

An Econometric Analysis of Intra-daily Stock Market Liquidity, Volatility and News Impacts

DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum politicarum (Dr. rer. pol.)
im Fach Volkswirtschaftslehre

eingereicht an der
Wirtschaftswissenschaftlichen Fakultät
Humboldt-Universität zu Berlin

von
Axel Groß-Klußmann, M.Sc.

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Wirtschaftswissenschaftlichen Fakultät:
Prof. Dr. Ulrich Kamecke

Gutachter:

1. Prof. Dr. Nikolaus Hautsch
2. Prof. Dr. Christian Conrad

eingereicht am: 02.03.2012

Tag der mündlichen Prüfung: 14.05.2012

Abstract

In this thesis we present econometric models and empirical features of intra-daily (high frequency) stock market data. We focus on the measurement of news impacts on stock market activity, forecasts of bid-ask spreads and the modeling of volatility measures on intraday intervals.

First, we quantify market reactions to an intraday stock-specific news flow. Using pre-processed data from an automated news analytics tool we analyze relevance, novelty and direction signals and indicators for company-specific news. Employing a high-frequency VAR model based on 20 second data of a cross-section of stocks traded at the London Stock Exchange we find distinct responses in returns, volatility, trading volumes and bid-ask spreads due to news arrivals.

In a second analysis we introduce a long memory autoregressive conditional Poisson (LMACP) model to model highly persistent time series of counts. The model is applied to forecast quoted bid-ask spreads, a key parameter in stock trading operations. We discuss theoretical properties of LMACP models and evaluate rolling window forecasts of quoted bid-ask spreads for stocks traded at NYSE and NASDAQ. We show that Poisson time series models significantly outperform forecasts from ARMA, ARFIMA, ACD and FIACD models in this context.

Finally, we address the problem of measuring volatility on small 20 second to 5 minute intra-daily intervals in an optimal way. In addition to the standard realized volatility approaches we construct volatility measures by integrating spot volatility estimates that include information on observations outside of the intra-daily intervals of interest. Comparing the alternative volatility measures in a simulation study we find that spot volatility-based measures minimize the RMSE in the case of small intervals.

Zusammenfassung

In dieser Dissertation befassen wir uns mit ökonometrischen Modellen und empirischen Eigenschaften von Intra-Tages (Hochfrequenz-) Aktienmarktdaten. Der Fokus liegt hierbei auf der Analyse des Einflusses, den die Veröffentlichung von Wirtschaftsnachrichten auf die Aktienmarktaktivität hat, der Vorhersage der Geld-Brief-Spanne sowie der Modellierung von Volatilitätsmaßen auf Intra-Tages-Zeitintervallen.

Zunächst quantifizieren wir die Marktreaktionen auf Marktneuigkeiten innerhalb eines Handelstages. Zu diesem Zweck benutzen wir linguistisch vorab bearbeitete Unternehmensnachrichtendaten mit Indikatoren über die Relevanz, Neuheit und Richtung dieser Nachrichten. Mit einem VAR Modell für 20-Sekunden Marktdaten der London Stock Exchange weisen wir durch Nachrichten hervorgerufene Marktreaktionen in Aktienkursrenditen, Volatilität, Handelsvolumina und Geld-Brief-Spannen nach.

In einer zweiten Analyse führen wir ein long memory autoregressive conditional Poisson (LMACP)-Modell zur Modellierung hoch-persistenter diskreter positivwertiger Zeitreihen ein. Das Modell verwenden wir zur Prognose von Geld-Brief-Spannen, einem zentralen Parameter im Aktienhandel. Wir diskutieren theoretische Eigenschaften des LMACP-Modells und evaluieren rollierende Prognosen von Geld-Brief-Spannen an den NYSE und NASDAQ Börsenplätzen. Wir zeigen, dass Poisson-Zeitreihenmodelle in diesem Kontext signifikant bessere Vorhersagen liefern als ARMA-, ARFIMA-, ACD- und FIACD-Modelle.

Zuletzt widmen wir uns der optimalen Messung von Volatilität auf kleinen 20 Sekunden bis 5 Minuten Zeitintervallen. Neben der Verwendung von realized volatility-Ansätzen konstruieren wir Volatilitätsmaße durch Integration von spot volatility-Schätzern, sodass auch Beobachtungen außerhalb der kleinen Zeitintervalle in die Volatilitätsschätzungen eingehen. Ein Vergleich der Ansätze in einer Simulationsstudie zeigt, dass Volatilitätsmaße basierend auf spot volatility-Schätzern den RMSE minimieren.

Contents

Acknowledgements	1
1 Introduction	3
1.1 On intraday stock market activity	3
1.2 Outline of the thesis	4
2 Quantifying High Frequency News-implied Market Reactions	7
2.1 Introduction	7
2.2 Data	9
2.3 Unconditional Effects of News Items	14
2.3.1 Impact on Volatility and Liquidity	14
2.3.2 Trading Profitability	18
2.4 Market Dynamics around News Items	21
2.4.1 Econometric Methodology	21
2.4.2 Estimation Results	23
2.4.3 Impulse Response Analysis	26
2.5 Conclusions	29
A Appendix	31
3 A Forecasting Framework for Bid-Ask Spreads	35
3.1 Introduction	35
3.2 Properties of Bid-Ask Spreads	36
3.2.1 The Bid-Ask Spread as a Liquidity Measure	36
3.2.2 Empirical Properties	37
3.3 An Econometric Model for Bid-Ask Spread Dynamics	42
3.3.1 The Autoregressive Conditional Poisson model	42
3.3.2 Long Memory Autoregressive Conditional Poisson Models	44
3.4 Forecasting Bid-Ask Spreads	47
3.4.1 Computation of Forecasts	47
3.4.2 Additional Predictors based on Market Microstructure Theory	48
3.4.3 Forecast Benchmarks	49
3.4.4 Point and Directional Forecast Evaluation	50
3.5 Results	51
3.5.1 Estimation Results and RMSE Performance	51
3.5.2 A Trading Schedule based on Spread Forecasts	57
3.5.3 Robustness of the Results	59
3.6 Conclusions	59

Contents

B	Appendix	60
B.1	Technical appendix	60
B.2	Evolution of Coefficient Estimates	61
4	Modelling Volatility Measures on Intra-Daily Time Intervals	65
4.1	Introduction	65
4.2	Intraday volatility measures	67
4.2.1	Realized measures of volatility	67
4.2.2	Volatility measures based on spot volatility estimators	71
4.3	Comparing Intraday Volatility Estimators in a Monte Carlo Simulation Study	77
4.4	Empirical comparison of intraday volatility measures	86
4.4.1	Data and model calibration	86
4.4.2	Empirical Properties of Intraday Volatility Measures	88
4.4.3	A model for intraday returns and volatility	90
4.4.4	The tail probabilities of intraday returns	92
4.5	Conclusions	97
C	Appendix	98
C.1	Finite sample adjustment of the pre-averaging estimator	98
C.2	Kernel functions	98
C.3	Simulation of the SV models with seasonality	98
C.4	Optimal bandwidth according to the grid-search procedure (2-factor model)	99
C.5	Simulation results for the 2-factor model with microstructure noise of magnitude $\omega^2 = 0.00001$	101
C.6	Sample autocorrelations of four volatility measures	103
C.7	Basic descriptive statistics for the AAPL.OQ volatility measures .	105
C.8	Real Data results: RV Garch based on t-distribution, $\alpha = 10\%$.	107
Bibliography		119
List of Figures		122
List of Tables		123

Acknowledgments

First of all I would like to express my deep gratitude towards my supervisor Nikolaus Hautsch for his exceptional support and guidance over the past years. Without his encouragement, comments and suggestions this thesis could not have been written. I would also like to thank Christian Conrad for kindly accepting the task of being co-examiner as well as for his comments on the draft.

This thesis has benefited hugely from the expertise and atmosphere at the chair of econometrics at HU Berlin. I would like to thank my friends and colleagues at the chair of econometrics for valuable advice and many pleasant discussions, especially Bernd Droege, Jonas Haase, Gustav Haitz, Ruihong Huang, Franziska Lottmann, Peter Malec, Tomas Polak, Julia Schaumburg and Melanie Schienle.

The main fraction of the thesis was written at the Deutsche Bank Quantitative Products Laboratory (QPL), a research collaboration between Deutsche Bank, Humboldt Universität and Technische Universität Berlin. Over several years I enjoyed the working atmosphere, financial support and especially the DB data infrastructure. Many thanks are thus due to the directors of QPL, Peter Bank, Almut Birsner and Ulrich Horst. Moreover, I thank my friends and colleagues at the QPL for valuable advice and the pleasant atmosphere, in particular Gökhan Cebiroglu, Peter Kratz, Lada Kyj, Markus Mocha and Nicholas Westray.

I also wish to thank the financial market practitioners from Deutsche Bank and Thomson Reuters who commented on the work and gave me the opportunity to discuss and showcase the results at conferences. In particular I thank Richard Brown, Boris Drovetsky, Kerr Hatrick and Roel Oomen.

Finally, I thank my family and friends for their patient support during the last years.

1 Introduction

1.1 On intraday stock market activity

The last years have witnessed a drastic increase in market activity at major stock exchanges. Chordia et al. [2011] report an increase of the average monthly shares traded per outstanding shares from 5% to 26% in 1993 - 2003, while turnover virtually stagnated in the decade before 1993. Most notably, the number of daily transactions even increased ninetyfold during 1993-2008, indicating a fundamental change in the way of trading.

The changes in market activity are undoubtedly linked to the past and ongoing automatization of trading. The technological change dramatically reduced the costs of market participation and ultimately led to the advent of so-called algorithmic trading. Algorithmic trading is commonly defined as the use of computer algorithms to automatically submit, organize and manage orders for the purpose of trading (see Hendershott et al. [2011]). The automated trading strategies rely heavily on the statistical analysis of daily and intra-daily market activity data, the availability of which has become industry standard. An important subset of algorithmic trading in this context is high frequency trading, where trading decisions are based on data at very high intraday frequencies.

The importance of algorithmic trading on financial markets cannot be underestimated. Hendershott et al. [2011] report that about 73 % of the trading volume in the United States are due to algorithmic trading. However, the consequences of this dramatic development are less clear. While basic liquidity cost measures have declined considerably over the past decade, the effect of algorithmic trading on the overall market quality is virtually unexplored. The reported improvements in the stock market liquidity are typically short-lived and concentrated on the stocks with very high market capitalization. Moreover, the events of the May 6th 2010 "flash crash" at exchanges in the US have raised serious concerns about the risks associated with high frequency and algorithmic trading. In the time-span of half an hour on this day the Dow Jones Industrial Average declined by 1014.14 points, the biggest intraday decline in its history, before rebounding rapidly. Surprisingly, no fundamental new information or change in market sentiment could be attributed to the sudden \$ 1 trillion drop in market value (Easley et al. [2011]). Research documents of the US securities and exchange comission (see SEC and CFTC [2010]) as well as studies such as Kirilenko et al. [2011] confirm that the overreliance on high frequency trading and computerized systems crucially contributed to the crash. The reason for the crash was likely a wrongly configured trading program which triggered the reaction of other trading programs, leading in the end to rapid price drops. In view of the tail risks it is thus still doubtful whether the algorithmic trading industry contributes to the stability of financial markets and the overall welfare or can rather be

1 Introduction

considered a dangerous arms race

While the subsequent chapters of this thesis do not go into detail on the risks and merits of algorithmic trading, they contribute to a better understanding of the empirical features of intraday (high frequency) market data, which is of interest for regulators of markets, market participants and academics interested in the econometric analysis of high frequency data. The thesis focuses on three empirical aspects of intraday trading activity that have not received attention so far, but are important to the statistical support of trading decisions.

First, we present econometric models and empirical results on the impact of public news on stock markets. As yet another indication of the fast technological change, news vendors offer an immediate automated interpretation of various news texts and headlines in the instant of a millisecond. Whenever machine-readable news are generated around the world, programs based on linguistic pattern recognition techniques derive trading signals from the news text that can be fed to trading algorithms instantaneously. Our analysis draws on signals from such a news engine in order to quantify the impact of company-specific news on financial markets using high frequency data.

Second, we propose a novel model and forecasting framework for bid-ask spreads which are important indicators of market liquidity on stock markets. The time series of bid-ask spreads poses interesting challenges for an econometric model as the spread is discrete and typically exhibits a long memory time series behavior.

Third, we model intraday volatility on small intraday intervals. The literature on volatility measures so far covers mainly daily frequencies. However, the increasing intra-day trading activity requires risk measures for intra-daily intervals. In this context we propose to use spot volatility based measures of volatility to overcome the small sample problem encountered in case of small intraday intervals.

1.2 Outline of the thesis

In chapter 2 we examine high-frequency market reactions to an intraday stock-specific news flow.¹ Using unique pre-processed data from an automated news analytics tool based on linguistic pattern recognition we analyze relevance, novelty and direction signals and indicators for company-specific news. Employing a high-frequency VAR model based on 20 second data of a cross-section of stocks traded at the London Stock Exchange we find distinct responses in returns, volatility, trading volumes and bid-ask spreads due to news arrivals. We show that a classification of news according to its relevance indicator as given by the linguistic pre-processing is crucial to filter out noise and to identify significant effects. Moreover, sentiment indicators have predictability for future price trends though the profitability of news-implied trading is deteriorated by increased bid-ask spreads.

In chapter 3 we introduce a long memory autoregressive conditional Poisson (LMACP)

¹A version of this chapter is published in the *Journal of Empirical Finance* (see Groß-Klußmann and Hautsch [2011a]).

1.2 Outline of the thesis

model to model highly persistent time series of counts.² The model is applied to forecast quoted bid-ask spreads, a key parameter in stock trading operations. It is shown that the LMACP nicely captures salient features of bid-ask spreads like the strong autocorrelation and discreteness of observations. We discuss theoretical properties of LMACP models and evaluate rolling window forecasts of quoted bid-ask spreads for stocks traded at NYSE and NASDAQ. We show that Poisson time series models significantly outperform forecasts from ARMA, ARFIMA, ACD and FIACD models. The economic significance of our results is supported by the evaluation of a trade schedule. Scheduling trades according to spread forecasts we realize cost savings of up to 13 % of spread transaction costs.

In the last chapter we address the problem of measuring volatility on small 20 second to 5 minute intra-daily intervals in an optimal way. In addition to the standard realized volatility approaches we construct volatility measures by integrating spot volatility estimates that include information on observations outside of the intra-daily intervals of interest. Comparing the alternative volatility measures in a simulation study we find that spot volatility-based measures minimize the RMSE in the case of small intervals. In an empirical analysis we propose a multiple component realized GARCH framework for intraday returns and evaluate the tail probabilities of the return distributions implied by competing volatility measures. The analysis shows that standard realized measures of volatility and volatility measures constructed from spot volatility estimates contain information beyond squared returns and improve the tail fit of the multiple component realized GARCH.

²See the working paper Groß-Klußmann and Hautsch [2011b].

2 Quantifying High Frequency News-implied Market Reactions

2.1 Introduction

Trading on financial markets is strongly influenced by public company-specific, macroeconomic or political information flows. Markets react sensitively to textual information updates – so-called “news” – which are announced on a regular and irregular basis. However, due to the enormous amount of news continuously released by modern electronic communication media nowadays it becomes increasingly difficult to process all news related to a certain financial asset. Particularly nonscheduled news are noisy and often hard to quantify and to interpret. It is not trivial to separate information from noise and to distinguish between relevant and less relevant news. Consequently, empirical studies typically focus on specific and easily identifiable news events such as scheduled macroeconomic announcements, political interventions or earnings announcements.

This chapter addresses the challenge of linking a virtually continuous and nonscheduled asset-specific news flow to intraday market activity. The fundamental objective of this study is to analyze to which extent high-frequency movements in returns, volatility and liquidity can be explained by the underlying mostly nonscheduled news arrivals during a day. To overcome the major difficulty of structuring and filtering news we employ the trading signals of an automated news engine. Such engines are technological innovations fueled by the algorithmic trading industry which computerize the interpretation of news based on linguistic pattern recognition techniques. The news engines are designed to provide signals on the meaning and the relevance of news items for future price movements as well as for future volatility and liquidity situations.

To our best knowledge, the present study is the first one systematically analyzing data from an automated news engine. We use the Reuters NewsScope Sentiment Engine which classifies firm-specific news according to positive, neutral and negative author sentiments based on linguistic pattern analysis of the respective news story. A further crucial feature of the engine is a numeric indicator classifying the relevance of news as well as a variable indicating the novelty thereof. Exploiting these numeric indicators of news sentiment, relevance and novelty we relate the firm-specific news to high-frequency returns, volatility, trading intensity, trade sizes, trade imbalances, spreads and market depth.

In specific, we aim to answer the following research questions:

- (i) Are there significant and theory-consistent market reactions in high-frequency returns, volatility and liquidity to the intraday news flow?

2 Quantifying High Frequency News-implied Market Reactions

- (ii) Is trading on news-driven, machine-generated trading signals profitable?
- (iii) Is the machine-indicated relevance of news empirically supported by corresponding market reactions?

Question (i) addresses a gap in the empirical finance literature which still lacks evidence on the impact of intra-daily (nonscheduled) news on high-frequency market dynamics. Therefore, this study sheds some light on the question whether it is possible to empirically link linguistically pre-processed and filtered asset-specific news to the intraday trading process. We are particularly interested in the question whether there are market reactions beyond the effects induced by company-specific earnings releases which are well-known to have strong impacts. Hence, we explicitly discard all news on earnings announcement days and focus on effects which are predominantly driven by mostly non-scheduled and inhomogeneous news items. In this context, it is of interest to study not only reactions in returns but also in volatility, bid-ask spreads as well as liquidity demand and supply (represented by trading volume and market depth, respectively). Specifically the high-frequency news-driven effects on bid-ask spreads and market depth are widely unexplored. To our best knowledge, only Fleming and Remolona [1999] and Lee et al. [1993] report mainly summary statistics as some evidence of news-induced reactions in liquidity.

Research questions (ii) and (iii) are about the usefulness and effectiveness of machine-generated news feeds in intraday trading. While (ii) is addressed by testing for the significance of abnormal returns, question (iii) is investigated based on a classification of news into important and less important news items. The answers to (ii) and (iii) show to which extent linguistic analyses can help news vendors and traders to automatically structure the news flow. Finally, these questions provide also a first piece of evidence whether news engines have the potential to become building blocks in algorithmic trading strategies and thus major driving forces in market activity.

Compared to the vast literature on earnings announcements, only very few studies try to measure the market response to firm-specific intraday news. This is mainly because high-frequency news items are typically considered to be too noisy due to the interference with other sources of information. The work of Berry and Howe [1994] is an early attempt to link intraday market activity to aggregated measures of news like, e.g., the number of news. A similar approach is taken by Kalev et al. [2004] who document a positive relationship between the number of intraday news and stock market volatility. In an alternative intraday study, Busse and Green [2002] consider the impact of news released via television to test market efficiency. Ranaldo [2008] is the only analysis providing descriptive statistics on the impact of singular firm-specific news items on intraday trading processes. However, all studies show that the impact of news on intraday trading activity is only very weak and not identifiable anymore if earnings announcements are discarded. Moreover, typically news have to be aggregated to reduce the influence of noisy and non-informative news. This is also confirmed by Mitchell and Mulherin [1994] reporting weak impacts of public news on a daily level. Finally, our study is also related to approaches based on the quantification of news texts. For instance, Tetlock [2007] and Tetlock et al. [2008] perform linguistic analyses of daily Wall Street Journal stories.

Similarly, Antweiler and Frank [2004] link daily stock market activity to textual information from internet stock message boards. However, none of these approaches employ machine-processed and filtered textual news items.

Using 20 second aggregates of transaction data from 39 liquid stocks traded at the London Stock Exchange (LSE), we study news' impacts on abnormal returns, volatility, trading volume, average trade sizes, spreads, trade imbalances and market depth. While most studies analyze news effects based on fixed windows around the event dates, we model the complete underlying trading process. To avoid spurious regression results due to neglected dynamics and cross-dependencies between the variables, we employ a high-frequency Vector Autoregressive (VAR) model which is augmented by news-specific explanatory variables and explicitly accounts also for the naturally high proportion of zero variables arising from non-trading in a 20-second interval.

A major finding of our study is that high-frequency trading activity indeed significantly reacts to intraday company-specific news items which are identified as relevant. The fact that earnings announcements are discarded makes these results quite remarkable. We show that the observed market reactions well match theoretical predictions. By capturing dynamics and cross-dependencies in the VAR framework we find strongest effects for volatility and cumulative trading volumes. Bid-ask spreads, trade sizes and market depth do not necessarily directly react to news but indirectly through the cross-dependencies to volumes and volatility and corresponding spillover effects. Two findings confirm the usefulness of the linguistic pre-processing of news. First, we find that the indicated sentiments have predictive contents for price movements prior to news arrivals. However, simultaneously rising spreads during these periods reduce the profitability of potential trading strategies. Second, only little market impact is found for news that are classified as not being relevant, while strong and significant effects are shown for relevant news. This result shows that news engines have the potential to successfully pre-structure news and to filter out noise.

The remainder of the paper is organized as follows. In the next section, we describe the underlying data set and present descriptive statistics. Section 2.3 reports empirical evidence for unconditional effects of published news items without explicitly controlling for time series dynamics and cross-dependencies in the processes. In Section 2.4, the econometric framework and corresponding results based on a high-frequency VAR model are given. Section 2.5 concludes.

2.2 Data

In order to facilitate the processing of new information, several news vendors offer software environments capturing particular characteristics of news in realtime. These tools electronically analyze available textual information using linguistic pattern recognition algorithms. Words, word patterns, the novelty of a news item, its type and other characteristics are translated into indicators of the relevance, novelty as well as of the tone of the item.

We use pre-processed news data from a news-analytics tool of the Reuters company,

2 Quantifying High Frequency News-implied Market Reactions

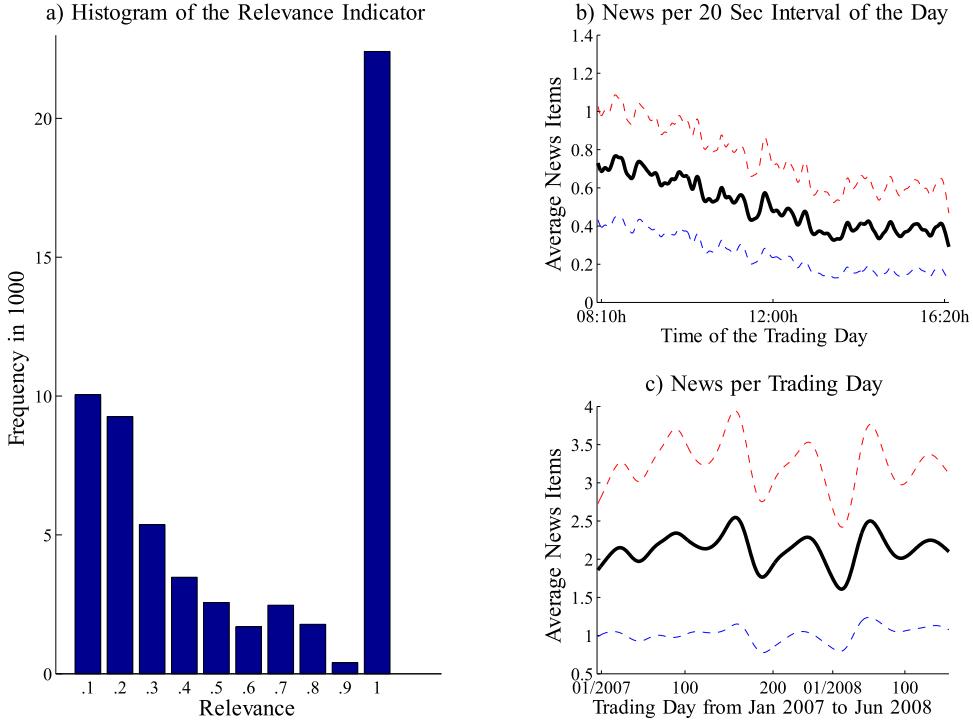


Figure 2.1: a) Distribution of the Relevance Indicator. b) Distribution of news over a day and c) over the time span January 07 to June 08 (averages). Confidence bounds dotted. Smoothed via kernel regression.

the Reuters NewsScope Sentiment Engine. The data contain 29,497 news headlines for the stocks we consider for 01/07 to 06/08 as observed on traders' screens. News arrival is recorded with GMT time stamps up to a millisecond precision. Each news message provides a sentiment, novelty and relevance indicator. The indicators are produced based on an automated linguistic analysis of news texts. The sentiment tags of the news are coded +1, 0 and -1 for a positive, neutral and negative tone of the underlying story, respectively. Relevance is given by a number in the [0, 1] interval. The novelty indicator reflects how many news with similar content have been published prior to a certain news item. It can thus be used to identify initial news topics (novelty= 0) and updates on a topic (novelty> 0).

We select 40 stocks traded at the LSE which are most active in terms of the number of published news items. As we require data availability for 355 trading days from 01/03/2007 to 06/01/2008, the sample is ultimately cut down to 39 stocks. The fact that the selected stocks are also very actively traded allows us to study market dynamics based on a high frequency. Covering 70 % of the market capitalization of the FTSE100, our sample can be considered representative of the FTSE100 (see Table 2.1).

RIC	M.Cap in % of FTSE100	Money Val. Traded	Return 1/7-6/8	Spread	Nr. of Trades	Nr. of Relevant News	Rel. Pos. News	Rel. Neg. News
AAL	2.41	58.44	0.37	2.41	85.87	1266	418	166
AV	1.36	84.21	-0.23	0.68	519.23	558	242	108
AZN	2.72	54.92	-0.19	1.43	175.34	861	391	142
BATS	1.90	38.33	0.31	1.40	232.87	312	131	81
BARC	3.08	391.72	-0.49	0.52	992.96	2243	874	390
BG	1.53	92.56	0.82	0.89	507.97	447	156	82
BGY	0.00	57.63	0.35	0.76	313.03	558	304	144
BLT	1.43	160.19	1.03	1.27	370.12	1300	279	122
BP	7.22	499.47	0.07	0.53	4546.40	2408	908	308
BSY	0.59	59.95	0.03	0.66	603.27	354	150	53
BT	1.61	251.73	-0.27	0.30	2097.74	508	170	85
CBRY	0.74	79.05	0.23	0.70	560.58	376	164	68
CNA	0.84	118.33	-0.18	0.34	855.75	348	134	66
DGE	1.76	64.39	-0.02	0.96	530.28	282	110	67
EMG	0.64	90.51	0.18	0.73	488.34	174	85	40
FP	0.30	134.54	-0.44	0.25	1134.61	469	223	110
GSK	5.00	131.54	-0.17	1.13	855.07	1219	408	127
HBOS	2.75	170.43	-0.65	0.81	650.74	927	320	99
HSBA	6.96	349.53	-0.09	0.56	1806.28	1986	695	293
III	0.30	22.21	-0.12	1.26	177.50	272	62	33
IMT	0.88	20.49	-0	1.98	97.36	397	207	118
ITV	0.27	175.74	-0.45	0.17	3149.78	378	190	53
LLOY	2.08	220.87	-0.33	0.49	1175.85	528	181	71
LSE	0.00	10.84	-0.22	2.67	99.64	700	281	144
MKS	0.78	105.32	-0.47	0.60	565.14	475	152	49
NG	1.29	65.86	0	0.67	499.17	312	124	56
NXT	0.26	22.78	-0.36	1.90	97.74	323	153	65
PRU	1.10	111.43	-0.05	0.68	512.25	417	173	88
RBS	4.05	450.73	-0.89	0.69	1046.44	1859	560	234
RIO	1.80	59.42	1.20	3.21	70.75	1223	362	199
RR	0.52	81.26	-0.06	0.54	502.99	397	238	125
SAB	1.14	36.72	0.10	1.41	254.56	368	181	98

SBRY	0.45	90.66	-0.14	0.54	718.96	652	307	169	92
SL	0.40	43.67	-0.16	0.45	454.73	321	189	82	77
STAN	1.33	54.13	0.24	1.45	213.59	735	328	173	96
TSCO	2.07	213.45	0.02	0.33	1034.07	506	190	95	55
ULVR	1.20	41.52	0.17	1.34	299.38	330	108	61	29
VOD	4.82	1426.32	0.15	0.14	508.25	1700	559	315	137
XTA	1.55	56.23	0.56	2.80	76.53	1008	326	164	120
Sum	69.16				29497	11033	4943	3891	

Table 2.1: Descriptive statistics of the FTSE stocks

Note: RIC denotes the Reuters Identifier Code. The second column is the % market cap, defined as the shared price times the number ordinary shares in issue for 2007, as fraction of the market cap for the FTSE100. Money Value (traded) is computed as the trade size times the respective price (traded turnover total in 01/2007 to 06/2008 in 100000). Return refers to the % price change from 01/03/07 to 06/01/08. Spread and Nr. of trades (in 100000) are averages per trading day. The News column refers to the number of news items per firm without overnight news and duplicate entries. Relevant news items are classified to be the ones with a relevance indicator equal to one. Rel. Positive and Rel. Negative give the numbers of relevant positive and negative items, respectively.

The underlying transaction data is aggregated to 20 second intervals. We consider this aggregation level to be a good compromise between exploiting a maximum of information on the one hand and making the analysis still computationally tractable (given 1.5 years of data). To reduce the impact of market opening and closing effects, we discard the first ten and last ten minutes of a trading day. Intraday returns, volatility and liquidity are captured by the following variables computed over 20 second intervals:

- (i) money value traded, defined as trade sizes in the intervals weighted by the corresponding mid-quotes,
- (ii) average trade size, defined as the cumulated trade size divided by the corresponding number of trades per interval,
- (iii) bid-ask spread, evaluated at the endpoint of each interval,
- (iv) mid-quote returns over each interval,
- (v) depth, defined as the volume pending at the best bid and ask level, evaluated at the endpoint of each interval,
- (vi) volatility, defined as the sum of squared mid-quote transaction returns over each interval,
- (vii) trade imbalance, defined as the difference of cumulated sizes of buyer and seller-initiated trades (identified by the Lee and Ready [1991] algorithm), normalized by the cumulated trade size,
- (viii) absolute trade imbalance.

As shown by Figure 2.13 (see the appendix), all volatility and liquidity variables exhibit pronounced intraday trading patterns. To capture the seasonality, we standardize all processes by the yearly average of the corresponding underlying 20 s interval. We compute standardized variables according to

$$x_{jd}^* := \frac{x_{jd}}{1/n \sum_{d=1}^n x_{jd}},$$

where j denotes the specific interval of the trading day d and x represents the corresponding variable.

We delete all news on days of earnings announcements to minimize the influence of scheduled earnings releases on the results. This allows us to focus on the yet unexplored data of widely unscheduled intraday news driven mostly by random events. In addition, we only consider the news flow within a trading day and do not examine overnight news. Incorporating the latter would considerably increase the complexity of the study.

After pre-filtering, the number of news range from a minimum of 174 to a maximum of 2408 disclosures per stock for the 01/2007 to 06/2008 period (see Table 2.1). We observe that news tend to cluster in the first half of a day. As shown by Figure 2.1 b),

the news intensity peaks at the beginning of the trading period and decreases during the rest of the day. Figure 2.1 c) reveals that there is no pronounced yearly pattern of news arrival.

In order to identify potentially market-moving news, we distinguish between relevant and less relevant news according to the linguistic pre-analysis. Since we expect the reported relevance tag of news to be a relatively noisy measure, we classify items with an indicator value at (below) the maximum 1 as relevant (irrelevant) news (see Figure 2.1 a)).

2.3 Unconditional Effects of News Items

2.3.1 Impact on Volatility and Liquidity

Quantifying the unconditional impact of news without controlling for market dynamics and cross-dependencies between variables already provides important insights and serves as a basis for the econometric modelling in Section 2.4. We analyze 720 20-second intervals around the arrival of news items capturing 180 intervals before each disclosure and 540 thereafter.

Figure 2.2 illustrates the timing of the intervals. I_0 denotes the specific 20-second interval around the news item, whereas T_1 and T_2 are the numbers of intervals before and after the disclosure, respectively. For each stock, we compute the average market reaction and corresponding standard errors over all event windows. For sake of brevity, we refrain from showing results for individual stocks but report pooled averages over the cross-section of stocks. Correspondingly, by denoting the market reaction of variable X to news item i during interval I_j as X_{iI_j} , the pooled average across all news events and all stocks is computed as $\bar{X}_{I_j} = 1/n \sum_{i=1}^n X_{iI_j}$, where n is the total number of news for all stocks. Given that the stocks have quite similar empirical characteristics (see Table 2.1), this proceeding allows us to highlight the results common to all stocks. Assuming (approximative) normality, the 95% confidence intervals of \bar{X}_{I_j} are computed as two times the standard errors of \bar{X}_{I_j} . Since these standard errors reflect variations across all event windows as well as across the market they capture overall news responses and statistical confidence thereof. Two robustness checks underscore the validity of the inference. First, the confidence intervals closely match those obtained from a parametric bootstrap. Second, to account for the fact that stocks with a high number of news naturally have a stronger weight in \bar{X}_{I_j} , we perform a robustness check using a group-means estimator instead of a pooled average. The corresponding results are qualitatively identical.¹

Figures 2.3 to 2.5 show the money value traded, realized volatility, bid-ask spreads, market depth, average trade sizes and absolute trade imbalances around relevant and less relevant news items. Note that by construction of the seasonality adjustment the mean of each series equals one.

¹See the Appendix for more details on the computation of standard errors.

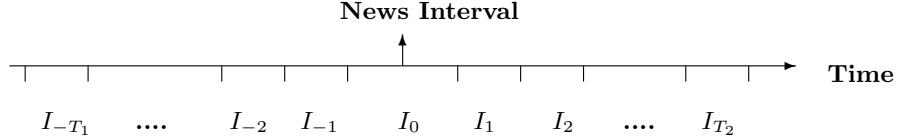


Figure 2.2: Intervals around news arrival

The following findings can be summarized: First, we identify significant upward movements in money value traded, average trade sizes and volatility around the releases of news items. Hence, volatility and trading activity clearly increase when news are published. This supports a 'micro-foundation' of the mixture-of-distribution hypothesis as postulated by Clark [1973] and put forward by Tauchen and Pitts [1983] and Karpoff [1986]. In this framework, price changes are essentially driven by trading on pieces of news, whereas uninformed traders tend to trade when they see large price movements. In consequence, the theory indeed predicts co-movements of volatility and volume. The observed effects are also well-supported by market microstructure theory providing several explanations for higher trading activity during news announcements: (i) Larger trade sizes due to execution by better informed market participants according to Easley and O'Hara [1987], (ii) increased trading due to news-induced information asymmetry among market participants as advocated in the models of Kim and Verrecchia [1991] and Kim and Verrecchia [1994], (iii) trading because of differences in opinion of traders on news' topics as in Harris and Raviv [1993] as well as Kandel and Pearson [1995], and (iv) trading as a consequence of the attention grabbing behavior of investors as documented by Barber and Odean [2008]. Beyond overall increases in volumes, we also observe slight increases in absolute trade imbalance reflecting that trading activity on the two sides of the market tends to become also more asymmetric in periods of information dissemination.

Second, the release of a news item significantly increases bid-ask spreads but does not necessarily affect market depth. Hence, liquidity suppliers predominantly react to news by revising quotes but not by offered order volumes. This is well supported by asymmetric information based market microstructure theory (see, e.g., Easley and O'Hara [1992]) where specialists try to overcompensate for possible information asymmetries. Though on an electronic market there are no designated market makers, the underlying mechanism is similar: Liquidity suppliers reduce their order aggressiveness in order to avoid being picked off (i.e., being adversely selected) by traders who are better informed. For earnings announcements, such effects are also reported by Krinsky and Lee [1996].

Third, the machine-indicated relevance of a news item is clearly supported by corresponding market reactions. All variables respond significantly stronger if information is indicated to be of highest relevance. Actually, for less relevant news we cannot identify

2 Quantifying High Frequency News-implied Market Reactions

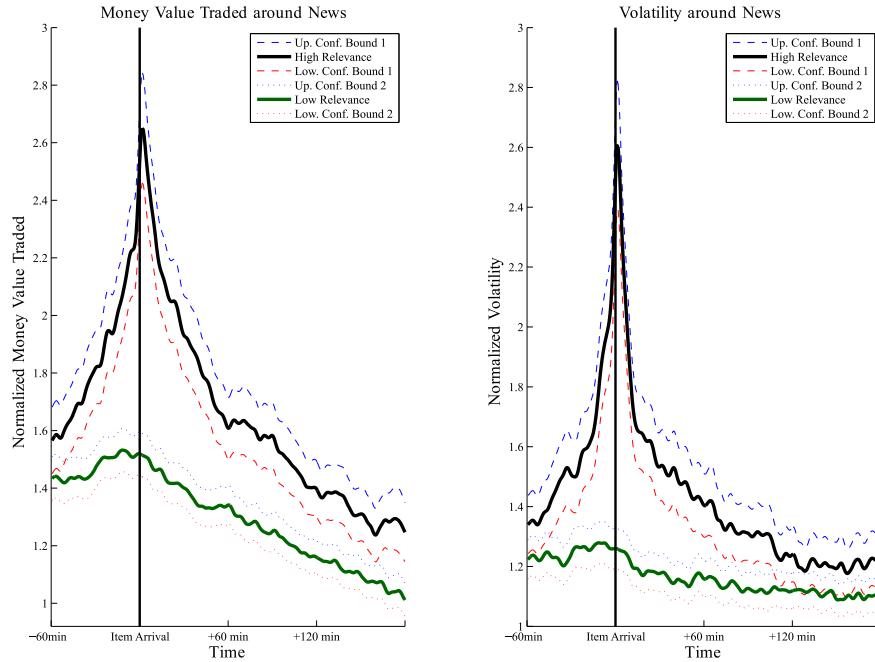


Figure 2.3: Money Value and Volatility around news arrivals. Smoothed via kernel regression.

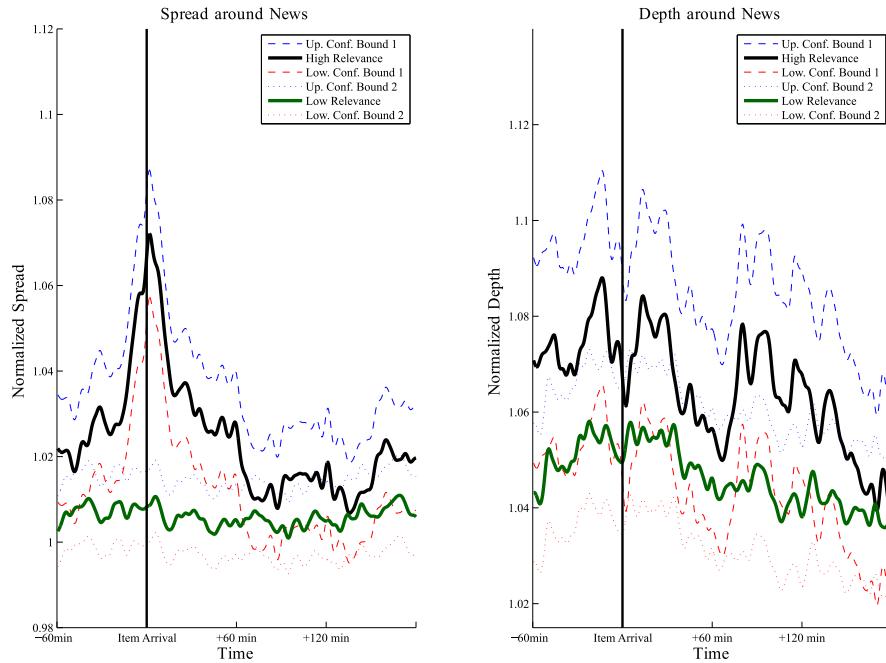


Figure 2.4: Spread and depth around news arrivals. Smoothed via kernel regression.

