Applications of Point Processes in Empirical Economics and Finance

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Wirtschaftswissenschaftlichen Fakultät

der Eberhard-Karls-Universität Tübingen

vorgelegt von

Kerstin Kehrle

aus Augsburg

2010

Dekanin:	Prof. Dr. rer. pol. Kerstin Pull
Erstberichterstatter:	Prof. Dr. rer. pol. Joachim Grammig
Zweitberichterstatter:	Prof. Dr. rer. pol. Martin Biewen
Tag der mündlichen Prüfung:	16. Dezember 2009

Acknowledgments

Und jedem Anfang wohnt ein Zauber inne,
der uns beschützt und der uns hilft zu leben.

A magic dwells in each beginning,
protecting us tells us how to live.
(Hermann Hesse)

Although, I officially started my PhD in October 2005, statistics and econometrics sparked my interests and fascination already during my first semesters at the University of Tübingen and Katholieke Universiteit Leuven. I want to thank my undergraduate teachers Gerd Ronning and Robert Jung for introducing me to the field, for employing me as a student assistant and for allowing me to collect my first teaching experiences as an undergraduate tutor.

As a PhD student I had the pleasure to present my work at various seminars and conferences. I gratefully acknowledge financial support from the German Economic Association, the International Institute of Forecasters and the Graduate School of Economics at the University of Tübingen. I thank all seminar and conference participants for useful comments. In particular, I want to mention Marcelo Fernandes, Tilmann Gneiting, Alexander Kempf and Winfried Pohlmeier. I am grateful to Martin Biewen and Rainer Schöbel for serving as members in my thesis committee.

Although, many teachers taught me, no one has had such a profound influence on my academic thinking and working as my supervisor Joachim Grammig. I am deeply indebted to him for sharing his knowledge with me and for supporting me. I thank him for working with me on a joint project. What I learned from him is invaluable and without his encouragements,

comments and suggestions, this PhD thesis would not have been written. I could not have wished for a better boss, coach and coauthor.

Many people accompanied me during the years of my PhD project at the chair of Econometrics, Statistics and Empirical Economics. I want to thank our fantastic student assistants who always delivered more than one expected (please forgive me if I forgot you): Irina Dyshko, Tati Figueiredo, Benjamin Friedrich, Tobias Gummersbach, Benedikt Heid, Tobias Langen, Felix Prothmann, Jantje Sörensken, Jan Starmans, Natascha Wagner and Franziska Weiss. I thank also the secretaries Sylvia Bürger and Angelika Hutt. I was fortunate to work with great colleagues: Thomas Dimpfl, Stefan Frey, Luis Huergo, Stephan Jank, Franziska J. Peter, Peter Schmidt, Miriam Sperl, Oliver Wünsche. In particular, I owe Thomas Dimpfl and Stefan Frey my gratitude for installing an excellent IT infrastructure. Special thanks are also reserved for Franziska J. Peter, my doctoral sister. She did not hesitate to start a common project and put all effort and passion in writing a joint paper with me. I really appreciated working with her and hope that there will be future prospering projects and the same unspoken thinking and understanding between us. I am very grateful to Oliver Wünsche, who always gave me the feeling of being understood, who provided a shoulder to cry on, who comforted me and confirmed my self-confidence in times I doubted, and who made me laugh - also when there was nothing to laugh about. I cannot imagine that anybody under the sun could possibly have a better team than I had during the time I wrote this thesis.

Beyond academia there exists another world. I am grateful to my friends who are simply there for me. Thanks Ina, Lisa, Manu, Solveig and Tina. Finally, I thank my family. My parents, Anneliese and Karl and my brother, Jan-Michael. What would I be without them? They are the best and I thank them for their constant love, indestructible trust and invaluable advice.

Often, PhD students - including me - live a life on a rollercoaster that is marked by failure and success and some more failure. Those who are able to share these ups and downs with others are lucky - lucky like me.

Kerstin Kehrle, Tübingen, 15. January 2010

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Standard statistical methods in the empirical economics and finance literature are mostly applicable to data that is aggregated on equally spaced time points. However, a key characteristic of many economic and financial data is their random occurrence and irregular spacing in time. Since the pathbreaking work of Robert Engle in the last years of the 20^{th} century, there are new approaches that do not require aggregated data but are able to account for their irregular timing nature. A new field of empirical research was born and a vast amount of work followed.

These developments were mainly supported by the increasing availability of high frequency transaction data due to the implementation of electronic order recording systems at stock exchanges all over the world. Typically, financial markets data are irregularly observed along the time axis. As pointed out by Hautsch (2003) and Bowsher (2007), the time series analysis of fixed time interval data annihilates the natural timing dependence of transaction data and possibly neglects relevant information. Further, the selection of inappropriate equidistant aggregation schemes and the exclusion of data points might lead to misspecifications. Easley and O'Hara (1992) and Dufour and Engle (2000) are prominent references in the market microstructure literature that show the importance of time in the transaction process. Hence, the inclusion of all events in an empirical analysis provides additional information about the timing relation of transaction variables and allows to revisit old and to analyze new questions delivered by financial markets theory.

The statistical modeling framework to account for characteristics of irregularly spaced event data is provided by the theory of *point processes*. A point process statistically describes

the history of events that occur consecutively in time. A process consisting of points at which we simultaneously observe variables that "mark" the points is conceived as *marked point process*.

The major leap forward in financial market data modeling using point processes is the work of Engle and Russell (1997, 1998). They develop the autoregressive conditional duration (ACD) model that describes the waiting time between consecutive events by a dynamic parametric conditional mean function. The ACD models the serial dependencies of durations and has experienced a large number of specifications and extensions in the subsequent literature.[1] An important example is the study of Grammig and Maurer (2000), who focus on an alternative distributional assumption for durations. The authors improve the empirical fit of the ACD model by replacing the exponential or Weibull distribution by a Burr distribution. Jasiak (1998) suggests a fractionally integrated ACD that incorporates long memory dependence and Zhang et al.'s (2001) threshold ACD allows for structural breaks in the duration process.

A significant contribution to the ACD literature is the work of Bauwens and Giot (2000). They introduce a logarithmic ACD specification that ensures positive durations in the presence of additionally included explanatory variables. Since the ACD updates in event time, covariates enter the logarithmic ACD model simultaneously at the arrival of financial markets events. However, if relevant information arrives within subsequent events modeling expected conditional durations in an ACD framework is limited. These shortcomings of the ACD are removed by Hamilton and Jordà's (2002) autoregressive conditional hazard (ACH) model. Hamilton and Jordà (2002) extend the ACD by proposing a discrete time intensity based approach which is able to include information within a duration spell. In contrast to the ACD, the ACH model is defined on calendar time and gives an estimate of the probability that an event occurs within the next fixed time interval.

Hence, if the point process is not only driven by its past history but also by information arriving irregularly between the points, the formulation of a conditional intensity function becomes useful. In particular, this holds in a multivariate setting when the interdependencies of two or more point processes are relevant. Engle and Lunde (2003) present a bivariate ACD that models the arrival of trade and quote durations in discrete time. However, as

[1] An extensive presentation of model modifications to the original ACD of Engle and Russell (1998) can be found in Bauwens and Giot (2001).

pointed out by Bowsher (2007) due to the model structure some information of the trading process is neglected and extensions to this model class seem to be difficult. These drawbacks are overruled by Russell (1999), Bauwens and Hautsch (2006) and Bowsher (2007). They address the question of relations and interdependencies of transaction processes by focusing on multivariate conditional intensities. In a continuous-time multivariate intensity model intensities interdepend and update immediately as new information arrives. New information includes either events in one of the individual processes or covariates that occur irregularly in time and have a direct impact on the intensities.

The starting point for the multivariate intensity models in the empirical finance literature is the work of Russell (1999). He introduces the autoregressive conditional intensity (ACI) model and applies it to quote and trade arrival rates. Hall and Hautsch (2006) use the ACI to model arrival times of buy and sell trades on a limit order book market. Bauwens and Hautsch (2006) extend the ACI by an underlying common latent factor. Recently, Bowsher (2007) introduced a multivariate generalized Hawkes intensity model for the timing interaction of trades and quotes.

This thesis's aim is to present new univariate and multivariate empirical point processes applied in the field of financial and monetary econometrics. In particular, we analyze the following topics. In the second chapter we suggest a univariate discrete marked point process model for the federal funds rate target and investigate its point and probability forecast performance. Chapter 3 presents a model for daily return variation that is disentangled into a continuous and jump variation component. While daily continuous variation is modeled by an autoregressive conditional time series model, irregularly occurring jumps are conceived as a univariate marked point process. Finally, the fourth chapter introduces a new information share that measures the home and foreign market share in price discovery. For this purpose, a multivariate point process based on high frequency transaction data is used.

Since the main focus of the thesis is to show point processes from an empirical applicants perspective, mathematical concepts are introduced whenever required. Karr (1986) gives an extensive statistical description of point processes and their inference and an excellent overview of point processes applied to financial transaction data is given by Hautsch (2003).

A New Marked Point Process Model for the Federal Funds Rate Target - Methodology and Forecast Evaluation

Although the ACD model was primarily developed in the context of high frequency tick-by-tick transaction data, the idea of "time matters" spread into research that uses non-aggregated data observed on lower frequencies, as well. Chapter 2 studies the US Federal Reserve Bank's (Fed) monetary policy on weekly data. The Fed's main policy tool to regulate the demand and supply of money is to set a *target* interest rate for the effective federal funds rate which is the rate at which depository institutions lend reserves at the Fed to other depository institutions overnight. In the meetings of the Federal Open Market Committee (FOMC), the Fed decides based on macroeconomic and financial indicators whether the target interest rate changes. Hence, an irregular spaced time series of target changes emerges due to the Fed's institutional framework. Hamilton and Jordà (2002) propose to model the target changes as points in time and to combine them with an ordered probit (OP) that accounts for the discreteness of target change sizes. In other words, Hamilton and Jordà (2002) model the target changes as a marked point process.

Chapter 2 of this thesis draws on Hamilton and Jordà's (2002) seminal work. We present a new marked point process model for the federal funds rate target by combining Hamilton and Jordà's (2002) autoregressive conditional hazard (ACH) and Russell and Engle's (2005) autoregressive conditional multinomial (ACM) model. Further, the chapter compares the forecast performance of the proposed model with the ACH-OP and other approaches. We also suggest a method to assess the quality of probability forecasts delivered by this model class and apply it to forecasts of the federal funds rate target. By improving goodness of fit and point forecasts, the ACH-ACM qualifies as a sensible modeling framework. Furthermore, our results show that discrete marked point process models deliver reasonable probability function forecasts at short and medium term horizons.

Forecasting Return Volatility with Continuous Variation and Jumps

Chapter 3 presents an application of a point process in the context of volatility forecasting using daily data. Due to the importance of accurate volatility forecasts for the valuation of

derivatives, portfolio management and risk management, it plays a central role in financial econometrics. Recently, Andersen et al. (2003) and Andersen and Bollerslev (1998) introduced a nonparametric approach to measure and model daily return volatility. Based on high frequency data they suggest the sum of intra-daily squared returns as realized volatility measure that converges to the quadratic variation of a continuous-time price process.

The chapter is linked to Andersen et al.'s (2007b) and Bollerslev et al.'s (2009) work who disentangle return volatility into a continuous and a jump component and model realized volatility by a reduced form time series approach. In this chapter continuous variation is described by an autoregressive time series model and jump variation is conceived as a marked point process. Daily variation jumps occur irregularly spaced in time and at each jump event (points) we immediately observe the size of the jump (mark). Further, the chapter takes up the idea and extends the method of forecast evaluation of Chapter 2 to density forecasts of realized volatility. Diagnostics as well as point and density forecast results show that the suggested approach qualifies as a useful forecast model for daily return variation.

International Price Discovery in Stock Markets - A Unique Intensity Based Information Share

Chapter 4 gives an application of a multivariate point process in the field of international price discovery. Investors' decision to invest and companies' intention to list their stocks on a stock exchange depends on the ability of an exchange to provide a prospering trading environment. In particular, within the context of international cross-listed stocks, it is of paramount concern for a national stock exchange to remain the dominant market with regard to price discovery.

In Chapter 4 we use Russell's (1999) autoregressive conditional intensity model (ACI) and develop a new information share that measures the home and foreign market share in price discovery. By using a bivariate intensity approach, we account for the informational content of time between consecutive trades and the timing interdependencies between two markets' transaction processes. In contrast to the commonly applied Hasbrouck (1995) methodology we account for the irregularity of the data and deliver a unique information share rather than lower and upper bounds.

We apply our information share to Canadian stocks that are traded on the Toronto Stock

Exchange (TSX) and cross-listed on the New York Stock Exchange (NYSE). We find that the TSX is the dominant market with an information share of 71%. Our results confirms previous findings by Phylaktis and Korczak (2007), Eun and Sabherwal (2003), and Grammig et al. (2005), who also analyze Canadian stocks. We also compare our results to the Hasbrouck (1995) information shares. On average over all sample stocks we find a larger home market contribution than indicated by the Hasbrouck midpoints.

Chapter 2

A New Marked Point Process Model for the Federal Funds Rate Target - Methodology and Forecast Evaluation

Forecasts of key interest rates set by central banks are of paramount concern for investors and policy makers. Recently it has been shown that forecasts of the federal funds rate target, the most anticipated indicator of the Federal Reserve Bank's monetary policy stance, can be improved considerably when its evolution is modeled as a marked point process (MPP). This is due to the fact that target changes occur in discrete time with discrete increments, have an autoregressive nature, and are usually in the same direction. We propose a model which is able to account for these dynamic features of the data. In particular, we combine Hamilton and Jordà's (2002) autoregressive conditional hazard (ACH) and Russell and Engle's (2005) autoregressive conditional multinomial (ACM) model. The paper also puts forth a methodology to evaluate probability function forecasts of MPP models. By improving goodness of fit and point forecasts of the target, the ACH-ACM qualifies as a sensible modeling framework. Furthermore, our results show that MPP models deliver useful probability function forecasts at short and medium term horizons.

This chapter is based on the article "A New Marked Point Process Model for the Federal Funds Rate Target Methodology and Forecast Evaluation" by J. Grammig and K. Kehrle (2008) published in the Journal of Economic Dynamics and Control.

2.1 Introduction

By setting a target for the effective federal funds rate, the executive body of the US Federal Reserve Bank influences a widespread range of economic variables and financial markets. Therefore, if and how much the Fed changes the target is of paramount interest for policy makers and investors. The econometric modeling of the target change process has to account for specific data characteristics due to institutional structures. Central banks tend to prefer small target changes in the same direction rather than a large one-time target change. As a result, target changes take place in discrete time with discrete increments and have an autoregressive nature. To address these features of the data, Hamilton and Jordà (2002) propose a dynamic model, the autoregressive conditional hazard (ACH) model, that accounts for the irregular spacing of the target changes in time and combine it with an ordered probit (OP) in order to model the discreteness of target change sizes. In this paper we draw on Hamilton and Jordà's (2002) seminal work. We present a new model for the federal funds rate target and compare its forecast performance with the ACH-OP model and other approaches. We also propose a method to assess the quality of probability forecasts delivered by this class of models and apply it to forecasts of the federal funds rate target.

Our paper is linked to the literature which focuses on the estimation of empirical reaction functions, i.e. the response of the Fed to economic developments (see Judd and Rudebusch 1998, Khoury 1990). For that purpose, other papers have employed vector autoregressive (VAR) models (e.g. Bernanke and Blinder 1992, Evans and Marshall 1998, Sack 1998). However, since target changes occur in discrete steps, and the time interval between change events is irregular, using a VAR can be criticized on methodological grounds (Rudebusch 1998, Evans and Kuttner 1998). A popular econometric approach that takes into account the discreteness of the target change sizes is the OP model. Analyses of the Fed's, the Bank of England's and ECB's monetary policy using OP models include Eichengreen et al. (1985), Davutyan and Parke (1995), Dueker (1999), Gerlach (2005), Jansen and De Haan (2006) and Carstensen (2006). Hamilton and Jordà's ACH-OP was the first model to take

into account both the irregular spacing in time and the discrete size of target changes. It can be classified as a *discrete marked point process* (DMPP) model in which the ACH explains the autoregressive dynamics of the durations between target changes.[1] It gives an estimate of the probability that a target change will occur within the next week. The OP on the other hand delivers the probability of observing a target change of a certain size given that a target change occurs. The ACH-OP methodology was a major leap forward in terms of improving forecast accuracy. Compared to a VAR, the mean squared error of the target forecast is considerably reduced at all forecast horizons.

Our paper offers two contributions to this literature. First, we motivate an alternative model for the target that combines the ACH with the autoregressive conditional multinomial (ACM) model introduced by Russell and Engle (2005) and compare its empirical performance with the ACH-OP. Russell and Engle's (2005) main objective was to provide a model for discrete transaction price changes.[2] Those tick changes do not only occur at a very high frequency (with only seconds between events), but also with irregular intervals between trade events. In this paper we deal with much longer durations between fewer events, but the similarities between the data generating processes are obvious.

Combining ACH and ACM to create a new DMPP model for the target is appealing from a methodological point of view. As pointed out by Liesenfeld et al. (2006), one major drawback of the OP model is that the parameters result from a threshold crossing latent variable model, in which the underlying continuous latent dependent variable has to be given some more or less arbitrary economic interpretation. Furthermore, Russell and Engle (2005) argue that the OP allows for a very limited dependence due to its Markov structure and is far less flexible regarding the impact of new information on the transition probabilities. The ACM model resolves these methodological shortcomings by allowing for more complex intertemporal dependencies. We conjecture that this flexibility is rewarded when modeling the autoregressive nature of target changes that usually take place in the same directions.

The second contribution of this paper is the evaluation of probability forecasts delivered by DMPP models. This is particularly interesting if the models are employed for value-at-risk and risk scenario analysis involving the federal funds rate target. For that purpose, we adapt the

[1] Due to its versatility, the ACH model enjoys increasing popularity (e.g. Demiralp and Jordà 1999, Zhang 2001, Dolado and María-Dolores 2002, Bergin and Jordà 2004, Davis and Hamilton 2004, Scotti 2005).

[2] Other applications of the ACM model can be found in Liesenfeld et al. (2006) and Prigent et al. (2004).

density forecast evaluation method popularized by Diebold et al. (1998).[3] However, a direct application of their method is infeasible, as it only applies to continuous forecast variables and does not readily extend to the probability function forecasts issued by DMPP models. This extension is delivered in the present paper. We also investigate at which horizons DMPP models deliver sensible in- and out-of-sample probability and point forecasts, and we offer recommendations for their practical use.

The main findings of this paper can be summarized as follows. The ACH-ACM model delivers encouraging results in terms of goodness of fit and out-of-sample point forecast performance. Given the relatively small number of target change events available for estimation, we argue that parsimony of the ACM specification is called for to avoid over-fitting. We therefore consider specifications which impose sensible restrictions on the responses to previous target changes. These specifications deliver economically plausible estimates and improve the benchmark model in terms of goodness of fit. In-sample probability forecast evaluations (conceived as goodness of fit diagnostics) underline the suitability of the DMPP approach towards modeling the evolution of the federal funds rate target. The out-of-sample point forecast evaluation confirms the suitability of DMPP models for target forecasting. Parsimoniously parameterized ACH-ACM specifications do a particularly good job. Bayesian type model averaging helps stabilizing the point forecast performance in subsamples. Furthermore, DMPP models deliver sensible out-of-sample probability function forecasts of the target for horizons up to six months. However, the federal funds rate target time series is relatively short. More data are needed before firm conclusions concerning longer probability forecast horizons can be given.

The remainder of this paper is structured as follows. Section 2.2 describes the institutional background and the data. Section 2.3 presents the ACH-ACM methodology and adapts techniques for the evaluation of density forecasts to assess the quality of the probability forecasts issued by DMPP models. Section 2.4 discusses estimation results, compares goodness of fit measures, evaluates in-sample point forecast performances and employs the diagnostic tools for the evaluation of probability forecasts for model comparisons. Furthermore, out-of-sample forecast evaluations are discussed. Section 2.5 summarizes the main findings and concludes.

[3] Sarno et al. (2004) assess the accuracy of density forecasts of alternative models for the federal funds rate (which is conceived as a continuous forecast variable).

2.2 Institutional Details and Data

The US Federal Reserve Bank (Fed) uses three principal tools to implement its monetary policy: the reserve requirement ratio, the discount rate and open market operations. The latter, the sales and purchases of government securities, is the most flexible and most frequently used. In the case of a purchase (sale) of securities by the Fed, the reserves increase (decrease) and money supply extends (contracts).[4] Meulendyke (1998) and recently Carpenter and Demiralp (2006) provide details on the Fed's monetary policy implementation and history. The executive organ of the Fed, the Federal Open Market Committee (FOMC), is responsible for the implementation of open market operations. Specifically, the FOMC sets a target for the effective federal funds rate which is the rate at which depository institutions lend reserves at the Fed to other depository institutions overnight.

As described by the Federal Reserve System (2005), the implementation of the monetary policy changed over time. Up to the mid-1980s the Fed influenced the effective federal funds rate indirectly by targeting the borrowed reserves, a period that was characterized by small and frequent target changes. Beginning in the mid-1980s, doubts about the financial health of some depository institutions induced a reluctance to borrow at the discount window and the link between borrowing and the federal funds rate weakened. Consequently, the Fed increasingly set a specific level of the federal funds rate rather than a targeted amount of borrowed reserves. Before 1994, the target level was not publicly announced. However, by observing the activity at the Domestic Trading Desk at the Federal Reserve Bank of New York, the objective of the Fed was inferred and speculations about the intended target level were publicized in press. Since the mid-nineties, the FOMC explicitly states its short term objective for open market operations by announcing a target level for the effective federal funds rate.

Figure 2.2.1 depicts the March 1984 to January 2006 time series of the federal funds rate target, its changes and the effective federal funds rate. Dates of the target changes are given in Table 2.2.1. The target data from March 1984 to April 2001 were obtained from O. Jordà's website. These official trading desk data were originally compiled by Rudebusch (1995) and updated by Volker Wieland.

[4] As a matter of fact, sales of securities are extremely rare. Instead of issuing new securities, the Fed rather redeems some maturing securities.

Date of Change	Target Value	Target Change	Duration in Days	Day of the Week	Date of Change	Target Value	Target Change	Duration in Days	Day of the Week
1 March 84	9.5			Thursday	11 July 85	7.6875	-0.0625	52	Thursday
15 March 84	9.875	0.375	14	Thursday	25 July 85	7.75	0.0625	14	Thursday
22 March 84	10	0.125	7	Thursday	22 August 85	7.8125	0.0625	28	Thursday
29 March 84	10.25	0.25	7	Thursday	29 August 85	7.875	0.0625	7	Thursday
5 April 84	10.5	0.25	7	Thursday	6 September 85	8	0.125	8	Friday
14 June 84	10.625	0.125	70	Thursday	18 December 85	7.75	-0.25	103	Wednesday
21 June 84	11	0.375	7	Thursday	7 March 86	7.25	-0.5	79	Friday
19 July 84	11.25	0.25	28	Thursday	10 April 86	7.125	-0.125	34	Thursday
9 August 84	11.5625	0.3125	21	Thursday	17 April 86	7	-0.125	7	Thursday
30 August 84	11.4375	-0.125	21	Thursday	24 April 86	6.75	-0.25	7	Thursday
20 September 84	11.25	-0.1875	21	Thursday	22 May 86	6.8125	0.0625	28	Thursday
27 September 84	11	-0.25	7	Thursday	5 June 86	6.875	0.0625	14	Thursday
4 October 84	10.5625	-0.4375	7	Thursday	11 July 86	6.375	-0.5	36	Friday
11 October 84	10.5	-0.0625	7	Thursday	14 August 86	6.3125	-0.0625	34	Thursday
18 October 84	10	-0.5	7	Thursday	21 August 86	5.875	-0.4375	7	Thursday
8 November 84	9.5	-0.5	21	Thursday	4 December 86	6	0.125	105	Thursday
23 November 84	9	-0.5	15	Friday	30 April 87	6.5	0.5	147	Thursday
6 December 84	8.75	-0.25	13	Thursday	21 May 87	6.75	0.25	21	Thursday
20 December 84	8.5	-0.25	14	Thursday	2 July 87	6.625	-0.125	42	Thursday
27 December 84	8.125	-0.375	7	Thursday	27 August 87	6.75	0.125	56	Thursday
24 January 85	8.25	0.125	28	Thursday	3 September 87	6.875	0.125	7	Thursday
14 February 85	8.375	0.125	21	Thursday	4 September 87	7.25	0.375	1	Friday
21 February 85	8.5	0.125	7	Thursday	24 September 87	7.3125	0.0625	20	Thursday
21 March 85	8.625	0.125	28	Thursday	22 October 87	7.125	-0.1875	28	Thursday
28 March 85	8.5	-0.125	7	Thursday	28 October 87	7	-0.125	6	Wednesday
18 April 85	8.375	-0.125	21	Thursday	4 November 87	6.8125	-0.1875	7	Wednesday
25 April 85	8.25	-0.125	7	Thursday	28 January 88	6.625	-0.1875	85	Thursday
16 May 85	8.125	-0.125	21	Thursday	11 February 88	6.5	-0.125	14	Thursday
20 May 85	7.75	-0.375	4	Monday	30 March 88	6.75	0.25	48	Wednesday
9 May 88	7	0.25	40	Monday	9 January 91	6.75	-0.25	21	Wednesday
25 May 88	7.25	0.25	16	Wednesday	1 February 91	6.25	-0.5	23	Friday
22 June 88	7.5	0.25	28	Wednesday	8 March 91	6	-0.25	35	Friday
19 July 88	7.6875	0.1875	27	Tuesday	30 April 91	5.75	-0.25	53	Tuesday
8 August 88	7.75	0.0625	20	Monday	6 August 91	5.5	-0.25	98	Tuesday
9 August 88	8.125	0.375	1	Tuesday	13 September 91	5.25	-0.25	38	Friday
20 October 88	8.25	0.125	72	Thursday	31 October 91	5	-0.25	48	Thursday

17 November 88	8.3125	0.0625	28	Thursday	6 November 91	4.75	-0.25	6	Wednesday
22 November 88	8.375	0.0625	5	Tuesday	6 December 91	4.5	-0.25	30	Friday
15 December 88	8.6875	0.3125	23	Thursday	20 December 91	4	-0.5	14	Friday
29 December 88	8.75	0.0625	14	Thursday	9 April 92	3.75	-0.25	111	Thursday
5 January 89	9	0.25	7	Thursday	2 July 92	3.25	-0.5	84	Thursday
9 February 89	9.0625	0.0625	35	Thursday	4 September 92	3	-0.25	64	Friday
14 February 89	9.3125	0.25	5	Tuesday	4 February 94	3.25	0.25	518	Friday
23 February 89	9.5625	0.25	9	Thursday	22 March 94	3.5	0.25	46	Tuesday
24 February 89	9.75	0.1875	1	Friday	18 April 94	3.75	0.25	27	Monday
4 May 89	9.8125	0.0625	69	Thursday	17 May 94	4.25	0.5	29	Tuesday
6 June 89	9.5625	-0.25	33	Tuesday	16 August 94	4.75	0.5	91	Tuesday
7 July 89	9.3125	-0.25	31	Friday	15 November 94	5.5	0.75	91	Tuesday
27 July 89	9.0625	-0.25	20	Thursday	1 February 95	6	0.5	78	Wednesday
10 August 89	9	-0.0625	14	Thursday	6 July 95	5.75	-0.25	155	Thursday
18 October 89	8.75	-0.25	69	Wednesday	19 December 95	5.5	-0.25	166	Tuesday
6 November 89	8.5	-0.25	19	Monday	31 January 96	5.25	-0.25	43	Wednesday
20 December 89	8.25	-0.25	44	Wednesday	25 March 97	5.5	0.25	419	Tuesday
13 July 90	8	-0.25	205	Friday	29 September 98	5.25	-0.25	553	Tuesday
29 October 90	7.75	-0.25	108	Monday	15 October 98	5	-0.25	16	Thursday
14 November 90	7.5	-0.25	16	Wednesday	17 November 98	4.75	-0.25	33	Tuesday
7 December 90	7.25	-0.25	23	Friday	30 June 99	5	0.25	225	Wednesday
19 December 90	7	-0.25	12	Wednesday	24 August 99	5.25	0.25	55	Tuesday
16 November 99	5.5	0.25	84	Tuesday	6 November 02	1.25	-0.5	329	Wednesday
2 February 00	5.75	0.25	78	Wednesday	26 June 03	1	-0.25	231	Thursday
21 March 00	6	0.25	48	Tuesday	30 June 04	1.25	0.25	369	Wednesday
16 May 00	6.5	0.5	56	Tuesday	10 August 04	1.5	0.25	40	Tuesday
3 January 01	6	-0.5	232	Wednesday	21 September 04	1.75	0.25	41	Tuesday
1 February 01	5.5	-0.5	29	Thursday	10 November 04	2	0.25	49	Wednesday
20 March 01	5	-0.5	47	Tuesday	14 December 04	2.25	0.25	33	Tuesday
18 April 01	4.5	-0.5	29	Wednesday	2 February 05	2.5	0.25	49	Wednesday
15 May 01	4	-0.5	26	Tuesday	22 March 05	2.75	0.25	47	Tuesday
27 June 01	3.75	-0.25	42	Wednesday	3 May 05	3	0.25	41	Tuesday
21 August 01	3.5	-0.25	54	Tuesday	30 June 05	3.25	0.25	57	Thursday
17 September 01	3	-0.5	24	Monday	9 August 05	3.5	0.25	39	Tuesday
2 October 01	2.5	-0.5	14	Tuesday	20 September 05	3.75	0.25	41	Tuesday
6 November 01	2	-0.5	34	Tuesday	1 November 05	4	0.25	41	Tuesday
11 December 01	1.75	-0.25	34	Tuesday	13 December 05	4.25	0.25	41	Tuesday

Table 2.2.1: Calendar dates of federal funds rate target changes.

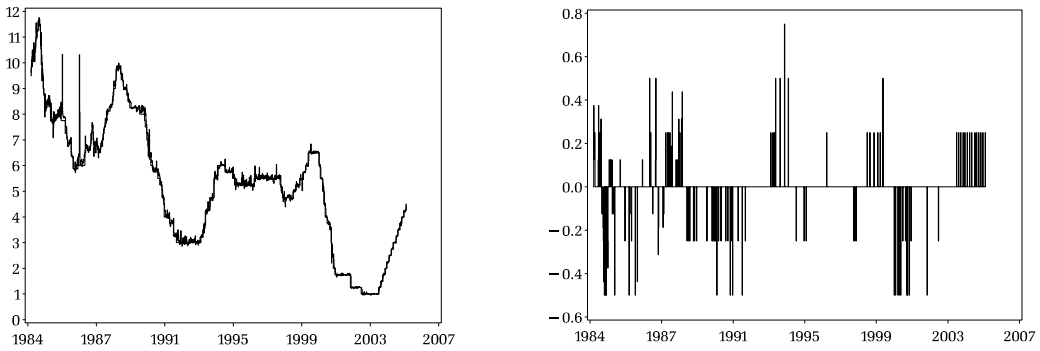

Figure 2.2.1: Federal funds rate target, effective federal funds rate and time series of target changes, March 1984 – January 2006. The left panel depicts the time series of the federal funds rate target (bold line) and the effective federal funds rate (FFR, thin line). All data are on weekly frequency. The right panel shows the time series of changes of the federal funds rate target.

Hamilton and Jordà (2002) transform the daily data into a weekly frequency by defining a seven-day period from Thursday until Wednesday. We extend the time series for the period May 2001 – January 2006 using data from the Federal Reserve Statistical Release. The distinguishing feature of the target time series is visible in the right panel of Figure 2.2.1. Target changes occur in discrete steps and are irregularly spaced in time. These characteristics of the data motivate their modeling as a DMPP.

As proposed by Hamilton and Jordà (2002), we consolidate the observed target changes ($y^{\#}$) into five categories (y) in the following way:[5]

$$y = \begin{cases} s_1 = -0.50 & \text{if} \quad -\infty < y^{\#} \leq -0.4375 \\ s_2 = -0.25 & \text{if} \quad -0.4375 < y^{\#} \leq -0.125 \\ s_3 = 0.00 & \text{if} \quad -0.125 < y^{\#} < 0.125 \\ s_4 = 0.25 & \text{if} \quad 0.125 \leq y^{\#} < 0.4375 \\ s_5 = 0.50 & \text{if} \quad 0.4375 \leq y^{\#} < \infty \end{cases} . \tag{2.1}$$

The last column of Table 2.2.2 reports the resulting unconditional frequency distribution of the consolidated target changes.

[5]Before 1990, the Fed changed the target in increments of 6.25 basis points and later in increments of 25 basis points. One could therefore argue that using four categories, omitting the category that collects the small target changes, may be more appropriate for modeling the evolution of the target. All analyses of the paper were also carried out using four categories (omitting the mid-state). However, the results do not change the conclusions. We therefore stick to the five categories classification to be comparable to Hamilton and Jordà's (2002) analysis. The four category results are available in Appendix A.1.

		previous target change					uncond.
		-0.5	-0.25	0	0.25	0.5	frequency
		Panel A: March 1984 – April 2001					
target change	-0.5	33.3	16.3	25.0	0.0	11.1	13.9
	-0.25	40.0	65.1	25.0	14.3	11.1	37.4
	0.0	20.0	4.7	33.3	2.9	22.2	10.4
	0.25	6.7	14.0	16.7	65.7	22.2	30.4
	0.5	0.0	0.0	0.0	17.1	33.3	7.8
		Panel B: March 1984 – January 2006					
target change	-0.5	38.1	19.1	25.0	0.0	11.1	15.2
	-0.25	42.9	61.7	25.0	10.4	11.1	34.1
	0.0	14.3	4.3	33.3	2.1	22.2	8.7
	0.25	4.8	14.9	16.7	75.0	22.2	35.5
	0.5	0.0	0.0	0.0	12.5	33.3	6.5

Table 2.2.2: Conditional and unconditional relative frequency distributions of consolidated target changes (in %).

The frequency distributions of the target changes conditional on the previous target change are also reported in Table 2.2.2. Panel A reports these conditional frequencies for the period from March 1984 – April 2001 (Hamilton and Jordà's 2002 original sample period). We use these data for estimation and in-sample evaluations. Panel B reports the results for the complete sample period, including May 2001 – January 2006, the period used for out-of-sample evaluations. The large numbers on the diagonal of the transition matrices indicate persistence in the target change sizes. Table 2.2.2 also shows the rare occurrence of the event that a large negative target change is followed by a large positive target change and vice versa.

2.3 Econometric Methodology

2.3.1 The ACH-ACM Model

Conceiving the evolution of the federal funds rate target as a discrete marked point process, we specify a model that accounts for the time between successive target changes and a model for the magnitude of the target change. We retain Hamilton and Jordà's (2002) autoregressive conditional hazard (ACH) as the model for the point process. The ACH is

combined with Russell and Engle's (2005) autoregressive conditional multinomial (ACM) model which accounts for the dynamics of the sequence of target change sizes. Let us start with a brief review of the ACH and introduce some notation before combining ACH and ACM to form an alternative DMPP model.

The ACH entails a autoregressive specification for the expected time between two events conditional on previous durations. Hamilton and Jordà (2002) specify the model in discrete time. Given the empirical setup described in the previous section, the smallest time interval between events is one week. Let us denote by τ_n the duration in number of weeks between the n^{th} and $(n+1)^{th}$ target change. To provide a link between event time and calendar time it is convenient to introduce a step function, denoted $N(t)$, which counts the number of target changes that occurred as of week t. $N(t)$ jumps by one if a target change occurs during week t and remains the same as in week $t-1$ if no target change occurs. The sequence of conditional expected durations $\psi_{N(t)} \equiv \mathbb{E}\left[\tau_{N(t)}|\tau_{N(t)-1}, \tau_{N(t)-2}, ...\right]$ is assumed to evolve as an autoregressive process,

$$\psi_{N(t)} = \alpha\tau_{N(t)-1} + \beta\psi_{N(t)-1} \quad , \tag{2.2}$$

where α and β are parameters. Equation (2.2) implies that the expected duration is updated only if a target change occurs.

The conditional probability of a target change during week t given the information available in $t-1$ is referred to as the hazard rate,

$$h_t = \mathbb{P}[N(t) \neq N(t-1)|\Upsilon_{t-1}] \quad . \tag{2.3}$$

If the information set Υ_{t-1} only consists of past durations, the hazard rate will remain the same until the next target change occurs. Hamilton and Jordà (2002) show that in this case hazard rate and conditional expected durations are inversely related,

$$h_t = \frac{1}{\psi_{N(t-1)}} \quad . \tag{2.4}$$

To allow for an impact of predetermined variables $\mathbf{z}$ observed in $t-1$, Hamilton and Jordà

(2002) specify a hazard rate that varies in calendar time, viz

$$h_t = \frac{1}{\psi_{N(t-1)} + \boldsymbol{\delta}' \mathbf{z}_{t-1}} \quad , \tag{2.5}$$

where $\boldsymbol{\delta}$ is a parameter vector. Equations (2.2) and (2.5) constitute the ACH model.

Hamilton and Jordà (2002) employ an ordered probit to model the time series of target change sizes and refer to Hausman et al.'s (1992) analysis of transaction price changes as the classic reference for the OP used in a high frequency time series context. However, following Russell and Engle (2005) who motivate their ACM model by arguing that the dynamics of discrete transaction price changes are better captured by a time series model specifically designed for discrete variables, we conjecture that the ACM may be also better suited to model the dynamics of target size changes.

In the following, we show how the ACM methodology can be adapted to model the size of target changes occurring at infrequent event times. Let us first define a binary indicator x_t which takes the value one if a target change occurs during week t and is zero otherwise. Denote by y_t the size of the target change in t. y_t is either zero for a week with no target change (if $x_t = 0$) or takes one of k different ordered outcomes $s_1 < s_2 < ... < s_k$ if $x_t = 1$. Let us further denote by π_{jn} the probability that the n^{th} target change is equal to s_j and collect the complete set of k probabilities in a vector $\tilde{\boldsymbol{\pi}}_n = (\pi_{1n}, \ldots, \pi_{kn})'$. Since the columns of $\tilde{\boldsymbol{\pi}}_n$ have to sum up to one, an arbitrary target change size, say the r^{th} category, can be defined as a reference category. The probability of observing a target change in the reference category can then be calculated as $\pi_{rn} = (1 - \boldsymbol{\imath}' \boldsymbol{\pi}_n)$ with $\boldsymbol{\imath}$ a $(k-1) \times 1$ vector of ones. $\boldsymbol{\pi}_n$ is a $(k-1) \times 1$ vector that results from deleting π_{rn} from $\tilde{\boldsymbol{\pi}}_n$. To indicate the size of the n^{th} target change, it is convenient to introduce a $k \times 1$ vector $\tilde{\mathbf{x}}_n$. Its j^{th} element is equal to one if the size of the n^{th} target change is equal to s_j, the other elements of $\tilde{\mathbf{x}}_n$ are zero. Finally, define the $(k-1) \times 1$ vector $\mathbf{x}_n$ which results from deleting the r^{th} element (indicating a target change size within the reference category) from $\tilde{\mathbf{x}}_n$.

Adapting the ACM methodology to the present application, we allow for autoregressive dynamics of the size of the target changes and account for the impact of predetermined previous week variables, $\mathbf{w}_{t-1}$, on the probabilities of observing one of the k possible target

change sizes:

$$\ell\left(\boldsymbol{\pi}_{N(t)}\right) = \mathbf{A}\left(\mathbf{x}_{N(t)-1} - \boldsymbol{\pi}_{N(t)-1}\right) + \mathbf{B}\ell\left(\boldsymbol{\pi}_{N(t)-1}\right) + \mathbf{D}\mathbf{w}_{t-1}x_t \quad . \tag{2.6}$$

$\mathbf{A}$ and $\mathbf{B}$ are $(k-1)\times(k-1)$ parameter matrices. $\mathbf{D}$ is a $(k-1)\times m$ parameter matrix where m denotes the number of predetermined variables (including a constant). The logistic link function $\ell\left(\boldsymbol{\pi}_{N(t)}\right) = \ln\left(\boldsymbol{\pi}_{N(t)}/(1-\boldsymbol{\imath}'\boldsymbol{\pi}_{N(t)})\right)$ ensures that the resulting probabilities lie within the unit interval. The probabilities $\boldsymbol{\pi}_{N(t)}$ can be recovered by computing

$$\boldsymbol{\pi}_{N(t)} = \frac{\exp\left[\mathbf{A}\left(\mathbf{x}_{N(t)-1} - \boldsymbol{\pi}_{N(t)-1}\right) + \mathbf{B}\ell\left(\boldsymbol{\pi}_{N(t)-1}\right) + \mathbf{D}\mathbf{w}_{t-1}x_t\right]}{1 + \boldsymbol{\imath}'\exp\left[\mathbf{A}\left(\mathbf{x}_{N(t)-1} - \boldsymbol{\pi}_{N(t)-1}\right) + \mathbf{B}\ell\left(\boldsymbol{\pi}_{N(t)-1}\right) + \mathbf{D}\mathbf{w}_{t-1}x_t\right]} \quad . \tag{2.7}$$

The term $\mathbf{x}_n - \boldsymbol{\pi}_n$ in Equation (2.6) can be interpreted as the innovation associated with the n^{th} target change.

