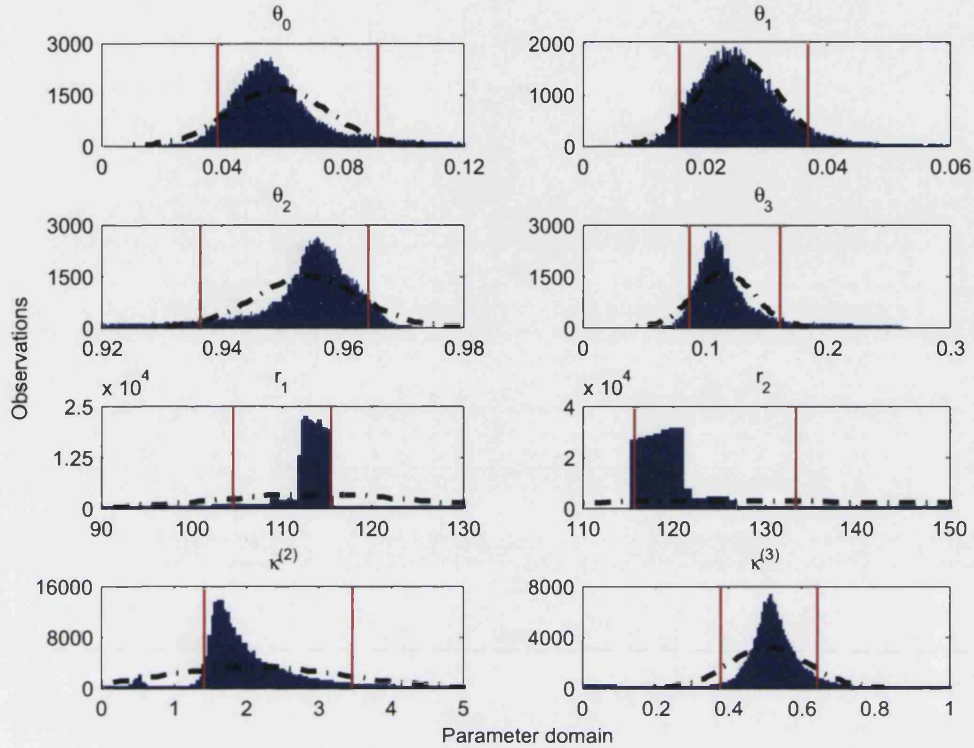

Figure 2.9: **SET-GJR-IGARCH-CAViaR posteriors ($\tau = 0.05$)**

This figure displays the posterior distribution histograms for parameter candidate values of the MCMC LTE of an SET-GJR-IGARCH-CAViaR model with $\tau = 0.05$ and a base quantile process given by (2.33) against 1980-2005 CRSP IBM holding returns. Also displayed as dash-dotted lines are fitted normal distributions with the same mean and variance as the respective parameter posterior. Additionally, the 95% confidence intervals bounds are shown as vertical lines.

linearity parameters $\hat{r}_1$, $\hat{r}_2$, $\hat{\kappa}^{(2)}$ and $\hat{\kappa}^{(3)}$ in particular as their asymptotic distribution cannot be expected to be standard (cf. e.g. Hansen, 1997, 2000).

This feature of the estimated SET-CAViaR models is further highlighted in figures 2.8 and 2.9: Both show the empirical posterior distributions of the parameter candidate draws after the tuning phase of chain. The histograms have been constructed from 250,000 parameter moves after the end of tuning phase and fitted Gaussian distributions

which is difficult to detect.

as well as the bounds of the 95% confidence intervals have been superimposed. From the figure it is evident that for both models the posteriors of $\hat{r}_1$, $\hat{r}_2$, $\hat{\kappa}^{(2)}$ and $\hat{\kappa}^{(3)}$ are non-standard. In particular the histograms for $\hat{r}_1$ and $\hat{r}_2$ do not have a familiar bell-shape form, and resemble a rugged uniform distribution. Indeed, estimation results obtained with longer chain tuning phases (not shown here) seem to confirm the progression towards a such a distribution. In contrast, the posteriors of the base quantile model parameters $\hat{\theta}_i, i = \{0, 1, 2, 3\}$ appear to be standard Gaussian, with the SET-AS-CAViaR model providing a closer match.

The figures further reveal that no constraints other than $\hat{r}_2 > \hat{r}_1$ need to be enforced during the estimation of both SET-CAViaR with $\tau = 0.05$ as the posteriors of all other parameters show no signs of being truncated.

2.7.5 Estimation Results for 99% VaR ($\tau = 0.01$)

The results of the estimation of the SET-AS-CAViaR and the SET-GJR-IGARCH-CAViaR models with the base quantile processes presented in section 2.7.3 for $\tau = 0.01$ are displayed in table 2.6:

As in the previous section, the parameter estimates for both SET-CAViaR models are in an order of magnitude, which intuitively 'makes sense': Once lagged absolute and squared returns cross a threshold of c. 3.5% and 12.6% (given by $\hat{r}_1$) in the case of SET-AS-CAViaR and SET-GJR-IGARCH-CAViaR respectively, both model enter into an explosive regime in which market risk (in this case given by 99% VaR) increases by a scale factor of c. 3.6 (as indicated by $\hat{\kappa}^{(2)}$). However, as soon as absolute and squared lagged returns increase beyond the second threshold ($\hat{r}_2$) of c. 3.7% and 13.4% for SET-AS-CAViaR and

Table 2.6: **SET-CAViaR parameter estimates ($\tau = 0.01$)**

This table shows parameter estimates as well as 95% confidence intervals for an MCMC LTE of an SET-AS-CAViaR and an SET-GJR-IGARCH-CAViaR model with $\tau = 0.01$ and base quantile processes given by (2.32) and (2.33) respectively against 1980-2005 CRSP IBM holding returns. Parameter estimates are displayed as the posterior mean and mode constructed from the Markov chain. Also shown are the loglikelihood as well as the AIC and BIC at the estimated parameter values.

$\mathbf{Q}_\tau^{BASE}(\mathbf{y}_t) = (\theta_0 + \theta_1 Q_\tau(y_{t-i}) + (\theta_2 + \theta_3 \mathbb{I}[y_{t-1} < 0])\,\lvert y_{t-1}\rvert)$			
Metric	**Estimates**		**95% Conf interval**
($\tau = 0.01$)	**Posterior mean**	**Posterior mode**	
$\hat{\theta}_0$	−0.058	−0.055	(−0.093, −0.028)
$\hat{\theta}_1$	−0.060	−0.050	(−0.112, −0.016)
$\hat{\theta}_2$	0.930	0.936	(0.910, 0.949)
$\hat{\theta}_3$	−0.240	−0.229	(−0.316, −0.169)
$\hat{r}_1$	3.546	3.555	(3.518, 3.559)
$\hat{r}_2$	3.674	3.666	(3.660, 3.709)
$\hat{\kappa}^{(2)}$	3.630	3.479	(3.112, 4.312)
$\hat{\kappa}^{(3)}$	0.906	0.892	(0.807, 1.019)
Loglikelihood	−365.667	−364.634	
AIC	747.334	745.269	
BIC	801.647	799.581	

$\mathbf{Q}_\tau^{BASE}(\mathbf{y}_t) = (1 - 2 \cdot \mathbb{I}[\tau < 0.5])\sqrt{(\theta_0 + \theta_1 (Q_\tau(y_{t-i}))^2 + (\theta_2 + \theta_3 \mathbb{I}[y_{t-1} < 0]) y_{t-1}^2)}$			
Metric	**Estimates**		**95% Conf interval**
($\tau = 0.01$)	**Posterior mean**	**Posterior mode**	
$\hat{\theta}_0$	0.305	0.115	(0.136, 0.529)
$\hat{\theta}_1$	0.052	0.007	(0.004, 0.138)
$\hat{\theta}_2$	0.914	0.956	(0.877, 0.946)
$\hat{\theta}_3$	0.834	0.474	(0.500, 1.255)
$\hat{r}_1$	12.578	12.610	(12.396, 12.668)
$\hat{r}_2$	13.481	13.417	(13.401, 13.687)
$\hat{\kappa}^{(2)}$	3.619	3.956	(3.078, 4.269)
$\hat{\kappa}^{(3)}$	0.808	0.857	(0.722, 0.906)
Loglikelihood	−369.226	−367.421	
AIC	754.452	750.842	
BIC	808.764	805.154	

SET-GJR-IGARCH-CAViaR respectively, a calming state is reached and market risk is damped by a factor of c. 0.8 (as given by $\hat{\kappa}^{(3)}$). The explosive behaviour of market risk is therefore more pronounced for $\tau = 0.01$, i.e. for the more extreme return quantile, in comparison to 95% VaR ($\tau = 0.05$). Moreover, the explosive regime sets in faster, with the thresholds for entry into the regime being considerable lower than in the case of 95% VaR in the previous section.

In comparison to the results for 95% VaR, the estimates for the asymmetry parameter $\hat{\theta}_3$ in the case of $\tau = 0.01$ also suggest a more pronounced asymmetric response to lagged returns than in the case of 95% VaR in the previous section: For $\tau = 0.01$, i.e. for more extreme return quantiles, negative lagged returns increase market risk measured by VaR much more than for the milder 95% VaR level as indicated by a larger $\hat{\theta}_3$ (in absolute terms) in both SET-CAViaR models.

Whereas there is close resemblance in the parameters estimates between the two models as in the case of 95% VaR (again, as suggested by theory, $\hat{r}_1$ and $\hat{r}_2$ correspond strikingly closely across the two models via the square root), the AIC and BIC once more favour the SET-AS-CAViaR model. Moreover, in the case of SET-GJR-IGARCH-CAViaR with $\tau = 0.01$ in this section, there is a now a need to enforce a positivity parameter restriction on $\hat{\theta}_1$, i.e. $\hat{\theta}_1 > 0$, as without restriction this parameter estimate turns out to be negative, which is not feasible given the square root in the GJR-IGARCH-CAViaR base quantile process. Consequently, as can be seen in table 2.6, $\hat{\theta}_1$ is very close to 0 in the case of SET-GJR-IGARCH-CAViaR.

The Relation of the Base Model to GARCH and Stationarity Issues

As in the case of 95% VaR, the order of magnitude of the quantile base model parameter estimates $\hat{\theta}_i, i = \{0, 1, 2, 3\}$ is in the range that is to be expected if the underlying return generating model were a Gaussian SET-GARCH model such as the ones in (2.14) and (2.17). The reasoning here is the same as in the case of 95% VaR, with the difference that $Q^{\mathcal{N}}_{0.01}(\varepsilon) = -2.326$, the 1% quantile of the standard Gaussian would have to be used here instead of $Q^{\mathcal{N}}_{0.05}(\varepsilon) = -1.645$ in order to deduct the corresponding SET-GARCH model parameters. Close inspection reveals that in the case of 99% VaR the corresponding Gaussian GARCH base volatility processes would not be stationary given the posterior mean and mode quantile base model parameter estimates both for SET-AS-CAViaR as well as for SET-GJR-IGARCH-CAViaR. However, in conjunction with the multiplicative scale parameter estimate for the calming regime ($\hat{\kappa}^{(3)}$) being well below unity, the corresponding entire SET-GARCH setup would, however, still be geometrically ergodic in all cases (cf. section 2.4.3).

The Performance of the MCMC LTE Routine for 99% Value-at-Risk ($\tau = 0.01$)

The analysis of the average ACFs of the parameter candidate draws during the tuning phase in figure 2.10 mirrors the difficulty in the estimation of the SET-GJR-IGARCH-CAViaR model for the 99% VaR level:

In the case of SET-AS-CAViaR the stopping point in the MCMC LT estimation of the SET-AS-CAViaR model can be determined reasonably well as the ACFs start to flatten out and rise after c. 80 batch rounds and are indeed very flat on a high level for very a long tun-

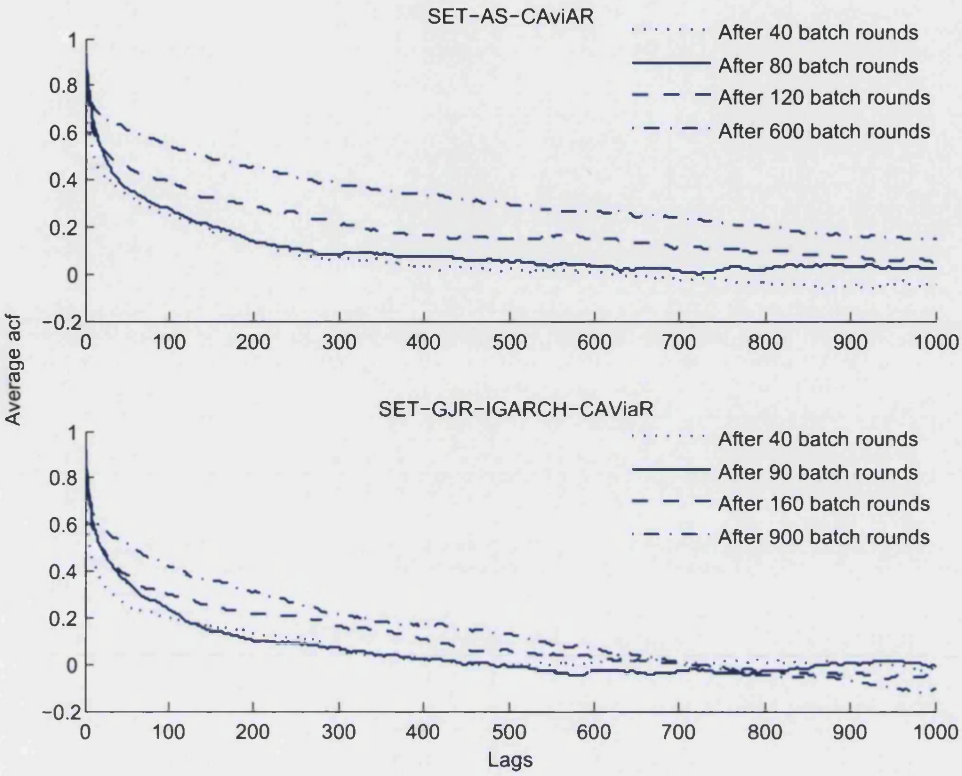

Figure 2.10: **Average autocorrelation functions (τ = 0.01)**

This figure displays the average autocorrelation functions (ACFs) for parameter candidate value in an MCMC LTE of the SET-AS-CAViaR and SET-GJR-IGARCH-CAViaR models with $\tau = 0.01$ and base quantile specifications given by (2.32) and (2.33) respectively against 1980-2005 CRSP IBM holding returns. Following tuning phases of different lengths, the average ACFs are computed for the first 1000 lags.

ing phase after 600 batch rounds. In the case of SET-GJR-IGARCH-CAViaR, the picture is not so clear-cut, rendering the choice of a good stopping point more difficult: After c. 90 batch rounds, the average ACF is by and large similar to the one after 40 batch rounds, albeit it seems more elevated and steeper for the first 100 lags. After 160 batch rounds of temperature cooling the ACFs is flatter and has risen slightly. Thereafter, albeit not shown, the average ACF does not move much and it is only after a very long tuning phase of 900 batch rounds

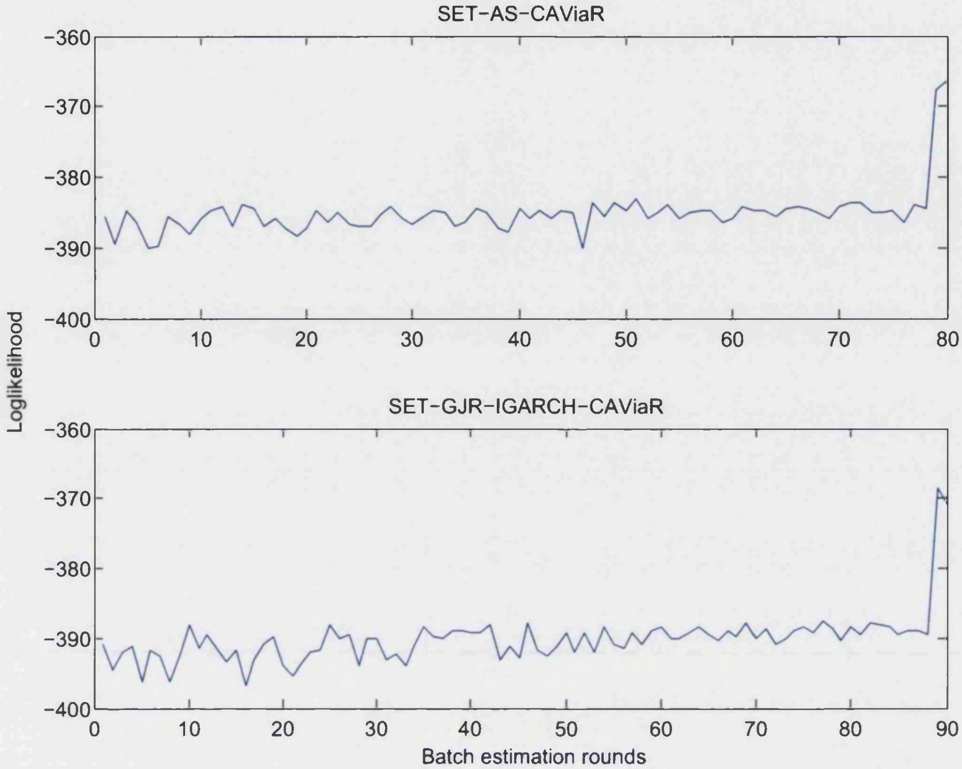

Figure 2.11: **SET-CAViaR likelihood evolution ($\tau = 0.01$)**

This figure displays the evolution of loglikelihood during the MCMC LTE of the SET-AS-CAViaR and SET-GJR-IGARCH-CAViaR models with $\tau = 0.01$ and base quantile specifications given by (2.32) and (2.33) respectively against 1980-2005 CRSP IBM holding returns. The loglikelihood is displayed after the completion of each batch estimation round during the tuning phase of the chain.

and a corresponding extremely low temperature that a stronger autocorrelation structure emerges. Still, after such a long phase of temperature cooling and chain tuning, the ACFs is not as flat and elevated as for SET-AS-CAViaR after 600 batch runs. Based on the above reasoning, 80 and 90 batch rounds are picked as stopping points for the tuning phase for SET-AS-CAViaR and SET-GJR-IGARCH-CAViaR respectively. In comparison to the results for $\tau = 0.05$, the analysis of the average ACFs in this section suggests that on the higher 99%

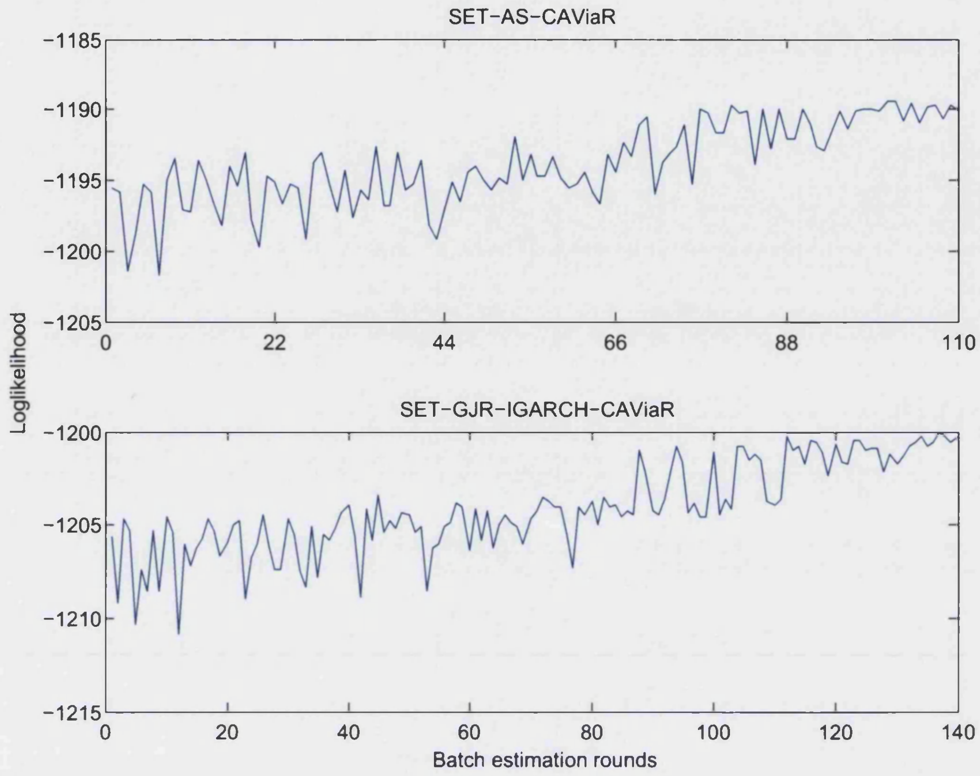

Figure 2.12: **SET-CAViaR likelihood evolution ($\tau = 0.05$)**

This figure displays the evolution of loglikelihood during the MCMC LTE of the SET-AS-CAViaR and SET-GJR-IGARCH-CAViaR models with $\tau = 0.05$ and base quantile specifications given by (2.32) and (2.33) respectively against 1980-2005 CRSP IBM holding returns. The loglikelihood is displayed after the completion of each batch estimation round during the tuning phase of the chain.

VaR level, the MCMC estimation of the SET-CAViaR models can be accomplished faster as the ACFs warrant stopping the tuning phase earlier for both models in the case of $\tau = 0.01$. A likely reason for this circumstance is the way quantile models are estimated with the objective function in (2.2), which involves finding a suitable quantile cutoff point in the empirical distribution of the data according to the piecewise loss function in (2.1): For more extreme quantiles (τ close to 0 or unity), i.e. the more one moves towards the ends of the data

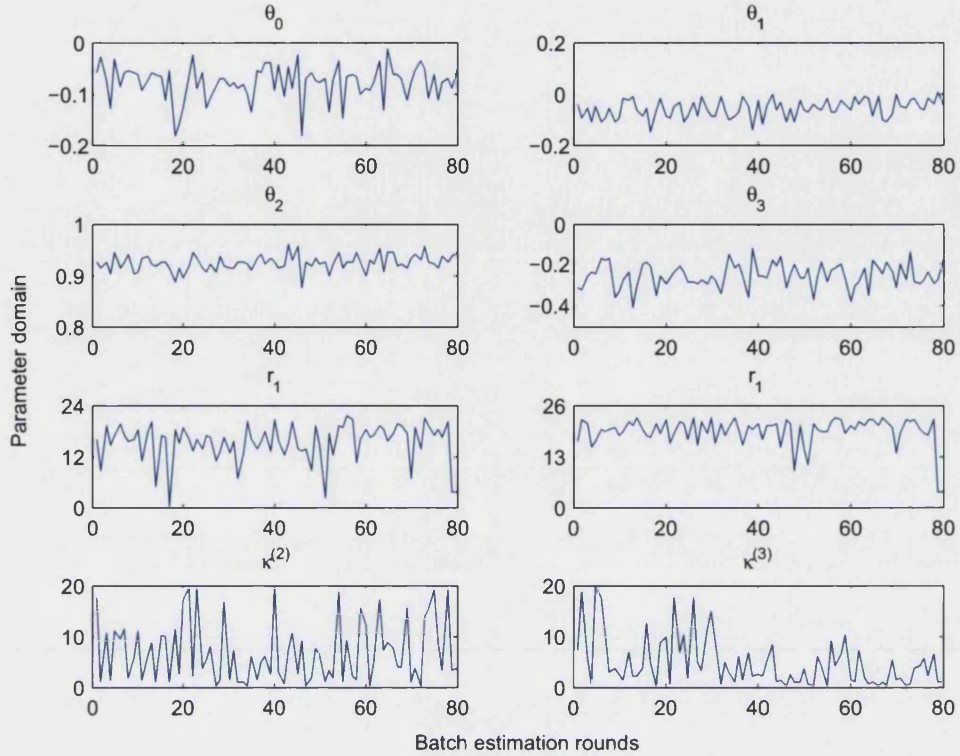

Figure 2.13: **SET-AS-CAViaR para's (τ = 0.01)**

This figure shows the evolution of the parameter candidate values in the MCMC LTE of the SET-AS-CAViaR model with $\tau = 0.01$ and a base quantile process given by (2.32) against 1980-2005 CRSP IBM holding returns. The parameter candidate value draws are plotted after each batch estimation round during the tuning phase of the Markov chain.

spectrum, there are less points to choose from which comply with the criterion, thus prompting a faster, albeit not more accurate, estimation procedure.

This reasoning also manifests itself in the evolution of the loglikelihood, shown for both SET-CAViaR models on both 95% and 99% VaR levels in figures 2.12 and 2.11: Due to datapoints becoming more and more scarce as well as more spread out towards the tails of the empirical data distribution, the estimation of the SET-CAViaR model with the more extreme quantile $\tau = 0.01$ compared to the milder

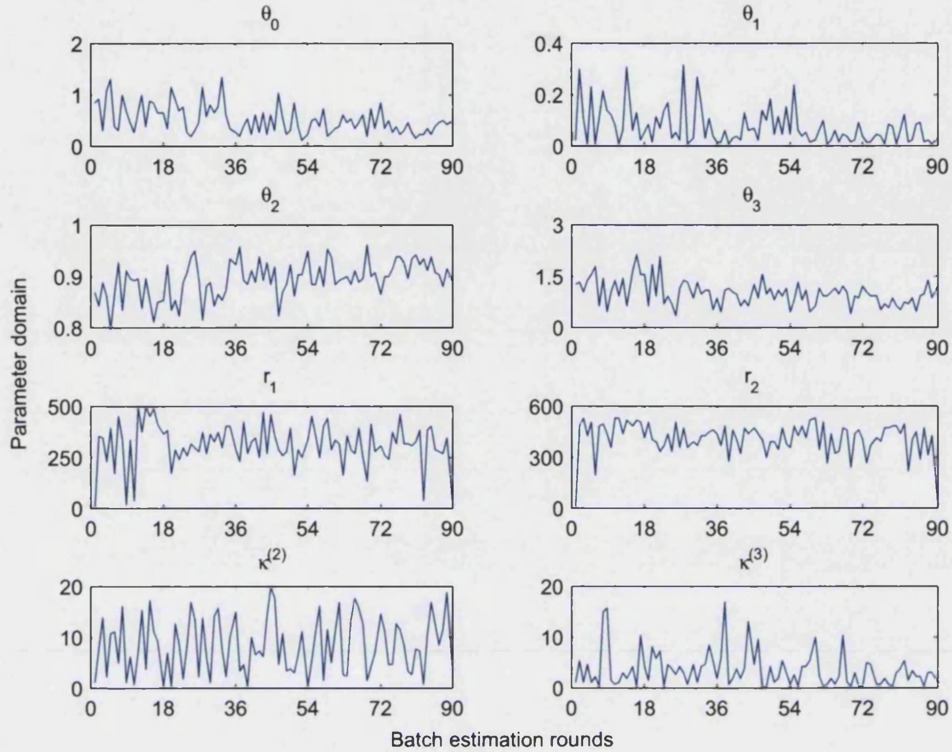

Figure 2.14: **SET-GJR-IGARCH-CAViaR para's ($\tau = 0.01$)**

This figure shows the evolution of the parameter candidate values in the MCMC LTE of the SET-GJR-IGARCH-CAViaR model with $\tau = 0.01$ and a base quantile process given by (2.33) against 1980-2005 CRSP IBM holding returns. The parameter candidate value draws are plotted after each batch estimation round during the tuning phase of the Markov chain.

$\tau = 0.05$ should result in a more 'bumpy' and rugged evolution of the likelihood as the algorithm has comparatively less data points to 'work with' in order to determine the quantile cutoff. Figure 2.11 supports this reasoning as one notes that the evolution of the loglikelihood of both SET-CAViaR models exhibits larger 'swings' as well as a marked incline towards the end of the tuning phase compared to the corresponding estimation with $\tau = 0.05$ in figure 2.12.

Again, as also suggested by the AIC and BIC, the MCMC LTE routine for the SET-AS-CAViaR model progresses towards to a slightly

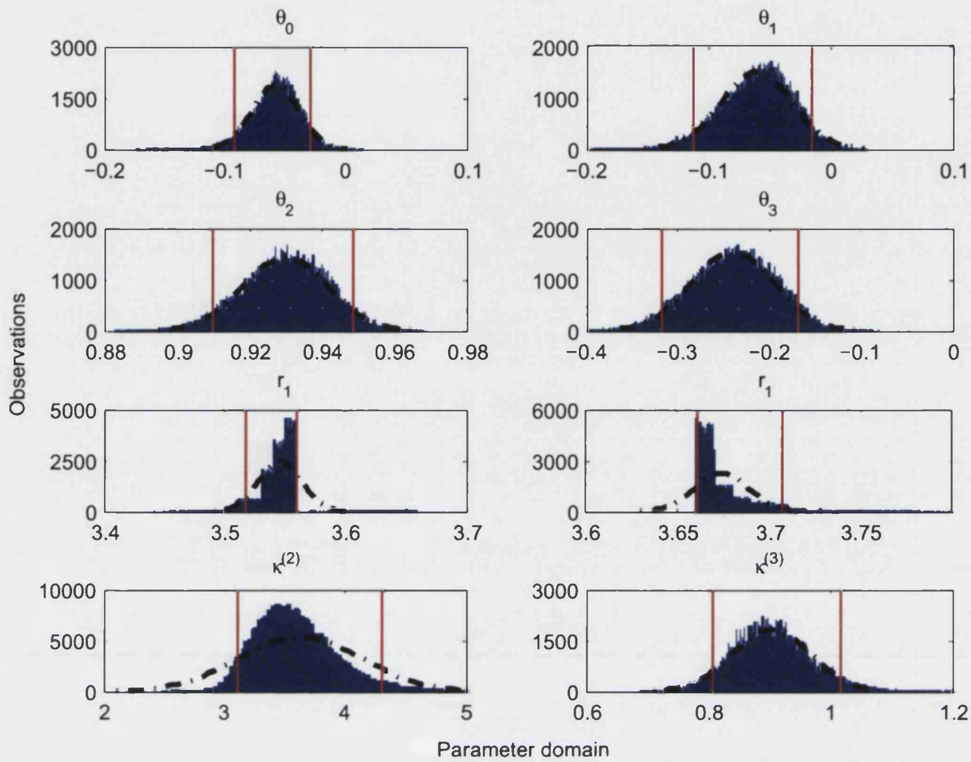

Figure 2.15: **SET-AS-CAViaR posteriors ($\tau = 0.01$)**

This figure displays the posterior distribution histograms for parameter candidate values of the MCMC LTE of an SET-AS-CAViaR model with $\tau = 0.01$ and a base quantile process given by (2.32) against 1980-2005 CRSP IBM holding returns. Also displayed as dash-dotted lines are fitted normal distributions with the same mean and variance as the respective parameter posterior. Additionally, the 95% confidence intervals bounds are shown as vertical lines.

higher loglikelihood level, indicating a better fit compared to SET-GJR-IGARCH-CAViaR. However, a comparison of loglikelihood levels with the results for $\tau = 0.05$, i.e. across τ, is *not* at apposite since different quantile levels τ translate into different tick-exponential density functions of different order altogether and therefore result in different likelihood functions that are not comparable to each other (cf. Komunjer, 2005).

The evolution of the parameter estimate draws for both SET-CAViaR models during the tuning phase displayed in figures 2.13 and 2.14 does

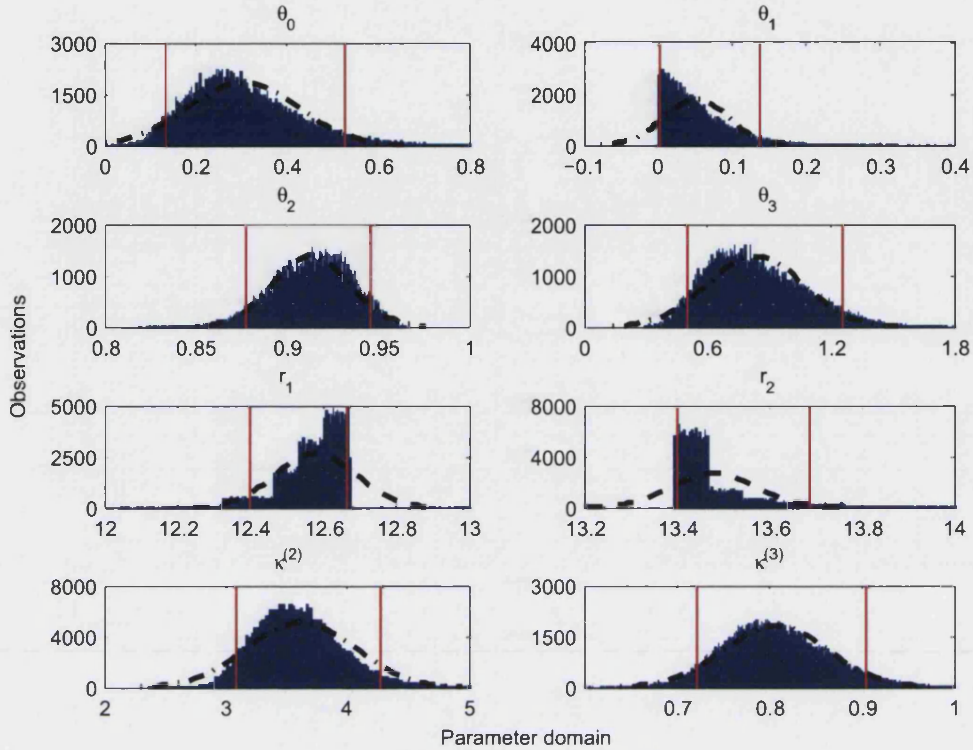

Figure 2.16: **SET-GJR-IGARCH-CAViaR posteriors ($\tau = 0.01$)**

This figure displays the posterior distribution histograms for parameter candidate values of the MCMC LTE of an SET-GJR-IGARCH-CAViaR model with $\tau = 0.01$ and a base quantile process given by (2.33) against 1980-2005 CRSP IBM holding returns. Also displayed as dash-dotted lines are fitted normal distributions with the same mean and variance as the respective parameter posterior. Additionally, the 95% confidence intervals bounds are shown as vertical lines.

not exhibit much controversy as the shape of the chain for the base quantile model parameters $\hat{\theta}_i, i = \{0, 1, 2, 3\}$ is again, as in the previous section, in line with other MCMC LTE applications. The evolution of the non-linearity parameters estimates $\hat{r}_1$, $\hat{r}_2$, $\hat{\kappa}^{(2)}$ and $\hat{\kappa}^{(3)}$ is similar to the 95% VaR case for both SET-CAViaR models, with large initial variation in the parameter draws, particularly in the case of $\hat{r}_1$ and $\hat{r}_2$ in SET-GJR-IGARCH-CAViaR. There is, however, not a long phase of subdued variation following the large initial swings as for $\tau = 0.05$. This is due to ending the tuning phase earlier in accordance with the

ACF criterion laid out above. Yet, if the tuning phase is prolonged (not shown here), the evolution of the parameters estimates $\hat{r}_1$, $\hat{r}_2$, $\hat{\kappa}^{(2)}$ and $\hat{\kappa}^{(3)}$ is similar in that there is subdued variation after c. 80 and 90 batch runs during the tuning phase of the chain for SET-AS-CAViaR and SET-GJR-IGARCH-CAViaR respectively.