2.3 Unconditional Effects of News Items

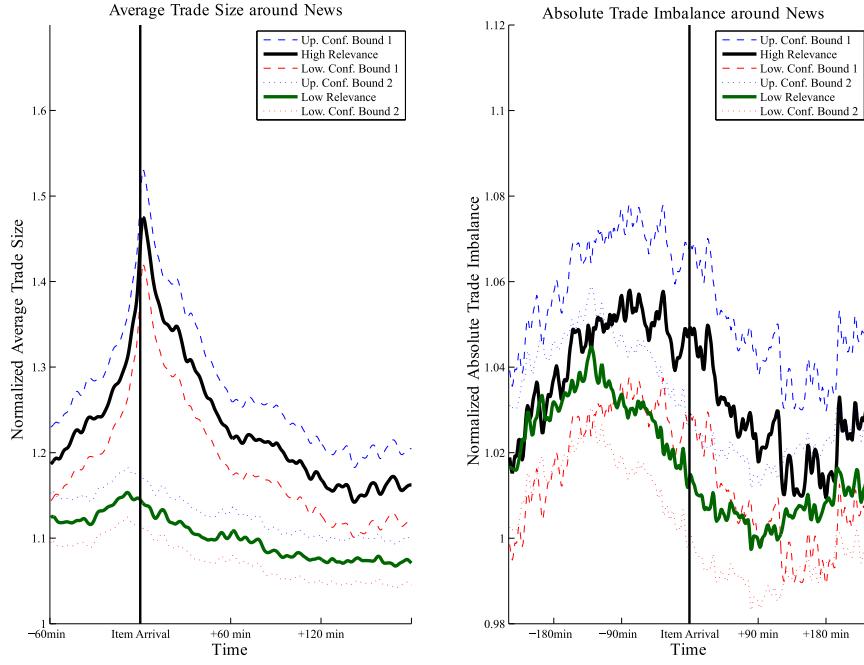


Figure 2.5: Average trade sizes and absolute trade imbalance around news. Smoothed via kernel regression.

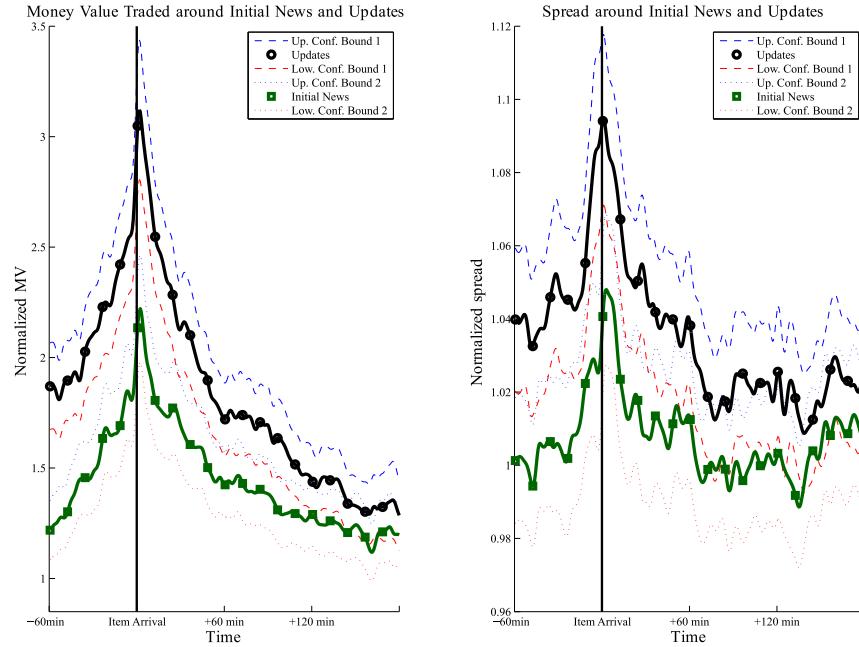


Figure 2.6: Money value and spread around initial news and updates. Smoothed via kernel regression.

significant deviations of the analyzed trading variables from their pre-news levels. This finding is economically in line with Blume et al. [1994] who argue that higher volumes reflect a higher quality of news signals. Moreover, it shows that an appropriate filtering and structuring of the news flow (as provided by the news engine) is crucial to identify systematic market responses. In fact, the noisiness of less relevant news items is the major reason for the yet missing empirical evidence on statistically significant relationships between intraday news flow and high-frequency market activity.

Fourth, for most variables, above-average activities start already more than sixty minutes before the item arrival. This phenomenon is also known in case of periodically scheduled earnings releases. According to the model in Kim and Verrecchia [1994], trading prior to news depends on the degree of information leakages. Our results show that some market participants seem to have additional and partly more timely channels of information. Besides information leakage we attribute the pre-news reaction to a clustering of news. The clustering is an effect inherent in the production of news stories: Beginning with an alert about the news content, subsequent updates ultimately culminate in a full-blown story. Typically, the time between updating steps is small. Indeed, the novelty indicator of the news data allows us to separate between 'news' (in its true sense) and updates on the topic published later. Figure 2.6 shows money value traded and bid-ask spreads around 'initial' news and subsequent updates. Most strikingly, we find that trading on updated news is much more pronounced than trading on the initial news which strongly supports the notion of news clustering. Hence, market reactions become stronger if signals on news are repeated, refined and possibly enforced. This confirms the importance of accounting for the full history of news. Given that later published full stories are more precise than initial alerts, the findings support also the theoretical model of Tetlock [2010] who argues that the magnitude of the volatility and volume change at public news disclosures are positively related to the precision of the news.

2.3.2 Trading Profitability

To test for the profitability of trading on news items we employ an event study framework as outlined in Campbell et al. [1997]. As a model for 'normal' returns we assume the market model

$$R_{it} = \alpha_i + \beta_i R_{mt} + \gamma_i R_{i,t-1} + \varepsilon_{it}, \quad \varepsilon_{it} \sim (0, \sigma_i^2), \quad (2.1)$$

where t denotes the underlying (20 second) intervals, R_{mt} is the market return, computed as the return of the FTSE 100 index, and R_{it} is the return for stock i . To capture return dynamics on high frequencies we also include lagged returns. Model (2.1) is estimated based on the complete 20-second return time series *without* including the event windows. Using the resulting parameter estimates, we compute the abnormal returns $\widehat{AR}_{it} := R_{it} - \widehat{\alpha}_i - \widehat{\beta}_i R_{mt} - \widehat{\gamma}_i R_{i,t-1}$ during the event windows. Let $\widehat{\mathbf{AR}}_i^k$ denote the $((T_1 + T_2 + 1) \times 1)$ vector of abnormal returns for event k of stock i computed between time points I_{-T_1} and I_{T_2} in Figure 2.2. Let γ_j be a $((T_1 + T_2 + 1) \times 1)$ vector consisting of

2.3 Unconditional Effects of News Items

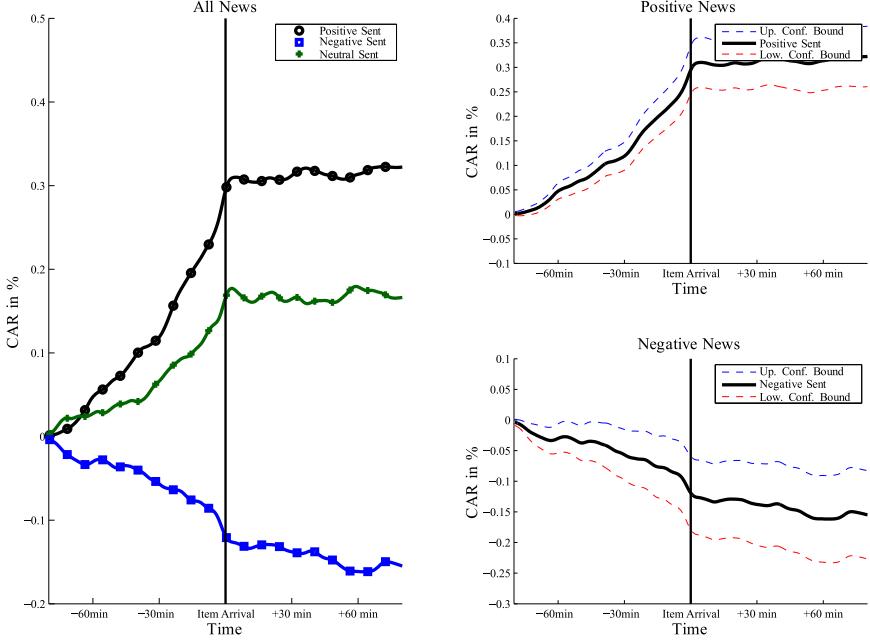


Figure 2.7: Cumulated abnormal returns around relevant positive, negative and neutral news. Smoothed via kernel regression.

ones in the first j places and zero else. Then, we define the cumulated abnormal return for interval j after the I_{-T_1} interval as

$$\widehat{CAR}_{ij}^k := \gamma_j' \widehat{\mathbf{AR}}_i^k. \quad (2.2)$$

Averaging \widehat{CAR}_{ij}^k yields

$$\widehat{CAR}_j = \frac{1}{n} \left(\sum_i \sum_k \widehat{CAR}_{ij}^k \right), \quad (2.3)$$

where n is the total number of events over all stocks. Assuming (asymptotic) normality, 95% confidence intervals are computed as two times the standard deviation of the estimates \widehat{CAR}_j .

Figure 2.7 shows the averaged cumulated abnormal returns (ACAR) \widehat{CAR} around relevant news. Starting 90 minutes *before* the disclosure we observe significantly positive (negative) cumulated abnormal returns as reactions to positive (negative) news items. The news engine obviously allows to establish a significant relationship between a stories' sentiment and the corresponding sign of price trends. However, we observe already significant price movements prior to news releases but only little return reactions thereafter. Though private pre-release information might be present, we conjecture that

2 Quantifying High Frequency News-implied Market Reactions

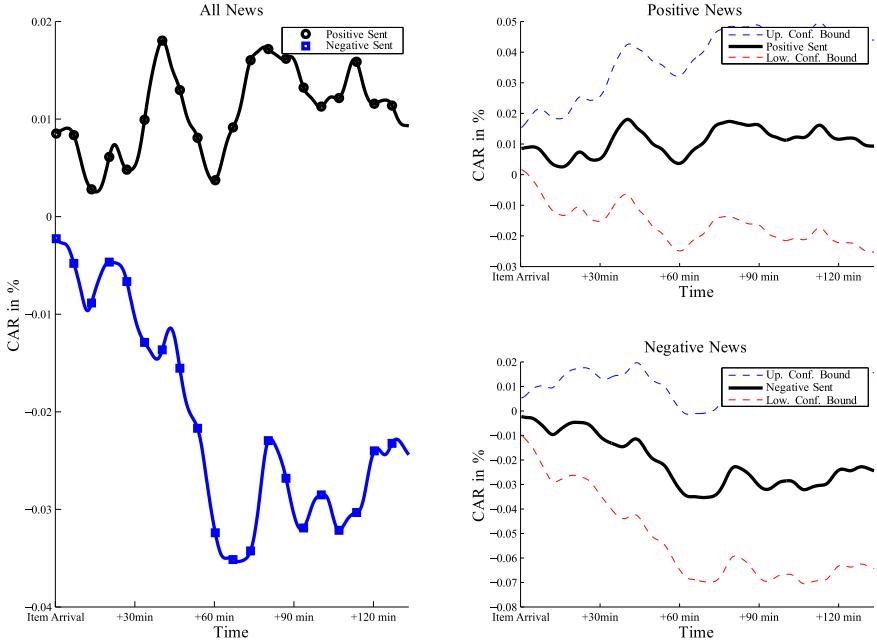


Figure 2.8: Cumulated abnormal returns after relevant positive and negative news.
Smoothed via kernel regression.

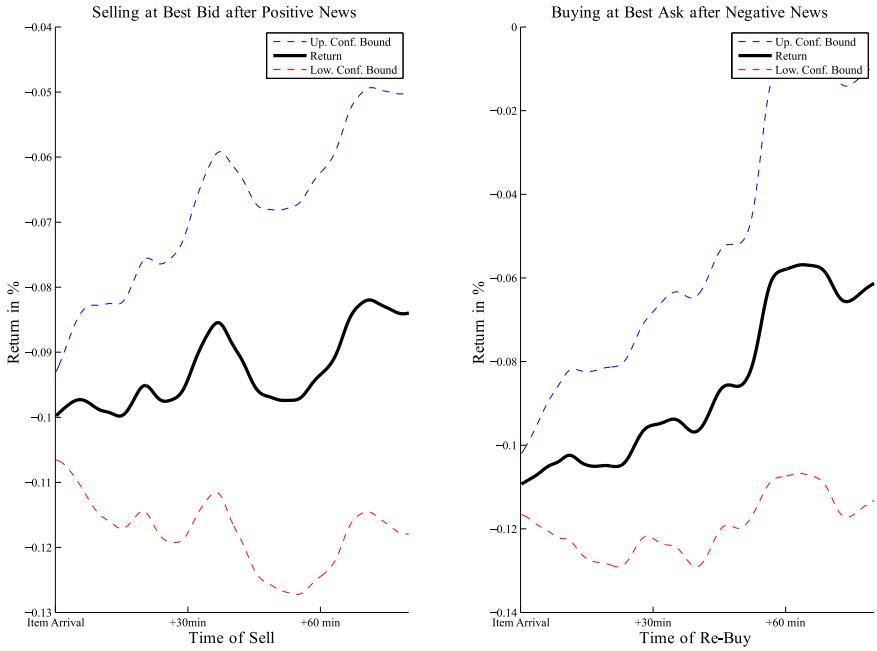


Figure 2.9: Profitability of trading on the sentiments. Smoothed via kernel regression.

the availability of other sources of information and an induced clustering of news items is mainly responsible for pre-announcement effects. Figure 2.8 depicts the ACARs explicitly starting at news disclosure. We observe that sentiment indicators of news items have some predictability for future price movements. Nevertheless, we find the abnormal returns to be mostly insignificant.

In order to provide more specific evidence on trading profitability, we test a stylized trading strategy based on the sentiment information. Following a positive news item returns are computed by buying at the best ask at the item arrival and selling later at the bid. Conversely, after negative news, the asset is sold at the best bid and re-bought later at the ask. As shown by Figure 2.9, we observe that the corresponding returns increase with the underlying horizon but are generally negative. This result shows that the abnormal returns of maximally 3.5 basis points in case of negative news (cf. Figure 2.8) are too low to overcompensate increased bid-ask spreads around news and to provide economic gains of the underlying trading strategies. Still, the fact that return trends are positive (though never crossing the zero line), might be exploited via trading strategies.

2.4 Market Dynamics around News Items

2.4.1 Econometric Methodology

The unconditional analysis of the previous section provides strong indications for -driven market reactions to news disclosures. However, we observe significant autocorrelations as well as cross-correlations between the variables. Figure 2.10 shows corresponding autocorrelation functions.² The autocorrelation functions reveal a high persistence of the individual processes. Geweke and Porter-Hudak [1983] estimates of the fractional integration parameter (not shown in the paper) indicate that some series exhibit long range dependence and are overall covariance stationary. In order to avoid spurious results, the dependencies and interdependencies have to be explicitly taken into account. We suggest a six-dimensional model for the realized variance, the money value traded, the bid – ask spread, market depth, average trade size and absolute trade imbalance. As high-frequency volatility and liquidity variables are only weakly related to (signed) returns we refrain from including the latter in the model. Furthermore, in a separate analysis we find that the signed trade imbalances do not react to signed news (results given in the appendix). Nevertheless, to capture the order flow we consider the absolute trade imbalance instead. Accordingly, the vector of endogenous variables consists of the money value traded, the volatility, the absolute trade imbalance, the bid-ask spread, the market depth and the average trade size.

Since even for liquid stocks there is not necessarily a transaction in every 20 second interval we observe a non-trivial fraction of zero observations for money value traded and realized volatility. In particular, there are no trades in 46% of all 20 second intervals on average. To capture this finding, we suggest explicitly differentiating between the cases of trading, $y_{1t} > 0$, and no trading, $y_{1t} = 0$, in interval t . Correspondingly, the log

²Sample cross-correlations are given in the appendix.

2 Quantifying High Frequency News-implied Market Reactions

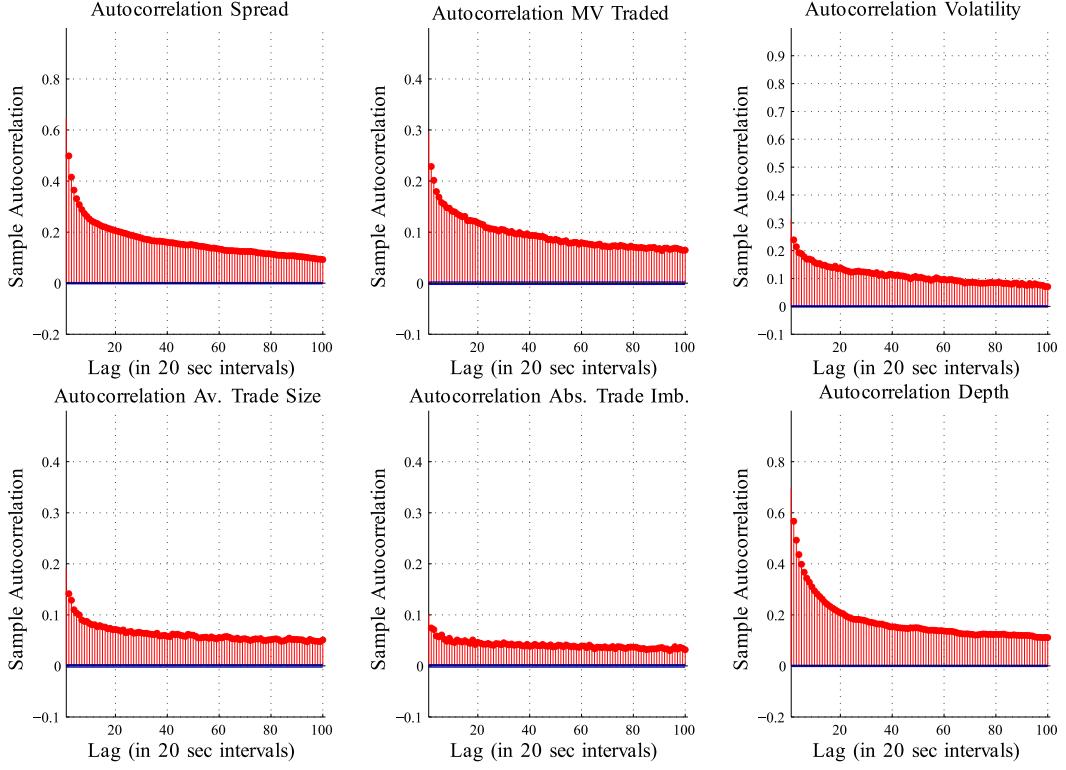


Figure 2.10: Typical sample autocorrelations for the variables of interest (case: AV.L)

likelihood function is given by

$$\begin{aligned} \ln \mathcal{L}(\mathbf{y}; \boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \boldsymbol{\theta}_3) = & \sum_{t=1}^T \{ \ln f(\mathbf{y}_t | y_{1t} > 0; \boldsymbol{\theta}_1) + \ln P(y_{1t} > 0; \boldsymbol{\theta}_2) \} \cdot \mathbb{1}(y_{1t} > 0) \\ & + \sum_{t=1}^T \{ \ln P(y_{1t} = 0; \boldsymbol{\theta}_2) + \ln f(\mathbf{y}_t | y_{1t} = 0; \boldsymbol{\theta}_3) \} \cdot \mathbb{1}(y_{1t} = 0), \end{aligned}$$

where $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$ and $\boldsymbol{\theta}_3$ denote corresponding parameter sets.

As long as the parameter sets $\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$ and $\boldsymbol{\theta}_3$ are disjoint, the likelihood components can be maximized separately. Since the case of no trading, $f(\mathbf{y}_t | y_{1t} = 0; \boldsymbol{\theta}_3)$, is not in the core of our interest, we leave it unspecified. In addition, we refrain from explicitly modeling the long range dependence in some individual time series since this is infeasible in our case of 530000 observations per variable and stock. To parameterize $f(\mathbf{y}_t | y_{1t} > 0; \boldsymbol{\theta}_1)$, we suggest a VAR specification given by

$$\mathbf{y}_t | y_{1t} > 0 = \mathbf{c} + \sum_{i=1}^p (\boldsymbol{\Gamma}_i \mathbf{y}_{t-i} + \boldsymbol{\Psi}_i Z_{t-i}) + \boldsymbol{\Xi} \cdot \mathbf{D}_t + \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \boldsymbol{\Omega}), \quad (2.4)$$

where $\boldsymbol{\Gamma}_i$ and $\boldsymbol{\Xi}$ denote (6×6) and $(6 \times (p_1 + p_2 + 1))$ coefficient matrices, where $p_1 > 0$ and $p_2 > 0$ are integers. Lags of the dummy $Z_t := \mathbb{1}_{(y_{1t}=0)}$ capture previous periods of nontrading with corresponding (6×1) coefficient vectors $\boldsymbol{\Psi}_i$. To capture the impact of news we define the dummy variable d_t with value one in case of relevant news in t and zero otherwise. Then, $\mathbf{D}_t := (d_{t+p_1} \dots d_{t-p_2})'$ is a vector of time dummies indicating the arrival of *relevant* news and covering p_1 intervals before and p_2 intervals after news disclosures. Model (2.4) can be consistently (though not necessarily efficiently) estimated equation by equation using ordinary least squares.

For the money value equation, the conditional probabilities for the occurrence of zero observations (i.e., no trading) in period t , $P(y_{1t} = 0; \boldsymbol{\theta}_2)$, are parameterized in terms of a probit specification. Let \mathbf{x}_t contain all right-hand side variables of equation (2.4), i.e., $\mathbf{x}'_t := [1 \mathbf{y}'_{t-1} \dots \mathbf{y}'_{t-p} Z'_{t-1} \dots Z'_{t-p} \mathbf{D}_t]$. Assuming a normally distributed latent process $y_{1t}^* \sim N(\mathbf{x}'_t \boldsymbol{\theta}_2, 1)$ underlying the trading "decision", we have

$$P(y_{1t}^* > 0) = \Phi(\mathbf{x}'_t \boldsymbol{\theta}_2), \quad \text{if } y_{1t}^* > 0 \Leftrightarrow y_{1t} > 0, \quad (2.5)$$

$$P(y_{1t}^* \leq 0) = 1 - \Phi(\mathbf{x}'_t \boldsymbol{\theta}_2), \quad \text{if } y_{1t}^* \leq 0 \Leftrightarrow y_{1t} = 0, \quad (2.6)$$

for the binary decision $y_{1t} > 0$ vs. $y_{1t} = 0$. The probit model is straightforwardly estimated by maximum likelihood.

The model is applied to each stock in our sample. Depending on the number of underlying trading days, the individual time series for the 39 stocks contain up to 533,000 observations. In order to obtain equal lag structures in all equations which eases cross-sectional comparisons and the computation of cross-sectional averages, we choose a universal lag length of 10 for all stocks. This lag length is sufficiently close to the individually optimal lag length according to the Bayes Information Criterion and does not restrict the validity of the results discussed below. In the following we show the cross-sectional averages of point estimates and corresponding standard errors.

2.4.2 Estimation Results

Table 2.2 reports averaged estimates of the VAR model which is augmented by dummies indicating relevant news. For sake of brevity, we do not show coefficients for lags of the dependent variables greater than two. Likewise, coefficient estimates for the dummies Z_t are not reported.³ News dummies cover 40 seconds before the disclosure and 100 seconds thereafter.

Analyzing the dynamics of volatility and liquidity, we can summarize the following main results: First, all variables reveal significantly positive own dynamics. This is strongly expected given the underlying autocorrelations reported above. Second, we observe a significantly positive relationship between money value traded and volatility. Hence, volatility and trading activity are closely dependent not only on a daily level as, e.g., shown by Clark [1973] and Tauchen and Pitts [1983], but obviously also on a high-frequency level, confirming, e.g., Hautsch [2008]. Third, bid-ask spreads increase if past trading periods reflect rising liquidity demand and volatility. This causality is well

³These results are available upon request from the authors.

2 Quantifying High Frequency News-implied Market Reactions

VAR (I)	Money Value	Volatility	Abs. Trade Imb.	Depth	Spread	Av. Trade Sz.	Probit (I) $\mathbb{1}(y_{it} > 0)$
c	-0.666*** (0.115)	-0.120*** (0.076)	0.522*** (0.010)	0.239*** (0.017)	0.342*** (0.022)	0.159*** (0.038)	1.847*** (0.048)
<i>mv_{t-1}</i>	0.224*** (0.003)	0.046*** (0.005)	-0.003*** (0.000)	-0.001* (0.000)	0.001*** (0.000)	0.020*** (0.002)	0.018*** (0.003)
Money Value	0.130*** (0.002)	0.011*** (0.002)	-0.001*** (0.000)	0.001* (0.000)	-0.000 (0.000)	0.008*** (0.001)	0.004*** (0.001)
<i>volat_{t-1}</i>	0.000 (0.001)	0.203*** (0.004)	-0.002*** (0.000)	0.000 (0.000)	0.004*** (0.000)	0.008*** (0.000)	0.011* (0.001)
<i>volat_{t-2}</i>	-0.000 (0.001)	0.095*** (0.002)	-0.001*** (0.000)	0.000 (0.000)	0.002*** (0.000)	0.004*** (0.000)	0.003 (0.000)
Abs. Trade Imb.	<i>ati_{t-1}</i> 0.140*** (0.017)	0.118*** (0.016)	0.060*** (0.001)	0.002* (0.002)	-0.008*** (0.002)	-0.053*** (0.007)	0.222** (0.005)
	<i>ati_{t-2}</i> 0.079** (0.010)	0.012 (0.015)	0.043*** (0.001)	-0.007 (0.001)	-0.003 (0.002)	-0.024 (0.005)	0.147** (0.004)
Depth	<i>dp_{t-1}</i> 0.594*** (0.025)	-0.302*** (0.024)	0.008*** (0.001)	0.445*** (0.010)	-0.025*** (0.001)	0.398*** (0.001)	0.008*** (0.002)
	<i>dp_{t-2}</i> -0.028 (0.007)	0.008 (0.008)	-0.001 (0.000)	0.091*** (0.002)	-0.001 (0.001)	-0.002 (0.006)	0.008*** (0.001)
Spread	<i>spr_{t-1}</i> -0.359*** (0.016)	0.647*** (0.071)	0.006*** (0.001)	-0.097*** (0.006)	0.297*** (0.009)	-0.105*** (0.007)	-0.160*** (0.005)
	<i>spr_{t-2}</i> 0.053** (0.006)	0.044* (0.019)	-0.004*** (0.001)	-0.012*** (0.001)	0.070*** (0.004)	0.002 (0.003)	0.028*** (0.002)
Av. Trade Sz.	<i>ats_{t-1}</i> -0.030*** (0.003)	0.034*** (0.004)	-0.000 (0.000)	-0.007*** (0.001)	0.001 (0.000)	0.130*** (0.003)	0.005*** (0.002)
	<i>ats_{t-2}</i> -0.018** (0.003)	0.017** (0.002)	0.001 (0.000)	-0.000 (0.001)	0.000 (0.000)	0.094*** (0.003)	0.005*** (0.001)

2.4 Market Dynamics around News Items

VAR (II)		Money Value	Volatility	Abs. Trade Imb.	Depth	Spread	Av. Trade Sz.	Probit (II) $\mathbb{1}(y_{it} > 0)$
Dummy Leads	d_{t+2}	0.894 (0.329)	1.023 (0.354)	-0.013 (0.005)	0.018 (0.012)	0.004 (0.013)	0.211 (0.079)	0.030 (0.015)
	d_{t+1}	0.485* (0.145)	0.484* (0.182)	-0.009 (0.005)	0.002 (0.007)	-0.024 (0.011)	0.101 (0.055)	0.065 (0.018)
Item Dummy	d_t	1.036*** (0.224)	1.332*** (0.308)	-0.015 (0.004)	0.022 (0.012)	0.043 (0.014)	0.163 (0.042)	0.110 (0.019)
	d_{t-1}	1.244*** (0.236)	1.470*** (0.482)	-0.022 (0.004)	-0.011 (0.014)	0.035 (0.010)	0.201 (0.081)	-0.081 (0.016)
Lags	d_{t-2}	0.891*** (0.178)	0.925*** (0.330)	-0.009 (0.006)	-0.003 (0.012)	0.020 (0.010)	0.137 (0.048)	0.048 (0.015)
	d_{t-3}	0.588	0.909	-0.009	-0.015	-0.004	0.166	0.031
	d_{t-4}	0.434	0.413	-0.013	0.020	0.005	0.107	0.057
	d_{t-5}	0.307	0.549	-0.013	0.022	-0.000	0.076	0.024

Table 2.2: Average VAR Results: Dynamics ($y|y_{it} > 0$) and Dummies. *Note:* The first six columns show OLS estimation results of system (2.4). The last column shows the ML estimation results of the corresponding probit model (2.5). VAR (I) (Probit (I)) gives the dynamics and VAR(II) (Probit (II)) gives the estimates for the news dummies. Reported coefficients are averages of the estimates for each individual stock. Significance is reported based on average t-statistics. (Cross-sectional) standard errors of the averaged coefficients are given in parentheses below. (****) denotes significance of the average coefficient estimates at the 1 % level, (**) at the 5 % level, and (*) at the 10 % level.

confirmed by asymmetric information based market microstructure theory (e.g., Easley and O’Hara [1992]) where increased trading activity is an indicator for the existence of information and thus increased risks due to adverse selection. Our findings show that such situations are typically also characterized by increased trade sizes. Conversely, liquidity demand is reduced as a response to increased trading costs as induced by higher bid-ask spreads and reduced market depth.

Fourth, in contrast to the unconditional analysis in Section 2.3, significant effects induced by news items are only identifiable for volatility and cumulated trading volumes but not for spreads, absolute trade imbalances, average trade sizes and market depths. In particular, the insignificant spread dummies contradict corresponding results for earnings announcements. These results suggest that the (unconditional) reactions of these variables during news arrival periods as reported in Section 2.3 are mainly due to spillover effects arising from increased volatility and cumulated volumes but do not necessarily arise from news alone. Moreover, due to the persistence in the market dynamics, news-induced effects and pre-release trading activity are carried over to subsequent periods. It is therefore not surprising that the direct impact of news as captured by the dummy variables dies out relatively quickly. These findings show that ignoring dynamics and interdependencies can cause spurious results.

Fifth, confirming the results of section 2.3 we find significant effects only for relevant news. Indeed, filtering out noise and structuring news according to their relevance is even more important in a dynamic setting than in an unconditional framework.

The estimation results for the probit model widely confirm those for the VAR specification. However, the fact that all news dummy variables are insignificant indicates that the occurrence of a trade in a 20-sec interval is not significantly driven by news arrivals.

The averaged estimates capture the major features common to all assets, but most stocks still reveal idiosyncratic responses to news. Even though, for instance, the *average* spread reaction is insignificant, we still observe significant individual spread responses for 19 out of 39 stocks in the sample. Figure 2.11 shows that the significant (positive) dummies for most stocks center around the item arrival interval. Accordingly, there is evidence for news-implied reactions in spreads, depth and average trade sizes which are, however, diffuse across the stock universe. Stock-specific effects for the money value and volatility are much more in line with the average results as we detect significant reactions after news arrivals for all but three stocks.

2.4.3 Impulse Response Analysis

To provide more insights into news-implied market responses in a dynamic system, we perform an impulse response analysis. Here, a ’news shock’ is defined as a change in the corresponding news dummies. As the arrival of news generally stimulates trading activity, it is sufficient to conduct the analysis under the assumption that there is trading activity throughout post-release periods. In the following we thus use the conditional mean specification (2.4) for the periods $t, \dots, t + s$, where t denotes the news item arrival time.

2.4 Market Dynamics around News Items

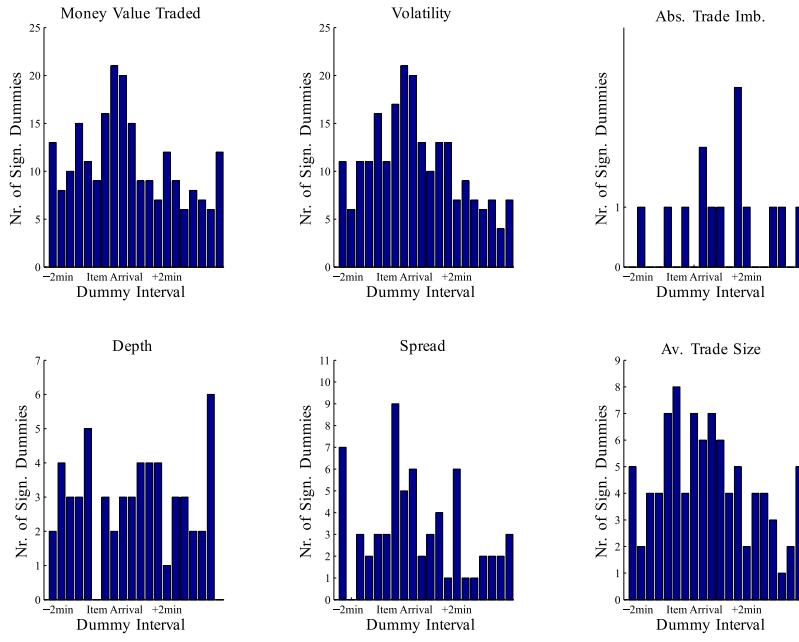


Figure 2.11: Numbers of significant dummy variables in the intervals around the news disclosure. Dummies cover 7 20-second intervals before and 13 after the news arrival.

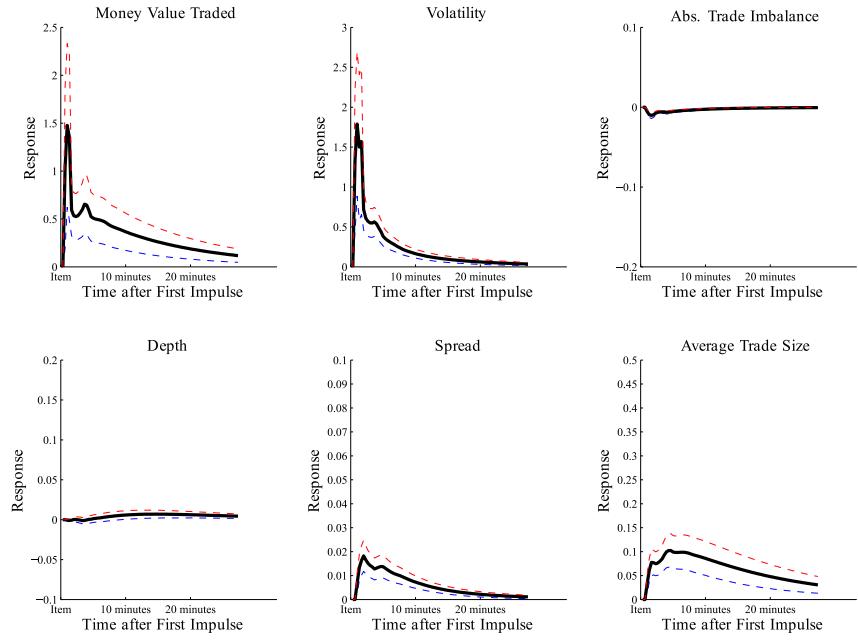


Figure 2.12: Response analysis of a change in the news dummies for highly relevant news (95% confidence intervals as dotted lines)

2 Quantifying High Frequency News-implied Market Reactions

Accordingly, the response after s periods to a news arrival in t is computed as

$$\Delta_s(\boldsymbol{\theta}_1) := \mathbb{E}[\mathbf{y}_{t+s} | \Omega_{t-1}, d_t = 1; \boldsymbol{\theta}_1] - \underbrace{\mathbb{E}[\mathbf{y}_{t+s} | \Omega_{t-1}, d_t = 0; \boldsymbol{\theta}_1]}_{(*)}, \quad (2.7)$$

where Ω_{t-1} represents the history of the multivariate process at t and the second term $(*)$ removes the effect of constants and initial values on the response function. Let $p_1 = 0$, $p_2 > 0$ and $\widehat{\boldsymbol{\Xi}}_{\cdot i}$ denote the i -th column of $\widehat{\boldsymbol{\Xi}}$. Coefficients in the second to p_2 -th columns of $\widehat{\boldsymbol{\Xi}}$ that are not significantly different from zero at the 5% level are assumed to be zero throughout. Initially we have

$$\begin{aligned} \Delta_0 &= \mathbb{E}[\mathbf{y}_t | \Omega_{t-1}, d_t = 1; \boldsymbol{\theta}_1] - \mathbb{E}[\mathbf{y}_t | \Omega_{t-1}, d_t = 0; \boldsymbol{\theta}_1] \\ &= \widehat{\mathbf{c}} + \sum_{i=1}^p (\widehat{\boldsymbol{\Gamma}}_i \mathbf{y}_{t-i} + \widehat{\boldsymbol{\Psi}}_i Z_{t-i}) + \widehat{\boldsymbol{\Xi}}_{\cdot 1} - \left(\widehat{\mathbf{c}} + \sum_{i=1}^p (\widehat{\boldsymbol{\Gamma}}_i \mathbf{y}_{t-i} + \widehat{\boldsymbol{\Psi}}_i Z_{t-i}) \right) = \widehat{\boldsymbol{\Xi}}_{\cdot 1}. \end{aligned}$$

Since the initial conditions, constants and Z_t cancel out, the responses in $t+s$, $s = 1, 2, \dots$, to the dummy impulse in t are given as

$$\Delta_1 = \widehat{\boldsymbol{\Gamma}}_1 \Delta_0 + \widehat{\boldsymbol{\Xi}}_{\cdot 2}, \quad \Delta_2 = \widehat{\boldsymbol{\Gamma}}_1 \Delta_1 + \widehat{\boldsymbol{\Gamma}}_2 \Delta_0 + \widehat{\boldsymbol{\Xi}}_{\cdot 3}, \dots$$

Standard errors of the response function are derived using the delta method. Accordingly, Δ_s is asymptotically distributed according to

$$\Delta_s(\widehat{\boldsymbol{\theta}}_1) \xrightarrow{d} N(\Delta_s(\boldsymbol{\theta}_1), (1/T) \mathbf{G}_s (\boldsymbol{\Omega} \otimes \mathbf{Q}^{-1}) \mathbf{G}'_s),$$

where $\mathbf{Q} = \mathbb{E}[\mathbf{x}_t \mathbf{x}'_t]$ and $\mathbf{G}_s = \frac{\partial \Delta_s(\boldsymbol{\theta}_1)}{\partial \boldsymbol{\theta}_1}$. Estimates for $\boldsymbol{\Omega}$ and \mathbf{Q} are readily available from the VAR estimates. Following Hamilton [1994], we construct the columns of \mathbf{G}_s based on finite differences according to

$$\frac{\partial \Delta_s(\widehat{\boldsymbol{\theta}}_1)}{\partial \theta_{1i}} \approx \frac{\Delta_s(\widehat{\boldsymbol{\theta}}_1 + \mathbf{e}_i h) - \Delta_s(\widehat{\boldsymbol{\theta}}_1)}{h},$$

where h is some small number, θ_{1i} denotes the i -th element of $\boldsymbol{\theta}_1$ and \mathbf{e}_i is the i -th unity vector.

Figure 2.12 shows the impulse response to news-induced dummy variable changes based on the averaged VAR estimates. The depicted reaction to relevant news mimics the unconditional market responses of volatility, money value traded, average trade sizes and bid-ask spread very well (cf. Figures 2.3 to 2.5). The percentage increase of volatility and money value traded over the pre-news period is roughly 150 % (given the mean one for each series) and thus corresponds to the unconditional effects. Again it turns out that responses in bid-ask spreads and trade sizes are induced through dynamic spill-overs from news effects in volatility and cumulative volumes.

Overall, we conclude that the dynamic analysis strongly confirms the unconditional effects shown above. Obviously, volatility and trading volume are most sensitive to news arrival. Reactions in money value traded and volatility are rather idiosyncratic and

due to spillovers. In order to check the robustness of our results, we have estimated several alternative specifications, in particular (i) a simple VAR model based on 20 second aggregates (without explicitly accounting for zero observations), and (ii) the corner-solution model by Cragg [1971] for the conditional density based on 20 second aggregates. For sake of brevity we refrain from reporting the corresponding estimates in the paper. However, it turns out that our findings are qualitatively quite stable across the individual specifications.