The combination of Equations (2.2), (2.5) and (2.6) constitutes the ACH-ACM model as an alternative DMPP model for the federal funds rate target. Setting up the conditional likelihood function is straightforward. The probability of observing a target change of size y_t conditional on $\mathbf{w}_{t-1}$ and $x_t = 1$ can be written as $\tilde{\mathbf{x}}'_{N(t)}\tilde{\boldsymbol{\pi}}_{N(t)}$. This implies that the joint probability function of target change indicator x_t and target change size y_t is given by

$$\begin{aligned} f(x_t, y_t|\Upsilon_{t-1}; \boldsymbol{\theta}_{ACH}, \boldsymbol{\theta}_{ACM}) &= g(x_t|\Upsilon_{t-1}; \boldsymbol{\theta}_{ACH}) q(y_t|x_t, \Upsilon_{t-1}; \boldsymbol{\theta}_{ACM}) \\ &= \{h_t\}^{x_t}\{1-h_t\}^{(1-x_t)}\{\tilde{\mathbf{x}}'_{N(t)}\tilde{\boldsymbol{\pi}}_{N(t)}\}^{x_t} \quad , \end{aligned} \tag{2.8}$$

where the ACH parameters $\boldsymbol{\delta}, \alpha, \beta$ are collected in the vector $\boldsymbol{\theta}_{ACH}$ and the vectorized ACM parameter matrices $\mathbf{A}, \mathbf{B}, \mathbf{D}$ in $\boldsymbol{\theta}_{ACM}$.

The ACH-ACM log-likelihood function,

$$\mathscr{L}(\boldsymbol{\theta}_{ACH}, \boldsymbol{\theta}_{ACM}) = \sum_{t=1}^{T}\{x_t \ln(h_t) + (1-x_t)\ln(1-h_t)\} + \sum_{t=1}^{T} x_t \ln\left(\tilde{\mathbf{x}}'_{N(t)}\tilde{\boldsymbol{\pi}}_{N(t)}\right) \quad , \tag{2.9}$$

can be maximized with respect to the unknown parameters $(\boldsymbol{\theta}_{ACH}, \boldsymbol{\theta}_{ACM})$. If the parameters $(\boldsymbol{\theta}_{ACH}, \boldsymbol{\theta}_{ACM})$ are variation free as defined in Engle et al. (1983), and if the parameters of interest are contained in $\boldsymbol{\theta}_{ACH}$, then maximum likelihood estimates can be delivered by

maximizing:

$$\mathscr{L}_1(\boldsymbol{\theta}_{ACH}) = \sum_{t=1}^{T}\{x_t \ln(h_t) + (1-x_t)\ln(1-h_t)\} \tag{2.10}$$

(Engle 2000). Furthermore, if the parameters of interest are in $\boldsymbol{\theta}_{ACM}$, then x_t is weakly exogenous and maximum likelihood estimates are obtained from maximizing:

$$\mathscr{L}_2(\boldsymbol{\theta}_{ACM}) = \sum_{t=1}^{T} x_t \ln\left(\tilde{\mathbf{x}}'_{N(t)}\tilde{\boldsymbol{\pi}}_{N(t)}\right) \quad . \tag{2.11}$$

2.3.2 Evaluating Probability Function Forecasts of DMPP Models

The DMPP models considered in this paper deliver forecasts of the complete probability distribution of the forecast variable. It is thus tempting to use these models for value-at-risk and risk scenario analyses involving the federal funds rate target. This section proposes a methodology to evaluate the quality of the probability forecasts delivered by DMPP models.

A probability forecast is a probability function defined for a one-step or κ-period ahead observation of a (discrete) variable of interest, given the information at time t. The ACH-ACM one-step probability function forecast is readily available as a byproduct of the construction of the likelihood function in Equation (2.8),

$$f(i_{t+1}|\Upsilon_t) = \begin{cases} \mathbb{P}(i_{t+1} = i_t|\Upsilon_t) & = & 1 - h_{t+1} \\ \mathbb{P}(i_{t+1} = i_t + s_j|\Upsilon_t) & = & h_{t+1}\pi_{jN(t+1)} \quad j = 1, 2, ..., k \quad . \end{cases} \tag{2.12}$$

The probability function is zero for all other values of i_{t+1}. The expression for the probability function in Equation (2.12) is the same for ACH-OP with the only difference that the conditional probabilities $\pi_{jN(t+1)}$ originate from an OP model.

Let us briefly review the basic idea of Diebold et al.'s (1998) method for the evaluation of density forecasts and assume for the moment that the target is a continuous random variable. Denote by $\{f(i_t \mid \Upsilon_{t-1})\}$ a sequence of density forecasts and by $\{p(i_t \mid \Upsilon_{t-1})\}$ the sequence of true densities. Diebold et al. (1998) show that the correct density is weakly superior to all other forecasts. It will be preferred, in terms of expected loss, by all forecast users regardless of their loss functions. This suggests that forecasts can be evaluated by testing the null

hypothesis that the forecasting densities are correct, i.e. whether

$$\{f(i_t \mid \Upsilon_{t-1})\} = \{p(i_t \mid \Upsilon_{t-1})\} \quad . \tag{2.13}$$

At first sight, testing whether Equation (2.13) holds appears infeasible because $p(i_t \mid \Upsilon_{t-1})$ is unobserved. However, the distributional properties of the probability integral transform (PIT),

$$z_t = \int_{-\infty}^{i_t} f(u|\Upsilon_{t-1})du = F(i_t|\Upsilon_{t-1}) \quad , \tag{2.14}$$

provide the solution to this problem. Diebold et al. (1998) extend Rosenblatt's (1952) classic result by showing that under the null hypothesis the distribution of the sequence of probability transforms $\{z_t\}$ is iid $U(0,1)$.

In the present application we cannot rely on iid uniformity of the PIT sequence. The reason is that the PIT theorem only holds for continuous random variables. It applies to density function forecasts but not probability function forecasts. To address this problem, we adopt a methodology proposed by Denuit and Lambert (2005) who derive a discrete analog of the PIT theorem. For notational convenience assume that $s_{j+1} - s_j = c$ for $j = 1, 2, ..., k-1$. Equation (2.1) implies that $k = 5$ and $c = 0.25$. Transferring Denuit and Lambert's (2005) results, we "continue" the discrete target value i_t by adding an independent uniformly distributed random variable with support $[-c, 0]$, viz

$$i_t^* = i_t + c(u_t - 1) \quad , \tag{2.15}$$

where u_t is iid $U(0,1)$. Denuit and Lambert (2005) show that the PIT of the continued variable i_t^* can be computed as

$$z_t^* = F^*(i_t^*|\Upsilon_{t-1}) = F(i_t - c|\Upsilon_{t-1}) + f(i_t|\Upsilon_{t-1})u_t \quad . \tag{2.16}$$

The discrete analog of the PIT theorem states that z_t^* is $U(0,1)$ if the forecast probability function is correctly specified. Having obtained the z_t^* sequence, it is possible to apply the diagnostic tools proposed by Diebold et al. (1998) to evaluate probability function forecasts of the target.

The continuation principle extends to multi-step probability forecasts. Here the object of

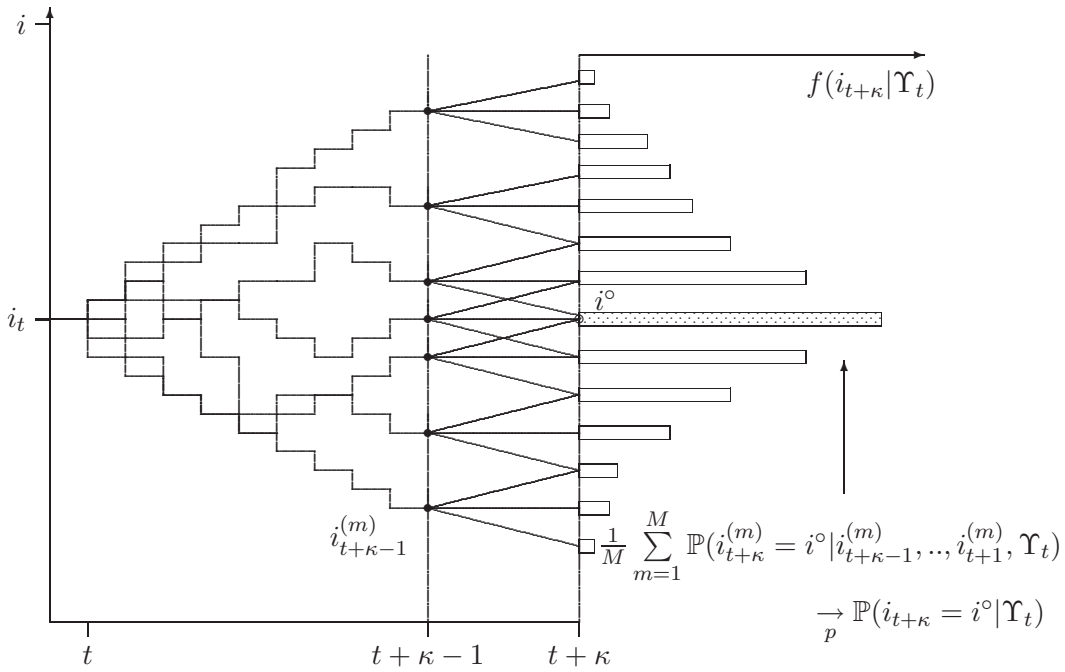

Figure 2.3.1: Simulation of probability forecasts for the ACH-ACM model. The figure illustrates the simulation procedure to obtain κ-step ahead probability function forecasts. Details are provided in the Appendix A.2. Starting from week t, the figure depicts $M = 7$ simulated sample paths that reach different federal funds rate target values in week $t + \kappa - 1$. Given the target value in $t + \kappa - 1$, three different target values can be attained in period $t + \kappa$. We assume three outcomes to keep the figure clear. The conditional probability of each outcome is computed via Equation (A.3) in Appendix A.2. Some values in $t + \kappa$, like i°, may have non-zero probabilities in many replications, while some values may be assigned non-zero probabilities only once, and some never. Summing conditional probabilities for each possible value over the M replications and dividing by M yields the estimate of the $t + \kappa$ period ahead probability forecast $f(i_{t+\kappa}|\Upsilon_t)$ sketched on the right hand side of the figure.

interest is the probability forecast $f(i_{t+\kappa}|\Upsilon_t)$ where $\kappa > 1$. We adopt a simulation strategy, since the analytic computation of the probability function $f(i_{t+\kappa}|\Upsilon_t)$ and the corresponding continued PIT sequence is numerically intractable. Figure 2.3.1 illustrates the procedure and shows how a κ-step ahead probability function forecast is conducted. The Appendix A.2 describes the procedure in detail as it applies to the ACH-ACM. The procedure works in an analogous manner for the ACH-OP.

Some remarks concerning the methodology are in order. First, the PIT sequence is not iid for $\kappa > 1$ even if the probability forecast is correct. It exhibits a MA($\kappa - 1$) autocorrelation structure. To account for this fact, we follow Diebold et al. (1998) and partition the continued PIT sequence into subseries for which we expect iid uniformity if the forecasts were correct. For instance, for correct two-step ahead probability forecasts, the subseries $\{z_1^*, z_3^*, z_5^*, \ldots\}$

and $\{z_2^*, z_4^*, z_6^*, \ldots\}$ should each be iid $U(0,1)$, although the complete series is not. Tests for iid uniformity are then based on the minimum or maximum of the test statistic within the subseries. Critical values are obtained by dividing the significance level by the number of subseries. Second, it is informative to augment formal tests of iid $U(0,1)$ with additional diagnostics. We adopt the iid uniformity test used in Bauwens et al. (2004) which compares the frequencies in the PIT histogram bins to expected values if the data were iid $U(0,1)$. However, iid uniformity tests alone are nonconstructive. When rejection of the null hypothesis occurs, they do not provide guidance about the reasons why. Diebold et al. (1998) suggest to augment formal tests of iid uniformity by visual inspection of histograms and autocorrelograms of the continued PIT sequences which assists in detecting particular forecast failures.

2.4 Estimation Results and Diagnostic Checks

2.4.1 Empirical Setup

This section describes alternative specifications of ACH-ACM models we consider for our comparison and the basic empirical setup. We split the data into an estimation sample and a part that is reserved for out-of-sample evaluation. The estimation period is the same as in Hamilton and Jordà (2002), March 1984 – April 2001. The out-of-sample period is May 2001 – January 2006. All models considered in this paper employ the ACH to model the duration between target changes. Since the sample period is the same, we adopt the ACH specifications reported by Hamilton and Jordà (2002). They document extensively the search procedure that leads to these specifications. We report the parameter estimates in Appendix A.3. Hamilton and Jordà's (2002) ordered probit estimation results can also be found in Appendix A.3.

The ACM specifications considered in this paper are based on the $k = 5$ target change categories defined in Equation (2.1). Hence, target changes occur with fixed increments $c = 0.25$. The third category (smallest absolute target changes) is chosen as the reference category. Hamilton and Jordà (2002) identify the previous week's spread between the six-month treasury bill rate on the secondary market ($TB6$) and the effective federal funds rate (FFR) as a predictor for next week's target change size. Accordingly, we use $\mathbf{w}_{t-1} = SP_{t-1}$, where $SP = TB6 - FFR$. The ACM variants considered in the following are thus restricted versions

of

$$\ell\left(\boldsymbol{\pi}_{N(t)}\right) = \mathbf{c} + \mathbf{A}\left(\mathbf{x}_{N(t)-1} - \boldsymbol{\pi}_{N(t)-1}\right) + \mathbf{B}\ell\left(\boldsymbol{\pi}_{N(t)-1}\right) + \mathbf{d}\mathrm{SP}_{t-1}x_t \tag{2.17}$$

$$= \begin{pmatrix} c_1 \\ c_2 \\ c_4 \\ c_5 \end{pmatrix} + \begin{bmatrix} a_{11}\, a_{12}\, a_{14}\, a_{15} \\ a_{21}\, a_{22}\, a_{24}\, a_{25} \\ a_{41}\, a_{42}\, a_{44}\, a_{45} \\ a_{51}\, a_{52}\, a_{54}\, a_{55} \end{bmatrix} \left(\mathbf{x}_{N(t)-1} - \boldsymbol{\pi}_{N(t)-1}\right) + \begin{bmatrix} b_{11}\, b_{12}\, b_{14}\, b_{15} \\ b_{21}\, b_{22}\, b_{24}\, b_{25} \\ b_{41}\, b_{42}\, b_{44}\, b_{45} \\ b_{51}\, b_{52}\, b_{54}\, b_{55} \end{bmatrix} \ell\left(\boldsymbol{\pi}_{N(t)-1}\right) + \begin{pmatrix} d_1 \\ d_2 \\ d_4 \\ d_5 \end{pmatrix} \mathrm{SP}_{t-1}x_t.$$

OP and ACM are estimated on the sequence of target change events. This time series contains much less observations than Russell and Engle (2005) had available for their original application of the ACM model. Their time series of transaction price changes contain many thousands of events while our sample contains only 115 target changes. Hence, parsimony of the ACM specification is called for to avoid in-sample over-fitting. To reduce the number of parameters, Russell and Engle (2005) advocate ACM specifications that imply symmetries in the responses to innovations $\mathbf{x}_{N(t)-1} - \boldsymbol{\pi}_{N(t)-1}$. They call a matrix $\mathbf{A}$ "response symmetric" if its elements are constrained in the following way:

$$\mathbf{A} = \begin{bmatrix} a_{11}\, a_{12}\, a_{14}\, a_{15} \\ a_{21}\, a_{22}\, a_{24}\, a_{25} \\ a_{41}\, a_{42}\, a_{44}\, a_{45} \\ a_{51}\, a_{52}\, a_{54}\, a_{55} \end{bmatrix} = \begin{bmatrix} a_1 & a_5 & a_8 & a_4 \\ a_2 & a_6 & a_7 & a_3 \\ a_3 & a_7 & a_6 & a_2 \\ a_4 & a_8 & a_5 & a_1 \end{bmatrix} \quad . \tag{2.18}$$

Persistence of the state probabilities $\boldsymbol{\pi}_n = (\pi_{1n}, \pi_{2n}, \pi_{4n}, \pi_{5n})'$ can be allowed for by a non-zero parameter matrix $\mathbf{B}$. For the sake of parsimony we focus our attention on diagonal $\mathbf{B}$ matrices. Consider an ACM specification that combines a response symmetric matrix $\mathbf{A}$ with a diagonal, but otherwise unrestricted diagonal matrix $\mathbf{B}$. We refer to this specification as *Partially Response Symmetric ACM* (PSACM). Another variant where the diagonal matrix $\mathbf{B}$ is also response symmetric is referred to as *Response Symmetric ACM* (SACM). Two more parsimonious specifications are also considered. The *Diagonal ACM* (DACM) restricts all off-diagonal elements in the matrices $\mathbf{A}$ and $\mathbf{B}$ equal to zero, but leaves the diagonal elements unrestricted. Our most parsimonious ACM specification, the *Response Symmetric Diagonal ACM(1,1)* (SDACM), constrains both $\mathbf{A}$ and $\mathbf{B}$ to be diagonal and response symmetric.

We also consider a simple non-dynamic alternative. The *UNConditional* model issues a probability forecast for a target change of a given size that is equal to the unconditional

relative frequencies reported in Table 2.2.2. The UNC model can be viewed as a special case of Equation (2.17) where all parameters except the vector **c** are set to zero.

2.4.2 Estimation Results and Goodness of Fit

Table 2.4.1 reports the value of the maximized log-likelihood, the Akaike information criterion (AIC) and a pseudo-R^2 measure (mean maximized likelihood value). We provide this information for the ACM specifications discussed above as well as for the ordered probit and UNC model.

	A	**B**	n_{par}	$\mathscr{L}_{max}$	AIC	R^2_{pseudo}
SACM	resp sym/8	resp sym diag/2	18	-114.8	2.31	0.369
PSACM	resp sym/8	diag/4	20	-114.5	2.34	0.370
OP	-	-	6	-137.1	2.49	0.303
SACM(1,0)	resp sym/8	-/0	16	-129.3	2.53	0.325
SDACM	resp sym diag/2	resp sym diag/2	12	-134.0	2.54	0.312
DACM	diag/4	diag/4	16	-133.1	2.60	0.314
UNC	-	-	4	-165.5	2.95	0.237

Table 2.4.1: Summary of estimation results. All models are estimated on March 1984 – April 2001 data. $\mathscr{L}_{max}$ is the maximized log-likelihood value, AIC is the Akaike information criterion computed as $-2 \cdot \frac{\mathscr{L}_{max}}{N(T)} + 2 \cdot \frac{n_{par}}{N(T)}$ where $N(T)$ denotes the total number of target change events and $R^2_{pseudo} = \exp\left(\frac{\mathscr{L}_{max}}{N(T)}\right)$. The models are sorted in ascending order by AIC. The total number of free parameters in each model is reported in the column n_{par}. The ACM specifications are special cases of

$$\ell\left(\boldsymbol{\pi}_{N(t)}\right) = \mathbf{c} + \mathbf{A}\left(\mathbf{x}_{N(t)-1} - \boldsymbol{\pi}_{N(t)-1}\right) + \mathbf{B}\ell\left(\boldsymbol{\pi}_{N(t)-1}\right) + \mathbf{d}\mathrm{SP}_{t-1}x_t \quad ,$$

where SP_{t-1} denotes the spread between the six-month treasury bill rate and the federal funds rate. The columns **A** and **B** provide information about the restrictions placed on the parameter matrices **A** and **B**. *resp sym* denotes a response symmetric and *diag* a diagonal structure of the respective matrix. The figure after the / gives the number of free parameters in the respective matrix.

Table 2.4.1 is sorted in ascending order by AIC, so models that appear on top of the list are preferred based on that criterion. The highest pseudo-R^2 (0.37) is delivered by the PSACM, in terms of AIC this model is ranked second. The SACM, which additionally imposes response symmetry of **B**, delivers an only marginally smaller pseudo-R^2 and is superior in terms of AIC. All ACMs deliver a higher pseudo-R^2 than the ordered probit, but due to its parsimony the OP is ranked third in terms of AIC. Parameter estimates of the five ACM specifications are reported in Table 2.4.2 and Table 2.4.3.

$\mathbf{c}'$	c_1	-2.407 (1.367)	c_2	1.278 (0.676)	c_4	1.766 (0.659)	c_5	-0.138 (0.630)
$\mathbf{A}$	a_{11}	1.434 (1.297)	a_{12}	2.729 (1.024)	a_{14}	-10.991 (3.723)	a_{15}	1.316 (2.355)
	a_{21}	1.050 (0.723)	a_{22}	3.532 (0.756)	a_{24}	1.911 (0.769)	a_{25}	0.777 (0.687)
	a_{41}	0.777 (0.687)	a_{42}	1.911 (0.769)	a_{44}	3.532 (0.756)	a_{45}	1.050 (0.723)
	a_{51}	1.316 (2.355)	a_{52}	-10.991 (3.723)	a_{54}	2.729 (1.024)	a_{55}	1.434 (1.297)
$\mathbf{B}$	b_{11}	0.470 (0.102)	b_{12}	0	b_{14}	0	b_{15}	0
	b_{21}	0	b_{22}	0.403 (0.125)	b_{24}	0	b_{25}	0
	b_{41}	0	b_{42}	0	b_{44}	0.430 (0.098)	b_{45}	0
	b_{51}	0	b_{52}	0	b_{54}	0	b_{55}	0.689 (0.069)
$\mathbf{d}'$	d_1	-1.128 (1.177)	d_2	0.453 (0.720)	d_4	1.772 (0.807)	d_5	3.705 (1.266)

Table 2.4.2: Maximum likelihood estimates of the PSACM. The estimation period is March 1984 – April 2001. Standard errors are reported in parentheses.

The ACM estimates are sensible from an economic point of view. For the sake of brevity let us focus on the PSACM results reported Table 2.4.2. As can be seen from the first row in the table, the estimates of the state specific constant are higher for the "inner states" (medium size positive or negative target changes) which is in accordance with the empirical frequency distribution of target changes reported in Table 2.2.2. The estimates of $\mathbf{d}$ (last row of Table 2.4.2) imply that an increase of the spread of the six-month treasury bill rate and the effective federal funds rate increases the probability of observing a subsequent positive target change (especially in the highest state) and reduces the probability of observing a negative target change next. This is in line with Hamilton and Jordà's (2002) ordered probit results (see Equation (A.9) in Appendix A.3).

Positive and significant estimates of the diagonal elements of the matrix $\mathbf{B}$ (third panel in Table 2.4.2) indicate persistence in the state probabilities. Persistence is highest in the categories indicating a large target change ($|y| = 0.5$). Medium size target changes ($|y| = 0.25$) are less persistent. The estimates of the matrix $\mathbf{A}$ (see second panel of Table 2.4.2) are plausible, but difficult to interpret directly. Due to the nonlinearity of the model, the marginal effect of a target change shock depends on the prevailing state probabilities. To give an idea of the economic significance of the parameter estimates, Figure 2.4.1 illustrates the response of

	SACM					SACM(1,0)			
$\mathbf{c}'$	-3.659	1.334	1.653	-2.037	$\mathbf{c}'$	-0.421	1.917	2.321	0.534
	(1.127)	(0.602)	(0.646)	(1.363)		(1.243)	(0.829)	(0.817)	(1.067)
$\mathbf{A}$	1.230	2.821	-17.010	2.393	$\mathbf{A}$	1.625	2.687	-1.068	-1.480
	(0.828)	(0.918)	(5.726)	(1.329)		(1.115)	(1.225)	(1.935)	(1.470)
	0.549	3.112	1.764	0.589		0.771	3.092	2.149	0.485
	(0.692)	(0.737)	(0.755)	(0.706)		(0.860)	(1.028)	(1.049)	(0.919)
	0.589	1.764	3.112	0.549		0.485	2.149	3.092	0.771
	(0.706)	(0.755)	(0.737)	(0.692)		(0.919)	(1.049)	(1.028)	(0.860)
	2.393	-17.010	2.821	1.230		-1.480	-1.068	2.687	1.625
	(1.329)	(5.726)	(0.918)	(0.828)		(1.470)	(1.935)	(1.225)	(1.115)
$\mathbf{B}$	0.494	0	0	0	$\mathbf{B}$	0	0	0	0
	(0.047)								
	0	0.375	0	0		0	0	0	0
		(0.077)							
	0	0	0.375	0		0	0	0	0
			(0.077)						
	0	0	0	0.494		0	0	0	0
				(0.047)					
$\mathbf{d}'$	-1.305	0.503	1.385	4.470	$\mathbf{d}'$	-0.613	0.396	1.794	3.384
	(0.969)	(0.727)	(0.811)	(0.978)		(1.207)	(0.841)	(0.884)	(1.116)

	SDACM					DACM			
$\mathbf{c}'$	0.316	0.720	0.899	0.585	$\mathbf{c}'$	0.223	0.814	0.901	0.752
	(0.507)	(0.512)	(0.545)	(0.526)		(0.241)	(0.563)	(0.492)	(0.421)
$\mathbf{A}$	0.847	0	0	0	$\mathbf{A}$	0.966	0	0	0
	(0.744)					(0.449)			
	0	1.402	0	0		0	1.255	0	0
		(0.338)					(0.436)		
	0	0	1.402	0		0	0	1.548	0
			(0.338)					(0.467)	
	0	0	0	0.847		0	0	0	0.917
				(0.744)					(0.708)
$\mathbf{B}$	0.605	0	0	0	$\mathbf{B}$	0.774	0	0	0
	(0.227)					(0.080)			
	0	0.597	0	0		0	0.551	0	0
		(0.145)					(0.224)		
	0	0	0.597	0		0	0	0.610	0
			(0.145)					(0.147)	
	0	0	0	0.605		0	0	0	0.438
				(0.227)					(0.151)
$\mathbf{d}'$	0.381	0.401	1.064	1.673	$\mathbf{d}'$	0.284	0.454	1.094	2.198
	(0.724)	(0.573)	(0.778)	(0.990)		(0.333)	(0.540)	(0.565)	(0.726)

Table 2.4.3: Maximum likelihood estimates of alternative ACM specifications. The estimation period is March 1984 – April 2001. Standard errors are reported in parentheses.

the state probabilities to a target change shock. Assuming identical initial state probabilities, and setting the prevailing spread equal to its sample mean, we use Equation (2.7) to compute the change of state probabilities due to a target change shock in the respective categories.

Figure 2.4.1 shows that a positive small target change shock increases the probability of observing another small positive target change, while the probability of observing a subsequent

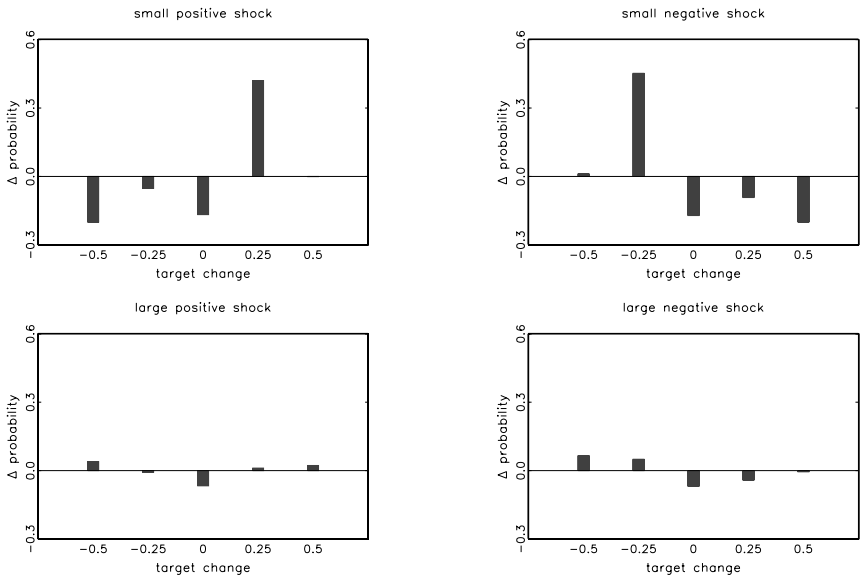

Figure 2.4.1: Effect of a target change shock on state probabilities. The figure depicts the change of state probabilities $\tilde{\boldsymbol{\pi}}_{N(t)} - \tilde{\boldsymbol{\pi}}_{N(t)-1}$ in response to a shock in the innovation term $\mathbf{x}_{N(t)-1} - \boldsymbol{\pi}_{N(t)-1}$. These computations are based on the PSACM estimates in Table 2.4.2. A small positive shock implies $\mathbf{x}_{N(t)-1} = (0,0,1,0)'$, a small negative shock implies $\mathbf{x}_{N(t)-1} = (0,1,0,0)'$, a large positive shock implies $\mathbf{x}_{N(t)-1} = (0,0,0,1)'$, and a large negative shock $\mathbf{x}_{N(t)-1} = (1,0,0,0)'$. To compute the state probabilities, SP_{t-1} is set to its sample mean, and $\boldsymbol{\pi}_{N(t)-1} = (0.2, 0.2, 0.2, 0.2)'$, i.e. we assume identical state probabilities prior to the shock.

negative target change is reduced. These effects are mirrored for small negative target change shocks. By contrast, large target change shocks do not induce strong effects on the state probabilities.

2.4.3 Comparing Short Term Interest Rate Forecasts

Hamilton and Jordà (2002) made a strong case for DMPP modeling by showing that their ACH-OP model improves the accuracy of the short term interest rate forecasts considerably compared to a forecast delivered by a standard vector autoregressive (VAR) model. Panel A of Table 2.4.4 reports the MSEs of one to twelve months forecasts of the effective federal funds rate using the VAR from Evans and Marshall (1998) and the ACH-OP model. These results are taken from Table 7 in Hamilton and Jordà (2002). For the ACH-OP (and the other DMPP models considered below) the model's target forecast is used as a forecast for the effective rate.[6] Interestingly, the simple DMPP specification ACH-UNC is also able to outperform the

[6]More precisely, the forecast of ACH-OP and the other DMPP models is the conditional expected target value for the respective horizon. We use the simulation techniques described in Section 2.3.2 to compute these target forecasts. Each DMPP model includes the same estimated ACH model reported in Equations (A.7) and (A.8)

	Panel A: Apr 1985-Apr 2001 (in-sample)					Panel B: May 2001-Jan 2006 (out-sample)				
h	UNC	OP	PSACM	DACM	VAR	UNC	OP	PSACM	DACM	BTMA
1	**0.03**	0.04	**0.03**	**0.03**	0.21	**0.01**	**0.01**	**0.01**	**0.01**	**0.01**
2	0.10	0.10	**0.08**	0.09	0.59	0.08	**0.05**	0.07	**0.05**	**0.05**
3	0.19	0.17	**0.14**	0.17	0.95	0.21	**0.13**	0.19	**0.13**	**0.13**
4	0.31	0.26	**0.23**	0.28	1.24	0.39	**0.24**	0.36	0.25	**0.24**
5	0.45	0.37	**0.34**	0.40	1.45	0.63	**0.40**	0.59	0.41	**0.40**
6	0.61	0.48	**0.47**	0.53	1.62	0.94	0.66	0.90	**0.64**	0.65
7	0.81	0.61	**0.61**	0.68	1.77	1.28	0.96	1.26	**0.91**	0.93
8	1.04	**0.77**	**0.77**	0.87	1.89	1.66	1.30	1.66	**1.21**	1.24
9	1.26	**0.92**	0.94	1.05	2.00	2.07	1.67	2.07	**1.52**	1.58
10	1.48	**1.08**	1.11	1.22	2.18	2.51	2.09	2.52	**1.88**	1.96
11	1.70	**1.24**	1.28	1.39	2.42	2.96	2.55	2.98	**2.26**	2.38
12	1.92	**1.39**	1.44	1.56	2.70	3.40	3.01	3.45	**2.64**	2.80

	Panel C: May 2001-May 2004					Panel D: Jun 2004-Jan 2006				
h	UNC	OP	PSACM	DACM	BTMA	UNC	OP	PSACM	DACM	BTMA
1	**0.01**	**0.01**	0.02	**0.01**	**0.01**	0.01	**0.00**	**0.00**	**0.00**	**0.00**
2	0.08	**0.06**	0.11	**0.06**	**0.06**	0.08	0.02	**0.01**	0.03	0.02
3	0.22	**0.17**	0.27	**0.17**	**0.17**	0.20	0.06	**0.04**	0.07	0.06
4	0.40	0.31	0.51	**0.30**	0.31	0.37	0.11	**0.08**	0.15	0.12
5	0.65	**0.51**	0.84	**0.51**	**0.51**	0.59	0.19	**0.13**	0.24	0.20
6	0.99	0.87	1.28	**0.80**	0.84	0.84	0.27	**0.20**	0.35	0.28
7	1.37	1.29	1.80	**1.15**	1.22	1.11	0.36	**0.28**	0.46	0.37
8	1.80	1.76	2.35	**1.54**	1.65	1.41	0.46	**0.37**	0.59	0.47
9	2.25	2.27	2.93	**1.95**	2.12	1.72	0.56	**0.47**	0.73	0.58
10	2.75	2.85	3.56	**2.42**	2.64	2.06	0.68	**0.58**	0.88	0.71
11	3.26	3.49	4.22	**2.93**	3.22	2.40	0.79	**0.68**	1.02	0.82
12	3.77	4.16	4.89	**3.45**	3.81	2.72	0.90	**0.80**	1.14	0.93

Table 2.4.4: Mean squared errors for one to twelve months forecasts of the effective federal funds rate, in- and out-of-sample. In-sample OP and VAR results are taken from Table 7 in Hamilton and Jordà (2002). All models are estimated on March 1984 – April 2001 data and use the ACH specification in Equations (A.7) and (A.8) as the model for the point process. The column titled BTMA reports the MSEs of a Bayesian type model averaging of ACH-OP and ACH-DACM. The forecast horizon h in months is given in the first column. Bold faced numbers indicate the lowest MSE at the respective horizon.

VAR which emphasizes the point that modeling the time between target changes matters.

in Appendix A.3. The VAR model is based on monthly data, while ACH-OP and the other models considered in this paper are estimated on weekly data. To ensure comparability and avoid giving the DMPP models a head start, we follow Hamilton and Jordà (2002) and compute the DMPP and VAR forecasts based on the same conditioning information. Specifically, the monthly DMPP models' forecasts are based on end-of calendar month information even if newer weekly data would be available. Furthermore, we need forecast values of the spread (SP) to compute the target forecast of ACH-OP and ACH-ACM models. For that purpose, we use the following specification estimated by Hamilton and Jordà (2002):

$$\underset{}{\mathrm{SP}_t} = \underset{(0.032)}{0.129} + \underset{(0.083)}{0.228}\, i_t - \underset{(0.083)}{0.267}\, i_{t-1} + \underset{(0.023)}{0.723}\, \mathrm{SP}_{t-1} \tag{2.19}$$

with standard errors in parentheses.

We extend the analysis by letting ACH-ACM specifications enter the competition. Let us focus on the ACH-PSACM and the ACH-DACM. The other ACH-ACM specifications deliver comparable results. Panel A of Table 2.4.4 shows that the ACH-PSACM improves the mean squared errors of the ACH-OP target forecast at forecast horizons up to eight months while the ACH-DACM improves on the OP at forecast horizons up to four months. The ACH-ACM approach seems particularly useful for forecast horizons of up to six months. The ACH-OP is preferable beyond a horizon of three quarters of a year. Averaged over the first three months, the ACH-OP MSE is reduced by 14% by the ACH-DACM model and by 22% by the ACH-PSACM model. Although this is less impressive than the 82% MSE reduction of the ACH-OP compared to the VAR, it is still a favorable result for the approach proposed in this paper.

2.4.4 In-sample Probability Forecasts

The probability forecast evaluation techniques outlined in Section 2.3.2 can be conveniently used for diagnostic checking. Table 2.4.5 reports results of tests for iid uniformity of the continued PIT sequences produced by ACH-OP, ACH-PSACM and ACH-DACM. The probability forecast horizon κ ranges from one to 60 weeks. The analysis is performed in-sample, i.e. the period used for parameter estimation is also the period for which we compute the in-sample probability forecast. We employ the iid uniformity test suggested by Bauwens et al. (2004). Their test compares the number of observations in the bins of the PIT histogram with the expected values if the z^* sequence would indeed be iid $U(0,1)$. The caption of Table 2.4.5 explains computational details.

Comparing the values of the in-sample test statistics (left panel of Table 2.4.5) with the $10\%/\kappa$ critical values, we cannot reject the hypothesis that the three DMPP models deliver correct probability forecasts. It should be noted, however, that the number of observations on which the test is based shrinks for longer forecast horizons. As already explained above, the z^* sequence exhibits a MA$(\kappa-1)$ autocorrelation structure even if the probability forecast is correct. The necessary thinning into κ subseries which are iid $U(0,1)$ under the null reduces the number of observations and power. For instance, the 52 week ahead forecast evaluation is based on about 16 observations in each subseries. Diebold et al. (1998) advocate the use of autocorrelograms and histograms of the PIT sequences as diagnostic tools to detect

	IN-SAMPLE Apr 1985 – Apr 2001			OUT-OF-SAMPLE May 2001 – Jan 2006			CRITICAL VALUES		
κ	OP	PSACM	DACM	OP	PSACM	DACM	$\frac{1\%}{\kappa}$	$\frac{5\%}{\kappa}$	$\frac{10\%}{\kappa}$
1	8.6	5.7	6.2	16.9	12.9	18.0	21.7	16.9	14.7
4	11.5	14.1	15.3	28.6	23.5	19.6	25.5	21.0	19.0
8	12.3	12.3	9.7	17.7	46.1	24.8	27.3	23.0	21.0
12	12.9	12.0	20.1	18.0	23.0	16.0	28.4	24.1	22.2
16	13.4	12.6	15.7	17.7	25.7	15.0	29.1	24.9	23.0
20	19.7	18.8	18.8	18.0	21.3	9.7	29.7	25.5	23.6
24	14.8	19.5	19.5	18.0	24.0	12.0	30.1	25.9	24.1
26	16.1	19.2	16.1	16.6	18.8	12.1	30.3	26.2	24.3
28	14.1	14.8	20.3	14.5	20.3	12.0	30.5	26.4	24.5
32	15.5	20.2	14.8	14.4	20.1	14.4	30.9	26.7	24.9
36	14.0	15.7	14.8	14.0	14.0	14.0	31.2	27.0	25.2
40	13.0	19.0	14.0	14.0	17.3	14.0	31.4	27.3	25.5
44	16.3	17.3	14.2	17.0	17.0	13.0	31.7	27.5	25.7
48	25.9	22.4	14.2	17.0	17.0	17.0	31.9	27.8	25.9
52	17.8	15.2	16.5	21.0	21.0	21.0	32.1	28.0	26.2
60	17.0	10.8	15.5	21.0	21.0	21.0	32.4	28.4	26.5

Table 2.4.5: Results of iid uniformity test for continued PIT sequence. For each forecast horizon κ the continued PIT sequence is split into κ subseries which are iid $U(0,1)$ under the null hypothesis of a correct probability forecast. Bauwens et al.'s (2004) test statistic for iid uniformity is computed for each subseries. The test is based on the result that under the null of iid $U(0,1)$ behavior of the (continued) PIT sequence the joint distribution of the heights of the PIT histogram is multinomial, i.e. $f(n_i) = \binom{n}{n_i} p^{n_i} (1-p)^{n-n_i}$ where n gives the number of observations (in each subseries), n_i the number of observations in the i^{th} histogram bin and $p = 1/m$ with m the number of histogram bins. We use $m = 10$. The statistic $\sum_{i=1}^{m} \frac{(n_i - np)^2}{np}$ is under the null hypothesis asymptotically $\chi^2(m-1)$ distributed. The table reports the largest test statistic computed from κ subseries. The critical values are computed by dividing the significance levels by κ. The forecast periods are April 1985 – April 2001 (in-sample) and May 2001 – January 2006 (out-of-sample), respectively.

specification problems associated with a model's density forecasts. For instance, $\cup$-shaped PIT histogram would indicate that we would observe too many large and small future target values compared to what is predicted by the model. Significant serial correlation of the PIT series would indicate that the model is not able to account properly for the dynamics of the federal funds rate target. Figure 2.4.2 depicts ten-bin histograms of the continued PIT sequence for forecast horizons ranging from one to 60 weeks implied by the ACH-PSACM. The histograms for ACH-OP and ACH-DACM look quite similar. Due to space constraints we refrain from their presentation. The histogram is based on the original continued PIT sequence only for the one week horizon (i.e. one-step forecast). For multi-step forecasts ($\kappa > 1$), we plot the minimum and the maximum relative frequency of the thinned κ subseries in each of the ten histogram bins.