This also manifests itself in figures 2.15 and 2.16 displaying the empirical posterior distributions for the model parameter estimates after the end of the tuning phase: Again, as in the case of $\tau = 0.05$, the posteriors for $\hat{r}_1$, $\hat{r}_2$, $\hat{\kappa}^{(2)}$ and $\hat{\kappa}^{(3)}$ are not standard, however, their variance is evidently not as large as one might gather from figures 2.13 and 2.14. Further, as for 95% VaR, the quantile base model parameter estimate posteriors for both SET-CAViaR models closely resemble Gaussian distributions, with one notable exception: In the case of SET-GJR-IGARCH-CAViaR, the posterior of $\hat{\theta}_1$ is truncated at 0, owing to the binding positivity constraint. The need to enforce this parameter restriction thus indicates a suboptimal fit of the SET-GJR-IGARCH-CAViaR model for $\tau = 0.01$ favouring the SET-AS-CAViaR model for the given dataset, albeit the other parameters estimates, in particular the threshold and scale parameters are in close correspondence across the two SET-CAViaR models.

2.7.6 Discussion of Results

Even though the reinforcing mechanisms in Daníelsson and Shin's (2003) concept of endogenous risk have received considerable theoretical attention (e.g. Morris and Shin, 2004), explosive dynamics have so far been largely ignored in conventional empirical risk modelling and in practical applications. Recent experience from the financial markets, however, suggests that this traditional way of risk modelling might fall short of reality and can be improved upon:

The Dow Jones Stoxx 600 Index for example saw a drop of 3% during stock market fall on February 27^{th}, 2007, the "steepest (...) percentage decline since May 29^{th}, 2003" amidst trading conditions in which selling of securities was "almost like (...) a cascade" according to *The Wall Street Journal Europe, edn. February 28^{th}, 2007,* p. 1. Regarding the risk management techniques in the face of such a dramatic and unforeseen decline, *The Economist, issue March 3^{rd}, 2007* comments on p. 83:

> "Investment banks use 'value-at-risk' models which mean that, when volatility rises, they cut the capital they allocate to trading. This usually means selling assets. So a sudden jump in volatility tends to generate further volatility (...) According to Goldman Sachs, the latest jump in Vix (a measure of stockmarket volatility) took it eight standard deviations from its average. If conventional models are correct, such an event should not have happened in the history of the known universe (...) Perhaps modellers do not know the universe as well as they think."

The empirical results obtained in this chapter represent an attempt at estimating a new empirical model of market risk that is able capture such explosive endogenous risk dynamics against real data - and, it might be argued, a successful one:

We obtain a good model fit and sensible parameter estimates that allow for a quantitative description of endogenous market risk. And while an obvious point of criticism of the results obtained with the our approach is overfitting, we would argue against such allegations.

For one, the SET-CAViaR models used, despite their 'unorthodox' structure and non-linear features, are fairly parsimoniously parameterised. Also, given its semi-parametric nature, SET-CAViaR excludes the possibility of data-fitting via an assumed parametrised error distribution. Rather, as an empirical approach, it 'lets the data speak for itself'. Moreover, when obtaining the parameter estimates, we attach great attention to not allowing the Markov chain to degenerate and thus over-fitting the parameters. Also, in (2.23) we assume a non-dogmatic, flat prior, which further limits the scope for overfitting. Lastly, a strong argument against overfitting comes from the parameter estimates themselves. Rather than just being outcomes of a fitting procedure without any direct meaning in terms of magnitude, as is for example the case in 'Principal Components Analysis' (PCA)-type factor models or 'Neural Networks', the estimated parameters for the SET-CAViaR models above actually 'make sense'. Not only do the estimated of the base line CAViaR process correspond to typical parameters levels in related well-known ARMA-GARCH-type models, the estimates of the whole model, including scales and thresholds are also in line with what the theory literature outlined in section 2.4.2 would predict[26]:

Once a certain trigger point ($\hat{r}_1$) is passed, risk dynamics become self-exciting, thus capturing the explosive reinforcing mechanics introduced by Morris and Shin (2004). More concretely, the results suggest that on the 95% and 99% VaR levels, market risk approximately doubles and quadruples respectively, albeit after different trigger points as indicated by the estimates of the explosive scale $\kappa^{(2)}$. Granted that

[26]It is also notable, that with the exception of SET-GJR-IGARCH-CAViaR for 99% VaR ($\tau = 0.01$), the estimates are obtained without having to enforce any parameter restrictions other than the ordering of the thresholds.

these are one-off estimation results obtained with a new model, and therefore have to be interpreted with caution, it can nonetheless be argued that a practitioner might want to watch the behaviour of returns beyond the estimated threshold more closely and should expect a scale increase in the level of risk similar to the estimates.

Further, given the estimated parameter setup for the two SET-CAViaR models on both levels of VaR, market risk (as measured by the returns quantile) is prevented from exploding all the way to (negative) infinity, despite the explosive dynamics beyond threshold $\hat{r}_1$: Once the second threshold $\hat{r}_2$ is passed, the SET-CAViaR models enter into into a calming regime in which market risk is suppressed by a the scale factor $\hat{\kappa}^{(3)}$ less than unity. This result is, again, in support of the theory laid out in section 2.4.2: Loosely speaking, the explosive build-up of endogenous risk as for example described in Daníelsson and Shin (2003) needs to be expected to stop once the market has experienced enough 'exuberance' and market risk ought to decline to more 'normal' levels according to the argument presented in Morris and Shin (2004, p. 14).

Also notable about the estimation results is that the estimated explosive scale $\hat{\kappa}^{(2)}$ increases for the more extreme quantile $\tau = 0.01$ (i.e. the higher 99% VaR level as opposed to 95% VaR), whereas the corresponding threshold for entry into the explosive regime ($\hat{r}_1$) decreases. The interpretation of this result has to be that on more extreme market risk levels endogenous risk builds up more fiercely and faster and also calms itself more quickly as indicated by the lower second threshold. Intuitively, this makes sense as, loosely speaking, one might argue that during catastrophic times things tend to get even more severely 'out of hand' and more quickly so than during a less poignant crisis. Qualitatively, the result is also in line with the 'self-exciting peaks-over-thresholds' (self-exciting POT) model presented in McNeil et al. (2005, p. 307): This approach to modelling very extreme levels of risk is able to capture dynamics, where, "in a period of excitement, both the temporal intensity of occurrence [of losses] and the magnitude (...) increase."

The relative difficulty of the estimation for $\tau = 0.01$ with the lack of stationarity for a corresponding Gaussian GARCH-type base volatility process and the need to enforce a positivity constraint on $\hat{\theta}_1$ in the case of SET-GJR-IGARCH-CAViaR also suggests that despite the intuitive appeal and the encouraging results above, SET-CAViaR models might not be the modelling choice for the analysis of extreme market risks (e.g. VaR levels of 99.9%). As also mentioned in the previous section, the quantile estimation objective function in (2.2) will produce less accurate results for more extreme quantiles, warranting the use of special modelling approaches such as EVT or the self-exciting POT model mentioned above for such levels of market risk[27].

However, even though the SET-CAViaR model class might not be the optimal choice for all levels of VaR, the above empirical findings suggest a wide range of applications: The methodology might be of relevance for risk management in financial institutions, particularly investment banks and funds with large trading operations, where it could be used to forecast daily trading VaR more accurately, or in stress-testing situations, in which a sophisticated benchmark model is needed. SET-CAViaR models can also be of help to trading practitioners when forecasting and judging the impact of trades to be carried

[27] A comprehensive overview modelling techniques for extreme levels of risk can be found in McNeil et al. (2005).

out with respect to the amount of risk incurred as well as for tailoring and calibrating trading models with regards to VaR limit setting. Particular interest should be directed towards the estimated scale and threshold parameters in order to be able to monitor market and risk dynamics at critical magnitudes of market risk.

In academic applications, SET-CAViaR models can be used to test theoretical hypotheses, e.g. about endogenous risk as well as for backtesting comparison studies involving alternative (time series) risk models, e.g. quantile models on the basis of ARMA-GARCH or Historical Simulation. Another obvious area of application is the identification and dating of (past) financial crises, for which there appears to be an imminent need as has been pointed out by a range of researchers in the empirical literature on the modelling and testing of financial contagion: Dungey et al. (2005, p. 67) for example state that

> "The choice of both crisis and non-crisis periods [for modelling and testing of financial contagion] is almost always ad hoc, although often the sample selection is based on ex-post rationalisations, making it difficult to compare studies, even those apparently conducted on the same crisis (...) Productive future work would be to find a more objective procedure for dating crises (...) based on data characteristics."

Lastly, it needs to be mentioned that the empirical results above were obtained with very specific, fairly parsimonious parameterisation of the SET-CAViaR model in order to facilitate estimation. Possible extensions or alternatives to the work above are the estimation of SET-CAViaR models with different quantile base processes, some of which are presented in Engle and Manganelli (2004). Moreover, the non-linear self-exciting setup can be modified into a richer model, albeit harder to estimate, by admitting more regimes or allowing regimes to be asymmetrical, e.g. by not centering the thresholds around 0 or having an uneven number of regimes. In a very extreme application, one might even think of having no scale parameters at all, but rather allowing all base quantile model parameters to differ across regimes, similarly to the work of Zhang et al. (2001) in the case of the 'Autoregressive Conditional Duration' (ACD) model.

2.8 Conclusion

The goal of this chapter is to establish an new empirical model of market risk that is capable of capturing explosive risk dynamics in financial markets, which Daníelsson and Shin (2003) have described as endogenous risk. The proposed setup is SET-CAViaR, a self-exciting process with endogenous thresholds and an autoregressive quantile base model, which due to its semiparametric and non-linear structure is estimated with a special variant of Chernozhukov and Hong's (2003) MCMC LTE method.

The empirical findings obtained from an estimation of SET-CAViaR models against a set of CRSP IBM holding returns show that the proposed model class is well capable of producing results that are very much in line with the concept of endogenous risk and the wider theoretical literature in this field, notably Morris and Shin (2004), in the sense that beyond a certain return threshold market risk (as measured by VaR) becomes explosive, but eventually calms itself again.

While the results in this chapter are a successful first attempt at empirically capturing endogenous risk in a tractable model, a few caveats are, however, apposite: The strength and flexibility of the semiparametric setup of the SET-CAViaR model comes at the price of a robust, but difficult to implement and lengthy MCMC LTE procedure. Future research, especially when directed towards the applicability of SET-CAViaR in practical situations, might therefore focus on the ease of estimation in general as well as estimation speed in particular. In a radical step, one might even sacrifice the semiparametric CAViaR process and the associated not continuously differentiable quantile regression objective function in (2.2) favour of a simpler GARCH model at the heart of the self-exciting setup, thereby facilitating more conventional estimation while still keeping the self-exciting, non-linear features.

Further, the results in this chapter were obtained with a very specific parametrisation of the SET-CAViaR model. Promising future research might therefore concentrate on estimating SET-CAViaR models comparatively with different quantile base processes or different parameterisations. Lastly, it is still unclear how the proposed SET-CAViaR model would fare in comparison to alternative, more established and commonly used models of market risk. Such studies appear worthwhile both from an academic as well as from a practitioner's point of view and are left for future research.

Bibliography

Andrews, D. W. (1993), 'Tests for parameter instability and structural change with an unknown changepoint', *Econometrica* **61**, 821–856.

Andrews, D. W., Lee, I. and Ploberger, W. (1996), 'Optimal change-point tests for normal linear regression', *Journal of Econometrics* **70**, 9–38.

Bai, J. (1994), 'Least squares estimation of a shift in linear processes', *Journal of Time Series Analysis* **15**, 453–472.

Bai, J. (1999), 'Likelihood ratio tests for multiple structural changes', *Journal of Econometrics* **91**, 299–323.

Bai, J., Lumsdaine, R. L. and Stock, J. H. (1998), 'Testing for and dating common breaks in multivariate time series', *Review of Economic Studies* **65**, 395–432.

Bai, J. and Perron, P. (1998), 'Estimating and testing linear models with multiple structural changes', *Econometrica* **66**(1), 47–78.

Bai, J. and Perron, P. (2003), 'Computation and analysis of multiple structural change models', *Journal of Applied Econometrics* **18**, 1–22.

Basel Committee on Banking Supervision (1996), Overview of the amendment to the Capital Accord to incorporate market risks, Report, Bank of International Settlements; available from `http://www.bis.org/bcbs`.

Bassett, G. and Koenker, R. (1978), 'Regression quantiles', *Econometrica* **46**, 33–49.

Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroscedasticity', *Journal of Econometrics* **31**, 307–327.

Brown, R. L., Durbin, J. and Evans, J. M. (1975), 'Techniques for testing the constancy of a regression relationship over time', *Journal of the Royal Statistical Society, Series B* **37**, 149–192.

Chernozhukov, V. and Hong, H. (2003), 'An MCMC approacch to classical estimation', *Journal of Econometrics* **115**, 293–346.

Chow, C. G. (1960), 'Test of equality between subsets of coefficients in two linear regression models', *Econometrica* **28**, 591–605.

Daníelsson, J. and Peñaranda, F. (2007), On the impact of fundamentals, liquidity and coordination on market stability. Discussion paper dp586, Financial Markets Group, London School of Economics; obtainable with password from `http://fmg.lse.ac.uk/pdfs/dp586.pdf`.

Daníelsson, J. and Shin, H. S. (2003), Endogenous risk, *in* P. Field, ed., 'Modern Risk Management', Risk Books, London, UK, pp. 297–316.

Ding, Z., Granger, C. W. J. and Engle, R. F. (1993), 'A long memory property of stock market returns and a new model', *Journal of Empirical Finance* **1**, 83–106.

Dungey, M., Fry, R., Gonzales-Hermosillo, B. and Martin, V. (2005), A comparison of alternative tests of contagion with applications, *in* M. Dungey and D. N. Tambakis, eds, 'Identifying International Financial Contagion', Oxford University Press, New York, chapter 3, pp. 60–85.

Engle, R. F. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the United Kingdom inflation', *Econometrica* **50**, 987–1007.

Engle, R. F. and Manganelli, S. (2004), 'CAViaR: Conditional Autoregressive Value-at-Risk by regression quantiles', *Journal of Business & Economic Statistics* **22**, 367–381.

Engle, R. F. and Ng, V. (1993), 'Measuring and testing the impact of news on volatility', *Journal of Finance* **48**, 1747–1778.

Ferguson, T. (1967), *Mathematical Statistics: A Decision Theoretic Approach*, Academic Press, New York.

Fox, M. and Rubin, H. (1964), 'Admissibility of quantile estimates of a single location parameter', *Annals of Mathematical Statistics* **35**, 1019–1030.

Gallant, A. R. and Tauchen, G. (2008), EMM: A program for efficient method of moments estimation, version 2.3, user's guide. Mimeo, Duke University, Fuqua School of Business; obtainable from `http://econ.duke.edu/webfiles/arg/emm/emm.tar`.

Ghysels, E., Guay, A. and Hall, A. (1997), 'Predictive tests for structural change of unknown break point', *Journal of Econometrics* **82**, 209–233.

Glosten, L. R., Jagannathan, R. and Runkle, D. (1993), 'On the relation between expected value and volatility of the nominal excess return on stocks', *Journal of Finance* **48**, 1779–1801.

Goldfeld, S. and Quandt, R. (1973), 'A Markov model for switching regression', *Journal of Econometrics* **1**, 3–16.

Hamilton, J. (1989), 'A new approach to the economic analysis of non-stationary time series', *Econometrica* **57**, 357–384.

Hamilton, J. D. and Susmel, R. (1994), 'Autoregressive conditional heteroscedasticity and changes in regime', *Journal of Econometrics* **64**, 307–333.

Hansen, B. E. (1992), 'Testing for parameter instability in regressions with I(1) processes', *Journal of Business and Economic Statistics* **10**, 321–336.

Hansen, B. E. (1997), 'Inference in TAR models', *Studies in Nonlinear Dynamics and Econometrics* **2**, 1–14.

Hansen, B. E. (2000), 'Sample splitting and threshold estimation', *Econometrica* **68**, 575–603.

Hansen, P. R. and Lunde, A. (2005), 'A forecast comparison of volatility models: Does anything beat a GARCH(1,1)?', *Journal of Applied Econometrics* **207**, 873–889.

Hastings, W. (1970), 'Monte Carlo sampling methods using Markov chains and their application', *Biometrika* **57**, 97–109.

Hentshel, L. (1995), 'All in the family: nesting symmetric and asymmetric GARCH models', *Journal of Financial Economics* **39**, 71–104.

Hjek, B. (1988), 'Cooling schedules for optimal annealing', *Mathematics of Operations Research* **13**, 311–329.

Hwang, C. (1980), 'Laplace's method revisited: weak convergence of probability measures', *Annals of Probability* **8**, 1177–1182.

Judd, K. L. (1998), *Numerical Methods in Economics*, MIT Press, Cambridge, MA.

Koenker, R. and Bassett, G. (1978), 'Regression quantiles', *Econometrica* **46**, 33–49.

Koenker, R. and Park, B. J. (1996), 'An interior point algorith for nonlinear quantile regression', *Journal of Econometrics* **71**, 265–283.

Komunjer, I. (2005), 'Quasi-maximum likelihood estimation for conid-tional quantiles', *Journal of Econometrics* **128**, 137–164.

Lee, O. and Shin, D. (2005), 'On stationarity and β-mixing properties of certain nonlinear GARCH(p,q models)', *Statistics and Probability Letters* **73**, 25–35.

McNeil, A. J., Frey, R. and Embrechts, P. (2005), *Quantitative Risk Management - Concepts, Techniques and Tools*, Princeton University Press, Princeton.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A. and Teller, E. (1953), 'Equations of state calculations by fast computing machines', *Chem. Phys.* **21**, 1087–1092.

Morris, S. and Shin, H. S. (1998), 'Unique equilibrium in a model of self-fulfilling currency attacks', *American Economic Review* **88**, 587–597.

Morris, S. and Shin, H. S. (1999), 'Risk management with interdependent choice', *Oxford Review of Economic Policy* **15**, 52–62.

Morris, S. and Shin, H. S. (2004), 'Liquidity black holes', *Review of Finance* **8**, 1–18.

Müller, P. (1991), A generic approach to posterior integration and Bayesian sampling. Technical Report 91-09, Statistics Department, Purdue University.

Nelson, D. B. (1991), 'Conditional heteroskedasticity in asset returns: a new approach', *Econometrica* **59**, 347–370.

Nelson, D. B. and Cao, C. Q. (1992), 'Inequality constraints in the univariate GARCH model', *Journal of Business & Economics Statistics* **10**, 229–235.

O'Hara, M. (1995), *Market Microstructure Theory*, Blackwell, Cambridge, MA.

Pincus, M. (1968), 'A closed form solution to certain programming problems', *Operations Research* **18**, 1225–1228.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P. (1992), *Numerical recices in C: The art of scientific computing. Second Edition*, Cambridge University Press, New York.

Price, K. V. and Storn, R. (1997), 'Differential evolution - a simple and efficient heuristic adaptive scheme for global optimization over continuous spaces', *Journal of Global Optimization* **11**, 341–359.

Quandt, R. E. (1960), 'Tests of the hypothesis that a linear regression system obeys two seperate regimes', *Journal of the American Statistical Association* **55**, 320–330.

Robert, C. P. and Casella, G. (2004), *Monte Carlo Statistical Methods, 2nd ed.*, Springer-Verlag, New York.

Roberts, G., Gelman, A. and Gilks, W. (1997), 'Weak convergence and optimal scaling of random walk Metropolis algorithms', *Annals of Applied Probability* **7**, 110–120.

Schwert, C. W. (1998), 'Why does stock market volatility change over time?', *Journal of Finance* **44**, 1115–1153.

Taylor, S. J. (1986), *Modelling Financial Time Series*, John Wiley and Sons, Chichester.

Tierney, L. and Kadane, J. (1986), 'Accurate approximations for posterior moments and marginal densities', *Journal of the American Statistical Association* **81**, 82–86.

Tierney, L., Kass, R. and Kadane, J. (1989), 'Fully exponential Laplace approximations to expectations and variances of non-positive functions', *Journal of the American Statistical Association* **84**, 710–716.

Tong, H. (1983), *Threshold models in nonlinear time series analysis*, Springer-Verlag, New York.

Tong, H. (1990), *Nonlinear time series: a dynamical systems approach*, Oxford University Press, Oxford.

Zakoan, J.-M. (1994), 'Threshold heteroskedastic models', *Journal of Economic Dynamics and Control* **18**, 931–955.

Zhang, M. Y., Russell, J. R. and Tsay, R. S. (2001), 'A nonlinear autoregressive conditional duration model with applications to financial transaction data', *Journal of Econometrics* **104**, 179–207.

Chapter 3

CARL: An Empirical Conditional Autoregressive Model of Market Liquidity

3.1 Introduction

Few market participants would deny the importance of liquidity for the financial system and trading activities. Yet, liquidity seems quite an elusive concept that is difficult to define, let alone quantify. In its biannual *Financial Stability Report* as of April 26^{th}, 2007, the Bank of England proposes a categorisation into funding and market liquidity: The former can be understood as the ease with which a firm can "meet its cashflow needs", for the latter the Bank provides the equally broad, common-sense definition as the ease of buying and selling financial assets in the market place.

This mirrors a definition by Brunnermeier and Pedersen (2008) who provide a theoretical model of the linkages between the two kinds of liquidity and their dynamics over time. Yet, while theoretical research activity into the *modelling* of funding as well as market liquidity has seen more activity over the recent years, empirical research seems to have taken a different route: Quite a number of papers (e.g. Amihud, 2002; Chordia et al., 2000*a*,*b*, 2001) have aimed at empirically exploring and explaining the *properties* of liquidity and its linkages across securities. Only comparatively little activity however appears to have been devoted to the explicit empirical modelling of the dynamics of market liquidity since the approaches by Engle and Russell (1998) or Engle and Lange (2001).

In this chapter, we aim to continue this line of research and focus on the empirical modelling of market liquidity over time while explicitly taking account of the fact that market liquidity is a concept characterised by multiple facets and notions that cannot be captured entirely by simple attributes such as execution time or spreads. We thus base our approach on a more complex measure of liquidity, the Hiu-Heubel (HH) liquidity ratio: This composite liquidity proxy (cf. Sarr and Lybek, 2002) consists of the percentage difference between the maximum and minimum price divided by the price-weighted turnover over a certain measurement period. Contrary to more 'traditional' measures of market liquidity such as durations, bid-ask spreads or transaction volumes that highlight liquidity along a single dimension (time, price, volume), the HH ratio - being a ratio of such measures - can be viewed as a multi-dimensional proxy for market liquidity: It is for example useful for market participants that attach great importance to the being able to sell and buy at almost the same prices, thus being interested in the *tightness* notion of liquidity (captured through the numerator) as well as equally to others that may favour a *broad*, *deep* and *resilient* market with an abundance of orders on either side, in which large scale transactions can be made with only a minimal impact on prices and shocks to prices away from fundamentals get corrected quickly (as measured by the ratio as a whole)[1].

Yet, despite its versatility as a liquidity measure the HH ratio has so far been conceived for use in a static context, e.g. say the monthly period measurement of liquidity in market. In this chapter, we seek to progress down a new route by exploiting the properties of the HH ratio as a multi-dimensional measure of market liquidity and including it in a dynamic, autoregressive setting. We do so by embedding the liquidity measure into the established time series concept of the 'Autoregressive Conditional Duration' (ACD) model by Engle and Russell (1998): The resulting model, which we dub 'Conditional Autoregressive Liquidity' (CARL) model uses the HH ratio instead of the duration as a liquidity variable in a multiplicative autoregressive setup.

By construction it is designed to pick up autoregressive, self-reinforcing patterns in market (il)liquidity as they are predicted by theoretical research e.g. by Morris and Shin (2004) (liquidity 'black holes') and Brunnermeier and Pedersen (2008) ('margin' and 'error spirals' of illiquidity). Yet, our approach not only builds on theoretical findings, but also seems well warranted from an empirical practitioner's point of view as self-reinforcing patterns in liquidity have recently received strong attention in this arena. *The Economist, edn. April 28th, 2007* for example finds that:

> "Liquidity is a self-reinforcing process; investors are more willing to buy an asset they know they can sell easily. But if liquidity suddenly dries up, some investors might end up owning assets they neither want nor can get rid of. This

[1]In its *2007 Financial Stability Report (edn April 2007)* the Bank of England also advocates the use of a return to volume-type ratio in order to capture market depth and resiliency.

might make a virtuous circle turn vicious."

As we show in this chapter, the CARL model can be viewed as a simple empirical reduced form modelling technique for such patterns and might therefore be of interest to a range of practical applications in the dynamic analysis of market liquidity. Moreover, it is straightforward to use and comprehend as, by construction, it shares most econometric properties with the ACD model and might thus look familiar to many users.

In this chapter we exemplify the versatility of the CARL model in an empirical study using data on Amazon stock from 1997 to 2006. We estimate the model via 'Quasi Maximum Likelihood' (QML) against data sampled on both daily and weekly frequencies and show how to obtain the best model fit. We also examine the in-sample forecasting properties of the chosen models via a Mincer-Zarnowitz-type test.

In general, our results are very encouraging and show that a good model fit can be obtained, even when resorting to QML estimation: The CARL model is able to pick up the autocorrelation structure in the data on both daily as well as weekly frequencies very well. We also detect evidence for threshold effects using only lagged negative returns, that are in accordance with theoretical research on market liquidity. Moreover, on both frequencies, the model appears to be able to forecast well, which ought to be a desirable property many practitioners.

The rest of the chapter now proceeds as follows: The following section provides a comprehensive overview of the recent theoretical as well as empirical research in financial liquidity. The section is intended as a stand-alone part and can also be read on its own to obtain a grounding in the general background on liquidity and its academic treatment. Section 3.3 introduces the two building blocks of the CARL model mentioned above - the HH ratio and its properties as well as the important dynamic framework of the ACD model. We present the 'assembled' CARL model and its features in section 3.4. The empirical application including a description of the data as well as discussion of the results is contained in section 3.5. Section 3.6 concludes.

3.2 Liquidity in Financial Markets

Generally speaking, liquidity and the associated liquidity risk are very complex concepts that are neither clearly defined nor accurately measurable. In its *Financial Stability Report 2006* for example the Bank of England states that (cf. Bank of England, 2006, p. 52)

> "...understanding, modelling and hence pricing liquidity risk is more difficult - and as a result less advanced - than for, say, market and credit risk because of the complexity and unpredictability of the interactions which may arise."

Yet, despite the ambiguity surrounding the concept, most market participants would acknowledge the important role liquidity plays in the in the financial markets: For example, there seems to be general consensus that market nowadays are much more 'liquid' than in the past. Similarly, liquidity risk, despite not yet precisely defined or formally quantified, is often cited as a major threat to financial markets (cf. e.g. Bank of England, 2007, p. 7).

In an effort to conceptualise liquidity (risk) for practitioners, regulators and academics, the classification of liquidity into market and funding liquidity has recently become more prominent: Brunnermeier and Pedersen (2008) for example describe the latter as the "ease with which a trader can obtain funding" for her operations, the former as the ease of trading an asset. In its *Financial Stability Report 2007*, the Bank of England recognises two types of risk associated with the above classifications (cf. Bank of England, 2007, p. 18): Funding liquidity risk describes the risk of market participants not being able to meet their cash-flow needs, market liquidity risk occurs if a position in an financial market asset cannot easily be offset or eliminated without significantly affecting the price.

Funding liquidity has recently received increased attention from a regulatory point of view: Both the Bank for International Settlements (BIS) as well as the Financial Services Authority (FSA) in the UK have proposed (preliminary) frameworks for the 'management of liquidity risk', seemingly addressing funding liquidity risk (s. e.g. Basel Committee on Banking Supervision, 2008; Financial Services Authority, 2003). In the light of the events in summer 2007 and throughout 2007/2008, when problems in the sub-prime credit markets lead to a subsequent fall-out in the intra-bank lending market, in which liquidity (in the funding sense) almost completely 'dried up', leading (inter alia) to the bank run on the UK mortgage provider Northern Rock, the demise of the investment bank Bear Stearns and funding problems at the US mortgage banks Freddie Mac and Fannie Mae, such a regulatory effort appears apposite. Arguably, funding liquidity problems can lead to the collapse of financial institutions and therefore contribute to systemic risk. The above-mentioned regulatory frameworks and the recent experience in the financial markets suggest however that the impact of funding liquidity risk is very much dependent on firms' capital provision, business models[2] and as well as network effects in the financial system. This also means that funding liquidity and the associated risk affect individual market participants differently and renders a more objective empirical treatment difficult.

This chapter therefore focuses on the modelling and analysis of market liquidity - also because this form of liquidity has a more ready interpretation and meaning to wider audiences. As mentioned above, market liquidity can generally be described as the ease with which an asset can be bought and sold. More concretely, market liquidity is commonly associated with the bid-ask spread, i.e. the cost of having to buy at the higher ask price compared to the lower bid price when selling. Research, however, suggests that this notion does not capture all dimensions of market liquidity.

Kyle (1985) for example identifies *tightness*, *depth* and *resiliency* as three important characteristics of liquid markets. Sarr and Lybek (2002) adopt and further refine this categorisation to the following:

1. *Tightness* - A tight market is characterised by low transaction costs when buying and selling assets. A narrow bid-ask spread indicates a tight and thus liquid market along this dimension.

2. *Immediacy* - This notion captures the speed at which transactions can be executed in markets and represents the efficiency of trading, transaction clearing and settlement.

3. *Depth* - A deep market is characterised by an abundance of buy and/or sell orders at prices narrowly below and above the current price at which the security trades.

[2]Northern Rock's business model of obtaining funding for its mortgage business through the capital markets as opposed to the traditional way of customer deposits has for example been described as particulary aggressive and vulnerable to funding liquidity problems

4. *Breadth* - This notion is closely linked to depth, but not the same. Breadth refers to orders being large in volume with minimal impact on prices.

5. *Resiliency* - Resilient markets are characterised by fast corrections of market imbalances, which divert transaction prices from those warranted by fundamentals.

Considerable research has been devoted to the different dimensions of market liquidity (cf. also Kyle, 1985, p. 1130) suggesting that neither of the above notions captures the concept of liquidity single-handedly (cf. e.g. Baker, 1996, p. 1). Moreover, the distinction between these dimensions is not clear-cut. Rather, they are overlapping to some extent: Depth and breath for example can compensate for each other. A deep market, with an abundance of small orders can mimic a broad market with few, but large orders close to the current trading price.

In addition, the importance of the different attributes of market liquidity tends to shift over time: In periods of turmoil more importance might be put on resilient markets in which divergence from fundamentals is corrected quickly and orders are carried out fast and efficiently in the immediacy sense. Market liquidity during such times might therefore be primarily judged along these dimensions. Equally, during calm times, it is conceivable that markets place more importance on low transaction costs, thereby viewing liquidity mainly in the sense of tightness.

The multi-dimensionality of market liquidity means that there is no single precise, (theoretically) correct way to define and describe it. While this is obviously a complicating circumstance for precise analysis, it has helped foster a variety of research approaches into the subject. We now present a brief overview of some of the theoretical and empirical treatments of market liquidity in the financial economics literature as they are relevant for this chapter.

3.2.1 The Theory Context

Many theoretical advances into explaining and modelling liquidity have been proposed over the years[3]. Early approaches are to found in the banking literature and comprise (inter alia) the well-known bank run model by Diamond and Dybvig (1983), in which the demand for liquidity of individually rational but 'impatient' bank depositors can lead to the collective 'irrational' equilibrium of a run and thus to a funding liquidity crisis for the bank having to pay out deposits.

