Essays on Intraday Volatility and Market Microstructure

by

Hanyu Zhang

Thesis submitted for the degree of

Doctor of Philosophy

ICMA Centre, Henley Business School

September 2016

University of Reading

For my parents

Bingke M.D. and Xiaoxin M.D.

Acknowledgements

First of all, I would like to thank Dr. Alfonso Dufour for giving me the chance to further my research skills in financial econometrics. I applied to his PhD at a very late time in early July 2012 when the application process was about to finish. I met him on a rainy day and talked to him about my enthusiasm for research and programming. He kindly gave me a paper to read, which became the cornerstone of my entire thesis. I often wonder the magic of time series and the dynamics of human behavior. I have grown from a student who only had some basic knowledge of econometrics to a young researcher who still has much to learn about the econometrics I love. I am also bounded to express my gratitude to my parents, especially to my mom who gave me endless supports throughout my whole PhD study. Special thanks to my colleague Xiaoyan and her supervisor Dr. Andreas Hoepner who helped me when I felt frustrated about my research. The thankful list is not complete without acknowledging Prof. Michael Clements for being my internal assessor and Dr. Emese Lazar for her valuable suggestions. Last but not least, I want to thank Jeff Bi and Dr. Zhiqiang Liu for being insightful and caring friends.

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Abstract

This work makes three main contributions to the financial econometrics literature. In Chapter 3, we study the intraday volatility of European government bonds under the framework of the multiplicative component GARCH model (Engle and Sokalska, 2012). We suggest a flexible and effective procedure for jointly filtering mid-quote prices and estimating volatility models and show that intraday data contain relevant information for daily volatility forecasts.

In Chapter 4, we show that a bond portfolio can reduce its intraday variance risk by including bonds from Italy and Spain. Furthermore, we demonstrate that the bivariate (scalar) DCC model is capable of computing an accurate VaR, providing correct conditional and unconditional coverage at lower than 1% (inclusive) confidence level and inducing lower losses.

In Chapter 5, we demonstrate that liquidity measures, such as the bid-ask spread and quantity available for trading at the best quotes, improve across maturities and countries after EuroMTS has allowed every market participant to post limit orders and not just designated market makers. In particular, we show that the relative bid-ask spread for trading 10 million bonds decreases with the rule change. The proportion of time when the relative bid-ask spread stays low also increases. The results suggest that greater competition amongst liquidity providers improves liquidity.

Note

Here is the list of abbreviations used:

Austria: AT

Belgium: BE

France: FR

Germany: DE

Italy: IT

the Netherlands: NL

Spain: ES

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Econometrics is potentially scientific precisely because alchemy is creatable, detectable and refutable.

– David F. Hendry, *Econometrics: Alchemy or Science?*

1.1 Motivation

With the recent European sovereign debt crisis, we have seen, probably for the first time since the last world war, several bond issuing countries running into debt repayment problems at the same time. With the repercussion of Brexit spreading throughout the whole Europe, the uncertainty about economics, financial markets and politics fuels the bond volatility of peripheral countries. On the very same day when the UK decided to quit the European Union (EU), the yield of the 10-year Spanish government bond ended up 17 basis points higher than the opening value while the 10-year German bund yield became more negative.[1] Four days later, it is the first time that the 10-year

[1] Financial Times, June 24, 2016 `https://next.ft.com/content/7a888f22-39a2-11e6-9a05-82a9b15a8ee7`

Spanish yield dropped below its Italian counterpart (a fall of 13 basis points) because a general election favors staying in the EU.[2] This recent example highlights the speed and volatility of investors' actions adjusting for various factors that might affect European bond markets. Studying and predicting volatility becomes ever more important in today's bond markets.

Intraday volatility is partly due to news and partly to trading, especially high-frequency trading – a concept that deeply worries regulators and government officials. In the other side of the world, the large price swing in the 2010 US Flash Crash is allegedly caused by computer algorithms that work on nanosecond intervals. A similar event has happened again in the US Treasury market, the most liquid bond market, 4 years after the Flash Crash. A round trip of 37 basis points for the yield of the benchmark 10-year US Treasury bond is unprecedented in history. It has been pointed out in the financial press that ”volatility is a concern as a lower appetite for Treasurys among investors could drive up borrowing costs not just to finance the U.S. budget deficit but also for corporations and individual mortgage loan holders”.[3] The turmoil was not finished when the Dow Jones Industrial Average plunged 1000 points along with the stock of JPMorgan dipping 20 percent because algorithms failed to set prices for stocks. With such frequent flash crashes and propagation of algorithmic trading, high-frequency volatility should be examined closely for the functioning of the fixed-income markets. In addition, liquidity as a crucial factor that influences treasury bond price dynamics should be investigated along with volatility.

Our empirical analyses have expanded the current literature on volatility and liquid-

[2] Bloomberg, June 28, 2016 `http://www.bloomberg.com/news/articles/2016-06-28/spain-s-bonds-extend-recovery-sending-yields-to-one-year-low`

[3] Wall Street Journal, July 13, 2015, `http://www.wsj.com/articles/u-s-report-finds-no-single-cause-of-oct-15-treasury-market-volatility-1436801464`

ity. In particular, we base our analysis on the high-frequency dataset of MTS, which is one of the largest European inter-dealer fixed-income market. There are over 500 unique counterparties and average daily volumes are over 100 billion Euros on the MTS platforms.[4] First, we suggest a methodology for selecting optimal data filters when estimating bond volatility from high-frequency data in Chapter 3. Our procedure is based on the multiplicative component Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model originally developed by (Engle and Sokalska, 2012). We estimate volatility for Euro-area sovereign bonds. The availability of quote and trade data has allowed researchers to develop more accurate measures of asset price volatility. Nelson (1990) shows that the conditional variance estimated with the ARCH model converges to the true variance when the duration between consecutive transactions goes towards zero. However, using all the available data is not always ideal because of the presence of microstructure noise documented in many articles (see, Zhang et al., 2005, and Bandi and Russell, 2008). The market microstructure noise arises because of frictions in the trading process such as price discreteness (price changes are measured as multiples of the tick size which is the minimum price variation)and large transaction costs (very large bid-ask spreads deter market participants from trading). In our government bond data, we often observe a sudden and temporary widening of bid and ask quotes, which generates large jumps and short-term fluctuations in mid-quote prices. Hence, frictions and illiquidity affect volatility estimation.

The current literature focuses mainly on achieving bias reduction in the realized variance by properly sampling the high-frequency data. Oomen (2005) discusses the optimal sampling frequency in calendar time based on a pure jump process of the transaction price. In order to use the full dataset, Zhang et al. (2005) propose to

[4]See `http://www.mtsmarkets.com/About-Us`

average realized variance estimates generated from a number of sampling grids. Bandi and Russell (2008) assume a general MA(1) structure for the noise and determine the optimal sampling frequency by minimizing the MSE of the realised variance estimator. However, the issue we face is different from the microstructure noise addressed by Bandi and Russell (2008) and related papers. We face jumps in the mid-quote price due to temporary frictions and illiquidity which may not be removed by simply changing the sampling frequency. Brownlees and Gallo (2006) have done some related work in duration modeling with irregularly spaced data.[5] The possible effects of the filters remain largely unknown and are not comparable to other filters. Many established filters have not been applied to equidistant data. Therefore, the current parametric modelling of volatility calls for a comprehensive evaluation of various filters and a method for finding the optimal filter based on a common benchmark.

Besides the data filtering application, we also illustrate the dynamics of intraday volatility itself and its role in forecasting daily variance, a field that attracts much academic and practitioners' attention. Estimating volatility of lower frequency from information obtained at higher frequency has a very long history (see French et al., 1987, Schwert, 1989, and Schwert, 1990). As the recent development of realized variance has suggested, high-frequency data is able to generate more accurate forecasts of daily volatility (Andersen et al., 2003a). There are many attempts to forecast realized variance (see e.g. Corsi et al., 2008), but only a few papers try to use intraday data to forecast daily volatility (see e.g. Andersen and Bollerslev, 1998, and Andersen et al., 2003b).

Second, as a natural extension to the univariate intraday volatility model in Chapter

[5]Engle and Russell (1998) ignore price changes larger than 4 ticks, which is also a filter to the dataset.

3, we examine the multivariate volatility of several European bond markets in Chapter 4. The motivation is to develop a model for studying the possible contagion effects and to illustrate its applications to risk management. Contagion, often defined as high correlation of asset returns (Forbes and Rigobon, 2002), is especially dangerous to portfolio executions. When securities are highly correlated, the classic diversification effect is highly reduced and a small perturbation to one asset leads to a chain of reactions. However, high-frequency correlation is not well examined in the literature. To the best of our knowledge, only Giot (2005) and Dionne et al. (2009) have analyzed the issue from the VaR perspective. With the newly developed DCC model, we are able to investigate how intraday correlation evolves over time and when correlation is the highest. Furthermore, we can directly assess the ability of the DCC model in measuring the intraday VaR. In the original paper of Engle (2002a), the author only examines 5% and 1% VaR and uses the Dynamical Quantile test of Engle and Manganelli (2004) to test the adequacy of various VaR for real datasets. An unconditional coverage test and a gauge for capital allocation efficiency is missing from Engle (2002a). Dionne et al. (2009) have computed intraday VaR using a complicated specification for duration and return from irregularly spaced data. Their computation relies on the availability of frequent transactions and generally good liquidity. A simpler model using regularly sampled data may suit better to the MTS markets because of the illiquidity found in Chapter 3 and the infrequent transactions of bonds.

Third, after considering univariate and multivariate volatility models for government bond prices, we now turn our attention to government bond liquidity in Chapter 5. Illiquidity affects government bond price dynamics. Regulators and practitioners are interested in developing good quality markets with high liquidity. We contribute to the discussion about the link between market structure and liquidity by studying the effect on liquidity of an important structural market change which transforms a particular trading platform from a dealership system to an order driven system. On November 15, 2012, MTS lifted the restriction for ordinary investors to submit limit orders in the EuroMTS platform where European benchmark bonds are traded. This event provides us with a unique opportunity to strengthen our understanding of how this change in market structure has affected bond market liquidity. The change is unique in at least three perspectives. First, it is exogenous to any security selections as the transition is initiated by the exchange. Therefore our study does not suffer from endogeneity and self-selection problem inherited in many earlier studies (see a sequence of papers that study the liquidity of stocks switching from one exchange to another: Christie and Huang, 1994, Clyde et al., 1997,Barclay, 1997, Huang et al., 2002, and Bennett and Wei, 2006). Second, it is a transition from a dealership market to an auction market in an electronic interdealer trading platform. The dealership is different from the traditional one seen in the NASDAQ or the LSE. Prior to the change, market participants can see all the quotes posted by dealers in the EuroMTS whereas dearlers' quotes are not disseminated to the public in the hybrid SETS in the LSE and were not in the NASDAQ. Third, the rule change was implemented when the crisis was still influencing the bond markets. While the overhaul of the NASDAQ is very efficient in improving liquidity proved by Barclay et al. (1999), the European sovereign debt crisis has certainly diminished the willingness of market makers to provide greater liquidity for European government bonds (see Figure 3.3 for the plots of monthly percentiles of the bid-ask spread and Fender and Lewrick (2015) for recent report on the overall liquidity of European fixed-income markets). Whether the measure taken by MTS is strong enough to take effect remains to be explored.

1.2 Main Contribution

In Chapter 3, we develop a modeling approach that filters out the noise and estimate various intraday volatility components at the same time based on the finding of Engle and Sokalska (2012) and Ghysels et al. (2014). With a piecewise linear spline, we estimate intraday periodicity jointly with intraday volatility as compared to a multi-step estimation in Engle and Sokalska (2012). The specification is less subject to the multi-step estimation error according to Newey and McFadden (1994). We construct a way of choosing an appropriate filter objectively. Bandi and Russell (2008) show how to reduce the effect of microstructure noise on volatility estimation by optimally choosing the sampling frequency. We use a similar approach for determining the optimal data filtering procedure rather than the sampling frequency. Stemming from illiquidity concerns, we generally avoid using standard deviation and directly filtering return in our filters, which may be damaging to the study of volatility, as it would lead to underestimation.[6] Importantly, we recognize the dynamic nature of the dataset and apply the filters accordingly. We highlight the reason why some filters which rely too heavily on neighbouring observations are not optimal for the MTS dataset. In addition, we show that our approach leads to the estimation of a model which uses intraday data and has better forecasting ability for daily volatility than a simple GARCH(1,1) model estimated on daily data. The comparison proves the usefulness of high-frequency information in the parametric modelling, given that GARCH(1,1) is still a dominant model in bond markets.

[6] Huang et al. (2002) throw out any observations with 10 standard deviations away from the daily mean of the mid-quotes and the daily mean of the bid-ask spreads.

The contribution of Chapter 4 is two-fold. First, we use multivariate volatility models for assessing contagion across Euro-zone treasury bond prices during the European government bond crisis. Second, we show how our model can be used for risk management purposes and for computing adequate VaRs. Moreover, we show that European treasury bond portfolios achieve a better diversification when Italian and Spanish bonds are included. To tackle the problem of nonsynchronous trading, we propose to fit a cubic spline to the correlation series. We choose the number of knots (less than 30) for the cubic spline using BIC criterion for the overall 39423 in-sample observations, which largely ignores the transient low correlation in between two consecutive knots. The methodology is simple enough as compared to a multivariate spline and still reveals the long-term trend. More importantly, we test whether bond pairwise correlation changes when ECB started acquiring debts of peripheral countries. Several articles examine the success of ECB's policy from different perspectives but none of them have examined correlation.[7] The ECB intervention appears to restore the correlation between Italy/-Spain and other European countries and bring back investors' confidence in southern European government bond issuers. In the second part, we prove that with a simple dynamics of correlation matrices of the DCC model, the intraday risk is correctly covered for lower than 1% (inclusive) VaR. The Kupiec (1995) chi-square test and the Dynamic Quantile test of Engle and Manganelli (2004) suggest that all lower than 1% VaRs generate an accurate unconditional and conditional coverage for market risk exposure, respectively. Also, the decaying weights of past returns in the bivariate DCC model give portfolio managers extra flexibility and efficiency in setting up risk capital.

When studying liquidity in Chapter 5, we intend to add more evidence to the merits

[7] See Pattipeilohy et al. (2013), Ghysels et al. (2014), Eser and Schwaab (2016), Babecka Kucharcukova et al. (2016), and Dufour et al. (2016).

of auction markets against dealership markets in the context of a pure electronic trading platform. Not only is a single measure of liquidity studied, but also combined statistics are formulated. Our results are consistent with the literature and the order-driven market generally has better liquidity than the quote-driven market. However, we find lower depth in the order-driven market. We believe that liquidity providers posting more frequently small-size orders is responsible for the decline in the quoted depth. Greater competition amongst liquidity providers may lead them to post aggressively priced limit orders which reduce the spread but for smaller quantities. In addition, we cover a broad range of assets in contrast with Albanesi and Rindi (2000) who concentrate on the Italian market. Albanesi and Rindi (2000) rely on the ample transactions of Italian bonds in three separate months whereas we use a continuous sample of quotations that lasts two years. Our research should also be relevant to practitioners and regulators who seek to improve liquidity supply in fixed income market to reduce transaction costs or avoid a flash crash.

1.3 Outline

The rest of the thesis is organized as follows. In Chapter 2, we give a comprehensive review covering a number of aspects related to high-frequency volatility and liquidity studies. We start from the seemingly trivial issue of data sampling, and we show that in fact it is crucial to adopt a proper data filtering method because this choice leads to very distinctive approaches to modeling volatility. We then discuss the relation between intraday volatility and the realized variance and provide some comments on the realized covariance, where many papers try to solve nonsynchronous trading problem. In particular, we consider the same problem when estimating multivariate DCC models in Chapter 4. We proceed to describe some stylized facts pertaining to daily volatility and some unique features of intraday volatility. Notably, we explain how researchers have detected recurrent intraday patterns in volatility, also called intraday periodicities or diurnal effects. We illustrate the alternative specifications used to model intraday periodicities in volatility. Three important intraday volatility models are discussed in detail in Section 2.5. We focus on the models of Andersen and Bollerslev (1998) and Engle and Sokalska (2012), which provide the background for our research as discussed in Chapter 3 and 4. A recent work by Liu and Maheu (2012) who exemplify the notion of Engle (2000)'s ACD-GARCH is included. Having reviewed most of the relevant issues with respect to intraday volatility, we introduce the concept of liquidity and show how liquidity is related to volatility through the Mixture of Distribution Hypothesis. Finally, Section 2.7 finishes the review with a brief survey of the high-frequency trading literature.

In Chapter 3, we propose the modified multiplicative component GARCH model of Engle and Sokalska (2012) in Section 3.2. We consider some detailed adjustments and variable constructions in Section 3.3.3. Under the assumption that intraday return follows the GARCH process, we introduce three groups of filters which target bid-ask spread and mid-quote price changes in Section 3.3.4. The evaluation is based on the benchmark MSE that we design in the spirit of Bandi and Russell (2008). We try to find the optimal filter minimizing the distance between the daily summation of intraday volatility and daily realized variance in Section 3.3.5. Robustness concerning the sampling frequency of the daily realized variance is examined for the filtering performance. The model estimation of daily volatility and intraday components are presented and a forecast evaluation for the daily GARCH(1,1) model is conducted in the Mincer-

Zarnowitz Regression. The forecast comparison between the daily GARCH(1,1) and the intraday GARCH begins in Section 3.5. We devise a forecasting scheme aiming to provide the same up-to-date information for the daily and intraday model. We use four criteria to gauge the accuracy of the point forecasts; one of them penalizes over- and under-prediction asymmetrically.