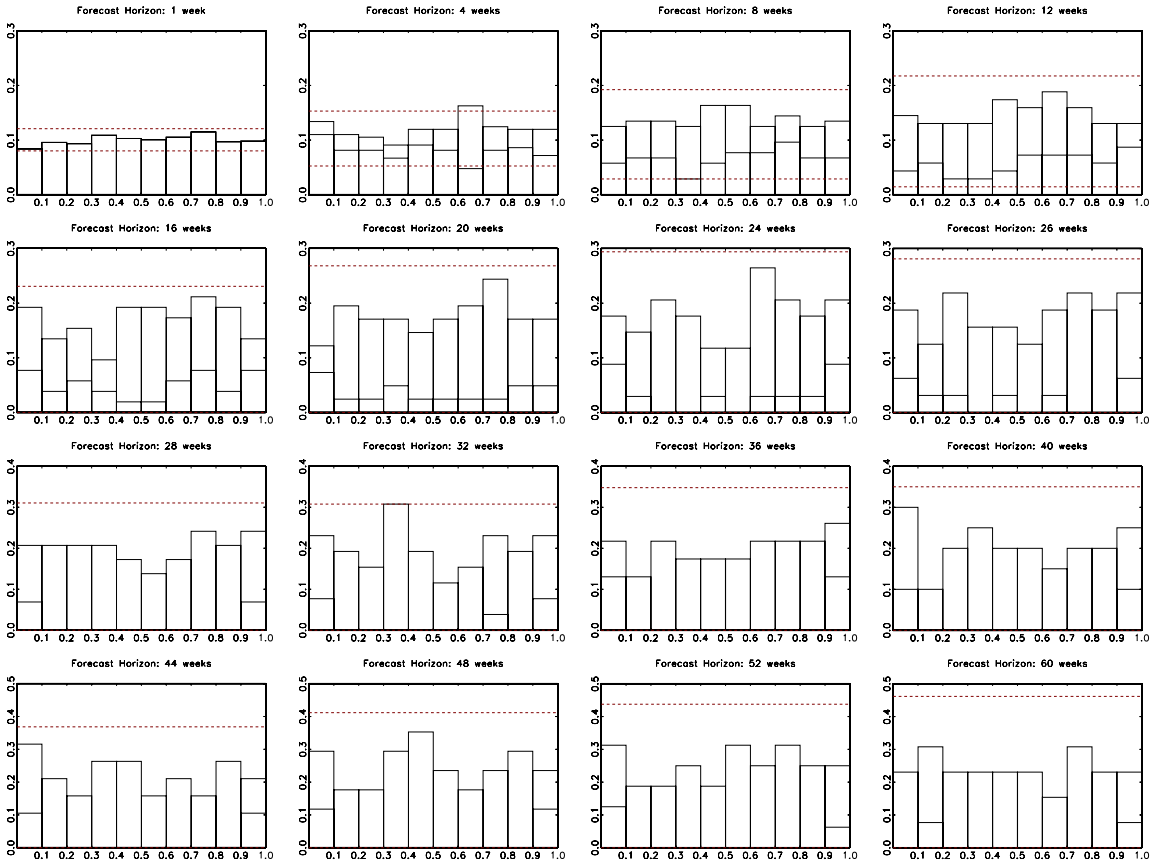

Figure 2.4.2: Histograms of the continued PIT sequence: ACH-PSACM, in-sample probability forecast. The figure shows ten-bin histograms of the continued PIT sequence for forecast horizons ranging from $\kappa = 1$ to 60 weeks. For $\kappa > 1$ the data are thinned into κ subseries which are iid $U(0,1)$ under the null hypothesis of a correct probability forecast. The horizontal solid lines show the minimum and the maximum relative frequency of the κ subseries in each of the ten histogram bins. Upper and lower bound (displayed in horizontal dashed lines) of the 95% confidence interval are computed from the $0.025/\kappa$ and $0.975/\kappa$ quantiles of a binomial distribution with $p = 0.1$ and number of draws equal to n, where n is the number of observations in each subseries. The estimation period is March 1984 – April 2001, the forecast evaluation period is April 1985 – April 2001.

Overall, the results are quite favorable for the DMPP approach. The histograms do not hint at violations of iid uniformity of the continued PIT sequence. As noted above, however, this finding should be taken with a pinch of salt. As a consequence of the thinning into subseries, the confidence bounds become wide for longer forecasting horizons (beyond nine months), so the power of the diagnostic test shrinks.

Figure 2.4.3 plots autocorrelograms of the continued PIT sequence for ACH-PSACM. Due to the necessary thinning, into subseries we refrain from computing the autocorrelograms beyond a 32 weeks horizon. The autocorrelograms suggest that some dynamics of the

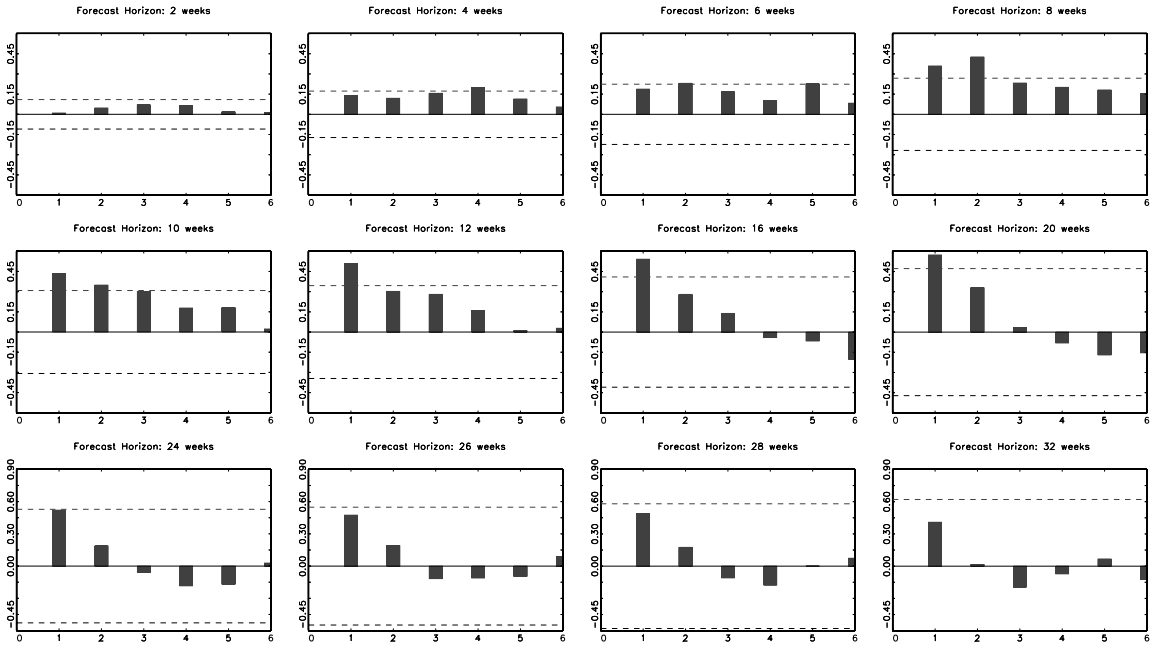

Figure 2.4.3: Autocorrelations of the continued PIT sequence: ACH-PSACM, in-sample forecast. For each forecast horizon κ the z^* sequences are split into κ subseries. The figures show the maximal autocorrelations of the κ subseries. The horizontal lines superimposed on the autocorrelograms mark the 95% confidence intervals. The estimation period is March 1984 – April 2001, the forecast period is April 1985 – April 2001.

probability forecasts at two to four month horizons are not entirely captured, so there seems to be room for further model improvement. The same holds true for ACH-OP and ACH-DACM for which the autocorrelograms look quite similar.

2.4.5 Out-of-sample Forecast Evaluation

This section extends the in-sample view by conducting an analysis of the out-of-sample point and probability forecast performance of DMPP models for the federal funds rate target. Section 2.4.3 analyzed the in-sample performance of DMPP models when employed for forecasting the effective federal funds rate. The out-of-sample results of such an exercise are presented in Panels B, C and D of Table 2.4.4.

The good in-sample forecasting performance of DMPP models for the target extends out-of-sample. Over the complete out-of-sample period May 2001 – January 2006 the parsimoniously specified ACH-DACM model delivers the best forecasts. It produces the smallest MSE in 10 out of 12 months (see Panel B of Table 2.4.4). At forecast horizons

up to three months, the out-of-sample MSEs of the best model (ACH-DACM) are even smaller than those of the best in-sample model (ACH-PSACM). In- and out-of-sample forecast performances remain comparable for up to six months horizons. Beyond, the out-of-sample MSEs become bigger than their in-sample counterparts. However, the out-of-sample ACH-DACM MSEs are still considerably smaller than the in-sample VAR MSEs. The results reported in Panel B of Table 2.4.4 show that using conditioning information matters. The ACH-UNC model's out-of-sample performance is the worst of all models.

None of the models is superior in terms of predicting the target equally well at all forecast horizons and subsamples (see Panels B, C and D of Table 2.4.4). A fashionable approach in such a situation is to apply (Bayesian) model averaging. We do not plunge into a fully fledged Bayesian analysis. Instead, we apply a Bayesian type model averaging procedure. Specifically, we refrain from accounting for parameter uncertainty and assume equal prior model probabilities. Model weights are then formed by the (predictive) likelihoods of the models. We describe the methodology in greater detail in the Appendix A.4. Table 2.4.4 reports in the columns titled BTMA the MSEs of a combination of ACH-OP and ACH-DACM. Generally, Bayesian model averaging entails probing all possible model combinations. Because of the computationally intensive simulation of multi-step forecasts, we focus on some selected combinations (see Appendix A.5 for additional modeling averaging MSEs results). The OP-DACM combination turns out to be the most successful. It does a good job in the overall out-sample and the performance in both subsamples is also satisfactory. However, the best models in each forecast sample cannot be outperformed.

The out-of-sample probability of forecast evaluation broadly confirms the conclusions of Section 2.4.4. Comparing the values of iid uniformity test using the complete out-of-sample continued PIT sequences of ACH-DACM with the $5\%/\kappa$ critical values (right panel of Table 2.4.5), the hypothesis that the probability forecasts are correct cannot be rejected (for $\kappa = 1$ and $\kappa = 8$ at $1\%/\kappa$). Figure 2.4.4 depicts the histograms of the continued PIT sequence for the ACH-DACM. As the corresponding figures for ACH-PSACM and ACH-OP look similar, they are omitted for the sake of brevity.

The ACH-DACM histograms are more ragged compared to their in-sample counterparts, but not enough to be at odds with the hypothesis that the out-of-sample probability forecast is correct. It should be noted, however, that the small sample problem at longer forecast

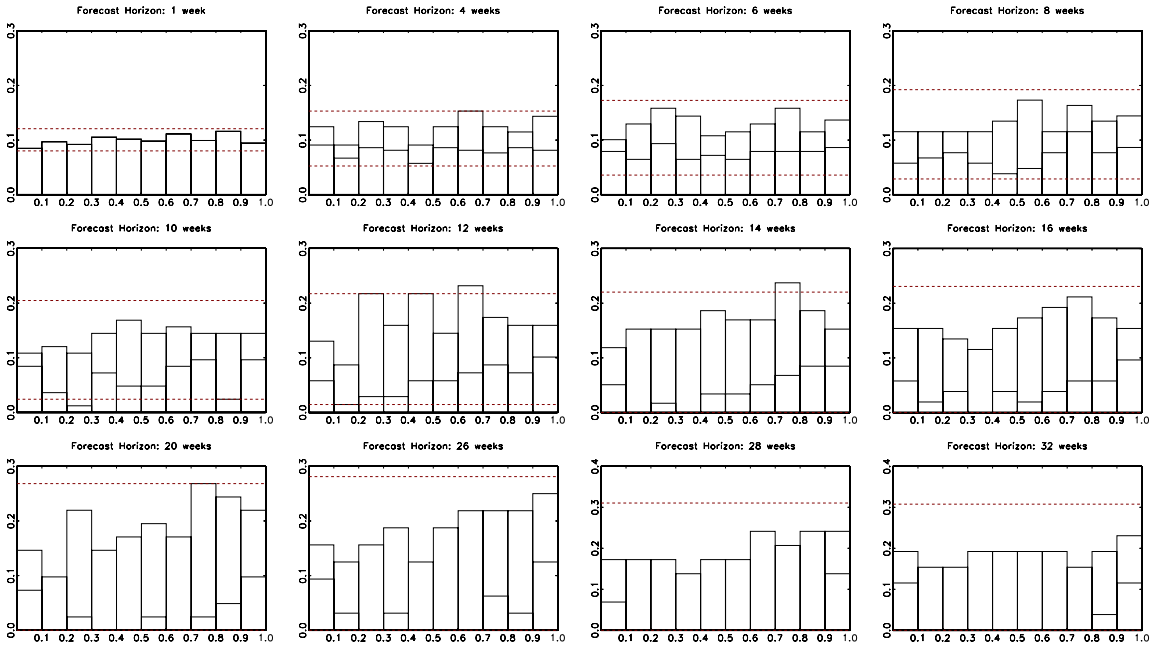

Figure 2.4.4: Histograms of the continued PIT sequence: ACH-DACM, out-of-sample probability forecast. The estimation period is March 1984 – April 2001, the forecast evaluation period is May 2001 – January 2006. See caption of Figure 2.4.2 for explanations.

horizons discussed above is aggravated as we have fewer data available for out-of-sample evaluation.

2.5 Conclusion and Outlook

Forecasts of the federal funds rate target are of key interest for investors and financial institutions. Hamilton and Jordà's (2002) finding that the quality of federal funds rate target forecasts can be substantially improved when the sequence of target changes is modeled as a discrete marked point processes (DMPP) highlighted the importance of this econometric approach for a central issue in monetary economics.

This paper introduces a new DMPP specification which combines the autoregressive conditional hazard (ACH) model put forth by Hamilton and Jordà (2002) and the autoregressive conditional multinomial (ACM) model developed by Russell and Engle (2005). From a methodological point of view, the ACM model seems the natural choice, as it is formulated as a time series model for discrete variables. This is precisely what is needed for modeling the evolution of the federal funds rate target.

The ACM offers a great amount of flexibility in accounting for target change dynamics. However, even in Russell and Engle's (2005) application of the ACM to transaction price changes (implying a much larger sample than we have at hand) the authors recommend exploring parsimonious specifications in order to avoid overfitting. By imposing diagonal and response-symmetric structures of the ACM parameter matrices one can provide straightforward testable restrictions. We show that such parsimonious ACH-ACM specifications deliver improved results in terms of goodness of fit and in-sample forecast performance. Hamilton and Jordà's (2002) study made a strong case for DMPP models showing their excellent in-sample forecast performance. We show that out-of-sample results are also promising. Parsimoniously specified ACH-ACM models do a good job and Bayesian type model averaging of ACH-ACM and ACH-OP robustifies the out-of-sample point forecast performance.

DMPP models deliver forecasts of the complete probability distribution of future target values. Hence, it seems intriguing to employ them for value-at-risk and risk scenario analysis. For that purpose, this paper proposes a methodology to evaluate the probability function forecasts of DMPP models. We show that parsimoniously specified DMPP models deliver useful probability function forecasts of the target for horizons up to six months. The available time series is yet too short to assess probability forecasts beyond that horizon.

Avenues for further research stretch in several directions. First, model averaging could be extended from point to probability forecasts within a fully fledged Bayesian framework. Second, it seems promising to adapt the specification test proposed by Hong and Li (2005) to evaluate multi-period probability forecasts of DMPP models. Their framework accounts for parameter estimation uncertainty within Diebold et al.'s (1998) density forecast evaluation. Third, the forecast evaluation could be extended towards the federal funds rate target's worldwide counterparts, like the European Central Bank's rate of main refinancing operations or the Repo Rate set by the Bank of England. These variables exhibit the same time series properties: discrete interest rate changes with irregularly spaced time intervals in between. An evaluation of the performance of the ACH-ACM framework using these data is left for further research.

Appendix A

A.1 Four Category ACH-ACM Model

Hamilton and Jordà (2002) introduce the ACH-OP model for target changes that are classified into five categories (see Section 2.2). In our analysis we focus on five category ACH-OP and ACH-ACM models as well to be comparable to their study. However, since the Fed changed its monetary policy implementation in 1990 and increased the minimum target change from 6.25 to 25 basis points, one could argue that using four categories, may be more appropriate for modeling the evolution of the target. In order to show the robustness of our conclusions, we consolidate the observed target changes ($y^{\#}$) into four categories in the following way,

$$y = \begin{cases} s_1 = -0.50 & \text{if} \quad -\infty < y^{\#} < -0.25 \\ s_2 = -0.25 & \text{if} \quad -0.25 \leq y^{\#} < 0 \\ s_3 = 0.25 & \text{if} \quad 0 < y^{\#} \leq 0.25 \\ s_4 = 0.50 & \text{if} \quad 0.25 < y^{\#} < \infty \end{cases} , \tag{A.1}$$

and carry all analyses of the paper out using four categories (omitting the mid-state). The reference state in the ACH-ACM and ACH-OP model is s_3.

A.1.1 Four Category Estimation Results

Using the consolidated target changes in Equation (A.1) for the ordered probit estimation (March 1, 1984 -April 26, 2001), we obtain for the latent target change equation:

$$y^*_{N(t)} = \underset{(0.407)}{2.544}\, y_{N(t)-1} + \underset{(0.211)}{0.541}\, \mathrm{SP}_{t-1} \cdot x_t \quad . \tag{A.2}$$

Standard errors are reported in parenthesis.

	A	**B**	n_{par}	$\mathscr{L}_{max}$	*AIC*	R^2_{pseudo}
ACM	full/9	diag/3	18	-104.9	2.14	0.402
OP	-	-	5	-120.6	2.19	0.350
SACM	resp sym/7	resp sym diag/2	15	-114.6	2.26	0.369
PSACM	resp sym/7	diag/3	16	-113.8	2.26	0.372
DACM	diag/3	diag/3	12	-128.2	2.44	0.328
UNC	-	-	3	-147.2	2.62	0.278

Table A.1.1: Summary of four category ACM estimation results. All four category ACM models, the unconditional and ordered probit model are estimated on March 1984 – April 2001 data. $\mathscr{L}_{max}$ is the maximized log-likelihood value, *AIC* is the Akaike information criterion computed as $-2 \cdot \frac{\mathscr{L}_{max}}{N(T)} + 2 \cdot \frac{n_{par}}{N(T)}$ where $N(T)$ denotes the total number of target change events and $R^2_{pseudo} = \exp\left(\frac{\mathscr{L}_{max}}{N(T)}\right)$. The models are sorted in ascending order by AIC. The total number of free parameters in each model is reported in the column n_{par}. The ACM specifications are special cases of

$$\ell\left(\boldsymbol{\pi}_{N(t)}\right) = \mathbf{c} + \mathbf{A}\left(\mathbf{x}_{N(t)-1} - \boldsymbol{\pi}_{N(t)-1}\right) + \mathbf{B}\ell\left(\boldsymbol{\pi}_{N(t)-1}\right) + \mathbf{d}\mathrm{SP}_{t-1}x_t \quad ,$$

where SP_{t-1} denotes the spread between the six-month treasury bill rate and the federal funds rate. The columns **A** and **B** provide information about the restrictions placed on the parameter matrices **A** and **B**. *resp sym* denotes a response symmetric and *diag* a diagonal structure of the respective matrix. The figure after the / gives the number of free parameters in the respective matrix.

We provide in Table A.1.1 summary estimation results for four categories ACM specifications, the ordered probit and the unconditional model. The table reports the value of the maximized log-likelihood, the Akaike information criterion (AIC) and a pseudo-R^2 measure (see Section 2.4.1 for model acronyms). Table A.1.1 is sorted in ascending order by AIC, so models that appear on top of the list are preferred based on that criterion. The highest pseudo-R^2 and lowest AIC criterion is delivered by a fully specified ACM model. Parameter estimates of the four ACM specifications using the target change classification of (A.1) are reported in Table A.1.2.

As for the five category ACM model, we find economic sensible parameter estimates. Table A.1.2 reports that an increase of the spread of the six-month treasury bill rate and the effective federal funds rate implies an increase in the probability of observing a subsequent positive target change and a reduction in the probability of observing a negative target change next. This is in line with the ordered probit results in (A.2). The significant estimates of the diagonal elements of the matrix **B** indicate persistence in the state probabilities. Higher categories ($|y| = 0.5$) indicate a large target change. Medium size target changes ($y = -0.25$) are less persistent.

	ACM				DACM		
$\mathbf{c}'$	-1.282	-0.879	-0.414	$\mathbf{c}'$	-2.428	-0.310	-0.149
	(0.780)	(0.421)	(0.492)		(0.769)	(0.279)	(0.164)
	5.011	4.785	2.778		-0.386	0	0
	(1.595)	(1.320)	(1.771)		(0.722)	0	
$\mathbf{A}$	2.943	3.065	-0.012	$\mathbf{A}$	0	1.367	0
	(1.056)	(0.875)	(0.569)			(0.400)	
	-2.695	-7.798	-0.826		0	0	0.944
	(4.058)	(4.261)	(0.845)				(0.608)
	0.511	0	0		-0.625	0	0
	(0.150)				(0.194)		
$\mathbf{B}$	0	0.497	0	$\mathbf{B}$	0	0.592	0
		(0.113)				(0.165)	
	0	0	0.746		0	0	0.745
			(0.077)				(0.146)
$\mathbf{d}'$	-1.110	-1.633	1.758	$\mathbf{d}'$	-2.004	-0.722	0.385
	(1.021)	(0.643)	(0.769)		(0.719)	(0.417)	(0.303)
	SACM				PSACM		
$\mathbf{c}'$	-0.746	-0.622	-0.966	$\mathbf{c}'$	-0.636	-0.534	-0.511
	(0.460)	(0.348)	(0.643)		(0.521)	(0.357)	(0.535)
	1.937	1.827	-1.050		1.777	1.941	-0.850
	(0.710)	(0.886)	(1.404)		(0.760)	(0.986)	(1.217)
$\mathbf{A}$	2.289	2.595	-0.259	$\mathbf{A}$	2.230	2.683	-0.348
	(0.879)	(0.732)	(0.783)		(0.910)	(0.833)	(0.837)
	-1.050	-6.044	1.937		-0.850	-4.494	1.777
	(1.404)	(3.314)	(0.710)		(1.217)	(2.733)	(0.760)
	0.589	0	0		0.482	0	0
	(0.125)				(0.277)		
$\mathbf{B}$	0	0.472	0	$\mathbf{B}$	0	0.465	0
		(0.128)				(0.156)	
	0	0	0.589		0	0	0.732
			(0.125)				(0.147)
$\mathbf{d}'$	-0.765	-1.108	1.183	$\mathbf{d}'$	-0.639	-1.153	0.900
	(0.647)	(0.532)	(0.582)		(0.742)	(0.566)	(0.542)

Table A.1.2: Maximum likelihood estimates of four category ACM specifications. The estimation period is March 1984 – April 2001. Standard errors are reported in parentheses.

A.1.2 In- and Out-of-sample Four Category ACH-ACM Forecast Results

Table A.1.3 depicts mean squared errors (MSEs) of one to twelve months forecasts of the effective federal funds rate. Each model uses the same estimated ACH equations reported in (A.7) and (A.8) in Appendix A.3. Considering the in-sample forecast performance in Panel A of Table A.1.3, the full parameterized four category ACH-ACM specification delivers the smallest MSEs up to eleven months among all models. The conclusions that five category DMPP models yield considerably smaller MSEs than the in-sample VAR MSEs (see Table

2.4.4) are confirmed for the four category DMPP models. Comparisons of the out-of-sample ability to forecast the effective federal funds rate are presented in the Panels B, C and D of Table A.1.3.

	Panel A: Apr 1985-Apr 2001 (in-sample)					Panel B: May 2001-Jan 2006 (out-sample)				
h	UNC	OP	ACM	PSACM	DACM	UNC	OP	ACM	PSACM	DACM
1	**0.03**	0.04	**0.03**	**0.03**	0.04	**0.01**	**0.01**	**0.01**	**0.01**	**0.01**
2	0.10	0.10	**0.09**	**0.09**	0.10	0.08	**0.05**	**0.05**	0.06	0.07
3	0.19	0.17	**0.16**	0.17	0.18	0.22	**0.13**	**0.13**	0.16	0.19
4	0.31	0.27	**0.24**	0.27	0.29	0.39	**0.23**	0.25	0.30	0.35
5	0.44	0.37	**0.34**	0.38	0.41	0.63	**0.39**	0.42	0.50	0.57
6	0.60	0.47	**0.45**	0.50	0.55	0.94	**0.66**	0.68	0.81	0.86
7	0.81	0.60	**0.58**	0.65	0.70	1.29	**0.99**	1.00	1.17	1.20
8	1.04	0.75	**0.73**	0.82	0.87	1.67	**1.35**	**1.35**	1.57	1.57
9	1.27	0.90	**0.88**	0.99	1.03	2.08	1.76	**1.72**	2.00	1.96
10	1.49	1.04	**1.03**	1.16	1.19	2.53	2.23	**2.14**	2.47	2.39
11	1.72	**1.18**	**1.18**	1.32	1.34	2.98	2.73	**2.59**	2.97	2.81
12	1.93	**1.32**	1.33	1.48	1.48	3.44	3.26	**3.07**	3.49	3.26
	Panel C: May 2001-May 2004					Panel D: Jun 2004-Jan 2006				
h	UNC	OP	ACM	PSACM	DACM	UNC	OP	ACM	PSACM	DACM
1	**0.01**	**0.01**	**0.01**	**0.01**	**0.01**	0.01	**0.00**	0.01	0.01	0.01
2	0.09	**0.06**	0.07	0.08	0.09	0.08	0.02	**0.01**	0.03	0.03
3	0.23	**0.16**	0.19	0.20	0.24	0.19	0.06	**0.03**	0.08	0.09
4	0.41	**0.30**	0.36	0.38	0.44	0.36	0.11	**0.05**	0.16	0.17
5	0.67	**0.49**	0.60	0.62	0.73	0.56	0.19	**0.09**	0.27	0.28
6	1.02	**0.87**	0.97	1.03	1.12	0.80	0.28	**0.14**	0.39	0.39
7	1.41	**1.32**	1.43	1.53	1.57	1.06	0.38	**0.20**	0.52	0.51
8	1.85	**1.82**	1.93	2.07	2.07	1.35	0.49	**0.27**	0.66	0.64
9	**2.32**	2.38	2.46	2.64	2.60	1.65	0.61	**0.35**	0.82	0.78
10	**2.82**	3.03	3.05	3.27	3.18	1.98	0.75	**0.45**	0.98	0.93
11	**3.36**	3.73	3.69	3.96	3.76	2.29	0.87	**0.54**	1.14	1.06
12	**3.90**	4.49	4.38	4.69	4.38	2.58	0.98	**0.65**	1.28	1.20

Table A.1.3: Mean squared errors for four category ACH-ACM forecasts of the effective federal funds rate, in- and out-of-sample. All four category ACH-ACM models, the unconditional and ordered probit model are estimated on March 1984 – April 2001 data and use the ACH specification in Equations (A.7) and (A.8) as the model for the point process. The forecast horizon h in months is given in the first column. Bold faced numbers indicate the lowest MSE at the respective horizon.

The good in-sample performance of the ACH-ACM models holds out-of-sample as well. Over the complete out-of-sample period May 2001 – January 2006 the ACH-ACM produces favorable results in terms of small MSEs on eight forcast horizons. The ACH-UNC model performs worst indicating that using conditioning information is important. The division of

	IN-SAMPLE Apr 1985 – Apr 2001				OUT-OF-SAMPLE May 2001 – Jan 2006				CRITICAL VALUES		
κ	OP	ACM	DACM	PSACM	OP	ACM	DACM	PSACM	$\frac{1\%}{\kappa}$	$\frac{5\%}{\kappa}$	$\frac{10\%}{\kappa}$
1	12.4	8.5	10.2	8.8	18.0	15.5	13.3	16.0	21.7	16.9	14.7
4	14.8	11.0	13.7	13.8	23.5	40.6	14.8	18.3	25.5	21.0	19.0
8	13.7	14.3	10.0	9.5	18.4	22.9	26.7	26.1	27.3	23.0	21.0
12	17.8	16.4	17.5	13.5	20.0	14.0	23.0	15.0	28.4	24.1	22.2
16	16.8	10.7	17.6	15.7	13.7	13.7	25.7	15.0	29.1	24.9	23.0
20	14.4	17.3	15.8	14.4	16.3	18.0	18.0	16.3	29.7	25.5	23.6
24	18.9	14.8	14.2	12.5	12.0	10.0	24.0	12.0	30.1	25.9	24.1
26	23.0	11.8	19.2	17.4	16.6	12.1	21.0	12.1	30.3	26.2	24.3
28	15.5	12.0	19.6	12.7	12.0	12.0	22.0	12.8	30.5	26.4	24.5
32	14.8	15.5	25.5	19.4	12.1	11.6	11.6	15.0	30.9	26.7	24.9
36	12.2	13.1	19.2	14.8	14.0	17.7	17.3	17.3	31.2	27.0	25.2
40	14.0	10.0	21.0	13.0	11.0	14.3	17.7	17.7	31.4	27.3	25.5
44	14.2	14.2	20.5	15.2	9.0	13.0	29.0	13.0	31.7	27.5	25.7
48	17.7	21.2	20.1	17.7	17.0	17.0	29.0	13.0	31.9	27.8	25.9
52	11.5	14.0	17.8	12.7	21.0	16.0	21.0	16.0	32.1	28.0	26.2
60	13.9	10.8	12.4	13.9	21.0	21.0	21.0	11.0	32.4	28.4	26.5

Table A.1.4: Results of iid uniformity test for continued PIT sequence using four category ACH-ACM. All four category ACH-ACM models, the unconditional and ordered probit model are estimated on March 1984 – April 2001 data. The forecast periods are April 1985 – April 2001 (in-sample) and May 2001 – January 2006 (out-of-sample), respectively. See caption of Figure 2.4.5 for explanations.

the out-of-sample periods shows that for the first subsample (May 2001 – May 2004) the ACH-OP achieves lowest MSEs up to the eighth month and for the second subsample (June 2004 – January 2006) the ACH-ACM produces smallest MSEs for all forecast horizons. These favorable four category ACH-ACM forecast results confirm the conclusions drawn from the five category ACH-ACM modeling approach.

Comparing the values of the in-sample uniformity test statistics in Table A.1.4 with the $5\%/\kappa$ for the ACH-ACM and ACH-OP model, we cannot reject the hypothesis that the models deliver correct probability forecasts. Same results are depicted in Figure A.1.1. The histogram bars lie within the 95% confidence bounds and the continued PIT sequence does not show violations of iid uniformity. However, as noted in Section 2.4.4, the thinning of the PIT sequence into κ subseries for multiperiod forecasts ($\kappa > 1$) implies wide confidence bounds for longer forecasting horizons.

The autocorrelograms in Figure A.1.2 of the four category ACH-ACM models' PIT sequence look quite similar to the autocorrelograms in Figure 2.4.3. They show that for predictions at two to four month horizons the dynamics of the probability forecast is not entirely captured.

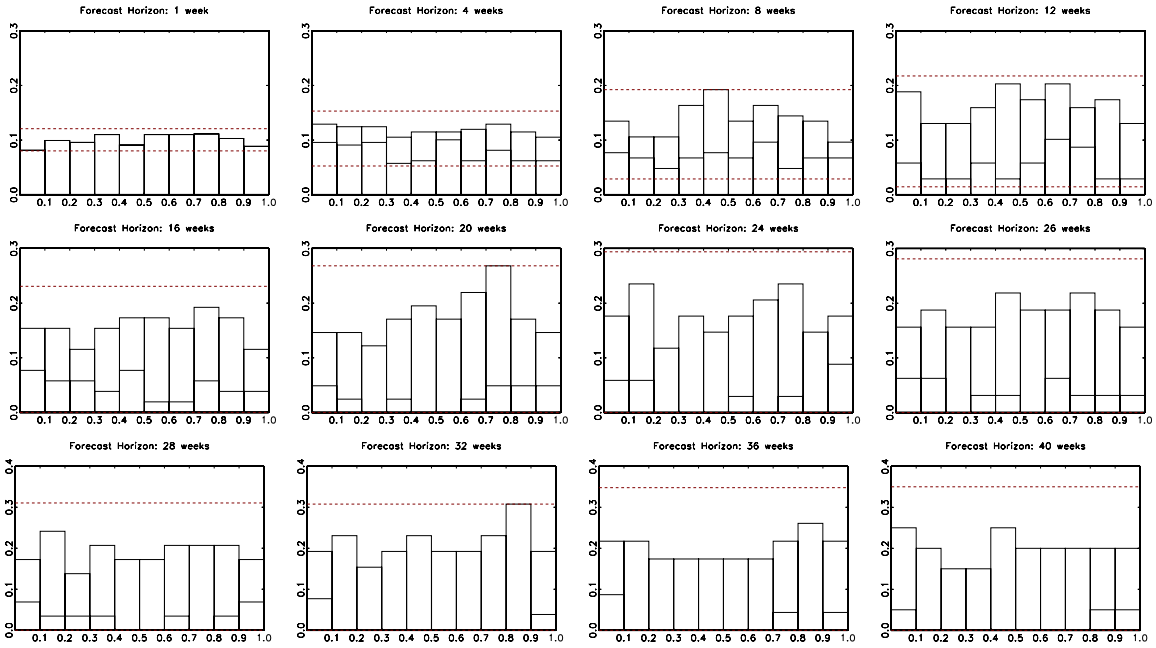

Figure A.1.1: Histograms of the continued PIT sequence: four category ACH-ACM, in-sample probability forecast. The estimation period is March 1984 – April 2001, the forecast evaluation period is May 2001 – January 2006. The horizontal lines superimposed on the histograms mark the 95% confidence intervals. See caption of Figure 2.4.2 for explanations.

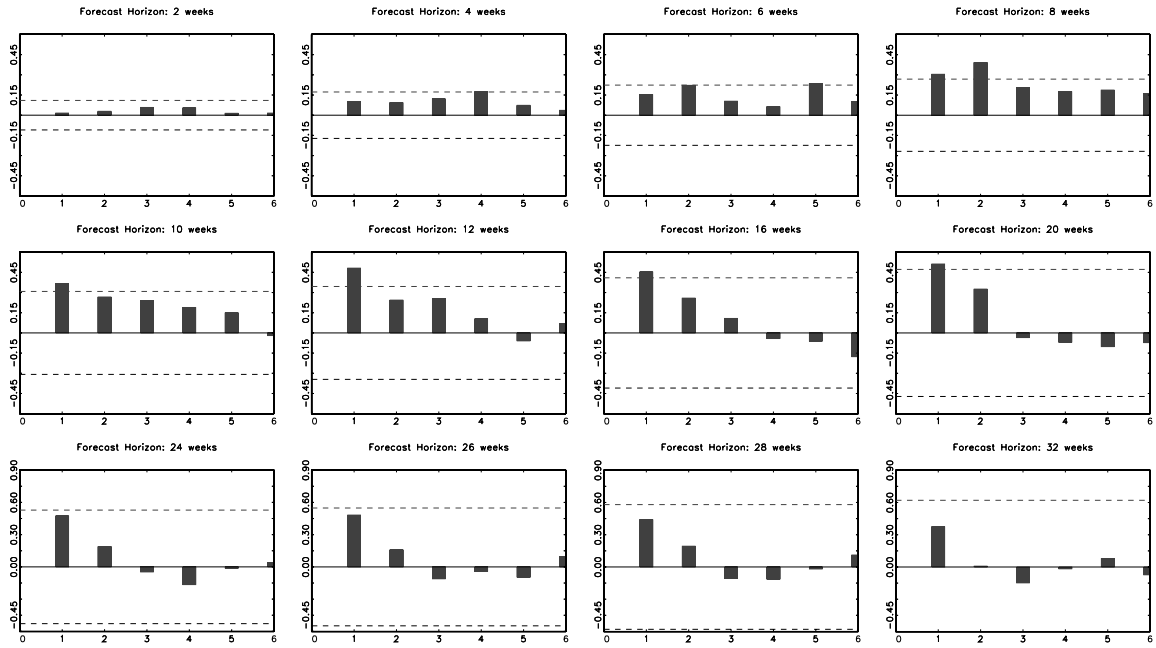

Figure A.1.2: Autocorrelations of the continued PIT sequence: four category ACH-ACM, in-sample probability forecast. For each forecast horizon κ the z^* sequences are split into κ subseries. The figures show the maximal autocorrelations of the κ subseries. The horizontal lines superimposed on the autocorrelograms mark the 95% confidence intervals. The estimation period is March 1984 – April 2001, the forecast period is April 1985 – April 2001.

Out-of-sample probability forecast evaluations in Table A.1.4 show that the hypothesis that the probability forecasts are correct cannot be rejected at a $1\%/\kappa$ significance level for all models (except for ACM $\kappa = 4$ forecast). The four category ACH-DACM histograms in Figure A.1.3 are more jagged than those in Figure A.1.1 but still support the hypothesis that the out-of-sample probability forecasts are correct.

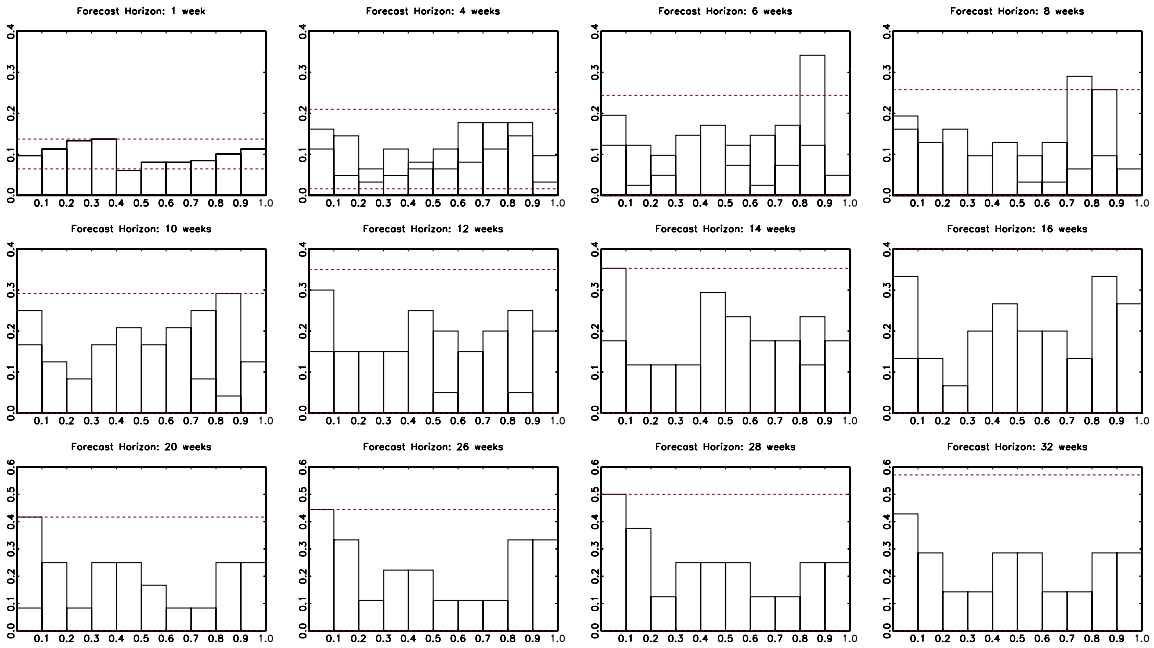

Figure A.1.3: Histograms of the continued PIT sequence: four category ACH-DACM, out-of-sample probability forecast. The estimation period is March 1984 – April 2001, the forecast evaluation period is May 2001 – January 2006. The horizontal lines superimposed on the histograms mark the 95% confidence intervals. See caption of Figure 2.4.2 for explanations.