While their paper may be more relevant for explaining funding liquidity risk, the mechanism at work has received considerable attention from various angles in the financial economics literature and, notably, has been employed as the basis for an approach aimed at explaining so-called 'liquidity black holes' by Morris and Shin (2004). Building on their findings in the area of higher order beliefs and 'coordination effects' in 'global games' (cf. Morris and Shin, 1998), they propose a model in which the market interaction of traders with privately known loss limits leads to said coordination effects, resulting in an equilibrium of sharp price drops once certain return (loss) 'trigger' thresholds are passed. Importantly, in their model, such a liquidity black hole, in which the sale of securities is very difficult, comes into existence

[3]The overview of theoretical literature on liquidity in this section provides is by no means exhaustive. We have for example left out the entire area of the effect of liquidity on asset pricing, which has seen which has been a growing field over the recent years. A good overview on this subject has been provided by Amihud et al. (2006).

without a fundamental shock to the value of asset. Consequently, following the liquidity black hole, the price eventually rebounds, aligning it with the fundamental value, and market liquidity in the asset returns to normal levels.

By using a third generation approach to risk/financial crises modelling, i.e. introducing a private knowledge information structure compared to common knowledge as for example in the second generation currency crises models (cf. e.g. Obstfeld, 1986), Morris and Shin (2004) are thus able to describe market (il)liquidity solely as the outcome of the strategic interplay between otherwise identically informed traders in the market.

In this respect, their model contrasts with earlier market microstructure models that explain the sources of illiquidity using an asymmetric information argument: Copeland and Galai (1983) as well as Glosten and Milgrom (1985) for example assume the existence a market maker facing both informed trade counterparties as well as noise/liquidity traders without being able to assess their type. The former trade based on private information about the fundamental value of the asset whereas the latter are less well informed and do not trade strategically. As already pointed out by Bagehot (1971), the outcome of the trading game under such an asymmetric information structure, is then that the market maker (usually assumed to be competitive and risk neutral) loses money on trades with informed traders and gains when trading with liquidity traders. In both Copeland and Galai (1983) as well as in Glosten and Milgrom (1985), albeit through different mechanisms, this gives rise to illiquidity in the form of a bid-ask spread as a compensation for the market maker's losses when trading with an informed counterparty. In Copeland and Galai (1983), the spread and thus illiquidity rises the higher the volatility of the asset, concurring with empirical evidence, whereas in Glosten and Milgrom (1985) a higher proportion of informed traders increases the spread.

Kyle (1985) assumes a similar asymmetric information structure, however, in his setting, illiquidity is more related to the depth and breadth notions: In his model, the transaction price depends on the order flow submitted by both informed and noise traders. The impact of the combined order flow on prices is lower the lower asset volatility and the higher the variance of the noise traders' orders. The market is thus broader and deeper the less uncertainty there is about the fundamental value of the asset and the more noise trading activity occurs.

Another important source of illiquidity has been documented to be the need for market makers to bear inventory in order to match buying and selling activity: As not all buying and selling counterparties tend to be present in markets at all times, market makers provide immediacy by being continuously facilitating trades from any direction and acting as an intermediate party. The market maker for example buys from a seller and later, when a suitable buying counterparty is present, sells the asset on. In the meantime, the asset is taken on as inventory, during which time the market maker faces the risk of adverse fundamental price changes. As pointed out by Stoll (1978), the market maker must be compensated for such risk, which Amihud and Mendelson (1980) and Ho and Stoll (1981, 1983) model as the market maker quoting bid-ask prices that depend on the inventory of the traded asset taken on. A wider bid-ask spread would thus correspond to more acute inventory risk and a more illiquid market.

Grossman and Miller (1988) also consider a market in which market makers face inventory risk. However, they point out that the bid-ask spread is likely to be an imperfect measure of the cost of trading in a dynamic environment in which traders should be more concerned with the change of quotes over time rather than the instantaneous spread. In their model, liquidity thus corresponds more to the immediacy notion mentioned above: They propose a dynamic trade setting with competitive market makers, in which traders' demand pressure for immediate trade execution and expectations about future transaction prices determine the inventory taken on by the market makers. In equilibrium, the amount of immediacy, sc. liquidity, provided in their model is an increasing function of the number of (competitive) market makers.

The notion of illiquidity stemming from inventory risk arising through the mismatch of buying and selling activity borders on another source of illiquidity risk, which is mostly prevalent in 'over-the-counter' (OTC) markets: In such markets, in which trade counterparties are sometimes scarce and bilateral trading is possible (i.e. traders can act as market makers and gain market power), illiquidity may also arise as a consequence of bargaining and search problems. Finding a suitable counterparty to trade with and negotiating prices in this setting causes frictions and leads to more illiquid trading activity, as documented inter alia by Duffie et al. (2005, 2006), Weill (2008) and Vayanos and Wang (2007).

Yet, while considerable effort in the theoretical literature has been directed at identifying causes of illiquidity in markets, less theoretical work has been devoted to modelling the dynamic properties of liquidity, which is however relevant for the purpose of this chapter: Notable results in this area can be found in the aforementioned paper by Morris and Shin (2004), who despite their focus being on the mechanisms triggering liquidity black holes, also shed light on the evolvement of liquidity through time. Their model indicates that situations of low liquidity tend to aggravate before returning to normal, which can be interpreted as 'illiquidity clustering', similar to the well-known stylised fact about volatility in markets.

The link between liquidity and volatility is also examined by Vayanos (2004) who considers a general dynamic equilibrium model with stochastic volatility and transaction costs. In his model, asset managers have a time-varying preference for liquid assets, which is correlated with volatility, resulting in a so-called 'flight to liquidity' during turbulent times.

This result is mirrored in Brunnermeier and Pedersen (2008), who add explicit structure to the dynamics of illiquidity and its link to volatility. They propose a setting in which market illiquidity is explicitly modelled as the price deviation from fundamental value and where 'liquidity spirals' can occur along the following lines: Low market liquidity increases the price impact of orders, leading to increased volatility in markets. Higher volatility then prompts higher margins for trades, leading to a worsened funding situation for traders, i.e. less funding liquidity. This 'margin spiral' in turn forces trades to be become more cautious so that they are more reluctant to take on positions, thereby further withdrawing market liquidity. Another feedback mechanism at work in their model is the so-called 'loss spiral', where low market liquidity and the associated price swings lead to losses on traders' existing positions and therefore to funding problems. This in turn, again, feeds into less trading and deteriorating market liquidity. Through these mechanisms, Brunnermeier and Pedersen (2008)provide a link of liquidity to volatility and are able to explain why market liquidity can suddenly dry up (similar to the situation of liquidity black holes above). For the purpose of devising an empirical time series model as in this chapter, their approach is especially interesting in that it highlights that illiquidity is self-reinforcing and therefore autocorrelated through time.

3.2.2 The Empirical Context

Empirical research into liquidity has gained in popularity over the last decade and is now a very active field. As a comprehensive overview is outside the scope of this chapter, in this section we again aim to document an excerpt of the literature as we think it is relevant for our purposes: For example, we narrow our focus down to empirical research on liquidity in equity markets and seek to only briefly touch on the arguably very important area of 'liquidity asset pricing'[4]:

An early empirical study of liquidity in markets has been provided by Amihud and Mendelson (1986), who analyse stock returns and bid-ask spreads of NYSE and AMEX stocks from 1960 to 1980 and find that the expected asset return is an increasing function of average 'categorised' illiquidity costs, i.e. the average bid-ask spread of one of seven portfolios sorted by their spread size. They also find evidence for a clientele effect, whereby in equilibrium less liquid assets are allocated to longer horizon investors, thus creating an increasing, but concave relationship between expected returns and bid-ask spreads.

A similar result is obtained by Brennan et al. (1998), who examine 'Center for Research in Security Prices' (CRSP) stock data from 1966 to 1995 and document that a stock's dollar trading volume as a

[4]For an extensive overview of this topic we again refer the reader to Amihud et al. (2006).

measure of liquidity acts as a determinant of expected returns. This conclusion is based on their finding that the trading volume is a statistically significant liquidity factor in an 'Arbitrage Pricing Theory' (APT) model along the lines of Ross (1976), even after the inclusion of other controlling characteristics such as for example size or the Connor and Korajczyk (1988) principal components.

In contrast to these two papers which address one dimension of liquidity explicitly, Amihud (2002) devises a composite liquidity measure, incorporating simultaneously the notions of tightness, depth and breath: He proposes the ILLIQ metric, the average daily ratio of absolute returns to dollar trading volume over some period (in his case a year) and finds that the lagged market illiquidity (i.e. the average ILLIQ measure across stocks) is a positive determinant of excess returns in a size-sorted sample of CRSP stock data from 1964 to 1996. While pointing out that ILLIQ is a robust measure of liquidity that can be easily derived for most securities from commonly used databases, Amihud (2002) also acknowledges that finer measures of liquidity are obtainable from intraday market microstructure data.

However, as Hasbrouck (2006) points out, such data is often not readily available[5], prompting the need for alternative methods: Hasbrouck (2006) proposes a Gibbs-filtered estimate of liquidity as a means to determine reliable and robust liquidity measures from low-frequency (daily, weekly or monthly) data. Using CRSP data from 1993-2005, he finds that his measure of liquidity is highly correlated with effective spreads (the difference between the transaction price and the mid-point of the bid-ask spread) calculated from intraday 'Trades and Quotes' TAQ data covering the same period. While he does not

[5]For example, high-frequency data is usually not obtainable prior to 1983.

find any evidence for a significant market-wide liquidity risk factor, the Gibbs-filtered measure per stock is documented to be a positive significant determinant in a Fama and French (1992)-style factor model against the full cross-section sample of CRSP equity returns from 1926 to 2006. This result is, however, not robust once the sample is split in January and non-January periods.

While the above papers relate levels of liquidity to stock returns, Pastor and Stambaugh (2003) as well as Acharya and Pedersen (2005) attempt to incorporate liquidity *risk* into a formal asset pricing approach: Using an asset pricing model based on Fama and French (1993), the former for example find that expected stock returns in a CRSP dataset covering 1966 to 1999 are related to the sensitivities of stock returns to innovations in aggregate illiquidity. However, their liquidity measure is the outcome of an econometric procedure that measures liquidity as the strength of volume-related return reversals and is as such not readily calculated from any given dataset.

Acharya and Pedersen (2005) follow a more familiar path of asset pricing, whereby they enhance the 'Capital Asset Pricing Model' (CAPM) with betas reflecting the pair-wise covariances between market liquidity, individual stock liquidity, market expected return and individual stock expected return. Using a liquidity measure based on Amihud's (2002) ILLIQ metric mentioned above, they find that in the cross-section of NYSE and AMEX stock returns from 1963 to 1999, their liquidity-adjusted CAPM fares better than the standard CAPM in terms of fit and in specification testing, even though both models equally use one degree of freedom.

The approaches of Pastor and Stambaugh (2003) as well as Acharya and Pedersen (2005) can thus both be viewed as the introduction of a systematic liquidity risk component into traditional asset pricing, whereby expected returns reflect exposure to this type of risk. As such, these models have been helped greatly by another recent important strain of empirical research into the so-called 'commonality' of liquidity:

Hasbrouck and Seppi (2001) for example study intraday order flows, returns and liquidity (as measured by bid-ask spreads, quote sizes and a combination of the two) in the cross-section of 30 'Dow Jones Industrial Average' (DJIA) stocks in 1994. While they find that common factors in order flow help explain returns, the evidence for common liquidity factors in their sample is weak. Rather, firm-specific effects seem to dominate common factors in explaining returns.

Using CSRP and intraday TAQ data for all NYSE stocks in 1996, Huberman and Halka (2001) on the other hand are able to identify a systematic time varying component in liquidity as measured by (percentage) bid-ask spreads and (dollar) quote depths. They also find that liquidity is negatively correlated with volatility, which is in support of the theoretical predictions above. While their results are robust to the inclusion of firm-specific as well as economy-wide control variables, such as e.g. trading volume or treasury yields, Huberman and Halka (2001) find it difficult to explain the common component in liquidity fully and acknowledge that there is "no established theory of time-series behaviour of liquidity proxies" (cf. p. 170). They attribute the commonality in liquidity behaviour over time to a systematic liquidity component that is the result of "the presence and effect of noise traders".

This result is also mirrored by Chordia et al. (2000*a*,*b*) who find co-movement in liquidity for NYSE stocks with market and industry-wide liquidity using intraday transaction data for 1992 and different measures of liquidity such as the (proportional) quoted spread, market depth, as well as the (proportional) effective spread. In robustness checks, the common effects remain significant after controlling for other stock-specific liquidity proxies such as volatility and volume. Nonetheless, they also struggle to pin down the source of commonality in their data. Contrary toHuberman and Halka (2001), who attribute the commonality in liquidity to noise trading activities across securities, Chordia et al. (2000*a*,*b*) suggest market-wide asymmetric information effects as well as inventory risks as common determinants of illiquidity.

Perhaps due to the aforementioned lack of clear-cut theories regarding its time series properties, the empirical time series modelling of liquidity has seen only modest activity over the years:

Chordia et al. (2001) for example focus on the time series properties of aggregate market liquidity, defined alternatively as either the average (percentage) quoted, or effective spreads, aggregate market depth or a composite ratio of quoted spreads across a sample of NYSE stocks from 1988 to 1998. They find that daily *changes* in aggregate liquidity as well as market activity - defined as aggregate (dollar) volume or number of trades per day - are highly volatile (more so than returns) and negatively serially correlated. Moreover, liquidity plummets significantly in down markets.

In a recent approach, Chordia et al. (2006) examine vector autoregressions (VARs) of the above liquidity proxies, returns and volatility for ten size-sorted portfolios of NYSE stocks as well as the value-weighted NASDAQ portfolio from 1988 to 2002. They find evidence for persistent liquidity, return and volatility 'spillovers' across size portfolios, with significant lead-lag effects. Large cap stocks for example appear to lead NASDAQ as well as small cap stocks in the transmission of liquidity shocks. Moreover, they find that stock returns and volatility in one sector can predict liquidity in the same sector as well as in others.

On the individual security level, the modelling approaches closest in spirit to the one employed in this chapter are given by Engle and Russell (1998) and Engle and Lange (2001): The latter devise the VNET measure of liquidity, which denotes the number of shares purchased minus the number of shares sold during a time interval in which prices move a pre-specified increment. VNET thus constitutes an intraday price duration-filtered measure of realised (net) depth associated with a certain price change. Using intraday data on NYSE stocks from November 1990 to January 1991 as well as for August to December 1997, they show that VNET is highly variable and in time series models can explained by past volume, number of transactions during previous price durations as well as past spreads. Importantly, VNET also exhibits high autocorrelation, a finding which is robust to the selection of the price increment.

While Engle and Lange's VNET model addresses the liquidity notions of depth and tightness, Engle and Russell (1998) focus on the immediacy notion of liquidity: They propose the so-called 'Autoregressive Conditional Duration' (ACD) model, whereby the time (duration) between consecutive trades is a point process and as such modelled as (multiplicative) autoregressive function of past durations and shocks. Their model builds on the empirical findings that intraday durations tend to cluster, similar to volatility, and lends itself to many empirical intraday applications such as the modelling of 'thinned point processes' like the arrival times of certain transaction volumes or price changes.

From the point of view of this chapter, the econometric modelling approach in the case of CARL is in many ways analogous to the ACD model, which will become more apparent below.

3.3 Two Building Blocks

Here, we present the two 'ingredients' for the CARL model. As mentioned before, our modelling relies on the ratio of the percentage difference in the minimum and maximum price to the price-weighted turnover, a measure that captures liquidity along several dimensions. We outline the properties of this ratio in more depth below. We then briefly present the well-established econometric time series concept of the ACD model, which will provide the dynamic framework for the CARL model.

3.3.1 A Multi-Dimensional Liquidity Measure

We use the so-called Hui-Heubel (HH) ratio as a measure of liquidity (cf. Sarr and Lybek, 2002). The HH ratio of liquidity (l_t) for period t is defined as

$$l_t = \frac{\frac{p_t^{\max} - p_t^{\min}}{p_t^{\min}}}{\frac{v_t}{\bar{p}_t \cdot s_t}}, \tag{3.1}$$

where

$$
\begin{aligned}
p_t^{\max} &\equiv \text{the maximum transaction price over the period}\\
p_t^{\min} &\equiv \text{the minimum transaction price over the period}\\
v_t &\equiv \text{the traded currency volume over the period}\\
\bar{p}_t &\equiv \text{the average price over the period}\\
s_t &\equiv \text{the average outstanding number of shares over the period.}
\end{aligned}
$$

There are a number of reasons why the HH ratio lends itself to use as a measure of liquidity for the purposes of this chapter:

- As it combines more than one dimension of liquidity in a single metric, the HH ratio can be considered a *composite* liquidity measure (cf. Kluger and Stephan, 1997). Liquidity is higher the lower l_t, i.e. the lower the numerator and the higher the denominator. Therefore, *ceteris paribus*, liquidity conditions are better the smaller the difference between $p_t^{\max}$ and $p_t^{\min}$ as a percentage of $p_t^{\min}$, which can readily be interpreted as a tighter market - *irrespective* of tick-sizes and price level. Equally, liquidity is higher, all other things equal, the higher price-weighted volume as a percentage of the outstanding price-weighted number of shares, i.e. the price-weighted turnover (share turnover has been prominently suggested and used as a separate, albeit one-dimensional measure of liquidity by Datar et al., 1998). Again, this makes intuitive sense as a higher price-weighted turnover rate (given the same tightness conditions) indicates a deeper market and thus more liquidity, albeit along a different dimension compared to the notion of tightness before. Compared to one-dimensional liquidity metrics such as for example the bid-ask spread, composite measures like the HH ratio have been shown to exhibit superior explanatory power for future expected returns as e.g. in Kluger and Stephan (1997).

- While also pointing out that bi-dimensional ratio-type measures of liquidity (given the different notions of liquidity applying to the numerator and denominator, the HH ratio can regarded as such) dominate one dimensional metrics in explaining commonalities in liquidity, Escribano et al. (2004) stress that changes in liquidity as measured by such a bi-dimensional metric can be ambiguous whenever liquidity dimensions do not reinforce each other. A particular occurrence of such an ambiguity is for example a situation in which the market tightens but less depth is provided, leaving the overall development of liquidity unclear (the interplay of the tightness and depth dimensions of liquidity is illustrated in figure 3.1).

  In the case of tightness (taken as the bid-ask spread) and depth Escribano et al. (2004) argue the case for incorporating a rate of substitution between the two *contemporaneous* facets of liquidity into bi-dimensional measures. While this renders the bi-dimensional analysis of liquidity in their setting more meaningful, incorporating it also compounds the empirical effort needed to derive such a metric. In the case of the HH ratio, it might be argued that fortunately this effort can be saved as the ratio, despite incorporating a notion of tightness, does not make use of the (change in the) instantaneous bid-ask spread as proxy for tightness nor of the (change in the) *contemporaneous* order book depth. Rather, by incorporating the price-weighted transaction volume and the associated difference between the maximum and

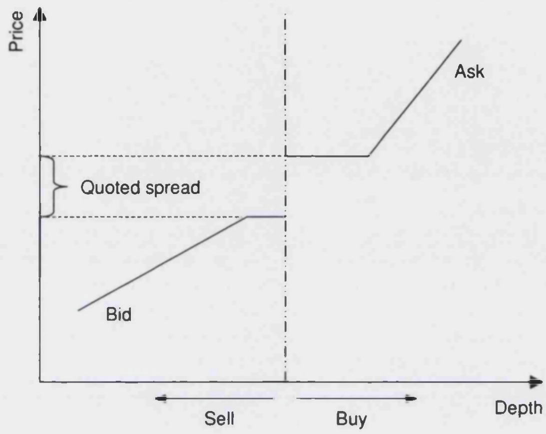

Figure 3.1: **Market liquidity - Tightness and depth**

This figure schematically displays the tightness as well as the depth dimension of market liquidity. The former is manifested by the bid-ask spread, the latter by the transaction volumes that can be bought and sold at different quoted prices.

minimum transaction price over some period, the HH ratio is to be considered a measure that relates volumes to their *impact* on prices and thus to resiliency, yet another dimension of liquidity. Moreover, it can be argued that the ratio also captures market breadth as a larger transacted percentage price-weighted volume per percentage maximum/minimum price difference, thus a lower l_t, indicates a broader market.

- Following the previous argument, the HH ratio is therefore closely aligned with Kyle's (1985) concept of liquidity, whereby prices respond to order flow and do so less in a broader, more liquid market. It has however been argued that this link and thus the meaningfulness of the HH ratio as an indicator of market liquidity can break down whenever price move instantaneously (and in large jumps) solely due to the announcement of news (cf. Sarr and Lybek, 2002, p. 13), say, with very little associated volume. Engle and Lange (2001, pp. 17), however, suggest that large price swings purely on the basis of new information are rare due to the presence of stale limit orders in the market and market makers charged with maintaining a steady price path. While conjecturing that news do (indirectly) move prices, they also stress that in 'normal' circumstances large movements in transaction prices do not stem from public announcements in an instantaneous fashion, but rather evolve over a longer time frame, with associated volumes driving the price evolvement. Thus, with the presence of new purely information-based trading in the market being fairly infrequent, the HH ratio as a measure of market liquidity is still meaningful.

- The incorporation of the percentage difference between the maximum and minimum price over the measurement period in the case of the HH ratio as opposed to e.g. the period return as in the ILLIQ measure used by Amihud (2002), also has the added benefit of linking the liquidity measure to the asset's volatility, as has been suggested by both the theoretical and empirical literature mentioned above: A period marked by a flat return, but large intra-period price swings would heuristically be classified as more illiquid than one with flat returns throughout, yet, ceteris paribus, a return-based measure would indicate an equal level of liquidity for both. Using the high and low prices of the period, however, the HH ratio on the other hand takes the higher level of volatility into account, flagging a lower level of liquidity for the former situation.

- By construction, also being a measure of the level of liquidity as opposed to the change in it, the HH ratio is always positive, which is particularly relevant for the econometric modelling as a positive, multiplicative autoregressive process. This point is highlighted in greater detail in section 3.4.

- Being the ratio of yet another two relative liquidity metrics, the HH ratio is independent of scales and units. This might be particularly relevant for researchers and practitioners seeking to employ the ratio to past data and markets that are characterised by non-decimal minimum tick sizes as for example on the NASDAQ before April 2001.

- Due to its simplicity, the HH ratio can be readily calculated from the most commonly used datasets. Contrary to other liquidity measures such as e.g. the duration as used in the ACD model, the HH ratio can be computed and thus used in a model at various frequencies, ranging from intradaily to annually and longer. This has the benefit of being able to compare higher frequency metrics to smoother, longer-run averages, a practice that is commonly used when analysing returns, but has to the author's knowledge not been prominent in the case of liquidity

Despite the above benefits of the HH ratio as a measure of market liquidity, it needs to be stressed that as such it is far from being perfect and is not appropriate for all applications. As suggested by various researchers (cf. e.g. Amihud, 2002, p. 35), given the multidimensionality of liquidity, it is doubtful that a single measure is able to capture all aspects of liquidity perfectly and in an exhaustive way. Nonetheless, for the purposes of this chapter, the HH ratio can be readily employed in a time series econometric context, which we lay out in the next section.

3.3.2 The Econometrics

As mentioned above Engle and Russell (1998) have conceptualised the 'Autoregressive Conditional Duration' (ACD) model, which constitutes a multiplicative, autoregressive process for the time series behaviour of trade durations, i.e. the intervals between successive transactions.

Being a so-called 'marked point process' like the well known Poisson process, the ACD model may be represented via the hazard function, a time series process that indicates the event rate at time a point in time, such as e.g. in the 'proportional hazard' (PH) modelling approach (i.e. the hazard function is taken to be a 'baseline hazard' multiplied by a function that depends on lagged durations and/or other exogenous variables), as well as by modifying the time scale directly, as in the 'accelerated time' (AT) time series approach used by Engle and Russell (1998):

They model durations as a baseline duration with a unity expectation times a function depending on lagged durations and/or other exogenous variables. From an econometric point of view, their approach is closely related to Bollerslev's (1986) GARCH model, however, in contrast to the returns in the latter, the ACD is modelled on a positive domain, with an error distribution taken from a family that fulfills this criterion, such as e.g. the exponential, Weibull or Gamma distribution.

It is mainly because of its positive domain multiplicative structure, that the AT ACD modelling technique provides a very useful econometric setup for the CARL model, in which we seek to model yet another positive variable, the HH ratio as a multiplicative autoregressive process through time. The next section provides more detail on the mechanics of the CARL model.

3.4 The CARL Model

Having laid out the building blocks for the CARL model above, we now turn to assembling the model: We present the full dynamic framework of the model as well as its statistical properties. We also touch on the issue of 'Maximum Likelihood Estimation' (MLE) as well the forecasting properties. Finally, we discuss a possible way of introducing exogenous past variables into the regressor equation to create a richer model that is also in accordance with theoretical research.

3.4.1 Model Setup

We model the univariate time series of the market liquidity of a security, proxied by the HH ratio, as a multiplicative autoregressive process akin to the ACD model for intraday trade durations. The resulting model is the 'Conditional Autoregressive Liquidity' (CARL) model, with the following more formal definition:

Definition 5 *l_t, the Hui-Heubel ratio at time t follows a CARL(p,q) process if*

$$l_t = \lambda_t \varepsilon_t, \textit{ with } \varepsilon_t \sim \textit{i.i.d.}, \mathbb{E}[\varepsilon_t] = 1, Var(\varepsilon_t) = \zeta_\varepsilon^2 < \infty, \forall t \in \{0, ..., T\},$$

and

$$\lambda_t = \omega + \sum_{i=1}^{p} \alpha_i \lambda_{t-i} + \sum_{j=1}^{q} \beta_j l_{t-j}, , \tag{3.2}$$

where $\omega > 0$ and $\alpha_i, \beta_j, \forall i, j$ obey the inequality conditions for the positivity of λ_t laid out in Nelson and Cao (1992).

The conditions on the i.i.d. innovations ε_t[6] ensure the that the expectation of the positive liquidity measure l_t conditional on the information set up to and including time $t-1$ (denoted $\mathcal{F}_{t-1}$) is

$$\mathbb{E}[l_t|\mathcal{F}_{t-1}] = \lambda_t. \tag{3.3}$$

The unconditional expectation of l_t is

$$\mathbb{E}[l_t] = \mu_l = \frac{\omega}{(1 - \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j)}, \tag{3.4}$$

which can be obtained by taking expectations on both sides of equation (3.2). Strictly speaking, the unconditional expectation only exists if under the above conditions, all roots of the associated difference equation polynomial

$$1 - (\alpha_1 + \beta_1)z - ... - (\alpha_m + \beta_m)z^m, \text{ with } m = \max(p, q) \tag{3.5}$$

lie outside of the unit circle. In the simple case of $p = q = 1$, with the above conditions on the roots of (3.5) holding, we also have

$$Var(l_t|\mathcal{F}_{t-1}) = \lambda_t^2 \tag{3.6}$$

[6]i.i.d. is taken short for '*i*ndependently and *i*ndentically *d*istributed'.

and

$$Var(l_t) = \sigma_l^2 = \mu_l^2 \zeta_\varepsilon^2 \left[\frac{1 - \alpha^2 - 2\alpha\beta}{1 - \alpha^2 - 2\alpha\beta - (1 + \zeta_\varepsilon^2)\beta^2} \right] \tag{3.7}$$

for the conditional and unconditional variance respectively. The autocovariance function (also cf. Bauwens and Giot, 2000) of the first order is given by

$$\gamma_1 = \omega\mu_l + \mathfrak{a}\alpha + \mathfrak{b}\beta - \mu_l^2, \tag{3.8}$$

with

$$\begin{aligned} \mathfrak{a} &= \mathbb{E}[\lambda_t^2] = Var(\lambda_t) + \mu_l^2 \\ &= \mu_l^2 \left[\frac{1 - (\alpha + \beta)^2}{1 - \alpha^2 - 2\alpha\beta - (1 + \zeta_\varepsilon^2)\beta^2} \right] + \mu_l^2 \\ &= \mu_l^2 \zeta_\varepsilon^2 \left[\frac{\beta^2}{1 - \alpha^2 - 2\alpha\beta - (1 + \zeta_\varepsilon^2)\beta^2} \right] \end{aligned} \tag{3.9}$$

and

$$\mathfrak{b} = \mathbb{E}[l_t^2] = \sigma_l^2 + \mu_l^2. \tag{3.10}$$

The autocovariance function γ_i for higher orders $i \geq 2$ is

$$\gamma_i = \omega\mu_l + \sum_{k=1}^{i-1} \mathfrak{D}_{i,k} + \mathfrak{a}\alpha^i + \mathfrak{b}\beta\alpha^{i-1} - \mu_l^2, \tag{3.11}$$

with

$$\mathfrak{D}_{i,k} = \alpha^{k-1}\beta(\gamma_{i-k} + \mu_l^2) + \omega\mu_l^2\alpha^k. \tag{3.12}$$

For higher order cases of p and q, the expressions for the unconditional variance and the autocovariance function increase in complexity very rapidly and are thus omitted here.

Setting $\nu_t = l_t - \lambda_t$, which is a martingale difference sequence, the CARL model can be re-written in the form of an ARMA(m, p) representation with non-Gaussian innovations and $m = \max(p, q)$:

$$l_t = \omega + \sum_{i=1}^{m} (\alpha_i + \beta_i) l_{t-i} + \sum_{j=1}^{p} \nu_{t-j} + \nu_t. \qquad (3.13)$$

From the representation, it can be seen that the condition for the existence of the unconditional expectation also ensures covariance stationarity of l_t: The latter is obtained iff all roots of the polynomial in (3.5) lie outside of unit circle. Given the above representation, forecasts for the CARL model are easy to compute using standard ARMA time series techniques. Moreover, the ARMA representation highlights the fact that in the CARL model the dependent liquidity variable l_t is modelled as an autoregressive function of its own lagged values. In this function the parameters α_i and β_i $(i = 1, ...m)$ govern the behaviour of the model: The closer the sum of the parameters is to unity, the less mean-reverting the process will be and the higher the autocorrelation in l_t. We now turn to the estimation and asymptotic properties of the process.

3.4.2 Asymptotic Properties and Estimation

The CARL model is analogous to the ACD model and as such has many similarities with the GARCH model class (cf. Engle, 1982; Bollerslev, 1986). Indeed, as in the case of ACD, many properties of GARCH carry over to the CARL model: We pick the CARL model in the (1,1) parameterisation as a 'natural' point of departure for our analysis in this section.

The most obvious (and arguably simple) candidate for an error distribution for the CARL model is, as in the case of the ACD model, the exponential distribution, with $\varepsilon_t \geq 0$ and $\mathbb{E}[\varepsilon_t] = Var(\varepsilon_t) = 1$. For 'exponential CARL', the ties to the GARCH model class are especially close: Importantly, the exponential CARL(1,1) model can be estimated via QML, similar to the GARCH(1,1) model. This important property is stated more clearly in the following proposition:

Proposition 6 *(Analogous to Engle and Russell, 1998) If a CARL(1,1) model exhibits the following properties*

I. $\mathbb{E}[l_t|\mathcal{F}_{t-1}] = \lambda_{0,t} = \omega_0 + \alpha_0\lambda_{t-1} + \beta_0 l_{t-1}$;

II. $\epsilon_t = \frac{l_t}{\lambda_{0,t}}$ *is strictly (a) stationary, (b) ergodic, (c) non-degenerate, (d)* $\mathbb{E}[\epsilon_t^2|\mathcal{F}_{t-1}] < \infty$, *(e)* $sup_t\mathbb{E}[\ln(\beta_0\epsilon_t + \alpha_0)|\mathcal{F}_{t-1}] < 0 a.s.$;

III. $\theta_0 = (\omega_0, \alpha_0, \beta_0)^\top$ *is in the interior of the parameter space* $\mathbf{\Theta}$;

IV. $\mathcal{L}(\theta) = -\frac{1}{T}\sum_{t=1}^{T} g_t(\theta)$, *where* $g_t(\theta) = \left[\log(\lambda_t) + \frac{l_t}{\lambda_t}\right]$,

$$\text{with} \begin{cases} \lambda_t = \omega + \alpha\lambda_{t-1} + \beta l_{t-1} \text{ for } i > 1 \\ \lambda_t = \frac{\omega}{(1-\alpha)} \text{ for } i = 1 \end{cases},$$

then:

1. $\theta^* = \arg\max[\mathcal{L}(\theta)] \rightarrow_p \theta_0$;
2. θ^* *is asymptotically normal and has the asymptotic covariance matrix* $\mathbf{V}_0 = \mathbf{B}_0^{-1}\mathbf{A}_0\mathbf{B}_0^{-1}$, *with* $\mathbf{B}_0 = -\mathbb{E}[\nabla^2 g_t(\theta_0)]$ *and* $\mathbf{A}_0 = \mathbb{E}[\nabla g_t(\theta_0)\nabla g_t(\theta_0)^\top]$.

Building on Lee and Hansen (1994), the proof of the above proposition is exactly analogous to Engle and Russell (1998, p. 1135) and is therefore omitted here.

The above results always hold for the stationary case with $\alpha+\beta < 1$ and for some integrated and explosive cases that meet condition *II.(e)*. More importantly, the proposition also covers situations in which $\epsilon_t = \frac{l_t}{\lambda_t}$ is not i.i.d., but only the weaker condition of strict stationarity as in *II.(a)* holds. The above also means that, as in the case of the ACD model, the CARL model can be conveniently estimated as a GARCH process by setting $\sqrt{l_t}$ as the input variable.

It should be stressed that the above only concerns the (1,1) case, while yet other parametrisation setups are conceivable. The extension of the analysis for these cases is far from trivial, as noted by both Engle and Russell (1998) and Lee and Hansen (1994). It is, however, to be expected that similar results hold.