In Chapter 4, we combine the univariate multiplicative component GARCH with the multivariate DCC model. The daily volatility component is removed from the GARCH model, and hence we can concentrate on the intraday correlation of the entire period from 2009 to 2013. We illustrate two types of the DCC model: a bivariate version and a multivariate version where all debts of 7 countries are included and the estimation of the multivariate DCC is carried out by the composite likelihood of Engle et al. (2007). The bivariate version is used to study the intraday correlation and a cubic spline fitted to the correlation series depicts the long-term trend. We investigate the conditional change induced by the European Central Bank (ECB) and the unconditional change during the policy time in a dummy variable regression in Section 4.5. We turn to VaR computation and backtesting for four methods: historical VaR, the Constant Conditional Correlation model of Bollerslev (1990) and the two aforementioned DCC models. Kupiec (1995) test of and the Dynamic Quantile test of Engle and Manganelli (2004) are implemented to verify the unconditional and conditional risk coverage, respectively. In addition, we measure the efficiency of the different VaR in the spirit of Lopez (1998).

In Chapter 5, we study the liquidity implications of the event that all market participants can submit one-sided limit orders in the EuroMTS platform. Benchmark government bonds are divided into three categories: short term, medium term and long term based on remaining time to maturity. Several liquidity measures are formulated:

the daily average time-weighted bid-ask spread, the daily average time-weighted depth, the daily average time-weighted spread for 10 million bonds, and the relative time length when the spread is lower than the maximum daily average time-weighted spread plus 3 maximum daily time-weighted standard deviation. The nonparametric Wilcoxon signed-rank test and a regression with control variables are applied to an symmetric sample where there are 220 observations before and after the event date. In view of a clear time trend, we detrend the series and carefully check the statistical properties to meet the assumptions of the Wilcoxon signed-rank test. The OLS regression includes a dummy variable for the event and macroeconomic announcements along with other control variables. Section 5.8.2.1 explains the smaller depth by investigating the relative frequency of the undersized orders. The results of the Wilcoxon test are consistent with the OLS regression. Section 5.9 tests the robustness of our results by considering only on-the-run bonds and exploiting a difference-in-difference approach to compare liquidity between the local MTS platforms and the EuroMTS market. 5.10 concludes the entire analysis.

Chapter 2

A Review of Intraday Volatility and High-Frequency Econometrics Literature

2.1 Introduction

Asset price volatility is conventionally associated with the standard deviation in mathematical or statistical terms, yet covers a much broader spectrum in financial market studies. The word 'variability' in common usage is usually interchangeable with volatility, which is defined as a summary measure of the deviations from the expectation. Generally, volatility describes a fluctuating pattern of a variable over time. Campbell et al. (1997) point out that asset price volatility often reflects the unpredictability of public and private information. As shown by Kyle (1985), volatility of insider information and noise traders' demand affect not only asset price volatility but also market liquidity, which is defined as the ease for uninformed traders to execute a large transaction in a short period of time without adversely affecting profits. In addition, private information is partly incorporated into prices, which is manifested as increased volatility following informative trades in Kyle (1985). Documented by Kim and Verrecchia (1991), public announcements appear to induce high volatility for several hours. In this survey, we summarize all the possible methods of sampling and aggregation for studies of high-frequency data. We show that calendar time sampling is the most reasonable and affordable sampling method for modeling intraday spot volatility using the MTS dataset. Besides the review of the realized variance literature, we outline the empirical features of intraday volatility, which is one focus of later chapters. We analyze possible model specifications and estimation methods, whose empirical discoveries are discussed in detail. In addition, we highlight the relation between the market microstructure literature of liquidity with volatility studies, which is a new direction for research.

The rest of the chapter is organized as the following. Section 2.2 discusses the sampling issues while Section 2.3 focuses on the realized variance. Section 2.4 describes some stylized facts of intraday volatility and Section 2.5 reviews the existing models. Section 2.6 indicates some links with other market microstructure literature. At last, Section 2.7 concludes the literature review.

2.2 Data Sampling Issues

The first issue in addressing high-frequency data when modeling volatility is the data sampled at irregularly spaced intervals. Three approaches are widely recognized and lead to a very distinctive modeling of volatility. First, Engle and Russell (1998) underline the importance of the transaction time between successive trades in modeling the trading process. They observe the duration, i.e. the time gaps, clustering in high-frequency intervals in a way that is similar to the clustering in lower frequency exhibited by volatility. Naturally, duration is modeled in an autoregressive way, so the model is called the Autoregressive Conditional Duration (ACD) model. By modifying the duration of the time gap between price changes, the volatility of the price can be modeled in a similar manner. Specifically volatility intensity function is described as an EACD (2, 2) process (Exponential ACD). A forecast could be derived in terms of event time. There are several extensions of ACD model in modeling the price process. Bauwens and Giot (2000) attempt to use a logarithmic ACD (log-ACD) model in order to provide an alternative approach. They claim that the log-ACD model is less restrictive and provides a more flexible approach enabling more exogenous variables to be added. Bauwens and Giot (2003) also add an asymmetric information content to the log-ACD model. The inclusion of information innovation in the modeling price process establishes its significant role in explaining market behavior. However, under the ACD framework, volatility alone is not the variable of interest. Duration is the endogenous variable of the trading process which contains volatility as one aspect. Only a few studies that have devoted to combining the ACD model with the GARCH model. Engle (2000) suggests ultra-high-frequency GARCH models with a duration appearing in the conditional variance equation. The ACD model is intrinsically different from event time sampling (see below), though the it might appear that high-frequency data is used without further processing in both methods. The ACD model extracts information from the length of time between consecutive transactions, whereas event time sampling does not.

The second method transforms the irregularly spaced data into regular ones. Most econometric analysis is based on a fixed frequency of sampling time. More specifically, aggregation and transformation are especially useful when we try to identify which factors are driving the volatility process. The sampling time and frequency can be defined and determined in different ways. The most intuitive way is to sample in calendar time at a predefined frequency. Andersen and Bollerslev (1998) construct 5-minute returns in examining Deutsche-Mark-Dollar volatility. The length of the interval is chosen in order to avoid any bid-ask bounce effect that might be observed in shorter intervals. They utilize a systematic approach which recognizes three main determinants of intraday return volatility: macroeconomic announcements, calendar effects, and daily volatility. Microstructure noise and bid-ask bounce effects are mostly considered in volatility measures other than parametric models. With the increasing use of automated trading systems, the noise caused by high-frequency trading has dampened the validity of the models' result. Discrete price changes or bid-ask spread may cause the real volatility to move away from its true value. As argued by Engle and Russell (1998), the selection of period length for transformation results in a loss of the high-frequency characteristic and heteroskedasticity. In addition, the robustness of the results is unknown as there are no criteria for selection.

The event time sampling records data whenever an event defined by the variable of interests happens, which is advocated by Hasbrouck in several seminal works (see Hasbrouck, 1991, Hasbrouck, 1993, and Hasbrouck, 1995). In Hasbrouck (1991), the author studies the interaction between trade and price using a VAR framework. The reason that he does not aggregate further is that he manages to capture every movement induced by trades. As the information components are represented by two residuals, the model is self-contained and does not require any exogenous variable. With the trading proceeds, the price changes are affected in terms of quote revisions as responses to trade innovations. However price changes in his paper do not directly relate to volatility, as the interaction is of primary concern. Hasbrouck (1991) concludes that trades lead price changes because private information first revealed by transactions is later incorporated into price. It is therefore useful in the identification of influential factors for volatility. Dufour and Engle (2000) bring duration into the framework, which develops a richer picture of the interaction because of the key status of duration in the trading process.

Oomen (2006) discusses three types of sampling schemes and various sampling frequencies, and proposes a new approach, called transaction time sampling, which obtains data when a fixed number of shares are traded in the market. Oomen (2006) views the transaction price process as a discontinuous jump process with finite variation, and increasing the sampling frequency should reveal the true efficient price process. Considering market microstructure noise, the sampling frequency in one sampling scheme should minimize the Mean Squared Error (MSE) between the realized variance and true integrated variance. The author shows that the transaction time sampling has a much smaller MSE than the calendar time sampling when the optimal sampling frequency is chosen in different sampling schemes.

Regarding the MTS datasets, the calendar time sampling has some advantages that no others possess. In the empirical chapters, we essentially study the univariate and multivariate volatility of returns generated from mid-quote prices. The event time sampling would capture the tick-by-tick returns, which incorporates too much microstructure noise, e.g. the bid-ask bounce of Roll (1984). While Oomen (2006)'s transaction time and business time sampling[8] are able to reduce the noise, they assume that transactions are very frequent so that the data can be sampled based on the number of traded shares, which is not the case for the MTS inter-dealer markets. Bond transactions are

[8] Business time sampling requires that the trading intensity of a Poisson process is constant between the two observations.

usually executed in blocks and one day can often witnesses one transaction (if any) per bond. The calendar time sampling can achieve the same cleanness as the transaction time sampling and at the same time remain feasible for the MTS dataset.

The third way of handling high-frequency data is to combine data of different frequencies together. A new mixed-frequency modeling approach has surpassed the equidistant sampling notion and originates from two distinctive streams, i.e. Ghysels et al. (2005) and Corsi (2009). The Ghysels et al. (2005) Mi(xed) Da(ta) S(ampling) regression (MIDAS regression) has opened a new arena for researchers to accommodate more empirical features and explicitly study the interactions of financial variables at different frequencies. The authors design a weighted sum of past squared returns of higher frequencies in order to forecast the variance of lower frequency. The weighting function has the flexibility to control the decay of historical shocks and reduce measurement errors simultaneously. Ghysels et al. (2005) confirm the existence of the compensation for the expected return from a high conditional variance by using a lag window of 252 days in the MIDAS regression. Ghysels et al. argue that the risk-return trade-off is most pronounced in monthly returns, where most of the previous literature failed to find such a relation because of lack of information from higher frequencies. The MIDAS concept and GARCH models seem to complement each other in a natural way. Engle et al. (2008) decompose the monthly volatility into two multiplicative components of high and low frequency and the low-frequency volatility, which is related to macroeconomic fundamentals, is modeled in the spirit of the MIDAS regression. Decaying weights, specified in the form of an exponential function and associated with long lags, are estimated in the likelihood of a GARCH model. Engle et al. (2008) conclude that adding a macroeconomic variable, such as inflation or industrial production growth, in the MIDAS part can improve the long-term forecasting of a pure time series GARCH-MIDAS model. More forecasting exercises and comparisons can be found in Clements and Galvao (2008).

In contrast with Ghysels et al. (2005), who seek to understand long-term low-frequency returns better based on high-frequency information, Corsi (2009) provides a simple way of modeling high-frequency realized variance and specifically targets the long-memory feature. Corsi's Heterogenous Autoregressive model of Realized Volatility (HAR-RV) is inspired by the Heterogenous ARCH (HARCH) mode of Müller et al. (1997) and Dacorogna et al. (1998). The HAR(3) model, containing only a lagged daily, weekly and monthly realized variance computed from rolling windows, yields some remarkable improvement in out-of-sample forecasts against the short-memory AR models of realized variance while remains much simpler than the long-memory Autoregressive Fractionally Integrated Moving Average (ARFIMA) model. Moreover, Corsi et al. (2008) suggest that a HAR-GARCH specification accounting for time-varying volatility of realized volatility can generate even more accurate forecasts. The HAR model has now become the benchmark for forecasting realized volatility in the literature (see Andersen et al., 2007, Bollerslev and Todorov, 2011, Busch et al., 2011, and Maheu and McCurdy, 2011).

2.3 Realized Variance and Realized Covariance

There are various definitions of volatility, due to the wide range of applications. The daily realized variance defines the daily volatility as the summation of the squared intraday return over one day. Obviously, the measurement covers a rather different area from intraday and daily volatility modeling as it is a model-free measure, which can approximate a true integrated variance in different stochastic processes of efficient price. Nonetheless, the literature has shown some salient empirical features of the high-frequency data. The thought of using squared daily returns to estimate monthly volatility dates back to French et al. (1987). Interestingly, this article may also be the earliest effort to harness the bipower variation of Barndorff-Nielsen and Shephard (2004b), who confined themselves to the effects of jumps in returns. Barndorff-Nielsen and Shephard (2002a) and Barndorff-Nielsen and Shephard (2002b) have given formal proofs of various properties of the realized variance and established the consistency of the estimator. However, the convergence of the realized variance is hampered by the microstructure noise. Many authors have concluded that an optimal sampling scheme and sampling frequency are needed for achieving consistency (see Zhang et al., 2005, and Bandi and Russell, 2008).

On the application of the realized variance, Andersen et al. (2001) discuss 'the distribution of realized stock return volatility' of the Dow Jones Industrial Average (DJIA). Their study exploits the richness of high-frequency data in order to derive a more robust conditional and unconditional distributions of daily volatility and daily return correlation. The result confirms the earlier observation of daily volatility, which follows a highly right-skewed conditional distribution. The academic value of this paper for the intraday volatility literature is that it highlights the relation between daily volatility and its high-frequency counterpart. It also justifies the use of the GARCH model, as it finds a strong temporal dependence of daily volatility aggregated from intraday returns. In terms of the asymmetric effects of positive and negative returns, Andersen et al. (2001) find that their influence only marginal in scale for individual stocks and is only strongly present in equity index returns (see Nelson, 1991, and

Glosten et al., 1993). In line with Andersen et al. (2001), Ghysels et al. (2005) also suggest that persistence instead of asymmetry in volatility is the key effect in the Intertemporal Capital Asset Pricing Model (ICAPM) of Merton (1973). Bollerslev et al. (2006) proves the validity of a two-factor model, where the long-term factor almost exclusively accounts for the extended response of hourly volatility to past negative returns.

Other than investigating daily volatility based on realized volatility, Andersen and Bollerslev (1998) define 5-minute volatility as the absolute value of returns, which they claim is a less noisy measure relative to the squared return described by a GARCH (1,1) model. In order to control for the various components of intraday volatility, noise is required to be maintained at a low level. The most common thought about variation is the R-squared, which represents the varying part of the dependent variable explained by model-fitted values. Balduzzi et al. (2001) gauge the effect of macroeconomic announcements on volatilities of securities by observing the R-squared of a regression equation containing surprise of news and securities' returns. As the variance represents the second moment of a random variable, squared returns is still a viable candidate.

For completeness of the survey and comparison with multivariate GARCH models, papers on realized covariance and correlation are covered in this section. Barndorff-Nielsen and Shephard (2004a) have not only shown the limiting distribution of the realized covariance and realized correlation estimators, but also proposed the validity of using a high-frequency realized regression. Some simulations in Barndorff-Nielsen and Shephard (2004a) sketch the convergence speed of the estimators, and the confidence interval of realized covariance varies significantly over time as the univariate conditional volatility moves. This suggests an intimate links between multivariate and univariate models. Following the nice theoretical properties of the covariance estimator in Barndorff-Nielsen and Shephard (2004a), several authors have sought to incorporate more realistic assumptions into the framework. Hayashi and Yoshida (2005) have designed so-called ”Cumulative Covariance” estimator, which multiplies a tick return of asset A with the overlapping tick returns of asset B, in order to alleviate the downward bias in the realized covariance, caused by the nonsynchronous trading of two assets. The estimator converges to the true covariance matrix as the sampling interval goes to zero regardless of the synchronicity of returns, and is attractive to practitioners because it is easy to implement. Hayashi and Kusuoka (2008) relax the assumption that the observation time is independent of the stochastic price process itself in Hayashi and Yoshida (2005) and prove the consistency under arbitrary stopping time.

However, the theoretical derivations and properties of Hayashi and Yoshida (2005) depend on the absence of market microstructure noise. Voev and Lunde (2007) show that such noise essentially biases the estimator and invalidates its consistency, which calls for the necessity of bias-correction methods. They propose a new estimator, which can be adjusted for the lead-lag relationship between asset prices, in a market where transactions are occasionally synchronized. The consistency of the new bias-correction estimator is achieved by subsampling in the spirit of Zhang et al. (2005). After correcting different realized covariance estimators according to the lead-lag effects, as in Voev and Lunde (2007), Griffin and Oomen (2011) rank the efficiency of those estimators based on the relative magnitude of the cross-asset correlation and the microstructure noise. Corsi and Audrino (2012) notice a problem pertaining to all the adjustment in Voev and Lunde (2007) and Griffin and Oomen (2011) – that is, the accuracy of the observed transaction time is subject to rounding, which influences the matching of concurrent returns. Other works on this subject dealing with the jumps in financial returns are beyond the scope of this survey (see e.g. Koike, 2016).