2.5 Conclusions

Recording and analyzing the overall news flow for a specific asset is challenging since the amount of news, the number of news sources and the speed of information dissemination is rapidly increasing over time. Due to the huge amount of information permanently published in all modern media, news are overlaid by substantial noise caused by irrelevant information. These effects make it difficult to identify significant links between high-frequency trading activity and the intraday news flow. As previous studies predominantly focus only on scheduled and homogenous types of news (typically earnings announcements), it is virtually unknown whether intraday trading activity, volatility and liquidity can be systematically linked to an intraday flow of news items other than regularly announced earnings figures.

To reduce the impact of noise, this study is the first one making use of unique data provided by an automated news analytics tool of the Reuters company. Designed for use in algorithmic trading applications and employing linguistic pattern recognition techniques, these novel news data allow us to disentangle relevant news from irrelevant ones and to identify the sign and the novelty of news items. Using this news engine our study explores the impact of news items on high-frequency returns, trading volume, volatility, depth and bid-ask spreads for a cross-section of stocks traded at the London Stock Exchange (LSE).

Analyzing the unconditional and conditional effects of news items on intraday trading activity we can summarize the following results. First, we identify significant unconditional market responses in volatility, money value traded, average trade sizes and bid-ask spreads. Given the fact that earnings announcements are explicitly discarded from the analysis these findings are remarkable and indicate that the news engine successfully filters the news flow. However, it turns out that only cumulative trading volumes and volatility are directly influenced by news releases, whereas reactions in bid-ask spreads and trade sizes are widely indirect and due to dynamic spillovers from volatility and volumes. Second, we confirm the usefulness of the machine-indicated relevance of news items. In fact, significant market responses to news are only observable for items which are identified as being relevant. Our results show that the classification is crucial to filter out noise and to identify significant relations between market activity and the news flow. Third, the news sentiment indicator has predictability for future price trends. However, significantly increased bid-ask spreads around public news arrivals render simple sentiment-based trading strategies rather unprofitable.

2 Quantifying High Frequency News-implied Market Reactions

Overall, our study shows that news engines are able to successfully structure and categorize the intraday news flow. This allows to deeply investigate the question to which extent high-frequency market activity is driven by information. Moreover, our findings provide first evidence on the usefulness of news engines in financial practice. In specific, the distinct volatility and liquidity effects around news are potentially relevant parameters in algorithmic trading applications.

A Appendix

A Note on the Computation of Standard Errors of Across-Market Averages

In the following we describe two ways of computing the mean reactions and their standard errors. The pooled average used in Section 2.3 is based on the model

$$X_i = \mu + \varepsilon_i, \quad \varepsilon_i \sim \text{i.i.d. } N(0, \sigma^2), \quad i = 1, \dots, n, \quad (2.8)$$

where we have suppressed the I_j index for the respective interval around the news item. Inference is based on the pooled estimator for the mean, $\bar{X} = 1/n \sum_{i=1}^n X_i$, where 95% confidence intervals are given as $\bar{X} \pm 2 * \hat{\sigma} / \sqrt{n}$ with $\hat{\sigma}^2 = \mathbf{e}'\mathbf{e}/(n - 1)$.

To account for the fact that the stocks have very different numbers of news items (see Table 2.1), we alternatively used group-specific means. Let n_s denote the number of news for stock s and let X_{sj} be the reaction of a certain (trading) variable of stock s to item j . For the average reaction of each of the n_n stocks (the group mean), $\bar{X}_s = 1/n_s \sum_{j=1}^{n_s} X_{sj}$, we assume

$$\bar{X}_s = \mu + \varepsilon_s, \quad \varepsilon_s \sim \text{i.i.d. } N(0, \sigma^2), \quad s = 1, \dots, n_n. \quad (2.9)$$

Then, inference is based on the estimator for the mean, $\bar{\bar{X}} = 1/n_n \sum_{s=1}^{n_n} \bar{X}_s$, where 95% confidence intervals are given as $\bar{\bar{X}} \pm 2 * \hat{\sigma} / \sqrt{n_n}$ with $\hat{\sigma}^2 = \mathbf{e}'\mathbf{e}/(n_n - 1)$.

Both approaches have their advantages. While the latter smoothes out the effect of a large number of news, it does not account for the within-group variation, which is captured by (2.8). The confidence intervals are slightly more conservative using (2.9). Nevertheless, all results of Section 2.3 hold using both procedures.

Figures

2 Quantifying High Frequency News-implied Market Reactions

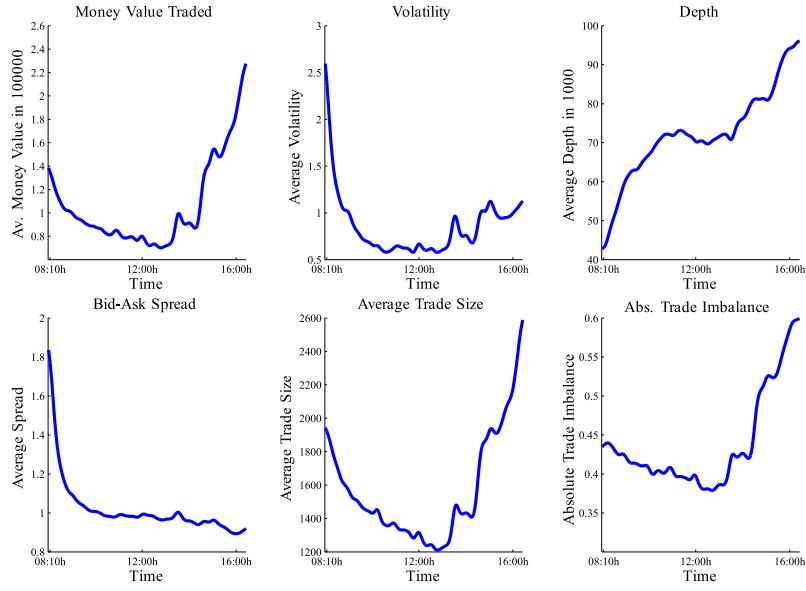


Figure 2.13: Average intraday seasonality patterns. Smoothed via kernel regression.

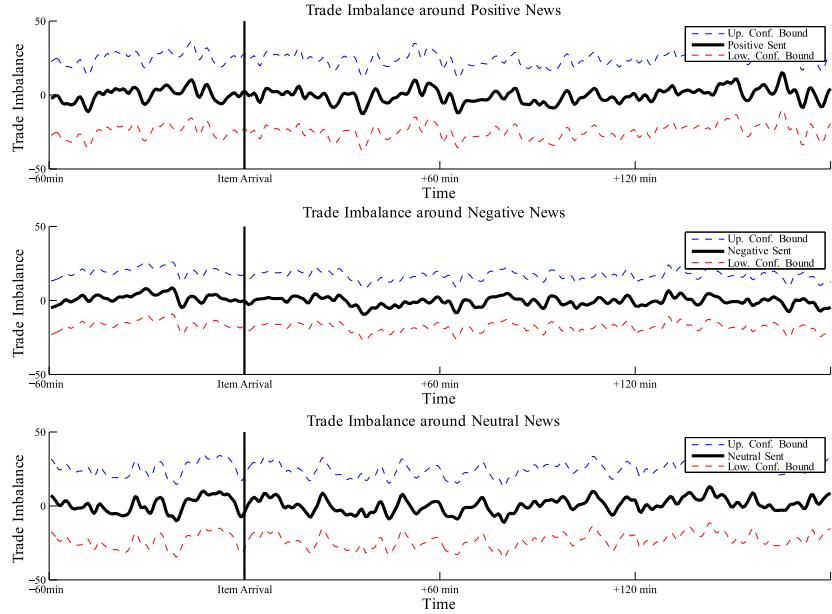


Figure 2.14: Signed trade imbalance around positive, negative and neutral news. Smoothed via kernel regression.

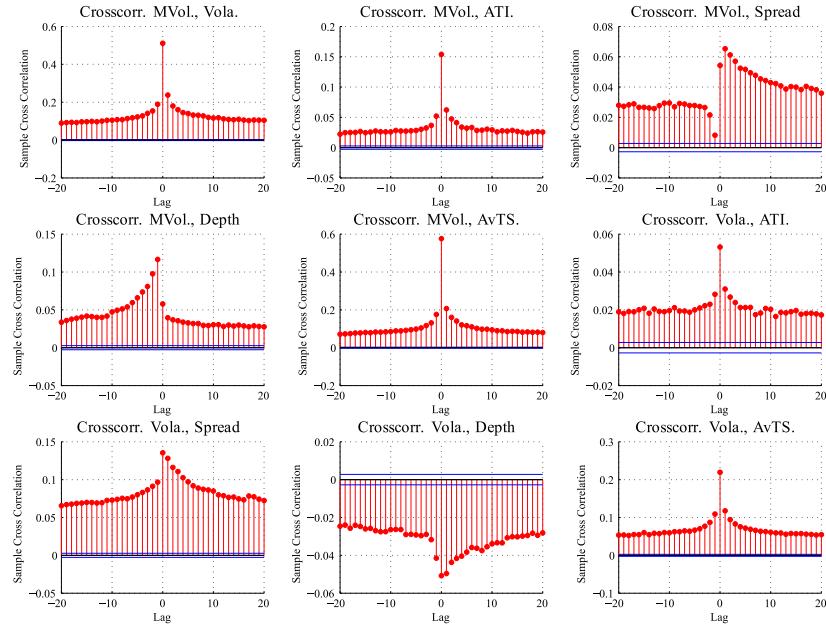


Figure 2.15: Sample average of cross-correlations (I)

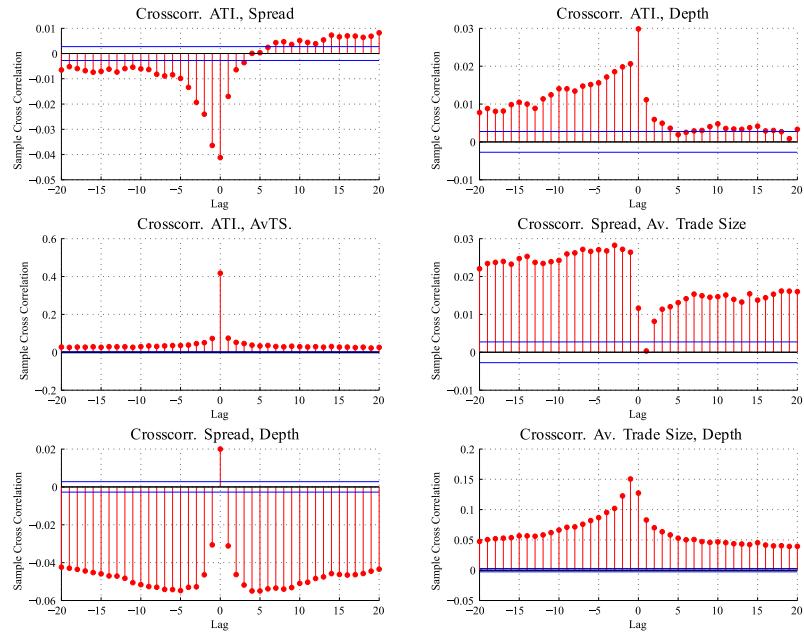


Figure 2.16: Sample average of cross-correlations (II)

3 A Forecasting Framework for Bid-Ask Spreads

3.1 Introduction

Bid-ask spreads reflect the fundamental costs of immediate trading, i.e., the cost of constantly guaranteeing a counterparty for trades to market participants. They are an important determinant of liquidity on stock markets and thus play a dominant role in the literature on market microstructure and stock trading. Theoretical models on market making along the lines of Copeland and Galai [1983], Kyle [1985], Glosten and Milgrom [1985] and Easley and O'Hara [1992] emphasize the adverse selection component in bid-ask spreads indicating information asymmetry among market participants. Based on game-theoretical dynamic models, Foucault [1999] and Foucault et al. [2005] argue that the bid-ask spread is *the* dominant parameter for the decision between different order types on stock markets. Traders can either be patient and submit limit orders or cross the spread and pay the bid-ask spread. Empirical studies confirm that limit and market order submission strategies indeed depend strongly on quoted spreads, see Biais et al. [1995], Harris and Hasbrouck [1996], Ranaldo [2004], Anand et al. [2005], Hall and Hautsch [2006] and Pascual and Veredas [2009].

Despite the importance of bid-ask spreads in trading applications and the relevance of spread predictions for the reduction of transaction costs, the question of how to statistically model bid-ask spreads has not been systematically addressed yet. Our study is the first contribution establishing a concise econometric methodology for modelling and forecasting bid-ask spreads on a high frequency.

Due to the discreteness of prices, bid-ask spreads are multiples of minimum price changes and hence form a time series of count variables. We observe that spreads reveal a pronounced seasonality pattern and are strongly serially dependent. Geweke and Porter-Hudak [1983] (GPH) tests indicate that the bid-ask spread time series exhibit long range dependence. To capture these empirical properties we introduce a novel count data model - the long memory autoregressive conditional Poisson (LMACP) model. Discussing empirical and theoretical properties we show that the model is suitable for the analysis and prediction of strongly persistent discrete time series.

Traditional approaches, such as Glosten and Harris [1988], George et al. [1991], Huang and Stoll [1997] and Bollen et al. [2004], decompose bid-ask spreads into their adverse selection, inventory holding and transaction components but are silent regarding their dynamic properties. Conversely, more recent time series approaches typically model bid-ask spreads implicitly in models for bid-ask quotes. Engle and Patton [2004] and Hautsch and Huang [2012], for instance, estimate error correction models for bid and

3 A Forecasting Framework for Bid-Ask Spreads

ask quotes with the bid-ask spread serving as cointegration relation. Hasbrouck [1999] models the dynamics of bid- and ask quotes separately but does not explicitly focus on the spread. Based on a vector autoregression framework, Taylor [2002] is the only approach deriving spread forecasts.

This chapter contributes to the recent literature on high frequency liquidity forecasting. See, e.g., Brownlees et al. [2011] on volume forecasts and Härdle et al. [2009] on forecasts of limit order book curves. Furthermore, the proposed LMACP model contributes to the literature on dynamic count data models as in Rydberg and Shephard [1999], Heinen [2003], Fokianos et al. [2009] and Ferland et al. [2006] among others. Finally, our study is related to count data predictions as in Sutradhar [2008], Jung et al. [2006] and Freeland and McCabe [2004].

Our forecast study is carried out based on representative stocks from the mid cap sector of the US Russell 3000 universe. We report rolling-window forecasts for quoted spreads on a 30 second frequency. The forecast evaluation of point and direction forecasts shows that LMACP models outperform competitors such as autoregressive moving average (ARMA), autoregressive fractionally integrated moving average (ARFIMA), autoregressive conditional duration (ACD) and fractionally integrated autoregressive conditional duration (FIACD) models. Four main results emerge from the analysis. First, we show the importance of explicitly addressing the discrete nature of bid-ask spreads. In particular, approaches based on continuous distributions are outperformed by Poisson models in terms of point, density and direction forecasts. Second, long memory specifications widely perform better than their short memory counterparts in terms of the root mean squared error and the directional accuracy. Third, additional predictors motivated from market microstructure theory, such as trading volume, volatility, first level depth and order imbalance improve forecasts. Fourth, an economic evaluation of a simple trading scheme reveals significant cost savings of up to 13 % when the trading schedules take bid-ask spread forecasts into account.

The remainder of this chapter is organized as follows. Section 3.2 gives descriptive statistics. In Section 3.3, we outline the econometric model. Section 3.4 describes the forecasting setup and corresponding evaluation criteria. In Section 3.5, the forecasting results are presented. Finally, Section 3.6 concludes.

3.2 Properties of Bid-Ask Spreads

3.2.1 The Bid-Ask Spread as a Liquidity Measure

Trading on financial markets requires the presence of a counterparty for trades. According to theoretical models along the lines of Copeland and Galai [1983] and Glosten and Milgrom [1985], liquidity suppliers like market makers, dealers or participants submitting limit orders act as intermediaries and mitigate the search costs by offering immediate trade execution.¹ Empirical studies like Glosten [1987], Glosten and Harris [1988] and Huang and Stoll [1997] suggest that quoted bid-ask spreads are measures of the costs

¹See Bessembinder and Venkataraman [2010] for an overview on spread-related trading costs.

of order processing, inventory holding and adverse selection incurred by these liquidity providers. Liquidity suppliers recoup their own costs in time t by purchasing at the bid price B_t and selling at a higher ask price A_t . As a measure of immediate trading and hence liquidity costs, the quoted bid-ask spread S_t in t is thus given by $S_t := A_t - B_t$, where the quotes A_t and B_t are given as multiples of 0.01 (price ticks). A closely related measure is the effective spread, given as $S_t^B := P_t - B_t$ for buyer-initiated trades and $S_t^A := A_t - P_t$ for seller-initiated trades, where P_t is the transaction price (in multiples of 0.01).

The importance of bid-ask spreads as liquidity measures for practitioners is reflected in limit order submission strategies employed by market participants to reduce trading costs. Limit and market order submission strategies depend strongly on quoted spreads and quoted depth as outlined in Biais et al. [1995], Harris and Hasbrouck [1996], Griffiths et al. [2000], Anand et al. [2005], Parlour [1998], Ranaldo [2004] and Pascual and Veredas [2009].

3.2.2 Empirical Properties

Histograms of spread distributions substantially vary over the Russell 3000 cross-section of the market at NYSE and NASDAQ. We show histograms for a large cross-section for January to February 2008 in the appendix to the working paper version, see Groß-Klußmann and Hautsch [2011b]. In our empirical analysis, we exclude stocks revealing "trivial" spread distributions with spreads being virtually constant to one tick. Rather, we focus on stocks with an average spread of more than two ticks revealing more dispersed distributions. The latter still cover a large fraction of the Russell 3000 index. Figure 3.1 shows that the average quoted spread for the constituents of the S&P 500 in 2010 is above 2. Moreover, as shown in Figure 3.1, average bid-ask spreads significantly vary over time. Particularly during the peak of the financial crisis in fall 2008, average spreads nearly doubled compared to the level before.

We employ national best bid and offer (NBBO) quotes and transaction data from the Trade and Quote (TAQ) database of the NYSE.² Bid-ask spreads are computed as end-point spreads based on a 30 second frequency. We omit the first and last ten minutes of the trading day to reduce the impact of trading starts and stops.

In the paper, we present results for four stocks from the mid-cap sector of the Russell index for January and February 2008 (GXP.N, EMN.N (traded at NYSE) and XRAY.OQ, EQIX.OQ (traded at NASDAQ)). The stocks are chosen to be representative for stocks below the big-cap sector of the Russell as well as for stocks of the big-cap sector with an average spread above two ticks. Figure 3.2 shows the distributions and snapshots of the time series evolution. Additional and robustifying results for a wide cross-section of stocks are provided in the appendix to Groß-Klußmann and Hautsch [2011b].

Due to the discreteness of price ticks (as multiples of 0.01), a time-ordered sequence of quoted bid-ask spreads multiplied by 100 forms a time series of count variables. Table

²According to the US Securities and Exchange Commission Regulation brokers are required to guarantee customers the best quoted prices across US-based exchanges.

3 A Forecasting Framework for Bid-Ask Spreads

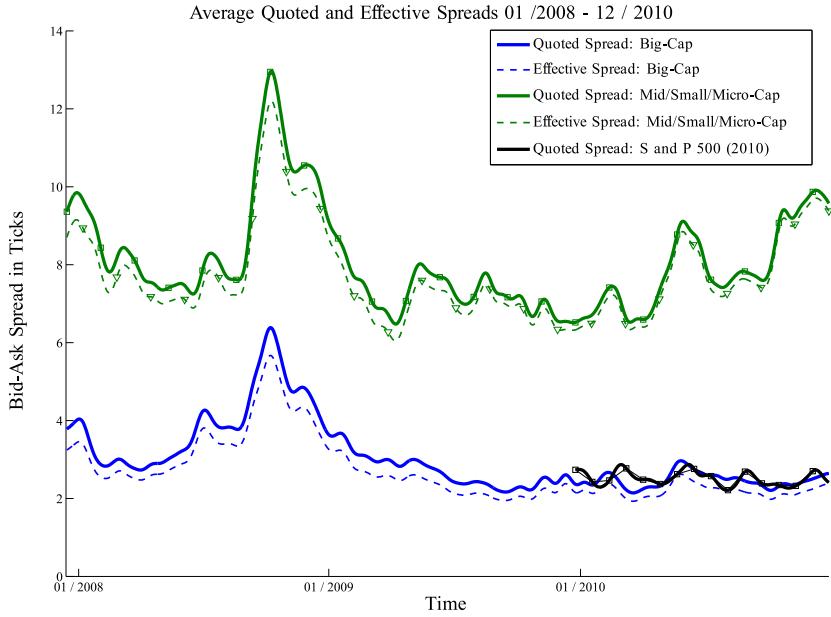


Figure 3.1: Evolution of the average quoted and average effective bid-ask spreads for 30 big-cap stocks and 80 mid-, small- and micro-cap stocks of the Russell 3000 as well as average quoted spreads of the S&P 500 for 2010. Smoothed via kernel regression

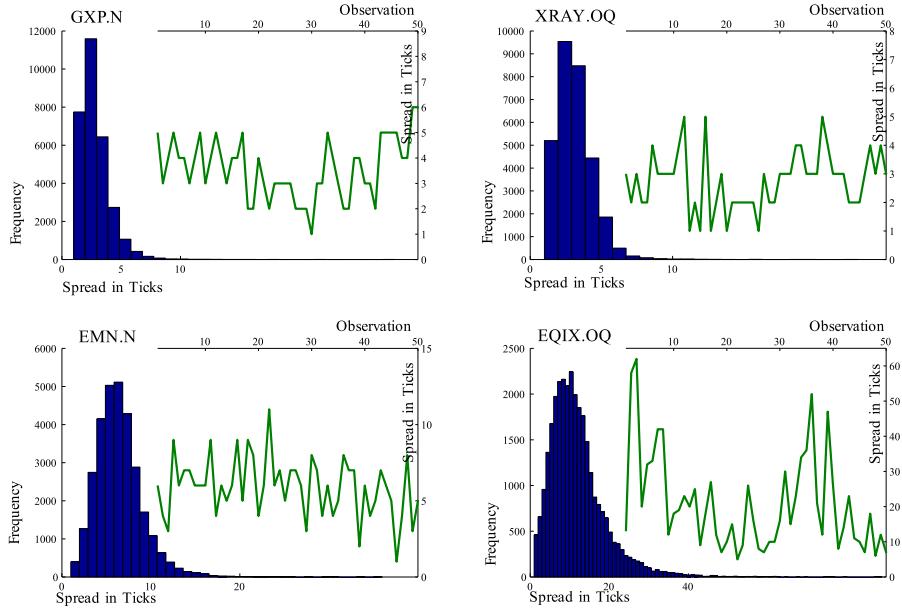


Figure 3.2: Histogram and typical pattern of the 30s bid-ask spreads for the four selected mid-cap stocks (histogram for Jan. and Feb. 2008)

3.2 Properties of Bid-Ask Spreads

	GXP.N	XRAY.OQ	EMN.N	EQIX.OQ
Mean	2.369	2.719	6.081	11.554
Variance	1.742	1.822	7.022	51.438
Max	28.000	23.000	36.000	89.000
Min	1.000	1.000	1.000	1.000
Median	2.000	3.000	6.000	10.000
10% Quantile	1.000	1.000	3.000	4.000
90% Quantile	4.000	4.000	9.000	20.000
LB_{20}	79853.9	71367.0	43658.7	46874.1
Average Mid-Quote	27.632	42.117	63.809	77.500
Relative Spread in %	8.6	6.5	9.5	14.9

Table 3.1: Descriptive statistics of the 30s quoted spread in ticks. LB_{20} denotes the Ljung-Box statistic for 20 lags. The relative spread is the spread fraction of the mid-quote price. Sample period: Jan.-Feb. 2008

3.1 gives descriptive statistics for the 30 second spread time series of the selected stocks. We observe that spread distributions can be both over- and underdispersed, i.e., have variance above or below the means. As shown in Figure 3.3, the autocorrelation functions (ACFs) of bid-ask spread series decay very slowly and indicate long range dependence. In light of the bid-ask spread as a cointegration relation between ask and bid quotes (see Engle and Patton [2004] and Hautsch and Huang [2012]), these results mean that deviations from equilibria are very persistent.

A time series formally is long range dependent if

$$\lim_{j \rightarrow \infty} \rho_j / (cj^{-\alpha}) = 1, \quad (3.1)$$

where $\alpha \in (0, 1)$, $c > 0$ and ρ_j denotes the j th order autocorrelation, see, e.g., Beran [1998]. An immediate consequence is that autocorrelations are not absolutely summable. Long range dependence is often found in financial market data (see, among many others, Ding and Granger [1996], Andersen et al. [2003] and Corsi [2009] for volatility data, Lux and Kaizoji [2007] for traded volumes and Deo et al. [2010] for trade durations) as well as in macroeconomic time series (see Bhardwaj and Swanson [2006] for an overview). The presence of long range dependence in spreads is supported by Figure 3.4 showing a convergence rate of the mean slower than \sqrt{n} for two of the four stocks. More formally, we conduct the Geweke and Porter-Hudak [1983] (GPH) test for long memory, which is based on the spectral regression

$$\ln(I(\omega_\lambda)) = a + b \ln \left(4 \sin^2 \left(\frac{\omega_\lambda}{2} \right) \right) + n_\lambda, \quad \lambda = 1, \dots, v, \quad (3.2)$$

where $I(\omega_\lambda)$ is the periodogram of the time series with sample size T at the frequencies $\omega_\lambda = \frac{2\pi\lambda}{T}$. Following Diebold and Rudebusch [1989] we select $v = \sqrt{T}$. Table 3.2

3 A Forecasting Framework for Bid-Ask Spreads

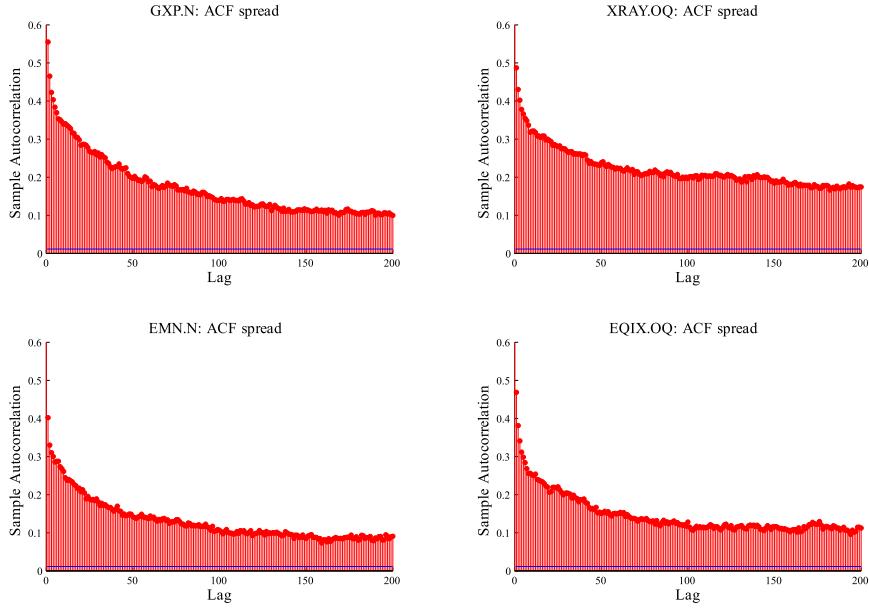


Figure 3.3: Autocorrelation functions of the 30s spreads for the mid-cap stocks (1-200 lags)

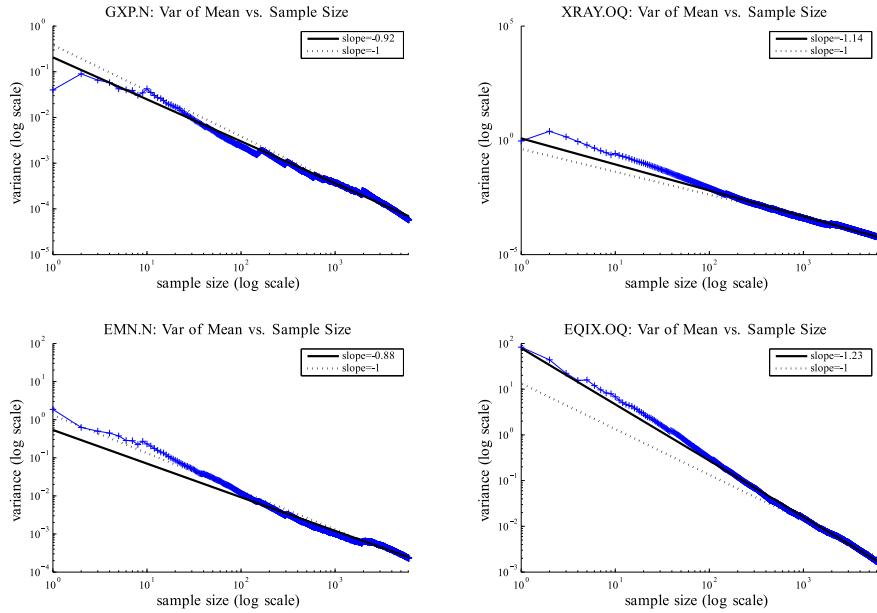


Figure 3.4: Log-log of variance of mean vs. sample size. The dotted line has slope -1 . The thick line is the regression line through the variance of the mean per sample size

3.2 Properties of Bid-Ask Spreads

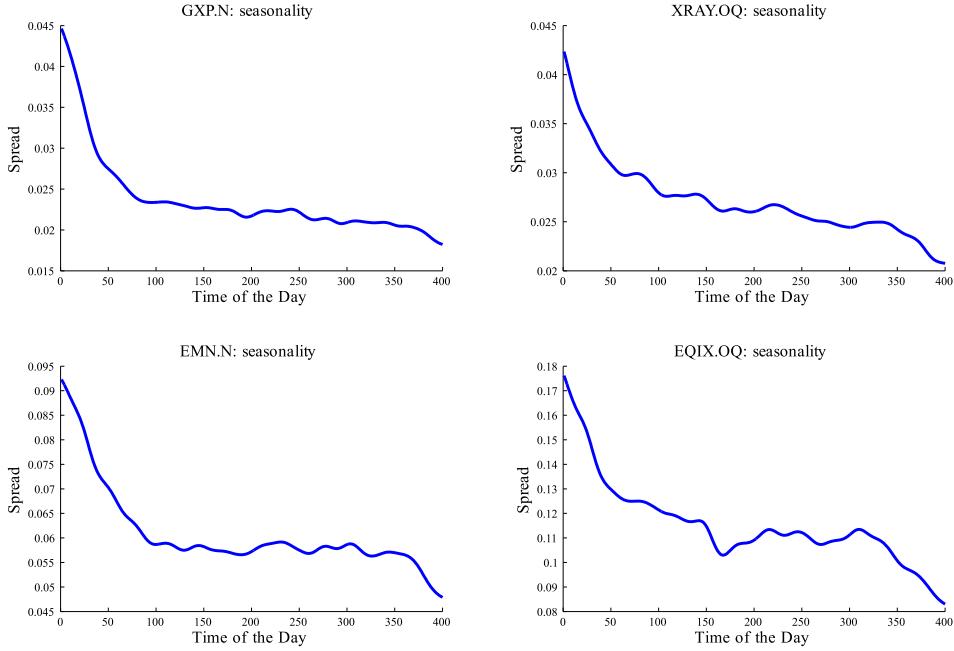


Figure 3.5: Intraday Periodicities for the 30s bid-ask spread series in January and February 2008 for the mid-cap stocks. Smoothed via kernel regression

	GXP.N	XRAY.OQ	EMN.N	EQIX.OQ
Estimate	0.455	0.439	0.470	0.303
T-Stat	8.868	8.559	9.148	5.901
P-Value	0.000	0.000	0.000	0.000

Table 3.2: GPH Test of the spread series

shows that the estimates of the long memory parameter b are significant and in the range $(0, 0.5)$. This indicates that the series is covariance stationary but long range dependent.

As reflected by Figure 3.5, we observe pronounced intraday periodicities in line with Chan et al. [1995] as well as Chung et al. [1999]. Bid-ask spreads are increased in the beginning of a trading day and decline in the course of the trading session. Such a pattern is explained by a higher adverse selection component in spreads due to the processing of overnight information in the morning.

3.3 An Econometric Model for Bid-Ask Spread Dynamics

Traditional approaches for time series of count variables are parameter-driven models based on Zeger [1988], Bayesian count data models in the spirit of Harvey and Fernandez [1989], Hidden Markov models (see MacDonald and Zucchini [1997]) or integer autoregressive moving average (INARMA) models (see McKenzie [2003]). Though conceptually elegant, these approaches suffer from tedious estimation procedures which makes them intractable in extensive high-frequency applications. As an alternative, the more recent literature focusses rather on observation-driven models, such as the autoregressive conditional Poisson (ACP) model as introduced by Rydberg and Shephard [1999] and put forward by Heinen [2003], Fokianos et al. [2009] and Ferland et al. [2006]. In contrast to the aforementioned models, ACP models are straightforward to estimate and tractable even for a large number of observations. Moreover, in contrast to Hidden Markov Models or the Autoregressive Multinomial Model proposed by Engle and Russell [2005], the ACP does not require to specify the states of the dependent variable prior to the estimation.

3.3.1 The Autoregressive Conditional Poisson model

Since bid-ask spreads are strictly positive but the Poisson distributions (and extensions thereof) are defined on $\mathbb{N} \cup \{0\}$, we follow Rydberg and Shephard [2003] and shift the spread distributions without loss of generality by one tick to the left. Accordingly, the spread process is re-defined as $S_t := \{(\text{spread in number of ticks}) - 1\}$.

Let $\mathcal{P}(\lambda_t)$ denote the Poisson distribution with mean λ_t and let \mathcal{F}_t denote the information available in t . Moreover, let two polynomials α and β be given as $\alpha(B) := \alpha_1 B + \alpha_2 B^2 + \dots + \alpha_q B^q$ and $\beta(B) := \beta_1 B + \beta_2 B^2 + \dots + \beta_p B^p$, where B is the backshift operator and $\alpha_i > 0$, $i = 1, \dots, q$, as well as $\beta_i > 0$, $i = 1, \dots, p$. Then, the autoregressive conditional Poisson process $\{S_t\}_{t \in \mathbb{Z}}$ is given by

$$\begin{aligned} S_t | \mathcal{F}_{t-1} &\sim \mathcal{P}(\lambda_t), \quad \forall t \in \mathbb{Z}, \\ \lambda_t &= c + \alpha(B)S_t + \beta(B)\lambda_t, \end{aligned} \tag{3.3}$$

where $c > 0$. Under (3.3) the conditional probability mass function of $S_t = s$, $s = 0, 1, 2, \dots$ is

$$\mathbb{P}(S_t = s | \lambda_t) = \frac{\lambda_t^s}{s!} e^{-\lambda_t}. \tag{3.4}$$

As shown by Ferland et al. [2006], the series $\{S_t\}$ is covariance stationary as well as strictly stationary if $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1$. Fokianos et al. [2009] derive the ergodicity conditions for a covariance stationary ACP process in the case $q = p = 1$.³

Rearranging (3.3), the ACP mean equation can be rewritten as ARMA ($\max\{p, q\}, p$)

³Alternative conditional mean specifications are given in the Poisson-based count data models outlined in Davis et al. [1999] and Davis et al. [2003].

3.3 An Econometric Model for Bid-Ask Spread Dynamics

specification of the form

$$(1 - \phi(B)) \left(S_t - \frac{c}{(1 - \phi(1))} \right) = (1 - \beta(B)) \nu_t, \quad (3.5)$$

with $\phi(B) := \alpha(B) + \beta(B)$ and $\nu_t := S_t - \lambda_t$ being a martingale difference sequence. As discussed by Ferland et al. [2006], the autocorrelation functions of the representations (3.5) and (3.3) are identical which makes it appropriate to interpret (3.3) as a conditional mean rather than a conditional variance model⁴.

While the mean and variance of the ACP process (3.3) are assumed to be conditionally equal, the variance of the ACP process is unconditionally greater or equal to the unconditional mean. As shown by Heinen [2003], in case of $q = p = 1$ we have

$$\mathbb{E}[S_t] = \frac{c}{1 - (\alpha_1 + \beta_1)}, \quad \text{Var}[S_t] = \frac{\mathbb{E}[S_t](1 - (\alpha_1 + \beta_1)^2 + \alpha_1^2)}{1 - (\alpha_1 + \beta_1)^2} \geq \mathbb{E}[S_t]. \quad (3.6)$$

Hence, the ACP specification can generate unconditional overdispersion and the marginal distribution of S_t is no longer Poisson.

To account for the possibility of both conditional as well as unconditional overdispersion and underdispersion, Heinen and Rengifo [2007] propose using the Double Poisson distribution proposed by Efron [1986] and defined by

$$\mathbb{P}(S_t = s | \lambda_t, \gamma) = c(\gamma, \lambda_t) \cdot \gamma^{1/2} e^{-\gamma \lambda_t} \left(\frac{e^{-s} s^s}{s!} \right) \left(\frac{e \lambda_t}{s} \right)^{\gamma s}, \quad s=0,1,2,\dots, \quad (3.7)$$

where the constants $c(\gamma, \lambda_t)$ can be numerically approximated by

$$\frac{1}{c(\gamma, \lambda_t)} = 1 + \frac{1 - \gamma}{12 \lambda_t \gamma} \left(1 + \frac{1}{\lambda_t \gamma} \right). \quad (3.8)$$

The Double Poisson distribution nests the Poisson for $\gamma = 1$.

Accordingly, a so-called Autoregressive Conditional Double Poisson (ACDP) model is given by

$$S_t | \mathcal{F}_{t-1} \sim \mathcal{DP}(\lambda_t, \gamma), \quad \forall t \in \mathbb{Z}, \quad (3.9)$$

where \mathcal{DP} denotes the Double Poisson distribution and λ_t is parameterized as in (3.3). The conditional variance of the ACDP model is given by $\text{Var}[S_t | \mathcal{F}_{t-1}] = \lambda_t / \gamma$ with $\gamma > 1$ ($\gamma < 1$) reflecting conditional underdispersion (overdispersion). As shown by Heinen [2003], the Double Poisson assumption can overcompensate the overdispersion generated by the dynamic mean specification resulting in unconditional underdispersion. In particular, in the case $p = q = 1$, the ACDP generates unconditional underdispersion if $\gamma > \frac{1 - (\alpha_1 + \beta_1)^2 + \alpha_1}{1 - (\alpha_1 + \beta_1)}$.⁵

⁴Recall, that under the Poisson assumption λ_t equals both the conditional mean and the conditional variance.

⁵An alternative generalization of the Poisson distribution is the Negative Binomial distribution. However, as it can only account for overdispersion in the data, it is less flexible.

3 A Forecasting Framework for Bid-Ask Spreads

Explanatory variables are easily included in the given setting using an exponential link function. Let x_t denote a k -dimensional vector of covariates without a constant and let γ denote the corresponding parameter vector. Then, the conditional mean is re-defined as $\mathbb{E}[S_t | \mathcal{F}_{t-1}] := \lambda_t \exp(x'_t \gamma)$, with λ_t given by (3.3).

AC(D)P models are straightforwardly estimated by maximum likelihood. In case of an ACDP specification the log likelihood function is given by

$$\ln \mathcal{L}(\cdot, \gamma) = \sum_{t=1}^T \left(\frac{1}{2} \ln(\gamma) - \gamma \lambda_t + S_t (\ln(S_t) - 1) - \ln(S_t!) + \gamma S_t \left(1 + \ln \left(\frac{\lambda_t}{S_t} \right) \right) \right), \quad (3.10)$$

where the constants $c(\gamma, \lambda_t)$ are omitted.

3.3.2 Long Memory Autoregressive Conditional Poisson Models

To account for long range dependence in spread series, we propose two types of long memory autoregressive conditional Poisson models. Both specifications capture hyperbolically decaying autocorrelation functions and are motivated by recent advances in long memory volatility models.