Additionally, while the proposition shows the QML estimator to be consistent, it is not necessarily efficient if the true error density is different from the exponential distribution. In these cases, 'Maximum Likelihood' (ML) estimation with the 'correct' density is to be preferred. For the ACD model, Engle and Russell (1998) for example propose the Weibull and the Gamma family of distributions as possible candidates for the error density. While such an extension is also feasible in the case of the CARL model, the focus of this chapter is more on the presentation of the new model and versatility in various data environments rather than on efficient estimation. Indeed, it is to be expected that the exponential model, with its lack of degrees of freedom in the parameterisation of the density, will not yield the most efficient estimates and is most likely to be dominated in this respect by setups with other error distributions. Yet, as per proposition 6, estimation results obtained with the exponential CARL model are still consistent, which is arguably sufficient for most practical applications.

It is also important to note that while QML estimation with the exponential model is consistent in the case of a 'wrong' density assumption, this does not hold in the case of a erroneous assumption about the entire data generating process (DGP). Thus, as also stressed by Lee and Hansen (1994) for GARCH, the above analysis rests on the assumption that the assumed CARL model setup is 'correct', which might however not strictly hold[7]. Unfortunately, an extension of the above results for the case of a 'wrong' CARL process assumption is outside the scope of this chapter and has to our knowledge not been established for the GARCH and ACD model class either.

3.4.3 Theoretical Underpinnings and Extensions

One of the most important features of the CARL model is its ability to pick up autoregressive patterns in the HH ratio, and thus arguably, in market illiquidity of a security over time. As can be seen by recursively substituting into the autocovariance function of the CARL(1,1) model given in equations (3.8) and (3.11), the process has decaying autocorrelations, whereby the rate of decay is governed by the magnitude of α. The higher closer the sum of α and β, the slower the autocorrelations die out. For the purpose of modelling market liquidity, this property provides an important link between the empirical CARL approach and theoretical research laid out in section 3.2.1: As high (low) market liquidity begets high (low) market liquidity during liquidity spirals according to the research reviewed (cf. Brunnermeier and Pedersen, 2008), the CARL model should be poised to pick up such patterns when estimated with real data.

[7] In the case of GARCH for example most research seems to (implicitly) agree that the model does only approximate the true DGP.

From the ARMA representation of CARL(1,1) in equation (3.13) it is further evident that in a stationary CARL process with $\alpha+\beta<1$ is mean-reverting, i.e. any (large, 'unusual') shocks to the process will eventually die out. From a theoretical point of view this feature is very much in line with e.g. Morris and Shin (2004) who postulate that liquidity black holes are eventually corrected and liquidity provision in markets will revert to normal.

Given the ARMA representation it is also easy to see that CARL model can be used for forecasting, i.e. predicting $l_{t+s}, s>0$ using variables in the information set $\mathcal{F}_t$ available at time t. In the case of CARL(1,1), this yields:

$$l(s) \equiv \mathbb{E}[l_{t+s}|\mathcal{F}_t] = \lambda_{t+s} = \frac{\omega[1-(\alpha+\beta)^{s-1}]}{1-\alpha-\beta} + (\alpha+\beta)^{s-1}l(1), \quad (3.14)$$

with

$$l(1) \equiv \mathbb{E}[l_{t+1}|\mathcal{F}_t] = \lambda_{t+1} = \omega + \alpha\lambda_t + \beta l_t.$$

Therefore, as $s \to \infty$, one has $\lim_{s\to\infty}[l(s)] = \mu_l = \frac{\omega}{1-\alpha-\beta}$, i.e. long-range forecasts essentially constitute the unconditional mean.

As established above, the CARL model can estimated as a fully specified probability model via QML. While this is a feasible result from a theoretical point of view, one usually performs *conditional* QML estimation in practice, i.e. maximising the average of the sum of the log of conditional densities:

$$\max_{\theta,\eta} T^{-1}\left(\sum_{t=2}^{T} \log f(l_t|\mathcal{F}_{t-1};\theta) + \log f(l_1;\eta)\right),$$

where $f(\cdot)$ denotes the error density used. Whereas in the definition of the basic CARL model in section 3.4, $\mathcal{F}_{t-1}$ simply includes lagged values of l_t, it may also comprise other exogenous variables that are non-random from a time t perspective - without changing the econometric properties of the model needed for estimation via conditional QML. One such example would be the following, 'augmented' CARL(p,q,r) model, in which the conditional mean λ_t is modelled as a function of lagged values of λ_t and l_t as well as $\mathcal{F}_{t-1}$-adapted indicator variables that return 1 whenever the past return y_{t-k} of period $t-k$ was negative and 0 otherwise:

$$\lambda_t = \omega + \sum_{i=1}^{p} \alpha_i \lambda_{t-i} + \sum_{j=1}^{q} \beta_j l_{t-j} + \sum_{k=1}^{r} \gamma_k \mathbb{I}[y_{t-k} < 0], \tag{3.15}$$

with $\gamma_k \geq 0, \forall k$.

Choosing this parameter setup introduces a threshold effect that is in accordance with the theoretical research laid out above: According to Brunnermeier and Pedersen (2008) negative returns (on equities) create losses on traders' positions, leading to funding problems, which in turn - through a so-called 'loss spiral' - translate into less trading and deteriorating market liquidity. Worsened market liquidity then feeds into yet more volatility and a more strained trading environment, possibly leading to even more losses. Thus, from an empirical point of view, negative (past) returns should inject over-proportionately more illiquidity, thereby aggravating the autocorrelation of market illiquidity, which the threshold variable in the augmented CARL(p,q,r) model above is designed to detect.

3.5 Empirical Application

In this section, we present a brief empirical application of the CARL model using the example of Amazon equity data. We estimate various versions of the model against daily as well as weekly data, highlighting the model's ability to cope with different frequency settings in the empirical analysis of liquidity.

3.5.1 Data

As mentioned above, the empirical application in this chapter is carried out using CRSP data of the NASDAQ-listed online retailer Amazon. We selected this particular stock due to its characteristics as a 'new economy blue chip': Amazon went public in 1997, but even though it arguably is a new economy stock, it dodged the burst of the 'dot-com bubble' in 2000 and emerged as one of the heavyweights in the NASDAQ index. It has since grown into one of the largest online retailers today. From the perspective of the analysis in this chapter, the stock thus has the desirable characteristics of having exposure to the volatile and 'nervous' technology segment of the market, which also reflects on liquidity being more volatile, while being sufficiently large and therefore defensive enough. Moreover, for Amazon the required data for the empirical analysis with the CARL model is readily available on both daily and monthly frequencies via CRSP (as well as on an intraday basis via the TAQ database).

In this chapter we estimate CARL models against daily as well as weekly data. The daily data covers the entire listing period of Amazon on the NASDAQ until the end of 2006, i.e. 15/05/1997-29/12/2006. We also extract the 'decimalised' period, i.e. the time span starting with the first date of decimalised trading[8] in Amazon stock, which in this case is 09/04/2001-29/12/2006. This is done to facilitate a comparison between 'tick regimes'.

The data comprises the daily holding return as per the CRSP database as well as the HH ratio. The latter is derived using the daily maximum and minimum transaction price given by CRSP as well as the shares traded and shares outstanding. As traded volume is not given on a transaction price weighted basis in the daily CRSP database and cannot be derived other than from intraday data, the denominator of the HH ratio is constructed by simply dividing trades shares multiplied by the daily average price (the average of the last and current day's closing price as per CRSP) by outstanding shares, again multiplied by the daily average price. As the latter drops out, this effectively amounts to using share turnover as the denominator. Even though this procedure does not strictly abide by the way the denominator of the HH ratio is defined, deriving the denominator from intraday data by weighing each intraday transaction size by its corresponding price involves high computational effort, which does not appear to significantly improve the accuracy of the ratio given the narrow spans of daily transaction prices.

The weekly data also comprises both the holding return and the HH ratio times and covers the period 19/05/1997-28/12/2006. The series are derived from daily CRSP data on a 'Wednesday-to-Wednesday' basis to avoid 'beginning-of-the-week' and 'end-of-the-week' effects: The maximum and minimum transaction prices for the week are used to construct the numerator of the HH ratio. For the denominator,

[8] Traditionally, the tick size on the NASDAQ was 1/8. That was changed to 1/16 on 02/06/1997 and later o gave way to fully decimalised trading in all stocks.

Table 3.1: **Amazon HH ratio, return and turnover data statistics**

This table shows statistics for the Amazon HH ratio, holding return and price-weighed share turnover calculated using CRSP data. The Ljung-Box autocorrelation test statistic for the first 50 lags is denoted as LB_{50}, the Jarque-Bera test statistic as JB. The statistics are presented for daily and weekly frequencies, with the former covering both the entire NASDAQ listing period from 15/05/1997-29/12/2006 (shown as 1997-2006) as well as the period starting with NASDAQ 'decimalisation' on 09/04/2001 (shown as 2001-2006). The weekly series have been derived from the 19/05/1997-28/12/2006 daily series on a 'Wednesday-to-Wednesday' basis.

Period	**1997-2006** (daily)	**2001-2006** (daily)	**1997-2006** (weekly)
HH ratio summary statistics			
Observations	2,423	1,440	502
Mean	2.747	2.298	1.304
Median	2.074	1.939	0.973
Maximum	23.587	13.623	6.978
Minimum	0.200	0.200	0.254
Std dev	2.715	1.382	0.896
Skewness	2.972	2.306	2.022
Kurtosis	18.053	11.773	8.903
LB_{50} (P-value)	34,836 (0.00)	16,203 (0.00)	3,600 (0.00)
Holding return summary statistics			
Mean	0.246	0.178	1.168
Median	−0.020	0.000	0.054
Maximum	34.470	34.470	46.053
Minimum	−24.770	−24.770	−30.200
Std dev	4.996	3.793	9.781
Skewness	0.790	1.202	0.901
Kurtosis	8.414	19.332	5.815
JB (P-value)	3,211 (0.00)	16,352 (0.00)	233 (0.00)
Price weighed share turnover summary statistics			
Mean	2.947	2.072	14.759
Median	1.972	1.744	10.334
Maximum	32.794	18.378	91.494
Minimum	0.171	0.343	2.003
Std dev	2.992	1.449	12.613

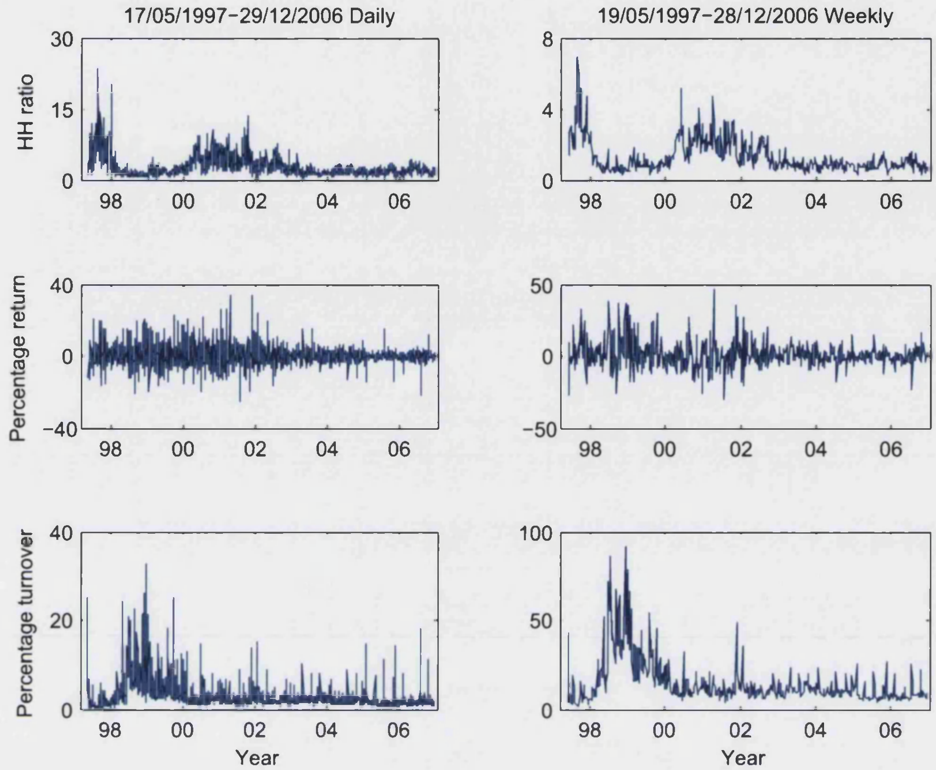

Figure 3.2: **Amazon HH ratio, return and turnover**

This figure displays Amazon HH ratio, holding return and price-weighed share turnover data for the 1997-2006 period. Both daily data from 15/05/1997-29/12/2006 as well as 'Wednesday-to-Wednesday' weekly data for the period 19/05/1997-28/12/2006 are shown.

we weigh both the trading volume as well as the shares outstanding on each trading day by the average transaction price of the day (the average of the current and previous day's closing prices) and then divide the so-weighted sum of the daily volume over the course of a week by the sum of the daily weighted shares outstanding over the week. The so-calculated price-weighted daily average turnover is then multiplied by 5, the number of trading days per week to construct a weekly average price-weighted turnover to be used as the denominator of the weekly HH ratio.

Summary statistics of the daily as well as weekly Amazon HH ratio, holding returns as well as the price weighted turnover can be found in table 3.1, while the time series of these metrics are displayed in figure 3.2. From the figure it is evident that, in general, the time series of the daily as well as weekly HH ratio exhibit a 'clustered' pattern (similar to volatility), indicating strong autocorrelation, with marked periods of low as well high liquidity: During 1997 up until 1998 for example, Amazon experienced a period of relatively low liquidity (as measured by a high HH ratio) - both on a daily as well as on a weekly measurement basis. Thereafter, liquidity in the stock in conjunction with turnover picked up strongly and deteriorated again during 2000-2002, possibly as a consequence of the faltering dot-com boom. From 2002 onwards, the HH ratio as well as turnover rates can be seen to be fairly stable on both frequencies, with daily turnover picking up slightly towards the end of the sample, a likely reflection of the growing importance of Amazon as a leading online retailing stock.

Comparing the daily to weekly data, one notices that market liquidity in general is higher for the weekly frequency, as indicated by a comparatively lower HH ratio. Given the definition of the ratio, this is does not come as a surprise as a market should be able to 'digest' more volume with the same price impact during a trading week than during a trading day - simply also because a week offers more time to trade than a day.

The table confirms this observation, as for example indicated by a lower mean HH ratio for the weekly series as opposed to the daily frequency. From the table, in particular the large 'Ljung-Box' (LB) statistics (calculated conservatively for 50 lags), it is also evident that HH ratio time series on both daily and weekly frequencies is indeed highly autocorrelated, confirming the impression given by figure 3.2.

The table further indicates that holding returns of the stock on both frequencies are strongly non-Gaussian (as shown by large 'Jarque-Bera' (JB) statistics), with strong kurtosis and skewness especially during the 2001-2006 period.

3.5.2 Estimation and Results

As indicated in section 3.4.2, estimation of the CARL model via QML can be carried out using conventional time series computer packages that include GARCH functionality by estimating the CARL model as a Gaussian GARCH process with $\sqrt{l_t}$ as the dependent variable. We resort to this technique below to obtain the results below.

Daily Frequency

First, we estimate the CARL model against the two daily Amazon datasets described above, comprising the HH ratio as well as the holding return for the 'full' 1997-2006 as well as the decimalised 2001-2006 period. The histogram and the autocorrelation function (ACF) of the HH ratio time series for the former and the latter period are shown in figures 3.3 and 3.4 respectively.

For both periods the histograms exhibit a skewed shape on the positive domain, reminiscent of e.g. a Gamma distribution. While this is surely a feasible choice, we use the exponential distribution here to facilitate QML estimation. The autocorrelation functions (ACFs) show quite persistent autocorrelation that dies out only after some time, c. after 250 and 300 lags for the 1997-2006 and 2001-2006 period respectively. We will return to the strong persistence in more detail in the discussion section below. One also notes that over the first, say, 50 lags the ACF for the 2001-2006 period is less high in magnitude

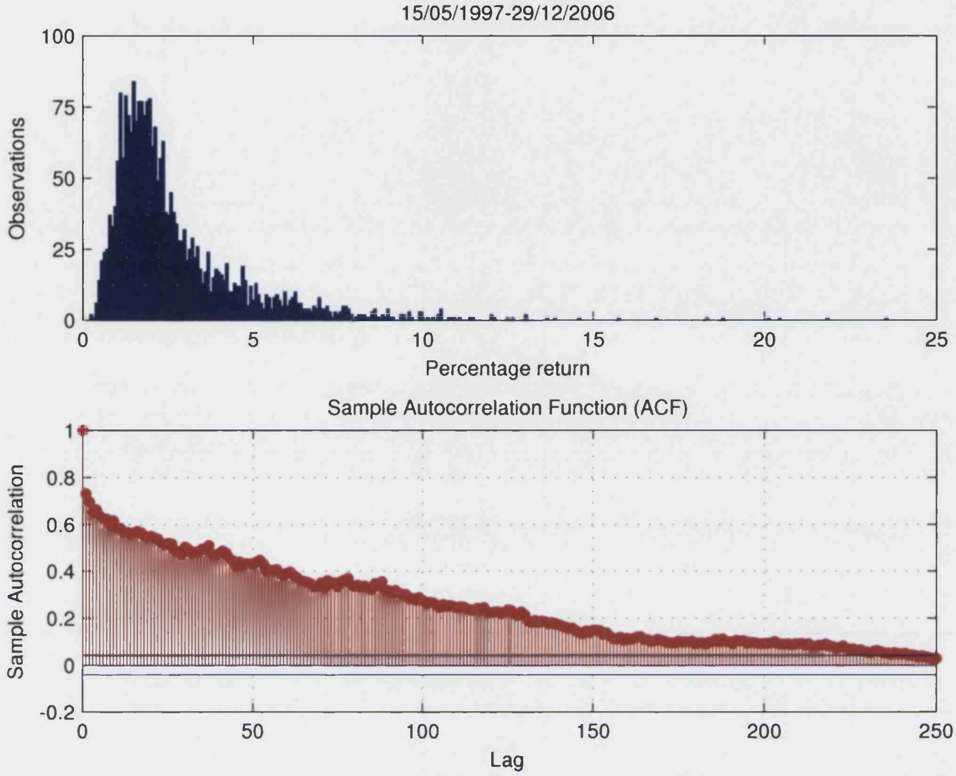

Figure 3.3: **Histogram and ACF - 1997-2006 daily Amazon HH ratio**

This figure displays the histogram (top panel) and the autocorrelation function (ACF; bottom panel) for the daily Amazon HH ratio series from 15/05/1997-29/12/2006. In addition to the ACF, 2-standard-deviation interval bounds are shown, corresponding roughly to a 95% confidence interval.

and also declines faster. A likely reason for this is that Amazon experienced a fairly consistent period of relatively low liquidity before 1998, followed by notably stronger liquidity, again fairly consistently, up until 2000 (cf. figure 3.2). Thereafter, especially after 2002, the end of the dot-com boom, the evolvement of the HH ratio appears less clustered, possibly contributing to a slightly lower and faster declining ACF for the decimalised period.

For the estimation of the CARL model against these datasets, we employ the following method: We estimate CARL(p,q) models for all

Table 3.2: **Estimation - 1997-2006 daily Amazon data**

This table shows results from the estimation of different selected CARL models against the daily Amazon data series from 15/05/1997-29/12/2006. The table displays parameter estimates, Bollerslev-Woolridge t statistics, goodness-of-fit measures as well as statistics for standardised residuals. * and ** denote significance at 1% and 5% levels respectively. The Ljung-Box autocorrelation test statistic for the first 50 lags is denoted as LB_{50}.

Model	**CARL(1,1)**		**CARL(1,2)**		**CARL(1,2,1)**	
Metric	**Estimate**	**t stat**	**Estimate**	**t stat**	**Estimate**	**t stat**
$\hat{\omega}$	0.044*	3.926	0.028*	3.455	−0.017**	−2.152
$\hat{\alpha}_1$	0.702*	42.745	0.795*	40.879	0.890*	71.678
$\hat{\beta}_1$	0.282*	18.716	0.344*	15.409	0.345*	15.217
$\hat{\beta}_2$			−0.150*	−4.992	−0.241*	−9.017
$\hat{\gamma}$					0.063*	4.268
Log-LH		−4,479		−4,478		−4,473
AIC		3.700		3.699		3.699
BIC		3.707		3.709		3.711
			Standardised residuals statistics			
Mean		1.000		1.002		1.002
Std Dev		0.370		0.368		0.367
LB_{50} (P-value)		96 (0.00)		59 (0.18)		58 (0.20)

Table 3.3: **Estimation - 2001-2006 daily Amazon data**

This table shows results from the estimation of different selected CARL models against the daily Amazon data series in the 'decimalised period' from 09/04/2001-29/12/2006. The table displays parameter estimates, Bollerslev-Wollridge t statistics, goodness-of-fit measures as well as statistics for standardized residuals. * and ** denote significance at 1% and 5% levels respectively. The Ljung-Box autocorrelation test statistic for the first 50 lags is denoted as LB_{50}.

Model	**CARL(1,1)**		**CARL(1,2)**		**CARL(1,2,1)**	
Metric	**Estimate**	**t stat**	**Estimate**	**t stat**	**Estimate**	**t stat**
$\hat{\omega}$	0.058*	3.522	0.048*	3.229	0.008	0.467
$\hat{\alpha}_1$	0.746*	34.587	0.781*	29.270	0.816*	35.637
$\hat{\beta}_1$	0.229*	12.269	0.292*	9.823	0.280*	9.401
$\hat{\beta}_1$			−0.095**	−2.491	−0.114*	−3.109
$\hat{\gamma}$					0.066*	2.982
Log-LH		−2586.365		−2585.936		−2583.674
AIC		3.596		3.597		3.598
BIC		3.607		3.612		3.616
Standardised residuals statistics						
Mean		1.000		1.001		1.001
Std Dev		0.353		0.352		0.352
LB_{50} (P-value)		68 (0.04)		57 (0.24)		56 (0.27)

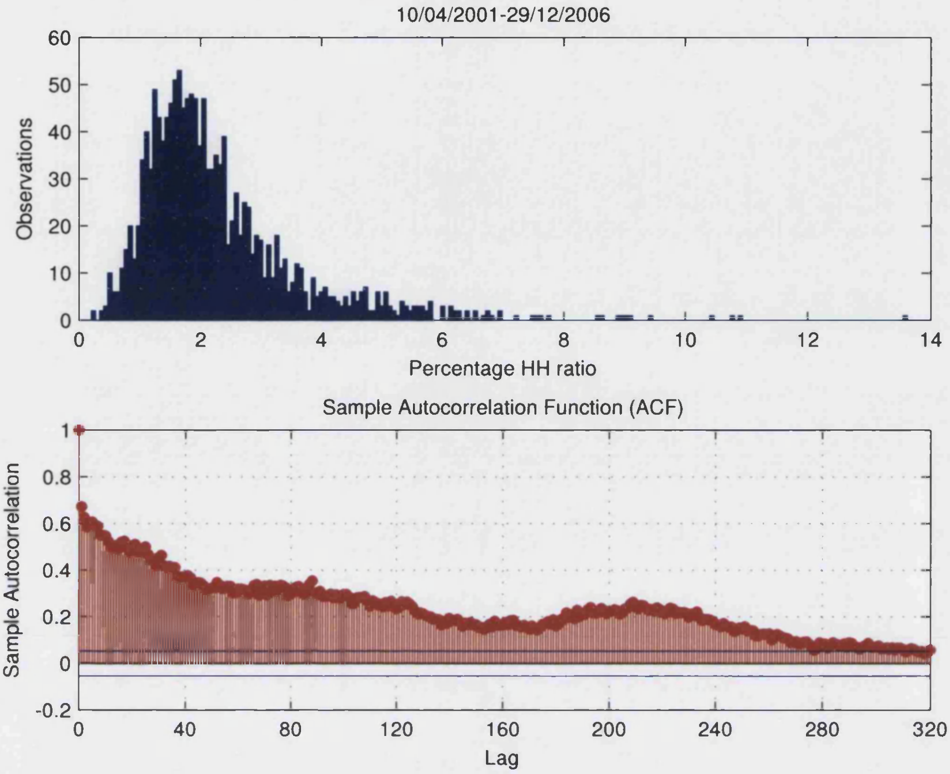

Figure 3.4: **Histogram and ACF - 2001-2006 daily Amazon HH ratio**

This figure displays the histogram (top panel) and the autocorrelation function (ACF; bottom panel) for the daily Amazon HH ratio series in the 'decimalised period' from 09/04/2001-29/12/2006. In addition to the ACF, 2-standard-deviation interval bounds are shown, corresponding roughly to a 95% confidence interval.

possible $p \in (0, 1, 2, 3)$ and $q \in (1, 2, 3)$ parametrisation combinations and compare the results in terms of goodness of fit using the AIC and BIC as well as the remaining (ideally low or not existent) autocorrelation of the standardised residuals. The CARL model offering the best fit in terms of these categories is then selected as optimal and 'augmented' with a single, lagged threshold variable according to the specification in (3.15). The augmented model is again estimated and contrasted with the non-augmented version as well as with a 'baseline' CARL(1,1) model. The reason for doing the latter is that for

GARCH models the simple $p = q = 1$ parametrisation has been found to dominate more complex versions of the GARCH model in certain applications (cf. Hansen and Lunde, 2005).

We report the estimation results for both daily datasets in tables 3.2 and 3.3 for the 1997-2006 and the decimalised period respectively. In both cases, the optimal model chosen is the CARL(1,2) model, for which the reduction in the autocorrelation of the standardised residuals in conjunction with the AIC and BIC is comparatively most favourable. Table 3.2 for example shows a LB statistic of 59 for the first 50 lags with a P-value of 0.18, indicating that the null of no autocorrelation over the first 50 lags cannot be rejected at a reasonable level of significance. For the decimalised period, the picture is even clearer, possibly reflecting less autocorrelation in the HH ratio time series in the first place (s. above), with an LB_{50} statistic of 57 and a corresponding P-value of 0.23. For both datasets, this compares with a highly significant LB statistic in the case of the CARL(1,1) baseline model, for which the null of no remaining autocorrelation in the standardised residuals over the first 50 lags has to be rejected, indicating inferior model fit. Moreover, also in terms of the AIC, the CARL(1,1) parametrisation compares unfavourably with the optimal CARL(1,2).

For both time horizons, augmenting the optimal CARL(1,2) with a past return threshold indicator variable according to (3.15) produces even better model fit in terms of the reduction of the autocorrelation of the standardised residuals: a reduction in the LB_{50} statistic for the 1997-2006 and 2001-2006 periods to 58 and 56 respectively. However, adding a further variable reflects unfavourably, albeit only very modestly, on the BIC (in the case of the 2001-2006 period also on the AIC). On the whole, however, it might be argued that including the return threshold indicator variable into the model specification slightly improves results, which is further confirmed by the fact that the coefficient of the variable is highly statistically significant for both horizons and has the 'correct' sign, i.e. $\hat{\gamma} > 0$. This result thus provides evidence for the existence of a return threshold effect along the lines of the theoretical research discussed in section 3.2.1. This point is debated further in the discussion section below.

Assessing the magnitude and the statistical significance of the parameters estimates further shows that in the case of the augmented model for both horizons the significance of the estimate of the constant term ω vanishes; this is to be expected as there is a linear relationship with the threshold variable. The magnitude of the threshold effect as measured by $\hat{\gamma}$ appears to be the same for both horizons.

For the 1997-2006 period, the model detects a slightly higher (initial) autocorrelation in (il)liquidity, indicated by a higher $\hat{\alpha}_1$ and $\hat{\beta}_1$ obtained for this period compared to 2001-2006, the effect of which however eventually gets damped by a more negative $\hat{\beta}_2$, resulting in a more persistent autocorrelation pattern for 2001-2006 period, as is also evident from figures 3.3 and 3.4. In both cases, the estimated CARL(1,2) model is stationary (because $\hat{\alpha}_1+\hat{\beta}_1+\hat{\beta}_2 < 1$) and, despite a negative $\hat{\beta}_2$, also adheres to the positivity criterion as the parameter estimates fulfill the (sufficient) conditions required by Nelson and Cao (1992), i.e. $0 \leq \alpha_1 < 1, \beta_1 \geq 0, (\alpha_1\beta_1 + \beta_2) \geq 0$.

We also test for the forecasting properties of the CARL by performing a forecast test according to Mincer and Zarnowitz (1969): For both horizons, we use the data and the parameter estimates of the estimated augmented CARL(1,2,1) model to produce an in-sample forecast series

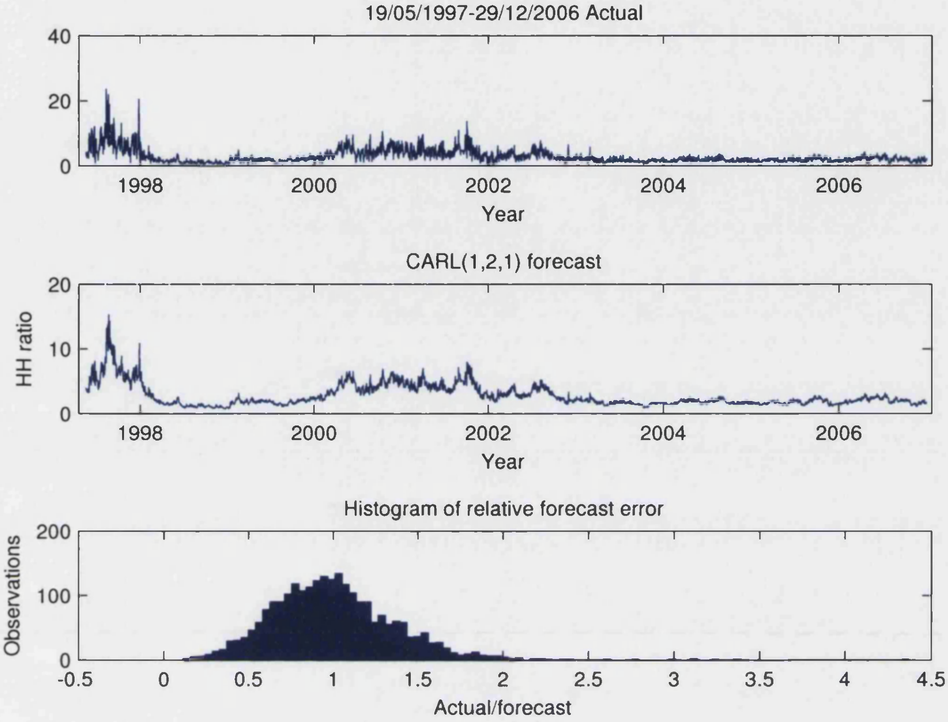

Figure 3.5: **Actual vs Forecast - 1997-2006 daily Amazon HH ratio**

This figure shows the actual daily Amazon HH ratio (top panel), the CARL model forecast (middle panel) as well as a histogram of the relative forecast error (bottom panel) for the 19/05/1997-29/12/2006 period. The forecast is derived using the fitted CARL(1,2,1) model.

$\hat{l}_t$. We then run a simple OLS regression of the sort $l_t = \theta_0 + \theta_1 \hat{l}_t + \epsilon_t$. The quality of the forecast can then be tested for by performing a Wald test on the joint parameter restriction $\theta_0 = 0 \cap \theta_1 = 1$. Under the null hypothesis of forecast optimality, this restriction holds .

The original series l_t, the forecast $\hat{l}_t$ from the CARL(1,2,1) model and the empirical relative forecast error $(\frac{l_t}{\hat{l}_t})$ distribution for the 1997-2006 and the 'decimalised' 2001-2006 is displayed in figures 3.5 and 3.6 respectively. The forecast series follows the actual one quite closely, albeit in a less extreme way, i.e. the amplitudes of extreme values of l_t are visibly forecasted to be lower. This is not surprising as essentially

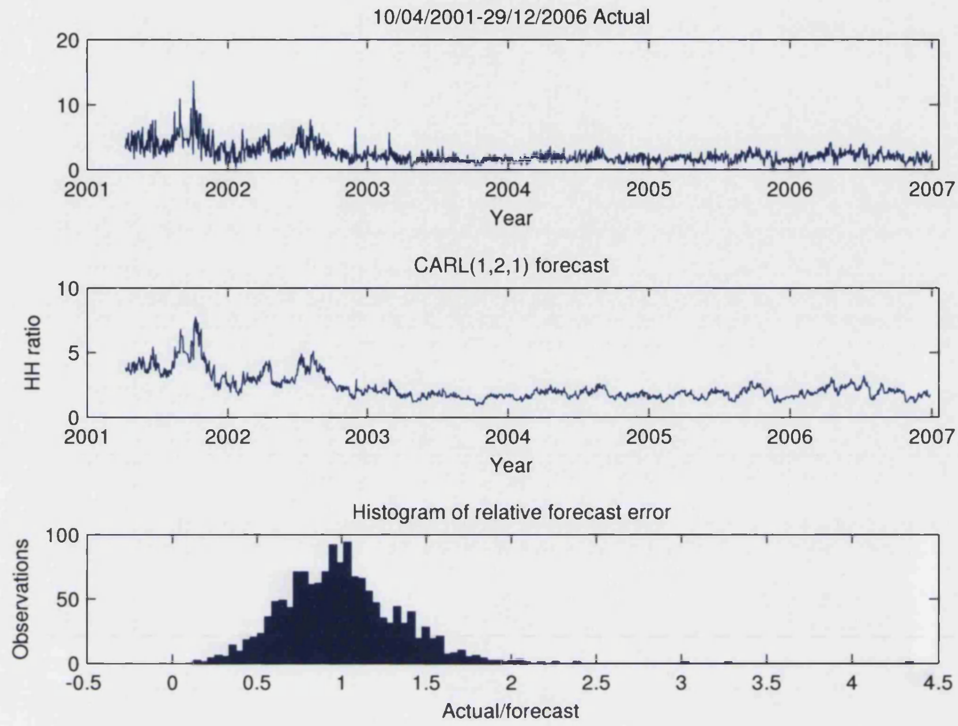

Figure 3.6: **Actual vs Forecast - 2001-2006 daily Amazon HH ratio**

This figure shows the actual daily Amazon HH ratio (top panel), the CARL model forecast (middle panel) as well as a histogram of the relative forecast error (bottom panel) for the 'decimalised period' from 10/04/2001-29/12/2006. The forecast is derived using the fitted CARL(1,2,1) model.

the CARL model, by construction, 'smoothens over' the data when constructing a forecast. The relative error distributions confirm this observation: While the forecast seems to be precise on average as indicated by the distributions being centered with their modes roughly at unity, the right hand tail of the distribution corresponds to instances where the forecast undershoots the actual observations. Equally, the distribution also shows that there are occurrences of the opposite, with the forecast being too high, as indicated by the left tail of the distribution; graphically, these instances are however more difficult to make out when comparing the figures displaying the actual and the

Table 3.4: **MZ - 1997-2006 daily Amazon data**

This table shows a Mincer-Zarnowitz regression of the form $l_t = \theta_0 + \theta_1 \hat{l}_t + \epsilon_t$ for the daily Amazon data series from 15/05/1997-29/12/2006. The forecast series $\hat{l}_t$ is obtained using the CARL(1,2,1) model, for which the in-sample fit is found optimal.