2.4 Stylized Facts Concerning Intraday Volatility

After resolving the preliminary problems of sampling data, econometricians attempt to shed light on any influential factors of the volatility process as well as to try to accommodate some regularity in high-frequency data. Compared to the whole world of daily- and lower-frequency volatility models created over the past twenty years, there are only a few papers devoted solely to modeling high-frequency volatility. Nonetheless, it is still valuable to look at the traditional field, which provides opportunity to bring the traditional model into the high-frequency arena. Engle and Patton (2001) suggest the quality of a good volatility model for daily data. They list several stylized facts which a volatility model must consider. High persistence in volatility should be the first characteristic involved in volatility modeling and the concept of persistence motivates the invention of the ARCH/GARCH class model (Engle, 1982, and Bollerslev, 1986). The ARCH/GARCH model captures the feature that an innovation in volatility can persist for more than one period. Apart from the high persistence in daily levels, the mean reversion phenomenon dominates in the long term. It demonstrates that eventually the conditional variance will converge to the unconditional variance and so will the volatility. However, microstructure noise often confounds the observability of mean reversion as is seen in Section 2.3, where realized variance can be an inconsistent estimator. Finally, some exogenous variables such as macroeconomic announcement, time-of-the-day effect or day-of-the-week effect might also influence volatility in the short term.

Usually, government bond markets respond to macroeconomic news instead of any specific news of companies. In a high-frequency framework, the volatility could be heightened for several minutes or hours, as confirmed by Ederington and Lee (1993) who examine the effects of several announcements on U.S. Treasury bond futures and reach the conclusion that the scheduled releases of economic data have significantly raised return volatility for the subsequent couple of hours. When one examines macroeconomic announcements, the impact could be different for various securities. Balduzzi et al. (2001) investigate the impacts of 26 types of macroeconomic news on different associated US Treasury securities in the GOVPX inter-dealer market. For instance, the price volatility of the 10-year Treasury bond reacts most actively to the unexpected component of the employment announcement and the Producer Price index (PPI) announcement whereas the price volatility of the 2-year one is best explained by the shock in the Civilian Unemployment and the Nonfarm Payrolls. Another significant contribution of Balduzzi et al. (2001) is that they separate the effects of simultaneous announcements, which Ederington and Lee (1993) fail to do, in that they capture the repercussion using dummy variables.

There are several features of Euro Area news releases, which make identification of any announcement effect difficult to implement. First, there is a prolonged announcement period for each type of economic data. The CPI of each country, for example, is announced gradually, starting from Germany and Italy, then Spain, finally to the Eurozone as a whole. Moreover, these CPIs are only estimates and subject to later correction, which would further complicates studies, as the correction can be postponed until the end of a quarter. The process implies that by the end of the releasing period the European CPI has already been predicted with some precision by investors and thus rarely causes a surprise to market expectations. The same implication is also documented by Andersson et al. (2009). They discover that in pre-crisis period 1996-2005 the German employment reports did not move the German Bund market because the employment figure had already been a common knowledge by the time of the release. Second, the number of macroeconomic news surged after the creation of Euro, but the importance of news varies greatly over time. German economic news always has the highest priority in market interpretation, in which French news ranks second. However, after the explosion of sovereign credit it is clear that some political risks dominate the market, e.g. Italian and Greek government elections. Third, definitions of the statistics may differ from their US or international counterparts. For instance, the German Bundsbank adds those who are looking for jobs into the unemployed whereas the international definition does not include that part of people in unemployment calculations. Due to the complications of the data release schedule, it is very difficult to separate and identify the individual effect of each announcement.

Along with those empirical properties exhibited in lower-frequency data, intraday volatility has shown many unique features. Periodicity, or known as the U-shape pattern, is the most famous one. As found in the empirical study of Wood et al. (1985) and Harris (1987), intraday volatility is high at the opening and closing period of a market and remains low for the rest of the trading day. Admati and Pfleiderer (1988)) are among the first who propose its existence and provide with theoretical support. They claim that the part of those liquidity traders who have the flexibility to schedule their trading tend to provide liquidity aggregately, which would attract privately informed traders. This concentration of activities may be observed as a U-shape curve in volatility diagram, and repeats every day. The U-shape curve is later argued to be a manifestation of traders' learning processes and inventory management instead of the incorporation of private information Hsieh and Kleidon, 1996. They examine the volatility pattern in Deutsche-Mark-Dollar foreign exchange market of New York and London simultaneously. Their results show that volatilities in two different markets do not interact with each other in overlapping trading hours, which accounts for a failure in any dissemination of information between the two markets. This empirical outcome disproves the conclusion of Admati and Pfleiderer (1988), who claim that U-shape volatility represents the aggregated informed trading in the opening and closing period. Further, lower volatility during the day is because of the inactivity of dealers who satisfy their needs rather than a lack of information incorporation. Nonetheless, Hsieh and Kleidon (1996) state that the disproof may be due to the special nature of the foreign exchange market, such as simultaneous trading in different area, a 24-hour market. The seasonal pattern of intraday volatility can be very different in fixed-income related market. Bollerslev et al. (2000) find out that there are two spikes caused by concurrent announcements in intraday volatility defined as the absolute return variation in the US bond futures market. The volatility pattern is similar to the finding of Ederington and Lee (1993), though the two discoveries are established from different sample periods.

Responding to earlier investigations into intraday seasonality, Andersen and Bollerslev (1997) discuss its implication for modeling volatility. The authors note that the direct application of an ARCH model on high-frequency returns results in different and potentially conflicting conclusions across different sampling frequencies. Dacorogna et al. (1993) and Müller et al. (1990) examine the intraday volatility pattern of Deutsche-Mark-Dollar market using a time invariant polynomial approximation. Their explanations and results may be very specific to foreign exchange trading due to the special nature of this 24-hour market. Engle and Russell (1998) suggest a cubic spline specification to smooth out periodicity from durations, which is also applied to returns by Giot (2000). Taylor and Xu (1997) remove the seasonal pattern by preestimating the factor from the realized variance and allowing different periodic shapes in announcement days and days without announcements. The parameter estimates are robust to this partition of days and unexpected jumps in returns, though Boudt et al. (2011) point out that idiosyncratic jumps can impair the estimation of periodicity, as the traditional method extracts information by simple aggregation.

The recent development of modeling periodicity is no longer restrictively referring to the U-shape pattern. More interest has been focused on nonparametric or semi-parametric studies of periodic behavior. There are various ways to model intraday periodicity, such as using dummy variables, the Fourier flexible form (FFF), wavelet form. Using dummy variables is the least efficient way of conducting the research, as it consumes too much data without giving precise pattern of them. Nonetheless, it is a simple way to generate a rough picture of the movement. Ranaldo (2009) studies the different segments of trading phase in the foreign exchange market and shows an empirical pattern of spot currency returns. Qualitatively, the result still confirms the U-shape pattern based on 4-hour intervals. The FFF approach pioneered by Gallant (1981), and extended by Andersen and Bollerslev (1997) allows a polynomial decay in the effects of macro news, which is widely used in the voluminous literature (see Martens et al., 2002, Bauwens et al., 2005, and Harju and Hussain, 2011). The very parsimonious structure can be easily applied not only to returns but also other to variables (see e.g. Hardle et al., 2012). Originated from the Fourier transform, wavelet technique in financial applications started to receive attention only a decade ago. Compared to Fourier transform, wavelet analysis is more adapted to the local properties of time intervals and the scale of returns (Gençay et al., 2001b).

At last, it is clear that different markets with different characteristics can heavily influence the final result. It is advisable to discuss the features of the largest electronic market for European government bond – the MTS market. The most distinctive feature of MTS is that one government bond can be traded in local MTS markets and also in the EuroMTS market – an international platform for all European benchmark securities (Dufour and Skinner, 2004). This specialty of MTS may lead to fragmentation of orders and higher short-term volatility (O'Hara and Ye, 2011). Other market features are also very important in volatility process. For example, the difference in trading hours may yield different intraday patterns. The US Treasury futures market usually closes at 15:00 Eastern Standard Time (EST), which renders some announcements scheduled at 16:30 unable to be examined directly. A U-shape pattern is less discernible in foreign exchange markets where trading is round-the-clock. For MTS market, the trading hour is pre-market 7:30am – 8:00am Central European Time (CET), pre-open 8:00am – 8:15am CET and regular trading 8:15am – 5:30pm CET (Dufour and Skinner, 2004). Therefore, volatility patterns in the MTS market should reflect most of the reactions to the different economic figures and monetary policies.

2.5 Models of Intraday Volatility

Considering the limited number of papers focusing on the systematic accounting of volatility for high-frequency data, the existing model deserves a deeper discussion. Nelson (1990) points out that the ARCH model can be seen as an approximation to a diffusion model, demonstrating the natural advantage of the ARCH model in utilizing high-frequency data. Giot (2000) applies the most frequently used GARCH (1, 1) and EGARCH model to US stock market. He claims that after controlling the intraday periodicity and seasonality the latent volatility process can be modeled under the temporal aggregation framework. The persistence parameters of both models turn out to be generally significant and amount to 0.9 and 0.95, respectively. The asymmetric response to negative returns has been shown in EGARCH model, where the additional parameter is significant and slightly negative.

Due to the similarity between the foreign exchange and fixed income markets, the volatility models of the two markets should receive equal attention. Andersen and Bollerslev (1998) examine the Deutsche-Mark-Dollar foreign exchange market and decompose determinants of the intraday volatility into three components: the daily ARCH effect, calendar effect, and macroeconomic news effect. The decomposition comes from a simple fact that the three main components best explain the deviation of returns from their expectations. They propose that the three components are indispensable when modeling volatility and any omission would lead to a distorted outcome (Andersen and Bollerslev, 1998). The reason behind their argument is that high volatility is usually accompanied by high volume, no matter whether daily or intraday data is used. Therefore, there must be some basic machinery behind the ARCH and calendar effect related to intraday returns. In other words, intraday movements must contain some information about this long-memory volatility. And from the investors' perspectives, they must take account of these effects when they react to market changes. The periodicity of intraday returns is realized by a Fourier flexible form (FFF):

$$
\begin{aligned}
&2\log\frac{|R_{t,n}-\bar{R}|}{\hat{\sigma}_{t,n}} = \\
&\hat{c}+\mu_0+\sum_{k=1}^{D}\lambda_k * I_k(t,n)+\sum_{p=1}^{P}(\delta_{c,p}\cdot\cos\ \frac{p2\pi}{N}n+\delta_{s,p}\cdot\sin\ \frac{p2\pi}{N}n)+\hat{u}_{t,n}
\end{aligned}
\tag{2.1}
$$

The categorical patterns of intraday volatility associated with calendars are captured by $\lambda_k * I_k(t,n)$, where λ_k represents the scale of the impacts and $I_k(t,n)$ is either a dummy variable for holidays and weekdays or a polynomial function of time. This superimposed restriction $I_k(t,n)$ allows a smooth decay of announcement effects and a efficient estimation of Equation (2.1). In order to model intraday volatility while controlling for long-term volatility, the daily GARCH forecast $\hat{\sigma}_{t,n}$ is incorporated for every intraday return $R_{t,n}$ by dividing the true forecast $\hat{\sigma}_t$ by the number of intraday intervals N. Their major findings are as following. First, among all macro news, the American Employment Report has the largest effect (15 % increase) on the daily cumulative absolute return, with the Advance Report on Durable Goods and the meeting of the German central bank following it. The authors argue that it is the different levels of controversy about those releases that sway the market heavily. The rest of the news typically has less than 5% influence. Second, they disprove the notion of a day-of-week effect; it appears that the scheduled announcements absorb the explanatory power of the day-of-week dummies. Third, the success of filtering out the intraday component produces volatility clustering – the ARCH effect – which confirms that daily volatility does have a long memory and suggests that it follows a fractionally integrated GARCH process. Finally, the periodic shape induced by normal trading activities explains much of intraday return variation in an in-sample forecast evaluation, although macroeconomic announcements play a central role in the foreign exchange market. One important element missing from Andersen and Bollerslev (1998) is that they do not include any out-of-sample forecast evaluation. The model has fitted the data closely but the complication of the model may hint at the possibility of it overfitting the data.

There has been a long history of decomposing daily volatility into different frequencies in the GARCH literature. Engle and Lee (1999) are among the first to separate a long-term component that has a rather slow mean-reverting rate from the transitory volatility component. Engle (2002b) creates a new class of the Multiplicative Error Model (MEM) for the conditional mean in order to provide better statistical support for various variables of interests. Combining with the earlier component based GARCH, we have seen the proliferation of the multiplicative component GARCH models.

Engle and Sokalska (2012) describe the intraday return as a multiplication of three components:

$$r_{t,i} = \sqrt{h_t s_i q_{t,i} \varepsilon_{t,i}} \tag{2.2}$$

where h_t is the daily variance, s_i is the diurnal (calendar) variance, $q_{t,i}$ is the intraday variance with mean 1, and $\varepsilon_{t,n}|\Phi_{t,n-1} \sim N(0,1)$. The diurnal adjustment s_i is estimated as the average squared return in the same minute bin over the entire sample. In order to control for the daily volatility component, it must be estimated separately. Andersen and Bollerslev (1998) estimate daily volatility using a MA(1)-GARCH (1, 1) model based on a daily sample longer than the span of high-frequency data. Engle and Sokalska (2012), among others, adopt a risk-factor model, incorporating industry and liquidity information into a time series analysis of volatility. After estimating the first two components, the final GARCH specification treat the multiplication of intraday volatility and the error term as the mean equation and $q_{t,i}$ as the conditional variance. Newey and McFadden (1994) show the consistency and efficiency loss of a multi-step estimation in the GMM framework. It is preferable to have estimations completed in one step. The results for individual stocks do not seem to reveal much a picture of volatility clustering. For a randomly selected stock of Valero Energy Corporation $\alpha+\beta$ is equal to 0.814, which is relatively low compared to the daily GARCH model. The authors seek to improve the poor performance by pooling the the stocks according to trading frequency. The intuition behind this is that any private or public information may result in a prolonged trading period for the actively traded stocks whereas thinly traded stocks may have a lower persistence of volatility. Liquidity and the industry code serve as two grouping criteria. Liquidity is measured by the average number of trades per day and 2721 companies are sorted into 54 industry groups. Not surprisingly, $\alpha+\beta$ is still relatively low after pooling according to companies' industry code, mostly ranging from 0.86 to 0.96. The liquidity sorting produces a decreasing trend for the GARCH parameter but an increasing trend for the ARCH one. Overall, the range of $\alpha+\beta$ is narrowed by this pooling; the magnitude of persistence remains the same as the industry code pooling.

An asymmetric response is a very common features for asset returns yet has not so far been modeled in high-frequency area. Engle and Sokalska (2012) give the estimation of a GJR specification (Glosten et al., 1993) for intraday volatility. Surprisingly enough, the leverage parameter γ has a negative sign, signifying that the positive returns have larger impacts. The cause of the problem remains unknown and is definitely a focus for further research. The forecasting performance is evaluated based on the MSE and likelihood based loss function. It turns out that the liquidity sorting parameters outrank the others in predicting the least liquid stocks whereas the parameters derived from one large-group GARCH perform the best in the forecasting of most liquid stocks. In addition, the forecasting performance is not improved by considering the asymmetric response.

Liu and Maheu (2012) extend the ACD-GARCH framework of Engle (2000) and apply the model to the American and the Chinese stock markets. The ultimate goal is to fuse the duration process with a multiplicative GARCH. They start with a basic ACD model with the Burr distribution (BACD) for the innovation. Specifically:

$$\chi_i = f(\psi_i) z_i \tag{2.3}$$

where

$$f(\psi_i) = \psi_i \frac{(\omega^2)^{(1+\frac{1}{\kappa})} \cdot \Gamma(\frac{1}{\omega^2} + 1)}{\Gamma(1 + \frac{1}{\kappa}) \cdot \Gamma(\frac{1}{\omega^2} - \frac{1}{\kappa})} \quad 0 < \omega^2 < \kappa \tag{2.4}$$

The Burr distribution of innovation z_i is

$$g(z_i) = \frac{\kappa z_i^{\kappa-1}}{(1 + \omega^2 z_i^{\kappa})^{(1/\omega^2)+1}} \tag{2.5}$$

The return simply follows an ARMA(1,1) process and the HAR-BACD nests the two models together, i.e. the Burr-ACD and the HAR model of Corsi (2009):

$$r_i = \rho r_{i-1} + u_i + \phi u_{i-1} \tag{2.6}$$

$$u_i = \sqrt{q_i} \zeta_i \quad \zeta_i \overset{i.i.d}{\sim} t_\nu(0,1) \tag{2.7}$$

$$q_i = \beta_0 + \sum_{m=1}^{M} \beta_m VC_{i-1,h_m} + \gamma_1 \chi_i^{-1} + \gamma_2 \frac{\chi_i}{\psi_i} + \gamma_3 \psi_i^{-1} \tag{2.8}$$

where

$$VC_{i-1,h_m} = \frac{u_{i-1}^2 + \ldots + u_{i-h_m}^2}{h_m} \tag{2.9}$$

VC_{i-1,h_m} is the realized variance associated with each different aggregation level h_m. By inheriting the ability of the HAR model, Equation (2.8) is able to produce the long-memory feature in a parsimonious way. Moreover, VC_{i-1,h_m} corresponds to a group of investors with an investment horizon h_m, a direct testimony of the Heterogeneous Market Hypothesis of Müller et al. (1997) when multiple h_m are included in Equation (2.8). A notable difference between Equation (2.8) and the Engle (2000)'s GARCH specification is the removal of the lag of q_i because of the HAR feature. The persistence of q_i is not easily interpreted from one parameter. Interestingly, Equation (2.6), (2.7), and (2.8) together resemble the realized GARCH of a concurrent work by Hansen et al. (2012). Both papers attempt to utilize the notion of a "realized" measure of variance such as the realized variance, the bipower variation, the realized kernel of Barndorff-Nielsen et al. (2008) since these measures reveal more information than the simple squared return. The in-sample and out-of-sample comparison are conducted between a GARCH-EACD model of Engle (2000), a HAR-EACD, and the HAR-BACD.[9] The models are estimated by the Bayesian Markov Chain Monte Carlo (MCMC) simulation method, which gives a natural advantage to the authors to compare the efficiency between nested and non-nested models.