A.2 Simulation of Multi-step Probability Forecasts

To compute multi-step ahead forecasts $f(i_{t+\kappa}|\Upsilon_t)$, we simulate future sample paths of an ACH-ACM process. We first need a realization of x_{t+1}, the random variable which indicates whether a target change occurs at $t+1$. Since the ACH model readily delivers $\mathbb{P}(x_{t+1} = 1|\Upsilon_t) = h_{t+1}$, this can be obtained by drawing from a Bernoulli distribution with success probability h_{t+1}. Let us denote by $x_{t+1}^{(1)}$ the result of that draw. The superscript indicates that this is the first of many replications. If $x_{t+1}^{(1)} = 0$, the target is unchanged, $i_{t+1}^{(1)} = i_t$. If $x_{t+1}^{(1)} = 1$, we update the counting function by computing $N_{t+1}^{(1)} = N_t + 1$, and

determine the size of the target change. This is done by drawing from the discrete distribution of the random variable y_{t+1} determining the size of the target change. For the ACM model, this probability distribution is given by $\boldsymbol{\pi}_{N(t)}$ (see Equation (2.7)). Having drawn the target change in $t+1$, $y_{t+1}^{(1)}$, we add it to i_t and obtain the simulated target value $i_{t+1}^{(1)}$. Iterating forward the ACH-ACM Equations (2.5) and (2.6) yields the probabilities $h_{t+2}^{(1)}$ and $\boldsymbol{\pi}_{N(t+2)}^{(1)}$. These probabilities are used to simulate $x_{t+2}^{(1)}$ and $i_{t+2}^{(1)}$. This works in the same way as just described except that we condition on time $t+1$ simulated values instead of time t observed values. The procedure is continued until $t+\kappa-1$. Conditioning on the simulated path of target values $i_{t+\kappa-1}^{(1)}, i_{t+\kappa-2}^{(1)}, \ldots, i_{t+1}^{(1)}$, we can then compute the one-step probability forecast analogously to Equation (2.12) as:

$$f(i_{t+\kappa}^{(1)}|i_{t+\kappa-1}^{(1)},..,i_{t+1}^{(1)},\Upsilon_t) = \begin{cases} \mathbb{P}(i_{t+\kappa}^{(1)} = i_{t+\kappa-1}^{(1)}|i_{t+\kappa-1}^{(1)},..,i_{t+1}^{(1)},\Upsilon_t) & = 1-h_{t+\kappa}^{(1)} \\ \mathbb{P}(i_{t+\kappa}^{(1)} = i_{t+\kappa-1}^{(1)}+s_j|i_{t+\kappa-1}^{(1)},..,i_{t+1}^{(1)},\Upsilon_t) & = h_{t+\kappa}^{(1)}\pi_{jN(t+\kappa)}^{(1)} \\ \quad \text{for } j=1,2,...,k \quad . & \end{cases} \tag{A.3}$$

In order to obtain the κ-period ahead probability forecast $f(i_{t+\kappa}|\Upsilon_t)$, we need to remove the conditioning on the sample path. For this purpose we exploit that

$$f(i_{t+\kappa},...,i_{t+1}|\Upsilon_t) = f(i_{t+\kappa}|i_{t+\kappa-1},...,i_{t+1},\Upsilon_t)\cdot f(i_{t+\kappa-1},...,i_{t+1}|\Upsilon_t) \tag{A.4}$$

and

$$f(i_{t+\kappa}|\Upsilon_t) \;=\; \sum_{i_{t+\kappa-1}} ... \sum_{i_{t+1}} f(i_{t+\kappa}|i_{t+\kappa-1},...,i_{t+1},\Upsilon_t)\cdot f(i_{t+\kappa-1},...,i_{t+1}|\Upsilon_t) \quad . \tag{A.5}$$

The computation of the multi-period probability forecast via Equation (A.5) is still computationally intractable. We address this problem by repeating the simulation of the sample paths described above M times. This delivers a vector sequence of M probability forecasts, each conditioned on the respective simulated sample path,

$$\{f(i_{t+\kappa}^{(m)}|i_{t+\kappa-1}^{(m)},\ldots,i_{t+1}^{(m)},\Upsilon_t)\}_{m=1}^{M} \quad .$$

Averaging over the M replications yields a consistent estimate of $f(i_{t+\kappa}|\Upsilon_t)$,

$$\frac{1}{M}\sum_{m=1}^{M} f(i_{t+\kappa}^{(m)}|i_{t+\kappa-1}^{(m)},\ldots,i_{t+1}^{(m)},\Upsilon_t) \underset{p}{\rightarrow} \mathbb{E}\left[f(i_{t+\kappa}|i_{t+\kappa-1},...,i_{t+1},\Upsilon_t)\right] = f(i_{t+\kappa}|\Upsilon_t) \quad . \tag{A.6}$$

A useful byproduct of this simulation strategy is the possibility to use the estimated probability forecast to produce multi-period point forecasts $\mathbb{E}\left[i_{t+\kappa}|\Upsilon_t\right]$ and conduct MSE comparisons.

A.3 ACH and OP Estimation Results

Taking into account that the objective of the Federal Reserve changed (see Section 2.2), Hamilton and Jordà (2002) estimate separate ACH specifications for two subperiods. For the first subsample, covering the period from March 1, 1984 to November 23, 1989, the estimated ACH equation reads

$$h_t = \left(\underset{(1.160)}{2.257} + \underset{(0.056)}{0.090}\ \tau_{N(t-1)-1} + \underset{(0.078)}{0.847}\ \psi_{N(t-1)-1} - \underset{(0.631)}{2.044}\ \mathrm{FOMC}_{t-1}\right)^{-1} \quad , \tag{A.7}$$

where FOMC_t is a dummy variable that equals one if there was a FOMC meeting in t and zero otherwise. The values in parentheses are standard errors. For the second subsample from November 30, 1989 to April 26, 2001, the estimated ACH equation is given by

$$h_t = \left(\underset{(7.119)}{30.391} + \underset{(0.024)}{0.067}\ \tau_{N(t-1)-1} - \underset{(7.295)}{23.046}\ \mathrm{FOMC}_t - \underset{(2.462)}{8.209}\ |\mathrm{SP}_{t-1}|\right)^{-1} \quad , \tag{A.8}$$

where $|SP|$ is the absolute value of the spread between the six-month treasury bill rate (TB6) and the effective federal funds rate (FFR), i.e. $SP = TB6 - FFR$.

Replicating Hamilton and Jordà's (2002) ordered probit estimation (March 1, 1984 -April 26, 2001), we obtain for the latent target change equation:

$$y_{N(t)}^{*} \quad = \quad \underset{(0.390)}{2.389}\ y_{N(t)-1} + \underset{(0.211)}{0.741}\ \mathrm{SP}_{t-1}\cdot x_t \quad . \tag{A.9}$$

Standard errors are reported in parenthesis.

A.4 Bayesian Type Model Averaging

Let $\mathscr{M} = \{M_1, ..., M_M\}$ be the set of models that yield a forecast of the variable of interest, in our application the value of the federal fund rate target, i. A target forecast that takes into account model uncertainty is then given by an average of the model forecasts weighted by their posterior model probabilities:

$$\mathbb{E}[i|data] = \sum_{j=1}^{M} \mathbb{E}[i|data, M_j] \cdot \mathbb{P}(M_j|data) \quad . \tag{A.10}$$

This posterior probability for model M_j after observing the data is proportional to the product of the marginal likelihood for model M_j and the prior probability for model M_j, viz

$$\mathbb{P}(M_j|data) = \frac{\mathbb{P}(data|M_j)\mathbb{P}(M_j)}{\sum_{l=1}^{M} \mathbb{P}(data|M_l)\mathbb{P}(M_l)} \propto \mathbb{P}(data|M_j)\mathbb{P}(M_j) \quad . \tag{A.11}$$

We refrain from accounting for parameter uncertainty and assume, in a Bayesian sense, equal prior model probabilities.

In a standard Bayesian model averaging approach, the marginal likelihood is used for the construction of the weights as a natural in-sample measure of fit (see for e.g. Garratt et al. 2006, Hoeting et al. 1999). Alternatively, Eklund and Karlsson (2005) suggest an out-of-sample measure of fit, the predictive likelihood. They split the sample into a training-sample used for parameter inference and a hold-out sample from which the predictive likelihood is computed. We follow their idea and use the predictive likelihood as weights in the Bayesian type model averaging.

A.5 Additional Bayesian Type Model Averaging Results

In Section 2.4.5 we consider in the last column of Panel B, C and D in Table 2.4.4 only the Bayesian type model averaging results of an ACH-OP and an ACH-DACM. We focus on this model combination since it is the most successful one in terms of smallest MSE among our selections.

Table A.5.1 gives additional Bayesian type model averaging MSEs results for four model combinations. See the caption of Table A.5.1 for details. All models use the ACH

specification in Appendix A.3 as model for the point process. Naturally, the ACH-DACM and ACH-PSACM combination is an interesting choice. However, it turns out that the MSEs of this combination are larger than the those of the ACH-OP and ACH-DACM reported in the Section 2.4.5.

	May 2001-Jan 2006				May 2001-May 2004				Jun 2004-Jan 2006			
h	BT1	BT2	BT3	BT4	BT1	BT2	BT3	BT4	BT1	BT2	BT3	BT4
1	**0.01**	**0.01**	**0.01**	**0.01**	**0.01**	**0.01**	**0.01**	**0.01**	**0.00**	**0.00**	**0.00**	**0.00**
2	**0.05**	**0.05**	**0.05**	**0.05**	**0.06**	0.07	**0.06**	0.07	**0.02**	0.03	**0.02**	**0.02**
3	**0.13**	0.15	0.14	0.14	**0.17**	0.19	0.18	0.19	**0.06**	0.07	**0.06**	**0.06**
4	**0.24**	0.28	0.26	0.27	**0.31**	0.35	0.34	0.35	**0.12**	0.14	**0.12**	**0.12**
5	**0.40**	0.47	0.44	0.45	**0.51**	0.59	0.57	0.58	**0.20**	0.24	**0.20**	**0.20**
6	**0.65**	0.73	0.69	0.70	**0.84**	0.93	0.91	0.93	**0.28**	0.35	**0.28**	**0.28**
7	**0.93**	1.03	0.99	1.00	**1.22**	1.33	1.32	1.34	**0.37**	0.46	**0.37**	**0.37**
8	**1.24**	1.36	1.32	1.33	**1.65**	1.77	1.77	1.79	**0.47**	0.58	**0.47**	**0.48**
9	**1.58**	1.71	1.67	1.68	**2.12**	2.24	2.26	2.27	**0.58**	0.72	**0.58**	**0.58**
10	**1.96**	2.10	2.06	2.07	**2.64**	2.76	2.80	2.81	0.71	0.87	**0.70**	0.71
11	**2.38**	2.50	2.48	2.49	**3.22**	3.31	3.38	3.39	**0.82**	1.01	**0.82**	**0.82**
12	**2.80**	2.91	2.91	2.90	**3.81**	3.86	3.97	3.97	**0.93**	1.14	**0.93**	**0.93**

Table A.5.1: Additional mean squared errors results for Bayesian type model averaging for one to twelve months forecasts of the effective federal funds rate, out-of-sample. The table depicts the mean squared errors for four Bayesian type model average combinations (BT1: OP-DACM, as in Table 2.4.4, BT2: PSACM-DACM, BT3: PSACM-DACM-OP, BT4: UNC-DACM-PSACM-OP). All models are estimated on March 1984 – April 2001 data. The forecast horizon h in months is given in the first column. Bold faced numbers indicate the lowest MSE at the respective horizon.

Chapter 3

Forecasting Return Volatility with Continuous Variation and Jumps

This paper introduces a time series model for realized volatility that accounts for continuous variation and jumps. Engle and Russell's (1998) autoregressive conditional duration approach is used to model continuous and jump size variation and Hamilton and Jordà's (2002) autoregressive conditional hazard model is applied to jump durations. We further suggest a methodology to evaluate density forecasts delivered by the model. Diagnostics as well as point and density forecast results show that this approach qualifies as a useful forecast model for daily return variation.

This chapter is based on the article "Forecasting Return Volatility with Continuous Variation and Jumps" by K. Kehrle (2008).

3.1 Introduction

Modeling and forecasting volatility is one of the major research topics in financial econometrics. Since the seminal introduction of Engle's (1982) ARCH model, we have observed an explosive growth in this field of finance. This development is due to the desire for a deeper understanding of financial markets and the importance of accurate volatility forecasts for the valuation of derivatives, portfolio management and risk management. This paper presents a new discrete time model for volatility and compares its forecast performance with other approaches discussed in the literature. It also proposes a methodology to evaluate the quality of density forecasts delivered by the model and applies it to forecasts of daily return variation.

Beyond the vast literature of discrete-time ARCH and GARCH modeling (Engle 1982, Bollerslev 1986)[1], two main strands emerged in the volatility literature over time. The first is based on continuous-time diffusion processes and dates back to the work of Merton (1980). More recent studies examine more flexible parametric diffusion processes that allow for time-varying jump diffusions. However, the empirical results based on daily data show difficulties and do not lead to precise findings, see e.g. Bates (2000), Chernov et al. (2003), Eraker (2004) and Pan (2002). The second strand of literature presents a nonparametric approach and uses high frequency data. Andersen et al. (2003) and Andersen and Bollerslev (1998), among others, suggest the sum of intradaily squared returns as measure for volatility that converges to the quadratic variation of the price process. Andersen et al. (2001a, 2001b) investigate this realized volatility measure and provide important empirical insights into the properties of daily return and volatility distributions. A major leap forward in the realized volatility literature is the work of Barndorff-Nielsen and Shephard (2004). They develop a volatility measure called bi-power variation that is immune to jumps and, therefore, enables a decomposition of realized volatility into continuous and discontinuous variation. Huang and Tauchen (2005), Andersen et al. (2007a) and Barndorff-Nielsen and Shephard (2006) show that the contribution of jumps to daily price variation is not negligible.

The present paper is linked to Andersen et al.'s (2007b) and Bollerslev et al.'s (2009) work that disentangles return volatility into a continuous and a jump component and models realized volatility by a reduced form time series approach. This paper introduces a new

[1] For an extensive forecast comparison of important GARCH specifications see Hansen and Lunde (2005).

model that is based on Engle and Russell's (1998) autoregressive conditional duration (ACD) model. The ACD accounts for transaction price durations that are strictly positive and serially correlated. Since continuous variation exhibits same stylized facts, we adopt Engle and Russell's (1998) framework to continuous variation. The jump variation process is conceived as a marked point process. As outlined by Engle (2000), marked point processes can be conveniently separated into a model for the duration between points in time (here: observing a jump) and a model for the marks, i.e. the variables which are observed when the event occurs (here: size of a jump). Thus, the occurrence of jumps is irregularly spaced in time and both jump durations and sizes have an autoregressive nature. To account for the time between successive jumps we employ, Hamilton and Jordà's (2002) autoregressive conditional hazard (ACH) model, as suggested by Andersen et al. (2007b). The ACH delivers an estimate of the probability that a jump will be observed during the next day. The time series features of jump sizes are addressed by using Engle and Russell's (1998) model. In general, an autoregressive conditional model structure is appealing since point and density forecasts are easily derived (see Engle and Russell 1997, Bauwens et al. 2004). Combining the continuous and jump variation models forms a joint model for total return variation.

Furthermore, we suggest a methodology to evaluate density forecasts delivered by this model class. For this purpose, we use forecast evaluation methods of Diebold et al. (1998) to assess the accuracy of realized volatility density forecasts. Density forecast evaluations, which are perceived as model diagnostics, are also applied to the models of continuous and jump variation. Since Diebold et al.'s (1998) method is only applicable to density forecasts, a direct implementation to probability forecast by ACH is infeasible. Hence, a discrete analog in the vein of Denuit and Lambert (2005) and Grammig and Kehrle (2008) is employed.

The main findings of the empirical results using high frequency intraday data of the DAX future, DJ Euro Stoxx 50 future, S&P 500 future and the General Motors stock can be summarized as follows. The estimation results for the continuous and jump variation models indicate high persistence and residual diagnostics do not detect model misspecifications. Probability and density forecasts are assessed and approve the suitability towards modeling realized, continuous and jump variation. For point forecast comparisons we estimate three alternative volatility models: a GARCH(1,1), an autoregressive conditional time series model for realized volatility and Corsi's (2004) HAR-RV model. Out-of-sample point forecast

comparisons show the usefulness of the introduced model for predicting daily return variation.

The remainder of this paper is structured as follows. Section 3.2 derives the nonparametric volatility measures. Section 3.3 describes the data and stylized facts. Section 3.4 presents the methodology and techniques for evaluating density forecasts. Section 3.5 discusses estimation results and employs diagnostic tools for density forecasts and compares point forecast performance. Section 3.6 concludes.

3.2 Theoretical Framework

The logarithmic price of a financial asset, denoted by p_t, is assumed to follow a continuous-time semimartingale jump diffusion process,

$$p_t = \int_0^t \mu(s)ds + \int_0^t \sigma(s)dW(s) + \sum_{s=1}^{N(t)} \kappa(s) \quad , \tag{3.1}$$

where $\mu(t)$ is a continuous mean process with local finite variation, $\sigma(t) > 0$ is the instantaneous volatility process which is càdlàg, $W(t)$ denotes a standard Brownian Motion and $N(t)$ counts the number of jumps occurring with size $\kappa(s)$ and (possibly) time-varying jump intensity $\lambda(t)$. The quadratic variation for the cumulative price process in (3.1) is[2]

$$[p,p](t) = \int_0^t \sigma^2(s)ds + \sum_{s=1}^{N(t)} \kappa^2(s) \quad , \tag{3.2}$$

where the first term on the right hand side is referred to as the integrated variation and the second as the sum of squared jumps.

Let the j^{th} intradaily continuously compounded return be denoted by $r_{t,j} = p_{t-1+\frac{j}{M}} - p_{t-1+\frac{(j-1)}{M}}$, with M as the number of returns per day. The nonparametric realized volatility measure is defined as the sum of M intraday squared returns,

$$RV_t = \sum_{j=1}^{M} r_{t,j}^2 \quad . \tag{3.3}$$

For increasing M, RV_t converges in probability to the increments of the quadratic variation

[2]The quadratic variation process of (3.1) is $[p,p](t) = \text{plim} \sum_{j=0}^{n-1} (p_{\tau_{j+1}} - p_{\tau_j})^2$, where $\tau_0 = 0 \leq \tau_1 \leq ... \leq \tau_n = t$ is a sequence of partitions with $sup_j\{\tau_{j+1} - \tau_j \to 0\}$ for $n \to \infty$.

process in (3.2) (see Andersen and Bollerslev 1998, Andersen et al. 2001a, Barndorff-Nielsen and Shephard 2001), viz

$$RV_t \underset{p}{\rightarrow} \int_{t-1}^{t} \sigma^2(s)ds + \sum_{s=N(t-1)+1}^{N(t)} \kappa^2(s) \quad \text{for } M \rightarrow \infty \quad . \tag{3.4}$$

The realized volatility measure is affected by the variation due to jumps. In order to disentangle integrated variation and jump variation, Barndorff-Nielsen and Shephard (2004) introduce the realized bi-power variation measure that is immune to jumps and converges to the integrated variation in (3.2) as M grows,

$$BV_t = \mu_1^{-2} \left(\frac{M}{M-1} \right) \sum_{j=2}^{M} |r_{t,j-1}||r_{t,j}| \underset{p}{\rightarrow} \int_{t-1}^{t} \sigma_s^2 ds \quad \text{for } M \rightarrow \infty \quad , \tag{3.5}$$

with $\mu_1 = \sqrt{2/\pi}$. Equations (3.3) and (3.5) define the daily jump variation measure as the residual between realized and bi-power variation,

$$JV_t = RV_t - BV_t \quad \underset{p}{\rightarrow} \sum_{j=N(t-1)+1}^{N(t)} \kappa^2(s_j) \quad , \tag{3.6}$$

that converges to the increments of the sum of squared jumps in (3.2) for $M \rightarrow \infty$. Any small deviation between RV_t and BV_t induces a jump size greater than zero. Even a negative jump variation becomes empirically possible due to measurement errors. This motivated Andersen et al. (2007a) to derive a test statistic that identifies significant positive jumps,

$$Z_t = \sqrt{M} \frac{(RV_t - BV_t)RV_t^{-1}}{\sqrt{(\mu_1^{-4} + 2\mu_1^{-2} - 5)\max\{1, TQ_t BV_t^{-2}\}}} \quad \underset{d}{\rightarrow} \quad N(0,1) \quad , \tag{3.7}$$

where extreme deviations of RV_t and BV_t are standardized by the realized tri-power quarticity,

$$TQ_t = M\mu_{4/3}^{-3} \left(\frac{M}{M-2} \right) \sum_{j=3}^{M} |r_{t,j-2}|^{4/3} |r_{t,j-1}|^{4/3} |r_{t,j}|^{4/3} \quad \underset{p}{\rightarrow} \quad \int_{t-1}^{t} \sigma^4(s)ds \quad , \tag{3.8}$$

with $\mu_{4/3} = 2^{2/3}\Gamma(7/6)/\Gamma(1/2)$ and $\Gamma(\cdot)$ denoting the gamma function.

If the nominator in (3.7) is large, the test statistic exceeds the critical value $\Phi_{1-\alpha}$ and identifies a significant squared jump, denoted by J_t. Then the jump occurrence indicator X_t

equals one and $J_t = RV_t - BV_t$. $\Phi_{1-\alpha}$ is the $1-\alpha$ quantile of the standard normal distribution with α as significance level. If $Z_t \leq \Phi_{1-\alpha}$, then the significant jump size in t is zero, i.e. $J_t = 0$ and $X_t = 0$. This formalizes to

$$J_t \equiv X_t(RV_t - BV_t) \quad \text{for} \quad t = 1, ..., T \quad . \tag{3.9}$$

Hence, positivity of squared significant jumps is ensured by (3.9). On a day without jumps, the continuous variation measure C_t is equal to the realized variance and on days with a jump C_t equals the bi-power variation, viz

$$C_t \equiv (1 - X_t)RV_t + X_t BV_t \quad \text{for} \quad t = 1, ..., T \quad . \tag{3.10}$$

3.3 Data

The estimated variation measures introduced in the previous section only converge to their continuous-time counterparts if the number of daily sampled returns increases infinitely. Hence, making use of a high data frequency is requested. However, variation measures computed on a tick-by-tick basis hinge on market microstructure noise.[3] Various approaches in the literature propose volatility estimators that are not affected by noise. Thomakos and Wang (2003), Andersen et al. (2001a) or Bollen and Inder (2002) suggest noise free variation measures based on filtered returns using MA or AR processes. Bandi and Russell (2008) and Aït-Sahalia et al. (2005) propose a bias-corrected version of realized volatility. The present paper follows Andersen et al. (2001b, 2007b) and solves the problem of market microstructure noise by computing variation measures on a coarser sampling frequency of five minutes.

The data contain intraday prices of the Eurex traded DAX and Euro Stoxx 50 future (DAX, ESX), the Chicago Mercantile Exchange traded S&P 500 future (SP) and the New York Stock Exchange traded General Motors (GM) stock. Futures data cover the period from July 1^{st}, 2002 to June 30^{th}, 2006 and GM data are available from January 1^{st}, 2001 to June 30^{th}, 2006.[4]

[3] The source of the market microstructure noise is manifold including price discreteness, bid-ask bounds, and measurement errors.

[4] All futures are denominated in local currencies. The maximum lifetime of a future contract is limited to nine months. Therefore, a continuous time series of transaction prices is not directly observed and needs to be constructed by using the last three months of the future life-span. Proceeding in that fashion assures a rather high trading frequency. GM data are extracted from NYSE Trade and Quotations (TAQ) bid and ask quotes.

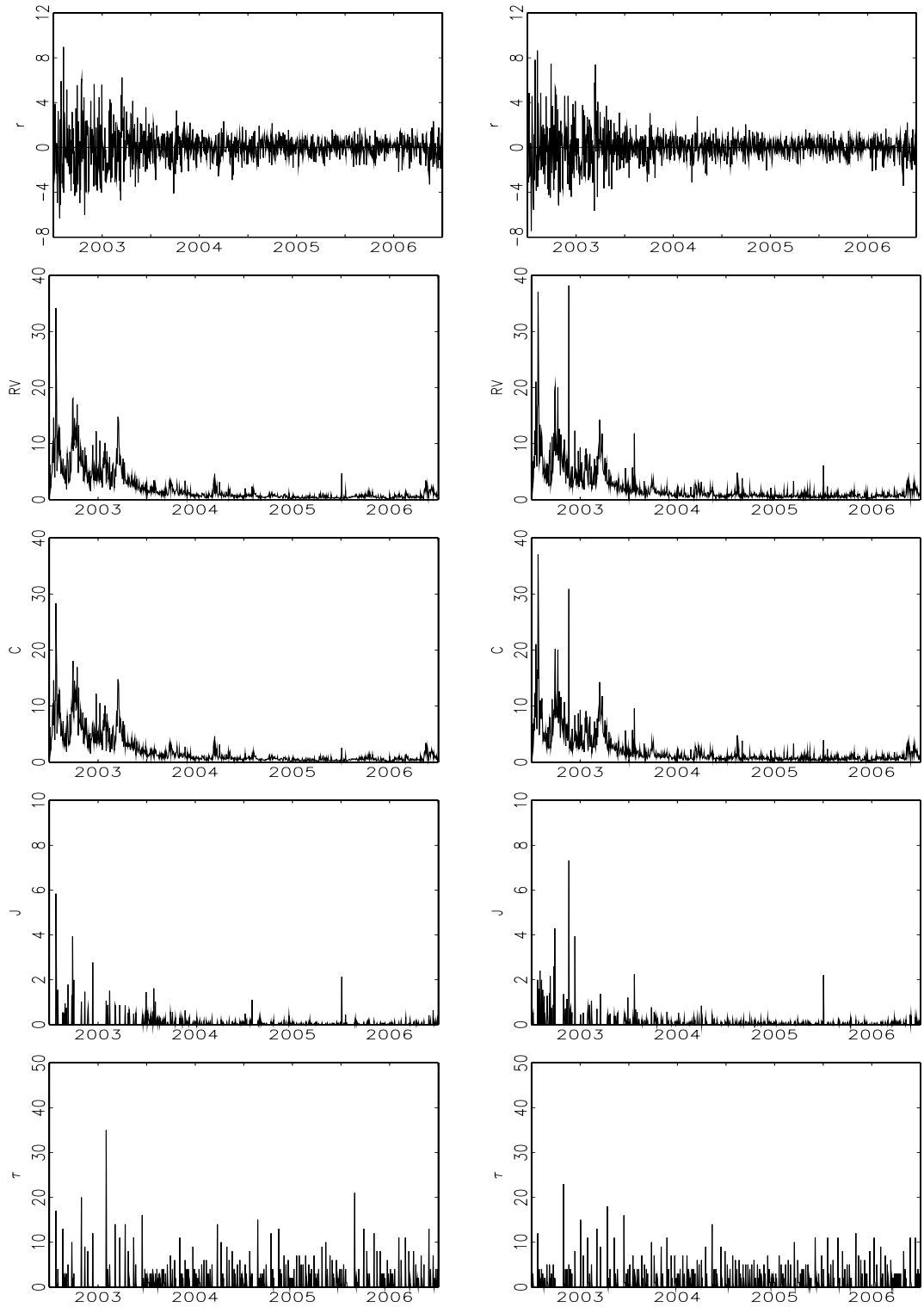

Figure 3.3.1: Time series of return and volatility measures for DAX and ESX. The left panels show daily time series for DAX and the right panels correspond to ESX. The top panel depicts daily returns, the second realized volatility, the third continuous variation, the fourth significant jumps computed using $\Phi_{0.05}$, and the bottom panel displays jump durations.

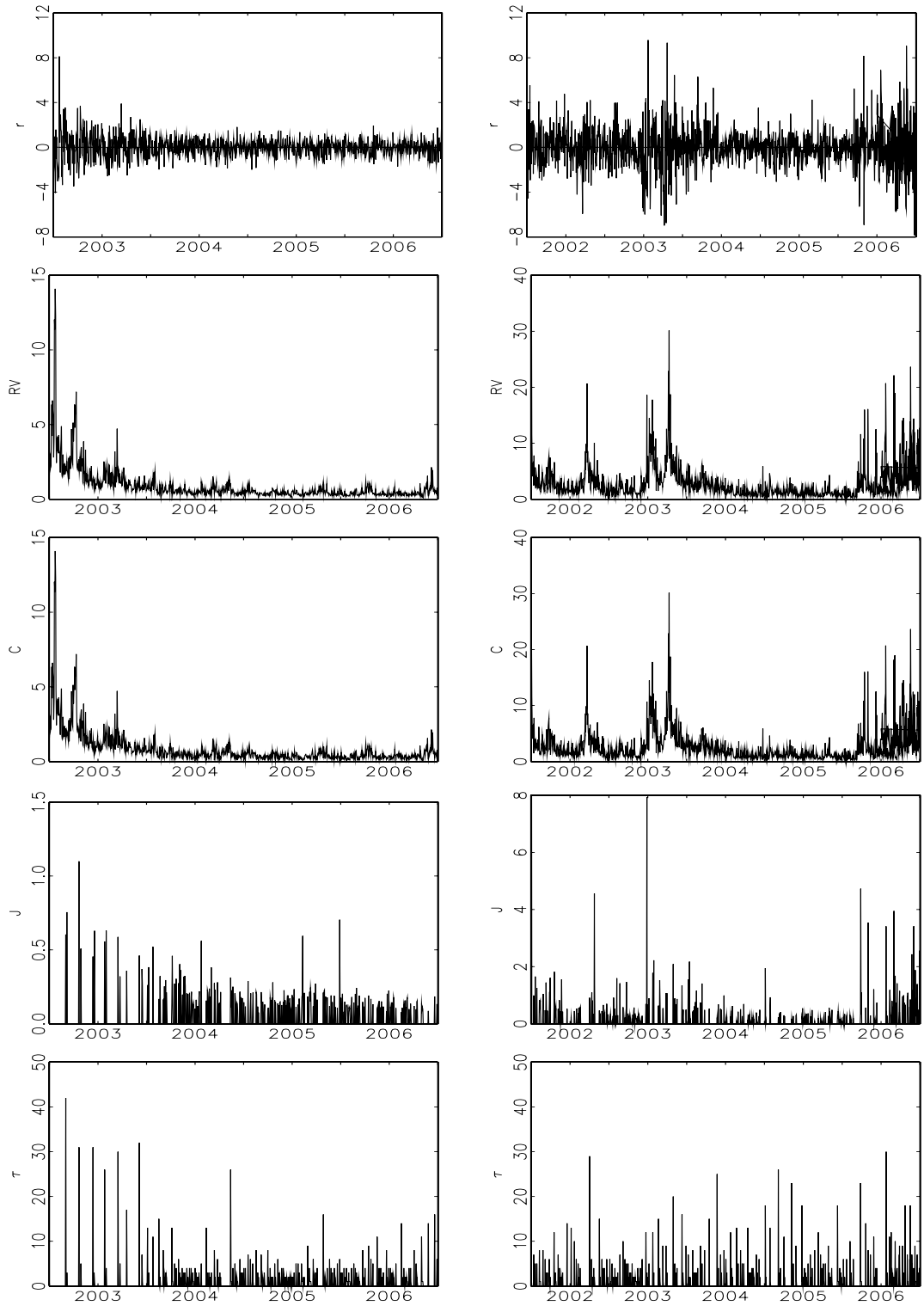

Figure 3.3.2: Time series of return and volatility measures for SP and GM. The left panels show daily time series for SP and the right panels correspond to GM. The top panel depicts daily returns, the second realized volatility, the third continuous variation, the fourth significant jumps computed using $\Phi_{0.05}$ for GM and $\Phi_{0.001}$ for SP, and the bottom panel displays jump durations.

We compute returns between the first prices in each five minute interval. Holidays and overnight returns as well as returns beyond 09:20am to 08:00pm (CET) for DAX and ESX, 08:30am to 3:15pm (CST) for SP, and 09:40am to 4:00pm (EST) for GM are excluded.

Time series plots of daily returns and volatilities are depicted in Figures 3.3.1 and 3.3.2. The returns in the figures show the well known volatility clustering which is also reflected in the series of realized and continuous variation. The figures also depict significant jumps that occur irregularly with varying size and exhibit a clustering.

The descriptive statistics in Table 3.3.1 summarize the stylized facts in the data. The empirical distributions of continuous and realized variation are highly skewed and leptocurtic.

	r_t	RV_t	C_t	J_n	τ_n
		Panel A: DAX			
$Mean$	-0.031	2.043	1.968	0.330	4.352
Std	1.449	3.011	2.882	0.621	4.236
$Skewness$	0.291	3.535	3.116	5.160	2.804
$Kurtosis$	4.39	20.33	13.83	34.53	12.89
$LB(10)$	28	5606	5996	326	20
		Panel B: ESX			
$Mean$	-0.034	2.186	2.092	0.356	3.752
Std	1.422	3.292	3.137	0.683	3.323
$Skewness$	0.489	4.632	4.432	5.887	2.041
$Kurtosis$	6.14	34.19	31.06	47.08	5.85
$LB(10)$	45	3987	4255	354	16
		Panel C: SP			
$Mean$	0.004	0.877	0.830	0.196	4.190
Std	0.937	1.170	1.182	0.131	5.675
$Skewness$	0.594	5.469	5.392	3.041	3.775
$Kurtosis$	7.08	43.28	42.27	12.60	16.68
$LB(10)$	14	4592	4579	353	113
		Panel D: GM			
$Mean$	-0.030	2.894	2.749	0.716	4.949
Std	1.830	3.091	2.915	1.147	5.266
$Skewness$	0.197	3.770	3.338	7.399	2.304
$Kurtosis$	2.99	23.32	16.65	74.92	6.14
$LB(10)$	16	2251	2558	13	15

Table 3.3.1: Summary statistics for return and volatility measures. The descriptive statistics are computed on the total sample for the return, r_t, realized volatility, RV_t, continuous variation, C_t, (significant) jump variation, J_n, and jump duration, τ_n. There are 1016 daily observations for the DAX, 1015 for ESX, 995 for the SP, and 1367 for the GM (denoted by the subscript t) with 233, 270, 237 and 276 significant jumps (denoted by the subscript n), respectively. The number of daily sampled returns for the DAX is 128, for ESX $M = 128$, for SP $M = 79$, and for GM $M = 75$. The first four rows of each panel report the sample mean, standard deviation, along with skewness and kurtosis. The last row labeled $LB(10)$ reports a Ljung-Box statistic up to lag ten.

Ljung-Box tests in Table 3.3.1 do not reject the null of no autocorrelation up to ten lags for realized and continuous variation. Weaker but still present serial dependence can be detected in jump sizes and durations (τ_n) between successive jumps.

3.4 Methodology

3.4.1 Econometric Model

Recently, Andersen et al. (2007a) and Bollerslev et al. (2009) proposed parametric models based on nonparametric volatility measures where daily realized volatility is decomposed into continuous and jump variation. We follow this idea and specify a model that accounts for the autoregressive nature in continuous variation, jump sizes and jump duration. We suggest to adopt Engle and Russell's (1998) approach to capture the time series dynamics in continuous variation and refer to it as autoregressive conditional continuous variation (AC-C) model.[5] The jump variation process is conceived as a marked point process. The time between two successive jumps are points at which we observe the mark, i.e. the jump size. To account for the durations between significant jumps, we employ Hamilton and Jordà's (2002) autoregressive conditional hazard (ACH) model. For jump sizes we apply Engle and Russell's (1998) framework and refer to it as autoregressive conditional jump size (AC-J) model. In the following we present the models for continuous variation, jump sizes and jump duration before we combine them to a joint model for daily return variation.

For the AC-C model it is assumed that daily continuous variation is specified as

$$C_t = \sigma_t \varepsilon_t \quad , \tag{3.11}$$

where ε_t is an iid innovation with constant expectation μ and constant variance such that the expected continuous variation conditional on the information set Υ_{t-1} is $\mathbb{E}\left[C_t|\Upsilon_{t-1}\right] = \sigma_t\mu$. σ_t follows a linear autoregressive process that updates daily,

$$\sigma_t = \delta_{\mathrm{C}} + \alpha_{\mathrm{C}} C_{t-1} + \beta_{\mathrm{C}}\sigma_{t-1} \quad , \tag{3.12}$$

[5] Engle and Russell's (1998) approach is also used by Lanne (2006) who applies a mixture multiplicative error model to realized volatility.

where $\delta_{\mathrm{C}} > 0$, $\alpha_{\mathrm{C}} \geq 0$ and $\beta_{\mathrm{C}} \geq 0$ which ensures positivity of σ_t. Given (3.11) and assuming an unconditional innovation distribution, $f_\varepsilon(\cdot;\boldsymbol{\theta}_\varepsilon)$, the conditional likelihood function is given by

$$f_{\mathrm{C}}(C_t|\Upsilon_{t-1};\boldsymbol{\theta}_{\mathrm{C}}) = f_\varepsilon\left(\frac{C_t}{\sigma_t};\boldsymbol{\theta}_\varepsilon\right)\sigma_t^{-1} \quad , \tag{3.13}$$

with unknown parameter vector $\boldsymbol{\theta}_{\mathrm{C}} = (\delta_{\mathrm{C}}, \alpha_{\mathrm{C}}, \beta_{\mathrm{C}}, \boldsymbol{\theta}_\varepsilon)$. The AC-C can be estimated straightforward by Maximum Likelihood (ML). In principle, any innovation distribution with positive support is possible. In this paper, we consider the exponential (AC-EC), Weibull (AC-WC) and Grammig and Maurer's (2000) Burr distribution (AC-BC) for $f_\varepsilon(\varepsilon_t;\boldsymbol{\theta}_\varepsilon)$.

The ACH model of Hamilton and Jordà (2002) describes a mean duration process and is based on the ACD model. Conditional on previous duration, it delivers an estimate of the probability that a jump will be observed the next day. Given Section 3.2, the smallest time interval between successive significant jumps is one day. Thus, τ_n denotes the duration in number of days between the n^{th} and $(n+1)^{th}$ jump. The sequence of conditional expected durations $\psi_n \equiv \mathbb{E}\left[\tau_n|\Upsilon_{n-1}\right]$ is assumed to evolve as

$$\psi_n = \alpha_{\mathrm{X}}\tau_{n-1} + \beta_{\mathrm{X}}\psi_{n-1} \quad , \tag{3.14}$$

where $\alpha_{\mathrm{X}} \geq 0$ and $\beta_{\mathrm{X}} \geq 0$. Equation (3.14) only updates if a jump occurs on day t and remains the same as in $t-1$ if no jump occurs. A discrete step function, $N(t)$, that increases by one if there is a jump in t and is unchanged if no jump occurs, provides a convenient link between event and calendar time. Hamilton and Jordà (2002) define the hazard rate as the probability of a jump in t conditional on $t-1$ information,

$$h_t = \mathbb{P}[N(t) \neq N(t-1)|\Upsilon_{t-1}] \quad . \tag{3.15}$$

If the information set in $t-1$ only consists of past durations, the hazard rate and conditional expected durations are inversely related, such that

$$h_t = \frac{1}{\psi_{N(t-1)}} \quad . \tag{3.16}$$

Hamilton and Jordà (2002) extend (3.16) by allowing for predetermined variables $\mathbf{z}_{t-1}$ and

formulate the hazard rate in calendar time. Hence, the hazard rate becomes

$$h_t = \frac{1}{\psi_{N(t-1)} + \boldsymbol{\delta}_{\mathrm{X}}' \mathbf{z}_{t-1}} \quad . \tag{3.17}$$

The conditional likelihood function of the ACH model in Equations (3.14) and (3.17) is

$$f_{\mathrm{X}}(X_t|\Upsilon_{t-1};\boldsymbol{\theta}_{\mathrm{X}}) = \{h_t\}^{X_t}\{1-h_t\}^{(1-X_t)} \quad , \tag{3.18}$$

which is maximized by ML with respect to $\boldsymbol{\theta}_{\mathrm{X}} = (\alpha_{\mathrm{X}}, \beta_{\mathrm{X}}, \boldsymbol{\delta}_{\mathrm{X}})$.

Analogous to (3.11), (3.12) and (3.13), we briefly introduce the notation for the jump size model, AC-J. Denote by $J_{N(t)}$ the size of the $N(t)^{th}$ jump, by $\phi_{N(t)}$ the conditional expected jump size, and by $\eta_{N(t)}$ an iid innovation such that $J_{N(t)} = \phi_{N(t)}\eta_{N(t)}$. $\phi_{N(t)}$ evolves as

$$\phi_{N(t)} = \delta_{\mathrm{J}} + \alpha_{\mathrm{J}} J_{N(t)-1} + \beta_{\mathrm{J}}\phi_{N(t)-1} \quad , \tag{3.19}$$

imposing $\delta_{\mathrm{J}} > 0$, $\alpha_{\mathrm{J}} \geq 0$ and $\beta_{\mathrm{J}} \geq 0$. The AC-J likelihood function conditional on $X_t = 1$ is

$$f_{\mathrm{J}}(J_{N(t)}|X_t = 1, \Upsilon_{t-1};\boldsymbol{\theta}_{\mathrm{J}}) = f_\eta\left(\frac{J_{N(t)}}{\phi_{N(t)}};\boldsymbol{\theta}_\eta\right)\phi_{N(t)}^{-1} \quad , \tag{3.20}$$

with $\boldsymbol{\theta}_{\mathrm{J}} = (\delta_{\mathrm{J}}, \alpha_{\mathrm{J}}, \beta_{\mathrm{J}}, \boldsymbol{\theta}_\eta)$. The innovation $\eta_{N(t)}$ is assumed to be either exponential (AC-EJ), Weibull (AC-WJ) or Burr (AC-BJ) distributed.