MZ regression		
Metric	**Estimate**	**t stat**
$\hat{\theta}_0$	−0.012	−0.118
$\hat{\theta}_1$	1.010*	23.713
Restriction test ($H_0 : \theta_0 = 0 \cap \theta_1 = 1$)		
Wald test stat (P-value)	0.367 (0.83)	

Table 3.5: **MZ - 2001-2006 daily Amazon data**

This table shows a Mincer-Zarnowitz regression of the form $l_t = \theta_0 + \theta_1 \hat{l}_t + \epsilon_t$ for the daily Amazon data series in the 'decimalisation period' from 09/04/2001-29/12/2006. The forecast series $\hat{l}_t$ is obtained using the CARL(1,2,1) model, for which the in-sample fit is found optimal.

MZ regression		
Metric	**Estimate**	**t stat**
$\hat{\theta}_0$	−0.030	−0.350
$\hat{\theta}_1$	1.015*	25.065
Restriction test ($H_0 : \theta_0 = 0 \cap \theta_1 = 1$)		
Wald test stat (P-value)	0.128 (0.94)	

forecast series.

Tables 3.4 and 3.5 provide the results of the Mincer-Zarnowitz (MZ) tests for the 1997-2006 and 2001-2006 horizons respectively: Inspection reveals that the null of forecast optimality cannot be rejected in both cases, albeit the test for the decimalised time period seems to provide the stronger result. However, caution is at hand to assess test results across both horizons as the MZ does not allow for relative comparisons.

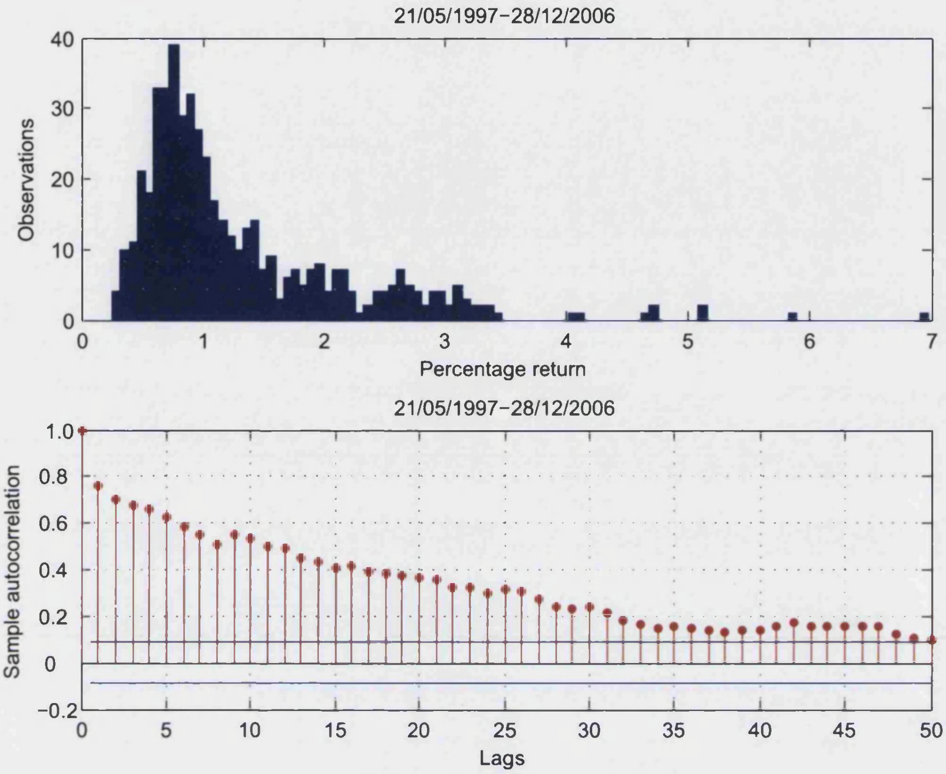

Figure 3.7: **Histogram and ACF - 1997-2006 weekly Amazon HH ratio**

This figure displays the histogram (top panel) and the autocorrelation function (ACF; bottom panel) for the weekly Amazon HH ratio series from 19/05/1997-28/12/2006. In addition to the ACF, 2-standard-deviation interval bounds are shown, corresponding roughly to a 95% confidence interval.

Weekly Frequency

The analysis here mirrors the one carried out for the daily data series above: Figure 3.7 shows the histogram and ACF for the weekly HH ratio series from 19/05/1997 to 28/12/2006. Similar to the daily series, the histogram exhibits a skewed shape on the positive domain. Inspection of the ACF reveals that in comparison to daily sampled data, the autocorrelation in the weekly HH ratio series dies out more quickly, i.e. while starting at a similar, of not slightly higher autocorrelation for the first lag (c. 0.75), the ACF then declines much more

Table 3.6: **Estimation - 1997-2006 weekly Amazon data**

This table shows results from the estimation of different selected CARL models against the weekly Amazon data series from 19/05/1997-28/12/2006. The table displays parameter estimates, Bollerslev-Woolridge t statistics, goodness-of-fit measures as well as statistics for standardised residuals. * and ** denote significance at 1% and 5% levels respectively. The Ljung-Box autocorrelation test statistic for the first 50 lags is denoted as LB_{50}.

Model	**CARL(0,1)**		**CARL(1,1)**		**CARL(1,1,1)**	
Metric	**Estimate**	**t stat**	**Estimate**	**t stat**	**Estimate**	**t stat**
$\hat{\omega}$	0.318	9.441	0.036**	2.367	−0.005	−0.285
$\hat{\alpha}_1$			0.675*	17.697	0.671*	16.973
$\hat{\beta}_1$	0.747	20.200	0.294*	8.433	0.292*	7.965
$\hat{\gamma}$					0.096*	4.138
Log-LH		−752		−748		−746
AIC		3.003		2.993		2.994
BIC		3.019		3.018		3.027
Standardised residuals statistics						
Mean		1.000		1.001		1.000
Std Dev		0.400		0.360		0.355
LB_{50} (P-value)		152 (0.00)		42 (0.80)		40 (0.84)

steeply and is barely significant at the 50^{th} lag. A likely explanation for this is obviously the lower weekly sampling frequency, which tends to 'smoothen over' and average out strong and persistent daily autocorrelations.

Table 3.6 presents the results of the estimation of various CARL models against the data. The procedure follows the one in the previous section: Here, however, the optimal model chosen is the CARL(1,1) parametrisation, for which the combination of goodness-of-fit in terms of AIC and BIC as well as the reduction in the autocorrelation of the standardised residuals was most favourable. Thus, for comparison, we have included the even simpler CARL(0,1) model, the equivalent of the ARCH(1) model. Although in terms of BIC, the CARL(1,1) models still fares equally well in comparison to CARL(0,1), it is far superior in terms of AIC and the reduction in autocorrelation. While

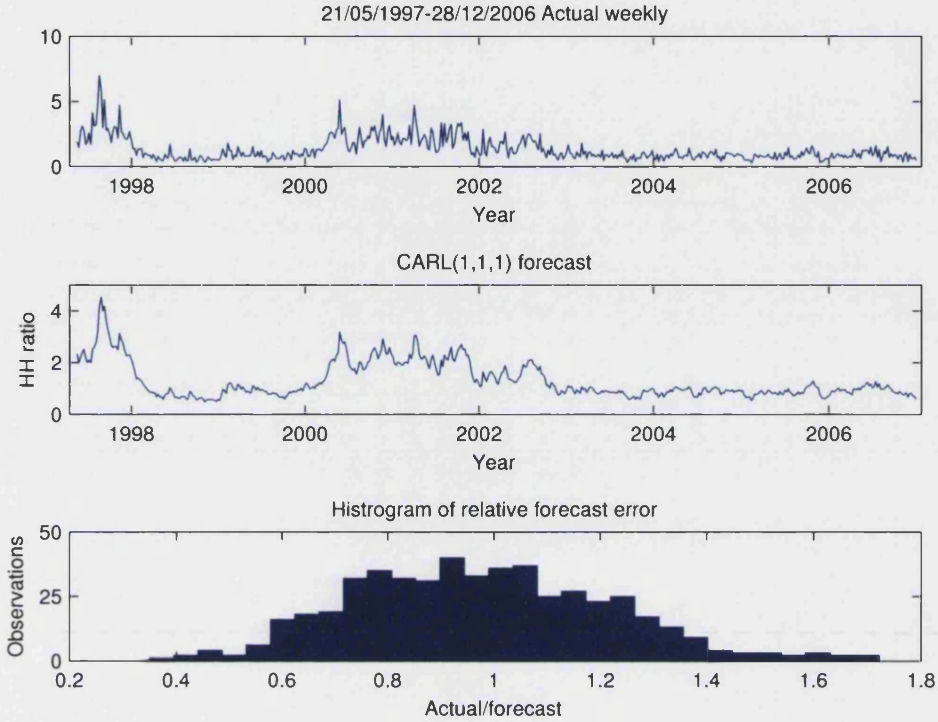

Figure 3.8: **Actual vs Forecast - 1997-2006 weekly Amazon HH ratio**

This figure shows the actual weekly Amazon HH ratio (top panel), the CARL model forecast (middle panel) as well as a histogram of the relative forecast error (bottom panel) for the 21/05/1997-28/12/2006 period. The forecast is derived using the fitted CARL(1,1,1) model.

the CARL(0,1) model still leaves very significant autocorrelation in the standardised residuals, the null hypothesis of no autocorrelation over the first 50 lags cannot be rejected for the CARL(1,1) model with a JB_{50} statistic of 42 and an associated P-value of 0.80.

When augmenting the CARL(1,1) model with a lagged return threshold variable in the same way as before, both AIC and BIC indicate a minor slip in model fit for the CARL(1,1,1) model, introduced by the punishment for the inclusion of the extra variable. There is however, yet another slight reduction in the autocorrelation of the standardised residuals with a JB_{50} statistic of 50 and a P-value of 0.84. Moreover, as in the case of daily data, the coefficient of the threshold variable $\hat{\gamma}$ is highly significant and, again, has the 'right' sign, i.e. $\hat{\gamma} > 0$. Equally, as to be expected, the estimate of ω loses its significance in the augmented model.

Judging the order of magnitude of the parameter estimates, one notices that in comparison to the estimates obtained with daily data, the sum of α_i and β_j is lower in magnitude, indicating a lower the persistence in (il)liquidity. This is well in line with more steeply declining ACF in figure 3.8.

The parameter estimates for the simple model further indicate stationarity (as $\hat{\alpha}_1 + \hat{\beta}_1 < 1$) and, with $\hat{\alpha}_1 > 0, \hat{\beta}_1 > 0$, also adhere to the positivity requirement.

The in-sample forecast $\hat{l}_t$ obtained with the parameter estimates of the augmented CARL(1,1,1) model, the original weekly HH ratio time series l_t as well as the empirical relative $\frac{l_t}{\hat{l}_t}$ forecast error distribution are displayed in figure 3.8: As in the case of the daily data, the forecast is a smoother version of the original series with less extreme amplitudes, but seems to be well precise on average with the relative error distribution exhibiting a clear and pronounced hump at unity. The Mincer-Zarnowitz regression, documented in table 3.7, shows that the null hypothesis of the optimality of the forecast cannot be rejected. In fact, given the P-value of 0.87 of the Wald test statistic, the MZ regression provides strong evidence in favour of forecast optimality.

3.5.3 Discussion

Reviewing the above results as a whole, the CARL model fares well in picking up the quite strong autocorrelation patterns present in the HH ratio liquidity time series on both daily as well as weekly frequen-

Table 3.7: **MZ - 1997-2006 weekly Amazon data**

This table shows a Mincer-Zarnowitz regression of the form $l_t = \theta_0 + \theta_1 \hat{l}_t + \epsilon_t$ for the weekly Amazon data series from 19/05/1997-28/12/2006. The forecast series $\hat{l}_t$ is obtained using the CARL(1,1,1) model, for which the in-sample fit is found optimal.

MZ regression		
Metric	**Estimate**	**t stat**
$\hat{\theta}_0$	-0.035	-0.600
$\hat{\theta}_1$	1.034*	16.504
Restriction test ($H_0 : \theta_0 = 0 \cap \theta_1 = 1$)		
Wald test stat (P-value)	0.289 (0.87)	

cies. The presence of such autocorrelation in the data is in accordance with the theoretical research laid out in section 3.2.1, according to which high market liquidity begets high liquidity and vice versa. The goodness-of-fit of the CARL model in the presence of such data thus strengthens our belief that the model is a valid empirical implementation of the above theory on market liquidity: The reduction in the autocorrelation in the standardised residuals after fitting the optimal model is significant, e.g. from a LB_{50} statistic of 34,826 in the original 1997-2006 daily time series to 59 in the standardised residuals after fitting a CARL(1,2) model to the same data. The in-sample forecasting capabilities are equally impressive, with the optimality of forecasts obtained with the best-fitting augmented CARL processes not being rejected when running a MZ regression.

Moreover, we find a significant threshold effect according to which a lagged negative return increases illiquidity even further compared to a positive past return. This finding is very much in line with theoretical research and might be counted towards evidence for the existence of loss-spirals in market liquidity as advocated by Brunnermeier and Pedersen (2008).

Heuristically, comparing the estimation results for the two daily data series above, we find no fundamental difference in the parameter estimates of the optimal CARL(1,2) and augmented CARL(1,2,1) model that would point towards a significant structural break due to the introduction of decimalised trading in the Amazon stock on 09/04/2001. Indeed, a separate analysis (not shown here) of the 'non-decimalised' period, i.e. 15/05/1997-08/04/2001, finds similar results compared to the decimalised period: The same optimal CARL(1,2) model is picked for the non-decimalised period, and parameter estimates are highly significant (with the exception of $\hat{\omega}$) and in the same order of magnitude, with slightly higher $\hat{\alpha}_1$ and $\hat{\beta}_1$ and a more negative $\hat{\beta}_2$, while still preserving stationarity and positivity. To us, these results indicate that there is no *fundamental* difference in the daily law of motion of liquidity[9] in the Amazon stock as a result of decimalised trading.

Liquidity Persistence

Comparing the results across sampling frequencies, we note that the estimation of the CARL model in the context of the weekly series results in the CARL(1,1) parametrisation being picked, while the richer CARL(1,2) is the optimal choice for daily data. Judging from the parameter estimates and also from the ACFs of the HH ratio time series at the two different frequencies it appears therefore, as already indicated above, that the autocorrelation for the lower weekly frequency is less persistent.

While this is a relative qualification across frequencies, the general

[9]The HH ratio as a measure of market liquidity in this context is not affected by tick-sizes and scales in the first place - thus excluding any *technical/measurement* effects that a new tick size might introduce.

level of persistence in the empirical HH ratio autocorrelations given the data in this chapter is rather high: As evident from figures 3.3, 3.4 and 3.7, autocorrelations lose significance only after about 250-300 and 50 lags for daily and weekly data respectively.

This artefact is very much reminiscent of the persistence in the autocorrelation in volatilities as documented e.g. by Bollerslev and Mikkelsen (1996). Moreover, with regards to liquidity, Engle and Russell (1998) also find that the persistence in trade durations is very high when devising their ACD model, with the sum of $\hat{\alpha}_i$ and $\hat{\beta}_j$ as in (3.2) being close to one, thus indicating a process bordering on non-stationarity. This finding has led to the conclusion of the existence of 'long memory' in these time series and prompted the development of fractionally integrated time series models, e.g. the FIGARCH for volatility by Baillie et al. (1996) and the FIACD for durations by Jasiak (1998).

While it appears that a similar conclusion might be at hand for the CARL model in this chapter, we would argue that the case is not as clear-cut: First, even though the ACFs of the series indicate strong persistence, such effect is arguably different across frequencies. It is not imminently clear how such a difference can be incorporated into a unified CARL framework: After all, the existing model appears to pick up the existing autocorrelations quite well across frequencies as witnessed by the lack of any apparent remaining autocorrelation in the standardised residuals after estimation for both data frequencies.

Moreover, the sum of the estimated $\hat{\alpha}_i$ and $\hat{\beta}_j$ coefficients in the optimal models is 0.989 and 0.978 for the daily 1997-2006 and 2001-2006 periods respectively as well 0.969 for the weekly estimation. While this is arguably close to unity, the conclusion of non-stationarity does not look immediately imperative as the 'gap' appears to be too big[10]. Further studies in this direction look warranted to establish the empirical stationarity properties of CARL across securities and markets.

Lastly, in comparison to the persistent autocorrelation in durations, the ACFs of the HH ratio in this chapter exhibit an important difference: Whereas the ACF of durations e.g. in the case of IBM data as documented e.g. by Jasiak (1998) starts off at rather low levels of c. 0.08 and then very slowly declines while still being significant at even a 1,000 lags, the structure of the autocorrelations of the HH ratio for Amazon data is rather different. Here, initial levels of autocorrelation are very high, starting at c. 0.8, a level ten times higher, and then fade more rapidly, losing significance after c. 250 and 50 lags for daily and weekly data respectively. It might therefore be argued that even though the decline may be slow, it can still be considered exponential as implied by the CARL model - albeit at a very low rate[11].

Distributional Assumptions

As pointed out above, our choice of using the exponential distribution to model the error term is governed by the fact that such an assumption facilitates the use of QML for estimation, producing consistent estimates while being able to estimate the CARL with GARCH functionality. However, if the true error distribution is not exponential, such a procedure, while still being consistent, will be non-efficient. We would argue that such concerns are of secondary nature for the objective of this chapter; nonetheless, for future work in this area,

[10]Estimating their exponential ACD model Engle and Russell (1998) for example obtain coefficients summing to 0.996, which arguably represents stronger evidence of non-stationarity.

[11]For example, an exponential factor of 0.989^x results in a level of 0.2 and 0.07 for $x = 150$ and 250 respectively, which roughly corresponds to the structure of the ACF for the daily 1997-2006 period if one views x as the lag input.

challenging the distributional assumption appears to be a promising route to further refine the CARL model: In figures 3.5, 3.6 and 3.8, the shape histograms of the empirical relative forecast errors, which incidentally also constitutes the standardised residuals of the estimation, indicates that, rather than the exponential, the hump-shaped Gamma or Weibull distributions might be appropriate choices for modelling the error term. However, while such a choice would affect the efficiency of the estimation, the qualitative nature of the results cannot be expected to change. Refinements of this kind are therefore left for future work.

3.6 Conclusion

In this chapter, we present the CARL model, an empirical univariate time series model of market liquidity that can be applied to both individual securities as well as to markets as a whole. As a measure of liquidity we use the HH ratio, a metric that combines several dimensions of liquidity in a single ratio and is easy to compute from most commonly available datasets on various frequencies. The CARL model combines this simple, yet meaningful metric with a multiplicative autoregressive process, similar to Engle and Russell's (1998) ACD model.

In an empirical application of the CARL model to Amazon equity data, we find that the model is straight-forward to estimate via QML, provides good fit to the data and is able to forecast market liquidity as measured by the HH ratio well. Importantly, these results hold for both daily and weekly data, demonstrating the versatility of the CARL model in applications involving different data frequencies; to our knowledge, other existent empirical models of liquidity are not equipped for such tasks.

Our empirical analysis also finds evidence for a return threshold effect as proclaimed by theoretical research on liquidity: In our results, negative returns trigger an even stronger increase in illiquidity, whereas the reaction to positive returns is neutral, thus pointing towards the existence of 'loss spirals' (as in Brunnermeier and Pedersen, 2008) in the dynamics of market liquidity.

On the whole, the results obtained with the CARL model in this chapter are encouraging and point towards several areas of application: On its own, the CARL model provides a good tool for the tracking and monitoring of liquidity in markets as a whole and on an individual security basis and might thus be of interest to central banks and traders as well as risk managers in banks, where it could be integrated into the wider risk management. Additionally, one may also think of integrating CARL into a vector autoregression (VAR), e.g. by taking logs, as such a resulting model seems well-poised for the analysis of the interaction of liquidity and returns. Other extensions, such as a multivariate setup, similar to multivariate GARCH are also conceivable and look promising, especially for the analysis of liquidity commonality (cf. section 3.2.2).

Further, we would like to stress that the focus of this chapter is on the presentation of the model and its properties. Thus, our empirical application provides a demonstration of its capabilities rather than a comprehensive empirical study. We deem this to be outside of the scope of the chapter and leave more detailed empirical applications for later work: An obvious area of more extensive work is for example the analysis of liquidity with the help of the CARL model on yet even more different data frequencies. Other conceivable research might comprise a detailed analysis of the out-of-sample forecast performance of the model, which we have not undertaken.

Lastly, the ML estimation of the model with different, more complex error distributions, such as the Weibull or the Gamma distribution mentioned above also looks worthwhile empirical exercise.

Bibliography

Acharya, V. V. and Pedersen, L. H. (2005), 'Asset pricing with liquidity risk', *Journal of Financial Economics* **77**, 375–410.

Amihud, Y. (2002), 'Illiquidity and stock returns: cross-section and time-series effects', *Journal of Financial Markets* **5**, 31–56.

Amihud, Y. and Mendelson, H. (1980), 'Dealership markets: market making with inventory', *Journal of Financial Economics* **8**, 21–53.

Amihud, Y. and Mendelson, H. (1986), 'Asset pricing and the bid-ask spread', *Journal of Financial Economics* **17**, 223–249.

Amihud, Y., Mendelson, H. and Pedersen, L. H. (2006), 'Liquidity and asset prices', *Foundations and Trends in Finance* **1**, 269–364.

Bagehot, W. P. (1971), 'The only game in town', *Financial Analysts Journal* **22**, 12–14.

Baillie, R. T., Bollerslev, T. and Mikkelsen, H. O. (1996), 'Fractionally integrated generalized conditional heteroskedasticity', *Journal of Econometrics* **74**, 3–30.

Baker, K. (1996), 'Trading location and liquidity: an analysis of the US dealer and agency markets for common stocks', *Financial Markets, Institutions and Instruments* **5**(4).

Bank of England (2006), Financial Stability Report, Issue no. 20, Bank of England; available from `http://www.bankofengland.co.uk/publications/fsr/2006/fsrfull0606.pdf`.

Bank of England (2007), Financial Stability Report, Issue no. 21, Bank of England; available from `http://www.bankofengland.co.uk/publications/fsr/2007/fsrfull0407.pdf`.

Basel Committee on Banking Supervision (2008), Principles for sound liquidity risk management and supervision, Report, Bank of International Settlements; available from `http://www.bis.org/publ/bcbs138.pdf`.

Bauwens, L. and Giot, P. (2000), 'The logarithmic ACD model: an application to the bid-ask quote process of three NYSE stocks', *Annales d'Economie et de Statistique* **60**, 117–149.

Bollerslev, T. (1986), 'Generalized autoregressive conditional heteroscedasticity', *Journal of Econometrics* **31**, 307–327.

Bollerslev, T. and Mikkelsen, H. O. (1996), 'Modeling and pricing long memory in stock market volatility', *Journal of Econometrics* **73**, 151–184.

Brennan, M. J., Chordia, T. and Subrahmanyam, A. (1998), 'Alternative factor specifications, security characteristics and the cross-section of expected stock returns', *Journal of Financial Economics* **49**, 345–373.

Brunnermeier, M. and Pedersen, L. H. (2008), Market liquidity and funding liquidity. Working paper, Princeton University; forthcoming in the *Review of Financial Studies*; available from `http://www.princeton.edu/~markus/research/papers/liquidity`.

Chordia, T., Roll, R. and Subrahmanyam, A. (2000*a*), 'Co-movements in bid-ask spreads and market depth', *Financial Analysts Journal* **56**, 23–27.

Chordia, T., Roll, R. and Subrahmanyam, A. (2000*b*), 'Commonality in liquidity', *Journal of Financial Economics* **56**, 3–28.

Chordia, T., Roll, R. and Subrahmanyam, A. (2001), 'Market liquidity and trading activity', *Journal of Finance* **56**, 501–530.

Chordia, T., Sarkar, A. and Subrahmanyam, A. (2006), Liquidity spillovers and cross-autocorrelations. Working Paper, UCLA; obtainable from `http://www.s1ubra.org/cscapn1.pdf`.

Connor, G. and Korajczyk, R. (1988), 'Risk and return in an equilibrium APT: application of a new test methodology', *Journal of Financial Economics* **21**, 255–290.

Copeland, T. and Galai, D. (1983), 'Information effects on the bid-ask spread', *Journal of Finance* **38**, 1457–1469.

Datar, V. T., Naik, N. Y. and Radcliffe, R. (1998), 'Liquidity and stock returns: An alternative test', *Journal of Financial Markets* **1**, 203–219.

Diamond, D. and Dybvig, P. (1983), 'Bank runs, deposit insurance, and liquidity', *Journal of Political Economy* **91**, 401–419.

Duffie, D., Gârleanu and Pedersen, L. H. (2005), 'Over-the-counter markets', *Econometrica* **73**, 1815–1847.

Duffie, D., Gârleanu and Pedersen, L. H. (2006), Valuation in over-the-counter-markets. Working Paper, forthcoming in the *Review of Financial Studies*; obtainable from `http://pages.stern.nyu.edu/%7Elpederse/papers/OTCvaluation.pdf`.

Engle, R. F. (1982), 'Autoregressive conditional heteroscedasticity with estimates of the United Kingdom inflation', *Econometrica* **50**, 987–1007.

Engle, R. F. and Lange, J. (2001), 'Predicting VNET: a model of the dynamics of market depth', *Journal of Financial Markets* **4**, 113–142.

Engle, R. F. and Russell, J. R. (1998), 'Autoregressive conditional duration; a new model for irregularly spaced transaction data', *Econometrica* **66**, 1127–1162.

Escribano, l., Pascual, R. and Tapia, M. (2004), 'On the bi-dimentionality of liquidity', *European Journal of Finance* **10**, 1–25.

Fama, E. and French, K. R. (1992), 'The cross-section of expected stock returns', *Journal of Finance* **47**, 427–465.

Fama, E. and French, K. R. (1993), 'Common risk factors in the returns of stocks and bonds', *Journal of Financial Economics* **33**, 3–56.

Financial Services Authority (2003), Liquidity risk in the Integrated Prudential sourcebook: a quantitative framework, Discussion paper, Prudential Standards Division at the Financial Services Authority (FSA); available from `http://www.fsa.gov.uk/pubs/discussion/dp24.pdf`.

Glosten, L. and Milgrom, P. (1985), 'Bid, ask and transaction prices in specialist market with heterogeneously informed traders', *Journal of Financial Economics* **14**, 71–100.

Grossman, S. J. and Miller, M. H. (1988), 'Liquidity and market structure', *Journal of Finance* **43**, 617–633.

Hansen, P. R. and Lunde, A. (2005), 'A forecast comparison of volatility models: Does anything beat a GARCH(1,1)?', *Journal of Applied Econometrics* **207**, 873–889.

Hasbrouck, J. (2006), Trading costs and returns for US equities: estimating effective cost from daily data. Working Paper, NYU Stern School of Business; obtainable from `http://pages.stern.nyu.edu/~jhasbrou/Research/GibbsEstimates2006/ctc10.pdf`.

Hasbrouck, J. and Seppi, D. (2001), 'Common factors in prices, order flows and liquidity', *Journal of Financial Economics* **59**, 383–411.

Ho, T. S. Y. and Stoll, H. (1981), 'Optimal dealer pricing under transactions and return uncertainty', *Journal of Financial Economics* **9**, 47–73.

Ho, T. S. Y. and Stoll, H. (1983), 'The dynamics of dealer markets under competition', *Journal of Finance* **38**, 1053–1074.

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

Jasiak, J. (1998), 'Persistence in intertrade durations', *Finance* **19**, 166–195.

Kluger, B. D. and Stephan, J. (1997), 'Alternative liquidity measures and stock returns', *Review of Quantitative Finance and Accounting* **8**, 19–36.

Kyle, A. S. (1985), 'Continuous auctions and insider trading', *Econometrica* **53**, 1315–1335.

Lee, S.-W. and Hansen, B. E. (1994), 'Asymptotic theory for the GARCH(1,1) Quasi-Maximum Likelihood estimator', *Econometric Theory* **10**, 29–52.

Mincer, J. and Zarnowitz, V. (1969), The evaluation of economic forecasts, *in* J. Mincer, ed., 'Economic Forecasts and Expectations', National Bureau of Economic Resarch, New York, pp. 3–46.

Morris, S. and Shin, H. S. (1998), 'Unique equilibrium in a model of self-fulfilling currency attacks', *American Economic Review* **88**, 587–597.

Morris, S. and Shin, H. S. (2004), 'Liquidity black holes', *Review of Finance* **8**, 1–18.

Nelson, D. B. and Cao, C. Q. (1992), 'Inequality constraints in the univariate GARCH model', *Journal of Business & Economics Statistics* **10**, 229–235.

Obstfeld, M. (1986), 'Rational and self-fulfilling balance-of-payments crisis', *American Economic Review* **76**, 72–81.

Pastor, L. and Stambaugh, R. F. (2003), 'Liquidity risk and expected stock returns', *Journal of Political Economy* **111**, 642–685.

Ross, S. A. (1976), 'The arbitrage theory of capital asset pricing', *Journal of Economic Theory* **13**, 341–360.

Sarr, A. and Lybek, T. (2002), Measuring liquidity in financial markets. Working Paper 02/232, International Monetary Fund; available from `http://www.imf.org/external/pubs/ft/wp/2002/wp02232.pdf`.

Stoll, H. R. (1978), 'The supply of dealer services in securities markets', *Journal of Finance* **33**, 1133–1151.

Vayanos, D. (2004), Flight to quality, flight to liquidity and the pricing of risk. Working Paper, London School of Economics; obtainable from `http://personal.lse.ac.uk/vayanos/WPapers/FQFLPR.pdf`.

Vayanos, D. and Wang, T. (2007), 'Search and endogenous concentration of liquidity in markets', *Journal of Economic Theory* **136**, 66–104.

Weill, P.-O. (2008), The liquidity premium in a dynamic bargaining market. Working Paper, UCLA, forthcoming in the *Journal of Economic Theory*; obtainable from `http://www.econ.ucla.edu/poweill/liquiditypremia.pdf`.

Chapter 4

Intraday Liquidity: A High-Frequency Application of the CARL Model

4.1 Introduction

Liquidity matters! Most participants in the financial markets would agree with this statement and at the same time not be able to write down a formula to calculate liquidity as we can for example to calculate volatility, another important financial market characteristic. Yet, most seem to have a fairly concrete opinion on the role of liquidity in the financial markets. Such views often tend to be fairly high-level: For example, most market participants would probably claim that the foreign exchange markets are more liquid than, say, the real estate market or emerging market shares with small capitalisations. It also seems an equally accepted view that liquidity problems in the interbank loan market were behind the crisis surrounding the British Bank Northern Rock in September 2007.

Statements like these again highlight the important role that liquidity plays for financial markets but show also how static and little differentiated the view on this abstract concept often is: It can be argued that the liquidity that fostered the crisis in the interloan bank market during the wake of the subprime debacle is of a very different nature than the liquidity underlying e.g. the afore-mentioned foreign exchange market. The former may be regarded as 'funding liquidity', whereas the latter can be described as 'market liquidity'[1]. Moreover, as for example Brunnermeier and Pedersen (2008) have shown, the interplay of the two concepts leads to dynamics in liquidity, whereby the liquidity of a market is not simply high or low, but rather varies through time and is subject to self-reinforcing mechanisms.