Liu and Maheu (2012) find that the most relevant investment horizon for the high-

[9]EACD stands for Exponential ACD. See Section 2.2 for a further discussion.

frequency transaction series is up to 1 hour and any long-term components ranging from above 1 hour to more than 1 day do not improve the model fitting. The HAR-BACD model has an astonishing performance against other specifications in fitting the data and density forecast. One extreme example is the stock of Sinopec, for which the HAR-BACD model is exp(3758) times better than the HAR-EACD at describing the volatility. The density forecast, which does not account for parameter uncertainty, consistently ranks the HAR-BACD at the top of the list. On the other hand, the improvement is only moderate in point forecast of the conditional variance q_i. The RMSE and MAE are only reduced in the fourth digit changing from the HAR-EACD to the HAR-BACD and the order of ranking reverses once for a Chinese real-estate stock. The negative sign of γ_2 shows that the interaction between duration and volatility tends to decrease volatility for heavily traded stocks due to noise traders who would continue trading regardless of new information. Without new information, volatility should barely move, or should decrease over time. The final empirical insight of the model lies in the influence of various volatility components VC_{i-1,h_m}. It turns out that the Chinese stocks have a longer memory than American stocks, with a memory length of up to 500 transactions. The phenomenon can be attributed to the organizational difference between the two markets: T+1 execution rule, no short sale, and daily return bounded by $\pm 10\%$ in the Chinese markets. Those restrictions largely limit investors' behavior in the very short term.

2.6 Links with Liquidity

The connection between volatility and liquidity can be traced back to Clark (1973) and the so-called "mixture of distribution hypothesis" (MDH). One important prediction in Harris (1987) is that the square of daily price change moves together with the daily trading volume as well as the number of transactions, because they are all proportional to the number of information events. We can easily see the reason that it bridges liquidity with volatility as the number of transactions or trading volume is one dimension of liquidity and the daily price change measures volatility. Liquidity, as defined in Section 2.1, has several dimensions. The bid-ask spread measures the width of liquidity, which represents the costs of a round-trip transaction. The volume of transactions refers to the depth of the liquidity. In view of this proportionality of liquidity measures and volatility, Lamoureux and Lastrapes (1990) argue that the strong persistence of daily return volatility is a manifestation of the daily autocorrelation of information events. In order to test the theory, they incorporate the daily trading volume as the proxy for the unobserved flow of information into the conditional variance equation of a GARCH(1,1) model. As the hypothesis predicts, the ARCH and GARCH parameter become mostly very small and insignificant when volume is included, suggesting that the conditional variance of returns is largely driven by the intensity of the information updates. The linkage between liquidity and volatility is not limited to the price-volume relationship (see Karpoff, 1987 for a review). Bollerslev and Domowitz (1993) include more market activity variables in the conditional variance equation, such as number of quote updates, bid-ask spreads, and duration between trades. They demonstrate that the lagged bid-ask spread has a strong and positive effect on current conditional variance, while the number of quote updates and duration plays little role in volatility.

Extensions of the earlier articles tend to give the variable of liquidity and volatility equal status. In the Blume et al. (1994)'s theoretical model, traders who learn information from past and current volume would trade more if precise information could be extracted from volume and *vice versa*. The current volume is a gauge of the quality and precision of information signals rather than representing the signal itself. Based on the original MDH, volume and price volatility are both driven by information flow and thus both are endogenous. Reflecting this argument, Foster (1995) builds a bivariate structural system determining volume and volatility simultaneously, contradicting Lamoureux and Lastrapes (1990) and providing strong support for the predictions in Harris (1987). Wang and Yau (2000) include the bid-ask spread in the structural VAR of Foster (1995), thereby formulating a trivariate system. They confirm the positive relationship between the bid-ask spread and volatility. In the light of the ACD model, Manganelli (2005) studies duration, volume, and volatility in a generalized VAR framework with an ACD, Autoregressive Conditional Volume (ACV) and GARCH model combined together, which allows more interdependence between the variables. He demonstrates strong empirical evidence of trading volume clustering regardless of trading frequencies and shows that intensive trading leads to greater volatility only for heavily traded stocks. In addition, volume itself is not influenced by lagged volatility or lagged duration in the model.

An important application of the MEM framework belongs to Engle et al. (2012), who discover the variation of the limit order book. Similar to earlier studies, a vector of variables is studied. The interaction between quoted market depth, price volatility, and depth volatility allows them to interpret various effects in the model, e.g. announcement effects, flights-to-safety effects, etc. In a baseline model where only the three variables are included, a two-way feedback between volatility and liquidity emerges at the top of the order book for one security: the market depth is reduced by high price volatility and high depth volatility and the lower depth predicts a future higher volatility, which is consistent with the theoretical work of Cespa and Foucault (2014) on multiple assets. Furthermore, a news impact curve can be described by allowing different effects of positive and negative changes. The authors find no asymmetry in price volatility in the presence of lagged depth and lagged depth volatility, which is close to the very small correlation suggested by the two-factor stochastic volatility model of Bollerslev et al. (2006). Last but not least, the persistence of depth is enhanced whereas the news impact is diminished on flights-to-safety days. Interestingly, news shocks become more important for price volatility relative to past volatility because of flights-to-safety in the US Treasury market, which is similar to the volatility dynamics when the ECB initiated bond purchasing, as examined in later chapters.

2.7 Concluding Remarks

Modeling high-frequency volatility plays an important role in understanding the lower-frequency process. Any monthly or quarterly process is an aggregation of high-frequency data. We have seen from the realized variance literature and intraday volatility models that intraday volatility contains much more relevant information for long-term returns. In addition, intraday volatility has its own merits in reflecting the fundamentals of the economy and the periodicity of investors' behavior. Furthermore, price volatility has a wide link with other liquidity measures in market microstructure studies. The intimate connection between high-frequency trading and intraday volatility is also an interest of regulators (see, Brogaard et al., 2014).

In this survey, we analyze various issues related to intraday volatility models, ranging from the basic data sampling approach to model specifications. Among possible data sampling schemes, we confirm that calendar time sampling with a 10-minute sampling frequency may be most suitable for the MTS market. We identify the value of realized variance papers to our studies and justify the specification of a multi-factor model. We comment on the realized covariance properties, including the adjustment for nonsynchronous trading, which is also seen in the parametric modeling of covariance matrices. The stylized facts of intraday volatility are described and it is shown that various ways of modeling intraday periodicity have their own advantages. There is no consensus on how to achieve the optimal results, which leaves a space for new exploration. Three classic intraday models are reviewed in depth as corresponding efforts to accommodate the stylized facts. Among the three models, the Engle and Sokalska (2012)'s method is the simplest yet still powerful in explaining intraday volatility. In the final section, we follow the development of the MDH as a route to outline the association between volatility and liquidity. Of course, high-frequency trading literature also brings the two strands of research together, but identifying high-frequency trading is beyond the scope of this survey.

Chapter 3

Modeling Intraday Volatility in European Bond Markets: A Data Filtering Application

3.1 Introduction

With the onset of the sovereign debt crisis raging through Europe, government bond volatility becomes a greater concern to researchers, regulators and practitioners. The study of interest rate volatility which is important for bond volatility dates back to the earlier studies of affine models. Longstaff and Schwartz (1992) are among the first to suggest yield change volatility is an important factor in explaining the term structure of interest rates. The roles and features of bond market volatility have been explored in numerous papers. Blume et al. (1991) investigate volatility risk of junk bonds relative to long term government securities. Jones et al. (1998) examine macroeconomic news effect on daily volatility and find different responses to a broad range of news using a GARCH(1,1) model (Bollerslev, 1986). De Goeij and Marquering (2004) estimate a multivariate model for bond and stock conditional variance using weekly data. Christiansen (2007) uses a GARCH model to study European bond markets before and after the introduction of the Euro and observes a substantial volatility spill-over effect from the aggregate European bond market to national markets.

High-frequency volatility remains less studied in contrast to the vast literature on daily and weekly volatility models (see Bollerslev et al., 1992; Poon and Granger, 2003). Taylor and Xu (1997) build a general ARCH model using hourly option returns and subsequently compare the information content of conditional variance, realized variance and implied volatility. Fleming and Lopez (1999) estimate a multivariate GARCH model on hourly returns for the US Treasury bond interdealer market. Bollerslev et al. (2000) adopt the flexible Fourier form (FFF) to model intraday seasonality and explicitly account for the macroeconomic news impact on 5-minute US Treasury bond futures volatility. They find long-memory effects and estimate an MA(1)-FIGARCH(1,d,1) model (Baillie et al., 1996) to forecast the daily variance. Deo et al. (2006) propose a long-memory stochastic volatility model and evaluate its forecasting performance against the component GARCH and ARFIMA (1,d,0) models. They introduce a gradually changing seasonal pattern to improve the forecasting performance of the model.

European markets have witnessed a dramatic increase in bond volatility since 2009. Early studies are mainly concerned with macro news and with the fundamental drivers of the bond market. During the crisis period, European sovereign debt markets exhibit a much higher volatility and are strongly influenced by the European Central Bank (ECB) intervention (see Eser and Schwaab, 2016 and Ghysels et al., 2014). Hence the importance of studying the bond return volatility associated with intense market activity during the times of economic uncertainty has greatly increased. The three fundamental questions we want to address in this study are: How can we accurately quantify the short-term fluctuations in bond returns? How can we properly filter out the temporary effects of liquidity dynamics on volatility models? Is intraday volatility important for predicting future daily volatility? We think our study helps portfolio managers and traders who want to quantify bond volatility and control for intraday bond risk. Giot (2005) applies the GARCH(1,1) model and the EGARCH(1,1) model (Nelson, 1991) to compute intraday Value-at-Risk (VaR) using 15-minute and 30-minute returns for New York Stock Exchange (NYSE) data. The study shows that intraday VaR exhibits similar features to their daily counterpart once intraday seasonality is taken into account. Almgren and Chriss (2001) develop the best execution strategy and efficient frontier concerning liquidation cost and volatility risk in a high-frequency trading environment. Engle and Ferstenberg (2007) view the problem of executing a portfolio transaction as a trade-off between the speed of trading and achieving a better price. And the variance of transaction cost also plays an important role in devising the optimized trading strategy.

We adopt the framework of Engle and Sokalska (2012) and introduce new specifications for each component of their multiplicative GARCH model. Engle and Sokalska (2012) focus on the forecasting performance of a simple intraday GARCH(1,1) model estimated for a large universe of US stocks. The three components of 10-minute return volatility are estimated separately, in three steps. For the daily variance the estimation relies on a commercial multiple factor model and daily periodicity is quantified as the mean of intraday return volatility for the same subinterval of the trading day across all sample days. Different ways of pooling stocks are considered and cross-section pooling appears to possess superior forecasting ability in frequently traded stocks. Liquidity conditions seem to play an important role in the estimation and forecasting of the intraday volatility of less frequently traded stocks. Ghysels et al. (2014) apply the same model of Engle and Sokalska (2012) with additional dummies in the conditional mean and conditional variance equations to study the effect of the Securities Markets Programme (SMP) implemented by the ECB.

Building on the findings of Engle and Sokalska (2012) and Ghysels et al. (2014), we turn our attention to the development of a better volatility modeling approach which simultaneously addresses the problems of filtering transitory liquidity effects, modeling intraday periodicity and estimating fundamental intraday volatility. We first choose to model the intraday periodicity as a piecewise linear structure in the spirit of the Spline-GARCH (Engle and Rangel, 2008) model. The daily volatility dynamics are captured by a simple GARCH(1,1) model. Second, our findings further improve our understanding of the European bond market during the sovereign debt crisis when the debt of distressed countries is no longer a safe asset, with serious repercussions for the whole economic environment. We study the volatility of benchmark, 10-year bonds for seven Euro area countries. With our sample, the dynamics of liquidity are paramount for understanding the short-term volatility of quoted prices and this poses a challenge to computing fundamental volatility. As is well known in the literature, high-frequency data often contain various errors and noise due to frictions and liquidity imbalances (Fleming, 2001, Bandi and Russell, 2008), which make proper data cleaning both necessary and challenging. Obviously, the data cleaning/sampling process will affect the computation of fundamental volatility (see Bandi and Russell, 2008). It is thus important to jointly address the data filtering and the volatility estimation problems. None of the previous studies have evaluated the effects of their filters. We consider several alternative data cleaning techniques and develop a procedure for choosing the filter which provides the best estimates of fundamental volatility. Last but not least, intraday data contain information that is helpful in estimating volatility at longer horizons as many papers from the realized volatility literature suggest, e.g. Barndorff-Nielsen and Shephard (2002a) Barndorff-Nielsen and Shephard (2002b). we show some empirical evidence that intraday data can help improving the forecasts of daily volatility.

The rest of the paper is organized as follows. Section 3.2 introduces the motivations and properties of our econometric high-frequency model. Section 3.3 explains our method for cleaning our sample of bond data and for constructing the return series. Section 3.4 presents the estimation results and interpretations. Section 3.5 carries out the forecasting comparison between the intraday GARCH and the daily GARCH(1,1). Finally Section 3.6 summarizes our findings.

3.2 A Multiplicative Error Model of Intraday Volatility

We denote the intraday log return by $r_{t,n}$ and the daily return by r_t. t represents the daily index ($t = 1, 2, \ldots, T$) and n is the intraday index ($n = 1, 2, \ldots N$). Each intraday time interval n is referred to as "bin" n. The log return $r_{t,n}$ is calculated as the difference in log mid-quote prices in a limit order market with designated market makers.

The multiplicative error model introduced by Engle (2002b) and adopted by Engle and Sokalska (2012) suggests that

$$r_{t,n} = \sqrt{h_t s_n q_{t,n}} \varepsilon_{t,n} \text{ and } \varepsilon_{t,n} | \Phi_{t,n-1} \sim N(0,1) \tag{3.1}$$

where h_t is the daily variance component

s_n is the intraday periodicity or diurnal component

$q_{t,n}$ is the intraday variance component with $E(q_{t,n}) = 1$

$\varepsilon_{t,n}$ is an error term

$\Phi_{t,n-1}$ denotes the set containing all the available information up to the bin preceding the current time interval. To avoid any confusion, we will refer in the subsequent analysis to the volatility of $r_{t,n}$ as intraday return volatility and $q_{t,n}$ as intraday volatility . Here we assume that the conditional distribution of the error term is standard normal, but this does not imply a normal distribution of returns. The overnight return $r_{t,0}$ is not specified here because with the diurnal component we are trying to model and explain the intraday volatility of fixed-interval returns and the overnight return is captured by the daily component.

3.2.1 Daily Model

Andersen and Bollerslev (1998) find a close relationship between the cumulative absolute intraday return and the MA(1)-GARCH(1,1) one-step-ahead volatility forecast in a one-year sample. The daily conditional variance forecast, which is not affected by short-term intraday volatility dynamics, represents a certain amount of *anticipated* intraday return variation. Failing to capture this lower-frequency component would distort the overall volatility computation. Hansen and Lunde (2005) confirms the superior predictive ability of the GARCH(1,1) model against more than 300 specifications for daily conditional variance of foreign exchange rates. As the forex market has a very similar structure to the sovereign bond market we study, we choose the GARCH(1,1) model as our forecast model for daily conditional variance.

During the crisis, sovereign bond volatility was affected by the ECB's actions through a series of interventions. The SMP was announced on May 10, 2010 along with several Longer-Term Refinancing Operations (LTRO) measures[10] to alleviate the heightened market tension. The programme was described as " *interventions in the euro area public and private debt securities markets to ensure depth and liquidity in those segments which are dysfunctional*".[11] With the first SMP the ECB purchased the government bonds of Greece, Ireland and Portugal and a second SMP was implemented to buy Italian and Spanish government bonds. The second SMP was announced on August 7 2011[12] and on the following day, the price of the 10-year Italian bond jumped by €5.7 to €96.32.[13] In a press release on February 21, 2013, the ECB disclosed the total amount of bonds acquired under the SMP and Italian and Spanish bonds accounted for two-thirds of those purchases.[14]

Ghysels et al. (2014) estimate a daily GARCH(1,1) model with a dummy accounting for SMP interventions to evaluate the success of the SMP. We adopt the same approach to control for the SMP effects when estimating volatility during the SMP window.[15]

Only the first two lags of returns are included in the conditional mean equations, as indicated by the t-test on the coefficients and by the Schwarz information criterion

[10]ECB provides liquidity to European commercial banks for holding illiquid assets via LTRO.
[11]See,`www.ecb.europa.eu/press/pr/date/2010/html/pr100510.en.html`
[12]See, `www.ecb.europa.eu/press/pr/date/2011/html/pr110807.en.html`
[13]See, `www.bloomberg.com/news/2011-08-08/`
[14]See,`www.ecb.europa.eu/press/pr/date/2013/html/pr130221_1.en.html`
[15]Since we only have weekly data for the SMP, we assume the purchase is achieved through the whole week.