Combining the models of continuous and jump variation to a joint model for realized volatility yields the joint density,

$$f(C_t, X_t, J_{N(t)}|\Upsilon_{t-1};\boldsymbol{\theta}) = f_{\mathrm{C}}(C_t|\Upsilon_{t-1};\boldsymbol{\theta}_{\mathrm{C}}) f_{\mathrm{X}}(X_t|\Upsilon_{t-1};\boldsymbol{\theta}_{\mathrm{X}}) f_{\mathrm{J}}(J_{N(t)}|X_t = 1, \Upsilon_{t-1};\boldsymbol{\theta}_{\mathrm{J}}) \ , \tag{3.21}$$

where $\boldsymbol{\theta}$ collects all parameters of the model. Henceforth, we will refer to (3.21) as the autoregressive conditional continuous jump (AC-CJ) model. Taking logs of (3.21) yields the log-likelihood function of the AC-CJ model

$$\mathscr{L}(\boldsymbol{\theta}) = \sum_{t=1}^{T} \ln f_{\mathrm{C}}(C_t|\Upsilon_{t-1};\boldsymbol{\theta}_{\mathrm{C}}) + \sum_{t=1}^{T} \ln f_{\mathrm{X}}(X_t|\Upsilon_{t-1};\boldsymbol{\theta}_{\mathrm{X}}) + \sum_{t=1}^{T} \ln f_{\mathrm{J}}(J_{N(t)}|X_t = 1, \Upsilon_{t-1};\boldsymbol{\theta}_{\mathrm{J}}). \tag{3.22}$$

As pointed out by Engle (2000), maximization of $\mathscr{L}(\boldsymbol{\theta})$ is equivalent to maximizing the three

terms on the right hand side of (3.22) separately if $\boldsymbol{\theta}_{\mathrm{C}}$, $\boldsymbol{\theta}_{\mathrm{X}}$ and $\boldsymbol{\theta}_{\mathrm{J}}$ have no parameters in common.

3.4.2 Forecast Setup

In this section we present the forecast methodology. We outline how single model forecasts for continuous and jump variation are combined to achieve point and density forecasts for total return variation. Further, we introduce the techniques of Diebold et al. (1998) to examine the accuracy of in-sample density forecasts. Their methods enjoy increasing popularity as they offer intuitive diagnostics to detect specification problems. Hence, the primary interest is to assess the models' fit and to conduct diagnostic checks by using density forecast evaluation methods.

Engle and Russell's (1998) framework as basis for AC-C and AC-J is appealing in terms of prediction, since point and density forecasts are obtained in a straightforward way. Point forecasts for continuous variation and jump variation result directly from iterating forward (3.12) and (3.19), respectively. Then, h-step ahead density forecasts are obtained by simply inserting the estimated parameters and the h-step ahead point forecast for C_{t+h} and $J_{N(t+h)}$ into the conditional densities in (3.13) and (3.20). The probability, of a jump in $t+h$, is computed using the ACH in (3.17). In the case of one-step forecasts this probability will be simply h_{t+1}. To obtain multiperiod ahead forecasts for the jump indicator, one needs to draw from the probability functions of $X_{t+1}, \cdots, X_{t+h}$. If there is no jump on day $t+1$ realized volatility is simply the continuous variation in $t+1$. A jump in $t+1$ implies $X_{t+1} = 1$ and realized volatility in $t+1$ is then the sum of continuous and jump variation, i.e. $RV_{t+1} = C_{t+1} + J_{N(t+1)}$. Point and density forecasts for RV_{t+h} are achieved by Monte Carlo simulation methods. The main idea is to simulate M forecast paths for the realized volatility from $h = 1$ up to the H^{th} horizon, which produces the sequence $\{RV_{t+h}^{(m)}\}_{m=1}^{M}$. From this sequence, point and density forecasts for each forecast horizon are computed. A detailed description for this forecasting procedure can be found in Appendix B.1.

Once a forecast density is achieved, Diebold et al.'s (1998) forecast evaluation methods can be applied. Denote by x_t a continuous forecast variable, by $\{p(x_t \mid \Upsilon_{t-1})\}$ a sequence of true densities and by $\{f(x_t \mid \Upsilon_{t-1})\}$ a sequence of density forecasts. As shown by Diebold et al. (1998) the correct density is weakly superior to all other forecasts. Hence, forecasters

will prefer the correct density irrespective of their loss function. Consequently, forecasts can be evaluated by testing the null hypothesis that the forecasting densities are equivalent to the true densities, i.e. whether

$$\{f(x_t \mid \Upsilon_{t-1})\} = \{p(x_t \mid \Upsilon_{t-1})\} \quad . \tag{3.23}$$

Obviously, the true densities are unobserved and testing whether Equation (3.23) holds seems impossible. However, this problem can be solved by using Rosenblatt's (1952) probability integral transform (PIT) result:

$$z_t = \int_{-\infty}^{x_t} f(u|\Upsilon_{t-1})du = F(x_t|\Upsilon_{t-1}) \quad . \tag{3.24}$$

Diebold et al. (1998) derive the distributional properties of of the probability transforms $\{z_t\}$ and show that under the null hypothesis the PIT sequence is iid $U(0,1)$ distributed.

Having obtained a $\{z_t\}$ sequence, it is possible to evaluate density forecasts of realized, continuous and jump variation by carrying out the following steps. First, we use the iid uniformity test proposed by Bauwens et al. (2004), which compares the number of PIT observations in classified bins to expected numbers if the data were indeed iid $U(0,1)$ distributed. Second, we follow Diebold et al. (1998) who suggest to augment formal tests of iid uniformity by diagnostic tools. Visual inspection of histograms and autocorrelograms of the PIT sequences help to detect the reasons for the rejection of the uniformity null hypothesis and assist to identify forecast failures. Third, since the PIT sequence possesses a MA$(h-1)$ autocorrelation structure, density forecasts for $h > 1$ will not be correct. Due to this, Diebold et al. (1998) recommend to remove the autocorrelation structure by partitioning the PIT sequence into subseries. For instance, for two-step ahead density forecasts, the subseries $\{z_1, z_3, z_5, \ldots\}$ and $\{z_2, z_4, z_6, \ldots\}$ should each be iid $U(0,1)$, although the sequence $\{z_1, z_2, z_3, z_4, \ldots\}$ is not. Iid uniformity tests are then based on the minimum and maximum test statistic within the h subseries. Critical values are computed by dividing the significance level by the forecast horizon. Finally, Diebold et al.'s (1998) idea is only applicable to continuous forecast variables. Due to this we require a discrete analog of the PIT to evaluate the probability function forecast delivered by the ACH model. To address this problem, we use Denuit and Lambert's (2005) methodology and "continue" the discrete variable X_t by

adding a uniform distributed random variable u_t, i.e. u_t is iid $U(0,1)$ and $X_t^* = X_t + (u_t - 1)$. Denuit and Lambert (2005) show that the PIT of the continued variable X_t^* is

$$z_t^* = F^*(X_t^*|\Upsilon_{t-1}) = F(X_t - 1|\Upsilon_{t-1}) + f(X_t|\Upsilon_{t-1})u_t \quad . \tag{3.25}$$

The discrete analog of the PIT theorem states that z_t^* is $U(0,1)$ if the probability forecast function is correctly specified.

3.5 Empirical Results

3.5.1 Estimation Results and Residual Diagnostics

As indicated in the methodology section AC-C, ACH and AC-J are estimated separately on an in-sample period from July 1^{st}, 2002 to January 31^{st}, 2006 for the futures and January 1^{st}, 2001 to January 31^{st}, 2006 for GM. The period from February 1^{st}, 2006 to June 30^{th}, 2006 is reserved for out-of-sample forecasts.

Table 3.5.2 reports the ML estimates of the AC-C, ACH and AC-J. The estimated AC-C coefficients α_{C} and β_{C} are statistically significant and vary between 0.18 and 0.36 for α_{C} and 0.59 and 0.82 for β_{C}. AC-C specifications indicate high persistence. The exponential, Weibull and Burr parameters in the AC-C and AC-J are statistically significant. The ACH estimates in Table 3.5.2 suggest strong persistence and serial correlation in durations with an average value over all time series of $\alpha_{\mathrm{X}} = 0.053$ and $\beta_{\mathrm{X}} = 0.855$.[6] The AC-J estimates α_{J} and β_{J} depend on the distributional assumptions for the innovation term. Generally, if a Burr distribution is assumed, β_{J} is higher and α_{J} is lower than the corresponding estimates when using an exponential or Weibull distribution.

As noted in Bauwens and Giot (2001), the imposed independence assumption of the error terms in the AC-C and AC-J models cannot be tested directly. However, autocorrelation tests on AC-C and AC-J residuals are conducted to detect potential specification errors. The residual autocorrelation tests in Table 3.5.3 show that the high correlation in C_t is removed for DAX, ESX, SP and GM. For the DAX some serial correlation is left in the higher order

[6] None of the included predetermined variables in Equation (3.17) appeared to be statistically significant. Since imprecise estimated parameters of exogenous variables might blur the forecast quality, the specifications in Table 3.5.2 are used for prediction.

	AC-EC	AC-WC	AC-BC	ACH	AC-EJ	AC-WJ	AC-BJ
				Panel A: DAX			
δ	0.000	0.000	0.000	2.743	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.769)	(0.000)	(0.000)	(0.000)
α	0.317	0.364	0.296	0.065	0.174	0.174	0.064
	(0.079)	(0.043)	(0.031)	(0.038)	(0.146)	(0.152)	(0.042)
β	0.663	0.587	0.708	0.827	0.750	0.748	0.924
	(0.083)	(0.047)	(0.030)	(0.090)	(0.193)	(0.201)	(0.041)
γ_1		2.034	4.362			1.004	4.176
		(0.045)	(0.243)			(0.042)	(0.633)
γ_2			1.235				2.510
			(0.149)				(0.628)
				Panel B: ESX			
δ	0.000	0.000	0.000	2.807	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(1.295)	(0.000)	(0.000)	(0.000)
α	0.195	0.200	0.181	0.016	0.307	0.350	0.051
	(0.046)	(0.028)	(0.023)	(0.021)	(0.115)	(0.104)	(0.012)
β	0.788	0.777	0.817	0.940	0.663	0.626	0.930
	(0.051)	(0.032)	(0.023)	(0.064)	(0.106)	(0.090)	(0.011)
γ_1		1.700	4.029			1.161	4.725
		(0.038)	(0.231)			(0.047)	(0.615)
γ_2			1.463				2.840
			(0.168)				(0.563)
				Panel C: SP			
δ	0.000	0.000	0.000	1.413	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.537)	(0.000)	(0.000)	(0.000)
α	0.309	0.312	0.314	0.100	0.258	0.362	0.167
	(0.063)	(0.027)	(0.030)	(0.033)	(0.218)	(0.113)	(0.037)
β	0.663	0.654	0.655	0.845	0.622	0.384	0.769
	(0.070)	(0.031)	(0.033)	(0.047)	(0.403)	(0.238)	(0.051)
γ_1		2.290	2.901			2.018	5.852
		(0.057)	(0.135)			(0.088)	(0.764)
γ_2			0.376				1.713
			(0.081)				(0.373)
				Panel D: GM			
δ	0.000	0.000	0.000	4.097	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.934)	(0.000)	(0.000)	(0.000)
α	0.262	0.270	0.247	0.032	0.485	0.501	0.088
	(0.055)	(0.037)	(0.031)	(0.040)	(0.167)	(0.171)	(0.032)
β	0.709	0.694	0.737	0.820	0.383	0.370	0.886
	(0.063)	(0.042)	(0.034)	(0.190)	(0.140)	(0.139)	(0.042)
γ_1		1.550	3.560			1.018	3.227
		(0.028)	(0.166)			(0.042)	(0.332)
γ_2			1.281				1.898
			(0.127)				(0.345)

Table 3.5.2: Maximum likelihood estimates of the AC-C, ACH and AC-J. The distributional assumption of the AC-C and AC-J innovations are exponential (E), Weibull (W) or Burr (B) with γ_1 and γ_2 as distributional parameters. Standard errors are reported in parentheses.

autocorrelations. The residuals of the AC-J do not exhibit any autocorrelation, indicating that the model accounts properly for jump size dynamics.

k	AC-EC	AC-WC	AC-BC	AC-EJ	AC-WJ	AC-BJ
			Panel A: DAX			
1	0.451	0.217	0.496	0.050	0.009	0.010
5	0.784	0.356	0.999	0.059	0.095	0.093
10	1.332	1.572	1.466	0.115	0.136	0.076
15	1.098	1.393	1.236	0.094	0.124	0.084
			Panel B: ESX			
1	0.065	0.037	0.035	0.147	0.202	0.041
5	0.387	0.380	0.327	0.375	0.433	0.243
10	0.447	0.462	0.381	0.290	0.319	0.159
15	0.556	0.536	0.579	0.263	0.285	0.140
			Panel C: SP			
1	0.520	0.327	0.226	0.151	0.003	0.426
5	0.302	0.234	0.196	0.201	0.138	0.277
10	0.351	0.298	0.281	0.210	0.268	0.229
15	0.559	0.545	0.505	0.301	0.284	0.402
			Panel D: GM			
1	0.326	0.317	0.442	0.340	0.208	0.001
5	0.638	0.568	0.716	0.260	0.210	0.158
10	0.522	0.436	0.640	0.203	0.171	0.179
15	0.592	0.576	0.615	0.347	0.316	0.218

Table 3.5.3: Autocorrelation tests of AC-C and AC-J estimated residuals. The table reports the result of the ratio $\frac{LB(k)}{\chi^2(k)}$ with $LB(k)$ denoting the Ljung-Box statistic and $\chi^2(k)$ the 5% critical value. The ratio is computed for the first, fifth, tenth and fifteenth order serial correlation (k). The null hypothesis of no autocorrelation is not rejected for values smaller than one.

3.5.2 Density Forecast Evaluation

The forecast techniques of Diebold et al. (1998) outlined in Section 3.4.2 can be conveniently used for diagnostic checking. To analyze the models' fit, we focus on an in-sample density forecast evaluations. As mentioned above, the rejection of the null hypothesis of uniformity in the PIT sequence does not provide guidance concerning the reasons. Therefore, Diebold et al. (1998) recommend to use autocorrelograms and histograms of the PIT sequences as diagnostic tools to detect specification errors associated with a model's density forecasts. For instance, inverse $\cup$-shaped PIT histograms suggest that we observe insufficient large and small future values of volatility compared to what is predicted by the model. Significant autocorrelation in the PIT series indicates that the model is not able to account properly for the dynamics of the variation measures.

Figure 3.5.3 depicts twenty-bin histograms of the PIT sequence for one-day ahead forecasts implied by AC-EC, AC-WC, AC-BC, ACH and AC-BJ. The histograms are based on the

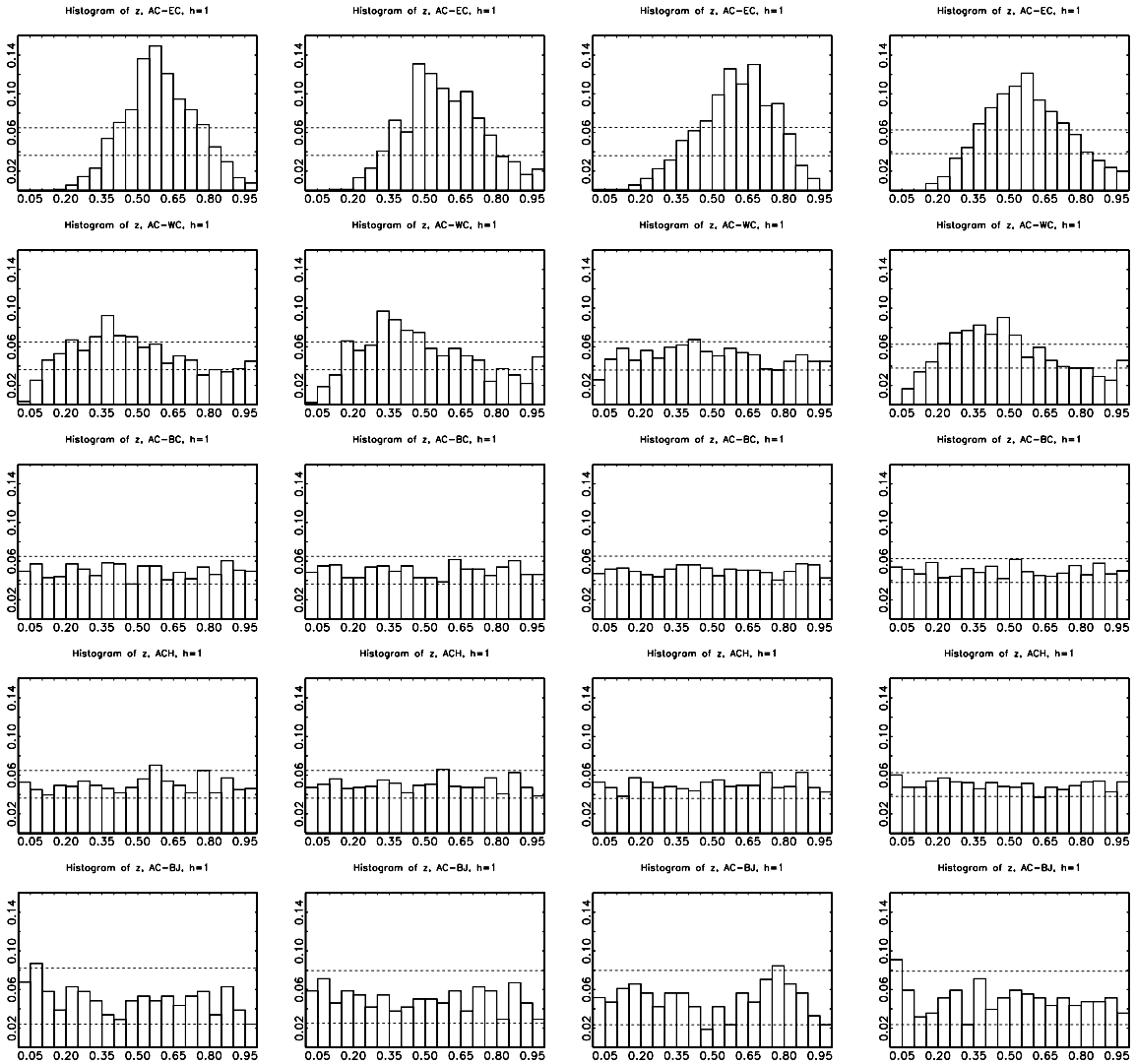

Figure 3.5.3: Histograms of the one-step ahead forecast PIT sequence for the AC-EC, AC-WC, AC-BC, ACH and AC-BJ. The figure shows twenty-bin histograms of the PIT sequence for one-step ahead forecast horizon, i.e. $h = 1$, for DAX (first column), ESX (second column), SP (third column) and GM (fourth column). Upper and lower bound (displayed in horizontal dashed lines) of the 95% confidence interval are computed from the 0.025 and 0.975 quantiles of a binomial distribution with $p = 0.05$ and number of draws equal to n, where n is the number of observations in each subseries. The first panel row depicts the histograms for the AC-EC, the second for AC-WC, the third for AC-BC, the fourth for ACH and the last for the AC-BJ model.

original PIT sequence for one-step forecasts. For multi-step forecasts ($h > 1$), the minimum and the maximum relative frequency of the thinned h subseries in each of the twenty histogram bins is plotted. Histograms for two- and four-day ahead forecasts are deferred to Appendix B.3. As seen from Figure 3.5.3, the histograms of the PIT sequence resulting from the AC-EC

are inverse ∪-shaped, while these from AC-WC are rather ragged. Thus, histograms of AC-EC and AC-WC point at specification errors due to incorrect distributional assumptions. In contrast, the histograms resulting from AC-BC do not show any deviations from uniformity implying that the dynamics of continuous variation are well captured. Histogram bars of the continued ACH PIT sequence lie mostly within the 95% confidence bounds (see also Figures B.3.1 and B.3.2 in Appendix B.3). Same conclusions can be drawn for the PIT histograms of the AC-BJ.

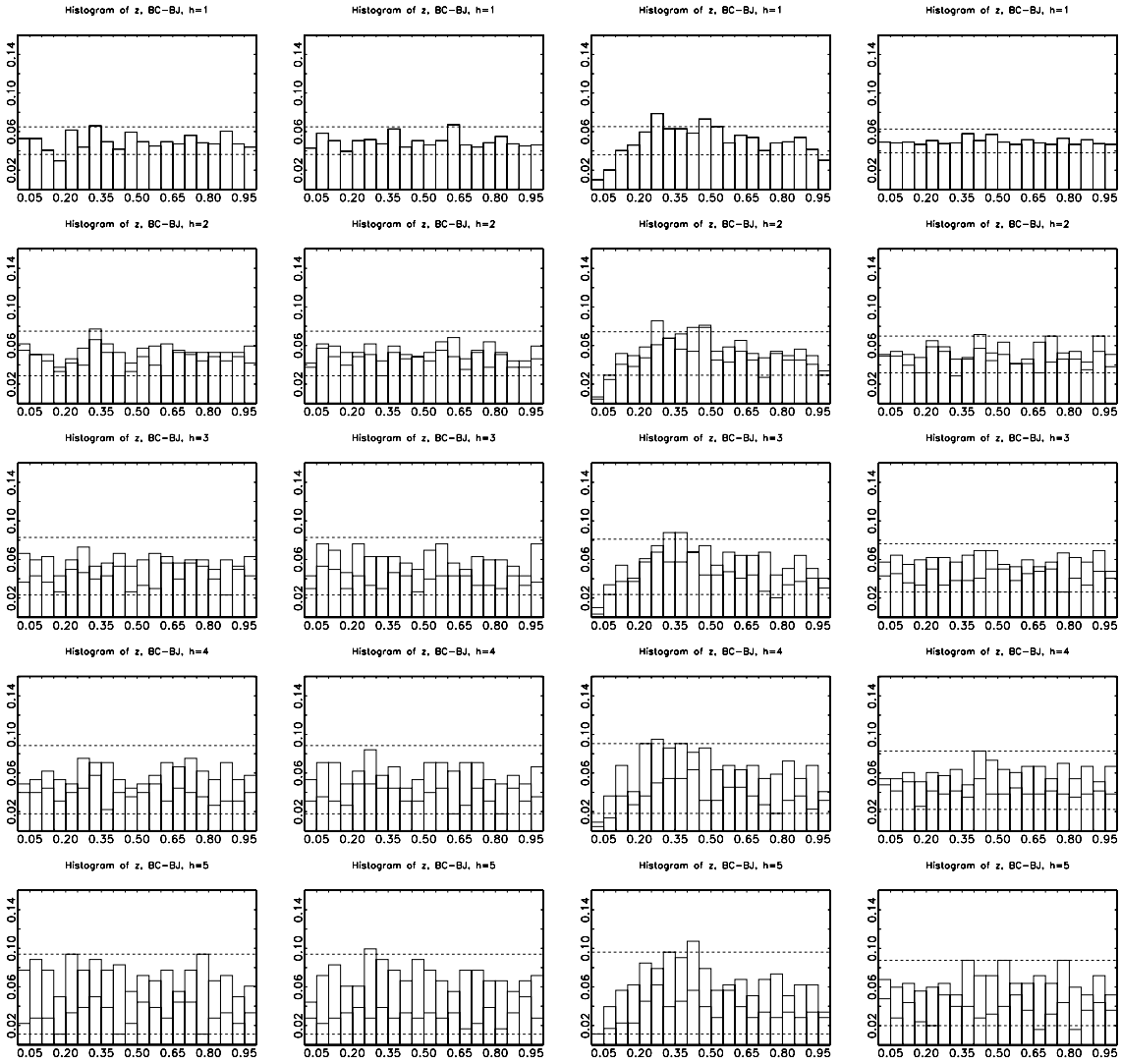

Figure 3.5.4: Histograms of the PIT sequence for the BC-BJ. The first panel column contains the results for the DAX, the second for ESX, the third for SP and the fourth for GM. See caption of Figure 3.5.3 for explanations.

From the single model evaluation it is inferred that the assumption of Burr distributed errors in the AC-C and AC-J fits the data best. Hence, we combine AC-BC, ACH and AC-BJ to the BC-BJ model and obtain density forecasts for realized volatility following the description in Section 3.4.2. The resulting PIT histograms are reported in Figure 3.5.4. For DAX, ESX and GM the figures do not exhibit deviations from uniformity. However, the histograms for SP suggest that insufficient small values of realized volatility are predicted by the model.

h	1	2	4	1	2	4
$1\%/h$	36.2	38.6	40.9	36.2	38.6	40.9
$5\%/h$	30.1	32.9	35.4	30.1	32.9	35.4
Model	Panel A: DAX			Panel B: ESX		
AC-BC	15.8	23.4	28.9	14.2	15.6	31.3
ACH	21.1	16.2	17.9	16.0	19.4	34.4
AC-BJ	17.3	18.4	23.1	12.3	16.3	26.8
BC-BJ	24.2	20.1	17.6	15.8	16.5	19.1
	Panel C: SP			Panel D: GM		
AC-BC	7.6	17.3	22.2	15.1	36.3	33.5
ACH	12.9	22.7	21.6	12.8	28.2	24.9
AC-BJ	22.6	20.0	35.7	20.6	22.6	32.6
BC-BJ	94.1	56.0	41.0	5.1	41.2	19.7

Table 3.5.4: Results of iid uniformity test for the PIT sequence. For each forecast horizon h the PIT sequence is split into h subseries which are iid $U(0,1)$ under the null hypothesis of a correct density forecast. Bauwens et al.'s (2004) test statistic for iid uniformity is computed for each subseries. The test is based on the result that under the null of iid $U(0,1)$ behavior of the PIT sequence the joint distribution of the heights of the PIT histogram is multinomial, i.e. $f(n_i) = \binom{n}{n_i} p^{n_i} (1-p)^{n-n_i}$ where n gives the number of observations (in each subseries), n_i the number of observations in the i^{th} histogram bin and $p = 1/m$ with m the number of histogram bins. We use $m = 20$. The statistic $\sum_{i=1}^{m} \frac{(n_i - np)^2}{np}$ is under the null hypothesis asymptotically $\chi^2(m-1)$ distributed. The table reports the largest test statistic computed from h subseries. The critical values are computed by dividing the significance levels by h.

Table 3.5.4 reports the uniformity test statistic suggested by Bauwens et al. (2004). The idea of Bauwens et al.'s (2004) test is to compare the number of observations in the bins of the PIT histogram with the expected values given that the PIT sequence would indeed be iid $U(0,1)$. The caption of Table 3.5.4 explains computational details. Since the assumed Burr distribution for the AC-C and AC-J innovations yields most accurate density forecasts for realized, continuous and jump size variation, we omit in Table 3.5.4 the results for the exponential and Weibull distribution. Bauwens et al.'s (2004) test is applied to the (continued) PIT sequences computed up to four-day ahead density forecasts and confirms the conclusions drawn from the visual inspection of the histograms. For DAX and ESX the null hypothesis

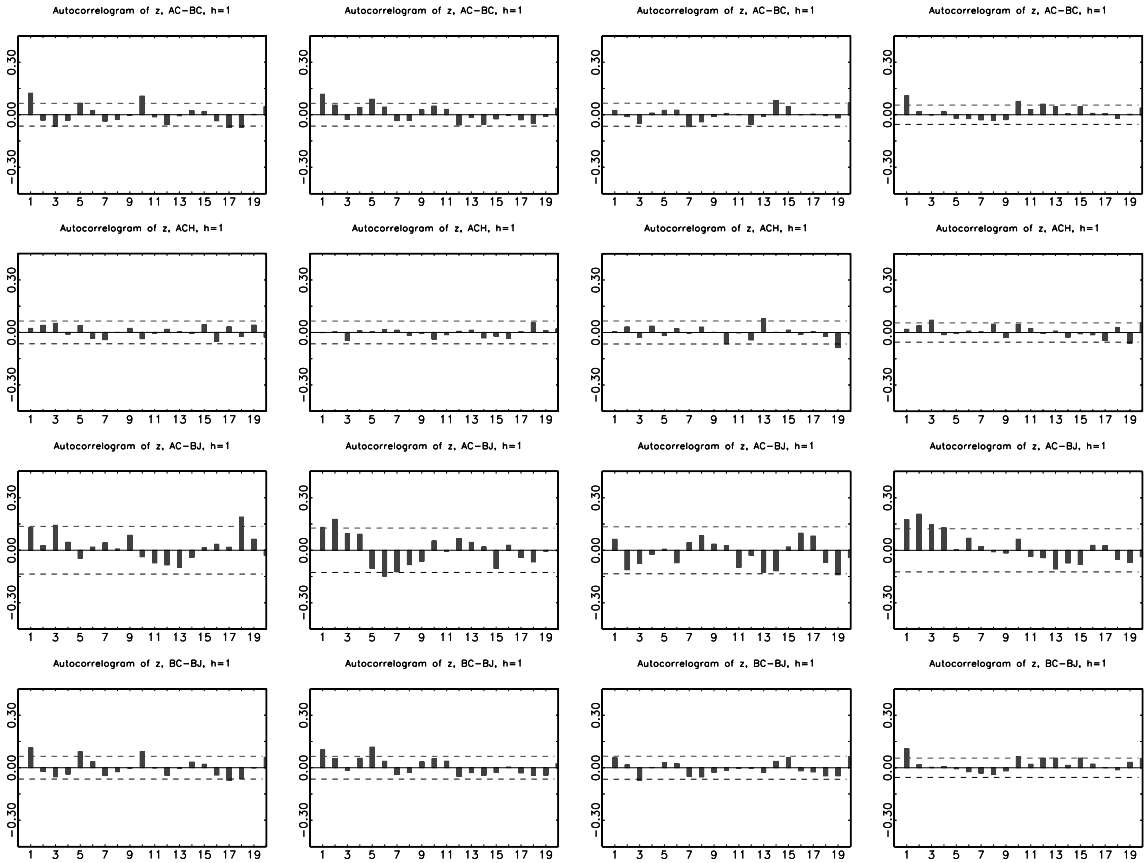

Figure 3.5.5: Autocorrelograms of the PIT sequence for the AC-BC, ACH, AC-BJ and BC-BJ. The figure shows autocorrelograms of the PIT sequence for one-step ahead forecast horizon, i.e. $h = 1$, for DAX (first column), ESX (second column), SP (third column) and GM (fourth column). The horizontal lines superimposed on the autocorrelograms mark the 95% confidence intervals. The first panel column contains the DAX , the second the ESX, the third the SP and the fourth the GM results. The first panel row depicts the autocorrelograms for the AC-BC, the second for ACH, the third for AC-BJ and the last for the BC-BJ model.

that the PIT sequences delivered by the AC-BC, ACH and AC-BJ are uniform cannot be rejected at any conventional significance level. For SP and GM the same conclusion can be drawn for a $1\%/h$ significance level. Tests for uniformity of PIT sequence implied by the BC-BJ do not detect misspecifications for DAX, ESX and GM. However, as already visible from the PIT histograms, the null of uniformity is rejected for SP.

Finally, we investigate the dynamics of PIT sequence obtained from AC-BC, ACH, AC-BJ and BC-BJ to detect specification errors. We compute autocorrelations of the PIT for a one-step ahead forecast horizon and plot them in Figure 3.5.5. Although minor autocorrelation is left in some PIT sequences, we conclude that the overall dynamics of the variation measures are captured quite well.

3.5.3 Out-of-sample Point Forecast Performance

In this section we analyze the model's ability of out-of-sample point predictions. To compare daily return forecasts delivered by the AC-CJ, we estimate three alternative volatility models. First, motivated by the work of Hansen and Lunde (2005), we consider a simple GARCH(1,1) model with normally distributed errors. The second model is the heterogeneous autoregressive realized volatility (HAR-RV) model suggested by Corsi (2004). This AR-type process includes lags corresponding to the time horizons of one day, one aggregated week and one aggregated month period. Finally, we estimate an autoregressive conditional realized volatility (AC-R) model, which is identical to the AC-C model in (3.11), (3.12) and (3.13), but the variable of interest is realized variation.[7] The estimation results of the benchmark models can be found in Appendix B.2.

To evaluate the point forecast performance, the sequence of volatility forecasts, $\{\widehat{RV}_{t+h}\}$, is compared to the sequence of actual realized volatility, $\{RV_{t+h}\}$. We use the root mean squared error (RMSE),

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}[RV_{t+h} - \widehat{RV}_{t+h}]^2} \quad , \tag{3.26}$$

as loss function. To test whether the difference between forecast errors of a null model and an alternative model is significant, we apply the modified version of the Diebold-Mariano (1995) test suggested by Harvey et al. (1997).

Table 3.5.5 contains out-of-sample RMSEs and modified Diebold-Mariano (DM) tests for one- and ten-day ahead volatility forecasts. The null model of the DM test is an AC-CJ model with exponential distributed innovations in the AC-C and Burr distributed innovations in the AC-J (EC-BJ). Hence, negative values for the DM test results from smaller forecast errors of EC-BJ compared to the alternative model. Considering the one-day ahead forecasts, the GARCH produces higher RMSEs compared to EC-BJ for all time series. The differences between the GARCH and the EC-BJ model are significant on a 5% level for ESX. For ten-day ahead forecasts the EC-BJ is more accurate than the GARCH for ESX and GM. One-day ahead AC-ER RMSEs are higher compared to EC-BJ for DAX and ESX and for ten-day

[7] The assumed distributions of the AC-R innovations are either exponential (AC-ER), Weibull (AC-WR) or Burr (AC-BR). Since the exponential distribution turns out to be most suitable in terms of forecasting, we refrain from presenting AC-WR and AC-BR results.

	1-day ahead			10-day ahead		
Model	RMSE	DM	p-value	RMSE	DM	p-value
			Panel A: DAX			
EC-BJ	0.455			0.624		
BC-BJ	0.464	-1.714	0.045	0.666	-1.603	0.056
GARCH	0.474	-0.591	0.278	0.605	0.654	0.743
AC-ER	0.458	-0.466	0.321	0.623	0.140	0.556
HAR-RV	0.494	-1.299	0.098	0.665	-0.817	0.208
			Panel B: ESX			
EC-BJ	0.567			0.694		
BC-BJ	0.575	-1.941	0.027	0.712	-1.246	0.108
GARCH	0.676	-2.350	0.010	0.778	-0.959	0.170
AC-ER	0.570	-0.873	0.192	0.700	-1.735	0.043
HAR-RV	0.591	-1.205	0.115	0.736	-0.942	0.174
			Panel C: SP			
EC-BJ	0.274			0.318		
BC-BJ	0.275	-0.032	0.487	0.319	-0.382	0.352
GARCH	0.286	-0.548	0.293	0.309	0.232	0.591
AC-ER	0.274	0.075	0.530	0.322	-0.535	0.297
HAR-RV	0.271	0.268	0.606	0.319	-0.021	0.492
			Panel D: GM			
EC-BJ	3.068			3.291		
BC-BJ	3.066	0.145	0.557	3.323	-0.403	0.344
GARCH	3.267	-1.166	0.123	3.501	-0.634	0.264
AC-ER	3.066	0.148	0.559	3.307	-0.634	0.264
HAR-RV	3.101	-0.511	0.305	3.283	0.051	0.520

Table 3.5.5: Point forecast evaluation for realized volatility: out-of-sample RMSEs. The table depicts root mean squared errors (RMSEs) of the daily realized variance forecast for one- and ten-day ahead horizons delivered by EC-BJ, BC-BJ, GARCH, HAR-RV and AC-ER. Further, the table contains the result of the modified Diebold-Mariano tests (DM) and corresponding p-values. The null model is the EC-BJ. Thus, negative values for the DM test corresponds to smaller RMSEs of the EC-BJ model.

ahead the AC-ER produces less precise forecasts then the EC-BJ for ESX, SP and GM. For the ten-day ahead ESX forecast the DM null hypothesis can be rejected on a 5% significance level. The EC-BJ is superior to the HAR-RV in terms of lower one- and ten-day ahead RMSEs, for DAX, ESX and GM (one-day ahead) and for DAX, ESX and SP (ten-day ahead), but the differences are not statistically significant. Generally, it can be concluded that the EC-BJ model is at least as accurate as the alternative models and in some cases even more accurate.

3.6 Conclusion

The present paper is linked to Andersen et al. (2007b) and Bollerslev et al. (2009) who disentangle return volatility into a continuous and a jump variation component and model

realized volatility by simple time series methods. We introduce a model that accounts for continuous variation and jumps as well. Since continuous variation is strictly positive and autocorrelated, we apply Engle and Russell's (1998) framework to continuous variation and refer to it as autoregressive conditional continuous variation (AC-C) model. The jump variation process is conceived as a marked point process. Time between successive jumps is modeled by Hamilton and Jordà's (2002) autoregressive conditional hazard (ACH) model. For the jump sizes, we apply the autoregressive conditional jump (AC-J) model. Combining the three models forms the AC-CJ model for total return variation. To assess the accuracy of density forecasts of realized, continuous and jump variation, the forecast evaluation methods of Diebold et al. (1998) are applied.

The main findings of the empirical results using high frequency intraday data of the DAX, ESX, SP and GM can be summarized as follows. The estimation of AC-C, ACH and AC-J models deliver sensible parameter estimates and encouraging results in terms of diagnostics. Especially AC-C and AC-J specifications with assumed Burr distributed errors (AC-BC and AC-BJ) are capable to capture the dynamics of continuous and jump variation. Density forecast evaluations confirm the suitability of the AC-BC and AC-BJ towards modeling the evolution of the continuous and jump size variation. Similarly, as shown by the probability forecast evaluation, the ACH qualifies as a suitable model for jump durations. For the point forecast evaluation it can be concluded that the EC-BJ model is at least as accurate as the GARCH(1,1), HAR-RV and AC-R model and in some cases even more accurate. Overall, these results suggest that the framework proposed in this paper is useful for modeling and forecasting volatility.