A building block of a long memory model is the fractional differencing operator $(1 - B)^d$ (see Hosking [1981]) which is a polynomial defined in terms of the hypergeometric function F and can be serially expanded according to

$$(1 - B)^d = F(-d, 1, 1; B) = \sum_{j=0}^{\infty} \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)} B^j. \quad (3.11)$$

The two types of models considered below differ in the way how $(1 - B)^d$ enters the conditional mean specification which has strong implications for the existence of first and second unconditional moments. Practically they differ by providing different forecasts which is analyzed in depth in Section 3.5.⁶

1) LMACP Type I

The so-called Long Memory ACP (LMACP) type I specification is obtained by augmenting the ARMA representation (3.5) by $(1 - B)^d$ resulting in

$$(1 - B)^d (1 - \phi(B)) (S_t - \omega) = (1 - \beta(B)) \nu_t, \quad (3.12)$$

where $\omega \in \mathbb{R}_0^+$. The polynomials ϕ and β are defined as in (3.5) with the roots of $(1 - \phi(B))$ and $(1 - \beta(B))$ lying outside the unit circle. In the GARCH case, a corresponding specification has been proposed by Karanasos et al. [2004] based on an ARMA representation of GARCH processes and is closely related to the models by Zaffaroni [2004], Koulikov [2003] and Giraitis et al. [2004].

⁶An alternative mean specification that accommodates both short and long memory behavior is the HYGARCH conditional mean specification. See Conrad [2010] for details.

3.3 An Econometric Model for Bid-Ask Spread Dynamics

Expressing (3.12) in terms of λ_t , a LMACP type I process is obtained by

$$\begin{aligned} S_t | \mathcal{F}_{t-1} &\sim \mathcal{P}(\lambda_t), \quad \forall t \in \mathbb{Z}, \\ \lambda_t &= \frac{(1 - \phi(B))(1 - B)^d}{(1 - \beta(B))} \omega + \Psi(B)S_t = \Psi(B)S_t, \end{aligned} \tag{3.13}$$

where the polynomial Ψ is given as

$$\Psi(B) := 1 - \frac{(1 - \phi(B))(1 - B)^d}{(1 - \beta(B))} = \sum_{i=1}^{\infty} \psi_i B^i \tag{3.14}$$

and $\frac{(1-\phi(B))(1-B)^d}{(1-\beta(B))}\omega = 0$. Accordingly, the long memory autoregressive conditional Double Poisson (LMACDP) model is given by $S_t | \mathcal{F}_{t-1} \sim \mathcal{DP}(\lambda_t, \gamma)$, $\forall t \in \mathbb{Z}$ with $\lambda_t = \Psi(B)S_t$ and nests the Poisson case for $\gamma = 1$.

Based on the representation

$$\lambda_t = \omega + (\Omega(B) - 1) \nu_t, \tag{3.15}$$

$$\Omega(B) := (1 - \Psi(B))^{-1} = \sum_{i=0}^{\infty} \omega_i B^i, \tag{3.16}$$

we observe that ω corresponds to the unconditional mean and is finite.

Conditions for the non-negativity of the conditional mean λ_t are identical to those for long memory GARCH (see Conrad and Haag [2006]) or fractionally integrated GARCH (FIGARCH) processes (Baillie et al. [1996]). In the empirically relevant case $0 < \beta_1 < 1$ and $p = q = 1$ we have:

Proposition 1. *Let $f_i = \frac{i-1-d}{i}$ for $i = 1, 2, \dots$ and let $\phi_1 = \alpha_1 + \beta_1$. The conditional mean of the LMACDP with order $p = q = 1$ is nonnegative a.s. if $0 < \beta_1 < 1$ and either $\psi_1 \geq 0$ and $\phi_1 \leq f_2$ or for $k > 2$ with $f_{k-1} < \phi_1 \leq f_k$ it holds that $\psi_{k-1} \geq 0$.*

Proof. See the proof in Conrad and Haag [2006] for FIGARCH models which directly applies to the mean specification of the long memory (Double) Poisson. \square

The next proposition establishes the unconditional variance $\text{Var}[S_t]$. Notably, in contrast to the fourth moment of long memory GARCH models the second moment of S_t exists without imposing additional conditions on the process S_t .

Proposition 2. *The unconditional variance of the long memory autoregressive conditional Double Poisson model type I is given by*

$$\text{Var}[S_t] = \frac{1}{\gamma} \mathbb{E}[\lambda_t] \sum_{j=0}^{\infty} \omega_j^2 < \infty. \tag{3.17}$$

Proof. See the appendix. \square

3 A Forecasting Framework for Bid-Ask Spreads

Accordingly, in the LMACP model ($\gamma = 1$), the dynamic specification induces unconditional overdispersion since $\sum_{j=0}^{\infty} \omega_j^2 \geq 1$ and thus $\text{Var}[S_t] = \mathbb{E}[\lambda_t] \sum_{j=0}^{\infty} \omega_j^2 \geq \mathbb{E}[\lambda_t]$.⁷ The overdispersion due to the dynamic specification is dependent on the parameter d via $\omega_j \approx C j^{d-1}$. However, the parameter γ of the Double Poisson distribution in the LMACDP model can generate underdispersion in case of $\gamma > \sum_{j=0}^{\infty} \omega_j^2$. Likewise overdispersion is captured if $\gamma < \sum_{j=0}^{\infty} \omega_j^2$.

Finally, from the representation $S_t = \omega + \Omega(B)\nu_t$ we straightforwardly get the n -order autocorrelation as

$$\rho_n(S_t) = \frac{\sum_{j=0}^{\infty} \omega_j \omega_{j+n}}{\sum_{j=0}^{\infty} \omega_j^2}. \quad (3.18)$$

Consequently, the LMACDP type I model is covariance stationary for $0 < d < 0.5$.⁸ Using $\omega_j \approx C j^{d-1}$ for high j we have $\rho_n(S_t) \approx C^* n^{2d-1}$ which in turn implies that $\lim_{n \rightarrow \infty} \sum_{k=0}^n |\rho_k(S_t)|$ is divergent.

2) LMACP Type II

The LMACP type II model is motivated by the specification of a FIGARCH model for volatility processes and the fractionally integrated autoregressive conditional duration (FIACD) model proposed by Jasiak [1998]. It builds on the following representation of the conditional mean,

$$(1 - \phi(B))(1 - B)^d S_t = \omega + (1 - \beta(B))\nu_t, \quad (3.19)$$

where $\omega \in \mathbb{R}_0^+$. Rearranging, the LMACP type II is then defined as

$$\begin{aligned} S_t | \mathcal{F}_{t-1} &\sim \mathcal{P}(\lambda_t), \quad \forall t \in \mathbb{Z}, \\ \lambda_t &= \frac{\omega}{(1 - \beta(B))} + \Psi(B)S_t, \end{aligned} \quad (3.20)$$

where Ψ is the polynomial (3.14). For the LMACDP type II we correspondingly change the conditional distribution assumption to $S_t | \mathcal{F}_{t-1} \sim \mathcal{DP}(\lambda_t, \gamma)$. The non-negativity of the conditional mean is guaranteed as long as the conditions of Proposition 1 are fulfilled (see Conrad and Haag [2006]).

The major difference to the type I specification is that the unconditional mean of the type II model is not finite since for $d < 0.5$ the coefficients of the power expansion of $(1 - B)^{-d}$ for $B = 1$ are not summable. This result is analogous to the difference between FIGARCH processes and long memory GARCH specifications proposed by Karanasos et al. [2004]. Hence, the expectation $\mathbb{E}[S_t] = \mathbb{E}[\Gamma(1)\omega + \Gamma(B)(1 - \beta(B))\nu_t] = \mathbb{E}[\Gamma(1)\omega]$ is not defined for $\Gamma(B) := (1 - \phi(B))^{-1}(1 - B)^{-d}$. The type II model is thus not covariance stationary and the long memory condition (3.1) cannot be directly verified since second

⁷ Details on why $\sum_{j=0}^{\infty} \omega_j^2 \geq 1$ can be found in the technical appendix.

⁸ Note that in the case $0.5 \leq d \leq 1$ $\lim_{k \rightarrow \infty} \sum_{i=0}^k \omega_i^2$ does not converge and thus the LMACDP is no longer covariance stationary.

moments do not exist.⁹

However, computing impulse response functions we show that the model can still capture long range dependence. The arguments follow the work of Conrad and Karanasos [2006] on long memory GARCH models. The impulse response function of the LMACDP type II is defined in terms of the sequence δ_k , $k = 0, 1, \dots$,

$$\delta_k := \frac{\partial \mathbb{E}[S_{t+k} | \mathcal{F}_t]}{\partial \nu_t} - \frac{\partial \mathbb{E}[S_{t+k-1} | \mathcal{F}_t]}{\partial \nu_t}. \quad (3.21)$$

Then, the cumulative impulse response function is given by $\lambda_k := \sum_{l=0}^k \delta_l$, $k = 0, 1, \dots$, where δ_k can be derived from the first difference in S_t ,

$$(1 - B)S_t = \frac{\omega}{(1 - \phi(B))(1 - B)^{d-1}} + \underbrace{(1 - B)\Omega(B)}_{=: \Delta(B)} \nu_t, \quad (3.22)$$

with $\Delta(B) := \sum_{j=0}^{\infty} \delta_j B^j$. The impulse response weights can also be recovered from the cumulative impulse responses by $\Delta(B) = (1 - B)\Lambda(B)$, where $\Lambda(B) := \sum_{k=0}^{\infty} \lambda_k B^k$. From (3.22) we deduce that $\Lambda(B) = \Omega(B)$. Thus, in the long run, shocks to the mean die out because $\Delta(1) = 0$, i.e.,

$$\lim_{k \rightarrow \infty} \sum_{l=0}^k \delta_l = \lim_{k \rightarrow \infty} \lambda_k = \Delta(1) = 0. \quad (3.23)$$

The shocks exhibit a slow, hyperbolic decay rate dependent on the parameter d since $\lambda_k = \omega_k \approx Ck^{d-1}$ for high k . Since this behavior is present also for $0.5 \leq d \leq 1$, we relax the restriction implied by the type I model and require $0 < d \leq 1$ in the type II specification.

3.4 Forecasting Bid-Ask Spreads

3.4.1 Computation of Forecasts

We evaluate out-of-sample forecasts based on a rolling window setup where the underlying model is re-estimated every 10 minutes to quickly adapt to potential changes in parameters. In particular, we conduct the following steps for the Jan./Feb. 2008 sample of 30s bid-ask spreads:

- (i) (*Estimation*) Estimate the econometric model based on an estimation window corresponding to five trading days of 30 second spread data.
- (ii) (*Forecasting*) Using the parameter estimates from (i), derive successive one-step ahead forecasts for the 10 minute horizon ahead of the estimation window.

⁹While Baillie et al. [1996] argue that the strict stationarity holds for FIGARCH models based on the results by Bougerol and Picard [1992], a similar argument does not hold for the LMACDP type II model as the conditional mean cannot be factored out from the Poisson distribution.

3 A Forecasting Framework for Bid-Ask Spreads

- (iii) (*Rolling forward the windows*) Move estimation and forecast window forward 10 minutes.

The models are estimated based on truncations of the (infinite) power expansion of $(1 - B)^d$. Pre-estimation analysis shows that a truncation point of 250 observations is sufficient to obtain reliable estimates which are widely independent of the truncation.

We evaluate point forecasts, $S_{t+1|t}$, and directional forecasts, $D_{t+1|t}$, defined as

$$S_{t+1|t} := \left[\widehat{\lambda}_{t+1} \right], \quad (3.24)$$

$$D_{t+1|t} := \mathbb{1}_{\{S_{t+1|t} > S_t\}} - \mathbb{1}_{\{S_{t+1|t} < S_t\}}, \quad (3.25)$$

where $\widehat{\lambda}_{t+1}$ is the mean forecast for $t + 1$ based on the conditional mean specification and $[\cdot]$ rounds its argument to the nearest integer. Hence, $D_{t+1|t} \in \{-1, 0, 1\}$ if spreads increase, are constant and decrease, respectively.

3.4.2 Additional Predictors based on Market Microstructure Theory

Decomposing the bid-ask spread into its components, Glosten and Harris [1988], George et al. [1991], Huang and Stoll [1997] and Bollen et al. [2004], among others, identify adverse selection and order processing costs as the main factors driving the spread. The adverse selection component of spreads is highly related to the amount of information asymmetry in the market. To capture states of high uncertainty and imbalances in the market, we include the realized volatility, given as the sum of squared mid-quote returns over each 30s interval and, alternatively, the absolute 30s mid-quote returns. Moreover, we construct measures of relative trade imbalance and relative depth imbalance to account for asymmetries in trading. The relative trade imbalance is given as

$$Timb_t := \frac{\left| \sum_{\tau=t_1}^{t_m} V_\tau \cdot \mathbb{1}_{\{q_\tau=-1\}} - \sum_{\tau=t_1}^{t_m} V_\tau \cdot \mathbb{1}_{\{q_\tau=1\}} \right|}{\sum_{\tau=t_1}^{t_m} V_\tau}, \quad (3.26)$$

where $\mathbb{1}_{\{\cdot\}}$ denotes the indicator function and V_1, V_2, \dots, V_m are the trade sizes corresponding to the time points of trades t_1 to t_m in a 30 second interval. The trade indicator q_t classifies trades into buys (+1) and sells (-1) according to the Lee and Ready [1991] algorithm. The relative depth imbalance is defined as

$$DPimb_t := \frac{|Adpt_t - Bdpt_t|}{Adpt_t + Bdpt_t}, \quad (3.27)$$

where $Adpt_t$ denotes the best ask depth and $Bdpt_t$ the best bid depth. As additional predictors we include the overall depth, given as the sum of the order book depth at best bid and ask level as well as the 30s cumulative trading volume serving as proxies for possible adverse selection in the market.

Finally, intraday periodicities in spreads are captured by a flexible Fourier form as

proposed by Gallant [1981],

$$s(t) = \delta^s \bar{t} + \sum_{j=1}^Q (\delta_{1,j}^s \cos(\bar{t}2\pi j) + \delta_{2,j}^s \sin(\bar{t}2\pi j)), \quad (3.28)$$

where $\delta^s, \delta_{1,j}^s$ and $\delta_{2,j}^s$ are parameters and $\bar{t} \in [0, 1]$ is the normalized intraday time defined as the time elapsed from the beginning of a trading day until observation t , divided by the length of the trading day.

In addition to the static inclusion of covariates, we alternatively conduct the following adaptive selection of the covariates in each step to allow for possible structural changes (see Blaskowitz and Herwartz [2009] for a related setup):

- (i) Estimate an AR model for spreads with all covariates based on observations within the estimation window. The AR setup is chosen here because (least squares) estimates can be computed in closed form which significantly reduces the computation burden in the rolling window framework.
- (ii) Discard predictors which are insignificant according to heteroscedasticity and autocorrelation consistent standard errors and execute steps (i) to (iii) from the scheme in 4.1 using only the remaining relevant predictors.

3.4.3 Forecast Benchmarks

To benchmark our approach, we compute forecasts using the following competing models:

- (i) A random walk model ("naïve" forecast) given by $S_t = S_{t-1} + \varepsilon_t$, where ε_t is white noise.
- (ii) An exponentially weighted moving average (EWMA) given by

$$S_{t+1} = \gamma_0 S_t + \gamma_1 S_{t-1} + \gamma_2 S_{t-2} + \dots, \quad (3.29)$$

where the weights are computed according to $\gamma_i = \alpha(1 - \alpha)^i$, $0 < \alpha < 1$ and the smoothing coefficient α is selected as the value minimizing the mean squared prediction error of one-step ahead forecasts.

- (iii) An ARMA(p, q) model for S_t , defined by the equation

$$(1 - \alpha(B))(S_t - c) = (1 - \beta(B))\varepsilon_t, \quad t \in \mathbb{Z}, \quad (3.30)$$

where α and β are lag polynomials as defined above and the errors ε_t are assumed to be normally distributed.

Moreover, we consider the ARFIMA(p, d, q) model put forward by Granger and Joyeaux [1980], Granger [1981] and Hosking [1981], given by

$$(1 - \alpha(B))(1 - B)^d(S_t - c) = (1 - \beta(B))\varepsilon_t, \quad t \in \mathbb{Z}, \quad c \in \mathbb{R}. \quad (3.31)$$

- (iv) The autoregressive conditional duration (ACD) model introduced by Engle and Russell [1998] and Engle [2000], which is the workhorse to capture serially dependent

3 A Forecasting Framework for Bid-Ask Spreads

positive-valued random variables, given by $S_t = \mu_t \cdot \varepsilon_t$ for $t \in \mathbb{Z}$ with conditional mean μ_t ,

$$\mu_t = \omega + \alpha(B)S_t + \beta(B)\mu_t, \quad \omega > 0. \quad (3.32)$$

The errors are assumed to be Weibull distributed, $\varepsilon_t | \mathcal{F}_{t-1} \sim \mathcal{W}(\mu_t, \gamma)$, with parameter γ .

Accordingly, the FIACD proposed by Jasiak [1998] is given by $S_t = \mu_t \cdot \varepsilon_t$ with $\varepsilon_t | \mathcal{F}_{t-1} \sim \mathcal{W}(\mu_t, \gamma)$ and

$$(1 - \phi(B))(1 - B)^d S_t = \omega + (1 - \beta(B))\nu_t, \quad \omega > 0, \quad (3.33)$$

where $\phi(B) := \alpha(B) + \beta(B)$ and $\nu_t := S_t - \mu_t$ is a martingale difference.

3.4.4 Point and Directional Forecast Evaluation

Let $\varepsilon_{t+1|t}^i := S_{t+1|t}^i - S_{t+1}^i$ be the forecast error of model $i \in \{1, 2\}$. To assess the basic forecast performance we report the root mean squared error (RMSE) of a series of forecast errors $\varepsilon_{t+1|t}^i$, $t = 1, \dots, T$. The predictive accuracy of competing forecast models, $i = 1, 2$, is tested using the test of Diebold and Mariano [1995] (DM), based on the loss differential

$$d_t := (\varepsilon_{t+1|t}^1)^2 - (\varepsilon_{t+1|t}^2)^2. \quad (3.34)$$

To test for differences in forecast performances, we test the null $H_0 : \mathbb{E}[d_t] = 0$. In the case of one-step ahead forecasts, the DM test statistic takes the form $DM := \bar{d}/\sqrt{\widehat{\text{Var}}(\bar{d})}$, where \bar{d} is the average of the d_t .

A minor modification of the DM test is necessary if nested models are compared. This is the case when we augment the models by additional predictors which inflate the RMSE due to additional estimation errors. Clark and West [2007] propose a test that explicitly accounts for the nested model structure. Let model 2 nest model 1 and let

$$\bar{a} := \frac{1}{T} \sum_{t=1}^T (S_{t+1|t}^1 - S_{t+1|t}^2)^2. \quad (3.35)$$

A suitable test statistic of the null is then given by

$$CW := \frac{\bar{d} - \bar{a}}{\sqrt{\widehat{\text{Var}}(\bar{d} - \bar{a})}}. \quad (3.36)$$

As the distribution of CW is non-standard, simulated critical values based on Clark and McCracken [2001] have to be used.

Note that Diebold-Mariano and Clark-West type tests can only compare two models. According to White [2000], a problem of such a sequential testing of competing models is that standard p-values may become invalid because of a possible spurious selection of the best model due to data snooping. Hansen [2005] propose a test for superior

predictive accuracy (SPA) that can account for this problem. Let $i = 1$ denote the benchmark model and let $d_t^i := (\varepsilon_{t+1|t}^1)^2 - (\varepsilon_{t+1|t}^i)^2$ be the loss differential to the rival model $i \in \{2, 3, \dots, m\}$. The null hypothesis of the SPA test is

$$H_0 : \mathbb{E}[d_t^i] \leq 0 \quad \forall i \in \{2, \dots, m\}. \quad (3.37)$$

Hence, H_0 is rejected whenever at least one of the competing models generates significantly better forecasts. The null can be tested based on the statistic

$$SPA := \max \left\{ \max_i \left\{ \widehat{\text{Var}}(\bar{d}^i)^{-1/2} \bar{d}^i \right\}, 0 \right\}, \quad (3.38)$$

where \bar{d}^i is the average of the d_t^i and $\widehat{\text{Var}}(\bar{d}^i)$ denotes the estimated variance of \bar{d}^i . The distribution of the SPA statistic has to be bootstrapped since the real distribution is nonstandard. Details of the stationary bootstrap procedure can be found in Hansen [2005].

Moreover, we evaluate direction forecasts, motivated by the fact that the direction of spread movements are equally important as point forecasts for trading decisions. The tests outlined above are straightforwardly applied based on the directional forecast errors $\varepsilon_{t+1|t} := D_{t+1|t} - D_{t+1}$, where $D_{t+1} := \mathbb{1}_{\{S_{t+1} > S_t\}} - \mathbb{1}_{\{S_{t+1} < S_t\}}$ is the realized direction of the spread movement. In addition to the SPA, DM and Clark-West tests for the squared directional error series, we report the directional accuracy of the forecasts which is given as

$$DA := \frac{\#\{t : D_{t+1} = D_{t+1|t}\}}{T}. \quad (3.39)$$

3.5 Results

3.5.1 Estimation Results and RMSE Performance

Model selection for all models is conducted by minimizing the RMSEs for one-step ahead forecasts of the January 2008 data. In this respect, we globally identify an ACP(1,1), ARMA(4,2), ACD(1,1) and Double ACP(1,1) as the best performing specifications across the stocks considered. To restrict the computational burden, the fractionally integrated and long memory specifications are restricted to $p = q = 1$. Diagnostics in terms of Probability Integral Transforms (PIT) based on probability mass function forecasts and autocorrelation functions are given in the appendix to Groß-Klußmann and Hautsch [2011b]. The diagnostics show that the LMACDP type II yields the best model fit in terms of capturing the spread distribution and dynamics.

Table 3.3 gives the median parameter estimates of the Double ACDP and LMACDP type II models over all estimates of the rolling window setup. Figure 3.6 shows the evolution of estimates for the fractional integration parameter d of the LMACDP type II model. We observe that the persistence in bid-ask spread clearly varies over time which makes it necessary to allow for parameter changes in a rolling window setup. The evolution of the estimates of the additional regressor coefficients is given in Figures

3 A Forecasting Framework for Bid-Ask Spreads

	GXP.N	XRAY.OQ	EMN.N	EQIX.OQ	
Median of Estimates for the LMACDP(1,d,1) model					
ω	1.414	0.175	1.802	0.327	5.298 1.084 11.098 1.562
ϕ_1	0.373	0.000	0.330	0.000	0.320 0.000 0.361 0.000
β_1	0.626	0.000	0.669	0.120	0.680 0.000 0.638 0.099
d	0.455	0.314	0.498	0.260	0.500 0.251 0.474 0.293
γ	1.512	1.565	1.484	1.541	1.023 1.043 0.341 0.359
Median of Estimates for the ACDP(1,1) model					
ω	0.120		0.141		0.304 1.257
ϕ_1	0.261		0.196		0.164 0.245
β_1	0.647		0.714		0.777 0.632
γ	1.541		1.528		1.041 0.353

Table 3.3: Median parameter estimates for the ACDP and LMACDP type I (on the left in each column) and II (on the right). The median is taken over the rolling window iterations. All variables significant at the 10% level in 95 % of the iterations. Notation is as in section 2

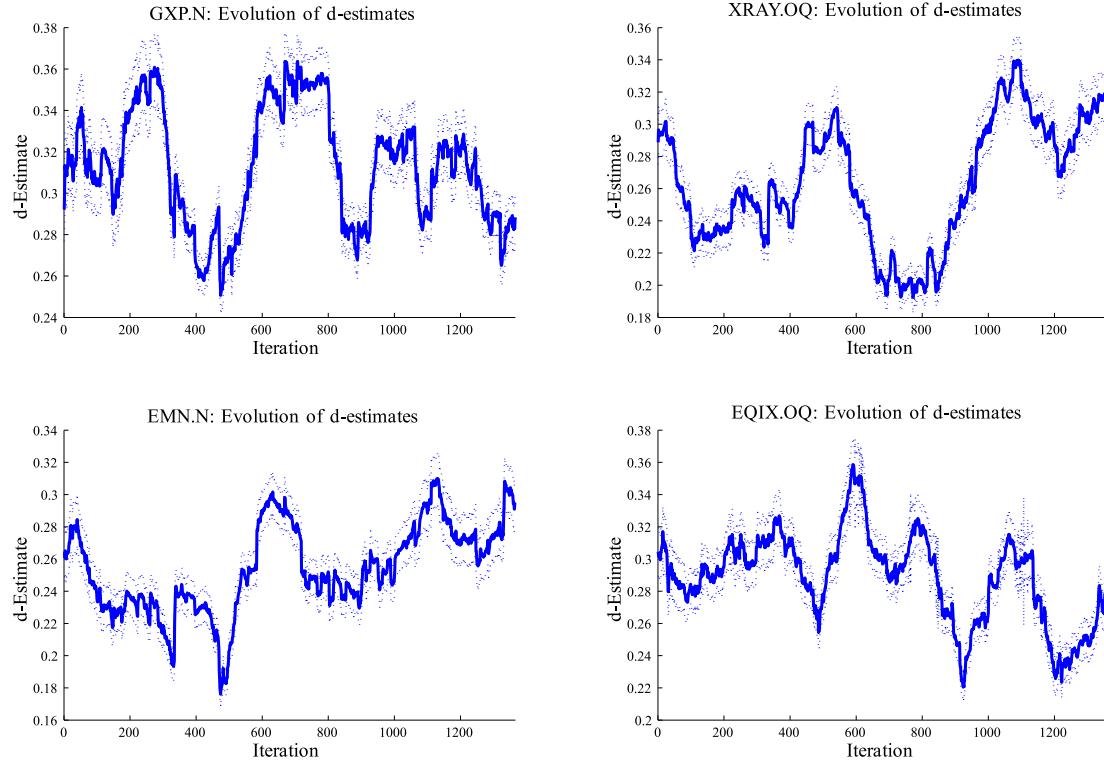


Figure 3.6: Evolution of the estimates of the fractional integration parameter in the LMACDP type II. 95 % confidence intervals dotted

3.5 Results

3.8 to 3.13 in the appendix. The signs of the parameter estimates are in line with economic theory. While volatility and trading volume are positively related to bid-ask spreads, the effects of trade and depth imbalances fluctuate around zero. Moreover, depth coefficients are widely negative, reflecting that deep markets are accompanied by low spreads reflecting periods of high liquidity.

The upper panel of Table 3.4 gives the one-step ahead RMSEs and the directional accuracy for all models without additional predictors. We observe that a Poisson-based model always performs best with RMSEs on average 18 % lower than those of the random walk benchmark. Likewise, the directional accuracy of Poisson-based forecasts improves on average by 25 % over the naïve benchmark. Overall, the LMACDP type I and II models outperform FIACD and ARFIMA benchmarks in terms of the RMSE and DA, with differences in the forecast performance being especially high for the directional forecasts. These results indicate that an appropriate distribution as implied by the Double Poisson distribution yields efficiency gains which lead to superior predictions. Moreover, in most cases we find forecasts of LMACDP type II specifications to be (marginally) superior compared to type I specifications.

The lower panel of Table 3.4 shows the RMSE and DA for the LMACDP type II model augmented by covariates. It turns out that the inclusion of predictors improves both point and direction forecasts. However, including all predictors ("All Predictors") or adaptive selections thereof ("Preselected Pred.") do not necessarily provide better forecasts than including only single predictors. Hence, potential forecast gains by the inclusion of several variables are obviously offset by a higher estimation uncertainty.

Table 3.5 shows Diebold-Mariano and Clark-West tests for forecast comparisons of the best and second best specifications in terms of the RMSE. We observe that point and directional forecasts of models without predictors are significantly different from each other. Hence, the superior performance of LMACDP models in terms of the RMSE and DA shown in Table 3.4 is statistically significant. The Clark-West tests indicate that the inclusion of trading characteristics in LMACDP type II models yields a significant improvement of the forecasting power over the basic LMACDP type II. However, among specifications including covariates, differences between squared prediction errors are often insignificant according to the Diebold-Mariano Test. This reflects that most of the covariates carry similar information about the adverse selection costs driving spreads.

To identify the overall best performing model in terms of point and direction forecasts we present the results of the SPA test in Table 3.6. In three out of the four cases we cannot reject the null at the 5 % level that a LMACDP type II model including covariates provides the best point forecast performance. The p-values are widely in accordance with the ordering of models according to the RMSE and DA results tabulated in Table 3.4.

RMSE DA	GXP.N	XRAY.OQ	EMN.N	EQIX.OQ
RMSE of Basic Models Directional Accuracy of Basic Models				
Naïve	1.2498 0.4756	1.3481 0.3740	2.9459 0.1950	7.1555 0.1444
EWMA	1.2867 0.5043	1.2381 0.5215	2.5649 0.5432	6.7206 0.5437
ARMA	1.1667 0.5321	1.2122 0.5131	2.5500 0.5231	6.3682 0.5287
ARFIMA	1.1056 0.3095	1.1281 0.3717	2.3906 0.4965	5.9638 0.5150
ACD	1.0880 0.3192	1.1217 0.3752	2.3692 0.4984	5.8634 0.5209
FIACD	1.0826 0.3181	1.1101 0.3731	2.3478 0.4975	5.8195 0.5246
ACP	1.0803 0.5516	1.1152 0.5369	2.3436 0.5430	5.8466 0.5400
ACDP	1.0803 0.5516	1.1153 0.5369	2.3436 0.5433	5.8466 0.5400
LMACDP type I	1.0805 0.5625	1.1153 0.5322	2.3440 0.5184	5.9553 0.5503
LMACDP type II	1.0720 0.5552	1.1086 0.5389	2.3448 0.5394	5.8026 0.5452
RMSE Directional Accuracy of LMACDP type II model plus additional predictors				
+All Predictors	1.0577 0.5534	1.1113 0.5411	2.3177 0.5458	5.7996 0.5489
+Preselected Pred.	1.0637 0.5549	1.1178 0.5380	2.3380 0.5416	5.8192 0.5484
+Seasonality	1.0678 0.5519	1.1022 0.5383	2.3281 0.5426	5.7924 0.5461
+Depth	1.0710 0.5550	1.1026 0.5427	2.3391 0.5420	5.8007 0.5443
+Depth Imb.	1.0727 0.5553	1.1086 0.5383	2.3431 0.5392	5.8034 0.5437
+Real. Vola	1.0800 0.5544	1.1203 0.5392	2.3369 0.5428	5.8038 0.5470
+Absolute Ret.	1.0637 0.5546	1.1181 0.5390	2.3384 0.5407	5.8071 0.5483
+Traded Vol.	1.0722 0.5556	1.1098 0.5388	2.3405 0.5408	5.8036 0.5460
+T.Vol. Imb.	1.0725 0.5556	1.1090 0.5389	2.3453 0.5394	5.8038 0.5445

Table 3.4: Upper panel: RMSE and DA values of models without additional predictors. Best model in terms of RMSE and DA highlighted. Lower panel: Models with additional predictors as defined in section 4.2. All predictors denotes inclusion of all variables. Preselected pred. refers to the preselection scheme outlined in section 4.2. Abbreviations are Imb. for imbalance, Real. Vola. for realized volatility, Ref. for return, Vol. for Volume and T.Vol. Imb. for Trading Volume Imbalance. Seasonality denotes the seasonality component (3.28) with $Q = 2$. Lowest RMSE and highest DA highlighted

DM and CW statistics	GXP.N	XRAY.OQ	EMN.N	EQIX.OQ
<hr/>				
DM test for equal forecast performance of Basic Model with lowest RMSE highest DA				
and naïve model	-19.23 -7.74 (0.00) (0.00)	-31.36 -15.12 (0.00) (0.00)	-37.71 2.58 (0.00) (0.00)	-31.69 14.66 (0.00) (0.00)
P-Value				
and 2nd best Basic model	-3.57 -3.64 (0.00) (0.00)	-3.26 -6.62 (0.00) (0.00)	0.34 5.87 (0.36) (0.00)	-7.55 -12.55 (0.00) (0.00)
P-Value				
<hr/>				
DM test for equal forecast performance of Model with additional Predictors with lowest RMSE highest DA				
and 2nd best with add. Predic.	-2.57 0.60 (0.01) (0.27)	-0.18 -2.72 (0.43) (0.00)	-2.94 1.28 (0.00) (0.10)	-1.57 0.44 (0.06) (0.33)
P-Value				
<hr/>				
CW test for equal forecast performance of Model with additional Predictors with lowest RMSE highest DA				
and the LMACDP type II	2.66* 0.23	2.24* 113.57*	4.81* -27.61*	1.93* 6.92*

Table 3.5: Diebold Mariano and Clark-West test results for point and direction forecasts. The * in the last row denotes significance at the 10% level

	% P-Values of SPA-Test	GXP.N	XRAY.OQ	EMN.N	EQIX.OQ
SPA test based on squared error series for point forecasts direction forecasts					
Naïve	0.00 0.00	0.00 0.00	0.00 5.25	0.00 0.00	0.00 0.00
EWMA	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00
ARMA	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00
ARFIMA	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00
ACD	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00
FIACD	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00
ACDP	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00	0.00 0.00
LMACDP type I	0.00 0.75	0.00 31.25	0.00 4.50	0.00 0.00	0.00 0.00
LMACDP type II	4.00 0.00	0.00 18.75	0.00 2.50	9.00 0.00	9.00 0.00
<i>LMACDP type II</i>					
+ All Pred.	0.75 0.00	14.25 1.75	0.75 2.25	25.25 0.00	
+ preselected Pred.	1.50 0.00	1.50 15.75	0.00 5.25	3.50 0.00	
+ Seasonality	4.00 0.00	49.00 1.00	0.00 0.00	23.25 0.00	
+ Depth	7.00 1.25	47.25 27.50	0.00 4.00	8.25 0.00	
+ Depth Imb.	5.25 0.25	0.25 13.00	0.00 0.50	7.25 0.00	
+ Real Vola.	1.50 0.00	1.25 11.75	0.00 3.00	13.00 0.00	
+ Abs. Return	1.00 0.00	7.25 2.50	0.00 5.50	21.00 0.00	
+ Traded Vol.	4.00 0.00	0.00 14.75	0.00 1.00	8.75 0.00	
+ T.Vol. Imb.	5.75 0.75	0.00 18.25	0.00 0.25	10.50 0.00	

Table 3.6: P-Values of the SPA test for all competing models. Highest value highlighted. Notation for the additional covariates as in Table 3.4

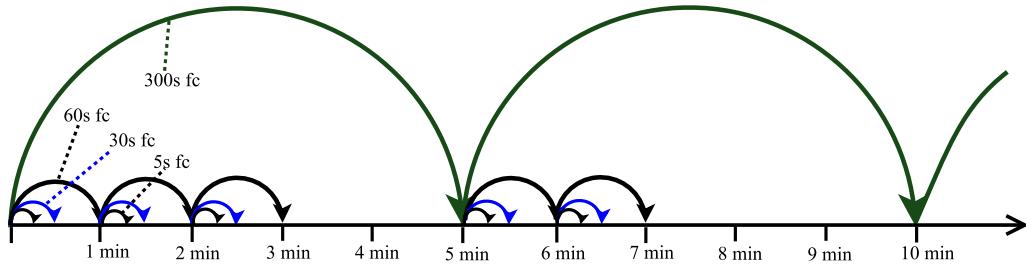


Figure 3.7: The principle of the trading schedule. 5s fc, 30s fc, 60s fc and 300s fc stand for the forecasts based on different data aggregation frequencies

We summarize the following main findings: First, the efficiency gains implied by (Double) Poisson modelling yield significantly superior forecast results in terms of the RMSE and DA criterion. Second, the forecast performance of the long memory specifications compared to their short-memory counterparts indicate the importance of accounting for the strong persistence in spreads. Third, the inclusion of predictors significantly improves point- and direction forecasts.

3.5.2 A Trading Schedule based on Spread Forecasts

To evaluate the potential economic gains implied by spread forecasts, we consider quantifying spread costs in trading schemes. The trading schedules are motivated by the fact that transaction costs of large trades can be reduced by splitting orders into smaller trades that are distributed over time. In such strategies, spread forecasts can improve trading algorithms by allowing to intensify trading in periods when spreads are expected to be small.

Suppose a benchmark trading schedule is based on trades occurring at the end of each 5 minute interval. In an alternative trading schedule the time of the trades is flexible within each interval and can be chosen in accordance with corresponding spread forecasts. Then, the resulting transaction costs serve as a measure of the implied economic gains.

To obtain a fine grid of spread forecasts within each 5 minute interval, we construct bid-ask spread forecasts on a 5, 30, 60 and 300 second frequency employing the LMACDP type II specification with seasonal component. Let fc_5^x , fc_{30}^x , fc_{60}^x and fc_{300} denote spread forecasts on a 5, 30, 60 and 300 second data aggregation frequency, respectively. Moreover, let the superscript $x \in \{1, 2, 3, 4, 5\}$ indicate the corresponding 1 minute subinterval within the 5 minute interval. Then, the timing of trades in the flexible schedule is chosen as follows.

- Starting from the left interval boundary, we search in the x successive subintervals for the smallest forecast out of $\{fc_5^x, fc_{30}^x, fc_{60}^x, fc_{300}\}$ until we arrive at the right

3 A Forecasting Framework for Bid-Ask Spreads

Schedule	GXP.N	XRAY.OQ	EMN.N	EQIX.OQ
Fast	7.74	7.47	8.17	11.60
PMF Info	11.85	10.51	10.58	12.90
Slow	14.42	8.22	10.51	11.90

Table 3.7: Spread cost savings as percentage of the benchmark schedule when employing forecasts in the schedules. Schedule names refer to the three rules in case of equal forecasts

boundary or a minimum is found. Once a minimum is found for $x \in \{1, 2, 3, 4, 5\}$, we do not consider subintervals $x^* > x$ since these are not known by a trader using one-step ahead forecasts on 5s, 30s, 60s and 300s frequencies. In case fc_{30}^x or fc_{60}^x is a new optimum, we optimize the time of trading within the corresponding subinterval using the forecasts on higher frequencies. Figure 3.7 illustrates the procedure.

- In case of equal forecasts, $\min\{fc_5^x, fc_{30}^x, fc_{60}^x\} = fc^{300}$, we choose the timing of trades according to one of the following three options.
 - (i) Motivated by traders' tendency to trade as fast as possible, we choose the time of trading to be closest to the left boundary of the interval and stop searching for further minima. The resulting schedule is labelled "fast".
 - (ii) We choose the time of trading to be closest to the right boundary and successively optimize the time of trading using forecasts on higher frequencies until the time of the new minimum forecast. The resulting schedule is labelled "slow".
 - (iii) We use the information from the probability mass function and weight the equal forecasts with the assigned forecasted probabilities $f(S_{t+1|t}, \hat{\lambda}_{t+1}, \hat{\gamma})$, where f denotes the Double Poisson probability mass function. Then, we choose the most probable forecast to be the new optimal one and optimize the trading in the subintervals until we arrive at another optimum. This schedule is labelled "PMF info".

After the timing of trades is chosen we sum up the incurred transaction costs (induced by crossings of the market) for the benchmark strategy and the three alternative schedules. Table 3.7 reports the percentage spread cost savings over the benchmark strategy. We find that the strategy exploiting the information from the pmfs of the Double Poisson assumption yields the highest average gains: spread costs reduce by 11.46% of the costs of the benchmark schedule. The "slow" schedule saves us 11.26% and the "fast" strategy still 8.74%. The better performance of the strategies "slow" and "PMF info" is obviously induced by the use of more optimization steps due to the rule specification in case of equal forecasts.