(BIC). The daily model is estimated via maximum likelihood. Specifically,

$$r_t = c_1 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \nu_t \quad \nu_t|\mathcal{F}_{t-1} \sim N(0, h_t) \tag{3.2}$$

$$h_t = w + a_1 \nu_{t-1}^2 + b_1 h_{t-1} \tag{3.3}$$

For Italian and Spanish bonds our daily GARCH becomes

$$r_t = c_1 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \sum_{p=1}^{4} d_p * dummy_p + \nu_t \tag{3.4}$$

$$h_t = w + (a_1 + a_2 * I(SMP_{t-1} > 0))\nu_{t-1}^2 + (b_1 + b_2 * I(SMP_{t-1} > 0))h_{t-1} \tag{3.5}$$

$$I(SMP_{t-1}) = \begin{cases} 1 & \text{if purchase amount>0 at t-1} \\ 0 & \text{amount=0} \end{cases}$$

The dummy $I(SMP_{t-1})$ controls for the regime shift associated with the purchasing of Italian and Spanish bonds by the ECB from August 08, 2011 to March 09, 2012 during the second round of the SMP. Four dummies are used to control for specific news corresponding to four dates with large daily returns caused by institutional announcements.[16] $dummy_1$ controls for the big drop in returns on May 06, 2010 when the ECB maintained its base rate unchanged with no action with respect to the Greek debt crisis.[17] $dummy_2$ and $dummy_3$ capture the two jumps in bond prices due to the activation of the SMP (see above). $dummy_4$ controls for the return of December 05, 2011 when former Italian Prime Minister Monti announced budget cut plans and all

[16] Controlling for one-time event with dummy variables is a common approach in volatility analysis, e.g. Andersen and Bollerslev (1998)

[17] See, `www.ecb.europa.eu/press/pr/date/2010/html/pr100506.en.html`

financial markets witnessed a big rally.

3.2.2 Intraday Seasonal Pattern and Volatility

Daily variance stays constant through intraday activities but the innovation in bond returns changes over time. We account for this periodicity using a piecewise linear structure while modeling intraday conditional variance using a unit GARCH model (i.e. the unconditional variance is 1). Our intraday model is implemented as follows:

$$s_n = \delta_0 * exp(\sum_{j=1}^{m} \delta_j (\Delta_n - k_{j-1})_+) \tag{3.6}$$

$$q_{t,n} = 1 - \alpha - \beta + \alpha \left(\frac{(r_{t,n-1})^2}{s_{n-1} h_t} \right) + \beta q_{t,n-1} \tag{3.7}$$

$$(\Delta_n - k_j)_+ = \begin{cases} (\Delta_n - k_j) & \text{if } \Delta_n > k_j \\ 0 & \text{otherwise} \end{cases}$$

$$\Delta_n = \frac{n}{N} \quad n = 1, 2, \ldots, N.$$

The specification has the advantage of estimating the intraday volatility and the diurnal component jointly and eliminates the need for a two-step estimation. In the original framework of Engle and Sokalska (2012), intraday seasonality is estimated with a simple average of returns for every interval of the trading day in a separate step. It can be shown that the statistical properties of a two-step estimator can be derived from the Generalized Method of Moment (GMM) by Newey and McFadden (1994). But there is an efficiency loss in the parameter estimation of the second step. Further, the linear spline equation has reduced the number of parameters substantially as compared to the original model. On the other hand, while Engle and Sokalska (2012) utilize a commercial forecast of daily volatility, we need to make a one-step-ahead forecast of daily conditional variance first. The consistency of the estimators in Equation (3.6) and (3.7) still holds according to the argument in the appendix to Andersen and Bollerslev (1998) while the possible autocorrelations and heteroskedasticity caused by including the daily GARCH volatility forecast should be adjusted. The autocorrelations can be alleviated by a longer sampling interval and heteroskedasticity is naturally controlled by the unit GARCH.

The exponential form guarantees the positivity of the diurnal component. k_j ($j = 1, 2, \ldots, m$) denotes a knot in the linear spline. The knots are set respectively at 9:00, 10:00, 11:00, 12:00, 13:00, 14:00, 15:00, 16:00, 17:00 and 17:30 (official closing time) for Belgium, Germany, Italy and Spain. Three nodes at 11:00, 12:00 and 13:00 are omitted for Austria, France and the Netherlands because the estimation of the exponential spline makes the optimization algorithm difficult to converge for these three countries' data. As can be seen in Figure 3.6 in Section 3.4, volatility stays low in the middle of the day for all major European countries. So we choose to remove the knots during 11:00–13:00 when the return does not vary significantly. It turns out that the diurnal patterns of the intraday return volatility are indeed not greatly affected by the omission for these three countries compared to those of other countries in Figure 3.6.[18] The spline we use is different from Engle and Rangel (2008) in terms of functional form and purpose. Their quadratic spline coupled with exogenous variables aims to incorporate the low-frequency volatility related to the macroeconomic environment. While our

[18]The starting value can be guessed by estimating a piecewise linear regression of the return divided by the daily conditional volatility forecast as a preliminary analysis.

linear spline has the same frequency as the intraday volatility and we assume it is not affected by exogenous variables. Obviously, this could easily be extended to allow exogenous factors to affect the diurnal pattern, for example if we wanted to distinguish information days, with relatively higher trading intensity, from normal days. Notice that $E(q_{t,n}) = 1$ implies that the unconditional variance of the stochastic component is one. Hence, the unconditional variance of high-frequency return is entirely dependent on the unconditional daily variance and the diurnal component, i.e.

$$E[(r_{t,n})^2] = s_n E(h_t) \tag{3.8}$$

Suppose that $E(h_t)$ is fixed by the GARCH(1,1) model, then the conditional volatility of the intraday returns will eventually converge to the diurnal component, which is a function of time. The deterministic pattern is exemplified by the squared return correlogram. The recurring cycle of intraday volatility appears in contrast with a geometric decay implied by the ARCH/GARCH model. Ignoring such patterns can produce some random results when the intraday GARCH model is applied (see Andersen and Bollerslev, 1997).

Many factors can induce a repeated pattern of intraday volatility. Ederington and Lee (1993) associate the spikes in volatility of fixed-income futures in the morning with several macro announcements. They also suggest that the speed of processing information, which is manifested in the decline of volatility, can serve as a test of market efficiency. Bollerslev et al. (2000) specifically explain the macro news effects using a dummy variable approach and confirm the finding by Ederington and Lee (1993) who show that the spike in the volatility of US bond futures is related to macro news. Furthermore, a periodic pattern represented by a Fourier series is still found to be significant in explaining the return variation in Bollerslev et al. (2000). Fleming and Lopez (1999) and Christiansen (2007) both find a volatility spill-over effect from US Treasury market to European trading centers. The US market opens at 14:00 Central European Time (CET) and may induce a prolonged period of increased volatility towards the end of the European trading day.

3.3 Data and Cleaning Procedures

Our high-frequency data contain 10-minute log returns constructed from the quote midpoints for ten-year benchmark government bonds from the MTS interdealer market. The intraday data runs from April 02, 2012 to December 30, 2013. We rely on a longer time series of daily data from 02 January, 2009 through December 30, 2013 to estimate the daily volatility component.[19]

3.3.1 Institutional Details

MTS is an electronic trading platform where unique counterparties trade various fixed-income securities including European government bonds, quasi-government bonds, corporate bonds and repurchase agreements. Here we describe the market features that are most relevant for our analysis. Detailed information on the MTS market structure is given in Darbha and Dufour (2013). There are two parallel platforms for benchmark bonds: the MTS domestic markets devoted to trading domestic bonds and the Euro MTS market where all benchmark securities across countries can be traded. Each platform has its own features in terms of trading rules, market participants, and market

[19]In the overlapping period of intraday and daily data, the daily volatility is computed as a one-step-ahead forecast

makers. The database has information on all changes in the best three quotes either in the ask side or in the bid side of the order book. Quote changes are due to either changes in the quote prices or in the quote sizes. Price discrepancies for the same bond due to the parallel trading structure can be eliminated by traders with access to both markets. Cheung et al. (2005) find that the liquidity conditions on domestic markets are very similar to those observed on the Euro MTS. Market makers are obliged to post two-way quotes called ”proposals” for the securities which are assigned to them by MTS. The limit orders they submit must satisfy a series of conditions including a minimum volume varying from €2.5 to €10 million, and a minimum tick value. MTS has made several modifications to their dealing obligations in order to introduce more liquidity since the beginning of the 2007 financial crisis. Before 2007, there were requirements for minimum quoting hours, and maximum spread during a trading day for market dealers. MTS now instead tracks the average duration of quoting and the average spreads pertaining to an individual market maker and makes sure that the averages are consistent with the market averages derived from all market makers. Price takers were only given permission to submit market orders against the best available quote before November, 2012. A single-sided limit order (either buy or sell) can be entered into the system by price takers since November, 2012. Trading is possible from 8:15 to 17:30 CET.

3.3.2 Variable Construction

We focus on major Euro-zone countries including Austria, Belgium, France, Germany, Italy, the Netherlands and Spain, which have benchmark ten-year bonds during the sample period. Since we concentrate on one maturity category, we choose on-the-run

10-year bonds defined as long-term bonds with a remaining time to maturity ranging from 8.5 years to 11.5 years. The lower bound for the selection is in accordance with the usual minimum remaining time to maturity for a bond to be qualified in a long-term bond futures contract (see the Eurex Exchange Long Term Bond Futures Contract). The upper bound is determined to have the same distance to 10 year as the lower bound. We select only one on-the run bond for each period of each country. Beber et al. (2009) have a tighter band of maturity (9.5–10.5 years) for 10-year bonds, as they want to study the relationship between credit default swaps (CDS) and sovereign yield spreads during a crisis. The CDS contracts are explicitly written on the breadth of bonds. Dunne et al. (2007) define long-term bonds with maturity of 6.6–13.5 years, which is broad enough to examine the benchmark status. We adjust the range of maturities according to the specific situation of a crisis during which European countries have a strikingly different issuing frequencies. For example, Germany has auctioned in total 10 bonds while Austria did not issue any new 10-year bonds from 2006 to 2011. Nevertheless, some bonds that were originally issued as 15-year bonds could be viewed as being in 10-year category by our definition.

The main concern when constructing a return series of bond data is to maintain constant maturity and a sufficient level of liquidity so that the mid-quote price is a good proxy for the underlying price. With the passage of time and new issues, the current benchmark bond loses its status. In order to have a broad view of the crisis period and maintain the quality of the study object, we have to change our benchmark bond whenever the existing benchmark bond does not comply with our maturity standard or there is a new auction. The rolling-over approach is a common solution for the periodic issues and changes in seasonality of benchmark bonds (see Fleming and Lopez, 1999 for

GOVPX data and Bollerslev et al., 2000 for US long-term bond futures data). On each switching date, the return is computed from the prices of the old bond and the returns are always computed using data from the same bond. We choose different policies to deal with switching bonds for liquidity or maturity reasons. If the maturity of the current benchmark bond falls below 8.5 years, the switching is triggered immediately.[20] However, if there is a new auction, we choose to delay the introduction of the new bond and the exclusion of the old bond by one month. According to Pasquariello and Vega (2009), there is a significant liquidity and price heterogeneity of newly issued benchmark bonds and the just off-the-run bonds across maturities in US market. They demonstrate that for 10-year US bonds, the liquidity condition of the on-the-run bonds is improved after 10 days since the auction. Diaz et al. (2006) also find that the liquidity measured by relative traded volume is different between off/on-the-run 10-year Spanish government bonds. The authors illustrate that an on-the-run bond does not instantly gain benchmark status. We therefore do not replace old bonds with new bonds immediately.

3.3.3 Data Preparation and Filter Evaluation

We follow a series of steps to prepare the intraday dataset from the start of 2009 to the end of 2013 for our analysis.[21] Firstly, we remove the quotes recorded outside the trading hours.[22] Following Fleming (2001), all quotes on October 22, 2009 are excluded from our dataset because the last quote update on that day was recorded at 15:26 and

[20]If the designated switching date is a market holiday, the switching will be postponed to the nearest market open date. If two bonds both qualify for this category we choose the most recent one.

[21]The cleaning covers all the daily and intraday sample as it helps us to estimate both models more accurately.

[22]Some pre-market quotes and post-market settlements are stored in the data set.

there were multiple transactions happening at different prices afterwards. Secondly, we compute the global best bid and offer prices across the two platforms for each country. Due to the parallel status of domestic MTS and the Euro MTS, quotes are often updated simultaneously on both platforms with recorded time stamps differing by a few milliseconds.[23] The adjustment is made for the delay and the overall best available quotes are computed from the simultaneous ticks. We also remove any quotes with a negative spread, only keeping the change in the best bid and ask prices. Thirdly, we apply a range of filters to remove temporary illiquidity effects and choose the best filter for each country. The procedure for the selection of the best filter is explained below. Finally, the longer daily sample and 10-minute sample are generated from the prepared data. The daily return is calculated as 100 times the log difference of 5 PM quote midpoints extracted from intraday data. The use of mid-point of quotes is discussed in Hasbrouck (1991) and can alleviate the temporary autocorrelation introduced by any bid-ask bounce. The reason we select the 5 PM mid-quote price instead of the closing one at 5:30 PM is that the quoting activity is less intensive for some days towards the end of the trading day. The final quote updates sometimes appear considerably earlier than 5:30 PM and thus the closing prices are often stale. The 10-minute returns from 8:15–9:00 (not included) are dropped from our dataset as the first quote is delayed on a few occasions.

With the increasing frequency of financial data, numerous errors are present and hard to clean. The problem is that market makers are obliged to keep their quotes on the system even when they have satisfied their quoting obligations. At times, this results in very large spreads which simply indicate to the market that dealers have temporarily withdrawn their competitive quotes. No rational traders would trade at these quotes.

[23] The delay varies from 1 millisecond to 995 milliseconds.

Table 3.1: Data preparation

Raw data are processed as the following. The quote updates recorded outside the trading hours (8:15-17:30 CET) are deleted. All ticks on 22 Oct 2009 of all countries are excluded because multiple transactions were recorded at various prices after the last quote was recorded. Simultaneous ticks due to parallel tradings are identified and adjusted. The best available bid and ask are selected from them. The observations with negative spreads are also dropped. Because ticks are recorded whenever there is a improvement to a level of order book, we only keep the change of best available bid and ask.

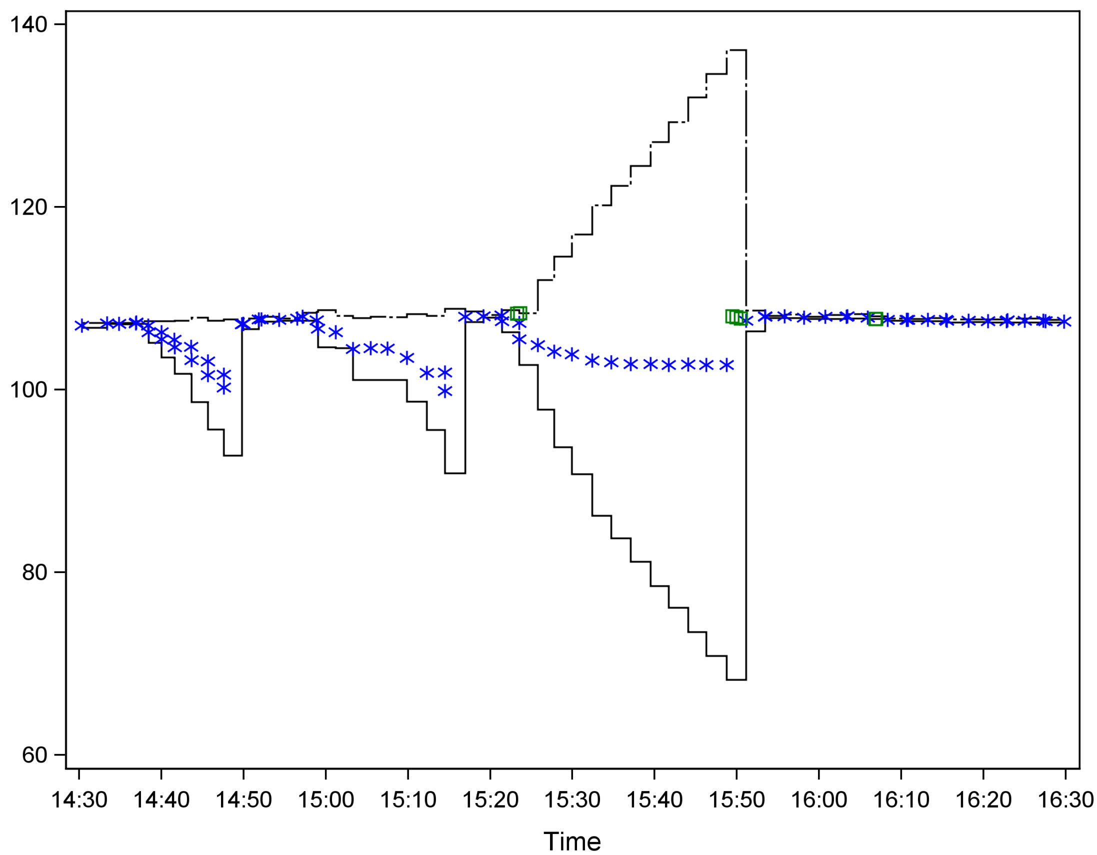
Operation	No. of obs.	Percentages(%) of the raw sample
Number of raw observations	13772614	100.0
Ticks outside trading hours	184407	1.3389
Ticks on 22 Oct 2009	9467	0.0687
Simultaneous ticks	4241931	30.7998
Negative spread	2194	0.0159
unchanged bid and ask prices	3705491	26.9048
Processed sample size	5629291	40.8731

Possible causes include macro news announcements, unscheduled ECB interventions on debt markets, human errors, and holiday effects. (Fleming, 2001). Attention must be paid to distinguishing transitory volatility due to illiquidity effects. A 3% jump in log returns is plausible if some macro news is released. A large jump would be suspicious in the absence of any observable information, especially when liquidity is scarce. Filtering is a way of categorizing abnormal outliers as errors. The temporary volatility caused by illiquidity is best illustrated by Figure 3.1.