This chapter takes account of this finding and concentrates on the empirical modelling of the dynamics of financial market liquidity. Moreover, we embed our analysis in an *intraday* context, the most 'fluid' data environment possible. The reasons for doing so are two-fold: For a pragmatic point of view, intraday data is available in large volumes, making a meaningful empirical analysis easier from a statistical point of view. More importantly though, the experience in financial markets has shown that intraday market liquidity - contrary to the static view on the liquidity conditions in markets - is very fast-changing and, if lacking, can distort the trading process considerably. It is arguable for example that the market crash in 1987 was very much a function of fading intradaily liquidity in the market place, with the problem becoming more and more aggravated throughout the day: Carlson (2007) for example points out that

[1]This categorisation has been proposed by the Bank of England in its *Financial Stability Report* on April 26^{th}, 2007: Funding liquidity is defined as the ease which a firm can "meet its cashflow needs", market liquidity broadly as the ease with which securities.can be bought and sold in the market place.

> "margin calls, as they were implemented during this period, were one factor that reduced market liquidity, especially in the futures markets, and likely contributed to the severity of the decline."

This mechanism, dubbed the 'margin spiral', has also been highlighted by Brunnermeier and Pedersen (2008) as a major factor behind the self-reinforcing dynamics in market liquidity. Thus, conceptualising an empirical intraday (time series) approach to market liquidity seems well warranted both from a trading as well as risk management/monitoring point of view:

In this chapter, we device such a model on the basis of the 'Conditional Autoregressive Liquidity' (CARL) model by Reusch (2008). While we keep the multiplicative econometric setup, we modify the liquidity measure used in the model for use in an intraday context: As a derivation of the Hui-Heubel (HH) ratio used in the original CARL is problematic in the context of tick-by-tick intraday data, we propose a filtering technique based on buy and sell volumes to partition the data into volume durations. We then derive the maximum percentage range measure, the difference of the highest buy less the lowest sell price as a percentage of the latter over these durations as our liquidity proxy for use in the CARL model: The result is an autocorrelated, positive series that shows 'liquidity clustering', similar to the behaviour of financial market volatility and in line with theoretical research on liquidity.

Using the derived metric as the liquidity proxy in the CARL setup, the resulting intraday model approaches the modelling of liquidity with a multi-dimensional perspective: Whereas other prominent intraday models of liquidity such as the 'Autoregressive Conditional

Duration' (ACD) model by Engle and Russell (1998) view liquidity along a single dimension such e.g. the immediacy of financial transactions as measured by the time span between successive trades, the intraday CARL approach, with its more complex liquidity measuring procedure, captures multiple facets of liquidity such as market tightness (the difference between selling and buying prices), market depth and breath (how much volume can be transacted at certain prices) as well as resiliency (how well is market able to digest transaction volumes without distortions).

In this chapter we formally define the intraday CARL model and lay out in detail the filtering procedure for the maximum percentage range liquidity measure. We then put the model to use in an empirical application using intraday data on the Amazon stock: We derive the best fitting model for data obtained with various filter sizes and demonstrate that the intraday CARL model is able to pick up the correlation structure in the data very well, rendering it a valid reduced-form empirical implementation of more complex theoretical liquidity models (such as for example Brunnermeier and Pedersen, 2008). We also find that the model forecasts well in-sample, which constitutes a rather encouraging result for practical applications in financial markets.

The rest of the chapter is organised as follows: We present the background on theoretical and empirical research in intraday market liquidity as well as on the econometric modelling in the next section. In section 4.3 we introduce and motivate the maximum percentage range intraday liquidity measure as well as the data filtering technique to obtain this metric. We then embed the measure in the context of the CARL model to yield our intraday version of the CARL. Section 4.4

presents the Amazon intraday data used for the empirical application and demonstrates the filtering procedure. The results of the empirical application with the filtered data series are shown in section 4.5. We conclude the chapter in section 4.6.

4.2 Intraday Market Liquidity

The analysis of (market) liquidity has seen rich research activity over the recent years, both on the theoretical as well as on the empirical side[2]. While most studies focus on the sources of liquidity as well as its linkages across markets (often dubbed 'commonality', cf. e.g. Chordia et al., 2000*a*,*b*, 2001; Huberman and Halka, 2001), little research has been carried on the time series properties of liquidity as well as on the associated empirical modelling of such properties for practical use.

Yet, recent theoretical research, e.g. by Morris and Shin (2004) and Brunnermeier and Pedersen (2008) suggests that liquidity in markets ought to exhibit autoregressive behaviour, with spells of illiquidity followed by further illiquidity and vice versa. Specifically, Brunnermeier and Pedersen (2008) show that illiquidity is subject to self-reinforcing processes, dubbed 'loss spirals' and 'margin spirals': According to the latter, during time of relative market illiquidity, the price impact of orders is higher, leading to heightened volatility in markets, which in turn increases margins on existing positions. These higher margins then potentially create funding problems that trigger the unwinding of positions or reluctance to take new positions in the market, further dampening liquidity. Similarly, in the case of the loss spiral, losses on existing positions lead yet to more funding problems, which in turn

[2]An fairly comprehensive overview can be found in Reusch (2008).

forces traders to liquidate positions and withdraw further liquidity from the market place. Thus, in Brunnermeier and Pedersen's (2008) setting, illiquidity begets more illiquidity and leads to even worse trading conditions. Furthermore, their model provides an explicit link of illiquidity to volatility, which are mutually reinforcing. Yet another takeaway from their framework for empirical liquidity modelling is that, mainly in the case of equities, illiquidity problems, the selling of positions and falling prices go hand-in-hand.

This point is further highlighted by Morris and Shin (2004) who propagate the concept of 'liquidity black holes' that come into existence once the price has fallen below a certain trigger threshold, which in their model can happen even without any apparent fundamental reason but merely as a result of a shock: Once there, driven by a concept analogous to Morris and Shin's (1998) global games approach, in their model the selling of positions by market participants then sparks further sell-offs by others, creating a (rapidly) falling market in which liquidity is being eroded and transactions cannot be concluded at 'reasonable' prices warranted by fundamentals. As in their model the liquidity black hole occurs without any fundamental reason, after falling for a 'sufficient' amount, the price eventually aligns again with the fundamental value and market liquidity is restored to 'normal' levels, creating 'v-shaped' price patterns. From an empirical modelling point of view, Morris and Shin (2004) thus provide yet another rationale for the existence of autoregressive patters in liquidity as well as a close association of illiquidity with falling prices. While they do stipulate the existence of feedback loops along the lines of Brunnermeier and Pedersen (2008), they also supply a reasoning for the reversion of liquidity to normal (average) levels following spells of strong illiquidity.

Being theoretical models of liquidity dynamics, neither Morris and Shin's (2004) nor Brunnermeier and Pedersen's (2008) approach explicitly specifies on which data frequency their results ought to hold - which is however important for the purpose of empirical modelling. Yet, their findings and model setup suggest that autoregressive liquidity patterns are characteristic both of intraday 'tick-by-tick' trading as the immediate outcome of the strategic interaction of market participants as well as of longer run market activity, e.g. when viewed on a daily, weekly or even monthly basis.

While the time series modelling of market liquidity using daily and weekly sampling has been addressed by Reusch (2008), this chapter is concerned with the empirical analysis of intraday liquidity.

For this purpose, we build on and modify Reusch's (2008) 'Conditional Autoregressive Liquidity' (CARL) model for the use in an *intraday* context. While we present our approach as well as a review of the CARL model and its properties further below, we now turn to briefly laying out some of the existing intraday liquidity modelling methodologies.

4.2.1 Econometric Model Background

One of the most well-known econometric time series models of intraday market liquidity is the 'Autoregressive Conditional Duration' (ACD) model by Engle and Russell (1998): They postulate a multiplicative, autoregressive process on a positive domain to describe the time series behaviour of trade durations, i.e. the time intervals that pass between successive trades in the marketplace. As such the ACD model represents a so-called 'marked point process' with 'conditionally orderly after-effects', i.e. a counting process characterised by a sequence of (stochastic) event arrival times[3], in which the arrival of new events depends on the history of event arrivals before and, loosely speaking, events occur successively, one at a time - with arbitrarily small time intervals in-between.

Naturally, such a process can also by represented by a 'conditional intensity' process or 'hazard function', i.e. a time series process that indicates the event rate at any point in time, conditional upon the history up to that point. In the case of the ACD model, such an event would be the end of a duration, i.e. a new trade. Durations can therefore be modelled both via the hazard function, with a popular approach there being the 'proportional hazard' (PH) model where the hazard function is taken to be a 'baseline hazard' multiplied by a function that depends on lagged durations and/or other exogenous variables, as well as by modifying the time scale directly, as in the 'accelerated time' (AT) time series approach used by Engle and Russell (1998). In the latter, the duration is modelled as a baseline duration distributed on a positive domain with a unity expectation times a function depending on lagged durations and/or other exogenous variables.

Over the recent years many extensions and refinements have been proposed to duration models, using both the PH and AT modelling approaches[4]: Work carried out in the latter category includes inter alia Ghysels and Jasiak (1998), Bauwens and Giot (2000), Bauwens and Giot (2002), Veredas et al. (2002), Bauwens et al. (2003), Bauwens et al. (2004), Bauwens and Veredas (2004) and Grammig and Fernan-

[3]In their case an event is a new transaction.

[4]A comprehensive overview can be found e.g. in(Bauwens and Giot, 2001).

des (2005, 2006). PH-type modelling techniques for intraday durations have been used by Gerhard and Hautsch (2000), Grammig and Maurer (2000), Gerhard and Hautsch (2002), Hautsch (2002), Heinen (2003), Bauwens and Hautsch (2006) as well as Gerhard and Hautsch (2006). Even more complex approaches, such as combining an ACD-type model with a Markov-switching setup by Hujer et al. (2005) or the inclusion of unobserved an stochastic components along the lines of stochastic volatility into a duration model by Gouriroux et al. (2004) have been proposed.

While these models have increased the understanding of the intraday trading process and facilitated testing of various hypotheses of market microstructure (as proposed e.g. in Engle and Russell, 1998), they are naturally limited to addressing liquidity along the single dimension of *immediacy*[5], i.e. the timeliness with a transaction can be executed

Yet, as documented by various authors (cf. e.g. Amihud, 2002, p. 35) liquidity in financial markets is hardly captured by one notion only, but is rather a more complex, multidimensional concept. Other dimensions of liquidity include for example the 'depth' (an abundance of orders close to the current price), 'breadth' (large volumes can be executed without much impact on prices) or 'resiliency' (fast corrections of price distortions from fundamentals) of a market, which the ACD-type models of trade durations do not capture.

While Bauwens and Giot (2001, pp. 45-48) for example propose to use ACD models on filtered 'price' or 'volume durations, defined respectively as the time that elapses between price moves of a certain size

[5]For an overview of the different dimensions of liquidity cf. e.g. Sarr and Lybek (2002) or Reusch (2008).

or until a certain order volume has been executed, this approach still relies on the premise of modelling liquidity as the waiting time between events: Thus, even though this ACD methodology incorporates other liquidity-relevant information, it primarily provides a one-dimensional liquidity measure in terms of immediacy.

Yet, despite the obvious limitation of modelling liquidity only with (filtered) durations, the price duration methodology has proved to be a helpful building block for Engle and Lange's (2001) 'VNET' model, which is related to the approach taken in this chapter: They propose the VNET intraday measure of liquidity which they define as the net buying volume in a security over a time interval during which the mid-quote price moves by a specified amount. They thus filter mid-quote prices into price durations, over which in turn the VNET net buying volume measure is defined. Linking price impact and volumes with VNET, they provide a multi-dimensional liquidity measure that addresses market depth. A comparatively low VNET for example is indicative of a shallow and thin market, whereby comparatively little volume is executed per (fixed) price change. Or, viewed from a different perspective, only comparatively low net buying volume is needed to move prices.

Using VNET on intraday cross-sectional data extracted from the 'Trades, Orders, Reports and Quotes' (TQRQ) database covering 144 NYSE stocks over the three months period from 01/11/1990-31/01/1991, Engle and Lange (2001) find that the VNET measure varies considerably over time, yet is highly autocorrelated and can be forecast and explained by volumes, transaction numbers, bid-ask spreads as well as expected price durations. Re-estimating the model with data taken from 'Trades and Quotes' (TAQ) database over the five months period between August and December 1997, they also show that the functional relationship between the VNET measure and its explanatory variables is relatively stable over time with the change in the estimated parameters being relatively modest from the early to the later period despite changes in the trading environment, minimum tick sizes as well as an increase in general trading activity.

They also use VNET to estimate market reaction curves, i.e. the VNET as a function of different price change threshold filters, showing that higher VNET metrics are associated with larger price changes. By construction, the fixed, exogenous variable of choice in this procedure is the price change threshold, which is then associated with VNET to give a measure of market depth.

While similar in spirit to VNET, this chapter takes the *opposite* route, i.e. associating a certain fixed transaction volume with a variable price change. We outline this procedure, its benefits and the tie-in with the CARL model in the next section.

4.3 Intraday Liquidity à la CARL

As mentioned above, the analysis in this chapter builds on the 'Conditional Autoregressive Liquidity' (CARL) model by Reusch (2008) and adopts it for an intraday context. As such, the approach taken is much related to both the ACD model described above - as CARL shares the same econometric foundations - as well as to Engle and Lange's (2001) VNET model: For the intraday context we adopt an approach of deriving a liquidity measure based on the 'Hui-Heubel' (HH) ratio used in the CARL model, whereby we filter the intraday data into volume durations first and then derive a metric on the filtered data set, a method that is similar to VNET, where data is filtered into price durations first over which the net volume is calculated. The procedure is outlined in more detail below.

4.3.1 The HH Ratio Revisited

The CARL model makes use of the Hui-Heubel ratio (described in Sarr and Lybek, 2002) as a measure of liquidity. Contrary to e.g. durations, that describe liquidity along the single time dimension, the HH ratio is a *composite* liquidity measure: It is defined as

$$l_t = \frac{\frac{p_t^{\max} - p_t^{\min}}{p_t^{\min}}}{\frac{v_t}{\bar{p}_t \cdot s_t}}, \tag{4.1}$$

where

$p_t^{\max}$ ≡ the maximum transaction price over the period
$p_t^{\min}$ ≡ the minimum transaction price over the period
v_t ≡ the traded currency volume over the period
$\bar{p}_t$ ≡ the average price over the period
s_t ≡ the average outstanding number of shares over the period

Conceptually, the HH ratio thus consists of the percentage range divided by the price-weighted turnover in the security over some period. As such, it addresses market liquidity along multiple dimensions: The numerator renders a feel for the tightness of the market, while the denominator might be considered a measure of market depth as well as breadth. In combination, the ratio gives a feel for the resilience of a market, i.e. how much a market might 'gap' when a certain (price-weighted) volume is executed[6].

[6]For a more complete discussion of the merits and properties of the HH ratio as a multi-

Yet, while the HH metric conveys a lot of information, it is by construction quite simple and therefore should be readily computable from commonly available datasets for various asset classes. Moreover, especially in comparison to other liquidity measures, such as e.g. durations, the HH metric can be used on various data frequencies: Reusch (2008) demonstrates the use of the HH ratio in CARL model using daily as well as weekly data; it can be considered a natural step to extend this kind of analysis to the growing field of intraday, high-frequency data analysis as well. Closer inspection, however, reveals that such an extension is not quite as straight-forward and needs a few modifications as shown in the next section.

4.3.2 The Maximum Percentage Range Measure

As shown above, the HH ratio is defined over a specific period, which in the case of equally-spaced data (e.g. daily or weekly) does not pose any difficulty for time series analysis. Intraday data, however, is usually not arranged in fixed time intervals, but is generated on a 'tick-by-tick' basis, with varying durations between ticks stemming from the trading behaviour of market participants. In order to construct an intraday HH ratio series and perform an analysis with the CARL model, however, one needs data that is partitioned into intervals over which the HH ratio can be derived.

While an obvious method in this context is equidistant sampling of intervals of a fixed, exogenously chosen length s, as used for example by Andersen and Bollerslev (1997), we have decided against this route as such a methodology often raises 'aggregation' issue about the appropriate choice of s, which we seek to avoid in this context.

dimensional liquidity measure the reader is referred to Reusch (2008).

Instead, we propose to partition the intraday data into intervals during which a volume of size m is both bought as well as sold. Over these intervals we then calculate the maximum percentage range, i.e. the maximum buying price less the minimum selling price as a percentage of the latter, yielding the series

$$r_{m,t} = \frac{p_{BUY,t}^{\max} - p_{SELL,t}^{\min}}{p_{SELL,t}^{\min}}, \tag{4.2}$$

where

$$\begin{aligned} m &\equiv \text{volume bought and sold over the interval} \\ p_{BUY,t}^{\max} &\equiv \text{maximum buying price recorded over the interval} \\ p_{SELL,t}^{\min} &\equiv \text{maximum selling price recorded over the interval.} \end{aligned}$$

Constructing the metric $r_{m,t}$, which we employ as the intraday measure of liquidity in the CARL model, in this way has several advantages:

- Like the HH ratio, $r_{m,t}$ is positive and, as a percentage measure, independent of tick sizes.

- Paired with the information about the buying and selling volume m, the metric captures multiple dimensions of intraday market liquidity, notably depth, breadth, tightness as well as resilience. For example, a high $r_{m,t}$ associated with a low m is indicative of a shallow, narrow, not very tight and resilient market.

- In spirit, the procedure for deriving $r_{m,t}$ is similar to the VNET approach, albeit our methodology takes the opposite route: Instead of filtering the data with a fixed price difference threshold and then calculating the net volume over the durations, we use a fixed *volume filter first* and then derive a percentage price range over the so-obtained intervals. We believe that conceptually, this way of doing things is closer to the perspective that market participants, especially traders, take: A trade is usually initiated with a known transaction volume in mind and the transaction price for the volume is determined subsequently during the actual trading process - (partly) as a function of the volume size, i.e. the causality is seen as running from volume to price change.

- By construction, $r_{m,t}$ uses real transaction prices instead of the artificial mid-quote point proxy. Similar to above, we believe that this provides close association with the actual trading process and thus makes the metric a more realistic measure of liquidity. Obviously, the use of returns from transaction prices at ultra high frequencies has been shown to be problem-laiden, due to market-microstructure effects such as the 'bid-ask bounce' (cf. Roll, 1984). While we turn to these issues in more depth in section 4.4 below it suffices to say here that our metric is not much affected by such complications as they mainly affect successive returns, whereas - loosely speaking - we are concerned with the maximum 'gap' of transaction prices that opens up over a filtered interval. Moreover, in our case, we also classify trades into buy and sell transactions, which further alleviates the problem (cf. Goodhart and O'Hara, 1997, p. 95).

- As part of the filtering process, we also derive the net sell size or 'order imbalance' series ($\text{OIB}_{m,t}$) as it also often called (cf. e.g. Chordia et al., 2002) over the intervals during which the volume m is sold and bought. This metric is constructed as the sell minus the associated buy volume over the filtered intervals and can be viewed as a useful 'by-product' of our methodology as it is indicative of market moves: A large positive $\mathrm{OIB}_{m,t}$ for example indicates a strong downward pressure on price. By itself, this by-product measure allows for a variety of interesting analyses and tests. Chordia et al. (2002) for example find that the absolute value of order imbalance is negatively associated with liquidity, i.e. $\mathrm{OIB}_{m,t}$ should be positively correlated with $r_{m,t}$. In accordance with the theoretical findings discussed above, one would also expect the metric $\left|\mathbb{I}_{\mathrm{OIB}_{m,t}>0} \cdot \mathrm{OIB}_{m,t}\right|$, i.e. the unsigned net *selling* volume that is taken to be zero when there is net buying, to be correlated with the liquidity measure $r_{m,t}$ as excessive selling can viewed as a component of Brunnermeier and Pedersen's (2008) 'loss spiral', leading to further illiquidity. Also, a liquidity black hole is characterised by rapid selling and fast-falling prices, which should also translate into such a correlation pattern.

- Contrary to VNET, where net volume is associated with a price change threshold without specifying the overall volume, our methodology thus provides three pieces of intraday liquidity information: The metric $r_{m,t}$, the associated buy and sell volume m, as well as the $\mathrm{OIB}_{m,t}$. This has the benefit of being able to make finer distinctions between liquidity situations. For example, while a VNET of some value $a > 0$ can ambiguously be the result of a both a sell size $(n + a)$ or $(k + a)$ and a buy size of magnitude n or k, with $n \neq k$ and $n, k > 0$ over some filtered interval respectively, $r_{m,t}$ is supplemented both by the information of the level of m as well as $\mathrm{OIB}_{m,t}$, allowing for a more informed analysis.

- Similar to the HH ratio, the meaningfulness of $r_{m,t}$ as liquidity measure is doubtful if price moves (in large jumps) occur purely as a consequence of news announcements. As discussed in more depth in Reusch (2008) however, Engle and Lange (2001, pp. 17) indicate that pure information-based trading is quite rare, and surely does not occur at all times during the day, which encourages us in our view of the usefulness of $r_{m,t}$ in an intraday context.

- Loosely speaking, the metric $r_{m,t}$ may also be interpreted as a liquidity risk measure along the lines of 'Value-at-Risk' (VaR_α): Whereas the latter constitutes the minimum loss that one could incur with a probability α on a bad day, $r_{m,t}$ can be viewed as the *maximum* percentage loss that one could have realised over an interval while trading (buying and selling) a volume m[7].

The interpretation laid out in the last point above is obviously helped by the fact that contrary to the HH ratio in Reusch (2008), we do not divide $r_{m,t}$ by price-weighted turnover over the filtered intervals. The reason is that we feel that dividing by 'unsigned' price-weighted turnover, i.e. ignoring the distinction between buy and sell volume, would water down the information content of the metric and not add more informational value, especially given that the such a distinction is vital for the derivation of $r_{m,t}$ in the first place. Similarly, dividing by (price-weighted) net selling turnover seems quite meaningless and in practice would ever so often entail a division by a number close to zero, thus distorting the meaningfulness of the metric.

Therefore, despite the difference in construction, $r_{m,t}$ being a positive metric, independent of tick-size, is very much akin to the HH

[7] Strictly speaking, this maximum loss over the interval at time t would materialize if one bought the entire volume m at $p^{\max}_{BUY,t}$ and sold at $p^{\max}_{SELL,t}$.

ratio both in statistical terms and in practical applications as the input to the intraday CARL model as shown below. We now turn to introducing the latter.

4.3.3 The CARL Model Revisited

As mentioned above, the maximum percentage range, like its HH ratio 'cousin', is a positive liquidity measure that is independent of tick-size. As such it is then possible to use it in place of the latter in Reusch's (2008) CARL model for intraday applications. The resulting intraday CARL(p,q) model looks as follows:

$$
\begin{aligned}
r_{m,t} &= \lambda_{m,t}\varepsilon_t, \text{with } \varepsilon_t \sim i.i.d., \mathbb{E}[\varepsilon_t] = 1; \forall t \in \{0, ..., T\} \text{ and} \\
\lambda_{m,t} &= \omega + \sum_{i=1}^{p} \alpha_i \lambda_{m,t-i} + \sum_{j=1}^{q} \beta_j r_{m,t-j}, \text{ with } \omega > 0; \alpha_i, \beta_j \geq 0, \forall i, j.
\end{aligned}
\tag{4.3}
$$

Statistically, the following properties hold:

$$
\begin{aligned}
\mathbb{E}[r_{m,t}|\mathcal{F}_{t-1}] &= \lambda_{m,t} \\
Var(r_{m,t}|\mathcal{F}_{t-1}) &= \lambda_{m,t}^2 \\
\mathbb{E}[r_{m,t}] &= \frac{\omega}{(1 - \sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j)},
\end{aligned}
$$

where $\mathcal{F}_{t-1}$ is the information set prevailing at time $t-1$. Moreover, in the simple CARL(1,1) case one obtains for the unconditional variance

$$
Var(r_{m,t}) = \frac{1 - \alpha - 2\alpha\beta}{(1 - \alpha - 2\alpha\beta - 2\beta^2)}.
$$

Reusch (2008) contains a more detailed discussion of the properties, features and estimation of the CARL model. Qualitatively, we would like to point out here that because of its multiplicative, autoregressive structure, the above intraday CARL model is well capable of the empirical modelling of the before mentioned self-reinforcing behaviour of market liquidity. We now turn to exploiting this property in an empirical application.

4.4 Data

The empirical application in this chapter focuses on intraday 'Trades and Quotes' (TAQ) data of Amazon, the online retailer on 24/07/2001: The stock, which can be considered a 'new economy blue chip' has been listed on the NASDAQ since 15/05/1997 and enjoys active and liquid trading. On 09/04/2001 decimalised trading in the stock was introduced.

Contrary to most intraday studies that usually cover multiple days of data, we concentrate the analysis on a single day. While we recognise that adding more days would allow for a richer application, we would like to stress that the aim of this chapter is a different one: First and foremost, we seek to establish the new methodology of the intraday CARL model and exemplify its merits as well as potential drawbacks in a show-case empirical application. A detailed empirical study, potentially covering a cross-section of securities over several days is outside of the scope of this work and is left for later work. Moreover, the focus on one day further allows to curtail the analysis to a manageable dataset - without having to worry about issues such as how intraday data stemming from different days is to be joined together or the re-occurrence of intraday seasonal effects.

In selecting the 24/07/2001 as our exemplary dataset, we have paid particular attention to a few desirable characteristics that the data on this day exhibits: We have picked the day on the basis of the occurrence of a large negative return, indicative of potential turmoil and sufficient trading activity in the stock. Moreover, liquidity black-holes along the lines of Morris and Shin (2004) are explicitly linked to (large) negative returns, which is why the day, bearing the largest price drop over the 2001-2007 period, is of special interest. Also relevant in the light of the above is the fact that most of the negative return occurred in the pre-market trading session before the official market 9:30 am opening, indicating that information was already in the market before the start of official trading, thus limiting the effect of pure, information-based trading during the day. Finally, the 24/07/2001 is characterised by a fairly large trading volume and a very large number of transactions on the day:

Figure 4.1 highlights the characteristics of the day relative to the 2001-2007 period. Summary statistics of the data on the day can be found in table 4.1.

The left hand column displays information relating to the 'raw' data on the day as obtained from the TAQ database. The right hand column shows the characteristics of the data after a cleaning procedure along the lines of Brownlees and Gallo (2006): Following their guidelines, for both transaction prices as well as quotes, we eliminate all negative prices as well as outliers, defined as data whose discrepancy from the 30-day forward and backward looking moving average lies outside the corresponding 3-standard deviation bound. We also purge quotes that seem to be generated by non-normal market activity, as

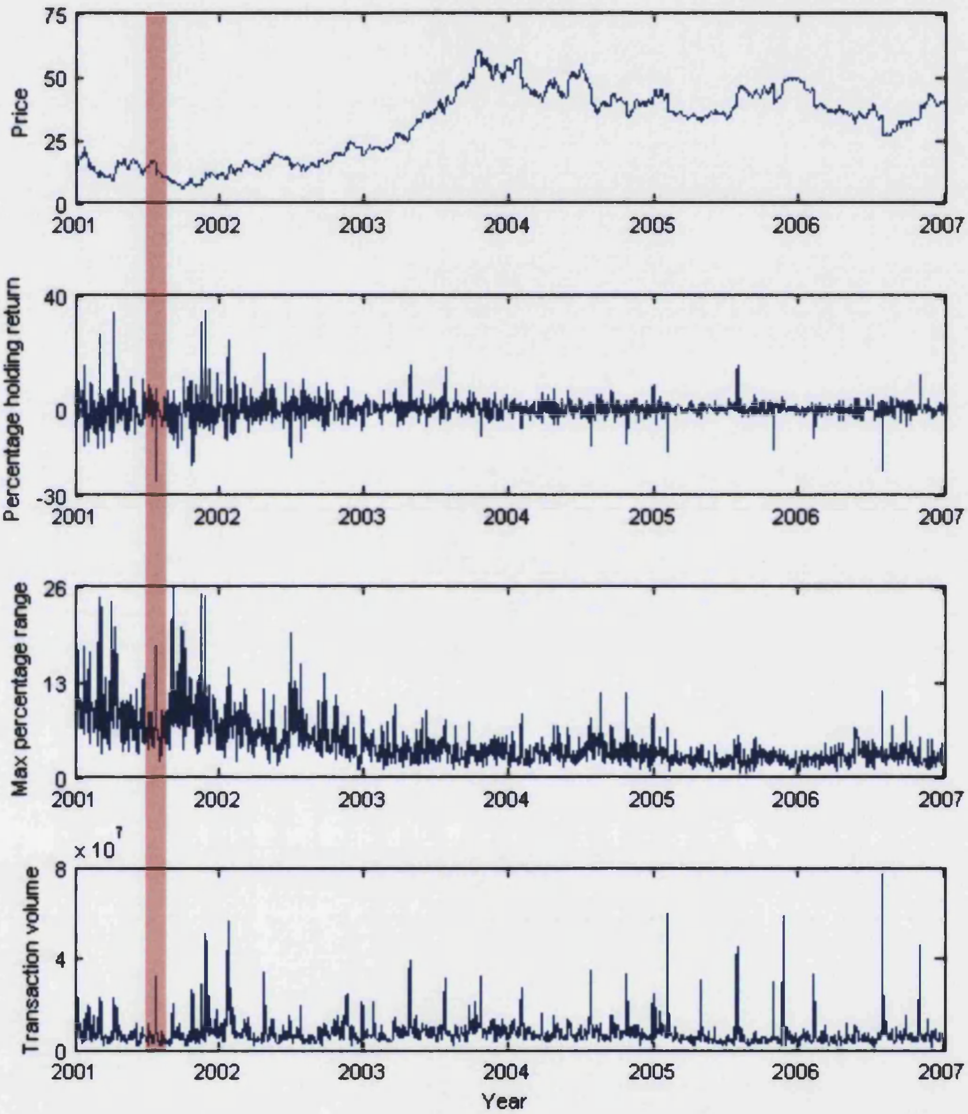

Figure 4.1: **Amazon daily price, percentage range, return and volume**

This figure displays the daily Amazon CRSP tape closing price (panel 1), maximum percentage range (panel 2), holding return (panel 3) and transaction volume (panel 4) from 02/01/2001-29/12/2007. The 24/07/2001 is highlighted.

Table 4.1: **Intraday Amazon data statistics (24/07/2001)**

This table shows statistics of TAQ intraday Amazon data covering 24/07/2001. Listed are both statistics for unadjusted raw as well as data data adjusted for outliers and 'erroneous' trades/quotes. The daily percentage return is the holding return calculated with closing prices of the actual and previous day.

Day	**24/07/2001**	
	Unadjusted	**Adjusted**
First transaction time	08:00:04	09:30:01
Last transaction time	18:25:59	16:10:28
First transaction price	14.84	13.61
Last transaction price	12.28	12.06
Maximum transaction price	15.06	13.65
Minimum transaction price	11.90	11.91
# of transactions	36,770	34,753
# of quotes	17,255	9,906
Shares traded	33,381,400	25,719,100
Percentage turnover	9.20	7.09
Closing price	12.06	
Previous day closing price	16.03	
Daily percentage return (using closing prices)	−24.77	

indicated by TAQ MODE values '1','2', '3', '6' and '18' (cf. TAQ documentation). Contrary to the suggestions by Brownlees and Gallo (2006) we do however not eliminate quotes on exchanges other than the NYSE. This mainly due to the fact that we are dealing almost exclusively with NASDAQ quotes, for which their procedure does not apply and also because we aim to refrain from decimating the quote 'pool' too much. Trades are included if they are reported regular and non-corrected according to TAQ code '0'; we eliminate trades if they occur outside of 'normal' NASDAQ trading hours, i.e. before 09:30 and after 16:00 and if they are reported late according to TAQ COND code 'Z'. We do however admit 6 trades that occurred after 16:00, but which are transactions that actually relate to an "obligation to trade at an earlier point in the trading day or that refer to a prior refer-

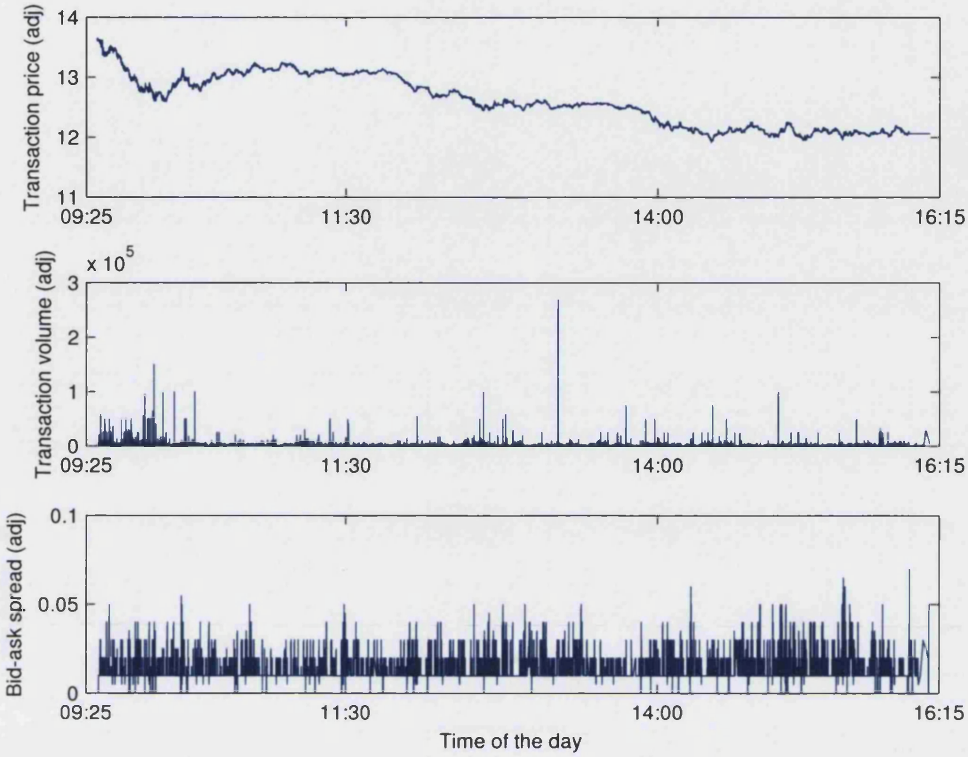

Figure 4.2: **Amazon adj transaction price, volume and bid-ask spread**

This figure shows the Amazon adjusted transaction price (panel 1), transaction volume (panel 2) and bid-ask spread (panel 3) on 24/07/2001 during NASDAQ market hours. The series have been derived from TAQ intraday data with adjustments made for outliers and 'erroneous' observations,.

enced price" - in this case the official closing price - according to TAQ COND code 'P'.