3.5.3 Robustness of the Results

Several robustness checks underscore the relevance of our study. First, we can confirm the reported RMSE results for a larger cross-section of stocks from the mid-cap segment of the Russell 3000. The appendix to Groß-Klußmann and Hautsch [2011b] shows descriptive statistics, point forecast and trading schedule evaluations for 46 stocks ordered according to their average spread. We find that the RMSE improvement of the LMACDP type II over the naïve benchmark model is 19.51% on average across stocks with average spread ticks $\in (2, 4)$ and 18.90% for stocks with mean spread ticks greater than 4. Interestingly, forecast gains are also possible in case of nearly constant stocks: RMSE improvements over the naïve model are on average 13.55% for stocks with mean spread ticks < 2 .

Second, the spread cost savings from the trading schedules tend to increase with the size of average spreads. Moreover, the specification of a flexible probability mass function in the Poisson models becomes more and more important with increasing spread sizes. While the "PMF info" trading schedule does not generate higher percentage cost savings than the alternative two schedules for stocks with spreads smaller than 4 ticks, the opposite is true when average spreads become larger than 4 ticks. We conclude that the flexible Double Poisson modelling is the more useful the larger the spreads and the more dispersed the spread series.

Third, the results are not only relevant for quoted spreads but also for alternative spread measures like, e.g., effective spreads. Effective spreads are closely related to quoted spreads (see Figure 3.1) and reveal very similar time series properties.

3.6 Conclusions

Motivated by the relevance of bid-ask spreads in trading decisions and market microstructure modelling this study is the first one systematically analyzing forecasts of quoted bid-ask spreads. To capture the empirical features of spread time series for Russell 3000 mid cap stocks traded at NYSE and NASDAQ we introduce a novel long memory autoregressive conditional Poisson (LMACP) model. The LMACP can accommodate highly persistent time series of count data and is thus suitable for modelling persistent discrete time series which are often found in high-frequency data applications.

We find that autoregressive conditional Poisson (ACP) models and their long memory extension well capture the distributional and dynamic properties of quoted bid-ask spreads. Generalizations of the Poisson distribution, such as the Double Poisson distribution, are able to account for both under- and overdispersion found in the data and underscore the good fit of the proposed model. Forecasting bid-ask spreads in a rolling window out-of-sample framework shows that long memory ACP models outperform competing benchmarks like ARFIMA, ARMA, ACD, FIACD and exponential moving average models in terms of the root mean squared error, directional accuracy and density forecasts. Implementing the spread forecasts in a simple trading algorithm we find that spread forecast can induce transaction cost savings of up to 12%.

B Appendix

B.1 Technical appendix

Proof of proposition 2:

Proposition 2. *The unconditional variance of the long memory autoregressive conditional Double Poisson model type I is given by*

$$\text{Var}[S_t] = \frac{1}{\gamma} \mathbb{E}[\lambda_t] \sum_{j=0}^{\infty} \omega_j^2 < \infty.$$

Proof. We obtain an expression for the unconditional variance of the errors ν_t of the Double Poisson specification from the following steps. We have

$$\begin{aligned} \mathbb{E}[\nu_t^2] &= \mathbb{E}[(S_t - \lambda_t)^2] = \mathbb{E}[S_t^2] - 2\mathbb{E}[\mathbb{E}[S_t \lambda_t | \mathcal{F}_{t-1}]] + \mathbb{E}[\lambda_t^2] \\ &= \mathbb{E}[S_t^2] - \mathbb{E}[\lambda_t^2], \end{aligned} \quad (3.40)$$

since λ_t depends only on past values of λ_t and S_t , and

$$\begin{aligned} \text{Var}[S_t] &= \mathbb{E}[S_t^2] - \mathbb{E}[S_t]^2 \\ &= \mathbb{E}\underbrace{[\text{Var}[S_t | \mathcal{F}_{t-1}]]}_{=\frac{\lambda_t}{\gamma}} + \text{Var}\underbrace{[\mathbb{E}[S_t | \mathcal{F}_{t-1}]]}_{=\lambda_t} = \frac{1}{\gamma} \mathbb{E}[\lambda_t] + \mathbb{E}[\lambda_t^2] - \mathbb{E}[\lambda_t]^2. \end{aligned} \quad (3.41)$$

Solving (3.41) for $\mathbb{E}[S_t^2]$ and substituting into (3.40) we get

$$\begin{aligned} \mathbb{E}[\nu_t^2] &= \mathbb{E}[S_t]^2 + \frac{1}{\gamma} \mathbb{E}[\lambda_t] - \mathbb{E}[\lambda_t]^2 + \mathbb{E}[\lambda_t^2] - \mathbb{E}[\lambda_t^2] \\ &= \mathbb{E}\underbrace{[\mathbb{E}[S_t | \mathcal{F}_{t-1}]]^2}_{=\lambda_t} + \frac{1}{\gamma} \mathbb{E}[\lambda_t] - \mathbb{E}[\lambda_t]^2 + \mathbb{E}[\lambda_t^2] - \mathbb{E}[\lambda_t^2] = \frac{1}{\gamma} \mathbb{E}[\lambda_t]. \end{aligned}$$

From the infinite moving average representation

$$S_t = \omega + \Omega(B)\nu_t \quad (3.42)$$

we obtain using $\text{Cov}(\nu_t, \nu_{t-1}) = 0$

$$\text{Var}[S_t] = \sum_{j=0}^{\infty} \omega_j^2 \text{Var}[\nu_t] = \frac{1}{\gamma} \mathbb{E}[\lambda_t] \sum_{j=0}^{\infty} \omega_j^2.$$

Furthermore, applying Stirling's formula we obtain $\omega_j \approx C j^{d-1}$ for high j (see Hosking [1981]), where C is a positive constant, such that $\lim_{k \rightarrow \infty} \sum_{i=0}^k \omega_i^2$ converges for $0 < d < 0.5$. \square

The following refers to the footnote on page 15.

Footnote 5. *The sum of the squared coefficients of $\Omega(B)$ is greater or equal than 1,*

$$\sum_{j=0}^{\infty} \omega_j^2 \geq 1.$$

Proof. We have

$$(1 - B)^{-d} = 1 + \delta_1 B + \delta_2 B^2 + \dots$$

(see Hosking [1981]) and

$$(1 - \alpha(B))^{-1} = 1 + a_1 B + a_2 B^2 + \dots$$

(see, e.g., Hamilton [1994]) such that the first coefficient of $\Omega(B)$ is $\omega_0 = 1$,

$$\begin{aligned} \Omega(B) &= \omega_0 + \omega_1 B + \omega_2 B^2 + \dots = (1 - \beta(B))(1 - \alpha(B))^{-1}(1 - B)^{-d} \\ &= 1 + \omega_1 B + \omega_2 B^2 + \dots \end{aligned}$$

and hence

$$1 \leq 1 + \sum_{j=1}^{\infty} \omega_j^2 = \sum_{j=0}^{\infty} \omega_j^2.$$

□

B.2 Evolution of Coefficient Estimates

3 A Forecasting Framework for Bid-Ask Spreads

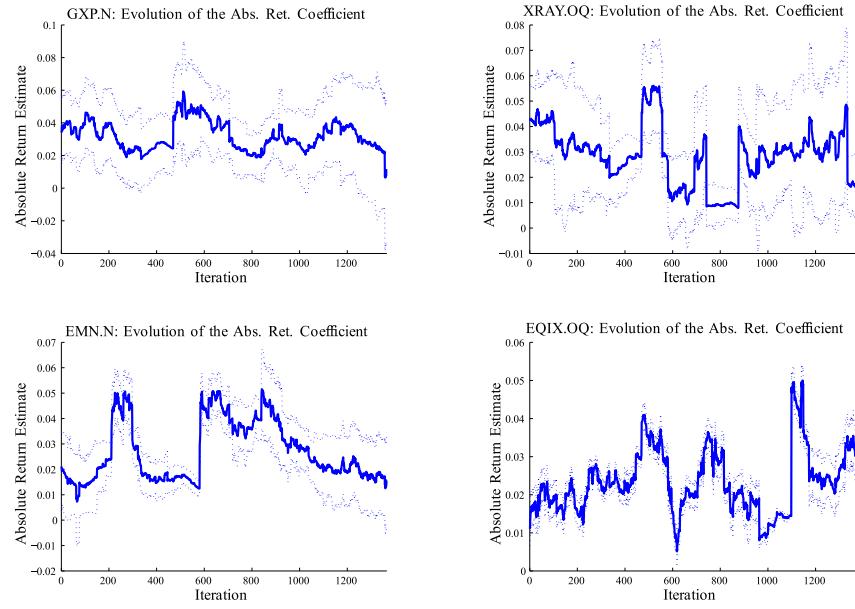


Figure 3.8: Evolution of the estimates of the absolute return coefficient. 95 % confidence dotted

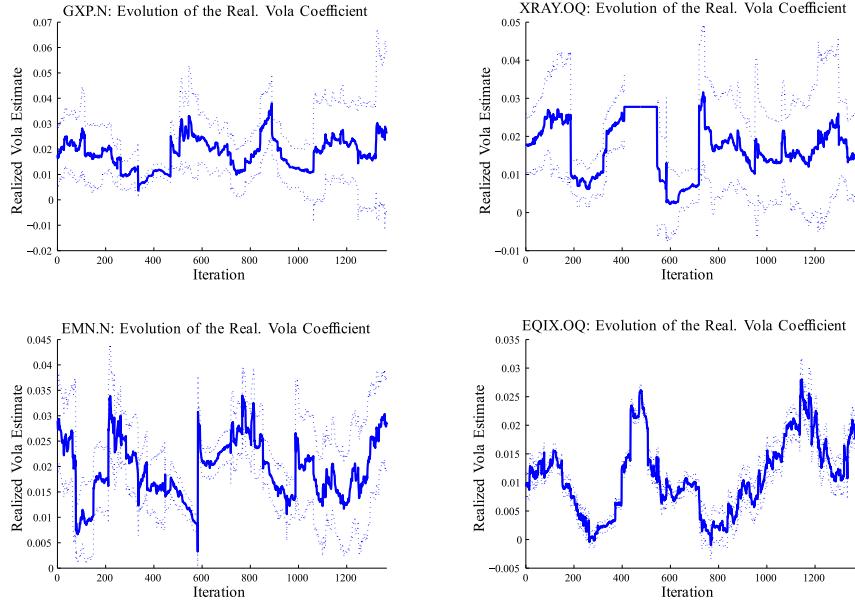


Figure 3.9: Evolution of the estimates of the realized volatility coefficient. 95 % confidence dotted

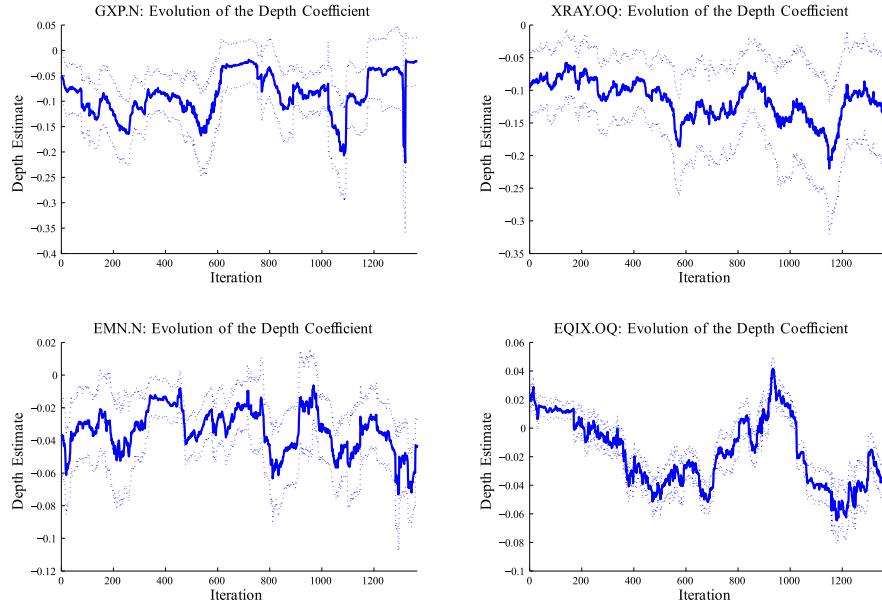


Figure 3.10: Evolution of the estimates of the depth coefficient. 95 % confidence dotted

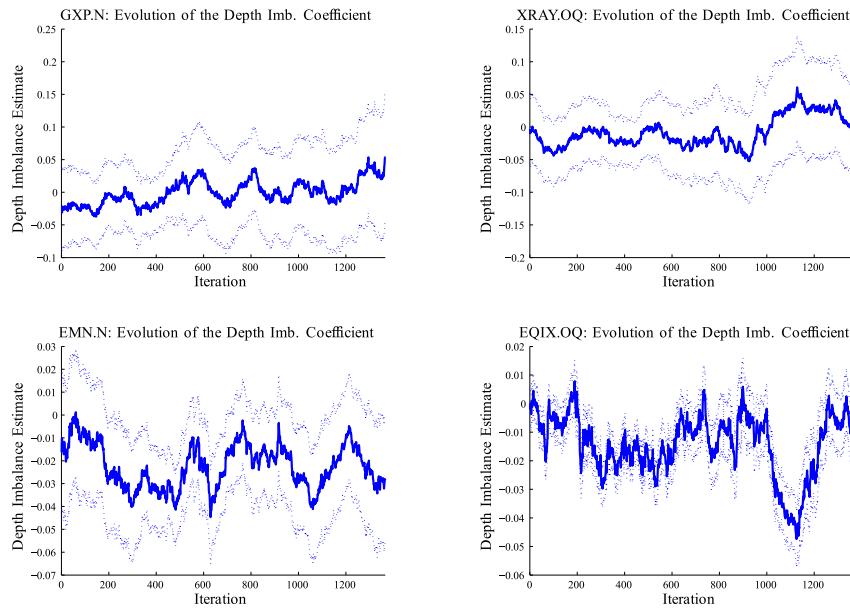


Figure 3.11: Evolution of the estimates of the depth imbalance coefficient. 95 % confidence dotted

3 A Forecasting Framework for Bid-Ask Spreads

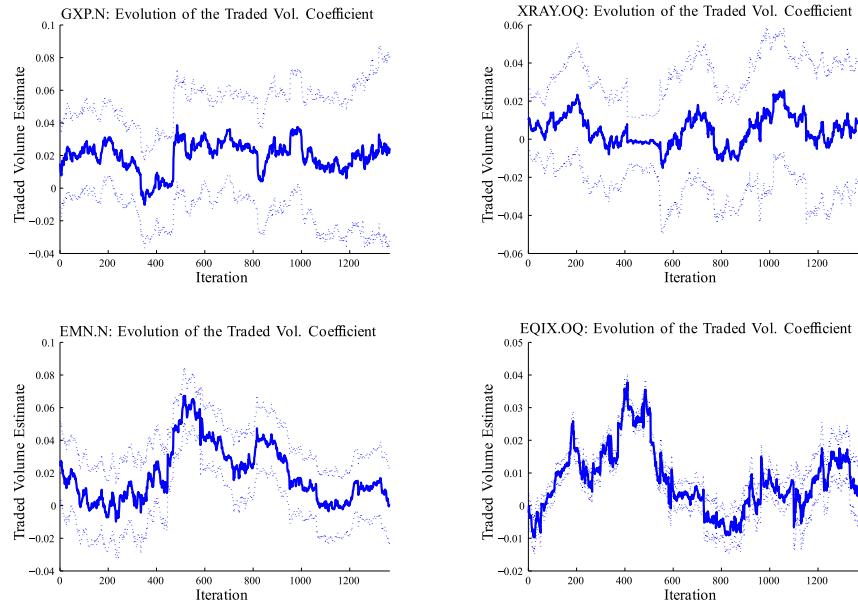


Figure 3.12: Evolution of the estimates of the traded volume coefficient. 95 % confidence dotted

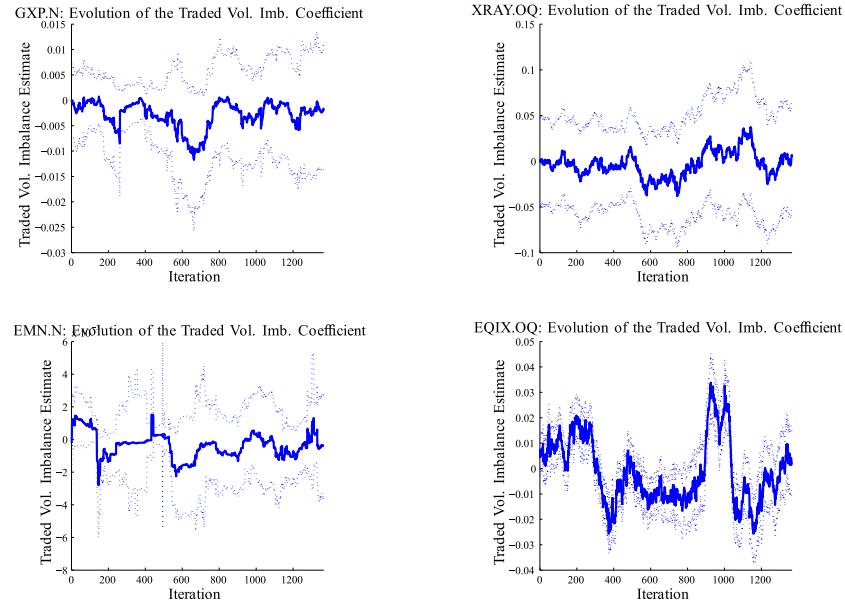


Figure 3.13: Evolution of the estimates of the traded volume imbalance coefficient. 95 % confidence dotted

4 Modelling Volatility Measures on Intra-Daily Time Intervals

4.1 Introduction

Estimating and modelling the latent volatility process of returns is important for financial market participants in a wide range of contexts such as option pricing as well as risk management, and has consequently inspired a very active area of research. However, in contrast to the extant literature on daily volatility and spot (or instantaneous) volatility modelling, little empirical research is devoted to measures of intraday volatility on fixed, often small time intervals, which are important for high frequency trading applications and have become increasingly important for intraday risk assessment. Hasbrouck and Saar [2002], Ranaldo [2004] and Hall and Hautsch [2006] for instance find that intraday volatility is a main determinant of limit order submission strategies in high frequency trading. Most importantly, Lo et al. [2002] show that limit orders are more likely to be filled when volatility is high. Easley and O'Hara [1992] and Glosten and Milgrom [1985] are examples of theoretical studies linking high frequency volatility to liquidity and trading activity. In a different branch of literature, Giot [2005] and Dionne et al. [2009] propose to use intraday volatility measures to derive an intraday value at risk.

Despite the importance for designing and evaluating trading strategies as well as for intraday risk measurement, the question of how to measure volatility on intra-daily intervals in an optimal way remains an unsolved issue. While there is a consensus that simple daily volatility measures like squared daily returns can be improved by using intraday transaction data, it is *a priori* unclear whether a similar finding also holds for intraday time intervals. Consequently, our paper is the first to systematically address the problem of measuring volatility on intra-daily intervals and comparing alternative approaches. In addition, we propose a multiple component realized Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model for intraday returns to exploit the information contained in the alternative volatility measures and benchmark it against the standard GARCH.

The traditional approaches to intraday volatility modelling typically use the squared returns based on the quotes at interval boundaries as measures of the volatility for intra-daily intervals. In order to utilize the additional within-interval information we employ realized volatility measures as well as measures based on spot volatility estimates.

The underlying theory of both realized volatility and spot volatility estimation requires the log price process X_t to be a continuous semimartingale, given for instance by an Itô

4 Modelling Volatility Measures on Intra-Daily Time Intervals

process,

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s, \quad (4.1)$$

where t denotes continuous calendar time and X_0 is the starting value. W_t is a standard Brownian motion adapted to a filtration \mathcal{F}_t , and μ_t as well as σ_t represent càglàg drift and càdlàg diffusion processes adapted to \mathcal{F}_t . A measure of the asset price volatility over $[0, t]$ in this case is given by estimates of the integrated volatility $\int_0^t \sigma_s^2 ds$. We compute estimates of the integrated volatility on intra-daily intervals along the lines of the realized volatility measures of e.g. Zhang et al. [2005], Barndorff-Nielsen et al. [2008a] and Jacod et al. [2009]. Moreover, we consider estimators of the (instantaneous) spot volatility σ_s^2 which can be numerically integrated to obtain estimators of the integrated volatility. This allows us to draw from a completely different strand of literature on both nonparametric (see e.g. Kristensen [2010], Boswijk and Zu [2010]) and parametric (see Engle and Russell [1998]) approaches to spot volatility estimation. The idea is that spot volatility estimation techniques are helpful for very small intraday intervals since the estimates are based on observations outside of the interval of interest in addition to the within-interval information. In contrast, the realized-type estimators of the integrated volatility are restricted to the data from the interval and thus potentially more prone to the finite sample problem encountered for small intervals.

As the finite sample properties of the proposed estimators are widely unclear we conduct a simulation study to compare the competing approaches to volatility measurement on intraday intervals. In order to assess the practical relevance of the estimators we conduct an analysis based on the time series of competing intraday volatility measures for the price data of four liquid NYSE and NASDAQ traded stocks. Accounting for the diurnal seasonality and inter-daily patterns we estimate a multiple component realized GARCH model which draws on the works of Engle et al. [2006] and Hansen et al. [2011] and analyze the tail probabilities of returns implied by the alternative volatility measure series.

The following main results emerge from our analysis. First, the simulation study confirms the theoretical usefulness of spot volatility estimators for the construction of integrated volatility estimators over small time intervals. Second, the proposed multiple component realized GARCH model captures salient features of the different volatility time series and the intraday returns. Third, the application of the outlined framework to real data shows that the tail probabilities are often better captured when the multiple component realized GARCH model includes standard realized volatility or spot-volatility-based measures instead of squared returns.

Our study contributes to the literature on intraday volatility modelling put forward by Andersen and Bollerslev [1997], Andersen and Bollerslev [1998], Tse and Yang [2010] and Engle et al. [2006] as well as to the literature on intra-daily risk management such as Giot [2005] and Dionne et al. [2009].¹ The study is related to the recent literature on high frequency data modelling for algorithmic trading applications in the spirit of

¹See in this context also Manganelli [2005] and Hautsch [2008] for multivariate approaches to high frequency volatility modelling.

Brownlees et al. [2011], Groß-Klußmann and Hautsch [2011b], Härdle et al. [2009] and Hautsch [2012].

The paper is organized as follows. Section 4.2 describes the competing approaches to the estimation of volatility on intraday intervals based on realized measures and spot volatility measurement. In section 4.3 we compare the performance of the competing estimators in a Monte-Carlo simulation study. In section 4.4 we describe the multiple component realized GARCH and compare the different volatility measures in terms of the tail probabilities of returns implied by the econometric model. Section 4.5 concludes.

4.2 Intraday volatility measures

4.2.1 Realized measures of volatility

A popular model under a no-arbitrage argument (see Dalbaen and Schachermeyer [1994]) requires log asset prices X_t to follow an Itô process defined by

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s, \quad 0 \leq t \leq t_n, \quad (4.2)$$

with notation as defined before, see (4.1). For the processes μ_t as well as σ_t we further require $\int_0^{t_n} |\mu_t| dt < \infty$ and $\int_0^{t_n} \sigma_t^2 dt < \infty$. In this setup the accumulated log price volatility from 0 to t is given as the integrated variance $\int_0^t \sigma_s^2 ds$. The integrated variance reflects the riskiness or volatility of the asset over a given time span and is thus of central interest to market participants.

Suppose X_t is observed on a grid of intraday time points,

$$\mathcal{G} := \{t_0 = 0, t_1, \dots, t_n\}, \quad (4.3)$$

which can be considered to span an intraday interval, for instance. To simplify the following notation we assume equidistant observations with distance $\Delta := t_{i+1} - t_i$. The quadratic variation of X_t relative to grid \mathcal{G} is defined by

$$[X, X]_t^{\mathcal{G}} = \sum_{t_{i+1} \leq t} (X_{t_{i+1}} - X_{t_i})^2. \quad (4.4)$$

In the above model one can show that $[X, X]_t^{\mathcal{G}} \xrightarrow{p} [X, X]_t$, where the limiting process $[X, X]_t$ is typically referred to as the quadratic variation of the process X_t . Specifically, the limit is given by

$$[X, X]_t = \int_0^t \sigma_s^2 ds \quad (4.5)$$

in the outlined log price model without jumps (see e.g. Andersen et al. [2003] and Mykland and Zhang [2010]). Hence, the integrated variance in $[0, t]$ can be consistently estimated by $RV_t := [X, X]_t^{\mathcal{G}} = \sum_{i: t_i \leq t} r_i^2$, where $r_i = X_{t_{i+1}} - X_{t_i}$ denotes the log-return for observation i . In the financial context the estimator RV_t is known under the name

realized volatility. The realized volatility can be shown to be asymptotically normal,

$$\sqrt{n} \left(RV_t - \int_0^t \sigma_s^2 ds \right) \xrightarrow{\mathcal{L}} U \times \left(2t_n \int_0^t \sigma_s^4 ds \right)^{1/2}, \quad (4.6)$$

where U is a standard normal random variable and \mathcal{L} denotes stable convergence. The term $t_n \int_0^t \sigma_s^4 ds$ is called the integrated quarticity which can be consistently estimated by the realized quarticity $RQ_t := 1/3 \sum_{i:t_i \leq t} r_i^4$.²

The true (or efficient) price process X_t observed on a high intraday frequency is typically thought of to be contaminated by so-called market microstructure noise stemming mainly from the discreteness of price quotes. Hence, we can only observe prices Y_{t_i} given by

$$Y_{t_i} = X_{t_i} + \varepsilon_{t_i}, \quad t_i \in \mathcal{G}, \quad (4.7)$$

where ε_{t_i} is a noise process. The properties of estimators of the integrated volatility depend on the exact assumptions for the noise. A standard assumption is given as ε_{t_i} i.i.d. and independent of X_{t_i} with $\mathbb{E}[\varepsilon_{t_i}] = 0$ and $\text{Var}(\varepsilon_{t_i}) = \omega^2$. However, these assumptions can be considered strict. Hansen [2005], for instance, give empirical evidence of autocorrelation in the noise.

In the following we shortly review some important approaches of realized volatility estimation in the presence of microstructure noise. Without loss of generality, let $[0, t_n]$ represent the interval of interest. To mimic the typical information set of market participants at t_n we do not consider observations after the right boundary of the respective interval of interest throughout the analysis. Hence, Y_{t_n} is the most recent observation used in estimating the integrated volatility over $[0, t_n]$.

(i) Sparse sampling

One early approach to deal with the microstructure noise was to employ sparse sampling. Suppose a subgrid of observations $\mathcal{H} := \{t_0, t_{0+K}, t_{0+2K}, \dots\} \subset \mathcal{G}$, $K \in \mathbb{N}$. The sparsely sampled realized volatility estimator is given by $RV_{t_n}^K := [Y, Y]_{t_n}^{\mathcal{H}}$. The idea is to limit the effect of the noise when there are fewer sampling points for large values of K while still preserving the convergence of the estimator to $[X, X]_{t_n}$. Motivated by the GARCH literature we also include the squared return based on the two prices at the left and right interval boundary as an alternative sparse sampling estimator in the study.

Noise-robust realized volatility estimators allow to derive estimates of the market microstructure noise variance ω^2 which will be used later to assess the degree of price contamination through noise. Hansen [2005] propose

$$\widehat{\omega}_1^2 = \frac{RV_{t_n} - \widehat{IV}}{2n}, \quad (4.8)$$

where \widehat{IV} denotes a consistent estimator of the integrated volatility. It can be shown that $\widehat{\omega}_1^2 \xrightarrow{p} \omega^2$, see Hansen [2005].

To account better for the empirically observed autocorrelation in the microstructure

²See Mykland and Zhang [2010] for details. Note that $[X, X]_t$ contains the cumulated jumps in case of a log price models with jumps.

4.2 Intraday volatility measures

noise, Barndorff-Nielsen et al. [2008b] propose an autocorrelation-robust estimator based on

$$\widehat{\omega}_2^2 = \frac{1}{q} \sum_{i=1}^q \widehat{\omega}_{(i)}^2, \quad (4.9)$$

with

$$\widehat{\omega}_{(i)}^2 = \frac{RV_{t_n}^i}{2n_i}, \quad i = 1, \dots, q, \quad (4.10)$$

where n_i denotes the number of zero returns used to compute the sparsely sampled $RV_{t_n}^i$, $i = 1, \dots, q$.

(ii) Two Scales Realized Volatility

To exploit the observations left out by the sparse sampling scheme, Zhang et al. [2005] propose the Two Scales Realized Volatility (TSRV) estimator based on averages of sparsely sampled price data. For the TSRV we assume an i.i.d. noise process independent of the efficient log price process. Suppose subgrids $\mathcal{G}_k := \{t_k, t_{K+k}, t_{2K+k}, \dots\}$ of \mathcal{G} are given for $k = 1, \dots, K$. Averaging realized volatility estimators across all subgrids gives the average realized volatility

$$ARV_{t_n}(Y, \mathcal{G}, K) = \frac{1}{K} \sum_{k=1}^K [Y, Y]_{t_n}^{\mathcal{G}_k}. \quad (4.11)$$

However, the average realized volatility estimator is still inconsistent. The TSRV estimator based on the average realized volatility is defined by

$$TSRV_{t_n} = ARV_{t_n}(Y, \mathcal{G}, K) - \frac{n - K + 1}{K} [Y, Y]_{t_n}^{\mathcal{G}}, \quad (4.12)$$

where the second term accounts for the bias in the ARV . It can be shown that $TSRV_{t_n} \xrightarrow{P} [X, X]_{t_n}$ under the noise assumption.³ The convergence rate is $n^{1/6}$.

Moreover, Zhang et al. [2005] derive an optimal parameter K^* defining the subgrids \mathcal{G}_k , $k = 1, \dots, K^*$. It is given by

$$K^* = \left(\frac{t_n}{12\omega^4} \int_0^{t_n} \sigma_t^4 dt \right)^{-1/3} n^{2/3}. \quad (4.13)$$

In case of i.i.d. microstructure noise, the integrated quarticity can be estimated by $t_n \int_0^{t_n} \sigma_t^4 dt \approx (RV_{t_n}^K)^2$, where $RV_{t_n}^K$ is a sparsely sampled realized variance estimator (see Barndorff-Nielsen et al. [2008b]). Estimates for ω^2 can be computed according to (4.8) and (4.9).

(iii) The Realized Kernel estimator

Barndorff-Nielsen et al. [2008a] propose an alternative noise-robust estimator of the quadratic variation that in contrast to the TSRV is also shown to be consistent under

³Under weaker noise assumptions Zhang [2006] proposes a multi scale realized volatility estimator.

Due to its closeness to the TSRV and since weaker noise assumptions are also covered by the realized kernel estimator we do not consider this approach here.

serially dependent microstructure noise and dependence between noise and efficient price. The basic underlying idea is to exploit the similarity between estimating a long-run variance of a serially dependent stationary time series and estimating the integrated variance from autocorrelated price observations. Barndorff-Nielsen et al. [2008a] employ kernel-based methods which are widely similar to those used for heteroscedasticity and autocorrelation robust covariance (HAC) estimation. Their realized kernel estimator is defined by

$$\text{RK}_{t_n} = \sum_{h=-H}^{h=H} k\left(\frac{h}{H-1}\right) \gamma_h, \quad (4.14)$$

where H is an additional bandwidth parameter and γ_h are the realized sample autocorrelations

$$\gamma_h = \sum_{i=h+1}^n (Y_{t_i} - Y_{t_{i-1}})(Y_{t_{i-h}} - Y_{t_{i-h-1}}). \quad (4.15)$$

The Parzen kernel k in (4.14) is given as

$$k(x) = \begin{cases} 1 - 6x^2 + 6x^3, & 0 \leq x \leq 1/2, \\ 2(1-x)^3, & 1/2 \leq x \leq 1, \\ 0, & \text{otherwise.} \end{cases} \quad (4.16)$$

The convergence rate of the estimator (4.14) is $n^{1/5}$.

For practical applications, Barndorff-Nielsen et al. [2008b] propose to use local averages of returns at the end points of the interval of interest to account for the end-effects due to the missing sample size adjustment in the computation of the realized sample autocorrelations. Moreover, the optimal bandwidth H can be computed based on

$$H^* := c^* \xi^{4/5} n^{3/5}, \quad \text{where } \xi^2 := \frac{\omega^2}{\sqrt{t_n \int_0^{t_n} \sigma_t^4 dt}}, \quad (4.17)$$

$c^* = 3.5134$ for the Parzen kernel and ω^2 denotes the variance of the microstructure noise ε . Estimates of the integrated quarticity and the noise variance can be derived as outlined above.

(iv) The Pre-averaging Estimator

In order to reduce the influence of microstructure noise, Jacod et al. [2009] propose to average the price observations before computing the realized variance. The pre-averaging approach is consistent under various noise specifications like e.g. autocorrelated noise. It is based on the averaged prices

$$\bar{Z}_i^n := \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right) \Delta_{i+j}^n Y, \quad \text{where } \Delta_i^n Y := Y_{t_i} - Y_{t_{i-1}}, \quad (4.18)$$

k_n is an integer sequence satisfying $k_n \Delta^{1/2} = \theta + o(\Delta^{1/4})$ with $\theta > 0$. The weighting function g is given by $g : [0, 1] \rightarrow \mathbb{R}$, $x \mapsto g(x) := \min(x, 1-x)$ in our empirical

4.2 Intraday volatility measures

analysis. We fix the integer sequence to $k_n = \theta\sqrt{\Delta}^{-1}$. The pre-averaging estimator of the integrated volatility is then defined by

$$\text{PreAv}_{t_n} = \frac{\sqrt{\Delta}}{\theta\psi_2} V(Z, 2)_{t_n} - \frac{\psi_1\Delta}{2\theta^2\psi_2} RV_{t_n} \quad (4.19)$$

with

$$V(Z, 2)_{t_n} := \sum_{i=0}^{[t_n/\Delta]-k_n} |\bar{Z}_i^n|^2 \quad (4.20)$$

and $\psi_1 = \int_0^1 (g'(s))^2 ds = 1$ as well as $\psi_2 = \int_0^1 (g(s))^2 ds = 1/12$. We apply the finite sample adjustments given in Hautsch and Podolskij [2010] which makes the estimator attractive in view of the finite sample problem in intraday volatility estimation. Details on the procedure can be found in the appendix.

The TSRV, the realized kernel and the pre-averaging estimators are essentially all based on averaging out the microstructure noise from the data. However, due to the computational differences of the estimators we expect different finite sample behavior.

(v) The Realized-Range Estimator

As an alternative volatility measure we consider the realized range of an interval $[0, t_n]$ introduced by Parkinson [1980] (see also Alizadeh et al. [2002]).⁴ It is given by

$$RR_{t_n} = \frac{1}{4\ln(2)} (\sup_{0 \leq t \leq t_n} Y_t - \inf_{0 \leq t \leq t_n} Y_t)^2. \quad (4.21)$$

In contrast to the estimators outlined above the realized range relies on just two informative intra-interval observations and thus potentially exploits intra-interval information in a very effective way.

4.2.2 Volatility measures based on spot volatility estimators

However, the interval of interest might be so small that the microstructure noise can hardly be averaged out using only data from within the interval. Moreover, both the TSRV and RK estimator induce boundary problems as they are based on lags of the data or sparsely sampled data (see (4.12) and (4.14)).

To solve the problems associated with applying realized volatility estimators to intraday intervals we consider recovering integrated volatility estimates from spot volatility estimators, i.e. from estimators of the instantaneous volatility σ_t^2 in the log-price model (4.1). Suppose that we are interested in the interval $[t_l, t_n]$, $t_l > 0$ such that data t_0, \dots, t_l outside of the interval, i.e. prior to the left interval boundary, is available. Spot volatility measures can use the information of price observations outside of the interval under consideration. Once the spot volatility estimates $\hat{\sigma}_t^2$, $t \in [t_l, t_n]$, are obtained, numerical integration yields estimators \widehat{IV} of the integrated volatility according to e.g. the

⁴Closely related realized range estimators are given by Christensen and Podolskij [2007] and Martens and van Dijk [2007].

4 Modelling Volatility Measures on Intra-Daily Time Intervals

Riemann sum,

$$\widehat{IV} := \Delta \sum_{t=t_l}^{t_n} \widehat{\sigma}_t^2. \quad (4.22)$$

The estimation error can be decomposed according to

$$\int_{t_l}^{t_n} \sigma_t^2 dt - \widehat{IV} = \underbrace{\int_{t_l}^{t_n} \sigma_t^2 dt - \Delta \sum_{t=t_l}^{t_n} \sigma_t^2}_{\text{integral approximation error}} + \underbrace{\Delta \sum_{t=t_l}^{t_n} (\sigma_t^2 - \widehat{\sigma}_t^2)}_{\text{local estimation error}}. \quad (4.23)$$

Andersen et al. [2008] show that the Riemann approximation error is asymptotically negligible (of order $O(n^{-1})$) in case the σ_t^2 process satisfies a Lipschitz condition

$$\exists \kappa > 0 : \sup_{s,t} |\sigma_s^2 - \sigma_t^2| < \kappa |s-t|. \quad (4.24)$$

Specifically, the vanishing integral approximation error means that we can construct consistent estimators of the integrated volatility based on consistent spot volatility estimates.

In the following, we review both nonparametric and parametric approaches to spot volatility estimation that differ in their convergence rates, their robustness to market microstructure noise as well as their finite sample properties.

(i) Nonparametric estimators based on Kristensen [2010]

In a nonparametric framework for spot volatility estimation, Kristensen [2010] proposes to take kernel weighted averages of squared returns in order to obtain estimates of the instantaneous volatility.⁵ In this context, the local linear estimator is given by

$$\widehat{\sigma}_{t,LL}^2 = \frac{\sum_{i=1}^n w_{t_{i-1}}(t)(Y_{t_i} - Y_{t_{i-1}})^2}{\sum_{i=1}^n w_{t_{i-1}}(t)}, \quad t \in [t_l, t_n], \quad (4.25)$$

with weights

$$\begin{aligned} w_{t_{i-1}}(t) &:= \Delta K_h(t_{i-1} - t) \{S_{n,2} - (t_{i-1} - t)S_{n,1}\}, \\ S_{n,k} &:= \Delta \sum_{i=1}^n K_h(t_{i-1} - t)(t_{i-1} - t)^k. \end{aligned} \quad (4.26)$$

The estimator is based on kernels K_h with bandwidth h which are defined as either normal, one-sided or Epanechnikov kernels (see the appendix for definitions). We choose not to restrict the analysis to just one kernel since the finite sample properties are not clear. Under $h \rightarrow 0$ and $nh \rightarrow \infty$ for $n \rightarrow \infty$, smoothness assumptions for the σ_t process as well as the absence of microstructure noise one can show the consistency, $\widehat{\sigma}_{t,LL}^2 \xrightarrow{p} \sigma_t^2$, of the local linear estimator.

⁵See also Bandi and Reno [2009] for nonparametric estimators. A special case of the kernel weighted estimators is the rolling-window estimator by Foster and Nelson [1996].

4.2 Intraday volatility measures

Alternatively, we consider the asymmetric beta kernel estimator by Chen [2000],

$$\hat{\sigma}_{t,beta}^2 = \frac{\sum_{i=1}^n K_b(t_{i-1}/T, t/T, h)(Y_{t_i} - Y_{t_{i-1}})^2}{\Delta \sum_{i=1}^n K_b(t_{i-1}/T, t/T, h)}, \quad t \in [t_l, t_n], \quad (4.27)$$

with kernels defined via the beta density

$$K_b(x, \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\text{Beta}(\alpha, \beta)}, \quad (4.28)$$

where $\text{Beta}(x, y) := \Gamma(x+y)/(\Gamma(x)\Gamma(y))$ is the beta function defined in terms of the gamma function Γ . The estimator is of the Nadaraya-Watson-type. However, unlike the standard Nadaraya-Watson kernel estimators the beta kernel estimator is consistent at the boundaries of the sample. The reason is that the beta kernel assigns weights according to the position of t in the sample with zero weight outside of the known observations.