Transactions are unlikely to occur when liquidity evaporates and the quoted price may be extreme. The so-called *stub quotes* defined by the literature[24] are exemplified

[24]See, Kirilenko et al. (2016)

Figure 3.1: Plot of best quotes for a 10-year benchmark French government bond after the processing in Section 3.3.3 (ISIN code: FR0011196856) on June 01, 2012 from 14:30:00 to 16:30:00. Tick-by-tick mid-quote prices (stars), transaction prices (square), best available bid prices (solid line), best available ask prices (dashed line).



in two ways. In Figure 3.1 we show that the dynamics of bid, ask and mid-quote prices for the 10 year French government bond on June 01, 2012. From 14:38:28 to 15:16:57, the bid price gradually moved away from the offer and reached its bottom around 90 twice while the offer price barely changed. From 15:21:24 onwards until 15:51:08, the two sides of the order book deviated from each other. In all cases, the mid-quote is affected by the temporary illiquidity and volatility is artificially boosted up due to the sudden adjustment in the bid and ask at the end of the two periods.

A preliminary analysis which involves sampling data from every 2 seconds to every 2 hours[25] for this particular day reveals that the daily realized variance can reach as high as 200! Even when we reduce the sampling frequency to every 50 minutes, the realized variance is still above 50. As another experiment, there are 104 2% 10-minute returns after 9:00 from the processed yet unfiltered intraday dataset and the average change in percentage spread is over negative 180 basis points! This means that most of the big jumps in 10-minute returns can be tied to the sudden decrease in the bid-ask spread as shown in Figure 3.1.[26] Obviously, using intraday data without filtering is not an appropriate way to study volatility. The asymmetric dynamics in bid and ask quotes mask the true volatility process, which is also observed in other markets (see for example Hasbrouck (2012) and Engle and Russell (1998) for the stock market). Interestingly, there is often a gradual deterioration in liquidity on one side of the market which is then promptly recovered in MTS markets, which is the opposite to the trade impacts observed and modeled by Hasbrouck (1991) in US stock markets. It is not plausible to infer fundamental values from these mid-quotes.

To establish the benchmark when evaluating various cleaning procedures, we resort to the study of Bandi and Russell (2008). They prove that the microstructure noise, which causes transitory volatility, heavily influences the estimation of the fundamental volatility. The optimal sampling frequency should minimize the mean squared error (MSE) of realized variance against the true variance under the MA(1) assumption of tick-by-tick returns. Specifically Bandi and Russell (2008) decompose the MSE into components of true integrated variance, the first four moments of noise, sampling frequency and the true integrated quarticity conditioning on the volatility path. The true

[25] The sampling interval is incremented by 1 second every time the daily realized variance is computed

[26] We delete the obvious errors when we do the experiment, e.g. the ask price is deleted when it is over 500 Euros.

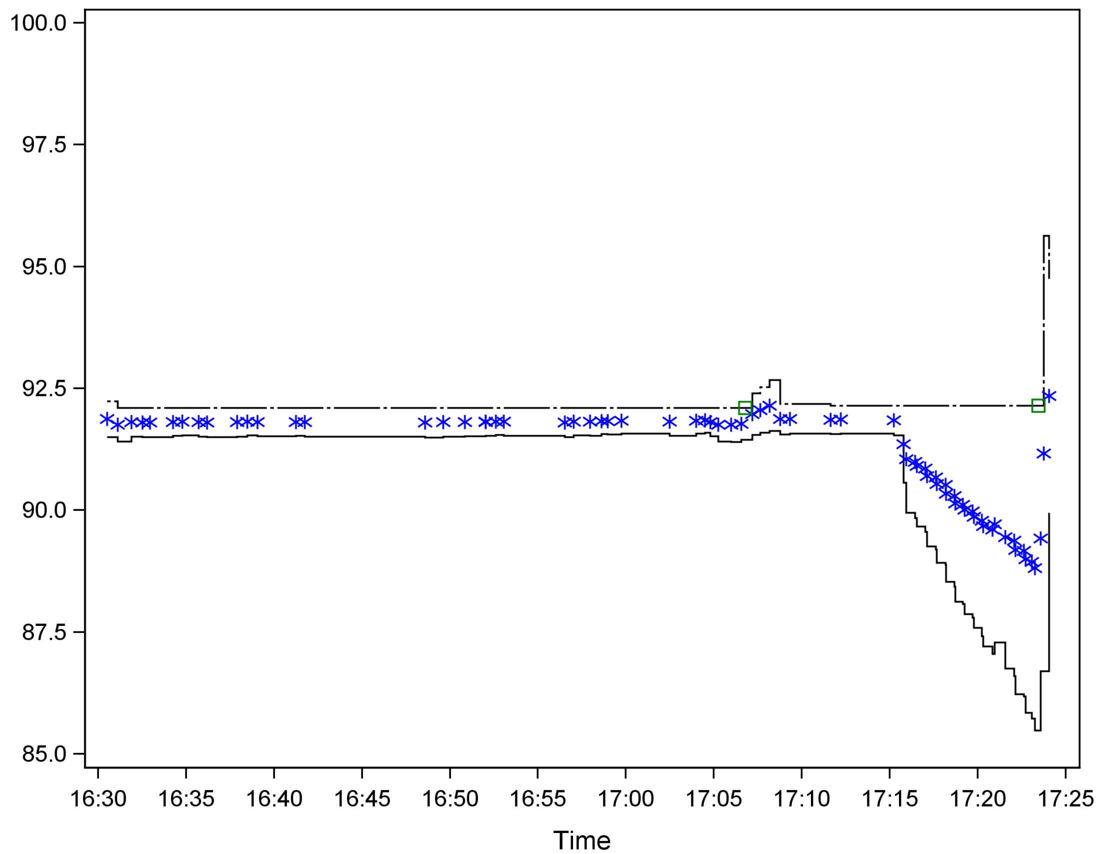
daily integrated variance is approximated by the realized variance of 15-minute squared returns. In the spirit of Bandi and Russell (2008) and in the interest of studying volatility, we try to find the best filtering procedure to minimize the effect of temporal noise on the modeling and computation of the bond volatility. We utilize the concept of the MSE and try to minimize the average difference between the daily summation of conditional variance of 10-minute return and realized variance derived from 2-hour returns. Specifically, let V denote the true daily integrated variance. The MSE $E(\sum_{n=1}^{N} h_t s_n q_{t,n} - V_t)^2$ is estimated by $\frac{1}{T}\sum_{t=1}^{T}(\sum_{n=1}^{N} \hat{h_t}\hat{s_n}\hat{q_{t,n}} - \hat{V_t})^2$ and the best filter should minimize this criterion. Our benchmark realized variance is a model-free measure of fundamental volatility. As it is seen in Figure 3.1, the 2-hour sampling interval is conservative enough to avoid including liquidity effects in the realized variance. In order to operationalize the benchmark, we assume that the returns of daily and intraday frequency follow a GARCH process as compared to the MA(1) structure of stock returns in Bandi and Russell (2008). It should be emphasized that the benchmark realized variance is computed from unfiltered data.

3.3.4 Choosing Filters

We now turn to describing the filtering methods. The methods can be categorized into three groups. The first group (A) belongs to the so called ”tradable spread” approach, as it involves both trades and quotes. The second group (B) is to compute quantiles of quoted percentage spread due to the well-known robustness of the statistics to outliers. The third group (C) can be dubbed as ”local window” filters, as they only concern the local properties of observations. The one rule that we apply to all filters is that we replace any deleted observation with the most recent valid one approved by the filter. The rule provides us the same number of observations across filters in order to compare them fairly. For the first two groups, we do not discard any observations with percentage spreads less than 50 basis points regardless of the corresponding threshold.

A. The maximum ”tradable spread” approach matches trades with their immediately preceding quotes in order to determine the maximum percentage spread where a trade can happen. The percentage spreads associated with real transactions are all deemed to be tradable. The maximum of all tradable spreads can serve as a threshold below which percentage spreads are reasonable enough to induce trades. The percentage spread, which is computed as a bid-ask spread divided by the mid-quote price, facilitates the comparison of different filters across assets. Filtering based on bid-ask spreads seems a natural choice, given it is a measure of the liquidity and quality of the market and market data (Hasbrouck 1993). This approach brings trades and quotes together and relies on the economic meaning of percentage spread. Harris (2002) illustrates that the posted spread represents a measure of transaction cost, which traders tend to minimize by searching smaller spreads. For a venue with a high trading frequency, e.g. the bid-ask spread in NYSE is usually very tight because of the fierce competition among liquidity providers. Matching trades with quotes can be a way of identifying erroneous trades. Barndorff-Nielsen et al. (2009) remove transactions based on quotes (see page C8, entry T4). Unlike the stock market, executions of bonds are fragmented and distributed across assets, as exemplified by Figure 3.1 in MTS markets. Nevertheless, the average daily volume of bond transaction on the MTS market was €8.7 billion in June 2012, which was much larger than that of the London Stock Exchange (LSE) during the same time (Darbha and Dufour, 2013). Hence using additional information of MTS trades could be appropriate for filtering.

Figure 3.2: Plot of best quotes for the 10-year benchmark Spanish government bond (ISIN code: ES00000123B9) on November 25, 2011 from 16:30:00 to 17:25:00. Tick-by-tick mid-quote prices (stars), transaction prices (square), best available bid prices (solid line), best available ask prices (dashed line).



On the other hand, there are some drawbacks of applying this method. Notice that there were two trades executed inside the bid and ask prices in Figure 3.1 and it is impossible to determine the bid-ask spreads for the trades. This demonstrates that the maxima can only be derived from matched trades which may lead to loss of information when filtering. Furthermore, a single trade is sometimes executed when the spread is large. Notice that the last execution in Figure 3.2

was buyer-initiated when the bid price was decreasing towards the bottom level. These abnormal records cast some doubt on the reliability of the maximum ”tradable spread”. It prompts us to look at alternative statistics such as percentiles. The 99th percentile of all traded percentage spreads may potentially give a more reliable estimate of the threshold within which trades will probably be executed.

B. The percentile filtering approach can be extended to be directly applied to all percentage spreads. However, percentage spreads are not stable during the sovereign debt crisis. According to Darbha and Dufour (2013), the spreads of European government bonds have gradually been increasing over the past few years. When defining the threshold for removing extreme percentage spreads, a successful filter needs to reflect the development of the liquidity condition. Specifically, we first compute the 90-99th percentile of the empirical distribution of percentage spreads belonging to one bond. We then remove the quotes with a percentage spread larger than the percentile. To accommodate the evolution in the liquidity condition, the computation and filtering are done each month. Arguably this approach is simple but *ad hoc*. Dropping any predetermined amount of data is purely mechanical and has no economic significance. Additionally, it is unlikely that any particular percentile uniformly outperforms the others for all countries. However, this approach still targets the liquidity measure and the benchmark devised in Section 3.3.3 is applicable. Due to the limited space we cannot present the detailed figure of each percentile each month but the patterns of percentiles would be the same. Figure 3.3 shows the 95th percentile by countries and the number of outliers can be inferred from this figure. We can see that liquidity is very volatile during the sample period. Germany has Treasury markets with the lowest spreads: roughly 95% of the spreads are below 50 basis points. For other countries, the outliers of the percentage spread are present in many months. Surprisingly, even French bonds have nearly 5% of their quoted percentage spread well above 100 basis points in late 2011. Austria, Belgium and Italy all have large spreads for a considerable time from 2009 to 2012. Spanish bonds experience the worst liquidity during December 2011, when 5% of the data possess percentage spreads more than 2000 basis points.

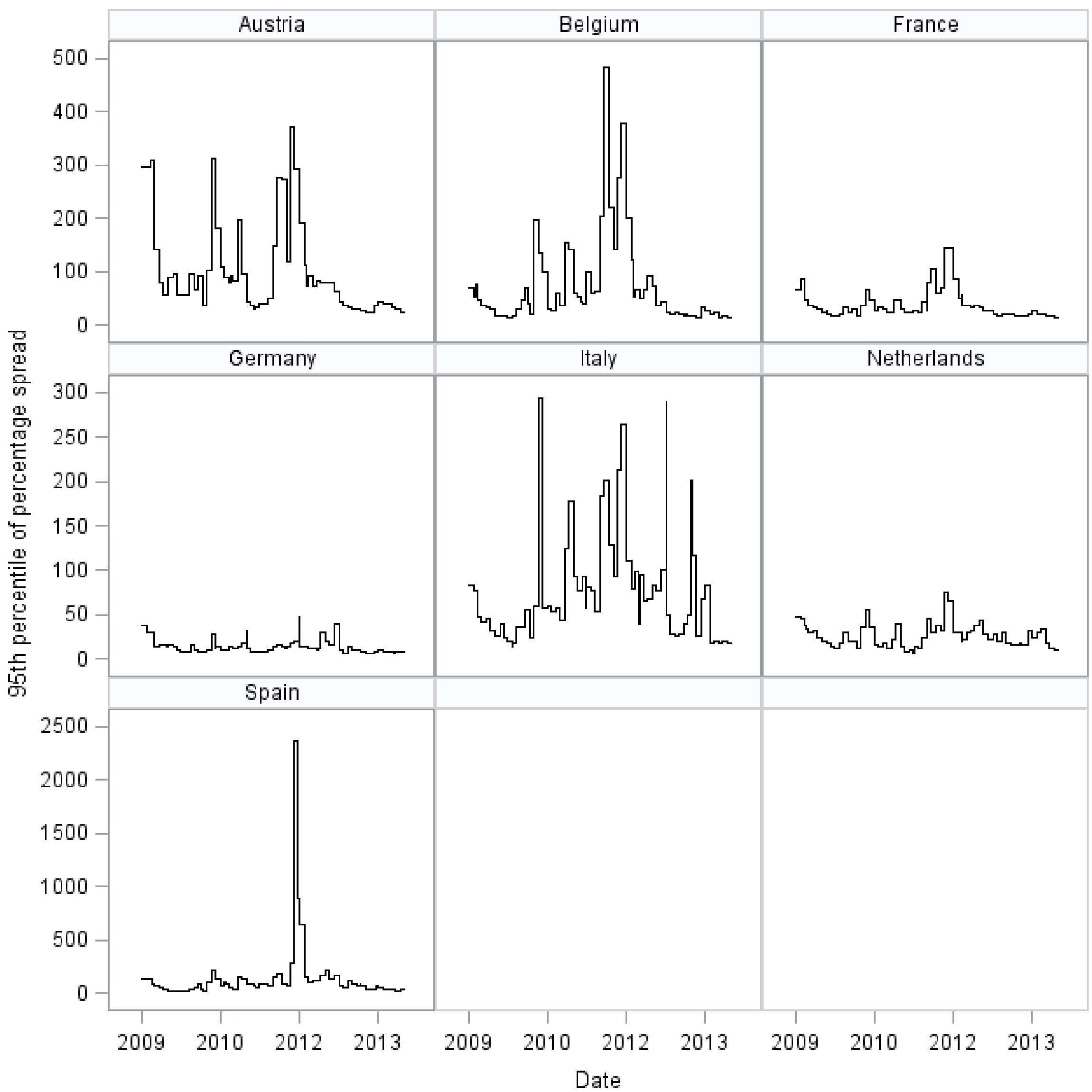
C. The third approach is to select price series. Some of the bid prices or offer prices deviate substantially from the quotes around them. Gençay et al. (2001a) propose a technique of detecting outliers, called "adaptive filtering". They suggest that a filter should learn from the series and develop its standard with a consideration of local properties. The same idea is also applied in two other papers, namely Brownlees and Gallo (2006) and Barndorff-Nielsen et al. (2009). Brownlees and Gallo (2006) devise a filter based on changes in transaction prices. The filter examines a local window of k trades near the current trade and computes the mean and variance of those trades after trimming the 10% tail values. Instead of cleaning transaction price, we apply Brownlees and Gallo (2006)'s core method to mid-quote price p_n. That is

$$(|p_n - \bar{p}_{-n}(k)| < 3\sigma_{-n}(k) + \gamma) = \begin{cases} True & \text{observation } n \text{ is kept} \\ False & \text{observation } n \text{ is removed} \end{cases}$$

where $\bar{p}_{-n}(k)$ and $\sigma_{-n}(k)$ are, respectively, the δ-trimmed mean and standard deviation of a length of k quotes around the current quote. The $-n$ subscript

Figure 3.3: 95th percentile of percentage spread by countries

The percentile is drawn from the empirical distribution of the percentage spread. The percentiles are real observations of data instead of the interpolated values. Notice the different scale of each row of each panel.



indicates that we exclude the current observation from calculating the mean and standard deviation. δ is kept as 10% and the k observations should belong to the same day as the current observation. Specifically, as in Brownlees and Gallo (2006), the local window of the first mid-quote price of a day should be the k quotes following it; the neighborhood of the last observation of a day is chosen as the k data points preceding it. In the middle of the day, we select the $k/2$ points before and after the current observation. k and γ are set to 60 and 0.02 as in the original paper, respectively.