Applying these cleaning procedures we eliminate around 2,000 transactions and a bit less than half of the c. 17,000 quotes. Most of the purged data is recorded before the official opening of the market, also revealing that most of the negative return on the day occurred around this time, as mentioned above. Figure 4.2 gives a graphic representation of the cleaned data.

4.4.1 Filtering

As mentioned above we apply a volume filtering technique to construct the maximum percentage range metric $r_{m,t}$ from the cleaned TAQ 'raw' data.

To this end, we first classify trades into sell and buy transactions. We do so using a modified variant of Lee and Ready's (1991) classification algorithm, according to which we classify a trade as a buy (sell) if the transaction price is higher (lower) than the mid-point of the prevailing bid and ask quotes. If a trade happens at a price exactly equal to the quote mid-point, we apply the 'tick rule': If the latest transaction price represents an 'uptick', i.e. the latest transaction price is higher than the previous one, we conclude that the transaction was a buy. Similarly, a downward-ticking price represents a sell transaction. In the case that the tick rule applies and the past transaction price is equal to the current one, we apply the tick rule backwards until we find a differing transaction price in order to classify the trade. As suggested by Lee and Ready (1991) we also implement their 'five second rule' whereby we match trades with quotes that are at least five seconds old in order to make sure that quotes, which tend to get revised more quickly than trades are posted, most accurately match the corresponding trades.

Figure 4.3 shows the results of the classification procedure: We note downward trending transaction prices for both buy and sell transactions and the relatively higher magnitude of sell orders in comparison the buys. This is not very surprising, given the overall downward direction of the market on the day. However, as mentioned above, the bulk of the day's selling order volume is recorded before the official market opening and is not displayed here.

In a second step, following the trade classification, we filter the data into intervals over which a minimum volume m is cumulatively both bought as well as sold. Over the so-obtained durations we then calculate our liquidity measure, the maximum percentage range $r_{m,t}$, as the difference between the highest buy price and the lowest selling price as a percentage of the latter.

Obviously, given the local direction of the market, the lengths of the filtered intervals can be quite different: For example, in a strong downward move, a cumulative sell volume m is 'reached' quite easily given the relative abundance of sell orders in a down-trend, while buy orders are harder to 'come by'. Consequently, while we set up the filtering procedure such that a *minimum* volume m is transacted on *both* sides, it may well be that there is an 'overhang' of volume on one side over the course of the filtered interval, as is for example to be expected in a situation such as the one described above. Equally, in a well-balanced, flat market situation filtered durations of this sort tend to be shorter as buy and sell orders are more evenly matched, resulting in less volume overhang on a particular side after filtering. We theme such a volume overhang over the duration at time t the order imbalance OIB_t, calculated as the difference between the actual volume sold and bought over the derived intervals. This series is a by-product of our analysis and can be viewed as a directional indicator: a strongly positive (negative) OIB_t for example indicates selling (buying) pressure on the security. Table 4.2 shows the results of the filtering procedure with a volume filter $m = 500$.

The filtering 'distills' roughly 35,000 transactions to c. 5,700 $r_{500,t}$ data points, each of which is obtained over volume filtered intervals

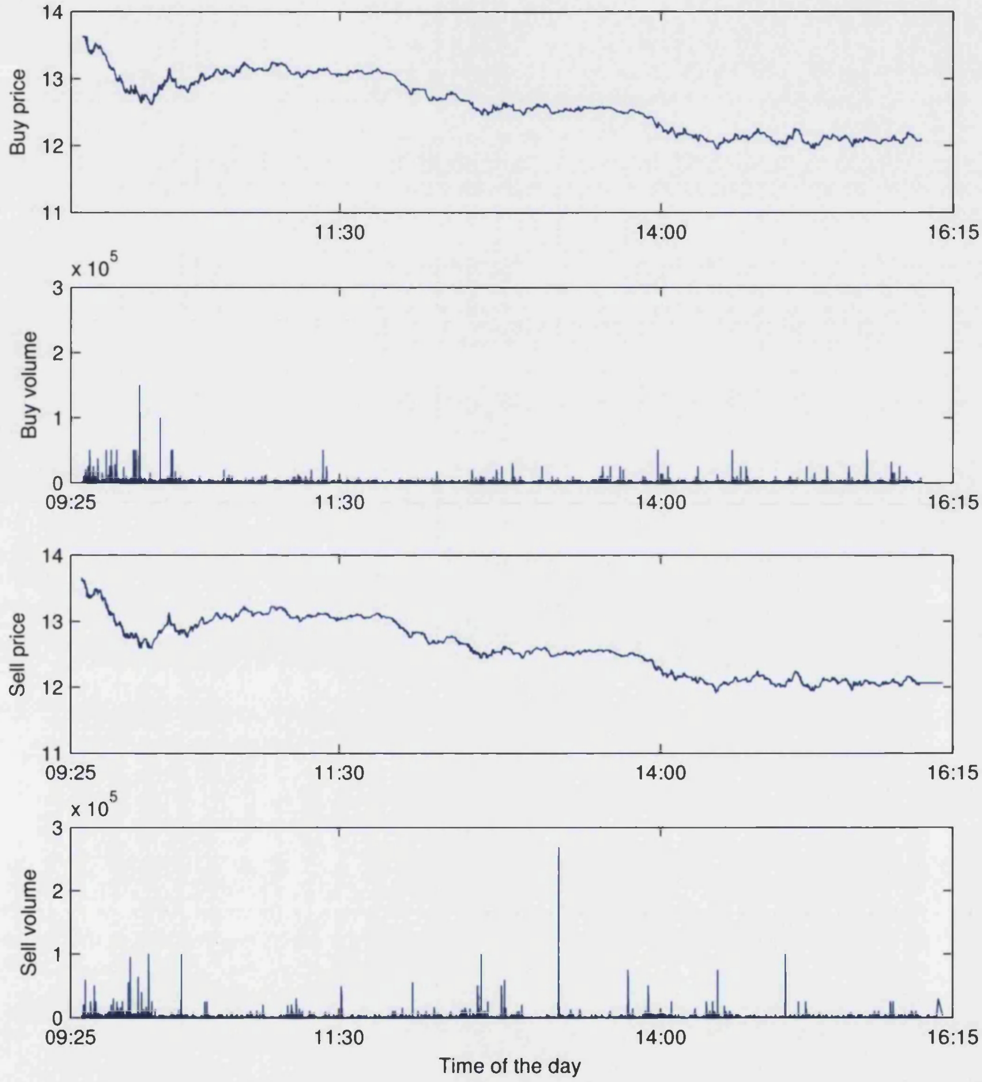

Figure 4.3: **Amazon intraday buy/sell price and volume**

This figure displays the filtered intraday Amazon buy price (panel 1), buy volume (panel 2), sell price (panel 3) and sell volume series (panel 4) on 24/07/2001 during NASDAQ market hours. The series have been obtained from adjusted TAQ data using Lee and Ready's (1991) classification algorithm.

Table 4.2: **Statistics - Filtered intraday Amazon $\mathbf{r_{m,t}}$ series ($\mathbf{m = 500}$)**

This table shows statistics of the filtered intraday Amazon $r_{m,t}$ series on 24/07/2001 during NASDAQ market hours. The series has been obtained using a volume filter size $m = 500$. The Ljung-Box autocorrelation test statistic for the first 50 lags is denoted as LB_{50}. The average filtered duration is the average length of the time interval during which the quantity m is sold and bought and the maximum percentage range realized. The series $\text{OIB}_{m,t}$ constitutes the difference between the filtered sell and buy volumes over the filtered intervals.

Filtered $\mathbf{r_{m,t}}$ series (24/07/2001)	
Transaction size filter	**m=500**
# of observations	5,692
Mean	0.164
Median	0.083
Maximum	1.264
Minimum	0.000
Std dev	0.139
Skewness	2.555
Kurtosis	12.092
$\text{Corr}(r_{m,t}, \lvert\text{OIB}_{m,t}\rvert)$	0.312
$\text{Corr}(r_{m,t}, \lvert\text{OIB}_{m,t-1}\rvert)$	0.118
$\text{Corr}(r_{m,t}, \lvert\mathbb{I}_{\text{OIB}_{m,t}>0} \cdot \text{OIB}_{m,t}\rvert)$	0.186
$\text{Corr}(r_{m,t}, \lvert\mathbb{I}_{\text{OIB}_{m,t-1}>0} \cdot \text{OIB}_{m,t-1}\rvert)$	0.064
LB_{50} (P-value)	8,966 (0.00)
Avg filtered duration (in sec)	4.114

that have an average duration of 4.1s. By construction, the obtained series has a positive domain and is, as expected, non-Gaussian with strong skewness and kurtosis. Moreover, and more relevant for our analysis, $r_{500,t}$ exhibits very strong autocorrelation, as indicated by a highly significant Ljung-Box (LB) statistic. This confirms us in our view that $r_{m,t}$ is a meaningful intraday liquidity proxy, as one would expect liquidity to be autocorrelated based on the findings in the theoretical research on liquidity laid out above. Figure 4.4 displays the autocorrelation function (ACF) of $r_{500,t}$ for the first 50 lags in the lower panel: Well outside the 2-standard deviation bound, the ACF slowly declines from a level bigger than 0.6 and only loses significance after the 25^{th} lag. The top panel of the figure also shows a histogram of the derived maximum percentage range documenting the non-Gaussian features of the empirical distribution.

In terms of magnitude the mean maximum percentage range is 0.16%, which can be interpreted as the expected maximum loss that one could have incurred by simultaneously buying and selling a volume $m = 500$.

We note strong positive correlation between $r_{500,t}$ and the contemporaneous OIB_t series, confirming the results of Chordia et al. (2002) who find evidence for a negative association between liquidity and the absolute value of the order imbalance[8]. While we also detect positive correlation using the lagged OIB_{t-1} series, we note that such an association is weaker; indeed, when we included the variable as a regressor in the CARL model in a test estimation against the above data (not documented here) it loses significance and holds no predictive power.

[8]This follows because the negative correlation involving liquidity translates in a positive correlation for $r_{m,t}$, which is in fact a measure of *il*liquidity.

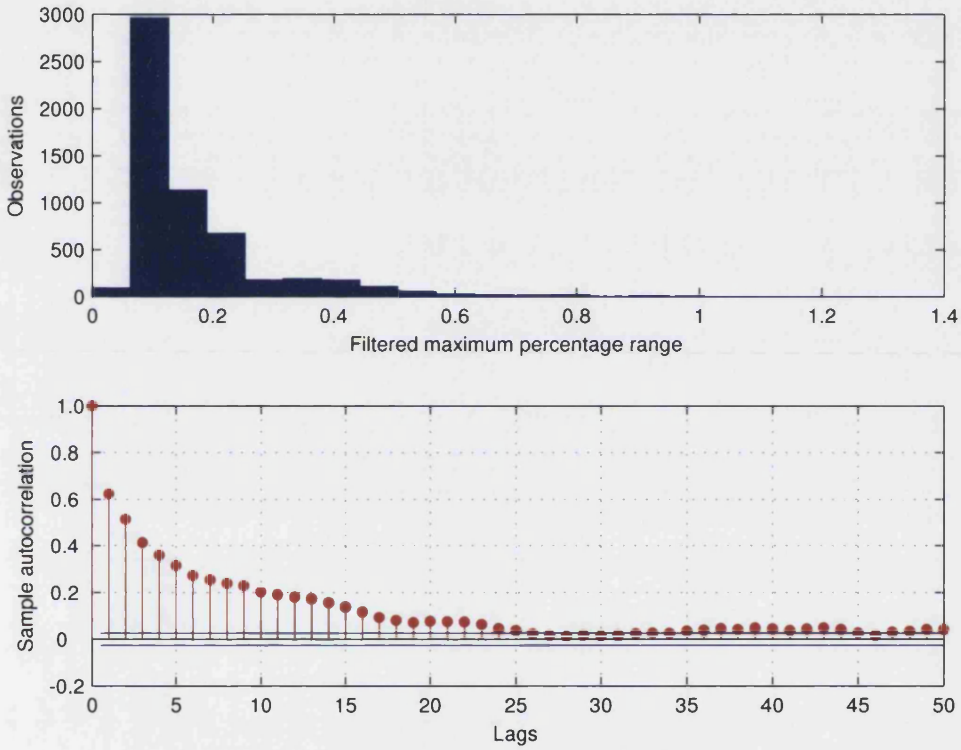

Figure 4.4: **Histogram & ACF - Amazon $\mathbf{r_{m,t}}$ (m = 500)**

This figure displays the histogram (top panel) and the autocorrelation function (ACF; bottom panel) for the filtered intraday Amazon $r_{m,t}$ series on 24/07/2001 during NASDAQ market hours. The series has been obtained using a volume filter size $m = 500$. In addition to the ACF, 2-standard-deviation interval bounds are shown as blue lines, corresponding roughly to a 95% confidence interval.

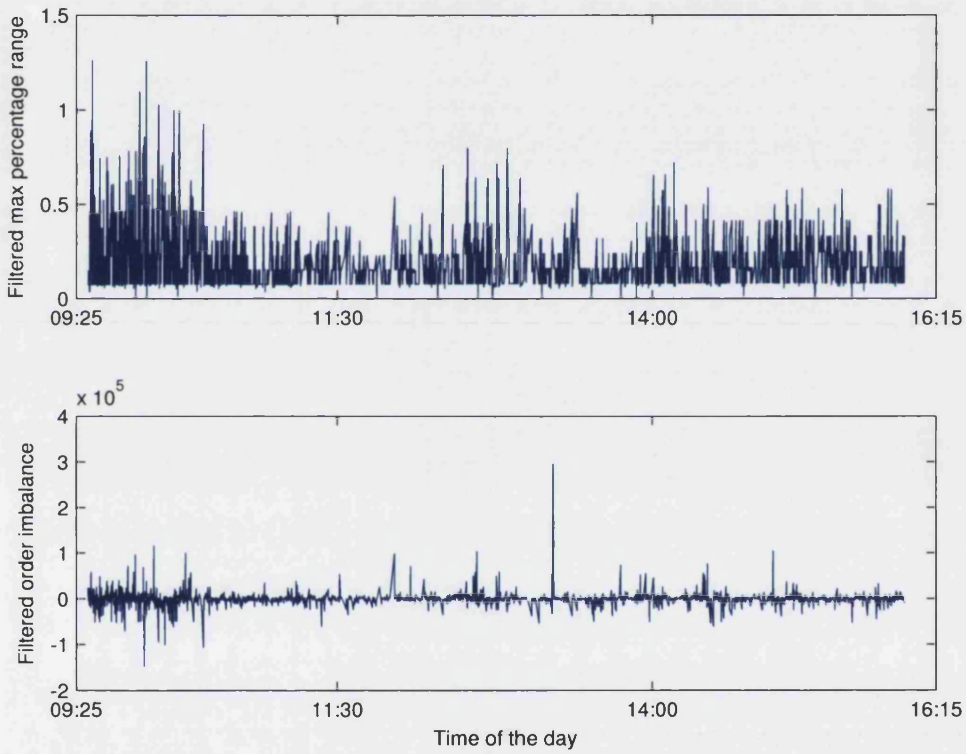

Figure 4.5: **Amazon $\mathbf{r_{m,t}}$ and $\mathbf{OIB}_{m,t}$ (m = 500)**

This figure shows the filtered intraday Amazon $r_{m,t}$ (top panel) and the order imbalance series $\text{OIB}_{m,t}$ (bottom panel) on 24/07/2001 during NASDAQ market hours. The series have been obtained from adjusted TAQ data using a volume filter size m=500.

We also find positive correlation between $r_{500,t}$ and $\left|\mathbb{I}_{\mathrm{OIB}_{m,t}>0} \cdot \mathrm{OIB}_{m,t}\right|$, i.e. the order imbalance series with the absolute value only recorded for a net sell volume overhang and zero otherwise. In the light of the theoretical research mentioned above, we count this as mild evidence towards liquidity loss spirals and the liquidity black hole mechanism, i.e. liquidity shrinking during selling environments.

Figure 4.5 displays the derived $r_{500,t}$ as well as the $\mathrm{OIB}_{m,t}$ series in the top and lower panel respectively. A rough glance at the filtered series shows (il)liquidity clustering, i.e. low liquidity (a high $r_{m,t}$) being followed by yet more low liquidity and vice versa - similar to the well-documented artefact of volatility clustering in financial return series. Obviously, the clustering property is very much linked to the above mentioned autocorrelation structure of the series. The $\mathrm{OIB}_{m,t}$ series depicted in the lower panel shows strong one-sided activity around the opening and during mid-day, whereas other periods of the day are more even in terms of buys and sells.

4.4.2 Market Microstructure Issues

When analysing and handling intraday data the issue of market microstructure effects often arises. Such effects usually comprise (i) intraday-seasonal ('diurnal') patterns in the data, such as e.g. shorter durations between trades or higher return volatility shortly after opening and before trading close compared to mid-day (cf. e.g. Andersen and Bollerslev, 1997), (ii) the well-documented 'bid-ask bounce' Roll (cf. 1984), whereby consecutive sells and buys 'bouncing' between bid and ask prices create the impression of negative serial correlation in asset returns (calculated from transaction prices), or (iii) asynchronous trading issues across securities leading to would-be lead-lag patterns.

With the increased use of intraday both for research as well as for practical purposes, such issues have drawn considerable attention and have been researched extensively. Comprehensive literature surveys on the topic can for example be found in Goodhart and O'Hara (1997) and Biais et al. (2005). The issue of diurnal patterns and market microstructure 'noise' has also been of importance to the recently very prominent field of realised/integrated volatility and its estimation from intraday data. Work in this area has been quite extensive and includes - inter alia - Aït-Sahalia et al. (2005), Andersen et al. (2003), Bai et al. (2004), Bandi and Russell (2006*b*), Bandi and Russell (2006*a*), Bandi et al. (2007), Barndorff-Nielsen et al. (2006) as well as Hansen and Lunde (2006).

While the importance and the complications for empirical analysis arising from market microstructure effects/noise cannot be denied, we would argue that our analysis in this chapter is not particularly prone to such influences. This is mainly due to the fact that we do not directly analyse return series or durations for which most of the above applies. Rather, we concentrate on the filtered maximum percentage range defined above as a proxy for liquidity, which despite being related to returns and volatility exhibits crucial differences: $r_{m,t}$ is defined as the *maximum* percentage difference in buy and sell transaction prices over a filtered interval and, as pointed out above, is therefore more of a limit risk measure than a return metric, which in turn means that for example the problem of the bid-ask bounce is not acute in the context of this chapter. Similarly, asynchronous trading effects should be negligible in the context of our analysis as such issues arise mainly in the context of multivariate analysis - this chapter however takes the route of a univariate approach. Moreover, we follow Lee and Ready's

(1991) advice of trades with quotes that are at least five seconds old in order to ensure that the quote-setting process is properly matched with trading decisions.

Furthermore, our filtering procedure limits possible diurnal issues in the data. Loosely speaking, applying the transaction volume filter effectively amounts to pre-filtering the data according to the breadth/depth dimension of liquidity; the maximum percentage range liquidity measure is then derived over the pre-filtered data, in which any existing diurnality is already 'accounted for'. The following example might serve as an illustration: Assume that a common diurnality pattern exists both for durations and volatilities, i.e. there are certain times of the day such as the opening of trading and shortly before the close during which trading durations, and for that matter also volume durations, are shorter and volatilities higher. During these times, compared to other times of the day, applying a volume filter of size m would result in shorter filtered durations over which the measure $r_{m,t}$ would then be derived. *Ceteris paribus,* one would expect the maximum percentage range $r_{m,t}$ sampled over a shorter intervals also to be smaller than over longer durations - simply because of the very fact that price variation is also a function of time.

This situation contrasts with one in which data is partitioned into equal time intervals over which $r_{m,t}$ can also be derived: In such a setting, the diurnal pattern of volatility will be very apparent as the price variation is given relatively more weight during the more volatile times at the opening and close of the day compared to mid-day, whereas in the case of volume duration pre-filtering the sampling intervals during high volatility 'regimes' would be shortened, limiting the impact of volatility on the liquidity measure. The volume filtering can therefore be viewed as a pre-sampling device that limits the impact of such seasonal patterns.

Figure 4.5 confirms this view. While the maximum percentage range measure exhibits clustering (e.g. around market opening), one would be hard-pressed to detect obvious diurnal patters as for example in the case of durations in Engle and Russell (1998): In their analysis, durations around mid-day are roughly 2 to 13 times higher than around the opening and close of the day respectively and exhibit a clear hump-shaped pattern. The opposite, poignant parabolic shape holds for volatilities as shown for example in Andersen and Bollerslev (1997). Yet, our analysis is not marked by such clear-cut patterns.

Moreover, varying the size of the volume filter constitutes yet another degree of freedom with which the analysis can be fine-tuned in the presence of market microstructure noise: A very small volume filter will in the limit permit tick-by-tick durations, during which the bid-ask bounce issue is likely to become more relevant. A very large volume filter for example can be used to allow entire trading days as durations in which case one is back to a similar situation as the one of a daily sampled data set as in Reusch (2008).

We provide a sensitivity analysis with different sized volume filters in the empirical application section below.

4.5 Empirical Application

In this section we estimate the CARL(p,q) model described above against Amazon intraday data on 24/07/2001. As described in more detail in Reusch (2008), the CARL model with an exponential error distribution can be estimated consistently via 'Quasi Maximum Likelihood' (QML): This in turn also means that one can use conventional software for the estimation of the more established GARCH-model class and estimate the CARL model as a Gaussian GARCH model with $\sqrt{r_{m,t}}$ as the dependent variable. In what follows, we first use the filtered dataset obtained with a volume filter $m = 500$. In the second part, we vary the filter size from $m = 100$ to $m = 10,000$. A discussion of the findings is provided at the end of the section.

4.5.1 Intermediate Volume (m = 500)

We carry out the estimation of the CARL model in the following manner, analogous to the analysis in Reusch (2008): We estimate the CARL(p,q) model with various lag parameterisation combinations of p and q. Specifically we take $p \in \{0, 1, 2, 3\}$ and $q \in \{1, 2, 3\}$. We then compare the results in terms of goodness of fit as measured by the AIC and BIC, as well as by the reduction in the autocorrelation of the standardised residuals after estimation. We then present the model with the best fit alongside a the most simplistic conceivable $p = 0, q = 1$ parameterisation as well as the CARL(1,1) version. We do the latter because in the case of the related GARCH model class, the very simple GARCH(1,1) setup has been found to be superior to more complex parameterisations in a number of applications (cf. Hansen and Lunde, 2005).

Table 4.3 displays the estimation results for $m = 500$.

We note that even though the $p = q = 1$ setup is a favourite for many GARCH applications, it might be argued that it can be improved upon in this context. In our estimation, we would claim that the best overall model fit is provided by the CARL(1,2) model, for which the AIC and BIC are comparable with the (1,1) setup, but the

Table 4.3: **Estimation - Filtered intraday Amazon $\mathbf{r_{m,t}}$ series ($\mathbf{m = 500}$)**

This table shows results from the estimation of different selected CARL models against the filtered intraday Amazon $r_{m,t}$ series on 24/07/2001 during NASDAQ market hours. The volume filter size $m = 500$. The table displays parameter estimates, Bollerslev-Woolridge t statistics, goodness-of-fit measures as well as statistics for standardized residuals. * and ** denote significance at 1% and 5% levels respectively. The Ljung-Box autocorrelation test statistic for the first 50 lags is denoted as LB_{50}.

Model	**CARL(0,1)**		**CARL(1,1)**		**CARL(1,2)**	
Metric	**Estimate**	**t stat**	**Estimate**	**t stat**	**Estimate**	**t stat**
$\hat{\omega}$	0.063*	26.889	0.025*	11.751	0.015*	6.261
$\hat{\alpha}_1$	0.616*	36.941	0.381*	15.641	0.600*	13.634
$\hat{\beta}_1$			0.470*	25.100	0.498*	25.073
$\hat{\beta}_2$					−0.185*	−4.722
Log-LH		−2,669		−2,644		−2,642
AIC		0.939		0.930		0.930
BIC		0.941		0.934		0.935
Standardised residuals statistics						
Mean		1.000		1.000		1.000
Std Dev		0.658		0.623		0.619
LB_{50} (P-value)		316 (0.00)		70 (0.03)		51 (0.43)

reduction in the autocorrelation of the residuals as measured by the Ljung-Box statistic over the first 50 lags (denoted LB_{50} in the table) is much larger. In fact, in the case of the CARL(1,1) model, despite having the similar AIC and BIC measures, one would have to reject the null hypothesis of no autocorrelation over any of the first 50 lags at a 5% significance level as indicated by a p-value of 0.03. For CARL(1,2), however, the autocorrelation is reduced significantly, with a p-value of 0.43 translating into not being able to reject the null at any reasonable level of significance.

When assessing the individual parameter estimates, one finds that for all models, the parameters are of commensurate magnitude and individually highly significant as measured by the robust t statistics reported in the table. Moreover, the estimates translate into covariance stationarity for all models, with $\hat{\alpha}_i + \hat{\beta}_j < 1, \forall i, j$. Moreover, the models also adhere to the positivity condition, i.e. the requirement to have $\lambda_{m,t} > 0$, because $0 < \hat{\alpha}_1 < 1, \hat{\beta}_1 \geq 0$ and $(\hat{\alpha}_1\hat{\beta}_1 + \hat{\beta}_2) \geq 0$ hold (the latter is relevant for CARL(1,2); cf. Nelson and Cao, 1992).

We also put the 'optimal' CARL(1,2) model to the test in a forecasting application according to Mincer and Zarnowitz (1969) (MZ). Following their procedure we assess the optimality of the in-sample forecast obtained with the chosen model by deriving the forecast series $\hat{r}_{500,t}$ with the fitted model and running the following OLS regression:

$$r_{500,t} = \theta_0 + \theta_1 \hat{r}_{500,t} + \nu_t.$$

The MZ test then allows for the assessment of the quality of the forecast by performing a joint parameter restriction Wald test of the sort $\theta_0 = 0 \cap \theta_1 = 1$. The null in this case is the optimality of the forecast, in the sense that the forecast is sufficient to predict the dependent

Table 4.4: **MZ - Filtered Amazon $\mathbf{r_{m,t}}$ series (m = 500)**

This table shows a Mincer-Zarnowitz regression of the form $r_{m,t} = \theta_0 + \theta_1 \hat{r}_{m,t} + \epsilon_t$ for the filtered intraday Amazon $r_{m,t}$ series on 24/07/2001 during NASDAQ market hours. The forecast series $\hat{r}_{m,t}$ is obtained using the CARL(1,2) model, for which the in-sample fit is found optimal.

MZ regression		
Metric	**Estimate**	**t stat**
$\hat{\theta}_0$	0.003	0.778
$\hat{\theta}_1$	0.991*	31.925
Restriction test ($H_0 : \theta_0 = 0 \cap \theta_1 = 0$)		
Wald test stat (P-value)	0.818 (0.66)	

variable. It should be noted, however, that the test does not produce a ranking of forecasts nor does it quantify the quality of the forecast. It may therefore not be used to judge one forecasting model against another one.

The results of the MZ test for the CARL(1,2) in the context of the data series obtained with a volume filter $m = 500$ are reported in table 4.4: The null hypothesis of forecast optimality cannot be rejected as the Wald test statistic is highly significant. This results confirms our conjecture that the intraday CARL model is an adequate statistical representation of the DGP underlying our maximum percentage range liquidity measure.

4.5.2 Varying the Transaction Filter Size

We now vary the size of the volume filter used to obtain $r_{m,t}$. More specifically, we twice increase the volume filter m by a factor 10, starting with a small volume $m = 100$, such that $m \in \{100,\ 1,000,\ 10,000)$. The effect of these different transaction volume filter sizes on the filtered results can be seen in table 4.6:

An obvious consequence of varying the transaction volume filter

Table 4.5: **Comparative statistics - Filtered Amazon $\mathbf{r_{m,t}}$ series**

This table shows statistics of three filtered intraday Amazon $r_{m,t}$ series on 24/07/2001 during NASDAQ market hours. The series have been obtained using a volume filter sizes $m = 100$, $1,000$ and $10,000$ respectively. The Ljung-Box autocorrelation test statistic for the first 50 lags is denoted as LB_{50}. The average filtered duration is the average length of the time interval during which the quantity m is sold and bought and the percentage range realized. The series $OIB_{m,t}$ constitutes the difference between the filtered sell and buy volumes over the filtered intervals.

Filtered $r_{m,t}$ series (24/07/2001)			
Transaction filter size	**m=100**	**m=1,000**	**m=10,000**
# of observations	19,640	3,452	605
Mean	0.134	0.191	0.409
Median	0.079	0.153	0.333
Maximum	1.262	1.264	1.544
Minimum	0.000	0.000	0.073
Std dev	0.118	0.155	0.263
Skewness	3.300	2.184	1.293
Kurtosis	17.953	9.618	4.888
$\mathrm{Corr}(r_{m,t}, \lvert OIB_{m,t}\rvert)$	0.260	0.347	0.399
$\mathrm{Corr}(r_{m,t}, \lvert OIB_{m,t-1}\rvert)$	0.186	0.202	0.043
$\mathrm{Corr}(r_{m,t}, \lvert \mathbb{I}_{OIB_{m,t}>0} \cdot OIB_{m,t}\rvert)$	0.175	0.133	0.190
$\mathrm{Corr}(r_{m,t}, \lvert \mathbb{I}_{OIB_{m,t-1}>0} \cdot OIB_{m,t-1}\rvert)$	0.117	0.072	0.028
LB_{50} (P-value)	$1.7 \cdot 10^6$ (0.00)	3,066 (0.00)	83 (0.00)
Avg filtered duration (in sec)	1.192	6.782	39.433

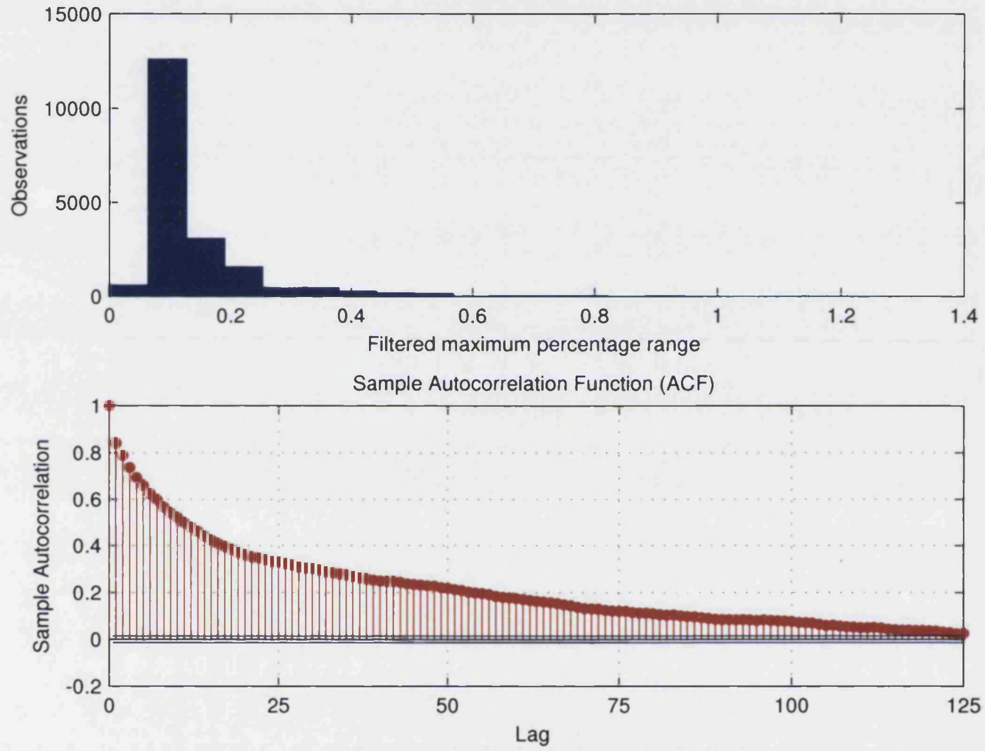

Figure 4.6: **Histogram & ACF - Amazon $\mathbf{r_{m,t}}$ (m = 100)**

This figure displays the histogram (top panel) and the autocorrelation function (ACF; bottom panel) for the filtered intraday Amazon $r_{m,t}$ series on 24/07/2001 during NASDAQ market hours. The series has been obtained using a volume filter size $m = 100$. In addition to the ACF, 2-standard-deviation interval bounds are shown, corresponding roughly to a 95% confidence interval.