Appendix B

B.1 Density Forecasts

Density forecasts for RV_{t+h} are achieved by using Monte Carlo simulation methods. Figure B.1.1 illustrates the simulation strategy. Conditional on time t information, one can draw a one-step ahead simulated value for the continuous variation $C_{t+1}^{(1)}$ from its conditional distribution using a draw for $\varepsilon_{t+1}^{(1)}$ and the estimated parameters $\hat{\boldsymbol{\theta}}_{\text{C}}$. The superscript indicates the first of M simulation steps. By drawing $X_{t+1}^{(1)}$ from the probability distribution of X_{t+1}, it is determined whether there is no jump in $t+1$, implying $RV_{t+1}^{(1)} = C_{t+1}^{(1)}$, or whether there is a jump in $t+1$. In the latter case, a simulated value for $J_{t+1}^{(1)}$ is obtained from a draw for $\eta_{t+1}^{(1)}$ using the estimates for $\hat{\boldsymbol{\theta}}_{\text{J}}$. Then realized volatility in $t+1$ is the sum of continuous and jump variation, i.e. $RV_{t+1}^{(1)} = C_{t+1}^{(1)} + J_{t+1}^{(1)}$ and $N(t+1)^{(1)} = N(t)+1$. Iterating conditional on time $t+1$ Equations (3.12), (3.17) and (3.19) forward, yields two-period ahead forecasts. Repeating this procedure results in a sequence $\{RV_{t+h}^{(1)}\}_{h=1}^{H}$. One can go back to the start and generate a second path, $\{RV_{t+h}^{(2)}\}_{h=1}^{H}$. Hence, drawing from the density of RV_{t+h},

$$
\begin{aligned}
& f(RV_{t+h}|\Upsilon_t) \qquad\qquad (\text{B.1}) \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(RV_{t+h}|RV_{t+h-1}, \cdots, RV_{t+1}, \Upsilon_t) f(RV_{t+h-1}|RV_{t+h-2}, \cdots, RV_{t+1}, \Upsilon_t) \cdots \\
&\qquad\qquad \cdots f(RV_{t+2}|RV_{t+1}, \Upsilon_t) f(RV_{t+1}|\Upsilon_t) \quad dRV_{t+h-1} \cdots dRV_{t+1} \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(RV_{t+h}, RV_{t+h-1}, \cdots, RV_{t+1}|\Upsilon_t)\, dRV_{t+h-1} \cdots dRV_{t+1} \quad ,
\end{aligned}
$$

M times produces a sequence of $\{RV_{t+h}^{(m)}\}_{m=1}^{M}$ for $h = 1, ..., H$. To conduct a density forecast evaluation, a h-step ahead empirical quantile distribution is computed using Monte Carlo

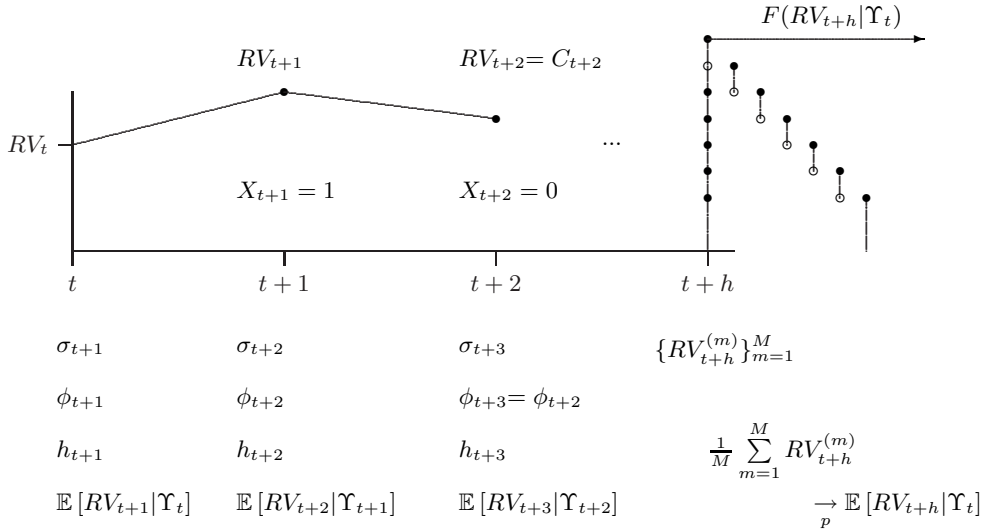

Figure B.1.1: Simulation of point and density forecasts for the AC-CJ model. The figure illustrates the Monte Carlo simulation to obtain h-step ahead point and density forecasts. Starting from day t, the figure depicts a schematic diagram of the Monte Carlo procedure. In t the single models deliver one-period ahead forecasts. A draw from the density of X_{t+1} results in a jump in $t+1$, hence, $X_{t+1} = 1$ and $RV_{t+1} = C_{t+1} + J_{N(t+1)}$ with $C_{t+1} = \sigma_{t+1}\varepsilon_{t+1}$ and $J_{t+1} = \phi_{t+1}\eta_{t+1}$. Forward iterations of Equations (3.12), (3.17) and (3.19) reveals a no-jump day in $t+2$ with $RV_{t+2} = C_{t+2}$ and $C_{t+2} = \sigma_{t+2}\varepsilon_{t+2}$. The repeated simulation of volatility paths attains a sequence of volatility values in $t+h$, $\{RV_{t+h}^{(m)}\}_{m=1}^{M}$. Averaging over these draws from $f(RV_{t+h}|\Upsilon_t)$ yields a point estimate and computing empirical quantile functions gives a distribution forecast sketched on the right hand side of the figure.

integration. The distribution function is given by,

$$\begin{aligned} F(RV_{t+h}|\Upsilon_t) &= \int_{-\infty}^{x_{t+h}} f(RV_{t+h}|\Upsilon_t)\, dRV_{t+h} \\ &= \int_{-\infty}^{\infty} \mathbb{1}(RV_{t+h} \leq x_{t+h}) f(RV_{t+h}|\Upsilon_t)\, dRV_{t+h} \end{aligned} \tag{B.2}$$

where $\mathbb{1}(\cdot)$ is an indicator function that takes value one if the expression $RV_{t+h} \leq x_{t+h}$ is true and zero otherwise. Plugging (B.1) into (B.2) yields a conditional expectation of $RV_{t+h} \leq x_{t+h}$

$$\begin{aligned} & F(RV_{t+h}|\Upsilon_t) \\ = & \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \mathbb{1}(RV_{t+h} \leq x_{t+h}) f(RV_{t+h}, RV_{t+h-1}, ..., RV_{t+1}|\Upsilon_t)\, dRV_{t+1}...dRV_{t+h-1}dRV_{t+h} \\ = & \mathbb{E}[\mathbb{1}(RV_{t+h} \leq x_{t+h})|\Upsilon_t] \quad . \end{aligned} \tag{B.3}$$

If M increases infinitely an estimator for (B.3) is simply the sample mean,

$$\frac{1}{M}\sum_{i=1}^{M}\mathbb{1}(RV_{t+h}^{(m)} \leq x_{t+h}) \quad \underset{p}{\rightarrow} \quad \mathbb{E}\left[\mathbb{1}(RV_{t+h} \leq x_{t+h})|\Upsilon_t\right] \quad . \tag{B.4}$$

A point forecast for RV_{t+h} is given by the conditional expectation of RV_{t+h},

$$\begin{aligned}\mathbb{E}\left[RV_{t+h}|\Upsilon_t\right] &= \int_{-\infty}^{\infty} RV_{t+h} f(RV_{t+h}|\Upsilon_t)\, dRV_{t+h} \\ &= \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} RV_{t+h} f(RV_{t+h}, RV_{t+h-1}, ..., RV_{t+1}|\Upsilon_t)\, dRV_{t+h} dRV_{t+h-1} ... dRV_{t+1}.\end{aligned} \tag{B.5}$$

As M grows, the sample mean over $RV_{t+h}^{(m)}$ gives a point forecast estimator that converges to the conditional expectation in (B.5),

$$\frac{1}{M}\sum_{i=1}^{M} RV_{t+h}^{(m)} \quad \underset{p}{\rightarrow} \quad \mathbb{E}\left[RV_{t+h}|\Upsilon_t\right] \quad \text{for} \quad m = 1, ..., M \quad . \tag{B.6}$$

B.2 Additional Table

AC-ER		GARCH		HAR-RV		AC-ER		GARCH		HAR-RV	
Panel A: DAX						Panel B: ESX					
δ_R	0.000	δ_G	0.000	c	-0.276	δ_R	0.000	δ_G	0.000	c	-0.409
	(0.000)		(0.000)		(0.129)		(0.000)		(0.000)		(0.164)
α_R	0.292	α_G	0.063	$\beta^{(d)}$	0.231	α_R	0.202	α_G	0.061	$\beta^{(d)}$	0.077
	(0.076)		(0.014)		(0.040)		(0.050)		(0.012)		(0.040)
β_R	0.687	β_G	0.929	$\beta^{(w)}$	0.496	β_R	0.782	β_G	0.931	$\beta^{(w)}$	0.598
	(0.080)		(0.016)		(0.068)		(0.054)		(0.013)		(0.074)
				$\beta^{(m)}$	0.245					$\beta^{(m)}$	0.281
					(0.055)						(0.062)
Panel C: SP						Panel D: GM					
δ_R	0.000	δ_G	0.000	c	-0.586	δ_R	0.000	δ_G	0.000	c	-0.623
	(0.000)		(0.000)		(0.182)		(0.000)		(0.000)		(0.218)
α_R	0.358	α_G	0.055	$\beta^{(d)}$	0.204	α_R	0.257	α_G	0.085	$\beta^{(d)}$	0.251
	(0.087)		(0.008)		(0.041)		(0.055)		(0.010)		(0.034)
β_R	0.619	β_G	0.934	$\beta^{(w)}$	0.559	β_R	0.717	β_G	0.903	$\beta^{(w)}$	0.377
	(0.093)		(0.012)		(0.068)		(0.062)		(0.010)		(0.060)
				$\beta^{(m)}$	0.177					$\beta^{(m)}$	0.299
					(0.052)						(0.054)

Table B.2.1: Estimates of the AC-ER, HAR-RV and GARCH(1,1) model. The GARCH(1,1) specification for the mean and variance equation for the daily return series is

$$\begin{aligned} r_t &= \epsilon_t \\ \sigma_t^2 &= \delta_G + \alpha_G \epsilon_{t-1}^2 + \beta_G \sigma_{t-1}^2 \quad , \end{aligned}$$

with ϵ_t being normally distributed with mean zero and variance σ_t^2. The parameters of the GARCH Model are estimated by ML. The HAR-RV model of Corsi (2004) is estimated using logarithmic realized volatilities and is specified as:

$$RV_{t+1}^{(d)} = c + \beta^{(d)} RV_t^{(d)} + \beta^{(w)} RV_t^{(w)} + \beta^{(m)} RV_t^{(m)} + \omega_{t+1d} \quad ,$$

where $RV_t^{(d)}$ is the daily realized volatility in t and a weekly aggregated realized volatility estimator is $RV_t^{(w)} = \frac{1}{5}\left(RV_{t-1} + \cdots + RV_{t-5}\right)$. A monthly aggregated $RV_t^{(m)}$ is computed analogously. The estimates δ_R, α_R and β_R of the autoregressive conditional realized volatility with exponential distributed innovations (AC-ER) are estimated by ML. Standard errors are reported in parentheses.

B.3 Additional Figures

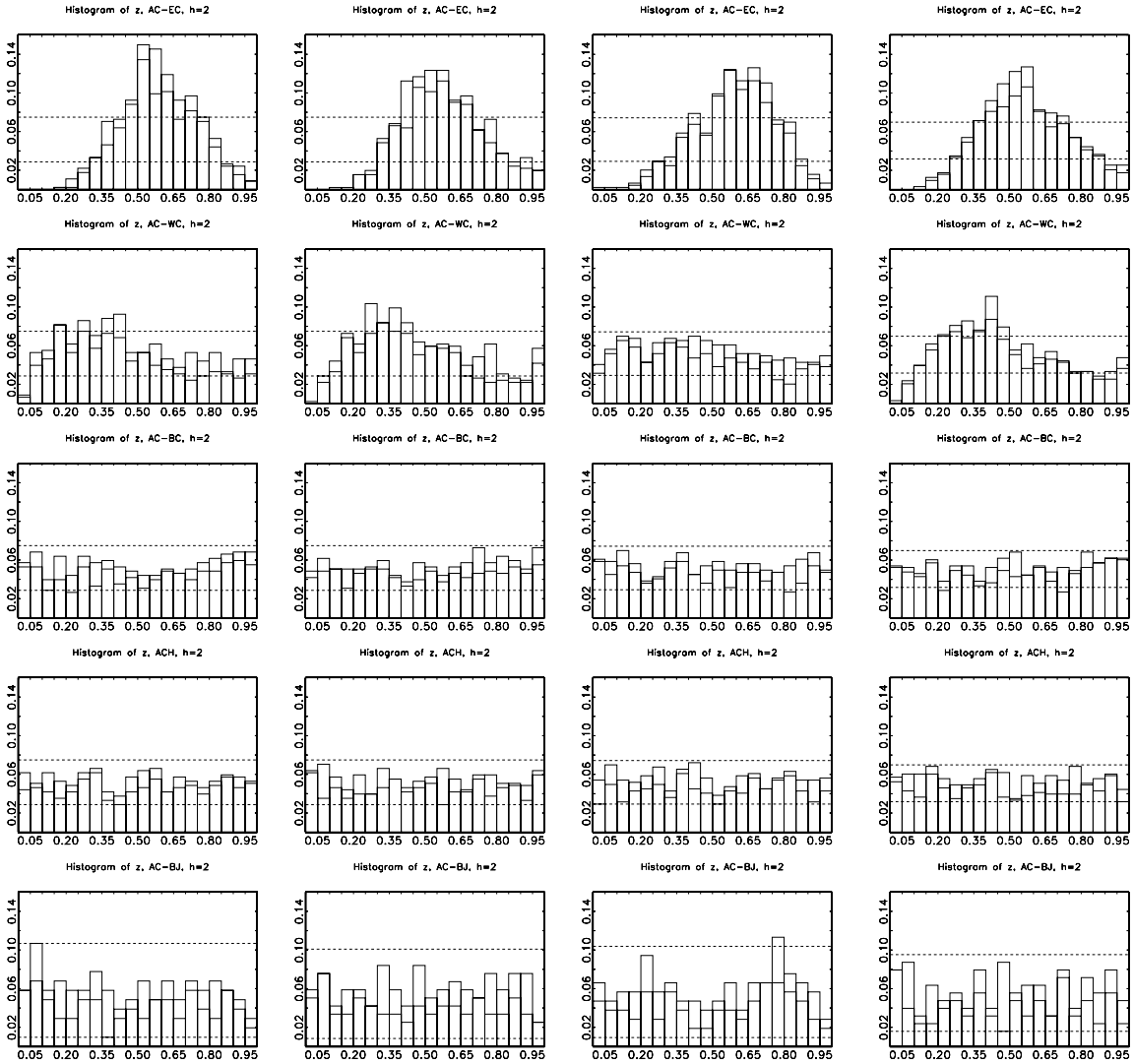

Figure B.3.1: Histograms of the two-step ahead forecast PIT sequence for the AC-EC, AC-WC, AC-BC, ACH and AC-BJ. The figure shows twenty-bin histograms of the PIT sequence for a two-step ahead forecast horizon, i.e. $h = 2$, for DAX (first column), ESX (second column), SP (third column) and GM (fourth column). For $h > 1$ the data are thinned into h subseries which are iid $U(0, 1)$ under the null hypothesis of a correct density forecast. The horizontal solid lines show the minimum and the maximum relative frequency of the h subseries in each of the twenty histogram bins. Upper and lower bound (displayed in horizontal dashed lines) of the 95% confidence interval are computed from the $0.025/h$ and $0.975/h$ quantiles of a binomial distribution with $p = 0.05$ and number of draws equal to n, where n is the number of observations in each subseries. The first panel row depicts the histograms for the AC-EC, the second for AC-WC, the third for AC-BC, the fourth for ACH and the last for the AC-BJ model.

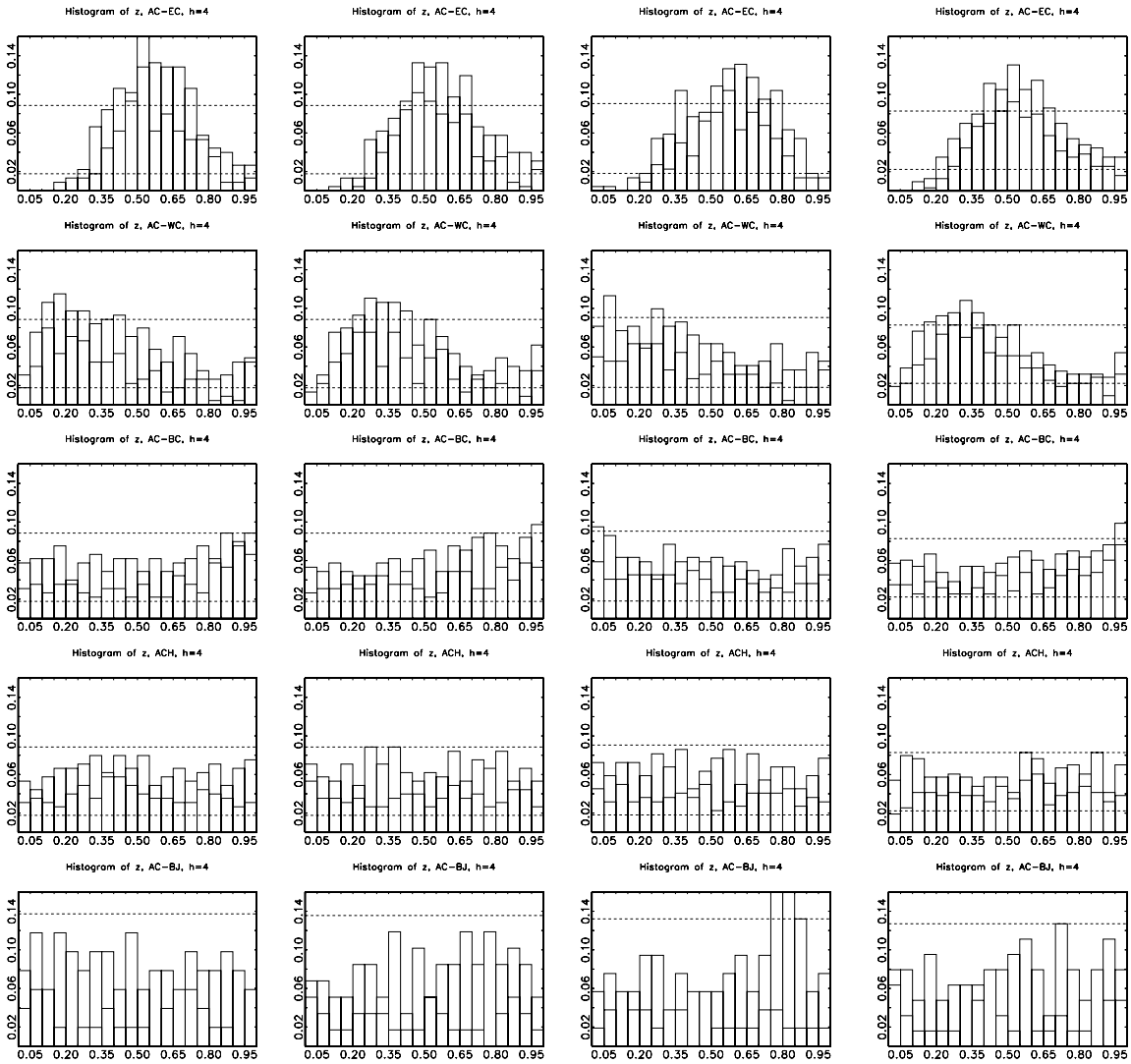

Figure B.3.2: Histograms of the four-step ahead forecast PIT sequence for the AC-EC, AC-WC, AC-BC, ACH and AC-BJ. The figure shows twenty-bin histograms of the PIT sequence for a four-step ahead forecast horizon, i.e. $h = 4$, for DAX (first column), ESX (second column), SP (third column) and GM (fourth column). See caption of Figure B.3.1 for explanations.

Chapter 4

International Price Discovery in Stock Markets - A Unique Intensity Based Information Share

This paper proposes a new information share for price discovery based on Russell's (1999) autoregressive conditional intensity model. While previous studies rely on equally spaced high frequency data, we use the information conveyed by trade intensities to determine a market's contribution to price discovery. Thereby, we account for the irregular nature of transaction data. Moreover, in contrast to the commonly applied Hasbrouck (1995) approach, which yields lower and upper bounds for information shares, our model delivers a unique measure. Our empirical application to US-listed Canadian stocks supports previous evidence for the home market leadership in price discovery.

This chapter is based on the article "International Price Discovery in Stock Markets - A Unique Intensity Based Information Share" by K. Kehrle and F. J. Peter (2009).

4.1 Introduction

According to Coffee (2002), increasing globalization and improved technology will lead to a decay in the number of securities exchanges around the world. Small national exchanges will lose their share in trading to large international exchanges, which provide a more efficient trading environment. Carpentier et al. (2007) examine this development for the Canadian stock exchanges with respect to the US markets. They report a rapidly growing share of US markets in trades of Canadian stocks, up to the point where interlisted stocks are absorbed by the foreign market and delisted on the home market. These developments foreshadow small national stock exchanges as markets for illiquid stocks that failed to attract investors on the large markets (see Gaa et al. 2002). Thus, within the context of international cross-listed stocks, it is of paramount interest to national stock exchanges to remain the dominant market with regard to price discovery.[1] The competition among smaller national and the giant US markets for the leadership in price discovery of interlisted stocks has therefore grown immensely and has stirred up an increasing field of research.

The main innovations of this paper are summarized as follows. We apply Russell's (1999) autoregressive conditional intensity model (ACI) and develop a new information share that measures the home and foreign market share in price discovery. By using a bivariate intensity approach, we account for the informational content of time between consecutive trades and the timing interdependencies between the two markets' transaction processes. In contrast to the commonly applied Hasbrouck (1995) methodology we account for the irregularity of the data and deliver a unique information share rather than lower and upper bounds. We apply our information share to Canadian stocks that are traded on the Toronto Stock Exchange (TSX) and cross-listed on the New York Stock Exchange (NYSE).

Evidence from previous studies suggests that the greatest part of price discovery for a cross-listed stock takes place in the home market. Eun and Sabherwal (2003) examine a sample of US listed Canadian stocks based on the relative adjustment of prices in a market to deviations from the equilibrium price. They conclude, while the contribution of the US market cannot be neglected, the home market leads price discovery. Adjustment coefficients as a measure for price discovery, however, are criticized, since they do not account for the contemporaneous correlations and variances of market's price innovation and focus only on the

[1] For a comprehensive study concerned with cross-listings in stock markets see Karolyi (2006).

price adjustment (see De Jong 2002, Baillie et al. 2002). The more commonly applied measure, the Hasbrouck (1995) information share, is given by the contribution of a market's price innovation to the variance of the efficient underlying price innovations. Grammig et al. (2005), Hupperets and Menkveld (2002), and Phylaktis and Korczak (2007), use the Hasbrouck (1995) methodology to estimate the home and foreign market share in price discovery for US listed stocks from various countries. They also conclude that trading on the home market stock exchanges contributes most to price discovery, while trading on the NYSE primarily takes place to offset arbitrage opportunities.

The main drawback of the Hasbrouck (1995) approach is that it does not yield a uniquely identified information share but merely upper and lower bounds. Hasbrouck's (1995) method requires equidistant sampled data to measure a market's contribution to price discovery. Depending on the sampling frequency the information share bounds can diverge considerably and conclusions concerning the leading market are rather vague (see Hupperets and Menkveld 2002, Phylaktis and Korczak 2007). However, transaction market events occur irregularly in time, where the time between consecutive transactions convey information (see Frijns and Schotman 2005, Dufour and Engle 2000). Hence, arbitrary sampling schemes to obtain regular spaced data induce an undesirable loss of information.

Our method overcomes these difficulties by modeling the instantaneous arrival rates (intensities) of the price processes in a bivariate intensity model. The intensity roughly gives the probability of a transaction event within the next instant. We use Russell's (1999) ACI model that allows for a flexible interaction between the two markets' conditional intensities. The autoregressive dynamics of the intensity function are driven by an innovation term that originates in one market and immediately affects the conditional intensities on both markets. The innovation is constructed as the difference between the realized and the expected arrival rates. According to economic theory we expect positive cross effects of one market's innovation on the other market's conditional intensity, since the arbitrage relation between prices of stocks listed on two trading venues force an immediate incorporation of information originating in one market in the second market's price. Hence, if the realized arrival rate of the process associated with the first market is higher than expected, it increases the conditional intensity function of the second market. This can be interpreted as information originating from the first market that triggers further events in the second market. The larger these adjustments

in the second market, the more the first market contributes to price discovery. We propose to relate these innovations' cross effects and use them as a new information share. This measure does not suffer from an identification problem inherent in the Hasbrouck (1995) approach and is therefore unique. Further, we examine the long run dynamics of an intensity shock by an impulse response analysis.

We empirically analyze the price discovery process of Canadian stocks, which are traded on the TSX and cross-listed on the NYSE. Our results show a clear leadership of the TSX in the price discovery process. With an average information share of 29%, the contribution of the NYSE is slightly less pronounced than indicated by Eun and Sabherwal (2003).

The remainder of the paper is organized as follows. Section 4.2 outlines the methodological details of the ACI and introduces the new information share. Section 4.3 describes the data preparation and addresses the question of deseasonalization. In Section 4.4 we show and discuss the results and implications of our empirical application. Section 4.5 concludes.

4.2 Methodology

4.2.1 The Autoregressive Conditional Intensity Model

Denote by $\{t_i\}_{i=1}^n$ a stochastic sequence of event times in calendar time t which represent a point process. Assume that the arrival times are strictly distinct, $0 < t_1 < t_2 ... < t_n$, and $N(t)$ counts the number of events through t. The internal filtration of $N(t)$ denoted by $\Im_t$ consists of the complete information path of $N(t)$. The $\Im_t$-intensity process that characterizes the evolution of $N(t)$ is then

$$\lambda(t; \Im_t) = \lim_{\Delta \to 0} \frac{1}{\Delta} \mathbb{P}\left[N(t+\Delta) - N(t) > 0 | \Im_t\right] \quad . \tag{4.1}$$

Equation (4.1) gives an instantaneous probability that an event occurs in the next instant conditional on the information set available at t.

In a multivariate point process setting, $N(t)$ describes a pooled S-dimensional process that orders and pools the arrival rates of S individual point processes that occur with arrival times $\{t_i^s\}_{i=1}^{n^s}$ for $s = 1, ..., S$. The associated counting functions for the s-type events are indexed by $N^s(t)$. Since the pooled point process is assumed to be strictly increasing, the individual processes will be strictly orderly, too. Figure 4.2.1 gives an illustration of a pooled

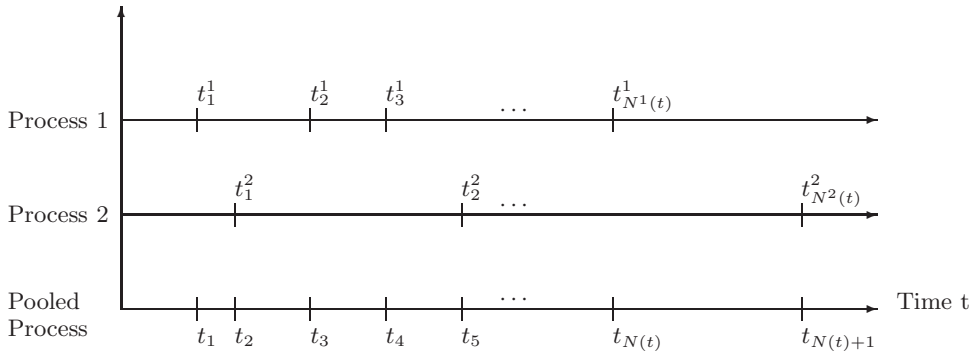

Figure 4.2.1: Pooled point process illustration. The figure gives an illustration of a simple point process $N(t)$ that consists of two individual counting processes $N^1(t)$ and $N^2(t)$. $\{t_i^1\}_{i=1}^{n^1}$ denotes the arrival times of events in process 1 and $\{t_i^2\}_{i=1}^{n^2}$ corresponds to event times in process 2. A time sequence $\{t_i\}_{i=1}^{n}$ containing both event time series is obtained by pooling and ordering the individual event times. As a consequence, an event occurring in the first process does not depend only on its own history but is allowed to depend on the history of process 2, as well, and vice versa.

point process $N(t)$ consisting of two individual processes $N^1(t)$ and $N^2(t)$.

The point process model adopted here is Russell's (1999) ACI model that parameterizes the s-type conditional intensity function as,

$$\lambda^s(t;\Im_t) = \lambda_o^s(t)\psi^s(t)\phi^s(t) \quad . \tag{4.2}$$

The process in (4.2) depends on a baseline intensity function $\lambda_o^s(t)$ and $\psi^s(t)$ that captures the dynamic structure of the conditional intensity. In many financial high frequency applications a daily seasonality function, $\phi^s(t)$, is of particular importance.

For the parametrization of the baseline intensity, we follow Russell (1999) and use a Weibull type hazard function that is based on the backward recurrence time, $x^s(t) = t - t_{N^s(t)}$, of process s,

$$\lambda_o^s(t) = \exp(w^s)\prod_{r=1}^{S} x^r(t)^{\gamma_r^s} \quad . \tag{4.3}$$

Hence, in the absence of events in the pooled process, γ_r^s gives the impact of the backward recurrence time of the r^{th} process on the conditional intensity in (4.2).

Collecting all ψ^s in a $S \times 1$ vector, $\boldsymbol{\psi}_i = (\psi^1, ..., \psi^S)'$, enables us to model the dynamic component in (4.2) as a vector autoregressive moving average (VARMA(1,1)) process,

$$\tilde{\boldsymbol{\psi}}_i = \mathbf{a}^s \varepsilon_{i-1}^s + \mathbf{B}\tilde{\boldsymbol{\psi}}_{i-1} \quad , \tag{4.4}$$

with $\psi^s(t) = \exp\left(\tilde{\psi}^s_{N(t)}\right)$ in order to ensure positivity of $\lambda^s(t;\Im_t)$. The right hand side of (4.4) contains a $S \times 1$ coefficient vector $\mathbf{a}^s$ multiplied by a scalar innovation, ε^s_i, that captures the arrival of new information originating from the s^{th} process. $\mathbf{B}$ is a $S \times S$ autoregressive coefficient matrix.

According to Russell (1999) the specification of the innovation in (4.4) is based on the integrated intensity which is computed by piecewise integration of $\lambda^s(t;\Im_t)$,

$$\Lambda^s(t^s_{i-1}, t^s_i) = \int_{t^s_{i-1}}^{t^s_i} \lambda^s(u;\Im_u)du = \sum_j \int_{\tilde{t}_j}^{\tilde{t}_{j+1}} \lambda^s(u;\Im_u)du \quad , \tag{4.5}$$

for j denoting all points with $t^s_{i-1} < \tilde{t}_j < \tilde{t}_{j+1} \leq t^s_i$. Using the random time change theorem any non-Poisson process can be transformed into a standard Poisson process which implies an iid standard exponential distributed integrated intensity, i.e. $\Lambda^s(t^s_{i-1}, t^s_i) \sim$ iid Exp(1) (see Hautsch 2003, Brémaud 1981, Bowsher 2007). Following Bauwens and Hautsch (2006), we then define the innovation in (4.4) as logarithm of an iid exponential variate centered by its unconditional expectation,[2]

$$\varepsilon^s_i = -0.5772 - \ln \Lambda^s(t^s_{i-1}, t^s_i) \quad . \tag{4.6}$$

Hence, an innovation in the ACI model has the interpretation of the deviation between the realized number of events minus the expected number of events within the interval $(t^s_{i-1}, t^s_i]$. This implies that positive values of ε^s_i indicate an underprediction of arrival rates and negative values an overprediction. Then the $S \times 1$ vector $\mathbf{a}^s$ in (4.4) gives the degree of adjustment to the new information originating from type s events that have an impact on the S conditional intensities.

The log-likelihood function of the S-dimensional ACI process can be expressed as

$$\ln L(\theta) = \sum_{s=1}^{S} \sum_{i=1}^{n} \{-\Lambda^s(t_{i-1}, t_i) + y^s_i \ln \lambda^s(t_i;\Im_{t_i})\} \quad , \tag{4.7}$$

where y^s_i is an indicator variable that takes the value one if the i^{th} point of the pooled process

[2] As indicated by Hautsch (2003), the logarithm of an iid exponential variate yields a minimum Gumbel variate, $\ln \Lambda^s(t^s_{i-1}, t^s_i)$, with mean -0.5772 and variance $\pi^2/6$.

is of type s and zero otherwise. Moreover, the first term on the right hand side of (4.7) corresponds to the s-type intensity integrated over $(t_{i-1}, t_i]$ and the second to the probability of the arrival times in the pooled process. The log-likelihood in (4.7) can be maximized by standard nonlinear optimization algorithms.

If the model is specified correctly, the resulting s-type residuals in (4.6), $\tilde{\varepsilon}_i^s = \Lambda^s(t_{i-1}^s, t_i^s)$, should be iid unit exponentially distributed. Hence, the dynamic and distributional properties of the estimated residuals can be evaluated by a Ljung-Box test and an overdispersion test suggested by Engle and Russell (1998). Their test statistic against excess dispersion, $\sqrt{\frac{n^s}{8}}(\sigma^2_{\tilde{\varepsilon}^s} - 1)$, follows asymptotically a normal distribution, where $\sigma^2_{\tilde{\varepsilon}^s}$ denotes the variance of the s-type residual series and n^s denotes the number of observations for process s.

4.2.2 Impulse Response Functions and Information Shares

In the context of international price discovery we are particularly interested in the cross effects of the conditional intensities associated with prices in the home and foreign market. In the following we therefore focus on a bivariate ACI model. Arbitrage relations between the two markets make sure that informative price changes originating in one market will subsequently be incorporated into the prices of the other. Consequently, if we observe unexpected trade events in one market this will trigger immediately events in the other market.

Henceforth, we denote an innovation originating in the home market (TSX) with superscript 1 and the corresponding coefficient vector becomes $\mathbf{a}^1$. The first element of $\mathbf{a}^1$, a_1^1, is then the impact of a TSX innovation, ε^1, on the TSX conditional intensity, λ^1. The second element, a_2^1, gives the cross effect of ε^1 on the NYSE conditional intensity λ^2. Thus, the subscript of a indicates the intensity of the market that is affected by the shock. Analogously, we denote NYSE associated shocks with a superscript 2. Following earlier studies (see Russell 1999, Bauwens and Hautsch 2006, Hall and Hautsch 2006), we restrict the autoregressive coefficient matrix $\mathbf{B}$ to be diagonal. Then, the diagonal elements of $\mathbf{B}$ determine the long run impact of a shock. As noted by Russell (1999), stationarity of the process is ensured if the eigenvalues of $\mathbf{B}$ lie inside the unit circle.

In order to examine the long run effect of an innovation shock on the dynamic component $\tilde{\psi}$ in (4.4) we compute impulse response functions and analyze the cumulated effects of a shock. The long run impact of a shock in period i in either market is computed by iterating

(4.4) h periods forward,

$$\tilde{\boldsymbol{\psi}}_{i+h} = \mathbf{a}^s \varepsilon^s_{i+h-1} + \mathbf{B}\mathbf{a}^s \varepsilon^s_{i+h-2} + ... + \mathbf{B}^{h-2}\mathbf{a}^s \varepsilon^s_{i+1} + \mathbf{B}^{h-1}\mathbf{a}^s \varepsilon^s_i + \mathbf{B}^h \tilde{\boldsymbol{\psi}}_i \quad .$$

To isolate the standard deviation shock in period i, $\sigma_\varepsilon = \sqrt{\pi^2/6}$, all subsequent shocks are set to their unconditional mean, $\mathbb{E}\left[\varepsilon^s_i\right] = 0$. Further, the unconditional mean $\mathbb{E}\left[\tilde{\boldsymbol{\psi}}_i\right] = 0$ is used as a starting value for $\tilde{\boldsymbol{\psi}}$. The impulse response functions are then given by,

$$IR^1(h) = \mathbf{B}^{h-1}\mathbf{a}^1\sigma_\varepsilon \qquad \text{and} \qquad IR^2(h) = \mathbf{B}^{h-1}\mathbf{a}^2\sigma_\varepsilon \quad , \tag{4.8}$$

where IR^1 denotes the bivariate impulse response function associated with a standard deviation shock on TSX and analogously IR^2 gives the bivariate impulse response function associated with a standard deviation shock on the NYSE. Summing up the effects in each period delivers the cumulative impulse response functions,

$$CIR^1(h) = \sum_{j=1}^{h} \mathbf{B}^{j-1}\mathbf{a}^1\sigma_\varepsilon \qquad \text{and} \qquad CIR^2(h) = \sum_{j=1}^{h} \mathbf{B}^{j-1}\mathbf{a}^2\sigma_\varepsilon \quad . \tag{4.9}$$

If the process is stationary, the effect of a shock in any of the markets dies out in the long run. Thus, the cumulative impulse response functions in (4.9) for $h \to \infty$ converge to a finite vector given below:

$$\lim_{h\to\infty} CIR^1 = [1 - \mathbf{B}]^{-1}\mathbf{a}^1\sigma_\varepsilon \qquad \text{and} \qquad \lim_{h\to\infty} CIR^2 = [1 - \mathbf{B}]^{-1}\mathbf{a}^2\sigma_\varepsilon \quad . \tag{4.10}$$

We focus on the cumulated cross effects of an unexpected shock in either market (CIR^1_2 and CIR^2_1) to determine the contributions to price discovery. If an unexpected positive deviation in the realized and expected arrival rates in the first market has a large impact on the second market's conditional intensity (implies large CIR^1_2), then this can be associated with information originating in the first market that has not yet been incorporated into the second market prices. A cumulated cross effect of zero ($CIR^2_1 = 0$) on the other hand indicates that the first market's conditional intensity does not react to informational events in the second market. This can be interpreted as evidence in favor of the first market leading the price discovery process.

Drawing on this, we propose to use the relative size of these cumulated cross effects as a measure for a market's contribution to price discovery. Due to characteristics of the trading process on the TSX and NYSE both markets might generally react differently to intensity shocks and we suggest to *standardize* the cumulative cross effects by the cumulative impact of a shock in the own market. Consequently, $\frac{CIR_1^2}{CIR_1^1}$ denotes the cross effect of a NYSE intensity shock on the TSX conditional intensity, standardized by the impact of a TSX shock on TSX's intensity. $\frac{CIR_2^1}{CIR_2^2}$ gives the analogue ratio for the NYSE.

Considering (4.10) it is obvious that terms cancel out and the ratio simplifies to $\frac{a_1^2}{a_1^1}$ and $\frac{a_2^1}{a_2^2}$. In a last step we confine our information shares to lie between zero and one by taking the standardized cross effects of each market relative to the sum of standardized cross effects:

$$IIS^1 = \frac{\frac{a_2^1}{a_2^2}}{\frac{a_1^2}{a_1^1} + \frac{a_2^1}{a_2^2}} \qquad \text{and} \qquad IIS^2 = \frac{\frac{a_1^2}{a_1^1}}{\frac{a_1^2}{a_1^1} + \frac{a_2^1}{a_2^2}} \quad . \tag{4.11}$$

Equation (4.11) then gives our unique intensity based information shares where IIS^1 denotes the TSX's and IIS^2 the NYSE's contribution to price discovery. Standard errors of the IIS in (4.11) can be computed via the delta method.

4.3 The Data

We use transaction data for 83 Canadian stocks that are traded on the TSX and cross-listed on the NYSE. The NYSE data are extracted from the Trade and Quote (TAQ) DVDs supplied by the NYSE. Toronto quote and trade data were taken from the Equity Trades and Quotes data set provided by the TSX. The sample period covers 62 trading days from 1^{st} of January 2004 to 31^{st} of March 2004. Continuous trading on both exchanges takes place from 9:30am to 4:30pm. We remove trades beyond these trading hours and overnight spells. We exclude 16 stocks due to erroneous data or extremely infrequent trading on one or both exchanges.[3]

There are 67 stocks remaining for estimation. Table 4.3.1 contains the stock tickers as well as the full company names.

[3] We omit BGM and ITN due to erroneous data. BEI, BR, CJR, CWG, EXEA, HBG, ITN, LAF, MDZ, MWI, OPY, RBA, RYG, TRA and VTS are excluded, since after thinning the data the resulting number of trading days on one or both exchanges is less then 62.

Ticker	Company Name	Industry
ABX	Barrick Gold	Gold Mining
ABY	Abitibi Consolidated Inc.	Paper
AEM	Agnico Eagle Mines Ltd.	Gold Mining
AGU	Agrium Inc.	Chemicals (Specialty)
AL	Alcan Inc.	Metals and Mining
BCE	BCE Inc.	Foreign Telecom.
BCM	Canadian Imp. Bank of Commerce	Bank
BMO	Bank of Montreal	Bank
BNN	Brascan Corp.	Real Estate Holding
BNS	Bank of Nova Scotia	Bank
BPO	Brookfield Properties Corporation	Real Estate Holding
BVF	Biovail Corp.	Pharmaceuticals
CCJ	Cameco Corp.	Nonferrous Metals
CGT	CAE Inc.	Aerospace
CLS	Celestica Inc.	Electronics
CNI	Canadian National Railway	Transport
CNQ	Canadian Natural Ressources	Petroleum (Producing)
COT	Cott Corp.	Soft Drinks
CP	Canadian Pacific Railway	Transport
DTC	Domtar Corp.	Paper
ECA	EnCana Corp.	Energy
ENB	Enbridge Inc.	Gas Distribution
ERF	Enerplus Resource Fund	Exploration and Production
FDG	Fording Canadian Coal Trust	Mining (Other Mines)
FFH	Fairfax Financial Holdings Ltd.	Property and Casualty Insurance
FHR	Fairmont Hotels Resorts Inc.	Hotels
FS	Four Seasons Hotels Inc.	Hotels
GG	Goldcorp Inc.	Gold Mining
GIB	CGI Group Inc.	Computer Services
GIL	Gildan Activewear Inc.	Clothing and Accessories
GLG	Glamis Golds Ltd.	Gold Mining
IDR	Intrawest Corp..	Hotels
IPS	IPSCO Inc.	Metals and Mining
IQW	Quebecor World	Publishing
ITP	Intertape Polymer Group Inc.	Containers and Packaging
KFS	Kingsway Financial Services Inc.	Insurance
KGC	Kinross Gold Corp.	Gold Mining
MDG	Meridian Gold Inc.	Gold Mining
MFC	Manulife Financial Corp.	Insurance
MGA	Magna International Inc.	Auto Parts
MHM	Masonite International Corp.	Building Products
MIM	MI Developments Inc.	Gambling
N	Inco Ltd.	Metals and Mining
NCX	Nova Chemicals Corp.	Commodity Chemicals
NRD	Noranda Inc.	Metals and Mining
NT	Nortel Networks	Foreign Telecom.
NXY	Nexen Inc.	Energy
PCZ	Petro-Canadian Com.	Integrated Oil and Gas
PDG	Placer Dome	Precious Metals
PDS	Precision Drilling Corp	Oil Equipment and Services
PGH	Pengrowth Energy	Exploration and Production
PKZ	PetroKazakhstan Inc.	Petroleum
POT	Potash Corp.	Chemical
PWI	Primewest Energy Trust	Energy
RCN	Radiant Communications	Telecommunications
RG	Rogers Publishing Limited	Publishing
RY	Royal Bank of Canada	Bank
SLF	Sun Life Financial Serv.	Insurance
SU	Suncor Energy	Petroleum
TAC	TransAlta Corp.	Conventional Electricity
TD	Toronto-Dominion	Bank
TEU	CP Ships Ltd.	Maritime
TLM	Talisman Energy	Energy
TOC	Thomson Corp.	Information Services
TRP	TransCanada Corp.	Energy
TU	Telus Corp.	Telecommunications
ZL	Zarlink Semiconductor Inc.	Semiconductors

Table 4.3.1: Sample stocks. The table shows the ticker symbols of the 67 Canadian sample stocks together with the full company name and their industry.