Barndorff-Nielsen et al. (2009) apply a similar idea to the quotes of stocks (see Barndorff-Nielsen et al. (2009) Section 3.1 on page C7, entry Q4), which is also used here.[27] Ticks are removed if their spreads are larger than 50 times the median spread on that day. In addition, the algorithm considers the average distance between the trade price and the median of the 50 trade prices in the neighborhood of the current price. It classifies as outliers observations where the distance between the trade price and the median of the 50 trade prices is greater than 10 times the average distance.

$$|p_n - median(p_{-n})| < 10 * \frac{1}{50}\sum_{j=1}^{50}|p_j - median(p_j)| = \begin{cases} True & \text{observation } n \text{ is kept} \\ False & \text{observation } n \text{ is removed} \end{cases}$$

Intuitively, these two methods do well when there are only "a few" quotes heavily deviating from others. However, it is difficult for this approach to filter out similar

[27] Actually, they apply a series of operations to clean data. We primarily apply the quote data and trade data rule.

outliers to those in Figure 3.1 because the local property of current observation is distorted due to the persistent enlargement of bid-ask spread. Also the parameters for identifying outliers rely on the discretion of econometricians. Brownlees and Gallo (2006) and Barndorff-Nielsen et al. (2009) choose parameters values related to the filters either through visual inspections or intensive experiments, without evaluating them against a benchmark. More examples can be found from other microstructure papers. For instance, Fleming and Lopez 1999, delete the ticks whose spread are larger than 50 times the median spread on that day. Engle and Russell (1998) filter the bid and ask of the IBM stock based on a simple threshold. They observe some disassociation of the bid and ask changes, which causes the mid-price to vary temporarily. They decide 4 ticks to be the minimum amount of change for bid and offer price to trigger a genuine price movement. There are no apparent reasons why 50 or 4 is an proper choice for filtering. This further underlies the need for a systematic evaluation of all filters based on one benchmark.

3.3.5 Cleaning Result based on the Benchmark

We attempt to remove the illiquidity effect by choosing the best filter which minimizes the distance between the fitted volatility and V_t, which is estimated by realized variance of 2-hour returns. In order to operationalize the benchmark, we assume that the returns of daily and intraday frequency follow a GARCH process as compared to the MA(1) structure of stock returns in Bandi and Russell (2008). Note that the realized variance is computed from unfiltered data.

In general, we do not see any danger that over-cleaning would be suggested by our benchmark. Barndorff-Nielsen et al. (2009)'s method appears to be suitable for

Table 3.2: Number of observations deleted by various methods

There are a total of 12 methods we apply to the processed sample. They can generally be put into three groups. The first group, which contains the first two methods, i.e. maximum tradable spreads and 99th percentile of all tradable spreads attempts to find a reliable threshold with the aid of transaction records. The second group, gathering the 7 percentiles of all percentage spreads, simply runs through the data month by month in order to ascertain outliers according to the empirical distributions of spreads. The third group, following the concept of local filtering, consists of two established methods on stock data from two published papers.

Cleaning method	Number of observations deleted	Percentage removed (%)
Maximum tradable spread	219246	3.8947
99th Percentile of tradable spreads	333351	5.9271
97th Percentile of percentage spreads	98597	1.7515
96th Percentile of percentage spreads	121806	2.1638
95th Percentile of percentage spreads	142513	2.5316
94th Percentile of percentage spreads	161729	2.8730
93th Percentile of percentage spreads	179681	3.1919
92th Percentile of percentage spreads	196836	3.4966
91th Percentile of percentage spreads	212915	3.7823
90th Percentile of percentage spreads	227838	4.0474
Brownlees and Gallo (2006)'s method	11046	0.20
Barndorff-Nielsen et al. (2009)'s method	5338	0.0948

Table 3.3: MSE of various filters

The benchmark $E(\sum_{n=1}^{N} h_t s_n q_{t,n} - V_t)$ is estimated by $\frac{1}{T}\sum_{t=1}^{T}(\sum_{n=1}^{N} \hat{h_t}\hat{s_n}\hat{q_{t,n}} - \hat{V_t})$. h_t is forecast by the GARCH(1,1) daily volatility model
$r_t = c_1 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \nu_t \;\; h_t = w + a_1 \nu_{t-1}^2 + b_1 h_{t-1}$. s_n is fitted by
$\delta_0 * exp(\sum_{j=1}^{m} \delta_j * (\Delta_n - k_j)_+)$ where $(\Delta_n - k_j)_+ > 0$ when $\Delta_n > k_j$ and
$(\Delta_n - k_j)_+ = 0 \; otherwise$, $t = 1, 2, \ldots, N$. $q_{t,n}$ is specified as
$1 - \alpha - \beta + \alpha \left(\frac{r_{t,n-1}^2}{s_n h_t}\right) + \beta q_{t,n-1}$. The score is a result of the ranking. The one that has the best performance on one country's data in terms of our benchmark receives the highest score, i.e. 12. Any tied ranking allocates the highest possible scores to the methods. The final column is the summation of all scores of one method.

Filtering method	Austria	Belgium	France	Germany	Italy	Netherlands	Spain	Sum
Maximum Tradable spread	0.020533	0.049768	0.130702	0.025478	0.33235	0.067315	4.326815	29
99th percentile of tradable spread	0.021192	0.049413	0.042893	0.025478	0.40676	0.060009	1.772955	31
90th percentile of percentage spread	0.019372	0.049347	0.042902	0.025472	0.40128	0.059920	1.750791	38
91th percentile of percentage spread	0.019122	0.049835	0.042744	0.025472	0.39348	0.059920	1.759567	42
92th percentile of percentage spread	0.019014	0.049667	**0.042453**	0.025472	0.38263	0.059920	1.735467	49
93th percentile of percentage spread	0.018788	0.049481	0.042872	0.025472	0.37201	0.059920	1.659913	45
94th percentile of percentage spread	0.018771	0.049092	0.042862	0.025472	0.36575	0.059920	1.653448	59
95th percentile of percentage spread	**0.018630**	**0.048796**	0.042871	0.025472	0.38935	0.060061	**1.506469**	56
96th percentile of percentage spread	0.018641	0.049311	0.042897	0.025472	0.38484	0.059973	1.679483	54
97th percentile of percentage spread	0.018691	0.050138	0.042896	0.025487	0.37845	**0.058922**	1.720455	50
Brownlees and Gallo (2006)'s method	12.011837	2.559600	12.601892	0.025435	0.30209	12.292285	44.393817	31
Barndorff-Nielsen et al. (2009)'s method	14.623279	3.245954	20.015327	**0.025369**	**0.30065**	10.185809	59.904429	30

Table 3.4: P-value of Wilcoxon rank sum test on the equality of MSE aginst the lowest one

The null hypothesis is that the median of the difference in the MSE generated by two methods is zero.

Cleaning method	Austria	Belgium	France	Germany	Italy	the Netherlands	Spain
Maximum tradable spread	0.66944	0.59580	0.02855**	0.88823	0.45793	0.00765***	0.32792
99th percentile of tradable spread	0.60441	0.88053	0.96039	0.88823	0.17381	0.44494	0.95368
90th percentile of percentage spread	0.62422	0.83688	0.97825	0.87015	0.13896	0.41328	0.53680
91th percentile of percentage spread	0.72121	0.85049	0.97404	0.87015	0.13514	0.41328	0.54080
92th percentile of percentage spread	0.71825	0.90573		0.87015	0.14533	0.41328	0.66829
93th percentile of percentage spread	0.75560	0.90469	0.96816	0.87015	0.16077	0.41328	0.59124
94th percentile of percentage spread	0.71314	0.98182	0.96312	0.87015	0.20074	0.41328	0.57102
95th percentile of percentage spread			0.95074	0.87015	0.25316	0.44087	
96th percentile of percentage spread	0.90490	0.95368	0.97636	0.87015	0.35711	0.41478	0.95032
97th percentile of percentage spread	0.97720	0.84223	0.96648	0.85813	0.36445		0.83647
Brownlees and Gallo (2006)'s method	0.15389	0.00004***	0.17331	0.98770	0.93879	0.00000***	0.00000***
Barndorff-Nielsen et al. (2009)'s method	0.00000***	0.00000***	0.01565**			0.00000***	0.00000***

Germany and Italy. Given that this procedure was originally designed to filter stock data, we can infer that the liquidity of German and Italian bonds resembles stock liquidity. But for other countries, the two local window filtering methods have the worst performance. Given the low deletion rate seen in Table 3.2, we may view the two cleaned datasets from two ”local window” filters as approximations to a raw dataset and we can conclude that further cleaning is definitely necessary. The most striking comparison comes from Spanish bond volatility. Due to the inadequate filtering, the fitted intraday return volatility diverges from the model-free daily realized variance. The first 10 filters, which concentrate on properties of percentage spread, yield similar result. The 95th percentile of all percentage spread turns out to be the best filter for Austria, Belgium and Spain. The 97th percentile wins in the Netherlands while the 92th is preferred by French bonds.

The closeness of MSE estimation prompts us to examine the statistical difference in various filters. Table 3.4 shows the Wilcoxon rank sum test on the equality of such MSE against the lowest one. Not surprisingly, the difference in most of the filters is statistically insignificant. In particular, it makes very little difference to choose one particular filter for German and Italian bonds. However, there is generally a huge gap between the performance of two local window filters and the rest. Interestingly, although the mean square error of Brownlees and Gallo (2006)’s method is thirty times more than that of the 92nd percentile method according to French bond data, they are not statistically different. A further investigation of the squared error series for France reveals that the large numerical difference arises from only a few observations. Therefore a rank sum test which is robust to outliers cannot reject the null hypothesis. On the other hand, the test surprisingly suggests that the maximum tradable spread method is statistically worse than the best methods based on French and Netherlands data. Judging from Table 4 and Figure 2, some doubtful matches of trades with quotes may influence the accuracy of identifying the maximum of ”tradable” spreads and lead to the inappropriate inclusion of some outliers.

If we rank the 12 methods and assign a score according to the ranking, the relative position of each is shown in the final column. The best method of each country will receive 12 points as there are twelve methods and any tied ranking is allocated the highest possible score. The summation of each method's score derived from each country is presented in the final column of Table 3.3. The 94th percentile becomes the top of our list, which indicates a generally acceptable filtering effect. The 95th percentile ranks in the second place, though it proves to be the best for three countries. From the ranking we can conclude that filtering based on spreads outperforms the mid-quote filtering. The reason may be that the latter ones are designed to clean stocks data which is very different from the bond data of the MTS dataset.

For a robustness check we compute the realized volatility by sampling the original tick-by-tick returns on a grid of 15, 30, 45, 60, 75, 90, and 105 minutes, respectively. The relative performance of these filters remains statistically the same when increasing the sampling frequency to every 105 or 90 minutes. But the results significantly change for all countries (except for Germany and Italy) when the sampling window is lower than 75 minutes, which proves the existence of a severe illiquidity issue.

3.4 Model Estimation Result

3.4.1 Daily Model Result and Evaluation

The subsequent results are all based on the best filters for their perspective countries. The summary statistics for sample series of daily returns are presented below in Table 3.5.[28]

Table 3.5: Summary statistics of daily series

The daily log returns are computed from 17:00 mid-quote price of cleaned series. The mean and standard deviation are in percentage point. The daily sample lasts from January 02, 2009 to March 30, 2012.

Country	N	Mean	St.D.	Skew.	Kurt. (excess)
Austria	827	0.0132	0.450	−0.393	3.560
Belgium	826	0.0096	0.489	−0.238	5.875
France	827	0.0116	0.411	−0.018	2.592
Germany	827	0.0167	0.462	0.165	1.559
Italy	827	0.0013	0.646	1.453	22.210
the Netherlands	827	0.0180	0.402	0.198	1.701
Spain	827	−0.0051	0.757	−0.184	50.320

The standard deviation is much larger than the mean for all seven countries and high kurtosis is present for Italy and Spain. Interestingly, Spain is the only country with a negative albeit not significant mean. Germany, France and the Netherlands

[28] We deleted one day of Belgian data because some of the filters eliminate January 02, 2009 entirely.

Table 3.6: Estimation of the daily GARCH(1,1) model

The GARCH(1,1) model for Austrian, Belgian, French, German and Dutch government bonds is $r_t = c_1 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \nu_t,\ \ h_t = w + a_1 \nu_{t-1}^2 + b_1 h_{t-1}$. The GARCH(1,1) model for Italian and Spanish government bonds is specified as

$r_t = c_1 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \sum_{p=1}^{4} d_p * dummy_p + \nu_t,\ \ h_t = w + (a_1 + a_2 * I(SMP_{t-1} > 0))\nu_{t-1}^2 + (b_1 + b_2 * I(SMP_{t-1} > 0))h_{t-1}$.

The conditional distribution of error $\nu_t|\mathcal{F}_{t-1}$ follows a normal distribution $N(0, h_t)$. The t-stat is in parentheses. The insignificant parameters are largely omitted. The in-sample period covers from January 02, 2009 to March 30, 2012. The daily return is generated from a mid-quote at 17:00:00 every day of the clean series.

country	c	ϕ_1	ϕ_2	w	a_1	b_1	a_2	b_2	d_1	d_2	d_3	d_4
Austria	0.0233		−0.0847	0.0038	0.0914	0.8879						
	(1.8928)		(−2.3031)	(2.4496)	(4.8150)	(39.7472)						
Belgium	0.0130	0.1437	−0.0933	0.0067	0.0972	0.8686						
	(1.0166)	(3.7280)	(−2.4822)	(2.8869)	(4.3066)	(29.7712)						
France	0.0188		−0.0983	0.0026	0.0659	0.9161						
	(1.5808)		(−2.7008)	(2.1455)	(4.4179)	(49.5816)						
Germany	0.0137	0.0773	−0.1085	0.0036	0.0623	0.9202						
	(0.9857)	(2.1916)	(−3.0450)	(1.8888)	(3.5788)	(40.9629)						
Italy	0.015769	0.1426	−0.1306	0.0041	0.1736	0.8182	0.4125	−0.2213	−1.716	2.1789	6.8597	5.5181
	(1.4350)	(3.8034)	(−3.5829)	(1.5359)	(4.0766)	(16.3389)	(2.3154)	(−2.2692)	(−5.2562)	(5.8314)	(7.7226)	(2.1629)
the Netherlands	0.0197		−0.0832	0.0022	0.0583	0.9277						
	(1.6213)		(−2.3055)	(1.8271)	(3.6269)	(46.0359)						
Spain	−0.0084	0.1338	−0.0803	0.0114	0.1271	0.8232	0.4945	−0.2459	−1.9745	3.4307	6.5602	
	(−0.5908)	(4.2407)	(−2.6381)	(2.1584)	(3.9550)	(17.0000)	(2.9308)	(−3.2145)	(−5.0229)	(9.2132)	(14.2098)	

possess kurtosis lower than that of a normal distribution. The GARCH result enriches the findings of Table 3.5. Some of the first order autoregressive coefficients are not significantly different from zero and therefore are not reported here. In the conditional variance equation, w is significantly different from zero except for Italy. Given that we control for the persistent increase in volatility due to the SMP with a slightly more complicated structure, the significance of w is of lesser importance.[29] We note the high persistence of volatility for France, Germany and the Netherlands (with estimated coefficients above 0.9) compared to the volatility of the other countries. The high a_1s of Italian and Spanish bonds clearly indicates that investors attach relatively more importance to volatility shocks. The low persistence (b_1) of the two distressed countries is consistent with Chou (1988) who examines the US stock market during the period 1967-1973 and finds a low persistence coefficient (delta=0.778) which characterizes this period of high uncertainty. During the period when the SMP was launched, we do find a 30% reduction in persistence for both Italy and Spain, which is confirmed by a significant and negative b_2. Further, the ECB's influence dominated the bond market with a surge in the coefficient measuring the effect of shocks ($a1 + a2$) to around 0.6, which provides striking evidence that conditional volatility is greatly affected by the ECB shocks. The sum of a_1, a_2, b_1, and b_2 exceeds 1 and thus this implies a non-stationary daily conditional variance. This temporary non-stationarity is successfully captured by our model.