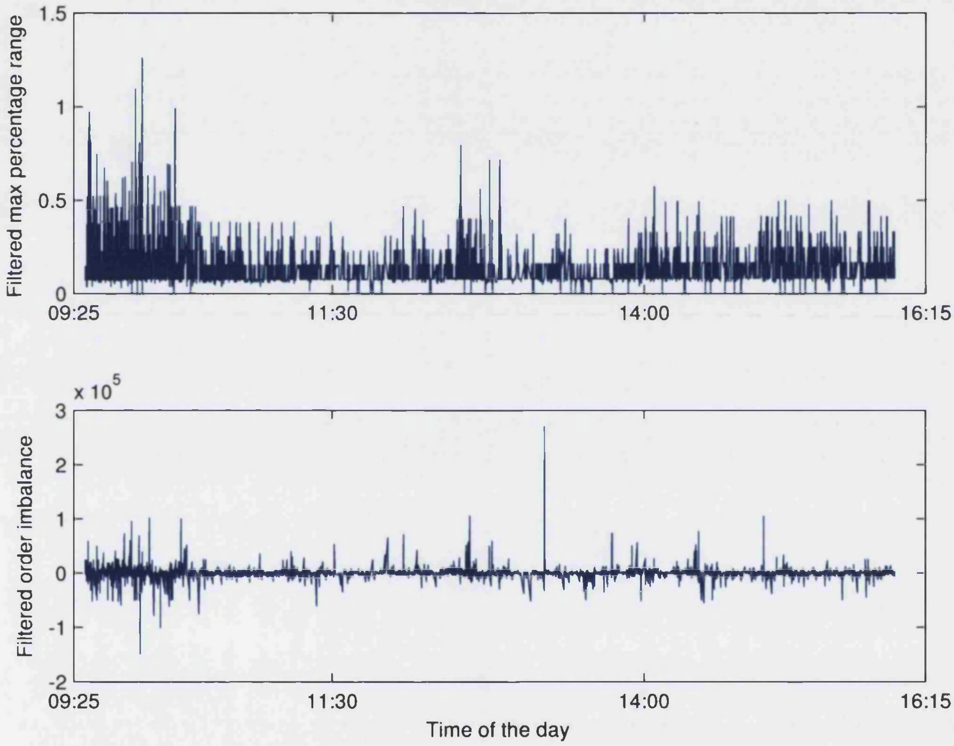

Figure 4.7: **Amazon $\mathbf{r_{m,t}}$ and $\mathbf{OIB}_{m,t}$ ($\mathbf{m = 100}$)**

This figure shows the filtered intraday Amazon $r_{m,t}$ (top panel) and the order imbalance series $\mathrm{OIB}_{m,t}$ (bottom panel) on 24/07/2001 during NASDAQ market hours. The series have been obtained from adjusted TAQ data using a volume filter size m=100.

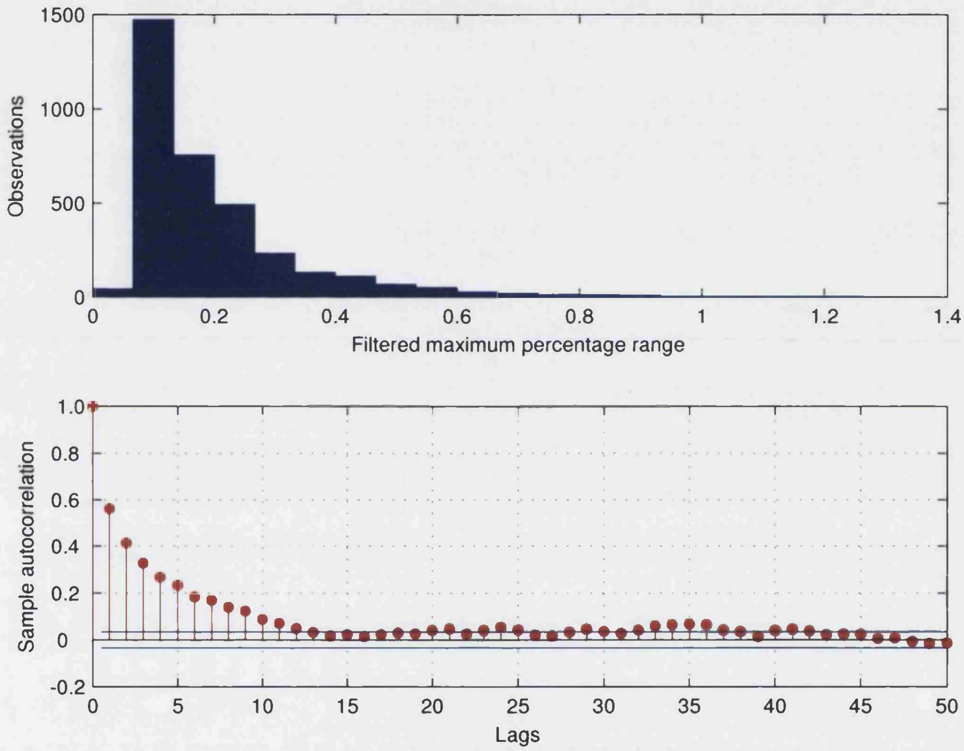

Figure 4.8: **Histogram & ACF - Amazon $\mathbf{r_{m,t}}$ ($\mathbf{m = 1,000}$)**

This figure displays the histogram (top panel) and the autocorrelation function (ACF; bottom panel) for the filtered intraday Amazon $r_{m,t}$ series on 24/07/2001 during NASDAQ market hours. The series has been obtained using a volume filter size $m = 1,000$. In addition to the ACF, 2-standard-deviation interval bounds are shown, corresponding roughly to a 95% confidence interval.

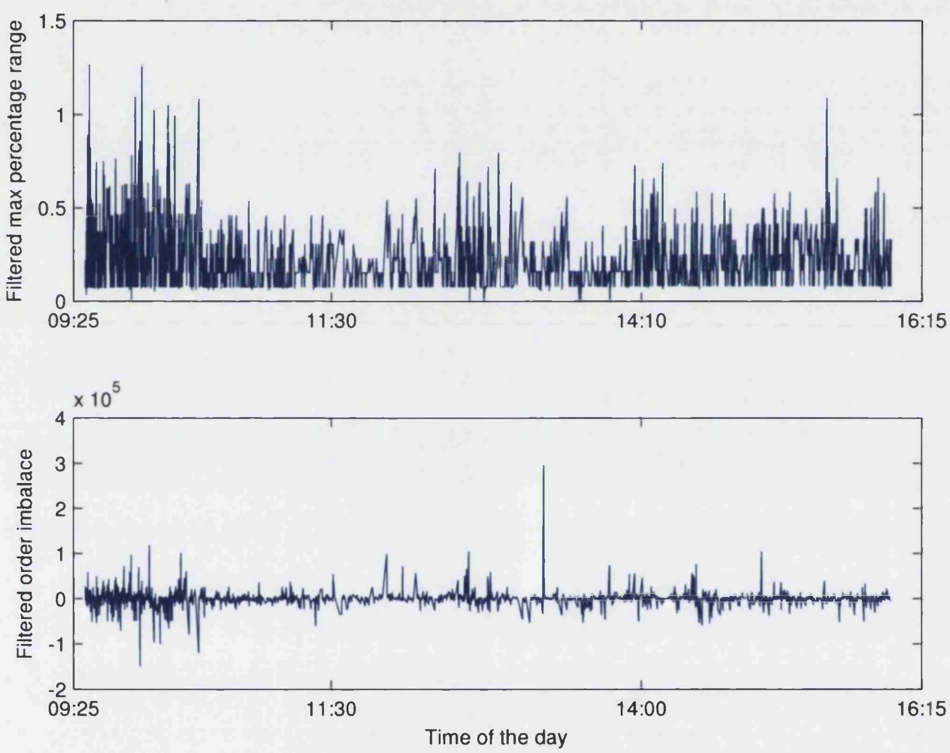

Figure 4.9: **Amazon $\mathbf{r_{m,t}}$ and $\mathbf{OIB}_{m,t}$ $\mathbf{(m = 1,000)}$**

This figure shows the filtered intraday Amazon $r_{m,t}$ (top panel) and the order imbalance series $\mathrm{OIB}_{m,t}$ (bottom panel) on 24/07/2001 during NASDAQ market hours. The series have been obtained from adjusted TAQ data using a volume filter size m=1,000.

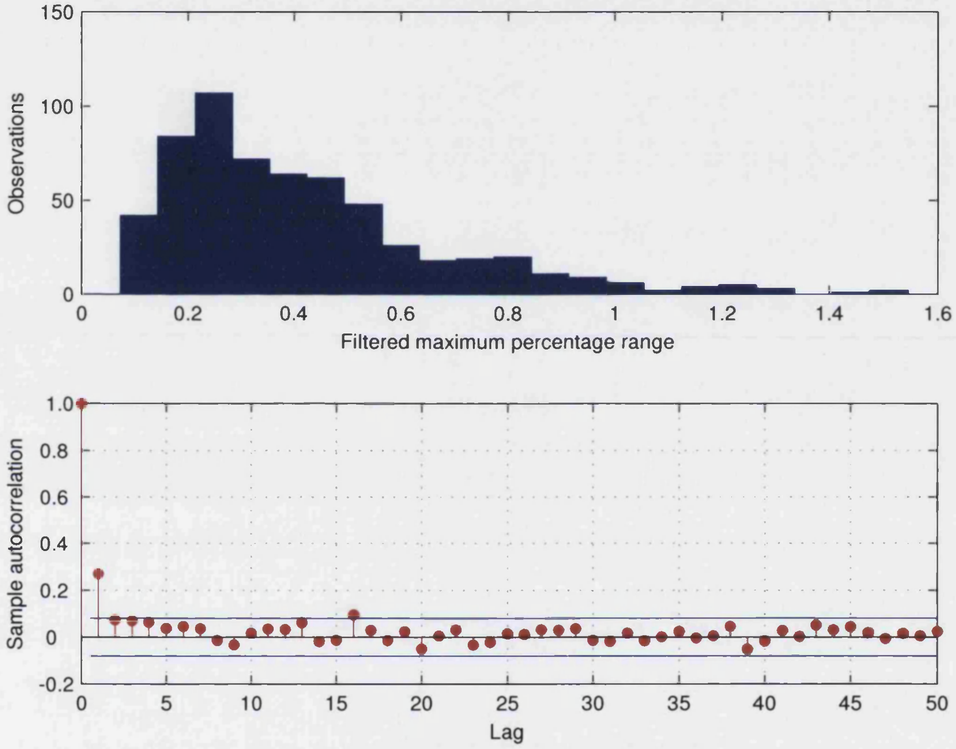

Figure 4.10: **Histogram & ACF - Amazon $\mathbf{r_{m,t}}$ ($\mathbf{m = 10,000}$)**

This figure displays the histogram (top panel) and the autocorrelation function (ACF; bottom panel) for the filtered intraday Amazon $r_{m,t}$ series on 24/07/2001 during NASDAQ market hours. The series has been obtained using a volume filter size $m = 10,000$. In addition to the ACF, 2-standard-deviation interval bounds are shown, corresponding roughly to a 95% confidence interval.

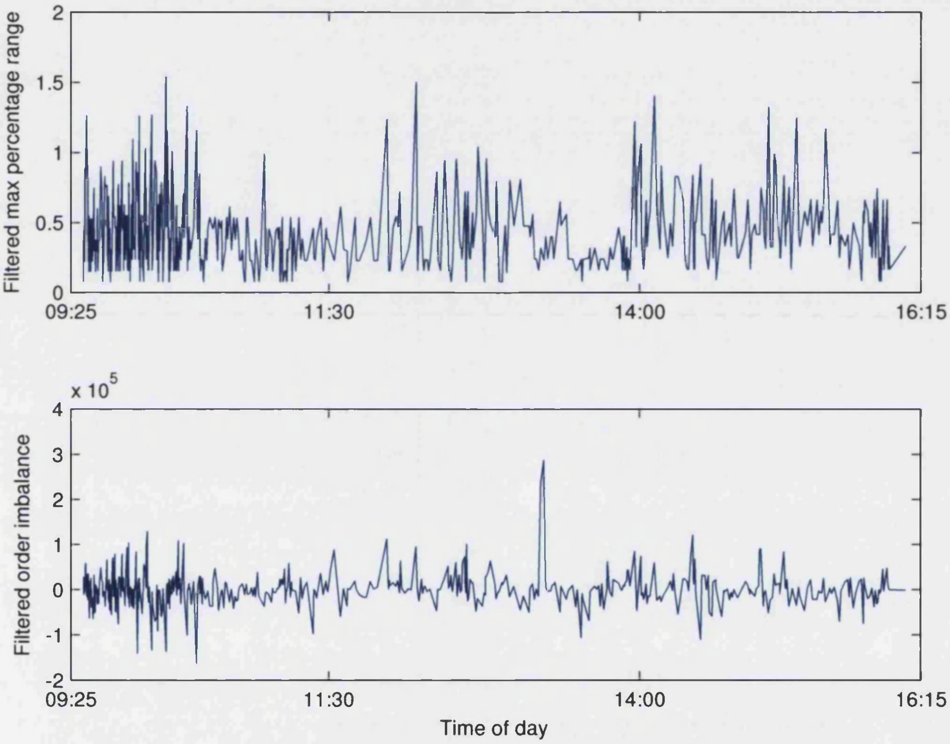

Figure 4.11: **Amazon $\mathbf{r_{m,t}}$ and $\mathbf{OIB}_{m,t}$ ($\mathbf{m = 10,000}$)**

This figure shows the filtered intraday Amazon $r_{m,t}$ (top panel) and the order imbalance series $\mathrm{OIB}_{m,t}$ (bottom panel) on 24/07/2001 during NASDAQ market hours. The series have been obtained from adjusted TAQ data using a volume filter size m=10,000.

size is the difference in the average length of the filtered durations. The smaller the volume, the shorter the average duration, ranging e.g. in the case of $m = 100$ from little more than 1s to close to 40s for the large $m = 10,000$ filter size. This also results in many more filtered observations - close to $20,000$ - for the smaller transaction size filter, which falls substantially to around 600 for $m = 10,000$.

As is to be expected, larger transaction size filters also translate into bigger mean maximum percentage ranges and a seemingly less skewed empirical distributions with less kurtosis. It also appears that the contemporaneous correlation with the $\mathrm{OIB}_{m,t}$ series as well as with $(\mathbb{I}_{\mathrm{OIB}_{m,t}>0} \cdot \mathrm{OIB}_{m,t})$ becomes stronger (still with the same, positive 'correct' sign) the bigger the filter size. The $\mathrm{OIB}_{m,t}$ is thus more of a meaningful directional indicator when larger volumes are transacted on either side. The picture is however less clear-cut across volumes when $\mathrm{OIB}_{m,t}$ is interacted with the downside indicator variable, albeit the correlation is still sizeable, which suggests to us that 'overhang' volume on the downside has a roughly equal effect when the volume filter is varied. Looking at the correlation with lagged (interacted) $\mathrm{OIB}_{m,t-1}$ we find that it actually declines the larger the transaction volume filter, which is indicative of buying/selling activity being more of a definitive factor for smaller volumes and more short-term focused trading activity than for longer horizons. The filtered $r_{m,t}$ series for $m \in \{100, 1,000, 10,000\}$ as well as the corresponding OIB_t are displayed in figures 4.7, 4.9 and 4.11 respectively. As in the case of $m = 500$ previously, while we detect more one sided activity around the opening and mid-day, we do not find any obvious diurnality patterns across the different filter size series.

In terms of the autocorrelation structure of the data, we note that even though the LB_{50} statistics are not strictly comparable across the different filtered series due to their varied samples sizes, the $r_{m,t}$ series obtained with $m = 100$ appears to exhibit much stronger autocorrelation than the ones obtained with a larger m. This is also confirmed by the ACFs displayed in the lower panels of figures 4.6, 4.8 and 4.10. In the case of $m = 10,000$ for example, the ACF loses significance already after the first lag, whereas it is strongly significant for up to 125 lags in case of the $r_{100,t}$ series. Moreover, the initial level of the ACF for the latter series is much higher than for the former, with $r_{1,000,t}$ ranging in the middle.

Based on the above results, we would regard the large filter size of $m = 10,000$ as the upper limit for our analysis given the data on this particular day. An even larger filter size would result in yet fewer filtered observations, making statistical inference more difficult. Moreover, it is to be expected that the autocorrelation structure might become less meaningful for yet larger volume filter sizes. Equally, given the very short filtered durations achieved with $m = 100$, we would argue against lowering the filter size much more as one might then enter into very high-frequency 'tick-by-tick territory', where market microstructure effects are much more pronounced. Moreover, it is very rare that such low volumes are actually bought and sold in the market place; most orders are of much larger size: as per table 4.1, the average transaction size calculated from the cleaned data for example is 740.

Estimating the intraday CARL model against the three $r_{m,t}$ series, we employ the same model selection procedure as above. This time, for ease of presentation, we do however only report the optimal model for each filter in table 4.6:

Table 4.6: **Estimation - Different filtered intraday Amazon $\mathbf{r_{m,t}}$ series**

This table shows results from the estimation of selected CARL models against three filtered intraday Amazon $r_{m,t}$ series on 24/07/2001 during NASDAQ market hours. The series have been obtained using volume filter sizes $m = 100,\ 1,000$ and $10,000$ respectively. The table displays parameter estimates, Bollerslev-Woolridge t statistics as well as statistics for standardized residuals. The CARL models shown in the table exhibit the best in-sample fit for the respective $r_{m,t}$ series. * and ** denote significance at 1% and 5% levels respectively. The Ljung-Box autocorrelation test statistic for the first 50 lags is denoted as LB_{50}.

Filter size	**m = 100**		**m = 1,000**		**m = 10,000**	
Model	**CARL(1,3)**		**CARL(1,2)**		**CARL(1,2)**	
Metric	**Estimate**	**t stat**	**Estimate**	**t stat**	**Estimate**	**t stat**
$\hat{\omega}$	0.000	1.043	0.028*	5.230	0.023	1.880
$\hat{\alpha}_1$	0.827*	13.662	0.527*	7.276	0.898*	16.764
$\hat{\beta}_1$	0.556*	28.082	0.471*	20.638	0.300*	7.080
$\hat{\beta}_2$	−0.233*	−5.981	−0.142*	−2.605	−0.254*	−5.255
$\hat{\beta}_3$	−0.150*	−4.780				
Standardised residuals statistics						
Mean		1.000		1.001		1.000
Std Dev		0.569		0.658		0.621
LB_{50} (P-value)		66 (0.06)		50 (0.48)		36 (0.93)

First, we note that across all three filter sizes the simple $p = q = 1$ parameterisation is not selected as the optimal model. Rather, as already in the case of $m = 500$, there seems to be a case for including more lags of the 'ARCH' component of the CARL model: The chosen 'optimal' model for the smallest transaction filter size is CARL(1,3), for the other two it is CARL(1,2). Again, these models were picked due to their comparatively superior model fit in terms of AIC and BIC[9] and based on their ability to pick of the autocorrelation structure of the data: For the largest filter size the reduction in the autocorrelation of the standardised residuals is substantial, as indicated by a very low LB_{50} statistic and a corresponding high p-value. To us, this is not very surprising as the only the ACF for this series is only significant for the first lag (cf. figure 4.10), which the model seems to pick up quite well. While in the case of $m = 1,000$, the CARL(1,2) model also takes account of the autocorrelation structure rather well ($LB_{50} = 50$, p-value $= 0.48$), the picture is less clear-cut for the smallest transaction size filter: Here, autocorrelation is much more persistent, and the optimal CARL(1,3) is only able to pick it up in a limited way. The 'remaining' autocorrelation in the standardised residuals is still sizeable and one would only not be able to reject the null hypothesis of no autocorrelation over any of the first 50 lags at a 'mild' significance level, as indicated by a LB_{50} statistic of 66 and a corresponding p-value of 0.06. This result suggests that an estimation with another error distribution providing more degrees of freedom while still having a positive domain and a unity expectation such as e.g. the Weibull or Gamma distribution, might yield better a better model fit as suggested by results obtained for the ACD model class in Engle and Russell (1998).

[9]Not reported this time as a comparison for different data series is meaningless

Since we feel however that such an analysis is beyond the scope of the present analysis, we leave it for later work.

We also note that with the exception of the intercept in the case of $m = 100$ and $m = 10,000$, all other parameter estimates are highly significant and, again, of 'reasonable' magnitude. However, for $m = 100$ the estimated setup is not stationary as $\hat{\alpha}_1 + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 > 1$, which again mirrors the difficulty of estimating the model with such a small transaction size filter and a resulting persistent autocorrelation structure. The estimated parameters in all cases do nonetheless adhere to the positivity criterion by Nelson and Cao (1992) since $0 < \hat{\alpha}_1 < 1, \hat{\beta}_1 \geq 0$ and $(\hat{\alpha}_1\hat{\beta}_1 + \hat{\beta}_2) \geq 0$ or $\hat{\alpha}_1(\hat{\alpha}_1\hat{\beta}_1 + \hat{\beta}_2) + \hat{\beta}_3 \geq 0$ for CARL(1,2) and CARL(1,3) respectively.

4.5.3 Discussion

The main result of our empirical application is that the proposed intraday CARL(p,q) approach is able to model the time series behaviour of the filtered maximum percentage range liquidity measure rather well. We base this claim on the good model fit and the ability of the model to pick up on the significant autocorrelation structure found in the data: The remaining autocorrelation in the standardised residuals is negligible for all filter sizes according to the LB_{50} statistic. Given that an autocorrelated data structure is to be expected based on the theoretical models of liquidity presented in section 4.2, the ability of the intraday CARL approach to incorporate such patters into a statistical model confirms us in our view that CARL is a valid and useful model for intraday empirical liquidity applications.

Moreover, the MZ test results document a very encouraging insample forecast performance; while we do not show the MZ test results when we vary the transaction size filter from $m = 100$ to $m = 10,000$ in the previous section for ease of presentation, we can report that the results are comparable to the case of $m = 500$, i.e. the null hypothesis of an optimal forecast cannot be rejected for all volume filter sizes.

Further, a high-level casual comparison of the results obtained with different filter sizes seems to suggest a pattern that is reminiscent of the findings in Reusch (2008): The smaller the filter size, i.e. the shorter the sampled time intervals which in turn translates into higher frequency data, the more persistent the autocorrelation structure in the liquidity measure[10] appears to be and the more there is a role for introducing more 'ARCH' terms into the model: At $m = 100$ and c. $20,000$ filtered durations with a mean of c. 1s for example the optimal chosen model is CARL(1,3), which becomes CARL(1,2) for $m = 10,000$ and c. 600 filtered observations with a mean duration of c. 39s. The latter model setup is also the preferred one for daily frequencies and becomes CARL(1,1) for yet lower frequency weekly data in Reusch (2008). It therefore appears that higher frequencies warrant a richer model parametrisation in order to pick up the autocorrelation structure in the data

4.6 Conclusion

This chapter introduces a version of the empirical CARL by Reusch (2008) for use in an intraday high-frequency context. The proposed model uses the same econometric setup as the 'conventional' CARL model for daily and lower frequency data and such as is very much akin to Engle and Russell's (1998) intraday ACD, which models liq-

[10]Reusch (2008) uses the 'full' HH ratio, i.e. the maximum percentage range divided by the price-weighted turnover.

uidity as the duration between trades. Whereas the latter model is only equipped for exclusive use in an intraday context, this chapter demonstrates that the CARL model, which has been shown to work with daily and weekly data *also* has a role in the intraday analysis of market liquidity when modified slightly: Our modification from the original CARL approach concerns the liquidity measure that is used as the dependent variable in the model. We propose a buy and sell volume filtering technique that allows us to partition the data into intervals over which we then derive the maximum percentage range liquidity measure, which is closely related to the HH ratio used in the original CARL model. The autocorrelated filtered series is strongly autocorrelated, which is very much in accordance with various theoretical models of market liquidity, and is then used as an input into the CARL 'machinery'. Again, similar to the original CARL model, the resulting intraday approach is able to capture multiple dimensions of liquidity and is therefore more versatile than e.g. the ACD model which explicitly only addresses the immediacy notion of liquidity.

In an empirical application, using intraday data on the Amazon stock, we exemplify the capabilities and properties of the intraday CARL model, which is able to pick up the autocorrelation structure of the data rather well and shows promising in-sample forecast ability as measured by a Mincer-Zarnowitz test. Moreover, a sensitivity analysis with different transaction volume filters reveals that the model performs best within certain volume filter bounds as especially low filter sizes produces data series whose autocorrelation structure the model is not able to pick up fully.

Yet, the results obtained with the model across various filter sizes are consistent with theoretical research into market liquidity and as such look very promising. One could for example envisage the use of the intraday CARL model in regulatory environments or intraday trading/risk management applications at financial institutions.

Based on the results in this paper, future research into the area also appears warranted. Such work might for example include the inclusion of different error distributions to achieve better model fit, especially for low filter sizes (as mentioned above), a detailed analysis of the model's out-of-sample forecast performance or an empirical application with a more extensive intraday dataset, possibly covering multiple days of intraday data. A multivariate extension of the model also looks conceivable and would allow for the modelling of common liquidity effects across assets.

Bibliography

Aït-Sahalia, Y., Mykland, P. and Zhang, L. (2005), 'How often to sample a continuous-time process in the presence of market microstructure noise', *Review of Financial Studies* **18**, 351–416.

Amihud, Y. (2002), 'Illiquidity and stock returns: cross-section and time-series effects', *Journal of Financial Markets* **5**, 31–56.

Andersen, T. G. and Bollerslev, T. (1997), 'Intraday periodicity and volatility persistence in financial markets', *Journal of Empirical Finance* **4**, 115–158.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2003), 'Modeling and forecasting realized volatility', *Econometrica* **71**, 579–625.

Bai, X., Russell, J. and Tiao, G. (2004), Effects on nonnormality and dependence on the precision of variance estimates using high frequency data. Working paper, Graduate School of Business, University of Chicago; obtainable from `http://faculty.chicagogsb.edu/jeffrey.russell/research/variance.pdf`.

Bandi, F. and Russell, J. (2006*a*), Market microstructure noise, integrated variance estimators, and the accuracy of asymptotic approximations. Working paper, Graduate School of Business, University of Chicago; obtainable from http://faculty.chicagogsb.edu/jeffrey.russell/research/BR_Kernel.pdf.

Bandi, F. and Russell, J. (2006*b*), 'Separating market microstructure noise from volatility', *Journal of Financial Economics* **79**, 655–692.

Bandi, F., Russell, J. and Yang, C. (2007), Realized volatility forecasting in the presence of ttme-varying noise. Working paper, Graduate School of Business, University of Chicago; obtainable from http://faculty.chicagogsb.edu/jeffrey.russell/research/forecasting.pdf.

Barndorff-Nielsen, O., Hansen, P., Lunde, A. and Shephard, N. (2006), Subsampling realised kernels. Working paper, Nuffield College, University of Oxford; obtainable from http://www.nuffield.ox.ac.uk/economics/papers/2006/w10/subsampling6.pdf.

Bauwens, L., Galli, F. and Giot, P. (2003), The moments of Log-ACD models. CORE Discussion Paper 2003/11, Universit Catholique de Louvain.

Bauwens, L. and Giot, P. (2000), 'The logarithmic ACD model: an application to the bid-ask quote process of three NYSE stocks', *Annales d'Economie et de Statistique* **60**, 117–149.

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

Bauwens, L. and Giot, P. (2002), 'Asymmetric ACD models: introducing price information in ACD models with a two state transition model', *Empirical Economics* **28**, 321–342.

Bauwens, L., Grammig, J. and Veredas, D. (2004), 'A comparison of financial duration models through density forecasts', *International Journal of Forecasting* **20**, 389–609.

Bauwens, L. and Hautsch, N. (2006), 'Stochastic conditional intensity processes', *Journal of Financial Econometrics* **4**, 450–493.

Bauwens, L. and Veredas, D. (2004), 'The stochastic conditional duration model: a latent factor model for the analysis of financial durations', *Journal of Econometrics* **119**, 381–412.

Biais, B., Glosten, L. and Spatt, C. (2005), 'Market microstructure: a survey of microfoundations, empirical results and policy implications', *Journal of Financial Markets* **8**, 217–264.

Brownlees, C. and Gallo, G. (2006), 'Financial econometric data analysis at ultra-high frequency: Data handling concerns', *Computational Statistics & Data Analysis* **51**, 2232–2245.

Brunnermeier, M. and Pedersen, L. H. (2008), Market liquidity and funding liquidity. Working paper, Princeton University; forthcoming in the *Review of Financial Studies*; available from `http://www.princeton.edu/~markus/research/papers/liquidity`.

Carlson, M. (2007), A brief history of the 1987 stock market crash: with a dicussion of the Federal Reserve response, The finance and economics discussion series (feds) papers, The Federal Reserve Board; available from `http://www.federalreserve.gov/pubs/feds/2007/200713/200713pap.pdf`.

Chordia, T., Roll, R. and Subrahmanyam, A. (2000*a*), 'Co-movements in bid-ask spreads and market depth', *Financial Analysts Journal* **56**, 23–27.

Chordia, T., Roll, R. and Subrahmanyam, A. (2000*b*), 'Commonality in liquidity', *Journal of Financial Economics* **56**, 3–28.

Chordia, T., Roll, R. and Subrahmanyam, A. (2001), 'Market liquidity and trading activity', *Journal of Finance* **56**, 501–530.

Chordia, T., Roll, R. and Subrahmanyam, A. (2002), 'Order imbalance, liquidity and market returns', *Journal of Financial Economics* **65**, 111–130.

Engle, R. F. and Lange, J. (2001), 'Predicting VNET: a model of the dynamics of market depth', *Journal of Financial Markets* **4**, 113–142.

Engle, R. F. and Russell, J. R. (1998), 'Autoregressive conditional duration; a new model for irregularly spaced transaction data', *Econometrica* **66**, 1127–1162.

Gerhard, F. and Hautsch, N. (2000), Determinants of inter-trade durations and hazard rates using proportional hazard ARMA models. CoFE Discussion Paper 00/20, Universität Konstanz.

Gerhard, F. and Hautsch, N. (2002), Semiparametric autoregressive proportional hazard models. Working Paper W02, Nuffield College, University of Oxford.

Gerhard, F. and Hautsch, N. (2006), A dynamic semiparametric proportional hazard model. Discussion Paper 2006-05, Finance Research Unit, Department of Economics, University of Copenhagen.

Ghysels, E. and Jasiak, J. (1998), 'GARCH for irregularly-spaced data: The ACD-GARCH', *Studies in Nonlinear Dynamics and Econometrics* **2**, 133–149.

Goodhart, C. A. E. and O'Hara, M. (1997), 'High frequency data in financial markets: issues and applications', *Journal of Empirical Finance* **4**, 73–114.

Gouriroux, C., Ghysels, E. and Jasiak, J. (2004), 'Stochastic volatility duration models', *Journal of Econometrics* **119**, 413–433.

Grammig, J. and Fernandes, M. (2005), 'Non-parametric specification tests for conditional duration models', *Journal of Econometrics* **127**, 35–68.

Grammig, J. and Fernandes, M. (2006), 'A family of autoregressive conditional duration models', *Journal of Econometrics* **130**, 1–23.

Grammig, J. and Maurer, K.-O. (2000), 'Non-monotonic hazard functions and the autoregressive conditional duration model', *Econometrics Journal* **3**, 16–38.

Hansen, P. and Lunde, A. (2006), 'Realized variance and market microstructure noise (with discussion)', *Journal of Business and Economic Statistics* **24**, 127–219.

Hansen, P. R. and Lunde, A. (2005), 'A forecast comparison of volatility models: Does anything beat a GARCH(1,1)?', *Journal of Applied Econometrics* **207**, 873–889.

Hautsch, N. (2002), Modelling intraday trading activity using Box-Cox ACD models. CoFE Discussion Paper 02/05, Universität Konstanz.

Heinen, A. (2003), Modelling time series count data: the autoregressive conditional poisson model. CORE Discussion Paper 2003/62, Université Catholique de Louvain.

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

Hujer, R., Kokot, S. and Vuletić, S. (2005), Modeling the trading process on financial markets using the MSACD model.

Lee, C. M. and Ready, M. J. (1991), 'Inferring trade direction from intraday data', *Journal of Finance* **46**, 733–746.

Mincer, J. and Zarnowitz, V. (1969), The evaluation of economic forecasts, *in* J. Mincer, ed., 'Economic Forecasts and Expectations', National Bureau of Economic Resarch, New York, pp. 3–46.

Morris, S. and Shin, H. S. (1998), 'Unique equilibrium in a model of self-fulfilling currency attacks', *American Economic Review* **88**, 587–597.

Morris, S. and Shin, H. S. (2004), 'Liquidity black holes', *Review of Finance* **8**, 1–18.

Nelson, D. B. and Cao, C. Q. (1992), 'Inequality constraints in the univariate GARCH model', *Journal of Business & Economics Statistics* **10**, 229–235.

Reusch, C. (2008), CARL: An empirical conditional autoregressive model of market liquidity. Working Paper, London School of Economics, Financial Markets Group (FMG); obtainable from the author upon request; email: `mail@christianreusch.com`.

Roll, R. (1984), 'A simple implicit measure of the effective bid-ask spread in an efficient market', *Journal of Finance* **39**, 1127–1159.

Sarr, A. and Lybek, T. (2002), Measuring liquidity in financial markets. Working Paper 02/232, International Monetary Fund; available from `http://www.imf.org/external/pubs/ft/wp/2002/wp02232.pdf`.

Veredas, D., Rodriguez-Poo, J. and Espasa, A. (2002), On the (intradaily) seasonality and dynamics of a financial point process: a semiparametric approach. CORE Discussion Paper 2002/23, Université Catholique de Louvain.