In contrast to the estimators (i) to (v) from subsection 4.2.1 applied to estimation of an interval $[t_l, t_n]$ we note that the spot volatility estimators (4.25) and (4.27) make use of *all* available squared returns $(Y_{t_i} - Y_{t_{i-1}})^2$, $i = 1, \dots, n$, with weights according to their distance to t . Specifically, they do not rely exclusively on the observations within the intraday interval of interest. The price of this potential advantage is the typically slow convergence of the spot volatility estimators which means they require a large sample of price observations around the intraday interval of interest. Kristensen [2010] reports a highest attainable convergence rate of $O_p(n^{-\gamma/(2\gamma+1)})$ for the estimators (4.25) and (4.27) with the parameter $\gamma > 0$ depending on the smoothness of the σ_t^2 process and chosen kernel.

The kernel estimation methods depend on a bandwidth h which needs to be chosen by the researcher. A data-driven bandwidth selection method is based on the integrated squared estimation error

$$ISE(h) = \int_{t_l}^{t_n} [\sigma_t^2 - \hat{\sigma}_t^2]^2 dt = \int_{t_l}^{t_n} \hat{\sigma}_t^4 dt + \int_{t_l}^{t_n} \sigma_t^4 dt - 2 \int_{t_l}^{t_n} \sigma_t^2 \hat{\sigma}_t^2 dt. \quad (4.29)$$

An optimal h can be obtained by the numerical minimization of the corresponding loss function defined by

$$CV(h) = \sum_{i=1}^n \mathbb{1}\{t_l \leq t_{i-1} \leq t_n\} \left[\hat{\sigma}_{-i,t_{i-1}}^4 \Delta - 2(Y_{t_i} - Y_{t_{i-1}})^2 \hat{\sigma}_{-i,t_{i-1}}^2 \right], \quad (4.30)$$

where $\hat{\sigma}_{-i,t_{i-1}}^2$ as well as $\hat{\sigma}_{-i,t_{i-1}}^4$ denote the leave-one-out estimators corresponding to $\hat{\sigma}_t^2$ and $\mathbb{1}\{\cdot\}$ is the indicator function.

(ii) Spot volatility estimators based on noise-robust realized measures of volatility

To account for potential microstructure noise pollution not covered by the Kristensen [2010] setup, Boswijk and Zu [2010] propose a spot volatility version of the TSRV esti-

mator as

$$\hat{\sigma}_{t,TSRV}^2 = \frac{\text{TSRV}_t - \text{TSRV}_{t-h}}{h}, \quad t \in [t_l, t_n], \quad (4.31)$$

where $h > 0$ is a bandwidth and both TSRV integrated volatility estimates for the intervals $[0, t]$ and $[0, t - h]$ depend on the additional parameter K_{spot} . The estimator represents the derivative of the TSRV volatility measure at t for $h \rightarrow 0$. Boswijk and Zu [2010] derive the consistency of the estimator (4.31) for $h \rightarrow 0$ and $n \rightarrow \infty$ with an convergence rate $n^{-1/12}$ under i.i.d. market microstructure noise and under assumptions on the true price and volatility process.⁶

Likewise, Bos et al. [2009] devise a spot volatility estimator based on the pre-averaging method (4.19) according to

$$\hat{\sigma}_{t,PreAv}^2 = \frac{\text{PreAv}_t - \text{PreAv}_{t-h}}{h}, \quad t \in [t_l, t_n], \quad (4.32)$$

with θ_{spot} as additional parameter for the pre-averaging estimators.

The localized version of the noise-robust realized kernel method utilizes - realized autocorrelations, γ_k^{loc} , defined by

$$\gamma_k^{loc}(t, h) = \sum_{i=k+1}^n K_1(t_{i-1}, th^{-2} + 1, (1-t)h^{-2} + 1)(Y_{t_i} - Y_{t_{i-1}})(Y_{t_{i-k}} - Y_{t_{i-k-1}}), \quad (4.33)$$

where K_1 is given by the beta kernel (4.28). Ikeda [2010] introduces the so-called double window realized kernel spot volatility estimator as

$$\hat{\sigma}_{t,RK}^2(h, H_{spot}) = \gamma_0^{loc}(t, h) + \sum_{k=-H_{spot}}^{H_{spot}} K_2\left(\frac{k}{H_{spot}}\right) \gamma_k^{loc}(t, h), \quad t \in [t_l, t_n] \quad (4.34)$$

where K_2 is the Parzen kernel (4.16). He argues that an optimal bandwidth H_{spot} can be computed based on (4.17).

(iii) A duration-based spot volatility estimator

The duration-based approach of Andersen et al. [2008] exploits the dual relationship between durations of price events and the volatility of prices. Let the price durations be defined as the time that elapses until the price change is above a threshold δ^{dur} . Given that the log price process has a constant variance in a local neighborhood of t , a local (spot) volatility estimator can be expressed in terms of the forward first exit time τ^+ and backward first exit time τ^- ,

$$\begin{aligned} \tau^+(t, \delta^{dur}) &= \inf\{c > 0 : |Y_{t+c} - Y_t| > \delta^{dur}\}, \\ \tau^-(t, \delta^{dur}) &= \inf\{c > 0 : |Y_{t-c} - Y_t| > \delta^{dur}\}. \end{aligned} \quad (4.35)$$

The idea of using both τ^+ and τ^- for volatility estimation is that boundary effects due to

⁶Moreover, the authors give a formula for the estimation of the optimal bandwidth which, however, relies on an ad hoc choice of an initial bandwidth.

missing duration data can be avoided. In this respect the corresponding spot volatility estimate is

$$\hat{\sigma}_{t,dur}^2 = \begin{cases} \frac{1}{\mu_1} \frac{(\delta^{dur})^2}{\tau^+}, & \text{if } t \text{ is in the first half of } [t_0, t_n], \\ \frac{1}{\mu_1} \frac{(\delta^{dur})^2}{\tau^-}, & \text{if } t \text{ is in the second half of } [t_0, t_n]. \end{cases} \quad (4.36)$$

The constant μ_1 can be shown to be equal to $2C = 1.8319$ where C is the Catalan constant (see Andersen et al. [2008]). The authors derive the consistency of the estimator in the case without microstructure noise. However, the estimator is shown to be widely robust to various specifications of market microstructure noise in a simulation study.

(iv) Spot volatility estimation in a parametric ACD framework

In view of the faster parametric convergence rates we employ the autoregressive conditional duration (ACD) model introduced by Engle and Russell [1998]. Let a price event be defined as the price change by an amount δ^{acd} . Starting from the first price observation we construct a time series of durations consisting of the time elapsing between price events. Let $x_i := t_{d,i} - t_{d,i-1}$, $i = 1, \dots, n_d$, denote the durations, where the $t_{d,i}$ are the time stamps associated with the occurrence of price events. The ACD model is defined by the equation

$$\varepsilon_i = \frac{x_i}{\Psi_i}, \quad i = 1, \dots, n_d, \quad (4.37)$$

where the ε_i are i.i.d. random variables with mean one and Ψ_i is the conditional mean of x_i .

The conditional mean is typically modelled as a linear function of the past information available,

$$\Psi_i = \omega + \sum_{j=1}^p \alpha_j x_{i-j} + \sum_{j=1}^q \beta_j \Psi_{i-j}. \quad (4.38)$$

ω , α_j , $j = 1, \dots, p$, and β_j , $j = 1, \dots, q$, are parameters to be estimated. The estimation is based on maximizing the likelihood corresponding to the distribution assumption for ε_i .

Engle and Russell [1998] establish the link between the ACD model and the spot volatility based on the log price series $Y_{t_{d,i}}$, $i = 1, \dots, n_d$, associated with the price events. The spot volatility at t , $t_{d,i} < t < t_{d,i+1}$, conditional on the information $\mathcal{F}_{t_{d,i}}$ is defined by

$$\begin{aligned} \sigma^2(t|\mathcal{F}_{t_{d,i}}) &= \lim_{\Delta_t \rightarrow 0} \left\{ \frac{1}{\Delta_t} \mathbb{E}[(Y_{t+\Delta_t} - Y_t)^2 | \mathcal{F}_{t_{d,i}}] \right\} \\ &= \lim_{\Delta_t \rightarrow 0} \left\{ \frac{1}{\Delta_t} \mathbb{P}(|Y_{t+\Delta_t} - Y_t| = \delta^{acd} | \mathcal{F}_{t_{d,i}}) (\delta^{acd})^2 \right\} \\ &= \lambda(t|\mathcal{F}_{t_{d,i}}, \delta^{acd}) (\delta^{acd})^2, \end{aligned} \quad (4.39)$$

where Δ_t is a time interval and λ denotes the conditional intensity function of the ACD process. The second equation of (4.39) uses that the expectation can be directly rewritten in terms of the probability to observe a price event. In the limit, this probability is given

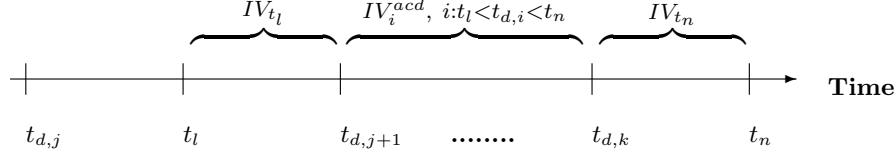


Figure 4.1: Example of duration time stamps $t_{d,i}$ not aligned with the interval boundaries t_l and t_n

by the conditional intensity function, which gives the instantaneous probability of a price change of δ^{acd} units. The conditional intensity function of the ACD based on the durations x_i is defined as

$$\lambda(t|\mathcal{F}_{t_{d,i}}) = \tilde{\lambda}\left(\frac{t - t_{d,i}}{\Psi_{i+1}}\right) \frac{1}{\Psi_{i+1}}, \quad (4.40)$$

with $\tilde{\lambda}$ denoting the hazard function of the innovations ε_i , $\tilde{\lambda}(\cdot) := f(\cdot)/(1 - F(\cdot))$, where f is the density of ε_i and F is its cumulative distribution function.

The integrated volatility over $[t_{d,i}, t_{d,i+1}]$ conditional on $\mathcal{F}_{t_{d,i}}$ can thus be obtained according to

$$IV_i^{acd} = \int_{t_{d,i}}^{t_{d,i+1}} \sigma_{t|\mathcal{F}_{t_{d,i}}}^2 dt = (\delta^{acd})^2 \frac{1}{\Psi_{i+1}} \int_{t_{d,i}}^{t_{d,i+1}} \tilde{\lambda}\left(\frac{t - t_{d,i}}{\Psi_{i+1}}\right) dt. \quad (4.41)$$

Depending on the distribution assumption for the innovation the integration is done numerically or analytically (see also Tse and Yang [2010]).

However, the time stamps of the price events might not be aligned with the interval boundaries of the interval of interest, $[t_l, t_n]$, such that problems at the boundary arise. We take a practical stance to this problem and compute estimates of the integrated volatility over $[t_l, t_n]$ as

$$IV_{[t_l, t_n]}^{acd} := \sum_{i: t_l < t_{d,i} < t_n} IV_i^{acd} + IV_{t_l} + IV_{t_n}, \quad (4.42)$$

where IV_{t_l} and IV_{t_n} for $t_{d,j} \leq t_l \leq t_{d,j+1}$ and $t_{d,k} \leq t_n$ are given as fractions of the volatility assigned to the durations containing the interval boundaries,

$$IV_{t_l} := \frac{t_{d,j+1} - t_l}{t_{d,j+1} - t_{d,j}} IV_j^{acd} \quad \text{and} \quad IV_{t_n} := \frac{t_n - t_{d,k}}{t_{d,k} - t_{d,k-1}} IV_{k-1}^{acd}. \quad (4.43)$$

The Figure (4.1) illustrates the problem and the outlined procedure.

4.3 Comparing Intraday Volatility Estimators in a Monte Carlo Simulation Study

We compare the competing integrated volatility estimators in simulation study setups with market microstructure noise of different magnitude added, respectively. In all cases we simulate 1,000 trajectories of log price processes with 23,400 observations each.

In the first case, the log price process is assumed to follow a 1-factor stochastic volatility model augmented by a seasonality component,

$$dY_t = \mu dt + \sigma_t dW_t, \quad \sigma_t = S_t \exp(\beta_0 + \beta_1 \tau_t), \quad (4.44)$$

$$d\tau_t = \alpha \tau_t dt + dB_t, \quad \text{Corr}(dW_t, dB_t) = \phi, \quad (4.45)$$

where W_t and B_t are standard Brownian motions and ϕ is the leverage parameter. S_t denotes the deterministic U-shaped seasonality component with $\int_0^1 S_t dt = 1$ and $S_t := S(t)$ given by

$$S(t) = \delta^s \bar{t} + \sum_{j=1}^Q (\delta_{1,j}^s \cos(\bar{t}2\pi j) + \delta_{2,j}^s \sin(\bar{t}2\pi j)), \quad (4.46)$$

where \bar{t} is the normalized intraday time. The seasonality component is based on the estimation of the seasonality for AAPL.OQ squared mid-quote return data in 2008. Moreover, we calibrate the model using the configuration of Barndorff-Nielsen et al. [2008a], $\beta_0 = \beta_1^2/(2\alpha)$ and $\mu = 0.03$, $\beta_1 = 0.125$, $\alpha = -0.025$, $\phi = -0.3$ and initial value $Y_0 = 20$. The process τ_t is initialized with its stationary limiting distribution $\tau_t \sim N(0, -1/(2\alpha))$.

In the second case the log price process is assumed to follow a 2-factor stochastic volatility model which allows for more pronounced heteroscedasticity. The 2-factor model in the spirit of Huang and Tauchen [2005] and Goncalves and Meddahi [2009] augmented by a seasonality component is given by

$$dY_t = \mu dt + \sigma_t dW_t, \quad \sigma_t = S_t \exp^*(\beta_0 + \beta_1 \tau_{1t} + \beta_2 \tau_{2t}), \quad (4.47)$$

$$d\tau_{1t} = \alpha_1 \tau_{1t} dt + dB_{1t}, \quad d\tau_{2t} = \alpha_2 \tau_{2t} dt + (1 + \phi \tau_{2t}) dB_{2t}, \quad (4.48)$$

$$\text{Corr}(dW_t, dB_{1t}) = \phi_1, \quad \text{Corr}(dW_t, dB_{2t}) = \phi_2, \quad (4.49)$$

where W_t , B_{1t} and B_{2t} are standard Brownian motions and S_t is the seasonality component. Moreover,

$$\exp^*(x) = \begin{cases} \exp(x), & x \leq \ln(1.5), \\ 1.5 \sqrt{1 - \ln(1.5) + x^2/\ln(1.5)}, & x \geq \ln(1.5), \end{cases} \quad (4.50)$$

which is numerically more stable than the exponential function. In order to calibrate the model we configure according to Huang and Tauchen [2005], i.e. $\mu = 0.03$, $\beta_0 = -1.2$, $\beta_1 = 0.04$, $\beta_2 = 1.5$, $\alpha_1 = 0.00137$, $\alpha_2 = -1.386$, $\phi = 0.25$, $\phi_1 = \phi_2 = -0.3$ and $Y_0 = 20$.

	AAPL.OQ		NYT.N		KFT.N		XOM.N	
	MQ	TR	MQ	TR	MQ	TR	MQ	TR
$\hat{\omega}_1^2$	0.0011	0.0013	0.0010	0.0013	0.0002	0.0006	0.0004	0.0007
$\hat{\omega}_2^2$	0.1912	0.1937	0.5116	0.5581	0.0917	0.1001	0.0829	0.0853

Table 4.1: Microstructure noise estimates in 0.000001 for the day 03/13/2008 for selected stocks at NYSE and NASDAQ in 2008. MQ denotes estimates based on log mid-quote prices and TR denotes estimates based on log transaction prices

The simulation of both model is based on the corresponding Euler schemes (see the appendix for details). The time between two observations is thought to be one second, normalized to 1/23,400. Hence, each simulated trajectory spans the interval $[0, 1]$, which is taken to be a (NYSE) trading day with 6.5 trading hours.

Figure 4.2 and 4.3 illustrate typical evolutions of the stochastic volatility and the corresponding price path implied by the 1-factor and 2-factor models. We observe a more volatile volatility process in case of the 2-factor model, which can be considered more realistic since volatility estimates from real data typically exhibit jumps and spikes. The reason is the additional factor of the 2-factor model which causes short run fluctuations of the volatility from the long-run trend and results in a more volatile volatility process.

We present simulation results for different degrees of noise. In the first scenario, we do not consider contamination by microstructure noise. Second, we add an i.i.d. noise process $\varepsilon_t \sim N(0, \omega^2)$ with variance $\omega^2 = 0.000001$ to the simulated log prices Y_t . The magnitude of the noise is motivated by the largest market microstructure noise estimates for four NYSE and NASDAQ traded stocks. Table 4.1 shows the estimates $\hat{\omega}_1^2$ and $\hat{\omega}_2^2$ based on the estimators (4.8) and (4.9). Third, we add an autocorrelated noise $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$ with innovations $u_t \sim N(0, (1 - \rho^2)\omega^2)$ such that the unconditional moments are given by $\mathbb{E}[\varepsilon_t] = 0$ and $\mathbb{E}[\varepsilon_t^2] = \omega^2$. In line with the simulation study of Andersen et al. [2008], who propose high autocorrelation coefficients, we choose $\rho = 0.8$. Empirical evidence of autocorrelation in the market microstructure noise is given by Hansen and Lunde [2006].

We apply the integrated volatility estimators to the estimation of 20 second, 1 minute and 5 minute intervals corresponding to 20, 60 and 300 observations. As the spot volatility estimators require data before the intervals of interest, we consider only mid-day intervals after 2 trading hours. Let IV_i denote the (true) integrated volatility for an interval $[t_l, t_r]$ in simulation run i , $i = 1, \dots, N$.⁷ Let the corresponding estimators of IV_i be denoted \widehat{IV}_i . Moreover, to account for the fact that the volatility paths are different in each simulation run i we normalize the estimators \widehat{IV}_i by the value IV_i yielding $\widehat{IV}_i^{norm} := \widehat{IV}_i / IV_i$. The normalized true integrated volatility has value one, $IV_i^{norm} := IV_i / IV_i = 1$. To assess the performance of the competing estimators we

⁷The true value IV_i is computed as Riemann sum using the simulated (true) spot volatilities σ_t^2 .

4.3 Comparing Intraday Volatility Estimators in a Monte Carlo Simulation Study

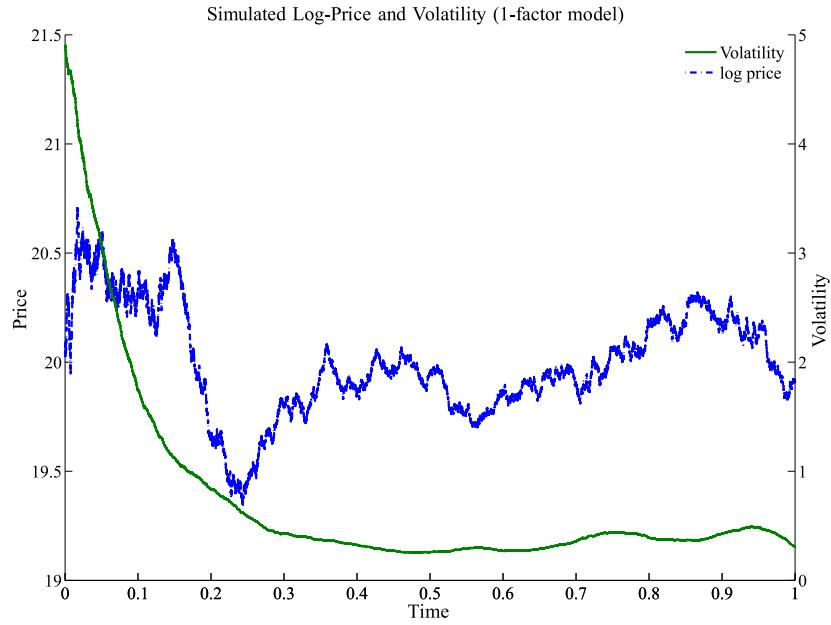


Figure 4.2: Example of a simulated log-price process realization and the corresponding spot volatilities based on the 1-factor model with diurnal seasonality component

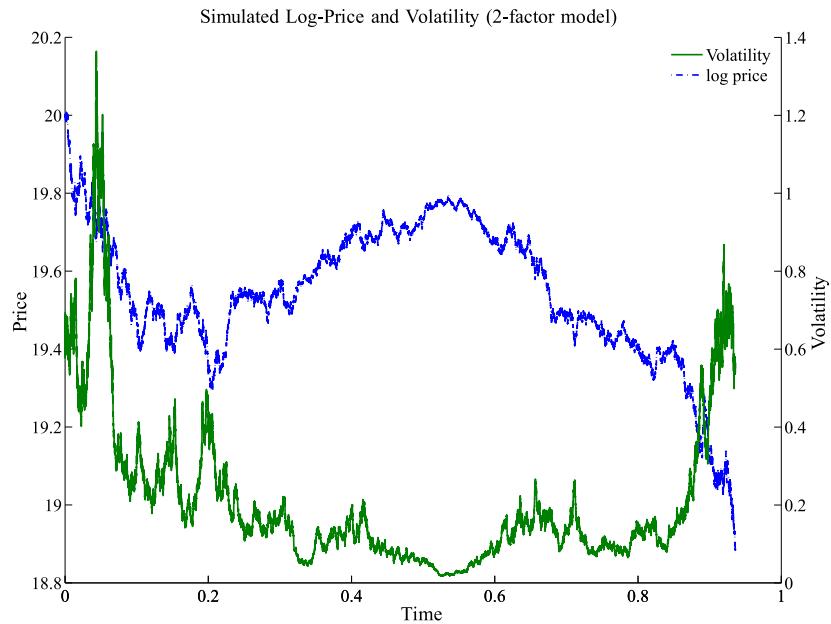


Figure 4.3: Example of a simulated log-price process realization and the corresponding spot volatilities based on the 2-factor model with diurnal seasonality component

4 Modelling Volatility Measures on Intra-Daily Time Intervals

report the root mean squared error of the normalized estimates,

$$\text{RMSE} := \sqrt{\frac{1}{N} \sum_{i=1}^N (\widehat{IV}_i^{norm} - 1)^2}. \quad (4.51)$$

The variance of the normalized estimators is given as

$$V := \frac{1}{N} \sum_{i=1}^N \left(\widehat{IV}_i^{norm} - \left(\frac{1}{n} \sum_{i=1}^n \widehat{IV}_i^{norm} \right) \right)^2. \quad (4.52)$$

Both the spot volatility based estimators and the estimators outlined in subsection 4.2.1 require to calibrate additional bandwidths and parameters. To select the bandwidths and parameters optimally we conduct a grid search based on a separate simulation study for each of the three noise specifications. The optimal bandwidths and parameters are defined as the values that minimize the RMSE for 5 minute intervals on a grid of possible parameter values. For the grid search we run 1,000 simulations. The grids are taken to be $\mathcal{G}_1 := \{10, 20, \dots, 300\}$ for the parameters K , K_{spot} , H , H_{spot} , and $\mathcal{G}_2 := \{0.05, 0.10, \dots, 0.7\}$ for the bandwidths h of the spot volatility estimators. The Grid for θ and θ_{spot} are $\mathcal{G}_3 := \{0.1, 0.2, \dots, 3\}$. Optimal values for δ_{acd} and δ_{dur} are based on the grid $\mathcal{G}_4 := \{0.01, 0.02, \dots, 0.2\}$. Table 4.2 gives an example of the optimal bandwidths and parameters for the 1-factor stochastic volatility model according to the grid search procedure. Corresponding values for the 2-factor model are given in the appendix. Comparing the spot volatility bandwidth for the two log price models we note that the bandwidths tend to be higher for the 1-factor model, reflecting the smoother underlying volatility process (see also Figure (4.2) and (4.3)).

ω^2	no noise	1E-06	1E-06 (ac)
Spot volatility based estimators			
bandwidths h in minutes			
LLnormal	3.85	3.85	3.85
LLonesided	7.70	3.85	3.85
LLepanechnikov	7.70	3.85	3.85
asym. Beta	19.25	26.95	23.10
TSRV _{spot}	15.40	23.10	19.25
PreAv _{spot}	23.10	15.40	23.10
RK _{spot}	7.70	7.70	7.70
additional parameters			
K_{spot}	30.00	40.00	40.00
θ_{spot}	1.40	1.60	1.50
H_{spot}	70.00	40.00	100.00
Estimators from subsection 4.2.1			
K	10.00	20.00	40.00

4.3 Comparing Intraday Volatility Estimators in a Monte Carlo Simulation Study

H	20.00	10.00	20.00
θ	1.50	1.00	1.10
δ_{acd}	0.005	0.007	0.005
δ_{dur}	0.014	0.014	0.014

Table 4.2: Optimal bandwidths and parameters for the 1-factor model according to the grid-search procedure. Abbreviations are LLnormal, LLonesided, LLepanechnikov for the local linear estimator with normal, one-sided and Epanechnikov kernel. asym Beta denotes the asymmetric beta kernel estimator

Tables 4.3 and 4.4 give the RMSEs and standard errors of the integrated volatility estimators for 20 s, 1 min and 5 min intervals in the three simulation setups for the 1- and 2-factor model, respectively. The bandwidths are chosen according to the grid search procedure.

We summarize the following main findings: First, in all noise setups for the 20 second intervals we find that spot volatility based estimators for the integrated volatility perform best in terms of the RMSE. Hence, the RMSE performance improves when estimators use information from outside of the 20 second interval of interest. The additional outside information outweighs the slow convergence rates of the spot volatility estimators. In contrast, the integrated volatility estimators from subsection 4.2.1 suffer from the small sample size of only 20 observations within each interval. However, the standard errors of the spot volatility based estimates are typically large due to the small optimal bandwidths. Testing for equality of the means of normalized estimates via t-statistics we find that the estimates which minimize the RMSE are indeed significantly different from all other estimate series at a 5% level.

Second, the RMSE gains of spot volatility estimators over the estimators from subsection 4.2.1 deteriorate with increasing interval lengths. In case of the 2-factor model the spot volatility based estimators have higher RMSE than competing models for 5 out of 6 simulation setups at the 1 minute and 5 minute frequency. This finding reflects the typically slow convergence of the spot volatility estimates which becomes more and more costly as the interval of interest grows. At the same time, the within interval sample size grows, thus alleviating the finite sample problem of the standard realized volatility estimators.

Third, the performance of the spot volatility estimators depends on the nature of the volatility process. We find that the Kristensen [2010] local linear estimator has the lowest RMSE in each setup of the 1-factor simulation study. However, the RMSE performance of the spot volatility based estimators compared to the competitors deteriorates when the volatility process changes from the smooth 1-factor specification to the more volatile 2-factor specification.

Simulation Results: RMSE of integrated volatility estimators for 20s, 1 and 5 minute intervals						
ω^2	0	1E-06	1E-06(ac)	0	1E-06	1E-06(ac)
20 seconds						
LLnormal	0.0055*	0.0137*	0.0078*	0.0055*	0.0136*	0.0042*
LLonesided	(0.0050)	(0.0091)	(0.0078)	(0.0049)	(0.0090)	(0.0064)
LLepanech.	0.0128	0.0296	0.0202	0.0132	0.0275	0.0169
	(0.0128)	(0.0235)	(0.0202)	(0.0132)	(0.0217)	(0.0169)
asym. Beta	0.0127	0.0255	0.0187	0.0110	0.0238	0.0144
	(0.0128)	(0.0201)	(0.0187)	(0.0110)	(0.0186)	(0.0144)
TSRV _{spot}	0.1602	0.2070	0.3793	0.1582	0.2054	0.3791
	(0.0011)	(0.0008)	(0.0003)	(0.0010)	(0.0007)	(0.0003)
PreAv _{spot}	0.4041	0.4328	0.3928	0.4038	0.4270	0.4019
	(0.0039)	(0.0105)	(0.0052)	(0.0034)	(0.0118)	(0.0056)
RK _{spot}	0.4227	0.4696	0.4150	0.4218	0.4631	0.4233
	(0.0061)	(0.0053)	(0.0041)	(0.0052)	(0.0058)	(0.0042)
ACD _{spot}	0.0538	0.0173	0.0496	0.0440	0.0200	0.0517
	(0.0534)	(0.0116)	(0.0470)	(0.0438)	(0.0125)	(0.0501)
DUR _{spot}	1.9770	2.3604	1.6231	1.4708	2.0819	1.2572
	(0.7579)	(0.8765)	(0.5364)	(0.2553)	(0.4099)	(0.1955)
1 minute						
PreAv	0.9770	0.9534	0.9811	0.9334	0.9107	0.9363
TSRV	0.3917	0.3635	0.3727	0.2313	0.2421	0.2181
RK	0.2936	0.2631	0.2785	0.1779	0.1786	0.1721
5 minute						
0						
1E-06						
1E-06(ac)						
Estimators based on Spot Volatility Models						
0.0042*						
0.004*						
0.004*						

Estimators from subsection 4.2.1

PreAv	0.9770 (0.0001)	0.9534 (0.0001)	0.9811 (0.0001)	0.9334 (0.0004)	0.9107 (0.0002)	0.9363 (0.0003)	0.8223 (0.0009)	0.7920 (0.0004)	0.8208 (0.0009)
TSRV	0.3917 (0.2645)	0.3635 (0.2276)	0.3727 (0.2491)	0.2313 (0.1836)	0.2421 (0.1825)	0.2181 (0.1706)	0.0597 (0.0434)	0.0567 (0.0417)	0.0546 (0.0414)
RK	0.2936	0.2631	0.2785	0.1779	0.1786	0.1721	0.0381	0.0352	0.0350

4.3 Comparing Intraday Volatility Estimators in a Monte Carlo Simulation Study

	(0.2332)	(0.2034)	(0.2233)	(0.1539)	(0.1459)	(0.1472)	(0.0371)	(0.0342)
RV ^K	0.5119	0.4812	0.4559	0.3093	0.3331	0.2974	0.0693	0.0675
	(0.4840)	(0.4493)	(0.4102)	(0.2911)	(0.3122)	(0.2819)	(0.0683)	(0.0674)
RV	0.1067	0.1300	0.1077	0.0312	0.0471	0.0335	0.0072	0.0129
	(0.1066)	(0.1200)	(0.1078)	(0.0312)	(0.0408)	(0.0334)	(0.0072)	(0.0079)
Real. Range	0.4498	0.3885	0.3859	0.3893	0.3646	0.3786	0.3862	0.3711
	(0.3916)	(0.3370)	(0.3187)	(0.3715)	(0.3428)	(0.3534)	(0.3789)	(0.3687)
Sq. Return	1.8105	1.6695	1.9484	1.7431	1.8658	2.6801	2.0371	2.1019
	(1.8116)	(1.6685)	(1.9486)	(1.7442)	(1.8676)	(2.6783)	(2.0384)	(2.1039)

Table 4.3: RMSE for 20 second, 1 minute and 5 minute (integrated) volatility estimates. Results are based on 1,000 simulated trajectories of the one factor stochastic volatility model with alternative noise specifications. Notation as above, (ac) denotes the autocorrelated noise. Lowest RMSE highlighted in each column

ω^2	Simulation Results: RMSE of integrated volatility estimators for 20s, 1 and 5 minute intervals					
	20 seconds	1E-06	1E-06(ac)	0	1E-06	1E-06 (ac)
Estimators based on Spot Volatility Models						
LLnormal	0.2846 (0.2835)	0.5787 (0.4653)	0.2620 (0.2559)	0.2943 (0.2926)	1.1706 (0.9159)	0.2599 (0.2393)
LLonesided	0.1903* (0.1905)	0.6130 (0.4915)	0.2246* (0.2151)	0.2052 (0.2053)	1.1674 (0.9151)	0.2380* (0.2156)
LLepanech.	0.2791 (0.2794)	0.5960 (0.4764)	0.2677 (0.2594)	0.2718 (0.2714)	1.1596 (0.9103)	0.2582 (0.2369)
asym. Beta	0.7099 (0.5775)	0.5732 (0.4399)	0.6072 (0.1980)	0.7331 (0.6979)	0.7383 (0.7288)	0.5708 (0.2150)
TSRV _{spot}	0.6599 (0.2243)	0.6636 (0.1957)	0.6502 (0.2690)	0.6436 (0.2407)	0.6543 (0.1931)	0.6189 (0.2660)
PreAv _{spot}	0.6746 (0.2333)	0.6843 (0.1651)	0.6614 (0.2562)	0.6570 (0.2399)	0.6785 (0.1583)	0.6206 (0.2482)
RK _{spot}	0.3358 (0.3341)	0.5086* (0.4322)	0.4640 (0.4477)	0.3190 (0.3153)	0.7777 (0.6137)	0.4773 (0.4410)
ACD _{spot}	0.9465 (0.8188)	1.0849 (0.9212)	1.2597 (1.0647)	0.9011 (0.8267)	0.8920 (0.8175)	1.0026 (0.8463)
DUR _{spot}	0.5693 (0.3323)	0.5245 (0.3135)	0.5739 (0.3490)	0.5298 (0.2818)	0.5199 (0.3229)	0.5380 (0.3117)
Estimators from subsection 4.2.1						
PreAv	0.9866 (0.0107)	0.9687 (0.0178)	0.9881 (0.0119)	0.9654 (0.0185)	0.9434 (0.0214)	0.9657 (0.0197)
TSRV	0.6041 (0.5548)	0.6221 (0.5964)	0.5994 (0.5675)	0.4880 (0.4515)	0.4777 (0.4432)	0.4807 (0.4460)
RK	0.5334	0.5666	0.5533	0.4230	0.4328* (0.4236)	0.4336 (0.4507)

4.3 Comparing Intraday Volatility Estimators in a Monte Carlo Simulation Study

	(0.5146)	(0.5671)	(0.5439)	(0.4017)	(0.4250)	(0.4218)	(0.4013)	(0.2704)	(0.3792)
RV ^K	0.6976	0.8664	0.7500	0.5869	0.5760	0.5888	0.6067	0.6231	0.6368
	(0.6941)	(0.8665)	(0.7498)	(0.5841)	(0.5732)	(0.5830)	(0.5699)	(0.5993)	(0.6146)
RV	0.3454	0.8157	0.3892	0.1816*	1.1419	0.2867	0.0834*	2.7108	0.3310
	(0.3397)	(0.6488)	(0.3794)	(0.1811)	(0.8674)	(0.2570)	(0.0835)	(2.1806)	(0.2823)
Real. Range	0.7021	0.7462	0.6496	0.6323	0.7392	0.6252	0.6025	0.6202	0.6056
	(0.6964)	(0.7431)	(0.6455)	(0.6291)	(0.7392)	(0.6204)	(0.5977)	(0.6203)	(0.6045)
Sq. Return	1.9010	1.9103	1.6048	1.5589	1.7373	1.4115	1.2837	2.9101	1.1912
	(1.8816)	(1.8575)	(1.5864)	(1.5496)	(1.7266)	(1.4104)	(1.0007)	(2.7149)	(0.9989)

Table 4.4: RMSE for 20 second, 1 minute and 5 minute (integrated) volatility estimates. Results are based on 1,000 simulated trajectories of the 2 factor stochastic volatility model with alternative noise specifications. Notation as above.
Lowest RMSE highlighted in each column

Fourth, similar to the differences due to the underlying true volatility processes the RMSE performance of estimators depends on the type of market microstructure noise added. Although the local linear estimator minimizes the RMSE throughout the 1-factor setup the difference to competing noise-robust estimators becomes smaller as microstructure noise is added to the simulation. Moreover, the noise-robust estimators outperform the Kristensen [2010] estimators when the noise variance is increased considerably to e.g. $\omega^2 = 0.0001$ (results not shown). Differences between estimators due to the noise specification are more visible in the 2-factor model with i.i.d. noise added to the simulation. In line with the theory we observe that in this case noise-robust estimators perform best in terms of the RMSE. However, we find for both the 20 second and 1 minute case of the 2-factor simulation study that the Kristensen [2010] local linear estimator has the lowest RMSE when we add autocorrelated market microstructure noise to the simulated processes. This is particularly surprising for the 1 minute intervals, where the local linear estimator does not provide favorable results in the case of i.i.d. microstructure noise. We interpret this finding as direct consequence of the induced autocorrelation in the log-price series contaminated by autocorrelated noise. The autocorrelation is picked up by the nonparametric estimators which are essentially weighted averages of past observations and thus fit the time series structure better.

Several robustness checks underscore the validity of the results. First, the results remain qualitatively stable under optimal bandwidths multiplied by two as well as divided by two. Similar conclusions are obtained under parameters chosen according to the theoretically correct values. Second, we find that noise variances larger than the empirical estimates do not change the conclusions for the 2-factor model. The appendix contains the results for a 2-factor model augmented by microstructure noise with variance $\omega^2 = 0.00001$.

4.4 Empirical comparison of intraday volatility measures

In the context of stochastic volatility models the results of the simulation study show that integrated volatility estimates based on spot volatility models are superior to standard approaches in case of small (20 second) intraday intervals. However, due to the latent nature of volatility it is not clear whether these results hold for real stock market price data. To address this issue we present an empirical study of intraday return distributions where competing volatility estimators are included in realized GARCH specifications. The evaluation of the corresponding tail fit of return distributions allows us to assess indirectly whether the volatility measures from subsection 4.2 contain additional information beyond standard volatility measures like squared returns.

4.4.1 Data and model calibration

Our empirical analysis is based on four NYSE and NASDAQ-traded stocks with ticker symbols AAPL.OQ, KFT.N, NYT.N and XOM.M. We use level one trade and quote data for 01/01/2008 - 10/01/2008. The data-cleaning procedures follow the lines of Brownlees and Gallo [2006], Barndorff-Nielsen et al. [2008b] and Hautsch and Podolskij [2010].

4.4 Empirical comparison of intraday volatility measures

	AAPL.OQ	KFT.N	NYT.N	XOM.N
Bandwidths h in minutes				
LLnormal	7.19	8.34	7.13	6.58
LLonesided	10.46	11.59	10.56	10.28
LLepanech.	7.18	6.92	6.74	6.74
Asym. Beta	6.82	7.06	6.97	7.41
TSRV _{spot}	8.27	9.49	9.65	9.68
PreAv _{spot}	6.81	8.32	8.66	7.96
RK _{spot}	6.83	7.00	7.03	7.29
additional Parameters				
H	72.51	89.24	90.61	73.31
K	17.17	20.88	20.35	17.03

Table 4.5: Average empirical bandwidth selection based on the grid search procedure and the cross-validation criterion. Notation for the spot volatility estimators as before. K and H denote the estimates for both K and K_{spot} as well as for both H and H_{spot} , respectively.

We aggregate the data to second frequency and compute log mid-quotes $\log(MQ_t) = \log((B_t + A_t)/2)$ as log prices in our framework.⁸ The mid-quote prices minimize the effect of the transaction price bounce between bid-quotes B_t and ask quotes A_t and thus reduce the microstructure contamination of the prices.

The empirical analysis focuses on volatility measures for 20 second, 1 minute and 5 minute intervals of the trading day. We do not construct volatility measures for the first 30 minutes of the trading day such that there are data before the first interval boundary that can be used by the spot volatility estimators. Similar to the simulation study we select bandwidths in a grid-search procedure. As bandwidths for the spot volatility estimators we select $h \in \mathcal{G}^{bw} := \{0.03, 0.08, 0.13, \dots, 0.73\}$ that minimize the cross validation loss function (4.30). Note that the Kristensen [2010] framework does not include market microstructure noise. However, given the low noise levels (see table 4.1) we argue that the results approximately hold in the empirical setup, too. To keep the computational burden low we select optimal bandwidths only twice a day and conduct the grid-search procedure based on the half-hour interval 10:00-10:30 (for the first half of the trading day) and on the interval 14:00-14:30 for the second half of the trading day. Table 4.5 gives the average bandwidths for the period 01/01/2008-10/01/2008. We note that the bandwidths differ slightly from the optimal bandwidths of the simulation study.

In case of 5 minute intervals, the parameters K , K_{spot} , H and H_{spot} are optimally chosen according to (4.13) and (4.17), respectively. Similar to the bandwidth selection

⁸In contrast to Oomen [2006] we do not focus on the differences of estimators under different price sampling schemes and employ the so-called calendar time sampling with a fixed one-second frequency.