	TSX						NYSE					
TICKER	Vol^1	$\#T^1$	$\bar{\tau}^1$	σ_{τ^1}	AC^1	p^1_{LB}	Vol^2	$\#T^2$	$\bar{\tau}^2$	σ_{τ^2}	AC^2	p^2_{LB}
ABX	35.46	368.2	63.4	94.3	0.16	0.00	34.71	835.0	27.7	39.7	0.11	0.00
ABY	15.98	201.6	115.6	188.9	0.13	0.00	1.45	122.4	188.0	271.7	0.13	0.00
AEM	6.05	137.7	168.5	257.3	0.13	0.00	6.65	377.1	61.7	96.8	0.13	0.00
AGU	7.76	94.2	244.8	366.2	0.14	0.00	4.74	223.7	101.7	157.2	0.13	0.00
AL	49.47	456.5	51.1	72.4	0.17	0.00	63.85	905.7	25.6	34.6	0.10	0.00
BCE	56.76	320.2	72.9	101.2	0.14	0.00	5.02	218.5	105.6	159.7	0.12	0.00
BCM	68.87	286.1	81.5	118.0	0.12	0.00	1.72	80.1	284.3	412.2	0.13	0.00
BMO	52.16	289.0	80.7	118.4	0.13	0.00	1.39	82.9	273.6	402.0	0.17	0.00
BNN	13.07	105.5	219.2	315.2	0.08	0.00	1.00	54.0	405.8	622.1	0.15	0.00
BNS	49.98	214.8	108.7	160.3	0.14	0.00	0.41	21.5	977.2	1333.9	0.09	0.00
BPO	2.63	34.0	655.8	953.6	0.13	0.00	2.13	82.6	275.0	420.1	0.13	0.00
BVF	11.21	289.7	80.4	126.1	0.22	0.00	22.33	565.2	40.8	65.1	0.15	0.00
CCJ	18.34	159.9	145.3	253.4	0.17	0.00	4.58	174.5	130.7	216.1	0.18	0.00
CGT	6.57	156.5	148.5	244.0	0.19	0.00	0.08	12.7	1409.7	2254.4	0.10	0.01
CLS	19.03	393.1	59.3	92.9	0.18	0.00	22.39	578.2	40.1	70.7	0.15	0.00
CNI	23.26	218.5	106.5	162.6	0.14	0.00	10.58	346.7	66.1	97.5	0.12	0.00
CNQ	26.07	208.8	111.4	173.5	0.16	0.00	3.58	157.3	145.5	216.9	0.12	0.00
COT	3.53	50.1	458.7	611.6	0.08	0.00	2.61	143.7	159.9	237.7	0.11	0.00
CP	14.58	171.7	135.5	196.7	0.14	0.00	1.79	131.5	174.0	247.4	0.11	0.00
DTC	10.95	127.8	182.1	278.3	0.15	0.00	0.52	55.7	395.7	574.3	0.10	0.00
ECA	55.61	383.5	60.8	89.4	0.16	0.00	11.93	374.8	61.5	84.7	0.11	0.00
ENB	13.09	98.0	237.0	371.4	0.09	0.00	0.34	22.9	848.4	1329.3	0.15	0.00
ERF	5.14	77.1	298.5	445.7	0.16	0.00	5.82	199.4	115.3	175.8	0.17	0.00
FDG	5.64	77.2	296.5	499.4	0.20	0.00	4.33	156.1	147.4	264.8	0.20	0.00
FFH	4.12	51.8	433.7	713.5	0.18	0.00	4.98	94.8	238.1	450.2	0.15	0.00
FHR	2.62	59.0	390.6	578.8	0.15	0.00	6.00	185.7	123.0	206.2	0.14	0.00
FS	0.91	24.2	891.6	1411.9	0.16	0.00	7.66	258.8	88.9	157.4	0.17	0.00
GG	9.51	255.6	91.1	134.4	0.16	0.00	15.55	584.2	39.7	55.8	0.13	0.00
GIB	4.06	97.8	236.6	373.8	0.12	0.00	0.14	25.2	867.6	1334.6	0.17	0.00
GIL	1.49	26.7	742.2	1290.9	0.09	0.00	0.35	23.0	834.5	1635.4	0.14	0.00
GLG	7.26	124.6	186.6	281.2	0.18	0.00	8.09	413.7	56.1	86.1	0.12	0.00
IDR	1.47	36.2	620.4	889.5	0.13	0.00	1.61	87.0	260.8	408.2	0.11	0.00
IPS	3.47	40.4	543.8	872.6	0.10	0.00	0.12	15.8	1118.5	1915.5	0.10	0.00
IQW	7.13	94.4	245.0	353.9	0.14	0.00	0.32	42.0	520.3	769.6	0.10	0.00
ITP	1.36	30.2	688.2	1119.9	0.08	0.00	0.43	51.5	422.8	713.5	0.10	0.00
KFS	3.45	85.4	268.6	428.4	0.21	0.00	0.42	42.9	511.6	831.6	0.15	0.00
KGC	14.91	284.0	82.2	125.5	0.18	0.00	6.91	352.4	66.1	97.3	0.14	0.00
MDG	3.55	106.0	218.4	351.3	0.16	0.00	5.69	382.3	60.7	91.2	0.11	0.00
MFC	50.17	280.8	83.0	120.5	0.14	0.00	13.34	267.0	86.4	139.2	0.11	0.00
MGA	11.31	107.1	216.5	304.0	0.17	0.00	15.47	335.3	68.7	100.6	0.10	0.00
MHM	6.25	40.4	554.9	944.3	0.13	0.00	0.31	21.8	979.4	1501.2	0.11	0.00
MIM	1.93	19.8	1098.2	1873.6	0.14	0.00	3.51	73.6	301.6	595.6	0.17	0.00
N	45.64	426.4	54.7	80.1	0.17	0.00	55.09	881.3	26.2	37.5	0.07	0.00
NCX	6.51	77.0	300.4	412.3	0.16	0.00	1.94	120.3	188.0	273.7	0.11	0.00
NRD	21.31	290.1	80.3	123.2	0.15	0.00	1.14	103.7	218.9	307.2	0.10	0.00
NT	147.31	959.5	24.3	42.4	0.22	0.00	134.59	555.9	41.9	69.5	0.17	0.00
NXY	26.59	208.5	111.0	185.5	0.13	0.00	2.01	99.9	222.9	330.8	0.16	0.00
PCZ	50.51	285.2	81.7	132.9	0.19	0.00	1.94	99.1	231.4	335.3	0.14	0.00
PDG	28.75	399.7	58.3	86.4	0.16	0.00	23.94	687.7	33.6	49.0	0.11	0.00
PDS	11.59	135.7	170.5	250.9	0.15	0.00	7.36	259.0	87.0	130.7	0.11	0.00
PGH	4.74	116.0	198.3	315.8	0.20	0.00	5.54	248.8	92.7	147.2	0.17	0.00
PKZ	5.33	115.1	199.3	362.2	0.24	0.00	8.68	307.7	74.5	126.8	0.14	0.00
POT	7.19	68.7	335.7	451.8	0.17	0.00	8.54	242.9	94.3	150.3	0.15	0.00
PWI	3.89	94.3	243.6	400.4	0.19	0.00	3.58	121.0	189.6	311.9	0.22	0.00
RCN	1.63	26.8	777.2	1329.9	0.12	0.00	0.34	31.9	645.2	1215.7	0.15	0.00
RG	14.61	111.2	208.1	353.4	0.13	0.00	0.42	51.9	431.3	628.0	0.19	0.00
RY	90.44	303.3	77.0	108.4	0.17	0.00	2.19	103.2	222.5	304.9	0.18	0.00
SLF	29.18	239.3	97.4	138.2	0.13	0.00	1.96	151.3	151.3	215.2	0.13	0.00
SU	35.78	336.8	69.3	103.2	0.14	0.00	11.08	386.0	59.4	80.3	0.09	0.00
TAC	5.16	112.7	206.1	284.4	0.15	0.00	0.08	9.1	1979.3	2759.9	0.07	0.12
TD	61.19	294.3	79.3	113.9	0.15	0.00	1.17	83.0	275.6	395.9	0.17	0.00
TEU	9.01	109.7	210.7	335.1	0.15	0.00	1.69	98.4	233.1	366.5	0.15	0.00
TLM	22.82	162.0	143.6	204.4	0.16	0.00	5.39	219.4	104.4	150.6	0.13	0.00
TOC	14.39	143.9	161.7	258.4	0.17	0.00	1.03	74.9	306.0	453.4	0.14	0.00
TRP	25.90	210.0	111.2	162.1	0.14	0.00	2.43	179.3	129.3	174.9	0.11	0.00
TU	8.99	114.7	202.1	321.1	0.15	0.00	0.22	25.3	848.1	1603.1	0.12	0.00
ZL	3.91	173.8	133.3	236.0	0.18	0.00	0.78	71.4	313.8	561.1	0.22	0.00

Table 4.3.2: Descriptive statistics. For TSX the superscript s equals 1 and for NYSE $s = 2$. The columns labeled Vol^s and $\#T^s$ give the average daily trading volume in million CAD$ and the average number of transactions per day, respectively. Columns $\bar{\tau}^s$ and σ_{τ^s} contain the average and the standard deviation of transaction durations in seconds for the whole sample period. Columns AC^s give the first order autocorrelation of the transaction duration, a corresponding Ljung-Box p-value is reported in the columns labeled p^s_{LB}. All statistics are calculated over the whole sample period (from January 1st to 31st of March 2004). For full company names see Table 4.3.1.

Prior to estimation the trade data were thinned based on price marks, in order to extract events associated with new information (see Engle and Russell 1998).[4] In detail, we retain those trade events that are associated with a volume equal or larger than the volume available at the best quote. This procedure filters those trades that induce a quote revision. Since the S dimensional intensity model introduced in the previous section assigns zero probability to the simultaneous occurrence of two events, trades with the same time stamp within one market are treated as one trade. Further, trades with the same time stamp in both markets are deleted.

Table 4.3.2 presents detailed stock specific descriptive statistics. It can be seen that the sample includes a range of stocks varying with respect to size and transaction frequency. The average trading volume in the seven hours of parallel trading ranges for TSX from CAD$ 0.9 million to CAD$ 147.3 million and for NYSE CAD$ 0.1 million to CAD$ 134.6 million. The average daily number of transactions ranges from 20 to 960 at TSX and from 9 to 906 at NYSE. Trades arrive on TSX over all stocks on average every 240 seconds and on NYSE on every 303 seconds. Transactions on the TSX occur on average every 24 seconds for the most frequent stock and every 18 minutes for the most infrequent stock. Accordingly, at NYSE trades occur in a range between 26 seconds and 33 minutes.

Several authors (see e.g. Engle and Russell 1997) point out that price durations exhibit an intraday pattern in the rate of arrival. We therefore deseasonalize our data prior to estimation. Assuming the separability of time function and stochastic function in (4.2), the elimination of the time of day effect proceeds in the following two steps. First, the typical time-of-day pattern (ϕ_i) is estimated by regressing the transaction durations ($\tau_i = t_i - t_{i-1}$) of the pooled process on polynomial and trigonometric time functions (see Eubank and Speckman 1990 and Appendix C.1). Second, dividing the durations by their estimated typical shape gives seasonally adjusted durations, i.e. $\tilde{\tau}_i = \frac{\tau_i}{\phi_i}$. Then, a seasonal adjusted transaction arrival series of the pooled process is achieved by setting the first arrival time of the day to zero and cumulating durations ($\tilde{\tau}_i$) for each day. Figure 4.3.1 shows the transaction duration for two randomly chosen stocks in our sample before removing the intraday pattern. As clearly visible from the figure, transaction durations exhibit the typical $\cap$-shape. The intraday pattern is captured by the estimated seasonal figure.

[4]Price marks are characteristics in the point process that are observed simultaneously with the price arrival times e.g. the size of the price change, bid-ask spread or volume traded.

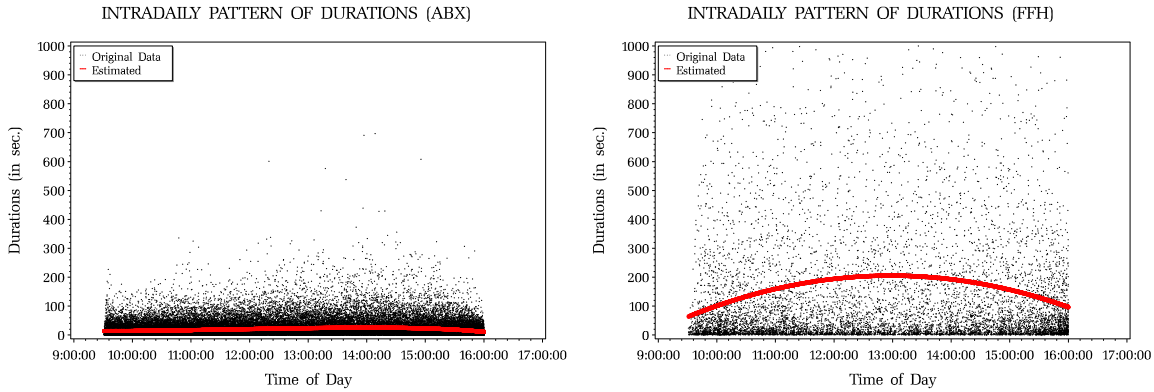

Figure 4.3.1: Intraday pattern of durations. The figure shows the transaction durations (Original Data: black dots) of the pooled process for two randomly chosen stocks (see Table 4.3.1). As visible from the figure, transaction durations exhibit a ∩-shape. The typical intraday pattern is captured by the estimated seasonal figure (Estimated: line).

4.4 Estimation, Information Shares and Results

4.4.1 Estimation Results and Diagnostics

As outlined in Section 4.2.1 estimation of our model parameters is done via maximizing the model's likelihood function in (4.7). Since we cannot allow previous day shocks to affect the next day's intensity, the likelihood function has to be re-initialized each day and becomes the sum of independent day-likelihoods. Therefore, the recursive process for the latent $\tilde{\psi}$ in (4.4) stops at the end of the trading day and is initialized the next morning. Table 4.4.1 contains sample stock descriptives for the estimated ACI coefficients. Stock specific results can be found in Table C.2.1 in the Appendix C.2.

First, we draw our attention on the estimates of the baseline intensity in Equation (4.3). Small average standard errors in Table 4.4.1 and standard errors of the estimates in Table C.2.1 in Appendix C.2 show that generally the constant coefficients of the baseline intensity (ω^s) are statistically significant on a 5% level of confidence. The ratio of constant baseline coefficients, $\frac{\exp(\hat{\omega}^1)}{\exp(\hat{\omega}^1)+\exp(\hat{\omega}^2)}$, estimates very well the ratio of daily average number of transactions in Table 4.3.2, $\frac{\#T^1}{\#T^1+\#T^2}$. Hence, the constants reflect the relative number of transactions in the two markets over the whole sample period. The estimated coefficients of the backward recurrence function $\hat{\gamma}_1^s$ and $\hat{\gamma}_2^s$ are negative and mostly significant for all stocks. Hence, an event in one market induces an upwards jump in the intensity functions and in the absence of new events in the pooled process the probability of an event will be downward sloping.

Estimates	M	Std	$Q25$	$Q75$	Min	Max	$M(SE)$	$\#sig$
$\hat{\omega}^1$	-0.657	0.413	-0.995	-0.354	-1.982	-0.044	0.0295	54
$\hat{\omega}^2$	-0.777	0.509	-1.143	-0.381	-2.267	0.031	0.0260	54
$\hat{\gamma}_1^1$	-0.223	0.055	-0.256	-0.189	-0.383	-0.052	0.0067	55
$\hat{\gamma}_2^1$	-0.088	0.049	-0.124	-0.051	-0.223	-0.026	0.0076	55
$\hat{\gamma}_1^2$	-0.198	0.082	-0.247	-0.144	-0.406	0.005	0.0078	54
$\hat{\gamma}_2^2$	-0.234	0.053	-0.271	-0.204	-0.362	-0.083	0.0074	55
$\hat{a}_1^1$	0.049	0.012	0.041	0.055	0.029	0.099	0.0056	55
$\hat{a}_2^1$	0.024	0.011	0.018	0.029	0.002	0.069	0.0053	52
$\hat{a}_1^2$	0.011	0.007	0.006	0.014	0.000	0.032	0.0040	38
$\hat{a}_2^2$	0.046	0.022	0.031	0.054	0.018	0.111	0.0068	54
$\hat{b}^1$	0.984	0.015	0.977	0.993	0.927	1.000	0.0046	55
$\hat{b}^2$	0.973	0.030	0.968	0.991	0.858	1.000	0.0085	55

Table 4.4.1: Estimation summary results. The table contains descriptive statistics for the estimated parameters of the ACI model in (4.3) and (4.4). The table displays the mean (M), the standard deviation (Std), the first ($Q25$) and third quartile ($Q75$), and the minimum (Min) and maximum (Max) of the estimated parameters over all sample stocks. $M(SE)$ is the average standard error of the estimates and $\#sig$ gives the number of significant estimates on a 5% significance level over the sample stocks. The descriptive statistics are computed over 55 stocks that have positive $\mathbf{a}^s$ estimates and $\mathbf{B}$ estimates smaller than one.

As reported in Table C.2.1 in Appendix C.2, we see that the parameter estimates for a_s^s are positive for 57 stocks indicating positively autocorrelated intensities. An underprediction of arrival rates in the previous interval of type s has for the majority of the stocks significant and positive effects on the conditional intensity of the same type in the next interval. According to Table 4.4.1, the effect of a shock in the own market tends to be higher than the cross effect of a shock in the other market. The short run innovation impacts from TSX on NYSE are on average larger than the effects of NYSE innovations on TSX. For 52 stocks the innovation effects from the TSX are significant. The spill over effects from NYSE on TSX are significant for 38 stocks. As expected from the duration modeling literature (see e.g. Engle and Russell 1998), we find strong persistence of innovation shocks. This is reflected in relatively large (on average 0.98) and significant autoregressive coefficients (b^s) for all stocks.

The results for the Ljung-Box test in Table 4.4.2 for the first autocorrelation of the estimated residuals are mixed. For some stocks the null hypothesis of no autocorrelation cannot be rejected. Further, the table shows that the mean and the standard deviation of the estimated residuals are on average close to one on both markets. Considering the standard deviation of the residuals some excess dispersion is still present.

	TSX					NYSE				
TICKER	$\bar{\tilde{\varepsilon}}^1$	$\sigma_{\tilde{\varepsilon}^1}$	OD^1	AC^1	LB^1	$\bar{\tilde{\varepsilon}}^2$	$\sigma_{\tilde{\varepsilon}^2}$	OD^2	AC^2	LB^2
ABX	1.01	1.07	7.67	0.05	47.18	0.99	1.09	14.88	0.03	52.42
ABY	1.00	1.06	4.69	0.03	8.76	0.97	0.99	-0.36	0.04	11.21
AEM	1.00	1.04	2.58	0.03	9.71	1.00	1.06	6.10	0.04	35.20
AGU	1.02	1.07	3.57	0.02	1.38	0.99	1.01	1.07	0.04	23.31
AL	1.00	1.08	9.31	0.04	42.22	0.98	1.05	9.33	0.03	56.35
BCE	1.01	1.06	5.76	0.04	31.56	0.97	1.01	1.10	0.03	14.01
BCM	1.01	1.05	5.13	0.03	15.63	0.96	0.95	-2.53	0.00	0.04
BMO	1.00	1.05	4.70	0.02	8.49	0.95	0.95	-2.39	0.04	7.00
BNN	1.02	1.08	4.86	0.01	0.39	0.95	0.97	-1.05	0.04	5.82
BNS	1.00	1.05	4.13	0.03	11.77	0.88	0.80	-4.50	0.01	0.18
CCJ	1.00	1.09	6.32	0.06	38.92	0.96	0.99	-0.85	0.04	13.23
CGT	0.97	1.04	2.65	0.06	30.67	0.83	0.89	-1.94	0.01	0.11
CLS	1.00	1.08	8.90	0.05	71.05	0.98	1.07	10.01	0.04	54.17
CNI	1.02	1.10	8.44	0.03	11.28	0.97	1.01	1.35	0.02	12.05
CNQ	1.01	1.07	5.58	0.04	20.09	0.98	0.98	-1.13	0.01	1.53
COT	1.02	1.13	5.47	0.02	1.51	0.99	1.03	1.79	0.02	3.76
DTC	1.02	1.05	2.95	0.05	15.69	0.95	0.93	-2.70	0.01	0.27
ECA	1.00	1.08	8.53	0.05	60.31	0.98	0.98	-2.25	0.02	9.56
ERF	1.01	1.05	2.38	0.03	3.59	0.99	1.01	0.63	0.05	31.02
FDG	0.99	1.03	1.69	0.04	8.76	0.97	1.07	4.84	0.04	14.45
FFH	0.97	1.05	2.01	0.03	3.32	0.99	1.08	4.53	0.04	7.87
FHR	1.01	1.04	1.79	0.03	3.96	0.99	1.02	1.82	0.03	13.25
FS	0.96	0.96	-1.15	0.01	0.24	0.99	1.07	5.83	0.05	31.07
GG	1.01	1.04	3.53	0.05	37.23	0.99	1.01	1.79	0.02	14.13
GIB	0.99	1.11	6.26	0.03	4.85	0.95	1.01	0.37	0.01	0.19
GIL	1.00	1.28	8.72	0.03	1.10	1.02	1.35	10.36	-0.01	0.09
GLG	1.01	1.05	2.99	0.03	6.13	0.99	1.05	5.52	0.03	29.54
IPS	1.03	1.14	4.86	0.02	0.75	0.90	0.97	-0.60	-0.04	1.14
IQW	1.01	1.04	2.25	0.04	10.16	0.96	1.05	1.90	0.04	3.48
KFS	0.99	1.13	6.90	0.05	14.42	0.98	1.21	8.50	0.06	10.45
KGC	1.00	1.07	7.23	0.05	45.12	1.00	1.05	4.84	0.03	26.22
MDG	0.99	1.00	0.01	0.05	13.28	1.00	1.03	3.10	0.02	9.05
MFC	1.01	1.07	6.75	0.03	19.48	0.99	1.04	3.50	0.02	10.05
NCX	1.02	1.04	1.71	0.02	2.80	0.99	0.98	-1.05	0.02	2.42
NRD	1.00	1.07	6.39	0.05	37.59	0.97	0.98	-1.05	0.03	4.31
NT	0.98	1.13	24.02	0.06	198.05	0.95	1.11	15.60	0.06	104.28
NXY	0.99	1.14	11.21	0.05	25.80	0.92	0.98	-0.99	0.03	7.26
PCZ	0.98	1.07	7.18	0.05	35.56	0.96	0.94	-3.34	0.02	1.77
PDG	1.01	1.07	8.38	0.05	52.02	0.99	1.05	7.06	0.03	27.65
PDS	1.02	1.10	7.00	0.05	17.36	0.98	1.00	0.05	0.02	3.64
PGH	0.99	1.06	3.94	0.05	14.27	0.98	1.07	6.18	0.02	4.33
PKZ	0.97	1.03	2.07	0.07	34.86	0.96	1.05	5.01	0.01	3.60
POT	1.01	0.99	-0.31	0.04	7.09	0.98	1.02	1.40	0.03	16.21
RCN	1.02	1.39	12.79	0.04	3.00	1.04	1.32	11.10	-0.02	0.75
RG	1.00	1.11	6.52	0.03	7.59	0.98	1.03	1.31	0.04	3.95
RY	1.01	1.07	6.72	0.06	65.88	0.96	0.96	-2.16	0.05	13.21
SLF	1.01	1.04	3.82	0.04	24.81	0.98	0.97	-1.75	0.02	5.45
SU	1.03	1.09	9.68	0.05	45.26	0.97	0.98	-2.03	0.01	2.78
TAC	0.97	1.04	2.22	0.05	14.48	0.80	0.79	-3.02	0.06	2.09
TD	1.00	1.06	6.14	0.03	21.08	0.97	0.95	-2.54	0.03	4.45
TEU	1.00	1.12	7.11	0.04	8.45	0.99	1.05	2.89	0.03	3.75
TLM	1.01	1.06	4.15	0.02	5.62	0.99	1.00	0.30	0.03	9.93
TOC	1.00	1.04	2.98	0.06	32.35	0.98	1.00	-0.19	0.01	0.92
TRP	1.00	1.08	6.24	0.05	27.77	0.99	0.99	-1.10	0.03	12.88
ZL	0.99	1.07	4.79	0.05	25.42	0.95	0.95	-2.13	0.08	25.72

Table 4.4.2: Residual diagnostics for the ACI model. The table presents residual diagnostics for the estimated residuals corresponding to TSX ($s = 1$) and NYSE ($s = 2$). AC^s denotes the value of the first order autocorrelation and columns labeled with LB^s contain the corresponding Ljung-Box statistic. $\bar{\tilde{\varepsilon}}^s$ and $\sigma_{\tilde{\varepsilon}^s}$ contain the mean and the standard deviation of the estimated residuals and OD^s gives the test statistic of the overdispersion test of Engle and Russell (1998). This statistics has a limiting normal distribution under the null with a 5% critical value of 1.645. The statistics are computed for 55 stocks that have positive $\mathbf{a}^s$ estimates and $\mathbf{B}$ estimates smaller than one. For full company names see Table 4.3.1.

4.4.2 Information Shares

Figure 4.4.1 shows cumulated impulse response functions in (4.9) for an exemplary sample stock. The left panel depicts the impact of a standard deviation shock on the TSX and its impact on TSX and NYSE processes. The right panel illustrates the impact of a NYSE standard deviation shock on both processes. Own market's shocks have larger impacts on

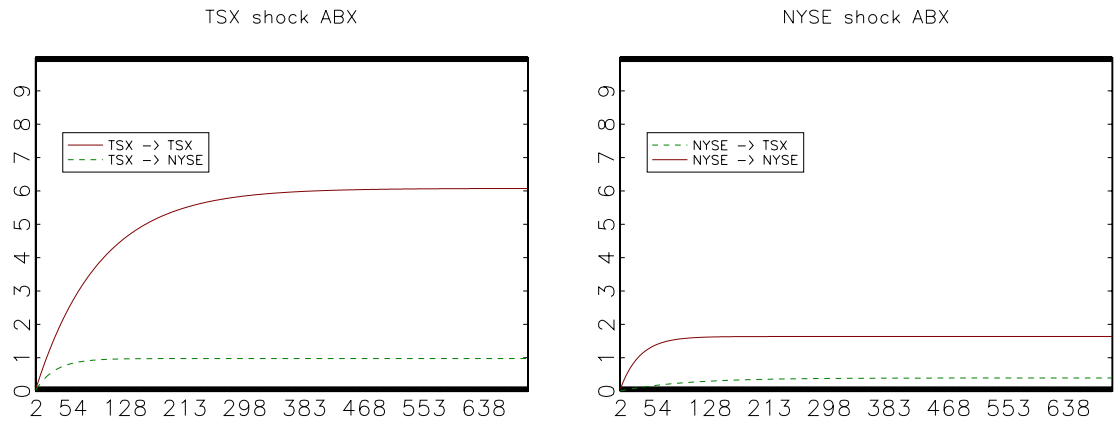

Figure 4.4.1: Cumulated impulse response function of a standard deviation innovation shock. The figure shows cumulated impulse response functions of (4.9) for the recursive process $\tilde{\boldsymbol{\psi}}$ in (4.4) for the sample stock ABX (see Table 4.3.1). The left panel depicts the impact of a standard deviation shock on the TSX and its impact on the TSX process (solid line) and on the NYSE (dashed line). Analogously, the right panel illustrates the impact of a standard deviation shock on the NYSE on the NYSE process (solid line) and on the TSX (dashed line).

$\tilde{\boldsymbol{\psi}}$ than shocks of the other market. Long run impacts of a TSX shock on $\tilde{\psi}^2$ are generally higher, than the effects of a NYSE shock on $\tilde{\psi}^1$.

Table 4.4.3 displays sample stock descriptives for the unique intensity based information share (IIS) according to (4.11). We also report Hasbrouck information shares (details are outlined in Appendix C.3). Stock specific results can be found in Table C.2.2 in the Appendix C.2.

As reported in Table 4.4.3 the average home market intensity based information share (IIS^1) amounts to 71%. This implies a clear leadership of the TSX in price discovery. It is even more pronounced than indicated by previous studies (see Grammig et al. 2005, Eun and Sabherwal 2003). For only 7 out of 55 stocks we observe a higher NYSE contribution. Considering the average Hasbrouck information share midpoint (HIS^1_{mid}) of 54% for the home market, we find on average higher intensity based information shares. However, the

	TSX				NYSE			
DESCRIPTIVE	IIS^1	HIS^1_{low}	HIS^1_{up}	HIS^1_{mid}	IIS^2	HIS^2_{low}	HIS^2_{up}	HIS^2_{mid}
M	70.9	30.2	77.3	53.7	29.1	22.7	69.8	46.3
$M(SE)$	11.3	2.6	2.5	2.3	11.3	2.5	2.6	2.3
Std	15.6	23.0	16.7	18.2	15.6	16.7	23.0	18.2
$Q25$	61.9	9.0	60.3	37.9	18.2	8.4	52.2	30.4
$Q75$	81.7	47.3	90.8	67.6	36.0	36.5	90.7	61.9
Min	30.9	0.1	21.7	10.9	0.7	2.9	11.6	7.7
Max	99.3	88.4	97.1	92.3	69.1	78.3	99.9	89.1

Table 4.4.3: Intensity based information shares – descriptives. The table presents descriptives computed over the information shares using an intensity based and the standard Hasbrouck approach in percent. The descriptives are the mean (M), the standard deviation (Std), the mean of the information share standard error ($M(SE)$), the 25% quantile ($Q25$), the 75% quantile ($Q75$), the minimum (Min) and maximum (Max) over the cross sectional information shares. The midpoint and the lower and upper bounds of Hasbrouck are denoted by HIS^s_{mid}, HIS^s_{low} and HIS^s_{up}, respectively. Columns labeled IIS^s give the unique intensity based information share. For TSX $s = 1$ and NYSE $s = 2$. The descriptives are computed over 55 stocks that have positive $\mathbf{a}^s$ estimates and $\mathbf{B}$ estimates smaller than one.

Hasbrouck information share bounds deviate considerably, with an average lower bound of 30% and an average upper bound of 77% for the home market share. Consequently, the lower and upper bounds differ by 47 percentage points on average and the midpoint can be considered a very imprecise proxy for the true information share, which emphasizes the need for a unique measure. We further find that the cross sectional variation of intensity based information shares corresponds to those of the Hasbrouck share midpoints. For 25 out of 55 stocks, the intensity shares lie outside the Hasbrouck share bounds.

4.5 Conclusion

Investors' decision to invest and companies' intention to list their stocks on a stock exchange depends on the ability of an exchange to provide a prospering trading environment. As a result of an increasing globalization and improved technology, small national exchanges fear to lose their attractiveness for investors and companies. In particular, within the context of international cross-listed stocks, it is of paramount concern for a national stock exchange to remain the dominant market with regard to price discovery.

We propose a new approach to measure the contribution of different trading venues to the price discovery process of internationally cross-listed stocks. We use a bivariate intensity approach as an alternative to the commonly applied vector error correction model in order to take into account the irregularity of the data. Based on the autoregressive conditional intensity model of Russell (1999), contributions to price discovery are determined by modeling

the interdependencies of the trading processes in both markets.

In contrast to the Hasbrouck (1995) approach, our new information share delivers unique results rather than upper and lower bounds. In our empirical application we examine Canadian stocks which are listed on the TSX as well as on the NYSE. We find that despite the concern of the TSX to lose its share in price discovery to the NYSE, trading on the TSX still plays the most important role. We show that the leadership of the TSX is even more pronounced than indicated by previous studies. The average TSX information share amounts to 71% , which confirms previous results by Phylaktis and Korczak (2007), Eun and Sabherwal (2003), and Grammig et al. (2005), who also analyze Canadian stocks. We also compare our results to the Hasbrouck (1995) information shares. On average we find a larger home market contribution than indicated by the Hasbrouck midpoints.

Generally, our intensity based unique information share offers an alternative approach to determine the leading market with respect to price discovery. Taking into account the irregular spacing of the transaction data, it presents an appealing alternative to the Hasbrouck (1995) information shares when analyzing price discovery using high frequency data.

Appendix C

C.1 Deseasonalization

In order to deseasonalize the transaction data, we follow Eubank and Speckman (1990) and regress the transaction durations of the pooled process on polynomial and trigonometric time functions. The regression equation reads for some integers $d \geq 0$ and $\delta \geq 0$ as follows,

$$\tau_i = \beta_0 + \sum_{j=1}^{d} \beta_j^p t_i^j + \sum_{j=1}^{\delta} [\beta_j^c \cos(jt_i) + \beta_j^s \sin(jt_i)] + \epsilon_i \quad , \tag{C.1}$$

where the transaction duration is $\tau_i = t_i - t_{i-1}$. The number of polynomial and trigonometric terms are selected by a generalized cross-validation measure defined as,

$$GCV = \frac{nRSS}{(n - 2\delta - d - 1)^2} \quad , \tag{C.2}$$

where RSS denotes the residual sum of squares and n the number of observations. In the selection we restrict d and δ to be smaller than five. To compute a typical time-of-day function we select the specification of (C.1) that minimizes the GCV in (C.2).

C.2 Additional Tables

Ticker	$\hat{\omega}^1$	$\hat{\omega}^2$	$\hat{a}_1^1$	$\hat{a}_2^1$	$\hat{a}_1^2$	$\hat{a}_2^2$	$\hat{b}^1$	$\hat{b}^2$	$\hat{\gamma}_1^1$	$\hat{\gamma}_2^1$	$\hat{\gamma}_1^2$	$\hat{\gamma}_2^2$
ABX	-0.991	-0.292	0.048	0.029	0.003	0.045	0.991	0.965	-0.189	-0.032	-0.060	-0.146
	(0.025)	(0.009)	(0.004)	(0.003)	(0.002)	(0.003)	(0.002)	(0.004)	(0.004)	(0.005)	(0.003)	(0.003)
ABY	-0.411	-0.898	0.058	0.010	0.015	0.043	0.965	0.984	-0.272	-0.093	-0.174	-0.260
	(0.018)	(0.024)	(0.005)	(0.004)	(0.004)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)	(0.007)	(0.007)
AEM	-1.137	-0.209	0.029	0.013	0.008	0.032	0.996	0.985	-0.228	-0.121	-0.135	-0.201
	(0.028)	(0.016)	(0.004)	(0.003)	(0.002)	(0.003)	(0.001)	(0.003)	(0.006)	(0.007)	(0.004)	(0.004)
AGU	-1.039	-0.287	0.063	0.032	0.002	0.044	0.975	0.976	-0.275	-0.131	-0.166	-0.271
	(0.024)	(0.019)	(0.008)	(0.005)	(0.004)	(0.004)	(0.006)	(0.005)	(0.008)	(0.009)	(0.006)	(0.005)
AL	-0.910	-0.293	0.045	0.023	0.006	0.022	0.995	0.992	-0.130	-0.036	-0.088	-0.108
	(0.030)	(0.014)	(0.003)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)	(0.003)	(0.003)
BCE	-0.446	-0.824	0.051	0.018	0.012	0.050	0.978	0.978	-0.158	-0.029	-0.145	-0.229
	(0.018)	(0.018)	(0.004)	(0.003)	(0.003)	(0.005)	(0.003)	(0.004)	(0.005)	(0.005)	(0.006)	(0.005)
BCM	-0.206	-1.404	0.047	0.037	0.021	0.111	0.971	0.946	-0.198	-0.041	-0.250	-0.213
	(0.018)	(0.025)	(0.004)	(0.007)	(0.005)	(0.012)	(0.004)	(0.011)	(0.005)	(0.006)	(0.009)	(0.009)
BMO	-0.214	-1.387	0.039	0.018	0.011	0.057	0.987	0.984	-0.189	-0.030	-0.225	-0.221
	(0.022)	(0.028)	(0.003)	(0.004)	(0.004)	(0.007)	(0.002)	(0.004)	(0.005)	(0.006)	(0.009)	(0.009)
BNN	-0.380	-1.048	0.046	0.010	0.011	0.034	0.952	0.989	-0.225	-0.064	-0.268	-0.273
	(0.020)	(0.032)	(0.007)	(0.005)	(0.006)	(0.007)	(0.012)	(0.006)	(0.008)	(0.009)	(0.010)	(0.010)
BNS	-0.087	-2.006	0.051	0.059	0.016	0.093	0.953	0.886	-0.207	-0.026	-0.254	-0.304
	(0.021)	(0.042)	(0.004)	(0.019)	(0.008)	(0.027)	(0.008)	(0.053)	(0.005)	(0.007)	(0.016)	(0.016)
BPO	-1.115	-0.286	0.051	0.046	-0.005	0.071	0.984	0.655	-0.262	-0.195	-0.229	-0.284
	(0.038)	(0.017)	(0.020)	(0.023)	(0.005)	(0.015)	(0.017)	(0.162)	(0.014)	(0.014)	(0.009)	(0.008)
BVF	-1.120	-0.646	0.048	0.023	0.013	0.024	0.999	1.001	-0.141	-0.080	-0.099	-0.172
	(0.036)	(0.028)	(0.002)	(0.002)	(0.002)	(0.001)	(0.000)	(0.000)	(0.004)	(0.005)	(0.004)	(0.003)
CCJ	-0.541	-0.571	0.063	0.035	0.023	0.038	0.997	0.999	-0.255	-0.098	-0.198	-0.226
	(0.042)	(0.037)	(0.004)	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)	(0.005)	(0.007)	(0.006)	(0.005)
CGT	-0.044	-2.041	0.059	0.037	0.013	0.075	0.991	0.975	-0.214	-0.046	-0.250	-0.362
	(0.043)	(0.060)	(0.005)	(0.011)	(0.009)	(0.020)	(0.003)	(0.010)	(0.006)	(0.009)	(0.021)	(0.019)
CLS	-0.697	-0.381	0.055	0.032	0.014	0.034	0.992	0.989	-0.169	-0.083	-0.115	-0.208
	(0.029)	(0.017)	(0.004)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.004)	(0.004)	(0.004)	(0.003)
CNI	-0.759	-0.419	0.034	0.017	0.000	0.023	0.996	0.990	-0.199	-0.075	-0.177	-0.213
	(0.034)	(0.017)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.005)	(0.006)	(0.004)	(0.004)
CNQ	-0.467	-0.790	0.051	0.023	0.014	0.040	0.985	0.978	-0.215	-0.079	-0.224	-0.221
	(0.025)	(0.020)	(0.004)	(0.004)	(0.003)	(0.006)	(0.003)	(0.006)	(0.005)	(0.006)	(0.006)	(0.006)
COT	-1.198	-0.235	0.029	0.008	0.002	0.034	0.991	0.982	-0.250	-0.125	-0.235	-0.232
	(0.029)	(0.022)	(0.007)	(0.005)	(0.003)	(0.004)	(0.006)	(0.005)	(0.011)	(0.012)	(0.007)	(0.006)
CP	-0.522	-0.832	0.053	0.016	-0.004	0.026	0.976	0.993	-0.213	-0.057	-0.235	-0.197
	(0.021)	(0.026)	(0.006)	(0.003)	(0.004)	(0.004)	(0.005)	(0.003)	(0.006)	(0.007)	(0.007)	(0.007)
DTC	-0.339	-1.098	0.064	0.022	0.002	0.056	0.927	0.940	-0.286	-0.079	-0.241	-0.280
	(0.019)	(0.025)	(0.009)	(0.009)	(0.007)	(0.011)	(0.021)	(0.023)	(0.006)	(0.008)	(0.010)	(0.010)
ECA	-0.574	-0.644	0.046	0.016	0.008	0.025	0.990	0.989	-0.165	-0.051	-0.142	-0.200
	(0.023)	(0.015)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.004)	(0.005)	(0.004)	(0.004)
ENB	-0.204	-1.573	0.049	-0.005	0.005	0.043	0.951	0.993	-0.276	-0.048	-0.307	-0.313
	(0.025)	(0.048)	(0.008)	(0.006)	(0.008)	(0.012)	(0.016)	(0.006)	(0.008)	(0.010)	(0.015)	(0.014)
ERF	-1.180	-0.379	0.037	0.017	0.014	0.028	0.997	1.000	-0.240	-0.192	-0.150	-0.222
	(0.033)	(0.029)	(0.004)	(0.003)	(0.003)	(0.002)	(0.001)	(0.001)	(0.008)	(0.009)	(0.006)	(0.005)
FDG	-1.059	-0.487	0.050	0.022	0.026	0.044	0.995	0.998	-0.249	-0.188	-0.224	-0.200
	(0.037)	(0.035)	(0.006)	(0.004)	(0.003)	(0.003)	(0.002)	(0.001)	(0.008)	(0.009)	(0.006)	(0.006)
FFH	-0.946	-0.400	0.056	0.032	0.010	0.035	0.991	0.988	-0.252	-0.145	-0.239	-0.291
	(0.040)	(0.031)	(0.009)	(0.006)	(0.004)	(0.006)	(0.004)	(0.005)	(0.011)	(0.011)	(0.008)	(0.007)
FHR	-1.245	-0.206	0.044	0.013	0.002	0.039	0.987	0.971	-0.284	-0.168	-0.162	-0.266
	(0.029)	(0.018)	(0.007)	(0.005)	(0.004)	(0.005)	(0.005)	(0.007)	(0.010)	(0.010)	(0.006)	(0.005)
FS	-1.982	0.031	0.099	0.019	0.002	0.049	0.979	0.975	-0.322	-0.223	-0.111	-0.294
	(0.050)	(0.023)	(0.024)	(0.008)	(0.007)	(0.004)	(0.012)	(0.004)	(0.015)	(0.015)	(0.006)	(0.004)
GG	-1.007	-0.286	0.040	0.023	0.017	0.026	0.994	0.991	-0.194	-0.081	-0.072	-0.184
	(0.026)	(0.017)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.005)	(0.006)	(0.004)	(0.003)
GIB	-0.221	-1.447	0.049	0.023	0.006	0.067	0.976	0.977	-0.258	-0.054	-0.302	-0.316
	(0.029)	(0.043)	(0.006)	(0.008)	(0.007)	(0.012)	(0.006)	(0.008)	(0.008)	(0.009)	(0.014)	(0.014)
GIL	-0.554	-0.728	0.045	0.029	0.002	0.093	0.975	0.935	-0.383	-0.139	-0.323	-0.325
	(0.037)	(0.042)	(0.011)	(0.014)	(0.009)	(0.015)	(0.016)	(0.023)	(0.013)	(0.015)	(0.015)	(0.015)
GLG	-1.267	-0.158	0.071	0.023	0.013	0.028	0.980	0.985	-0.231	-0.120	-0.135	-0.206
	(0.023)	(0.015)	(0.007)	(0.003)	(0.003)	(0.002)	(0.004)	(0.003)	(0.007)	(0.008)	(0.004)	(0.004)

Table C.2.1: Stock specific estimation results. The table contains estimated parameters of the ACI model in (4.3) and (4.4). Standard errors are reported in parentheses. For full company names see Table 4.3.1.