4 Modelling Volatility Measures on Intra-Daily Time Intervals

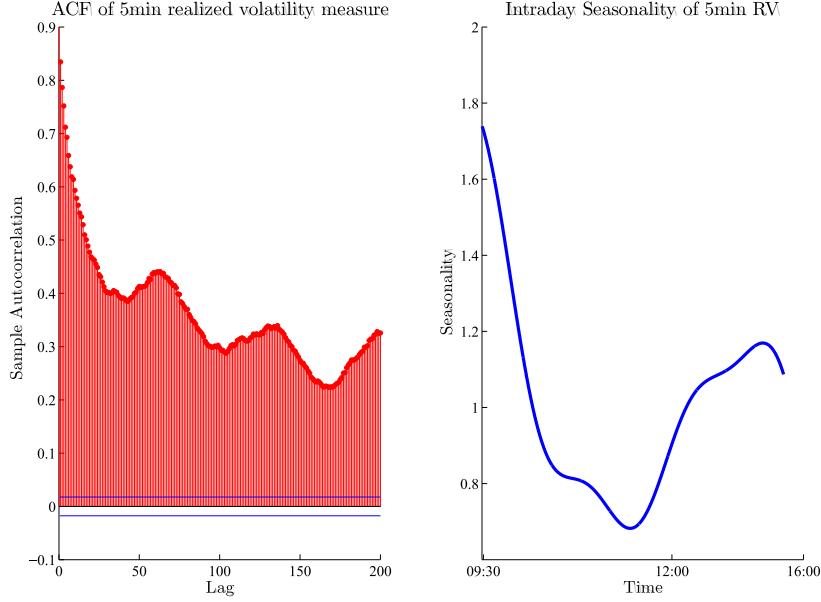


Figure 4.4: Sample autocorrelations and intraday seasonality of the standard realized volatility measure based on 5 minute intervals. Stock: AAPL.OQ for 02/2008-10/2008.

procedure, the values are estimated twice a day and used for the respective first and second half of a day, respectively. However, we find that the 20 second and 1 minute intervals are too small to compute the optimal estimates based on (4.13) and (4.17). We take a practical stance to this problem and set $H = H_{spot} = K = K_{spot} = 5$ for 20 second intervals and $H = H_{spot} = K = K_{spot} = 10$ for 1 minute intervals. Moreover, to minimize problems in estimating the ACD models we set δ^{acd} equal to the average spread divided by ten. Results for the duration-based volatility models are based on δ^{dur} equal to the average spread divided by two.

4.4.2 Empirical Properties of Intraday Volatility Measures

The volatility measures derived from spot volatility estimators and standard approaches from subsection 4.2.1 exhibit the typical U-shape volatility pattern as well as a pronounced autocorrelation. Figure 4.4 shows the seasonality and sample autocorrelation of the 5 minute realized volatility (*RV*) measure. The particular shape of the sample autocorrelation varies across the alternative measures. See the appendix for additional descriptive statistics and a comparison of sample autocorrelation functions for alternative volatility measures.

Notably, we find that the sample autocorrelations of the volatility measures are widely driven by the daily and intra-daily component. This finding is supported by the works of Engle et al. [2006] for squared returns as well as Brownlees et al. [2011] for intraday trad-

4.4 Empirical comparison of intraday volatility measures

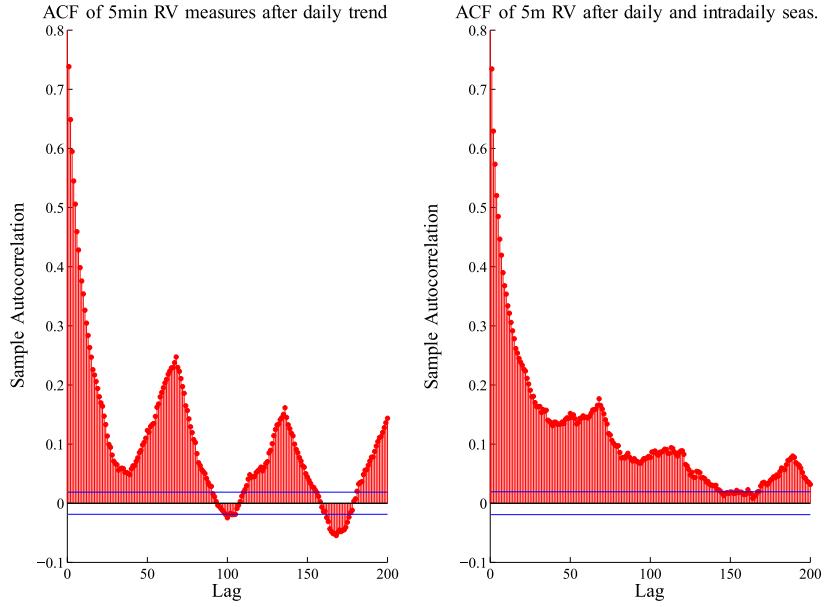


Figure 4.5: Sample autocorrelations of the realized volatility measure after removing the daily trend (left figure) as well as both daily trend and intra-daily seasonality (right figure)

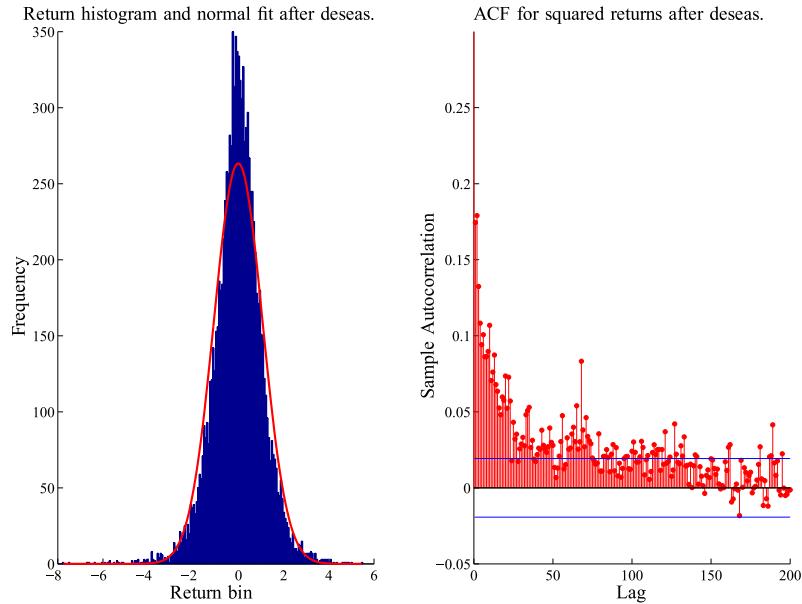


Figure 4.6: Histogram and fitted normal distribution for the 5 minute returns after daily and intra-daily trend cleaning (left) and sample autocorrelations of the corresponding squared returns (right)

ing volumes. To remove the daily and intra-daily seasonality we employ a rolling window scheme with 21 day estimation horizon to estimate the daily and intraday component of the volatility measures for each day. First, we estimate an exponential ACD(1,1) model for the daily (standard) RV and standardize the volatility measures by the estimated daily volatiliy trend. Second, we estimate the seasonality U-shape according to the flexible Fourier form, see equation (4.46), and standardize the volatility measures by the intraday seasonality trend. Figure 4.5 shows the sample autocorrelations of the 5 minute RV measure after intra-daily and daily de-trending, respectively.

For all volatility measures we observe a drastic reduction in the autocorrelations after removing daily and intra-daily trends.⁹ Similar to daily volatility measures, the sample autocorrelations are particularly strong for the standard integrated volatility and spot volatility based measures and considerably lower for the squared return over the 20 second, 1 minute or 5 minute intervals. Figure 4.6 shows the sample autocorrelations for the 5 minute squared return as volatility measure. The left panel in figure 4.6 shows the histogram of the detrended 5 minute returns and a fitted normal distribution. We observe overkurtosis in the data in line with findings for daily volatility measures.

4.4.3 A model for intraday returns and volatility

Motivated by the descriptive analysis of the volatility measures we propose a return model with a multiplicative structure for the conditional variance of returns. In spirit of the high frequency multiple component GARCH by Engle et al. [2006] we propose

$$r_{t,i} = \sqrt{D_t S_{t,i} \sigma_{t,i}^2} \varepsilon_{t,i}, \quad (4.53)$$

where $r_{t,i}$ is the mid-quote return of day t and intraday interval i .¹⁰ The innovation $\varepsilon_{t,i}$ is assumed to be i.i.d. with expectation zero and unit variance. The conditional variance of the return is thus a product of

- D_t , the daily component,
- $S_{t,i}$, the intra-daily seasonality component on day t , and
- $\sigma_{t,i}^2$, the intra-daily dynamic component.

As before, D_t is based on an ACD model for daily RV^K measures with $K = 300$ corresponding to 5 minute subsampling, i.e. we specify

$$RV_t^K = D_t \varepsilon_t^D, \quad (4.54)$$

with conditional mean specification $D_t = \omega_1 + \phi_1 D_{t-1} + \xi_1 RV_{t-1}^K$. The innovation ε_t^D is assumed to be i.i.d. exponentially distributed. Estimation of the seasonality component $S_{t,i}$ is based on the flexible Fourier form (4.46).

⁹Unlike for daily realized volatility measures we cannot confirm a long memory behavior of the time series after de-trending. Daily RV measures are typically thought of as long memory processes. See e.g. Andersen et al. [2003] and Corsi [2009].

¹⁰See also Brownlees et al. [2011] for a component MEM for intra-daily trading volumes.

4.4 Empirical comparison of intraday volatility measures

		Typical Parameter Estimates for AAPL.OQ					
		20 sec		1 min		5 min	
		Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.
<hr/>							
LLnormal	\hat{c}	0.35	(0.09)	0.22	(0.07)	0.16	(0.10)
	$\hat{\alpha}$	0.34	(0.09)	0.24	(0.08)	0.38	(0.16)
	$\hat{\beta}$	0.33	(0.13)	0.54	(0.11)	0.50	(0.21)
RV ^K	\hat{c}	0.42	(0.64)	0.02	(0.01)	0.03	(0.02)
	$\hat{\alpha}$	0.39	(0.67)	0.09	(0.03)	0.22	(0.04)
	$\hat{\beta}$	0.29	(0.12)	0.89	(0.04)	0.77	(0.05)
Sq. Ret.	\hat{c}	0.02	(0.01)	0.01	(0.01)	0.05	(0.02)
	$\hat{\alpha}$	0.04	(0.01)	0.05	(0.01)	0.11	(0.02)
	$\hat{\beta}$	0.93	(0.02)	0.94	(0.02)	0.85	(0.03)

Table 4.6: Average parameter and robust standard error estimates for 4 to 50 estimations of two normal RV-Garch specifications based on the local linear measure and the sparsely sampled realized variance as well as for the GARCH specification based on the normal distribution (squared return specification).

The component $\sigma_{t,i}^2$ is defined as the realized GARCH conditional mean process (see Hansen et al. [2011])

$$\sigma_{t,i}^2 = \omega + \sum_{j=1}^p \alpha_j V_{t,i-j} + \sum_{j=1}^q \sigma_{t,i-j}^2, \quad (4.55)$$

where $V_{t,i-j}$ is a volatility measure for day t and interval $i-j$ after daily and intra-daily de-trending.¹¹

We estimate the multiple component realized GARCH by first deriving detrended returns $\tilde{r}_{t,i} = r_{t,i}/\sqrt{\hat{D}_t \hat{S}_{t,i}}$ and volatility measures $\tilde{V}_{t,i} = V_{t,i}/(\hat{D}_t \hat{S}_{t,i})$, where the hats denote estimates, respectively. In the second step we maximize the likelihood based on a normal or student t-distribution assumption for the error $\varepsilon_{t,i}$.

We apply the outlined model to volatility measures based on the price data for the stocks AAPL.OQ, KFT.N, NYT.N and XOM.N during 01/01/2008 to 10/01/2008. The data of the 20 second, 1 minute and 5 minute aggregates is partitioned to blocks of roughly 3000 observations to reduce the computational burden and allow for parameter changes. The parameters of the multiple component RV GARCH models are calibrated on these subsamples. An estimation horizon of 21 days is used to estimate the daily trend as well as the intraday seasonality for each subsample.

¹¹See also Shephard and Sheppard [2010] for GARCH models augmented by realized volatility measures.

Table 4.6 gives the average estimates and t-statistics of the dynamic component $\sigma_{t,i}^2$ of the multiple component RV-GARCH based on the subsamples for AAPL.OQ. Compared to the standard multiple component GARCH (see the rows denoted Sq. Ret.) we typically observe a higher estimate for α when employing a RV GARCH with volatility measures based on integrated spot volatilities or realized measures of volatility. We interpret this finding as indication of the additional information contained in the RV-type measures over squared returns.

4.4.4 The tail probabilities of intraday returns

With intraday (trading) risk management applications in mind we assess the properties of the multiple component RV GARCH models in terms of the density and tail density fit.¹² Extreme returns are typically the main focus of quantitative risk management. Correspondingly, our main interest lies in the $\alpha = 10\%$ quantile q_t^α of the return distribution implied by the multiple component RV-GARCH model,

$$q_t^\alpha = F_t^{-1}(\alpha), \quad (4.56)$$

where F_t denotes the implied cdf of the return distribution at t .

We are particularly interested in the model's ability to match the empirically observed extreme event occurrence. Hence, we report an estimate of the conditional coverage probability $\hat{\alpha} = \mathbb{E}[\mathbb{1}\{r_t < q_t^\alpha\} | \mathcal{F}_{t-1}]$ based on its sample counterpart

$$\hat{\alpha} = \frac{1}{T} \sum_{t=1}^T \mathbb{1}\{r_t < \widehat{q}_t^\alpha\}, \quad (4.57)$$

where \widehat{q}_t^α is the quantile implied by the fitted RV-GARCH model at time t .

Kupiec [1995] proposes a formal test of the adequacy of the coverage probability which is based on the indicators $I_t^\alpha := \mathbb{1}\{r_t < \widehat{q}_t^\alpha\}$, $t = 1, \dots, T$. The failures $x := \sum_{t=1}^T I_t^\alpha$ follow a binomial distribution with parameters T and α . The likelihood ratio test statistic of $H_0 : \mathbb{E}[I_t^\alpha] = \alpha$ is given by

$$LR^{Kupiec} = 2 \log \left[\left(1 - \frac{x}{T}\right)^{T-x} \left(\frac{x}{T}\right)^x \right] - 2 \log[(1 - \alpha)^{T-x} \alpha^x]. \quad (4.58)$$

LR^{Kupiec} is χ^2 -distributed with 1 degree of freedom.¹³

Moreover, we consider the "check" loss function Q of quantile regression models for the evaluation of the tail fit. Koenker and Bassett [1978] define the function by $Q(\alpha) =$

¹²Related studies in this respect are Gerig et al. [2009] who examine the unconditional distribution of returns at different frequencies as well as the intraday risk framework outlined in Giot [2005], Dionne et al. [2009] and Colletaz et al. [2007]. A related comparison study of risk measures implied by different realized variance based estimates on a daily basis is given by Brownlees and Gallo [2010].

¹³Note that the Kupiec [1995] test is based on the null that the unconditional coverage probability is correct. Tests of the conditional coverage probability are given by e.g. Christoffersen [1998]. See also the dynamic quantile test by Engle and Manganelli [2004].

4.4 Empirical comparison of intraday volatility measures

$\mathbb{E}[(\alpha - \mathbb{1}\{r_t < q_t^\alpha\})(r_t - q_t^\alpha)]$ with sample counterpart

$$\widehat{Q}(\alpha) = \frac{1}{T} \sum_{t=1}^T (\alpha - \mathbb{1}\{r_t < \widehat{q}_t^\alpha\})(r_t - \widehat{q}_t^\alpha). \quad (4.59)$$

The model that minimizes the loss function \widehat{Q} is preferred.¹⁴

Alternative means of assessing the model fit and in particular the distributional fit are given by the probability integral transforms and the corresponding theory put forward by Rosenblatt [1952]. Let $F_t(\cdot)$ denote the (forecasted) cdf at time t . Under the null hypothesis of correct model specification the transformed observations U_t ,

$$U_t := F_t(r_t) = \int_{-\infty}^{r_t} f_t(u) du, \quad (4.60)$$

are i.i.d. uniform. Berkowitz [2001] proposes to transform the U_t into normally distributed observations Z_t based on $Z_t := \Phi^{-1}(U_t) = \Phi^{-1}(F_t(r_t))$. Under correct model specification the Z_t are i.i.d. standard normal. The corresponding test is based on the conditional mean specification $Z_t - \mu = \rho(Z_{t-1} - \mu) + \varepsilon_t$. Under normally distributed ε_t , the null $H_0 : \mu = 0, \rho = 0, \sigma^2 = 1$ can be tested with the likelihood ratio test statistic LR_{full} ,

$$LR_{full} := -2(L(0, 1, 0) - L(\widehat{\mu}, \widehat{\sigma}^2, \widehat{\rho})), \quad (4.61)$$

where the log-likelihood function is given by

$$\begin{aligned} L(\mu, \sigma, \rho) = & -\frac{1}{2} \log(2\pi) - \frac{1}{2} \left[\frac{\sigma^2}{(1-\rho^2)} \right] - \frac{(Z_1 - \mu/(1-\rho))^2}{2\sigma^2/(1-\rho^2)} - \frac{T-1}{2} \log(2\pi) \\ & - \frac{T-1}{2} \log(\sigma^2) - \sum_{t=2}^T \left(\frac{(Z_t - \mu - \rho Z_{t-1})^2}{2\sigma^2} \right). \end{aligned} \quad (4.62)$$

Under the null, LR_{full} is χ^2 -distributed with 3 degrees of freedom. Alternatively, one can test for the distributional fit in the tails of the return distribution. To detect misspecification in the $\alpha = 10\%$ tail, Berkowitz [2001] proposes the transformation

$$Z_t^* = \begin{cases} q^\alpha, & \text{if } Z_t \geq q^\alpha \\ Z_t, & \text{if } Z_t < q^\alpha, \end{cases} \quad (4.63)$$

where $q^\alpha := \Phi^{-1}(\alpha)$. A likelihood ratio test of misspecification in the first two moments of the tail is given as $LR_{tail} := -2(L(0, 1) - L(\widehat{\mu}, \widehat{\sigma}^2))$, with log-likelihood function

$$L(\mu, \sigma | Z^*) = \sum_{t: Z_t^* < q^\alpha} \log \left(\frac{1}{\sigma} \phi \left(\frac{Z_t - \mu}{\sigma} \right) \right) + \sum_{t: Z_t^* = q^\alpha} \log \left(1 - \Phi \left(\frac{q^\alpha - \mu}{\sigma} \right) \right). \quad (4.64)$$

Under the null $H_0 : \mu = 0, \sigma^2 = 1$ the test statistic is χ^2 -distributed with 2 degrees of

¹⁴See in this context Bao et al. [2006] on the implications for VaR forecast evaluation.

Real Data Results for AAPL.OQ: Normal-RV-GARCH model fit for 20 sec, 1 min and 5 min intervals

	20 seconds				1 minute				5 minute			
	CCP [CCP _{loss}]	D (D _{tail}) [Kupiec]	LL Value [LB ₁₀]	CCP [CCP _{loss}]	D (D _{tail}) [Kupiec]	LL Value [LB ₁₀]	CCP [CCP _{loss}]	D (D _{tail}) [Kupiec]	CCP [CCP _{loss}]	D (D _{tail}) [Kupiec]	LL Value [LB ₁₀]	
Estimators based on spot volatility models												
LLnormal	0.0897 [0.1803]	1 (0) [0.3743]	-226271 [0.2616]	0.0899 [0.1804]	1 (0) [0.3564]	-75982 [0.2383]	0.0941 [0.1847]	1 (0) [0.3205]	1 (0) [0.3205]	1 (0) [0.2788]	-15350 [0.2788]	
LLonesided	0.0896 [0.1804]	1 (0) [0.3648]	-227284 [0.2476]	0.0896 [0.1799*]	1 (0) [0.3415]	-76200 [0.2024]	0.0921 [0.1844]	1 (0) [0.3244]	1 (0) [0.3244]	1 (0) [0.2753]	-15380 [0.2753]	
LLepanech.	0.0892 [0.1804]	1 (0) [0.3554]	-227086 [0.2427]	0.0906 [0.1800]	1 (0) [0.3690]	-76125 [0.2177]	0.0936 [0.1858]	1 (0) [0.3293]	1 (0) [0.3293]	1 (0) [0.2603]	-15398 [0.2603]	
asym. beta	0.0905 [0.1805]	1 (0) [0.3906]	-225836 [0.2591]	0.0908 [0.1813]	1 (0) [0.3507]	-75767 [0.2552]	0.0930 [0.1848]	1 (0) [0.3105]	1 (0) [0.3105]	1 (0) [0.3123]	-15307 [0.3123]	
TSRV _{spot}	0.0904 [0.1803]	1 (0) [0.4276]	-225458 [0.2852]	0.0917 [0.1817]	1 (0) [0.4141]	-75498 [0.2878]	0.0924 [0.1830*]	1 (0) [0.3453]	1 (0) [0.3453]	1 (0) [0.3254]	-15236. [0.3254]	
PreAv _{spot}	0.0887 [0.1810]	1 (0) [0.2936]	-229569 [0.2310]	0.0889 [0.1826]	1 (0) [0.2477]	-77570 [0.2139]	0.0922 [0.1845]	1 (0) [0.3384]	1 (0) [0.3384]	1 (0) [0.3032]	-15604 [0.3032]	
RK _{spot}	0.0904 [0.1800*]	1 (0) [0.4382]	-224750 [0.2928]	0.0918 [0.1812]	1 (0) [0.3903]	-75257 [0.3264]	0.0945 [0.1832]	1 (0) [0.3648]	1 (0) [0.3648]	1 (0) [0.3790]	-15196 [0.3790]	
ACD _{spot}	0.0912 [0.1817]	1 (0) [0.3585]	-226440 [0.2659]	0.0911 [0.1826]	1 (0) [0.3222]	-76161 [0.2488]	0.0953 [0.1841]	1 (0) [0.2978]	1 (0) [0.2978]	1 (0) [0.3347]	-15207 [0.3347]	
DUR _{spot}	0.0898 [0.1808]	1 (0) [0.3485]	-226591 [0.2589]	0.0910 [0.1821]	1 (0) [0.3652]	-75996 [0.2523]	0.0940 [0.1846]	1 (0) [0.3347]	1 (0) [0.3347]	1 (0) [0.3046]	-15294 [0.3046]	
PreAv	0.0909	1 (0)	-225360	0.0919	1 (0)	-75464	0.0971*	1 (0)	1 (0)	1 (0)	-15183	

4.4 Empirical comparison of intraday volatility measures

	[0.1803]	[0.3933]	[0.2796]	[0.1805]	[0.4244]	[0.2738]	[0.1841]	[0.3912]	[0.3495]
TSRV	0.0908	1 (0) [0.4211]	-224918 [0.2944]	0.0921 [0.1814]	1 (0) [0.4339]	-75299 [0.3059]	0.0965 [0.1839]	1 (0) [0.3809]	-15166* [0.3594]
	0.0904	1 (0) [0.4079]	-225116 [0.2901]	0.0919 [0.1809]	1 (0) [0.4204]	-75390 [0.2921]	0.0960 [0.1837]	1 (0) [0.4028]	-15177 [0.3925]
RV ^K	0.0905	1 (0) [0.1802]	-224907 [0.4177]	0.0916 [0.2967]	1 (0) [0.1821]	-75336 [0.3166]	0.0960 [0.1846]	1 (0) [0.3731]	-15225 [0.3839]
	0.0908	1 (0) [0.1804]	-225319 [0.3957]	0.0913 [0.2794]	1 (0) [0.1816]	-75619 [0.2753]	0.0931 [0.1847]	1 (0) [0.3177]	-15269 [0.3468]
RV	0.0908	1 (0) [0.1802]	-224479* [0.4334]	0.0920 [0.3108]	1 (0) [0.1823]	-75142* [0.4570]	0.0951 [0.3574]	1 (0) [0.1834]	-15176 [0.4432]
	0.0910*	1 (0) [0.1805]	-224843 [0.4483]	0.0923* [0.3134]	1 (0) [0.1825]	-75434 [0.4625]	0.0930 [0.3453]	1 (0) [0.1834]	-15304 [0.3656]
Sq. Return									

Table 4.7: Diagnostics for the overall fit and 10% - tail risk fit of the Normal-RV-GARCH based on 1 min, 5 min and 20 second (integrated) volatility measures. Notation as above: CCP denotes the Conditional Coverage probability estimate, CCP_{loss} the corresponding loss function. D and D_{tail} denote the P-values for the density forecast backtest and the density forecast backtest for the tail, respectively. LL Value and LB_{10} give the log-likelihood value and the Ljung-Box test for ten lags. Sq. Return and R. Range denote the squared return model (=standard GARCH) and the RV-GARCH based on the realized range. Optimal values highlighted in each column. Data ranges from 01/20/2008 to 10/01/2008

4 Modelling Volatility Measures on Intra-Daily Time Intervals

	KFT.N	NYT.N	XOM.N
20 sec	CCP	DUR _{spot}	ACD _{spot}
	CCP _{loss}	LLnormal	RK _{spot}
1 min	CCP	R. Range	Sq. Return
	CCP _{loss}	RK _{spot}	RV ^K
5 min	CCP	ACD _{spot}	RV
	CCP _{loss}	RV	RK _{Diff}
		R. Range	Sq. Return

Table 4.8: The table shows optimal volatility measures for the CCP and CCP_{loss} criterium for $\alpha = 0.1$. Optimal values given are for the stocks KFT.N, NYT.N and XOM.N on 20s, 1 min and 5 min frequency. Notation as above.

freedom.

Table 4.7 gives the model diagnostics for the 20 second, 1 minute and 5 minute volatility measures in the case of AAPL.OQ. We report results for $\tilde{r}_{t,i} \sim N(0, \sigma_{t,i}^2)$ in the following since the tail fit for AAPL.OQ is found to be superior to the t-distribution. Corresponding results of the empirical analysis based on the t-distribution for the case AAPL.OQ can be found in the appendix.

First, in the 20 second and 1 minute case the conditional coverage probability (CCP) is closest to the nominal value 10% when we employ standard multiple component GARCH models (case Sq. Return) for the intraday component instead of RV GARCH models. The reason for the strong CCP performance of the standard GARCH models on 20 second and 1 minute aggregates is likely due to the low number of observations such that potential information gains from realized measures are small compared to the information already contained in the squared returns. In the 5 minute case the RV-GARCH models based on the pre-averaging estimator are optimal. Moreover, the standard GARCH model on 5 minute returns performs worse than most spot volatility based estimators. For all frequencies we find that the Kupiec [1995] test confirms the adequacy of the $\hat{\alpha}$ estimates implied by the forecasted quantiles. The null hypothesis is accepted on conventional levels.

Second, the spot volatility models minimize the loss function (4.59). However, mutually testing the difference of the loss function series with lowest mean to the loss series of alternative estimators using t-stats we find that the difference in the losses is not statistically significant.

Third, even though the normal distribution is not rejected by the Berkowitz [2001] test (4.61), the tail distribution is misspecified since the tail density forecast backtest rejects the null of correct model specification at all levels.

Fourth, in addition to the density forecast backtest (4.61) the multiple component RV GARCH is supported by the Ljung-Box test. We do not find correlation up to lag 10 in the squared return remainder $\tilde{r}_{t,i} := r_{t,i}/\sqrt{D_t S_{t,i} \sigma_{t,i}^2}$.

However, the results reported for AAPL.OQ do not necessarily hold for other stocks.

Table 4.8 shows the volatility measures for KFT.N, NYT.N and XOM.N that optimize the conditional coverage probability CCP and the tail fit loss function CCP_{loss} , respectively. Confirming the findings for AAPL.OQ for the 20s and 1 minute aggregates we find that the spot volatility based measures minimize the CCP_{loss} criterion. The integrated volatility measures based on subsection 4.2.1 predominantly optimize the CCP at the 1 and 5 minute aggregation level. The CCP is optimized by measures based on spot volatility estimates for 2 out of 3 stocks on 20s aggregates. Supporting the simulation study in section 4.3 this finding indicates that for some stocks it is worthwhile to include information from outside the 20s intervals for the volatility estimation.

4.5 Conclusions

The estimation and modelling of volatility is crucial to many financial applications like stock trading strategies and risk management. However, the latest surge in intraday trading activity requires market participants to measure volatility on small intra-daily intervals whereas the predominant fraction of the literature typically focuses on daily volatility measurement. Motivated by this imbalance, our study is the first to systematically compare alternative approaches to volatility measurement for small intra-daily intervals.

We construct volatility measures from two different strands of literature. First, we consider realized measures of volatility that have been successfully applied to daily volatility measurement. Second, we numerically integrate spot volatility estimators on small intervals. The spot volatility estimates are based on non-parametric and parametric methods. The idea of integrating spot volatility measures is to include information from the outside of the very small intervals of interest in order to compensate for the small sample problem when using only intra-interval observations.

A simulation study based on stochastic volatility models augmented by market microstructure noise of different magnitudes confirms the usefulness of the volatility estimators and reveals that volatility measures based on spot volatility estimates are favorable in the case of 20 second intervals.

To assess the properties of the estimators on stock market data we analyze the tail probabilities of 20 second, 1 minute and 5 minute return distributions implied by realized GARCH models which are based on numerically integrated spot volatility or standard realized volatility estimators. To account for daily and intra-daily trends as well as the autocorrelation properties of the volatility measures we propose a multiple component structure of the realized GARCH models. Analyzing the implied tail probabilities we find that both realized measures of volatility and integrated spot volatilities give more exact quantile estimates for the return distribution than only using the squared returns. The individual performance of the estimators depends on the particular stock under consideration.

C Appendix

C.1 Finite sample adjustment of the pre-averaging estimator

In order to apply the pre-averaging method to intraday intervals we implement the finite sample adjustments for PreAv_t proposed by Hautsch and Podolskij [2010],

$$\text{PreAv}_t = \left(1 - \frac{\psi_1^{k_n} \Delta}{2\theta^2 \psi_2^{k_n}}\right)^{-1} \left(\frac{t/\Delta \sqrt{\Delta}}{(t/\Delta - k_n + 2)\theta \psi_2^{k_n}} V(Z, 2)_t - \frac{\psi_1^{k_n} \Delta}{2\theta^2 \psi_2^{k_n}} RV_t \right), \quad (4.65)$$

with

$$\psi_1^{k_n} = k_n \sum_{j=1}^{k_n} \left(g\left(\frac{j+1}{k_n}\right) - g\left(\frac{j}{k_n}\right) \right)^2 \quad \text{and} \quad \psi_2^{k_n} = \frac{1}{k_n} \sum_{j=1}^{k_n-1} g^2\left(\frac{j}{k_n}\right). \quad (4.66)$$

C.2 Kernel functions

Three kernels are used: (i) the Epanechnikov kernel K_e ,

$$K_e(x) = \begin{cases} 3/4(1-x)^2, & |x| \leq 1, \\ 0, & \text{otherwise.} \end{cases} \quad (4.67)$$

(ii) The one-sided kernel K_o ,

$$K_o(x) = \begin{cases} 6(1+3x+2x^2), & -1 \leq x \leq 0, \\ 0, & \text{otherwise.} \end{cases} \quad (4.68)$$

(iii) The normal kernel K_n ,

$$K_n(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}. \quad (4.69)$$

C.3 Simulation of the SV models with seasonality

Following the appendix to Barndorff-Nielsen et al. [2008a] we simulate log price series consisting of 23,400 observations. Each price observation is thought of to be the prevailing price on an one second interval such that the total observation number spans a 6.5 hour trading day. The time stamps are normalized to $[0, 1]$ resulting in a discretization into sub-intervals of length $\Delta = 1/23,400$.

The simulated log price series of the 1-factor model are based on the iterative Euler scheme according to

$$\begin{aligned} Y_{i+1} &= Y_i + \mu \Delta + \sigma_i \sqrt{\Delta} \varepsilon_i^W, \\ \sigma_i &= S_i (\exp(\beta_0 + \beta_1 \tau_i)), \\ \tau_{i+1} &= \tau_i + \alpha \tau_i \Delta + \sqrt{\Delta} \varepsilon_i^B, \end{aligned} \quad (4.70)$$

where S_i is the intraday trend (4.46) based on parameter estimates for the stock AAPL.OQ in 01/01/2008 to 10/01/2008. The $(\varepsilon_i^W, \varepsilon_i^B)^T$ are drawn according to

$$\begin{pmatrix} \varepsilon_i^W \\ \varepsilon_i^B \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \phi \\ \phi & 1 \end{pmatrix} \right) \quad \text{i.i.d.}$$

The parameter values are as in subsection 4.3. The process is started with $\tau_0 \sim N(0, (-2\alpha)^{-1})$.

The simulated log price series of the 2-factor model (4.47) to (4.49) are based on the following discretization. We use the iterative scheme

$$\begin{aligned} Y_{(i+1)} &= Y_i + \mu\Delta + \sigma_i \sqrt{\Delta} \left(\phi_1 \varepsilon_i^{B_1} + \phi_2 \varepsilon_i^{B_2} + \sqrt{1 - \phi_1^2 - \phi_2^2} \varepsilon_i^W \right), \\ \sigma_i &= S_i(\exp^*(\beta_0 + \beta_1 \tau_{1,i} + \beta_2 \tau_{2,i})), \\ \tau_{1,(i+1)} &= \tau_{1,i} + \alpha_1 \tau_{1,i} \Delta + \sqrt{\Delta} \varepsilon_i^{B_1}, \\ \tau_{2,(i+1)} &= \tau_{2,i} + \alpha_2 \tau_{2,i} \Delta + (1 + \phi \tau_{2,i}) \sqrt{\Delta} \varepsilon_i^{B_2}, \end{aligned} \tag{4.71}$$

where the $(\varepsilon_i^W, \varepsilon_i^{B_1}, \varepsilon_i^{B_2})^T$ are drawn according to

$$(\varepsilon_i^W, \varepsilon_i^{B_1}, \varepsilon_i^{B_2})^T \sim N(\mathbf{0}, \mathbf{I}) \quad \text{i.i.d.}$$

Parameter values as given in subsection 4.3.

C.4 Optimal bandwidth according to the grid-search procedure (2-factor model)

ω^2	no noise	0.000001	0.000001 (ac)
Spot volatility based estimators			
		Bandwidths h in minutes	
LLnormal	3.85	3.85	3.85
LLonesided	3.85	3.85	3.85
LLepanechnikov	3.85	3.85	3.85
asym. Beta	15.40	53.90	32.85
TSRV _{spot}	23.56	15.40	28.23
PreAv _{spot}	26.95	15.40	28.23
RK _{spot}	3.85	7.70	7.70
additional Parameters			
K _{spot}	30.00	65.00	50.00
θ_{spot}	1.50	1.40	1.20
H _{spot}	80.00	35.00	155.00
Estimators from subsection 4.2.1			
K	10.00	30.00	40.00

4 Modelling Volatility Measures on Intra-Daily Time Intervals

H	20.00	20.00	10.00
θ	1.40	0.50	0.80
δ_{acd}	0.005	0.007	0.005
δ_{dur}	0.014	0.014	0.014

Table 4.9: Optimal bandwidths and parameters for the 2-factor model according to the grid-search procedure. Abbreviations are LLnormal, LLonesided, LLepanechnikov for the local linear estimator with normal, one-sided and Epanechnikov kernel. asym Beta denotes the asymmetric beta kernel estimator

C.5 Simulation results for the 2-factor model with microstructure noise of magnitude $\omega^2 = 0.00001$

Simulation Results: RMSE of integrated volatility estimators for 20s,1 and 5 minute intervals						
ω^2	no noise	1E-05	1E-05 (ac)	no noise	1E-05	1E-05 (ac)
20 seconds						
LLnormal	0.2846 (0.2835)	27.586 (23.301)	1.1847 (1.1757)	0.2943 (0.2926)	31.595 (27.889)	2.1294 (2.1319)
LLonesided	0.1903* (0.1905)	33.227 (28.231)	1.1964 (1.1844)	0.2052 (0.2053)	37.372 (32.251)	0.1368 (0.1358)
LLepanech.	0.2791 (0.2794)	24.767 (20.894)	1.3724 (1.3440)	0.2718 (0.2714)	28.755 (24.718)	0.2051 (0.2044)
asym. Beta	0.7099 (0.5775)	12.115 (10.421)	0.6042 (0.2296)	0.7331 (0.6979)	13.743 (12.001)	0.6469 (0.4836)
TSRV _{spot}	0.6599 (0.2243)	1.3309 (1.3171)	0.7570 (0.3581)	0.6436 (0.2407)	1.7348 (1.7201)	0.7594 (0.3630)
PreAv _{spot}	0.6746 (0.2333)	1.2632 (1.2563)	0.8143 (0.1381)	0.6570 (0.2399)	1.7244 (1.7157)	0.7999 (0.2326)
RK _{spot}	0.3358 (0.3341)	17.799 (15.022)	0.5153* (0.4996)	0.3190 (0.3153)	19.903 (17.116)	0.8398 (0.7816)
ACD _{spot}	0.9465 (0.8188)	19.270 (16.774)	0.6346 (0.6350)	0.9011 (0.8267)	22.600 (20.021)	0.6549 (0.6557)
DUR _{spot}	0.5693 (0.3323)	2.5915 (2.4800)	0.5328 (0.3458)	0.5298 (0.2818)	2.0886 (1.9287)	0.5300* (0.4452)
5 minute						
PreAv	0.9866 (0.0107)	0.9224* (0.7295)	0.9720 (0.0119)	0.9654 (0.0185)	0.8408 (0.5166)	0.9319 (0.0568)
TSRV	0.6041	2.0808	0.6722	0.4880	0.8219* 0.4914	1.0158 (0.4914)
1E-05 (ac)						
Estimators based on Spot Volatility Models						
LLnormal	28.341 (24.176)	0.1564 (0.1525)	3.3161 (3.1093)			
LLonesided	32.589 (27.918)	0.1368 (0.1358)	3.7579 (3.5144)			
LLepanech.	25.911 (0.1381)	0.1381 (0.1369)	2.0909 (22.093)			
asym. Beta	11.822 (0.7911)	0.7911 (0.1369)	0.6248 (2.0323)			
TSRV _{spot}	(10.250) (0.7827)	(0.7827) (0.7836)	(0.6115) (10.250)			
PreAv _{spot}	1.2165 (0.2005)	0.6206 (0.2005)	0.6958 (0.3096)			
RK _{spot}	1.1630 (0.2430)	0.2430 (0.2326)	0.7365 (1.1488)			
ACD _{spot}	1.1630 (0.6549)	0.6549 (0.6740)	0.7365 (0.5229)			
DUR _{spot}	1.3750 (0.4931)	0.4931 (0.4452)	0.4914 (0.4741)			
Estimators from subsection 4.2.1						
PreAv	0.7942 (0.1882)	0.9083 (0.0290)	0.7942 (0.1413)			
TSRV	1.1694 (0.3979*)	0.4914 (0.4741)	1.1694 (0.4741)			

RK	(0.5548)	(2.0364)	(0.6752)	(0.4515)	(0.8018)	(0.9974)	(0.4388)
	0.5334	5.4310	0.6249	0.4230	2.2188	1.1553	0.4507
	(0.5146)	(4.9200)	(0.6240)	(0.4017)	(2.0628)	(1.1166)	(0.4013)
RV ^K	0.6976	7.5756	0.9920	0.5869	3.8392	1.2163	0.6067
	(0.6941)	(7.0725)	(0.9781)	(0.5841)	(3.5422)	(1.1842)	(0.5699)
RV	0.3454	39.959	0.5639	0.1816*	42.717	1.7345	0.0834*
	(0.3397)	(34.841)	(0.4873)	(0.1811)	(37.629)	(1.6278)	(0.0835)
Real. Range	0.7021	5.3470	0.7940	0.6323	3.0458	1.2332	0.6025
	(0.6964)	(4.7134)	(0.7899)	(0.6291)	(2.6942)	(1.1987)	(0.5977)
Sq. Return	1.9010	4.0834	1.8407	1.5589	3.3663	2.2721	2.5800
	(1.8816)	(3.9204)	(1.8018)	(1.5496)	(3.3433)	(2.2528)	(0.0835)

Table 4.10: RMSE for 1,5 minute and 20 second (integrated) volatility estimates. Results are based on 1,000 simulated trajectories of the 2-factor stochastic volatility model with alternative noise specifications. Notation as above.
 Lowest RMSE highlighted in each column

C.6 Sample autocorrelations of four volatility measures

The following figures show the sample autocorrelograms for 20s, 1 minute and 5 minute volatility measures for AAPL.OQ. Measures are the nonparametric local linear estimator with normal kernel, the measure based on the two-scale realized variance spot volatility estimates, the volatility measure based on the ACD model and the pre-averaging volatility measure.

4 Modelling Volatility Measures on Intra-Daily Time Intervals

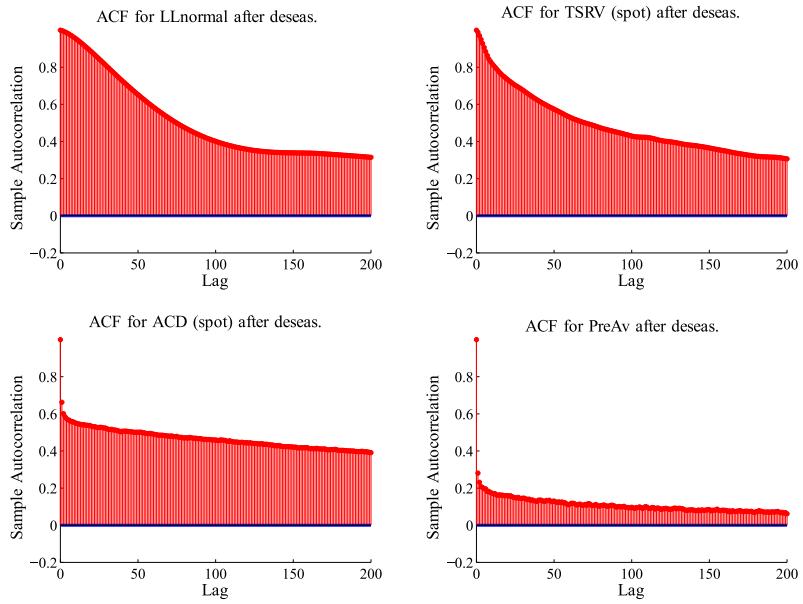


Figure 4.7: Sample autocorrelations of four 20 second volatility measures after removing the daily and intra-daily trends. In the above notation, the ACFs shown are for the measures LLnormal , $\text{TSRV}_{\text{spot}}$, ACD_{spot} and PreAv

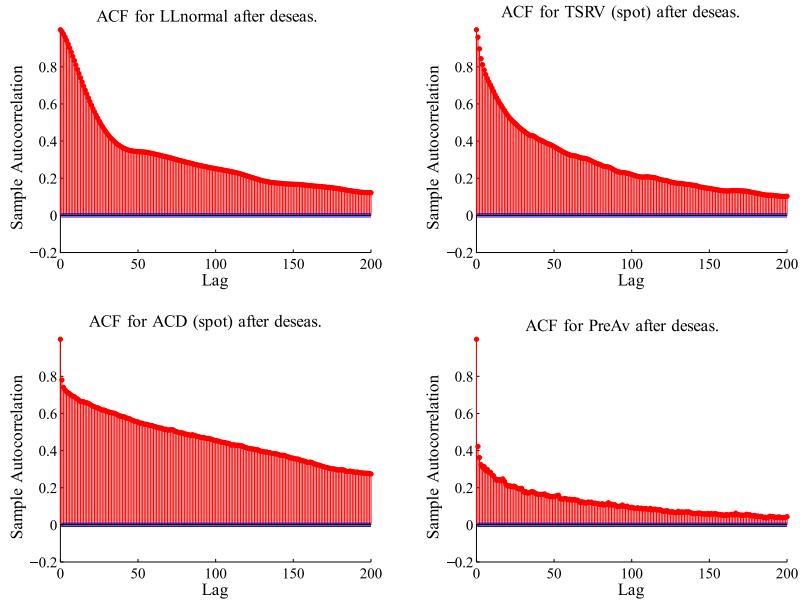


Figure 4.8: Sample autocorrelations of four 1 minute volatility measures after removing the daily and intra-daily trends.

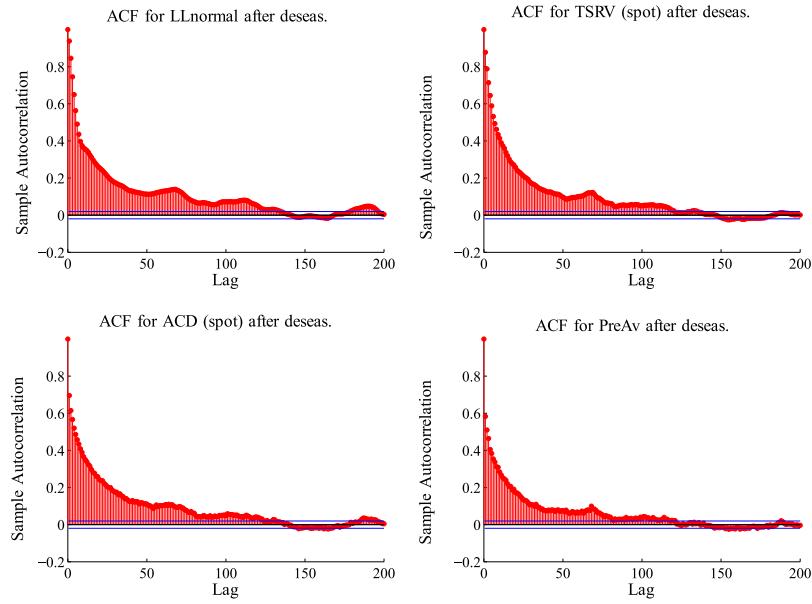


Figure 4.9: Sample autocorrelations of four 5 minute volatility measures after removing the daily and intra-daily trends.

C.7 Basic descriptive statistics for the AAPL.OQ volatility measures

	20 sec		1 min		5 min	
	Mean	Var	Mean	Var	Mean	Var
Estimators based on spot volatility models						
LLnormal	1.04	0.55	1.04	0.54	1.04	0.51
LLonesided	1.04	0.63	1.04	0.60	1.04	0.54
LLepanech.	1.04	0.62	1.04	0.61	1.04	0.54
asym. beta	1.04	0.47	1.04	0.46	1.04	0.41
TSRV _{spot}	1.04	0.64	1.04	0.63	1.04	0.55
PreAv _{spot}	1.14	3.77	1.14	3.75	1.22	5.26
RK _{spot}	1.04	0.61	1.04	0.60	1.05	0.53
ACD _{spot}	1.04	0.35	1.04	0.27	1.04	0.52
DUR _{spot}	1.05	0.40	1.05	0.33	1.05	0.71
Estimators from subsection 4.2.1						
PreAv	1.04	2.54	1.04	1.36	1.04	0.78
TSRV	1.04	2.40	1.04	1.50	1.05	0.81
RK	1.03	2.20	1.04	1.38	1.07	1.81
RV ^K	1.04	2.20	1.04	1.44	1.06	1.32

4 Modelling Volatility Measures on Intra-Daily Time Intervals

RV	1.04	1.59	1.04	0.96	1.04	0.58
R. Range	1.04	2.57	1.05	2.21	1.07	1.93
Sq. Return	1.05	4.98	1.06	5.01	1.10	5.24

Table 4.11: Mean and Variance of the volatility measures of AAPL.OQ in 01/01/2008 to 10/01/2008. Notation as before

C.8 Real Data results: RV Garch based on t-distribution, $\alpha = 10\%$

Real Data Results: t-RV-GARCH model fit for 20 sec, 1 min and 5 min intervals

	CCP [CCP _{loss}]	D [D _{tail}]	LL Value [LB ₁₀]	CCP [CCP _{loss}]	D [D _{tail}]	LL Value [LB ₁₀]	CCP [CCP _{loss}]	D [D _{tail}]	LL Value [LB ₁₀]
Estimators based on Spot Volatility Models									
LLnormal	0.0706 [0.1829]	1 0.0164	-137690 [0.2673]	0.0765 [0.1833]	1 [0]	-55341 [0.2157]	0.0778 [0.1860]	1 [0]	-15116 [0.3128]
LLonesided	[0.1831]	1 [0]	-159091 [0.2470]	0.0763 [0.1842]	1 [0]	-74150 [0.2378]	0.0766 [0.1862]	1 [0]	-15128 [0.3162]
LLepanech.	0.0700 [0.1830]	1 [0]	-184581 [0.2355]	0.0759 [0.1947]	1 [0]	-74704 [0.2230]	0.0766 [0.1859]	1 [0]	-15140 [0.2910]
asym. beta	0.0716 [0.1854]	1 [0]	-222046 [0.2680]	0.0788 [0.1856]	1 [0]	-74786 [0.2587]	0.0782 [0.1864]	1 [0]	-15104 [0.3133]
TSRV _{spot}	0.0726 [0.1827]	1 [0]	-146193 [0.2912]	0 [0]	0 [0]	0 [0]	0.0806 [0.1860]	1 [0]	-15071 [0.3308]
PreAV _{spot}	0.0699 [0.1924]	1 0.0164	-224915 [0.2429]	0.0737 [0.1845]	1 [0]	-76194 [0.2152]	0.0719 [0.1863]	1 [0]	-15310 [0.2868]
RK _{spot}	0.0726 [0.1846]	1 [0]	-221276 [0.3049]	0.0797 [0.1849]	1 [0]	-74470 [0.3298]	0.0820 [0.1859]	1 [0]	-15025 [0.3801]
ACD _{spot}	0.0737 [0.1840]	1 [0]	-176696 [0.2626]	0.0784 [0.1848]	1 [0]	-75149 [0.2465]	0.0818 [0.1857]	1 [0]	-15050 [0.3365]
DUR _{spot}	0.0712 [0.1860]	1 [0]	-137531 [0.2485]	0.0784 [0.1849]	1 [0]	-69426 [0.2614]	0.0780 [0.1856]	1 [0]	-15094 [0.3140]
PreAV	0.0731	1	-128072 0.0790	1 -51060	1 0.0823	1 1	-15037 1		
Estimators from subsection 4.2.1									

	[0.1858]	[0]	[0.2740]	[0.1841]	[0]	[0.2526]	[0.1855]	[0]	[0.3502]
TSRV	0.0734	1	-222170	0.0825	1	-53007	0.0835	1	-15026
	[0.1853]	[0]	[0.2994]	[0.1896]	[0]	[0.2810]	[0.1854]	[0]	[0.3578]
RK	0.0726	1	-220799	0.0799	1	-62122	0.0844	1	-15044
	[0.1848]	[0]	[0.2916]	[0.1837]	[0]	[0.2669]	[0.1859]	[0]	[0.3977]
RV ^K	0.0735	1	-97540	0.0797	1	-70401	0.0821	1	-15064
	[0.1846]	[0]	[0.2906]	[0.1853]	[0]	[0.3038]	[0.1860]	[0]	[0.3920]
RV	0.0725	1	-97630	0.0790	1	-58156	0.0803	1	-15077
	[0.1837]	0.0164	[0.2745]	[0.1838]	[0]	[0.2546]	[0.1861]	[0]	[0.3539]
R. Range	0.0739	1	-97240	0.0798	1	-55390	0.0823	1	-15031
	[0.1840]	[0]	[0.3148]	[0.1853]	[0]	[0.3401]	[0.1850]	[0]	[0.4665]
Sq. Return	0.0745	1	-142904	0.0792	1	-74683	0.0763	1	-15126
	[0.2006]	[0]	[0.3235]	[0.1867]	[0]	[0.3195]	[0.1851]	[0]	[0.5034]

Table 4.12: Diagnostics for the overall fit and 10% - tail risk fit of the t-RV-GARCH based on 1 min, 5 min and 20 second (integrated) volatility measures. Notation as above: CCP denotes the Conditional Coverage probability estimate, CCP_{loss} the corresponding loss function. D and D_{tail} denote the P-values for the density forecast backtest and the density forecast backtest for the tail, respectively. LL Value and LB_{10} give the log-likelihood value and the Ljung-Box test for ten lags. Sq. Return and R. Range denote the squared return model (=standard GARCH) and the RV-GARCH based on the realized range. Optimal values highlighted in each column

Bibliography

- S. Alizadeh, M. W. Brandt, and F. X. Diebold. Range-based estimation of stochastic volatility models. *Journal of Finance*, 57(3):1047–1091, 2002.
- A. Anand, S. Chakravarty, and T. Martell. Empirical evidence on the evolution of liquidity: Choice of market versus limit orders by informed and uninformed traders. *Journal of Financial Markets*, 8(3):289–309, 2005.
- T. G. Andersen and T. Bollerslev. Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, 4(2):115–158, 1997.
- T. G. Andersen and T. Bollerslev. Dm-dollar volatility: Intraday activity patterns, macroeconomic announcements and longer-run dependencies. *Journal of Finance*, 53(1):219–265, 1998.
- T. G. Andersen, T. Bollerslev, F. Diebold, and P. Labys. Modeling and forecasting realized volatility. *Econometrica*, 71(2):579–625, 2003.
- T. G. Andersen, D. Dobrev, and E. Schaumburg. Duration-based volatility estimation. Working Paper 034, Hitotsubashi University, 2008.
- W. Antweiler and M. Z. Frank. Is all that talk just noise? the information content of internet message boards. *The Journal of Finance*, 59(3):1259–1293, 2004.
- R. T. Baillie, T. Bollerslev, and H. O. Mikkelsen. Fractionally integrated generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 74(1):3–30, 1996.
- R.T. Baillie. Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, 73(1):5–59, 1996.
- F. M. Bandi and R. Reno. Nonparametric stochastic volatility. Discussion Paper gd08-035, Hitotsubashi University, 2009.
- Y. Bao, T. Lee, and B. Saltoglu. Evaluating predictive performance of value-at-risk models in emerging markets: A reality check. *Journal of Forecasting*, 25:101–128, 2006.
- B. M. Barber and T. Odean. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. *The Review of Financial Studies*, 21(2):785–818, 2008.

Bibliography

- O. E. Barndorff-Nielsen, P. R. Hansen, A. Lunde, and N. Shephard. Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise. *Econometrica*, 76(6):1481–1536, 2008a.
- O. E. Barndorff-Nielsen, P. R. Hansen, A. Lunde, and N. Shephard. Realised kernels in practice: Trades and quotes. *Econometrics Journal*, 4:1–32, 2008b.
- J. Beran. *Statistics for long-memory processes*. Chapman and Hall/ CRC, Boca Raton, Florida, 1998.
- J. Berkowitz. Testing density forecasts with applications to risk management. *Journal of Business and Economics Statistics*, 19:465–474, 2001.
- T. D. Berry and K. M. Howe. Public information arrival. *The Journal of Finance*, 49(4):1331–1346, 1994.
- H. Bessembinder and K. Venkataraman. Bid-ask spreads: Measuring trade execution costs in financial markets. In R. Cont, editor, *Encyclopedia of Quantitative Finance*. John Wiley and Sons, London, 2010.
- G. Bhardwaj and N. R. Swanson. An empirical investigation of the usefulness of arfima models for predicting macroeconomic and financial time series. *Journal of Econometrics*, 131(1-2):539–578, 2006.
- B. Biais, P. Hillion, and C. Spatt. An empirical analysis of the limit order book and the order flow in the paris bourse. *Journal of Finance*, 50(5):1655–1689, 1995.
- O. Blaskowitz and H. Herwartz. Adaptive forecasting of the euribor swap term structure. *Journal of Forecasting*, 28(7):575–594, 2009.
- L. Blume, D. Easley, and M. O’Hara. Market statistics and technical analysis: The role of volume. *The Journal of Finance*, 49(1):153–181, 1994.
- N. P. B. Bollen, T. Smith, and R. E. Whaley. Modeling the bid/ask spread: Measuring the inventory holding premium. *Journal of Financial Economics*, 72(1):97–141, 2004.
- C. S. Bos, P. Janus, and S. J. Koopman. Spot variance path estimation and its application to high frequency jump testing. Discussion paper, Tinbergen Institute, 2009.
- P. Boswijk and Y. Zu. Estimating spot volatility with high frequency financial data. Working paper, Tinbergen Institute, 2010.
- P. Bougerol and N. Picard. Stationarity of garch processes and of some nonnegative time series. *Journal of Econometrics*, 52(1-2):115–127, 1992.
- C. Brownlees and G. M. Gallo. Comparison of volatility measures: A risk management perspective. *Journal of Financial Econometrics*, 8(1):29–56, 2010.

- C. T. Brownlees and G. M. Gallo. Financial econometric analysis at ultra high-frequency: Data handling concerns. *Computational Statistics and Data Analysis*, 51(4):2232–2245, 2006.
- C. T. Brownlees, F. Cipollini, and G. M. Gallo. Intra-daily volume modeling and prediction for algorithmic trading. *Journal of Financial Econometrics*, 9(3):489–518, 2011.
- J. Busse and T. C. Green. Market efficiency in real time. *Journal of Financial Economics*, 65(3):415–437, 2002.
- J. Y. Campbell, A. W. Lo, and A. C. MacKinlay. *The econometrics of financial markets*. Princeton University Press, Princeton, New Jersey, 1997.
- K. C. Chan, W. G. Christie, and P. H. Schultz. Market structure and the intraday pattern of bid-ask spreads for nasdaq securities. *Journal of Business*, 68(1):35–60, 1995.
- S. X. Chen. Beta kernel smoothers for regression curves. *Statistica Sinica*, 10:73–91, 2000.
- T. Chordia, R. Roll, and A. Subrahmanyam. Recent trends in trading activity and market quality. *Journal of Financial Economics*, 101(2):243–263, 2011.
- K. Christensen and M. Podolskij. Realized range-based estimation of integrated variance. *Journal of Econometrics*, 141(2):323–349, 2007.
- P. F. Christoffersen. Evaluating interval forecasts. *International Economic Review*, 39(4):841–862, 1998.
- K. H. Chung, B. F. Van Ness, and R. A. Van Ness. Limit orders and the bid-ask spread. *Journal of Financial Economics*, 53(2):255–287, 1999.
- P. K. Clark. A subordinate stochastic process model with finite variance for speculative prices. *Econometrica*, 41(1):135–155, 1973.
- T. E. Clark and M. W. McCracken. Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics*, 105(1):85–110, 2001.
- T. E. Clark and K. D. West. Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, 138(1):291–311, 2007.
- G. Colletaz, C. Hurlin, and S. Tokpavi. Irregularly spaced intraday value at risk (isivar) models: Forecasting and predictive abilities. Working Paper 00162440, HAL, 2007.
- C. Conrad. Non-negativity conditions for the hyperbolic garch model. *Journal of Econometrics*, 157(2):441–457, 2010.
- C. Conrad and B. R. Haag. Inequality constraints in the fractionally integrated garch model. *Journal of Financial Econometrics*, 4(3):413–449, 2006.

Bibliography

- C. Conrad and M. Karanasos. The impulse response function of the long memory garch process. *Economics Letters*, 90(1):34–41, 2006.
- T. E. Copeland and D. Galai. Information effects on the bid/ask spread. *Journal of Finance*, 38(5):1457–1469, 1983.
- F. Corsi. A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 7(2):174–196, 2009.
- J. Cragg. Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica*, 39(5):829–844, 1971.
- F. Dalbaen and W. Schachermeyer. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300:463–520, 1994.
- R. A. Davis, W. T. M. Dunsmuir, and Y. Wang. Modelling time series of counts. In S. Ghosh, editor, *Asymptotics, nonparametrics and time series: A tribute to Madan Lal Puri*. Marcel Dekker, New York, 1999.
- R. A. Davis, W. T. M. Dunsmuir, and S. B. Streett. Observation-driven models for poisson counts. *Biometrika*, 90(4):777–790, 2003.
- R. Deo, M. Hsieh, and C. Hurvich. Long memory in intertrade durations, counts and realized volatility of nyse stocks. *Journal of Statistical Planning and Inference*, 140 (12):3715–3733, 2010.
- F. X. Diebold and R. S. Mariano. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13(3):253–265, 1995.
- F. X. Diebold and G. D. Rudebusch. Long memory and persistence in aggregate output. *Journal of Monetary Economics*, 24(2):189–209, 1989.
- Z. Ding and C. W. J. Granger. Modeling volatility persistence of speculative returns: A new approach. *Journal of Econometrics*, 73(1):185–215, 1996.
- G. Dionne, P. Duchesne, and M. Pacurar. Intraday value at risk (ivar) using tick-by-tick data with application to the toronto stock exchange. *Journal of Empirical Finance*, 16(5):777–792, 2009.
- D. Easley and M. O’Hara. Price, trade size and information in securities markets. *Journal of Financial Economics*, 19(1):69–90, 1987.
- D. Easley and M. O’Hara. Time and the process of security price adjustment. *Journal of Finance*, 47(2):577–605, 1992.
- D. Easley, M. M. Lopez de Prado, and M. O’Hara. The microstructure of the flash crash: Flow toxicity, liquidity crashes and the probability of informed trading. *The Journal of Portfolio Management*, 37:118–128, 2011.

Bibliography

- B. Efron. Double exponential families and their use in generalized linear regression. *Journal of the American Statistical Association*, 81(395):709–721, 1986.
- R. F. Engle. The econometrics of ultra-high frequency data. *Econometrica*, 68:1–22, 2000.
- R. F. Engle and S. Manganelli. Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics*, 22(4):367–381, 2004.
- R. F. Engle and A. J. Patton. Impact of trades in an error-correction model of quote prices. *Journal of Financial Markets*, 7(5):1–25, 2004.
- R. F. Engle and J. Russell. Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica*, 66(5):1127–1162, 1998.
- R. F. Engle and J. R. Russell. A discrete-state continuous-time model of financial transactions prices and times. *Journal of Business and Economic Statistics*, 23(2):166–180, 2005.
- R. F. Engle, M. E. Sokalska, and A. Chanda. Forecasting intraday volatility in the us equity market with a multiplicative component garch. Technical report, North American Meetings of the Econometric Society, 2006.
- R. Ferland, A. Latour, and D. Oraichi. Integer-valued garch process. *Journal of Time Series Analysis*, 27(6):923–942, 2006.
- M. J. Fleming and E. M. Remolona. Price formation and liquidity in the u.s. treasury market: The response to public information. *The Journal of Finance*, 54(5):1901–1915, 1999.
- K. Fokianos, A. Rahbek, and D. Tjostheim. Poisson autoregression. *Journal of the American Statistical Association*, 104:1430–1439, 2009.
- D. P. Foster and D. B. Nelson. Continuous record asymptotics for rolling sample variance estimators. *Econometrica*, 64(1):139–174, 1996.
- T. Foucault. Order flow decomposition and trading costs in a dynamic limit order market. *Journal of Financial Markets*, 2(2):99–134, 1999.
- T. Foucault, O. Kadan, and E. Kandel. Limit order book as a market for liquidity. *Review of Financial Studies*, 18(4):1171–1217, 2005.
- R. K. Freeland and B. P. M. McCabe. Forecasting discrete valued low count time series. *International Journal of Forecasting*, 20(3):427–434, 2004.
- R. A. Gallant. On the bias in flexible functional forms and an essential unbiased form: The fourier flexible form. *Journal of Econometrics*, 15(2):211–245, 1981.

Bibliography

- T. J. George, G. Kaul, and M. Nimalendran. Estimation of the bid-ask spread and its components: A new approach. *Review of Financial Studies*, 4(4):623–656, 1991.
- A. Gerig, J. Vicente, and M. A. Fuentes. Model for non-gaussian intraday stock returns. *Physical Review E*, 80:065102–1–065102–4, 2009.
- J. Geweke and S. Porter-Hudak. The estimation and application of long memory time series models. *Journal of Time Series*, 4(4):221–238, 1983.
- P. Giot. Market risk models for intraday data. *European Journal of Finance*, 11:309–324, 2005.
- L. Giraitis, P. M. Robinson, and D. Surgailis. Larch, leverage and long memory. *Journal of Financial Econometrics*, 2(2):177–210, 2004.
- L. R. Glosten. Components of bid-ask spread and statistical properties of transaction prices. *Journal of Finance*, 42(5):1293–1307, 1987.
- L. R. Glosten and L. E. Harris. Estimating the components of the bid/ask spread. *Journal of Financial Economics*, 21(1):123–142, 1988.
- L. R. Glosten and P. Milgrom. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14(1):71–100, 1985.
- S. Goncalves and N. Meddahi. Bootstrapping realized volatility. *Econometrica*, 77(1):283–306, 2009.
- C. W. J. Granger. Some properties of time series data and their use in econometric model specification. *Journal of Econometrics*, 16(1):121–130, 1981.
- C. W. J. Granger and R. Joyeaux. An introduction to long memory time series models and fractional differencing. *Journal of Time Series Analysis*, 1(1):15–39, 1980.
- M. Griffiths, B. Smith, A. Turnbull, and R. White. The costs and determinants of order aggressiveness. *Journal of Financial Economics*, 56(1):65–88, 2000.
- A. Groß-Klußmann and N. Hautsch. When machines read the news: Using automated text analytics to quantify high frequency news-implied market reactions. *Journal of Empirical Finance*, 18(2):321–340, 2011a.
- A. Groß-Klußmann and N. Hautsch. Predicting bid-ask spreads using long memory conditional poisson models. Working paper, Humboldt University Berlin, 2011b.
- A. D. Hall and N. Hautsch. Order aggressiveness and order book dynamics. *Empirical Economics*, 30(4):973–1005, 2006.
- J. D. Hamilton. *Time Series Analysis*. Princeton University Press, Princeton, New Jersey, 1994.

Bibliography

- P. R. Hansen. A test for superior predictive ability. *Journal of Business and Economic Statistics*, 23(4):365–380, 2005.
- P. R. Hansen and A. Lunde. Realized variance and market microstructure noise. *Journal of Business and Economic Statistics*, 24:127–218, 2006.
- P. R. Hansen, Z. Huang, and H. H. Shek. Realized garch: A complete model of returns and realized measures of volatility. *Journal of Applied Econometrics*, 2011. forthcoming.
- L. Harris and J. Hasbrouck. Market vs. limit orders: The superdot evidence on order submission strategy. *Journal of Financial and Quantitative Analysis*, 31(2):213–231, 1996.
- M. Harris and A. Raviv. Differences of opinion make a horse race. *The Review of Financial Studies*, 6(3):473–506, 1993.
- A.C. Harvey and C. Fernandez. Time series models for count or qualitative observations. *Journal of Business and Economic Statistics*, 7(4):407–417, 1989.
- J. Hasbrouck. The dynamics of discrete bid and ask quotes. *Journal of Finance*, 54(6): 2109–2142, 1999.
- J. Hasbrouck and G. Saar. Limit orders and volatility in a hybrid market: The island ecn. Working paper, Stern School of Business, 2002.
- N. Hautsch. Capturing common components in high frequency time series: a multivariate stochastic multiplicative error model. *Journal of Economic Dynamics and Control*, 32 (12):3978–4009, 2008.
- N. Hautsch. *Econometrics of financial high-frequency data*. Springer Verlag, Heidelberg, 2012.
- N. Hautsch and R. Huang. The dark side of the market: Detecting and modelling hidden liquidity. Working paper, Humboldt University Berlin, 2011.
- N. Hautsch and R. Huang. The market impact of a limit order. *Journal of Economic Dynamics and Control*, 36(3):501–522, 2012.
- N. Hautsch and M. Podolskij. Pre-averaging based estimation of quadratic variation in the presence of noise and jumps: Theory, implementation and empirical evidence. Discussion Paper 2010-038, SFB 649, 2010.
- A. Heinen. Modelling time series count data: An autoregressive conditional poisson model. Working Paper 1117187, Social Science Research Network, 2003.
- A. Heinen and E. Rengifo. Multivariate autoregressive modeling of time series count data using copulas. *Journal of Empirical Finance*, 14(4):564–583, 2007.

Bibliography

- T. Hendershott, C. M. Jones, and A. J. Menkveld. Does algorithmic trading improve liquidity? *Journal of Finance*, 66(1):1–33, 2011.
- J. R. M. Hosking. Fractional differencing. *Biometrika*, 68(1):165–76, 1981.
- W. Härdle, N. Hautsch, and A. Mihoci. Modelling and forecasting liquidity supply using semiparametric factor dynamics. Working Paper 1475168, Social Science Research Network, 2009.
- R. Huang and H. Stoll. The components of the bid-ask spread: A general approach. *Review of Financial Studies*, 10(4):995–1034, 1997.
- X. Huang and G. Tauchen. The relative contribution of jumps to total price variation. *Journal of Financial Econometrics*, 3(4):456–499, 2005.
- S. Ikeda. A kernel estimator of the spot volatility of security returns with stationary noise, diurnal seasonality and leverage effect. Discussion paper, Boston University, 2010.
- J. Jacod, Y. Li, P. Mykland, M. Podolskij, and M. Vetter. Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Processes and their Applications*, 119(7):2249–2276, 2009.
- J. Jasiak. Persistence in intertrade duration. *Finance*, 19:166–195, 1998.
- R. C. Jung, M. Kukuk, and R. Liesenfeld. Time series of count data: Modeling, estimation and diagnostics. *Computational Statistics & Data Analysis*, 51(4):2350–2364, 2006.
- P. S. Kalev, W. S. Liu, P. K. Pham, and E. Jarnebic. Public information arrival and volatility of intraday stock returns. *The Journal of Banking and Finance*, 28(6):1441–1467, 2004.
- E. Kandel and N. D. Pearson. Differential interpretation of public signals and trade in speculative markets. *The Journal of Political Economy*, 103(4):831–872, 1995.
- M. Karanasos, Z. Psaradakis, and M. Sola. On the autocorrelation properties of long memory garch processes. *Journal of Time Series Analysis*, 25(2):266–281, 2004.
- J. M. Karpoff. A theory of trading volume. *The Journal of Finance*, 41(5):1069–1087, 1986.
- O. Kim and R. E. Verrecchia. Market reactions to anticipated announcements. *Journal of Financial Economics*, 30(2):273–309, 1991.
- O. Kim and R. E. Verrecchia. Market liquidity and volume around earnings announcements. *The Journal of Accounting and Economics*, 17(1-2):41–67, 1994.

- A. A. Kirilenko, A. S. Kyle, M. Samadi, and T. Tugkan. The flash crash: The impact of high frequency trading on an electronic market. Working Paper 1853768, Social Science Research Network, 2011.
- R. Koenker and G. Bassett. Regression quantiles. *Econometrica*, 46(1):33–50, 1978.
- D. Koulikov. Modeling sequences of long memory non-negative covariance stationary random variables. Working Paper 331100, Social Science Research Network, 2003.
- I. Krinsky and J. Lee. Earnings announcements and the components of the bid-ask spread. *The Journal of Finance*, 51(4):1523–1535, 1996.
- D. Kristensen. Nonparametric filtering of the realized spot volatility. *Econometric Theory*, 26(1):60–93, 2010.
- N. H. Kupiec. Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, 3:73–84, 1995.
- A. S. Kyle. Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1336, 1985.
- C. M. C. Lee and M. J. Ready. Inferring trade direction from intraday data. *Journal of Finance*, 46(2):733–746, 1991.
- C. M. C. Lee, B. Mucklow, and M. J. Ready. Spread, depths and the impact of earnings information: An intraday analysis. *Review of Financial Studies*, 6(2):345–374, 1993.
- A. W. Lo, A. C. MacKinlay, and J. Zhang. Econometric models of limit-order executions. *Journal of Financial Economics*, 65(1):31–71, 2002.
- T. Lux and T. Kaizoji. Forecasting volatility and volume in the tokyo stock market: Long memory, fractality and regime switching. *Journal of Economic Dynamics and Control*, 31(6):1808–1843, 2007.
- I.L. MacDonald and W. Zucchini. *Hidden Markov and other models for discrete-valued time series*. Chapman and Hall, London, 1997.
- S. Manganelli. Duration, volume and volatility impact of trades. *Journal of Financial Markets*, 8(4):377–399, 2005.
- M. Martens and D. van Dijk. Measuring volatility with the realized range. *Journal of Econometrics*, 138(1):181–207, 2007.
- E. McKenzie. Discrete variate time series. In D. N. Shanbhag and C.R. Rao, editors, *Handbook of Statistics*. Elsevier, London, 2003.
- M. L. Mitchell and J. H. Mulherin. The impact of public information on the stock market. *The Journal of Finance*, 49(3):923–950, 1994.

Bibliography

- P. Mykland and L. Zhang. The econometrics of high frequency data. In M. Kessler, A. Lindner, and M. Soerensen, editors, *Statistical methods for stochastic differential equations*. Chapman and Hall, London, 2010. to appear.
- R. C. A. Oomen. Properties of realized variance under alternative sampling schemes. *Journal of Business and Economic Statistics*, 24:219–237, 2006.
- M. Parkinson. The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 53(1):61–65, 1980.
- C. A. Parlour. Price dynamics in limit order markets. *Review of Financial Studies*, 11 (4):789–816, 1998.
- R. Pascual and D. Veredas. What pieces of limit order book information matter in explaining order choice by patient and impatient traders? *Quantitative Finance*, 9(5): 527–545, 2009.
- A. Ranaldo. Order aggressiveness in limit order book markets. *Journal of Financial Markets*, 7(1):53–74, 2004.
- A. Ranaldo. Intraday market dynamics around public information disclosures. In F.-S. Lhabitant and G.G. Gregoriou, editors, *Stock Market Liquidity*, pages 199–226. John Wiley and Sons, New Jersey, 2008.
- R. Rosenblatt. Remarks on a multivariate transformation. *Annals of Mathematical Statistics*, 23(3):470–472, 1952.
- T. H. Rydberg and N. Shephard. Bin models for trade-by-trade data. modeling the number of trades in a fixed interval of time. Working Paper 0740, Nuffield College, 1999.
- T. H. Rydberg and N. Shephard. Dynamics of trade-by-trade movements: Decomposition and models. *Journal of Financial Econometrics*, 1(1):2–25, 2003.
- SEC and CFTC. Findings regarding the market events of may 6, 2010. Working paper, US Securities and Exchange Comission, 2010.
- N. Shephard and K. Sheppard. Realising the future: Forecasting with high frequency based volatility (heavy) models. *Journal of Applied Econometrics*, 25:197–231, 2010.
- B. C. Sutradhar. On forecasting counts. *Journal of Forecasting*, 27(2):109–129, 2008.
- George E. Tauchen and Mark Pitts. The price variability-volume relationship on speculative markets. *Econometrica*, 51(2):485–505, 1983.
- N. Taylor. The economic and statistical significance of spread forecasts: Evidence from the london stock exchange. *Journal of Banking and Finance*, 26(4):795–818, 2002.
- P. C. Tetlock. Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3):1139–1167, 2007.

Bibliography

- P. C. Tetlock. Does public financial news resolve asymmetric information? *The Review of Financial Studies*, 23(9):3520–3557, 2010.
- P. C. Tetlock, S. A. Macskassy, and M. Saar-Tsechansky. More than words: Quantifying language to measure firms’ fundamentals. *The Journal of Finance*, 63(3):1427–1467, 2008.
- Y. Tse and T. Yang. Estimation of high-frequency volatility: Autoregressive conditional duration models approach. Working Paper 1276, Research collection school of economics, 2010.
- H. White. A reality check for data snooping. *Econometrica*, 68(1):1097–1126, 2000.
- P. Zaffaroni. Stationarity and memory of arch(∞) models. *Econometric Theory*, 20(1):147–160, 2004.
- S.-L. Zeger. A regression model for time series of counts. *Biometrika*, 75(4):621–629, 1988.
- L. Zhang. Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli*, 12:1019–1043, 2006.
- L. Zhang, P. A. Mykland, and Y. Ait-Sahalia. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, 100(472):1394–1411, 2005.

List of Figures

2.1	Distribution of news and relevance indicator	10
2.2	Intervals around news arrival	15
2.3	Money Value and Volatility around news arrivals	16
2.4	Spread and depth around news arrivals	16
2.5	Average trade sizes and absolute trade imbalance around news	17
2.6	Money value and spread around initial news and updates	17
2.7	Cumulated abnormal returns around relevant positive, negative and neutral news	19
2.8	Cumulated abnormal returns after relevant positive and negative news . .	20
2.9	Profitability of trading on the sentiments	20
2.10	Typical sample autocorrelations for the variables of interest	22
2.11	Numbers of significant dummy variables in the intervals	27
2.12	Response analysis for highly relevant news	27
2.13	Average intraday seasonality patterns	32
2.14	Signed trade imbalance around positive, negative and neutral news . . .	32
2.15	Sample average of cross-correlations (I)	33
2.16	Sample average of cross-correlations (II)	33
3.1	Evolution of the average quoted and average effective bid-ask spreads .	38
3.2	Histogram and typical pattern of the 30s bid-ask spreads	38
3.3	Autocorrelation functions of the 30s spreads	40
3.4	Log-log of variance of mean vs. sample size	40
3.5	Intraday Periodicities for the 30s bid-ask spread series in January and February 2008	41
3.6	Evolution of the estimates of the fractional integration parameter in the LMACDP type II	52
3.7	The principle of the trading schedule	57
3.8	Evolution of the estimates of the absolute return coefficient	62
3.9	Evolution of the estimates of the realized volatility coefficient	62
3.10	Evolution of the estimates of the depth coefficient	63
3.11	Evolution of the estimates of the depth imbalance coefficient	63
3.12	Evolution of the estimates of the traded volume coefficient	64
3.13	Evolution of the estimates of the traded volume imbalance coefficient . .	64
4.1	Example of duration time stamps $t_{d,i}$ not aligned with the interval boundaries t_l and t_n	76

List of Figures

4.2	Example of a simulated log-price process realization and the corresponding spot volatilities	79
4.3	Example of a simulated log-price process realization and the corresponding spot volatilities	79
4.4	Sample autocorrelations and intraday seasonality of the standard realized volatility measure based on 5 minute intervals	88
4.5	Sample autocorrelations of the realized volatility measure	89
4.6	Histogram and fitted normal distribution for the 5 minute returns	89
4.7	Sample autocorrelations of four 20 second volatility measures	104
4.8	Sample autocorrelations of four 1 minute volatility measures	104
4.9	Sample autocorrelations of four 5 minute volatility measures	105

List of Tables

2.1	Descriptive statistics of the FTSE stocks	12
2.2	Average VAR Results: Dynamics and Dummies	25
3.1	Descriptive statistics of the 30s quoted spread in ticks	39
3.2	GPH Test of the spread series	41
3.3	Median parameter estimates for the ACDP and LMACDP type I and II .	52
3.4	RMSE and DA values of the forecasting study	55
3.5	Diebold Mariano and Clark-West test results for point and direction forecasts	55
3.6	P-Values of the SPA test for all competing models	56
3.7	Spread cost savings as percentage of the benchmark schedule	58
4.1	Microstructure noise estimates in 0.00001 for the day 03/13/2008 for selected stocks	78
4.2	Optimal bandwidths and parameters for the 1-factor model (grid-search procedure)	81
4.3	RMSE for 20 second, 1 minute and 5 minute (integrated) volatility estimates (1-factor model)	83
4.4	RMSE for 20 second, 1 minute and 5 minute (integrated) volatility estimates (2-factor model)	85
4.5	Average empirical bandwidth selection based on the grid search procedure and the cross-validation criterion	87
4.6	Average parameter and robust standard error estimates of two RV-Garch specifications	91
4.7	Diagnostics for the overall fit and 10% - tail risk fit of the Normal-RV-GARCH based on 1 min, 5 min and 20 second (integrated) volatility measures	95
4.8	Optimal volatility measures for the CCP and CCP_{loss} criterium for $\alpha = 0.1$.	96
4.9	Optimal bandwidths and parameters for the simulated 2-factor model .	100
4.10	Further RMSE results for a simulated log price process with noise	102
4.11	Mean and Variance of the volatility measures	106
4.12	Diagnostics for the overall fit and 10% - tail risk fit of the t-RV-GARCH .	108

Selbständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Berlin, den 02.03.2012

Axel Groß-Klußmann