TICKER	$\hat{\omega}^1$	$\hat{\omega}^2$	$\hat{a}_1^1$	$\hat{a}_2^1$	$\hat{a}_1^2$	$\hat{a}_2^2$	$\hat{b}^1$	$\hat{b}^2$	$\hat{\gamma}_1^1$	$\hat{\gamma}_2^1$	$\hat{\gamma}_1^2$	$\hat{\gamma}_2^2$
IDR	-1.104	-0.302	0.037	0.017	-0.004	0.034	0.981	0.967	-0.269	-0.200	-0.214	-0.290
	(0.032)	(0.022)	(0.012)	(0.007)	(0.005)	(0.006)	(0.013)	(0.011)	(0.013)	(0.013)	(0.008)	(0.008)
IPS	-0.273	-1.201	0.051	0.069	0.032	0.093	0.941	0.858	-0.306	-0.086	-0.364	-0.341
	(0.031)	(0.045)	(0.014)	(0.022)	(0.013)	(0.027)	(0.032)	(0.070)	(0.012)	(0.014)	(0.017)	(0.017)
IQW	-0.327	-1.122	0.036	0.016	0.012	0.052	0.990	0.981	-0.253	-0.076	-0.276	-0.246
	(0.030)	(0.035)	(0.004)	(0.006)	(0.005)	(0.010)	(0.003)	(0.008)	(0.008)	(0.009)	(0.012)	(0.012)
ITP	-0.888	-0.407	0.040	0.027	-0.001	0.060	0.973	0.917	-0.385	-0.173	-0.281	-0.303
	(0.032)	(0.025)	(0.013)	(0.011)	(0.006)	(0.009)	(0.021)	(0.022)	(0.012)	(0.014)	(0.010)	(0.010)
KFS	-0.444	-1.023	0.036	0.019	0.022	0.035	1.000	0.996	-0.210	-0.095	-0.246	-0.315
	(0.035)	(0.034)	(0.004)	(0.004)	(0.004)	(0.008)	(0.001)	(0.004)	(0.008)	(0.009)	(0.012)	(0.010)
KGC	-0.655	-0.503	0.051	0.030	0.013	0.048	0.990	0.969	-0.175	-0.085	-0.101	-0.184
	(0.026)	(0.013)	(0.004)	(0.003)	(0.002)	(0.004)	(0.002)	(0.004)	(0.005)	(0.005)	(0.005)	(0.004)
MDG	-1.295	-0.146	0.046	0.020	0.008	0.030	0.988	0.980	-0.284	-0.137	-0.123	-0.206
	(0.024)	(0.014)	(0.006)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.007)	(0.008)	(0.004)	(0.004)
MFC	-0.573	-0.630	0.052	0.028	0.012	0.054	0.975	0.961	-0.220	-0.062	-0.166	-0.253
	(0.017)	(0.013)	(0.004)	(0.004)	(0.003)	(0.005)	(0.004)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
MGA	-1.236	-0.181	0.053	0.033	-0.001	0.032	0.990	0.977	-0.222	-0.121	-0.168	-0.238
	(0.028)	(0.014)	(0.007)	(0.004)	(0.002)	(0.003)	(0.003)	(0.004)	(0.008)	(0.008)	(0.005)	(0.004)
MHM	-0.426	-0.998	0.074	-0.003	-0.005	0.039	0.711	0.985	-0.336	-0.061	-0.293	-0.330
	(0.025)	(0.042)	(0.022)	(0.008)	(0.017)	(0.013)	(0.149)	(0.017)	(0.011)	(0.014)	(0.015)	(0.015)
MIM	-1.340	-0.220	0.112	0.017	-0.004	0.064	0.945	0.976	-0.316	-0.262	-0.179	-0.356
	(0.045)	(0.039)	(0.019)	(0.009)	(0.011)	(0.009)	(0.016)	(0.010)	(0.016)	(0.016)	(0.009)	(0.008)
N	-0.998	-0.597	0.051	0.015	0.007	0.021	0.996	1.000	-0.161	-0.061	-0.082	-0.156
	(0.030)	(0.038)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.004)	(0.004)	(0.003)	(0.003)
NCX	-0.852	-0.469	0.062	0.036	0.004	0.042	0.952	0.936	-0.216	-0.124	-0.244	-0.272
	(0.022)	(0.017)	(0.010)	(0.008)	(0.005)	(0.006)	(0.015)	(0.019)	(0.009)	(0.010)	(0.007)	(0.007)
NRD	-0.243	-1.167	0.056	0.025	0.010	0.034	0.979	0.980	-0.213	-0.044	-0.180	-0.260
	(0.023)	(0.022)	(0.005)	(0.004)	(0.004)	(0.006)	(0.004)	(0.005)	(0.005)	(0.005)	(0.008)	(0.008)
NT	-0.421	-0.980	0.041	0.025	0.012	0.018	0.999	1.000	-0.052	-0.030	0.005	-0.083
	(0.038)	(0.032)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.002)	(0.003)	(0.004)	(0.003)
NXY	-0.373	-1.140	0.048	0.015	0.010	0.031	0.993	0.998	-0.255	-0.059	-0.203	-0.236
	(0.034)	(0.036)	(0.005)	(0.003)	(0.003)	(0.004)	(0.002)	(0.001)	(0.005)	(0.006)	(0.008)	(0.007)
PCZ	-0.316	-1.279	0.047	0.018	0.015	0.030	0.999	0.998	-0.195	-0.050	-0.214	-0.220
	(0.038)	(0.032)	(0.003)	(0.003)	(0.003)	(0.004)	(0.001)	(0.001)	(0.004)	(0.006)	(0.008)	(0.008)
PDG	-0.727	-0.341	0.044	0.026	0.011	0.034	0.996	0.988	-0.156	-0.055	-0.054	-0.166
	(0.038)	(0.015)	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.004)	(0.004)	(0.003)	(0.003)
PDS	-0.919	-0.340	0.042	0.031	0.006	0.032	0.989	0.977	-0.228	-0.103	-0.178	-0.216
	(0.026)	(0.016)	(0.005)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.007)	(0.007)	(0.005)	(0.005)
PGH	-1.139	-0.484	0.049	0.021	0.017	0.027	0.998	1.000	-0.180	-0.152	-0.123	-0.194
	(0.037)	(0.029)	(0.004)	(0.002)	(0.003)	(0.002)	(0.001)	(0.000)	(0.007)	(0.008)	(0.005)	(0.005)
PKZ	-1.353	-0.471	0.066	0.029	0.020	0.027	0.997	0.999	-0.233	-0.159	-0.149	-0.231
	(0.041)	(0.030)	(0.005)	(0.002)	(0.002)	(0.002)	(0.001)	(0.000)	(0.007)	(0.008)	(0.005)	(0.004)
POT	-1.341	-0.223	0.051	0.009	0.005	0.038	0.983	0.991	-0.215	-0.130	-0.186	-0.234
	(0.028)	(0.028)	(0.008)	(0.004)	(0.004)	(0.003)	(0.005)	(0.002)	(0.010)	(0.010)	(0.005)	(0.005)
PWI	-0.948	-0.750	0.035	0.028	0.013	0.026	1.001	1.001	-0.207	-0.167	-0.174	-0.228
	(0.032)	(0.032)	(0.003)	(0.003)	(0.002)	(0.002)	(0.000)	(0.000)	(0.008)	(0.008)	(0.007)	(0.007)
RCN	-0.690	-0.632	0.029	0.039	0.020	0.065	0.989	0.889	-0.289	-0.191	-0.406	-0.259
	(0.035)	(0.031)	(0.008)	(0.014)	(0.007)	(0.014)	(0.007)	(0.039)	(0.015)	(0.015)	(0.012)	(0.013)
RG	-0.358	-1.152	0.060	0.032	0.011	0.088	0.967	0.901	-0.302	-0.057	-0.358	-0.204
	(0.025)	(0.026)	(0.007)	(0.011)	(0.006)	(0.013)	(0.008)	(0.027)	(0.007)	(0.009)	(0.009)	(0.011)
RY	-0.209	-1.299	0.047	0.022	0.007	0.059	0.986	0.978	-0.145	-0.026	-0.218	-0.183
	(0.024)	(0.025)	(0.003)	(0.004)	(0.003)	(0.008)	(0.002)	(0.005)	(0.005)	(0.006)	(0.008)	(0.008)
SLF	-0.432	-0.899	0.044	0.025	0.007	0.061	0.976	0.949	-0.204	-0.048	-0.234	-0.236
	(0.018)	(0.016)	(0.004)	(0.005)	(0.003)	(0.007)	(0.004)	(0.010)	(0.005)	(0.006)	(0.006)	(0.006)
SU	-0.445	-0.539	0.039	0.023	0.006	0.032	0.996	0.980	-0.221	-0.057	-0.134	-0.189
	(0.054)	(0.014)	(0.003)	(0.003)	(0.002)	(0.003)	(0.001)	(0.003)	(0.004)	(0.005)	(0.004)	(0.004)
TAC	-0.110	-2.267	0.036	0.002	0.006	0.025	0.993	0.998	-0.153	-0.039	-0.328	-0.311
	(0.038)	(0.067)	(0.005)	(0.006)	(0.008)	(0.014)	(0.003)	(0.009)	(0.008)	(0.010)	(0.024)	(0.023)
TD	-0.145	-1.381	0.041	0.030	0.007	0.095	0.992	0.948	-0.153	-0.036	-0.244	-0.240
	(0.027)	(0.023)	(0.003)	(0.006)	(0.004)	(0.013)	(0.002)	(0.014)	(0.005)	(0.006)	(0.009)	(0.009)
TEU	-0.562	-0.721	0.061	0.018	0.009	0.049	0.981	0.976	-0.264	-0.112	-0.271	-0.253
	(0.028)	(0.024)	(0.008)	(0.005)	(0.004)	(0.007)	(0.005)	(0.007)	(0.007)	(0.008)	(0.008)	(0.008)
TLM	-0.738	-0.490	0.052	0.023	0.014	0.032	0.983	0.981	-0.189	-0.091	-0.185	-0.211
	(0.024)	(0.017)	(0.005)	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)	(0.006)	(0.007)	(0.006)	(0.005)
TOC	-0.396	-1.048	0.038	0.014	0.013	0.059	0.988	0.964	-0.270	-0.057	-0.301	-0.204
	(0.026)	(0.025)	(0.005)	(0.007)	(0.004)	(0.014)	(0.004)	(0.018)	(0.006)	(0.007)	(0.008)	(0.009)
TRP	-0.544	-0.725	0.052	0.021	0.008	0.029	0.986	0.991	-0.189	-0.051	-0.162	-0.195
	(0.024)	(0.022)	(0.004)	(0.003)	(0.003)	(0.004)	(0.002)	(0.002)	(0.005)	(0.006)	(0.006)	(0.006)
TU	-0.210	-1.597	0.086	0.001	-0.003	0.084	0.910	0.966	-0.326	-0.055	-0.367	-0.345
	(0.023)	(0.043)	(0.008)	(0.007)	(0.011)	(0.015)	(0.015)	(0.012)	(0.007)	(0.009)	(0.013)	(0.013)
ZL	-0.341	-1.172	0.057	0.028	0.018	0.036	0.992	0.996	-0.236	-0.102	-0.197	-0.280
	(0.037)	(0.037)	(0.005)	(0.004)	(0.004)	(0.007)	(0.002)	(0.002)	(0.005)	(0.007)	(0.009)	(0.008)

Table C.2.1: Stock specific estimation results continued. The table contains estimated parameters of the ACI model in (4.3) and (4.4). Standard errors are reported in parentheses. For full company names see Table 4.3.1.

	TSX					
TICKER	IIS^1	$SE(IIS)$	HIS^1_{low}	HIS^1_{up}	HIS^1_{mid}	$SE(HIS_{mid})$
ABX	91.0	4.5	3.4	58.2	30.8	2.0
ABY	46.7	12.1	39.4	85.8	62.6	1.6
AEM	59.8	8.7	3.4	51.3	27.3	1.5
AGU	95.6	7.6	9.3	58.3	33.8	1.6
AL	88.5	2.5	5.6	82.5	44.0	0.9
BCE	60.0	7.5	48.8	92.5	70.7	1.9
BCM	41.9	7.3	47.2	96.2	71.7	1.8
BMO	52.3	10.1	59.2	94.7	77.0	1.6
BNN	54.5	17.6	40.3	76.1	58.2	2.3
BNS	67.0	14.4	70.2	95.1	82.6	1.8
CCJ	71.9	3.8	18.6	83.6	51.1	1.5
CGT	69.2	17.3	88.4	96.1	92.3	1.9
CLS	77.9	3.1	7.3	82.2	44.7	1.2
CNI	99.3	6.8	11.4	85.2	48.3	22.3
CNQ	67.9	6.6	33.0	92.5	62.8	1.4
COT	81.6	31.1	11.9	59.5	35.7	1.9
DTC	91.0	24.3	49.0	83.7	66.3	2.3
ECA	78.8	5.0	25.8	93.6	59.7	1.3
ERF	62.5	6.7	11.1	64.5	37.8	1.5
FDG	48.0	6.3	8.7	58.5	33.6	1.9
FFH	83.9	6.2	16.9	54.6	35.7	2.6
FHR	88.3	18.6	5.3	45.3	25.3	1.6
FS	96.0	16.0	0.1	21.7	10.9	1.7
GG	68.0	3.8	6.4	69.9	38.1	1.2
GIB	74.6	23.8	57.0	83.5	70.3	2.7
GIL	87.1	49.3	34.1	52.6	43.4	4.7
GLG	81.8	4.5	3.2	60.6	31.9	1.5
IPS	54.5	13.2	45.0	57.4	51.2	4.2
IQW	47.0	13.6	51.4	82.7	67.0	3.1
KFS	47.4	8.3	65.1	90.3	77.7	2.7
KGC	70.7	4.4	17.5	77.0	47.2	1.5
MDG	80.4	6.3	4.0	56.6	30.3	1.6
MFC	69.4	6.5	21.6	78.8	50.2	4.4
NCX	93.0	8.8	17.8	70.7	44.3	1.9
NRD	79.7	7.2	61.4	96.9	79.1	1.4
NT	82.2	2.7	5.4	84.5	44.9	1.1
NXY	70.8	8.3	52.8	94.5	73.7	1.8
PCZ	64.6	6.3	45.2	97.1	71.2	1.7
PDG	74.9	3.3	9.7	82.3	46.0	1.5
PDS	86.9	5.3	13.1	76.5	44.8	1.2
PGH	69.4	4.7	9.2	59.1	34.1	2.3
PKZ	78.5	3.0	7.2	68.7	37.9	1.8
POT	73.5	16.8	4.7	54.3	29.5	1.5
RCN	46.5	12.6	15.3	69.4	42.3	3.2
RG	66.8	14.3	45.1	87.2	66.2	2.3
RY	69.7	10.0	61.8	94.9	78.4	1.5
SLF	72.8	10.9	40.5	89.2	64.9	1.6
SU	81.8	4.6	28.1	92.7	60.4	1.7
TAC	30.9	70.1	78.2	86.3	82.2	3.6
TD	66.5	13.6	61.0	96.1	78.6	1.7
TEU	70.9	10.9	30.2	78.2	54.2	1.7
TLM	73.1	5.7	22.5	83.9	53.2	1.3
TOC	41.4	12.2	39.3	90.7	65.0	1.9
TRP	81.6	5.7	47.5	91.3	69.4	1.6
ZL	71.6	6.7	43.3	84.9	64.1	2.1

Table C.2.2: Stock specific intensity based information shares. The table presents TSX stock specific information shares using an intensity based (IIS^1) and the standard Hasbrouck approach in percent. The midpoint and the lower and upper bounds of TSX Hasbrouck are denoted by HIS^1_{mid}, HIS^1_{low} and HIS^1_{up}, respectively. The standard error for IIS is reported in the column labeled $SE(IIS)$ and for HIS_{mid} it is $SE(HIS_{mid})$. NYSE information shares can be calculated by $IIS^2 = 100 - IIS^1$, $HIS^2_{low} = 100 - HIS^1_{up}$, $HIS^2_{up} = 100 - HIS^1_{low}$ and $HIS^2_{mid} = 100 - HIS^1_{mid}$. The information shares are computed for 55 stocks that have positive $\mathbf{a}^s$ estimates and $\mathbf{B}$ estimates smaller than one. For full company names see Table 4.3.1.

C.3 VECM and Hasbrouck Shares

According to the law of one price, prices in different trading venues that refer to the same underlying asset are cointegrated, meaning that they can only deviate from each other in the short run. Assume that TSX and NYSE price dynamics can be described by a bivariate vector autoregression of order q, we model price changes, $\Delta p_t = p_t - p_{t-1}$, as a bivariate vector error correction model (VECM),

$$\Delta p_t = \alpha\beta' p_{t-1} + \Gamma_1 \Delta p_{t-1} + \ldots + \Gamma_{q-1}\Delta p_{t-q+1} + u_t \quad , \tag{C.3}$$

where $p_t = (p_t^1, p_t^2)'$, Γ_1 to Γ_{q-1} are 2×2 parameter matrices. $u_t = (u_t^1, u_t^2)'$ is a white noise vector with zero means and covariance matrix Σ_u. The vector $\alpha = (\alpha^1, \alpha^2)'$ contains the coefficients associated with the speed of adjustment of each price series to deviations from the equilibrium. β denotes the 2×1 cointegration vector, which implies that there exists one common stochastic trend, which can be considered as the stock's underlying efficient price. Hasbrouck's information shares are then derived as the contribution of an innovation in one market's price series to the underlying efficient price innovations variance. Since the VECM innovations (u_t) tend to be contemporaneously correlated, the shares cannot be uniquely identified. To solve this dilemma, Hasbrouck applies the Cholesky decomposition to the covariance matrix of innovations ($\Sigma_u = CC'$).

With the home market ordered first, Hasbrouck information shares of TSX (HIS^1) and NYSE (HIS^2) can be computed as,

$$HIS^1 = \frac{[\xi' C_{[1]}]^2}{\xi' CC' \xi} \quad \text{and} \quad HIS^2 = \frac{[\xi' C_{[2]}]^2}{\xi' CC' \xi} \quad , \tag{C.4}$$

where $\xi' C_{[j]}$ denotes the j^{th} element of the vector $\xi' C$ and ξ gives the common row vector in the matrix of long run impacts ($\boldsymbol{\Xi}$) of time t idiosyncratic innovations on the efficient price. It is derived as,

$$\boldsymbol{\Xi} = \beta_\perp [\alpha_\perp' (\mathbf{I}_n - \sum_{i=1}^{q-1} \boldsymbol{\Gamma}_i) \beta_\perp]^{-1} \alpha_\perp' \quad . \tag{C.5}$$

The Cholesky decomposition implies that the contribution of the market ordered first is maximized and that of the market ordered second is minimized. Since there is no theoretical justification for such a hierarchy, the common solution is to permutate the ordering

of the markets. This yields upper and lower bounds of information shares. The main drawback of Hasbrouck's methodology is that these bounds can diverge considerably, as the contemporaneous correlation between the composite innovations u_t^1 and u_t^2 tends to increase with decreasing sampling frequency.

In our application to Canadian stocks we choose a one minute sampling frequency using transaction prices. Thereby the US stock prices are converted to Canadian Dollars. After testing for cointegration using the maximum eigenvalue and trace statistic, we confirm the existence of one cointegration relation. The number of lags in the VECM is determined by the Schwarz information criterion (see Schwarz 1978). Standard errors for Hasbrouck's information shares are derived by a nonparametric bootstrap as proposed by Grammig et al. (2004).

Chapter 5

Conclusion

Modeling data on their lowest observation level is an extremely active research field in the empirical econometrics literature. Since irregular spacing of many financial and economic data is a key characteristic, new methodologies are introduced that account for this data property. These approaches are based on point processes which describe the history of events that occur consecutively in time. A process consisting of points at which we observe simultaneously variables that "mark" the points is called a marked point process. In this thesis we present new univariate and multivariate empirical (marked) point process studies that use irregular observed monetary and financial data.

Since precise predictions of short term interest rates are of key interest to investors and financial institutions, we investigate a marked point process model for the federal funds rate *target* and its forecast performance in Chapter 2. By modeling the target as a marked point process, Hamilton and Jordà (2002) considerably reduce the forecast mean squared errors compared to a standard time series method that uses equidistant data.

In Chapter 2 of this thesis we present a new marked point process for the target and show that the suggested specifications deliver improved results in terms of goodness of fit and in-sample forecast performance. The proposed methodology to evaluate probability function forecasts reveals useful target probability forecasts up to a six months horizon. Out-of-sample results are promising as well and Bayesian type model averaging robustifies the point forecast performance. We conclude from our findings that the model seems to capture very well the target characteristics: target changes occur in discrete time with discrete increments and

have an autoregressive nature.

As Chapter 2, Chapter 3 of this thesis investigates the forecast ability of a marked point process model. We present a time series model to predict total return variation. Due to the importance of accurate volatility forecasts in the valuation of derivatives, portfolio management and risk management, the prediction of volatility plays a central role in financial econometrics literature.

Chapter 3 is based on the work of Andersen et al. (2003) and Andersen and Bollerslev (1998) who use high frequency intraday data to introduce the concept of a nonparametric realized volatility measure. As shown by Barndorff-Nielsen and Shephard (2004) realized volatility can be decomposed into a continuous and a jump variation part. Since continuous variation is serially correlated, we model it by an autoregressive conditional time series model. Daily variation jumps that occur irregularly in time are conceived as marked point process. Continuous and jump variation models are combined to assess the accuracy of point and density forecasts of total return variation. The main findings of the empirical section can be summarized as follows. The estimation of the models yields sensible results in terms of diagnostics and parameter estimates. Density forecast evaluations confirm the suitability of this approach with respect to modeling the evolution of realized, continuous and jump variation. A point forecast analysis shows that the suggested model yields at least as accurate and in some cases even more accurate volatility forecasts than standard models that use equally spaced data.

In Chapter 4 we extend the univariate point process of the previous chapters to a multivariate point process and propose a new information share that measures the home and foreign market share in price discovery. We apply a bivariate autoregressive conditional intensity approach that accounts for the irregularity of the data, the informational content of time between consecutive trades and the timing interdependencies between two markets' transaction processes. In contrast to the commonly applied Hasbrouck (1995) methodology that requires equidistant data we deliver a unique information share rather than lower and upper bounds.

Since national stock exchanges fear to lose their attractiveness for investors and are

ambitious to remain the leading market with regard to price discovery, the identification of an information share is of paramount concern for a trading venue. We apply Hasbrouck's (1995) and our intensity based information share to analyze the price discovery process of Canadian stocks, which are traded on the Toronto Stock Exchange (TSX) and cross-listed on the New York Stock Exchange (NYSE). We find that despite the concern of the TSX to lose its share in price discovery to the NYSE, trading on the TSX still plays the most important role. Further, we show that the leadership of the TSX is even more pronounced than indicated by previous studies. The average TSX information share amounts to 71%. We also compare our results to the Hasbrouck (1995) information shares. On average we find a larger home market contribution than indicated by the Hasbrouck midpoints.

Summarizing the results of Chapter 2 through 4 we find evidence that using (marked) point processes to model non-aggregated data helps to explore the timing related information. Further, we conclude that accounting for the irregular spacing of economic and financial variables can improve a model's fit and forecast or solve problems arising using equidistant data. Overall, the presented findings are promising and in favor for (marked) point processes compared to standard econometric methods that use equally spaced data.

Bibliography

Aït-Sahalia, Y., P. A. Mykland, and L. Zhang (2005): "How Often to Sample a Continuous-time Process in the Presence of Market Microstructure Noise," *Review of Financial Studies*, 18, 351–416.

Akaike, H. (1974): "A New Look at the Statistical Model Identification," *IEEE Transactions on Automatic Control*, 19, 716–723.

Andersen, T. G., and T. Bollerslev (1998): "Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts," *International Economic Review*, 39, 885–905.

Andersen, T. G., T. Bollerslev, and F. X. Diebold (2007a): "Roughing It up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility," *Review of Economics and Statistics*, 89, 701–720.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001a): "The Distribution of Realized Stock Return Volatility," *Journal of Financial Economics*, 61, 43–76.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001b): "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96, 42–55.

——— (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579–625.

Andersen, T. G., T. Bollerslev, and X. Huang (2007b): "A Reduced Form Framework for Modeling Volatility of Speculative Prices Based on Realized Variation Measures," Working Paper.

Baillie, R. T., G. G. Booth, Y. Tse, and T. Zabotina (2002): "Price Discovery and Common Factor Models," *Journal of Financial Markets*, 5, 309–321.

Bandi, F. M., and J. R. Russell (2008): "Microstructure Noise, Realized Volatility, and Optimal Sampling," *Review of Economic Studies*, 75, 339–369.

Barndorff-Nielsen, O. E., and N. Shephard (2001): "Econometric Analysis of Realised Volatility and Its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society, Series B*, 64, 253–280.

——— (2004): "Power and Bipower Variation with Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2, 1–37.

——— (2006): "Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation," *Journal of Financial Econometrics*, 4, 1–30.

Bates, D. S. (2000): "Post-87 Crash Fears in the S&P 500 Futures Option Market," *Journal of Econometrics*, 94, 181–238.

Bauwens, L., and P. Giot (2000): "The Logarithmic ACD Model: An Application to the Bid-ask Quote Process of Three NYSE Stocks," *Annales d'Économie et de Statistique*, 60, 117–149.

——— (2001): *Econometric Modelling of Stock Market Intraday Activity*. Kluwer Academic Publishers.

Bauwens, L., P. Giot, J. Grammig, and D. Veredas (2004): "A Comparison of Financial Duration Models via Density Forecasts," *International Journal of Forecasting*, 20, 589–609.

Bauwens, L., and N. Hautsch (2006): "Stochastic Conditional Intensity Processes," *Journal of Financial Econometrics*, 4, 450–493.

Bergin, P. R., and O. Jordà (2004): "Measuring Monetary Policy Interdependence," *Journal of International Money and Finance*, 23, 761–783.

Bernanke, B. S., and A. S. Blinder (1992): "The Federal Funds Rate and the Channels of Monetary Transmission," *American Economic Review*, 82, 901–921.

Bollen, B., and B. Inder (2002): "Estimating Daily Volatility in Financial Markets Utilizing Intraday Data," *Journal of Empirical Finance*, 9, 551–562.

Bollerslev, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 307–327.

Bollerslev, T., U. Kretschmer, C. Pigorsch, and G. Tauchen (2009): "A Discrete-time Model for Daily S&P500 Returns and Realized Variations: Jumps and Leverage Effects," *Journal of Econometrics*, 150, 151–166.

Bowsher, C. G. (2007): "Modelling Security Market Events in Continuous Time: Intensity Based, Multivariate Point Process Models," *Journal of Econometrics*, 141, 876–912.

Brémaud, P. (1981): *Point Processes and Queues, Martingale Dynamics*. Springer, New York.

Carpenter, S., and S. Demiralp (2006): "The Liquidity Effect in the Federal Funds Market: Evidence from Daily Open Market Operations," *Journal of Money, Credit, and Banking*, 38, 901–920.

Carpentier, C., J.-F. L'Her, and J.-M. Suret (2007): "Competition and Survival of Stock Exchanges: Lessons from Canada," CIRANO Working Paper.

Carstensen, K. (2006): "Estimating the ECB Policy Reaction Function," *German Economic Review*, 7, 1–34.

Chernov, M., A. R. Gallant, E. Ghysels, and G. Tauchen (2003): "Alternative Models for Stock Price Dynamics," *Journal of Econometrics*, 116, 225–257.

Coffee, J. C. (2002): "Racing Towards the Top?: The Impact of Cross-listings and Stock Market Competition on International Corporate Governance," *Columbia Law Review*, 102(7), 1757–1832.

Corsi, F. (2004): "A Simple Long Memory Model of Realized Volatility," Working Paper, University of Lugano.

Davis, M. C., and J. D. Hamilton (2004): "Why Are Prices Sticky? The Dynamics of Wholesale Gasoline Prices," *Journal of Money, Credit, and Banking*, 36, 17–37.

Davutyan, N., and W. R. Parke (1995): "The Operations of the Bank of England, 1890-1908: A Dynamic Probit Approach," *Journal of Money, Credit and Banking*, 27, 1099–1112.

De Jong, F. (2002): "Measures on Contributions to Price Discovery: A Comparison," *Journal of Financial Markets*, 5, 323–327.

Demiralp, S., and O. Jordà (1999): "The Transmission of Monetary Policy via Announcement Effects," UC Davis Working Paper No. 99-06.

Denuit, M., and P. Lambert (2005): "Constraints on Concordance Measures in Bivariate Discrete Data," *Journal of Multivariate Analysis*, 93, 40–57.

Diebold, F. X., T. A. Gunther, and A. S. Tay (1998): "Evaluating Density Forecasts with Applications to Financial Risk Management," *International Economic Review*, 39, 863–883.

Diebold, F. X., and R. S. Mariano (1995): "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13, 253–263.

Dolado, J. J., and R. María-Dolores (2002): "Evaluating Changes in the Bank of Spain's Interest Rate Target: An Alternative Approach Using Marked Point Processes," *Oxford Bulletin of Economics and Statistics*, 64, 159–182.

Dueker, M. J. (1999): "Measuring Monetary Policy Inertia in Target Fed Funds Rate Changes," Working Paper, Federal Reserve Bank of St. Louis Review.

Dufour, A., and R. F. Engle (2000): "Time and the Price Impact of a Trade," *Journal of Finance*, 55, 2467–2498.

Easley, D., and M. O'Hara (1992): "Time and the Process of Security Price Adjustment," *Journal of Finance*, 47, 577–605.

Eichengreen, B., M. W. Watson, and R. S. Grossman (1985): "Bank Rate Policy under the Interwar Gold Standard: A Dynamic Probit Model," *Economic Journal*, 95, 725–745.

Eklund, J., and S. Karlsson (2005): "Forecast Combination and Model Averaging Using Predictive Measures," Working Paper, Stockholm School of Economics.

Engle, R. F. (1982): "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation," *Econometrica*, 50, 987–1008.

——— (2000): "The Econometrics of Ultra-high Frequency Data," *Econometrica*, 68, 1–22.

Engle, R. F., D. Hendry, and J.-F. Richard (1983): "Exogeneity," *Econometrica*, 51, 277–304.

Engle, R. F., and A. Lunde (2003): "Trades and Quotes: A Bivariate Point Process," *Journal of Financial Econometrics*, 1, 159–188.

Engle, R. F., and J. R. Russell (1997): "Forecasting the Frequency of Changes in Quoted Foreign Exchange Prices with the Autoregressive Conditional Duration Model," *Journal of Empirical Finance*, 4, 187–212.

——— (1998): "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data," *Econometrica*, 66, 1127,1162.

Eraker, B. (2004): "Do Stock Prices and Volatility Jump? Reconciling Evidence from Spot and Options Prices," *Journal of Finance*, 59, 1367–1403.

Eubank, R., and P. Speckman (1990): "Curve Fitting by Polynomial-trigonometric Regression," *Biometrika*, 77, 1–9.

Eun, C. S., and S. Sabherwal (2003): "Cross-border Listings and Price Discovery: Evidence from U.S.-listed Canadian stocks," *Journal of Finance*, 58, 549–575.

Evans, C. L., and K. N. Kuttner (1998): "Can VARs Describe Monetary Policy?," Working Paper 9812, Federal Reserve Bank of New York.

Evans, C. L., and D. A. Marshall (1998): "Monetary Policy and the Term Structure of Nominal Interest Rates: Evidence and Theory," *Carnegie-Rochester Conference Series Public Policy*, 49, 53–111.

Federal Reserve System (2005): *Purposes and Functions.* Board of Governors of the Federal Reserve System.

Frijns, B., and P. Schotman (2005): "Price Discovery in Tick Time," Working Paper.

Gaa, C., S. Lumkin, R. Ogrodnik, and P. Thurlow (2002): "The Future Prospects of National Markets and Trading Centers," Working Paper, Bank of Canada 2001-10.

Garratt, A., G. Koop, and S. P. Vahey (2006): "Forecasting Substantial Data Revisions in the Presence of Model Uncertainty," *Statistical Science*, 14, 382–417.

Gerlach, S. (2005): "Interest Rate Setting by the ECB: Words and Deeds," Working Paper 4775, CEPR.

Grammig, J., and K. Kehrle (2008): "A New Marked Point Process Model for the Federal Funds Rate Target: Methodology and Forecast Evaluation," *Journal of Economic Dynamics and Control*, 32, 2370–2396.

Grammig, J., and K.-O. Maurer (2000): "Non-monotonic Hazard Functions and the Autoregressive Conditional Duration Model," *Econometrics Journal*, 3, 16–38.

Grammig, J., M. Melvin, and C. Schlag (2004): "The Role of U.S. Trading in Pricing Internationally Cross-listed Stocks," Working Paper, Tübingen University, Arizona State University and Frankfurt University.

——— (2005): "Internationally Cross-listed Stock Prices During Overlapping Trading Hours: Price Discovery and Exchange Rate Effects," *Journal of Empirical Finance*, 12, 139–164.

Hall, A. D., and N. Hautsch (2006): "Order Aggressiveness and Order Book Dynamics," *Empirical Economics*, 30, 973–1005.

Hamilton, J. D., and O. Jordà (2002): "A Model for the Federal Funds Rate Target," *Journal of Political Economy*, 110, 1135–1167.

Hansen, P. R., and A. Lunde (2005): "A Forecast Comparison of Volatility Models: Does Anything Beat a GARCH(1,1)?," *Journal of Applied Econometrics*, 20, 873–889.

Harvey, D., S. Leybourne, and P. Newbold (1997): "Testing the Equality of Prediction Mean Squared Errors," *International Journal of Forecasting*, 13, 281–291.

Hasbrouck, J. (1995): "One Security, Many Markets: Determining the Contributions to Price Discovery," *Journal of Finance*, 50, 1175–1199.

Hausman, J. A., A. W. Lo, and C. A. MacKinlay (1992): "An Ordered Probit Analysis of Transaction Stock Prices," *Journal of Financial Economics*, 31, 319–379.

Hautsch, N. (2003): *Modelling Irregularly Spaced Financial Data*. Springer.

Hoeting, J. A., D. Madigan, A. E. Raftery, and C. T. Volinsky (1999): "Bayesian Model Averaging: A Tutorial," Birbeck College, University of London and University of Strathclyde and Reserve Bank of New Zealand.

Hong, Y., and H. Li (2005): "Nonparametric Specification Testing for Continuous-time Models with Applications to Term Structure of Interest Rates," *Review of Financial Studies*, 18, 37–84.

Huang, X., and G. Tauchen (2005): "The Relative Contribution of Jumps to Total Price Variance," *Journal of Financial Econometrics*, 3, 456–499.

Hupperets, E. C., and A. J. Menkveld (2002): "Intraday Analysis of Market Integration: Dutch Blue Chips Traded in Amsterdam and New York," *Journal of Financial Markets*, 5, 57–82.

Jansen, D.-J., and J. De Haan (2006): "Does ECB Communication Help in Predicting Its Interest Rate Decisions?," CESifo Working Paper Series No. 1804.

Jasiak, J. (1998): "Persistence in Intertrade Durations," *Finance*, 19, 165–195.

Judd, J. P., and G. D. Rudebusch (1998): "Taylor's Rule and the Fed: 1970-1997," *Economic Review*, (3), 3–16, FRBSF Publication.

Karolyi, G. A. (2006): "The World of Cross-listings and Cross-listings of the World: Challenging Conventional Wisdom," *Review of Finance*, 10, 99–152.

Karr, A. F. (1986): *Point Processes and Their Statistical Inference*. Marcel Dekker, Inc.

Kehrle, K. (2008): "Forecasting Return Volatility with Continuous Variation and Jumps," Working Paper, University of Tübingen.

Kehrle, K., and F. J. Peter (2009): "International Price Discovery in Stock Markets - A Unique Intensity Based Information Share," Working Paper, University of Tübingen.

KHOURY, S. (1990): "The Federal Reserve Reaction Function: A Specification Search," in *The Political Economy of American Monetary Policy*, ed. by T. Mayer, pp. 27–41. Cambridge University Press, Cambridge,England.

LANNE, M. (2006): "A Mixture Multiplicative Error Model for Realized Volatility," *Journal of Financial Econometrics*, 4, 594–616.

LIESENFELD, R., I. NOLTE, AND W. POHLMEIER (2006): "Modeling Financial Transaction Price Movements: A Dynamic Integer Count Data Model," *Empirical Economics*, 30, 795–825.

MERTON, R. C. (1980): "On Estimating the Expected Return on the Market. An Exploratory Investigation," *Journal of Financial Economics*, 8, 323–361.

MEULENDYKE, A.-M. (1998): *U.S. Monetary Policy and Financial Markets*. New York: Federal Reserve Bank.

PAN, J. (2002): "The Jump-risk Premia Implicit in Options: Evidence from an Integrated Time Series Study," *Journal of Financial Economics*, 63, 3–50.

PHYLAKTIS, K., AND P. KORCZAK (2007): "Specialist Trading and the Price Discovery Process of NYSE-listed Non-US Stocks," Working Paper, Europa-University Viadrina Frankfurt.

PRIGENT, J.-L., E. RENAULT, AND O. SCAILLET (2004): "An Autoregressive Conditional Binomial Option Pricing Model," in *Mathematical Finance. Bachelier Congress 2000*, ed. by G. H., D. Madan, S. Pliska, and T. Vorst. Springer.

ROSENBLATT, M. (1952): "Remarks on a Multivariate Transformation," *The Annals of Mathematical Statistics*, 23, 470–472.

RUDEBUSCH, G. D. (1995): "Federal Reserve Interest Rate Targeting, Rational Expectations and Term Structure," *Journal of Monetary Economics*, 35, 245–274, Erratum. 36 (December 95): 679.

——— (1998): "Do Measures of Monetary Policy in a VAR Make Sense?," *International Economic Review*, 39, 907–931.

Russell, J. R. (1999): "Econometric Modeling of Multivariate Irregularly-spaced High-frequency Data," Working Paper.

Russell, J. R., and R. F. Engle (2005): "A Discrete-state Continuous-time Model of Financial Transactions Prices and Times: The Autoregressive Conditional Multinomial - Autoregressive Conditional Duration Model," *Journal of Business and Economic Statistics*, 23, 166 – 180.

Sack, B. (1998): "Does the Fed Act Gradually? A VAR Analysis," *Federal Reserve Board of Governors FEDS Working Paper*, 17.

Sarno, L., D. L. Thornton, and G. Valente (2004): "Federal Funds Rate Prediction," Working Paper 2002-005, Federal Reserve Bank of St. Louis.

Schwarz, G. (1978): "Estimating the Dimension of a Model," *The Annals of Statistics*, 6, 461–464.

Scotti, C. (2005): "A Multivariate Bayesian Analysis of Policy Rates: Fed and ECB Timing and Level Decisions," Ph.D. thesis, University of Pennsylvania.

Thomakos, D. D., and T. Wang (2003): "Realized Volatility in the Futures Markets," *Journal of Empirical Finance*, 10, 321–353.

Zhang, M. Y., J. R. Russell, and R. S. Tsay (2001): "A Nonlinear Autoregressive Conditional Duration Model with Applications to Financial Transaction Data," *Journal of Econometrics*, 104, 179–207.

Zhang, Z. (2001): "Speculative Attacks in the Asian Crisis," IMF Working Papers 01/189, International Monetary Fund.