We also want to examine the correlation of daily volatility forecasts with intraday activity. Theoretically, different types of traders and market makers may be exposed to and concerned about risk with different time horizons. Active fund managers and market makers attribute greater importance to short-term volatility, whereas pension

[29]We tested the change in w during the SMP period. The change turns out to be insignificant.

and passive fund managers are mainly concerned with long-term fluctuations. In addition, the increasing uncertainty relative to the macro environment and country credit risk may produce greater short-term bond price fluctuations which may affect intraday returns relatively more than daily returns. It is therefore always important to compare daily volatility with volatility computed from intraday returns, and assess whether it is necessary to include the daily variance component.

To study the relation between daily volatility forecasts and intraday volatility, we compute the *ex post* correlation, as in Andersen and Bollerslev (1998) between the daily volatility forecast and the cumulative squared intraday returns for the period from April 02, 2012 to December 30, 2013. Traditionally, R^2 of a Mincer-Zarnowitz (MZ), $r_t^2 = a + bh_t + u_t$, regression is used to evaluate the out-of-sample forecast performance of a GARCH type model. The R^2 is simply the square of the correlation between the regressor and the regressand. As noted by Engle and Patton (2001), squared daily returns are a noisy measure of the latent h_t. The noise could mask the true relationship of the forecast and the ”real” volatility. On the other hand, realized variance, which is the cumulative squared intraday return, proves to be able to provide a more efficient benchmark for the valuation of the volatility forecast.[30] Hence, we use the same approach for assessing the forecasting ability of our model.

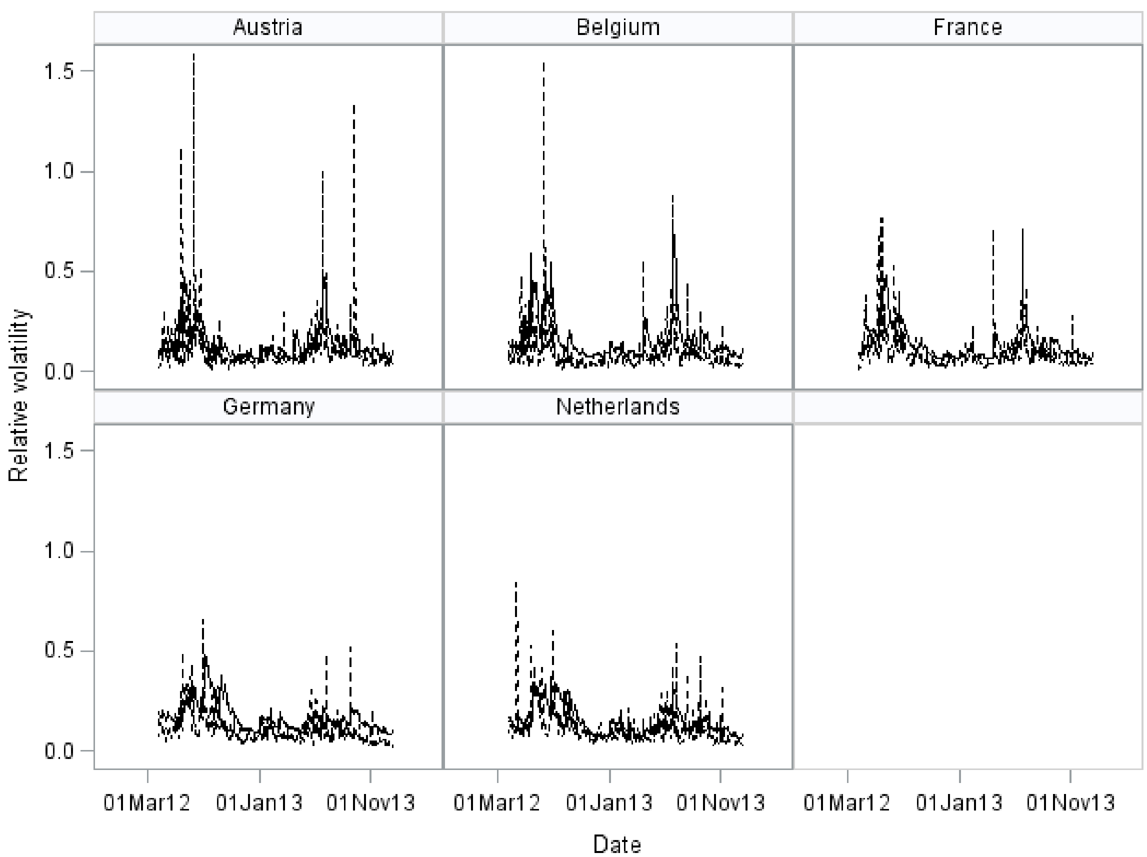
Table 3.7: Ex post correlations between forecasted daily volatility with cumulative squared 10-minute returns.

Austria	Belgium	France	Germany	Italy	the Netherlands	Spain
0.345	0.401	0.466	0.404	0.507	0.437	0.514

[30] Hansen and Lunde (2006) show a significant increase in R^2 when the realized variance is used in a MZ regression.

Figure 3.4: Daily volatility forecast and realized variance of the four safer countries

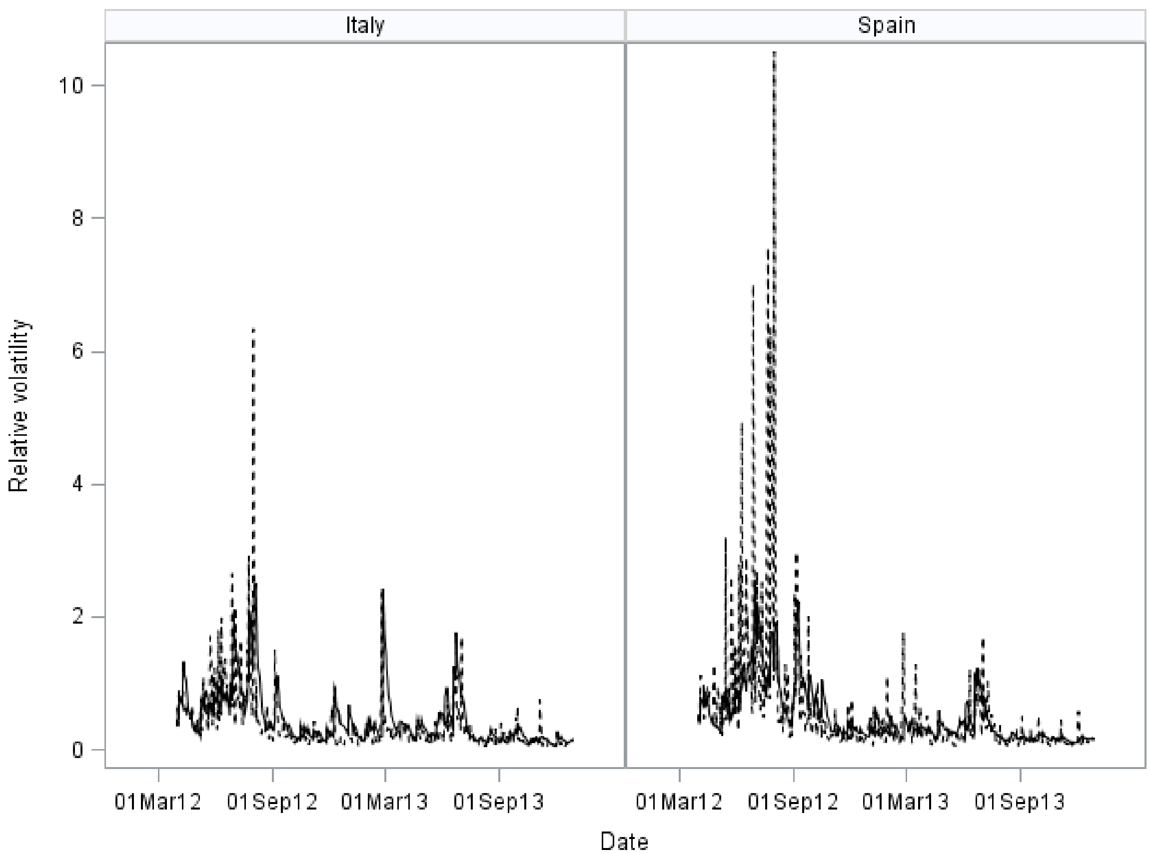
The solid line represents one-day-ahead daily conditional variance forecast. The dashed line is plotted according to cumulative 10-minute returns. The one-day-ahead daily conditional variance is generated from the GARCH(1,1) model $r_t = c_1 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \nu_t \quad \nu_t|\mathcal{F}_{t-1} \sim N(0, h_t),\ h_t = w + a_1\nu_{t-1}^2 + b_1 h_{t-1}$. The cumulative 10-minute return is computed as $\sum_{n=1}^{N} r_{t,n}^2$. The forecast period is from April 02, 2012 to December 30, 2013.



The correlation ranges from as low as 0.345 for Austria to as high as around 0.5 for Italy and Spain. A simple regression of cumulative squared returns on forecast conditional variances indicates that the forecast explains at least $0.345^2 = 0.12 = 12\%$ of the total intraday variation for the Austrian market. The Spanish and Italian markets show

Figure 3.5: Daily volatility forecast and realized variance of Italy and Spain

The solid line represents one-day-ahead daily conditional variance forecast. The dashed line is plotted according to cumulative 10-minute returns. The one-day-ahead forecast of daily conditional variance is generated from the GARCH(1,1) model
$r_t = c_1 + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \nu_t \quad \nu_t|\mathcal{F}_{t-1} \sim N(0, h_t),\ h_t = w + a_1 \nu_{t-1}^2 + b_1 h_{t-1}$. The cumulative 10-minute return is computed as $\sum_{n=1}^{N} r_{t,n}^2$. The forecast model of Italy and Spain no longer takes those dummy variables in both conditional mean and conditional variance equations into account since in the out-of-sample period the effect of those one-time events no longer exists. The forecast period is from April 02, 2012 to December 30, 2013.



a relatively high correlation between the volatility computed using intraday returns and the volatility predicted using daily returns(see Table 3.7). Apart from the big jump of daily volatility on August 02, 2012, we generally see that the two lines closely follow each other in Figure 3.4 and Figure 3.5. As the daily volatility is independent of the two intraday components, it does embody some degree of predictability, which could be explained by investors' risk preferences. Ignoring this daily effect would mistakenly attribute this part either to intraday periodicity or intraday volatility. However, high-frequency movement has definitely become a primary concern for investors, which is exemplified by the few peaks in each panel of Figure 3.4 and Figure 3.5. Instead of being subordinated as a secondary source of risk, the magnitude of intraday volatility is sometimes paramount.

3.4.2 Intraday Result

Table 3.8: Summary statistics of intraday 10-minute returns

The 10-minute returns are derived from the clean series. Moreover, the returns from 8:15-9:00 (excluded) are removed from the final series.

Country	N	Mean	St.D.	Skew.	Kurt. (excess)
Austria	22956	0.0003	0.046	1.236	65.278
Belgium	22961	0.0004	0.042	0.402	28.379
France	22968	0.0004	0.042	0.226	13.366
Germany	22957	0.0002	0.046	0.111	9.490
Italy	22979	0.0008	0.089	−2.366	108.35
the Netherlands	22943	0.0002	0.048	−0.100	14.118
Spain	22861	0.0007	0.105	−0.330	75.015

As expected, in the intraday data, Italy and Spain still have higher standard deviations, with twice the magnitude of the others in Table 3.8. The higher average of intraday returns tends to compensate the higher risk of Italian and Spanish government bonds. The signs of skewness seem not consistent with daily returns based on Table 3.5 and Table 3.8. The skewness of Austria, Belgium, France and Italy reverses its sign from the daily interval to the 10-minute interval. Nonetheless the kurtosis tells a consistent story in both daily and intraday data. Spain and Italy still have the most extreme kurtosis, with Austria and Belgium following them. Overall the kurtosis of the 10-minute returns is larger than that of the daily returns.

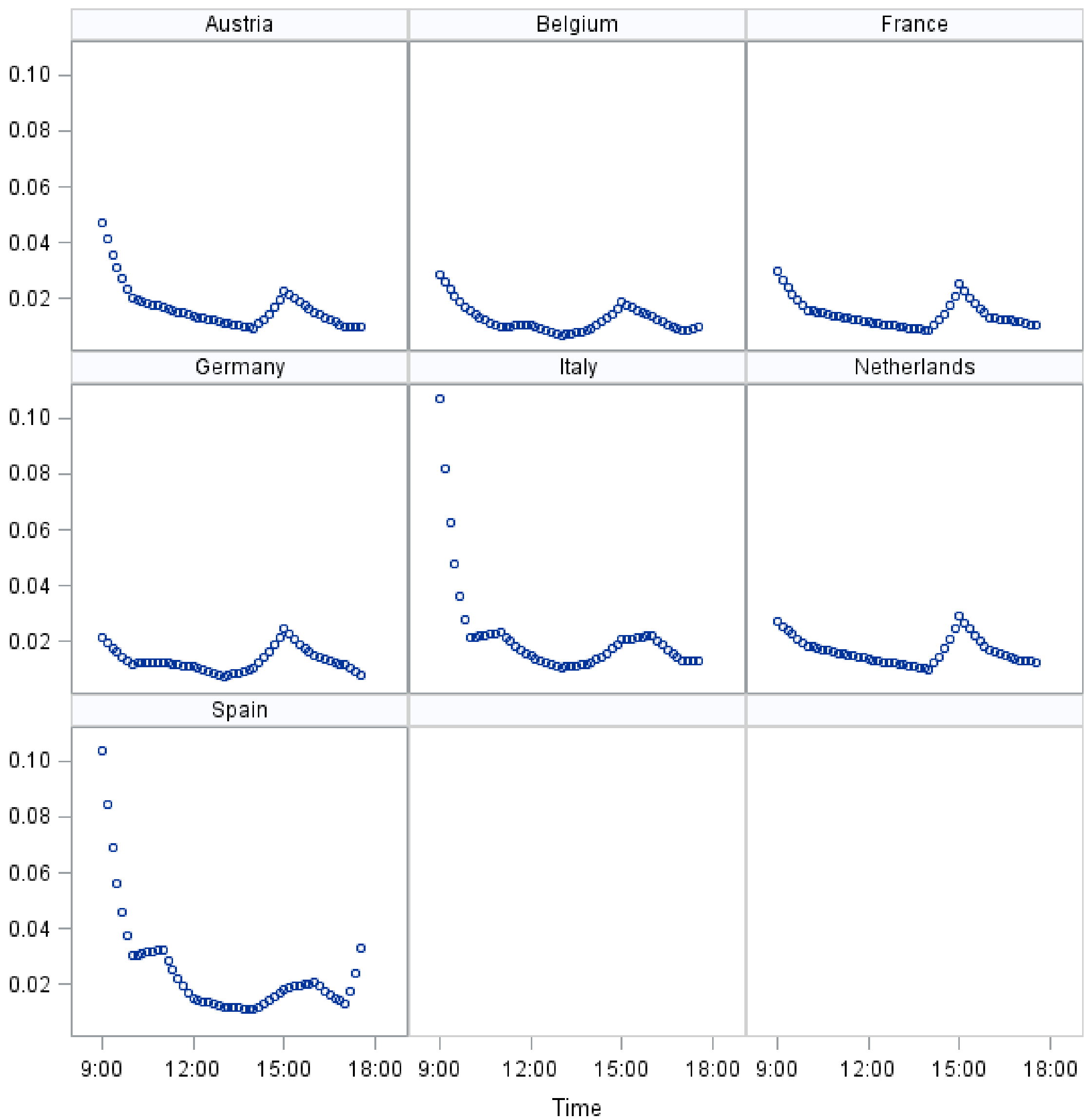
3.4.2.1 Diurnal Component

The intraday periodicity estimation consistently underlies the distinctive risk of Italian and Spanish government bonds. The results can be categorized into two groups. The typical patterns of Austria, Belgium, Germany, France constitute of one group. Even though there are three knots omitted in model specification of Austrian, French and Dutch government bonds for the estimation reason (see Section 3.2.2), the four countries still resemble each other. Since we remove the first 45 minutes of returns (see Section 3.3.3), the seasonal pattern starts from 9:00 to 17:30. The market volatility decreases rapidly in the first hour until 10:00, after which the decrease is dampened. The periodicity starts to pick up from 13:00 or 14:00 and peaks at 15:00, which is probably due to the opening of the US market and the volatility spill-over effect. The markets then adjust calmly towards the end of trading period without any further increase in volatility.

The other group naturally contains Italy and Spain. With an early spike in volatility near the opening time, the tension of Italian and Spanish government bonds is not overshadowed by American influence. The shift from a volatile period to a more stable one is achieved at 10:00. Though later volatility bounces back slightly, it trends down

Figure 3.6: Diurnal components of the seven European countries

The diurnal component is specified as $\delta_0 * exp(\sum_{j=1}^{m} \delta_j * (\Delta_n - k_j)_+)$ where $\Delta_n - k_j > 0$ when $\Delta_n > k_j$ and $\Delta_n - k_j = 0$ *otherwise*, $\Delta_n = \frac{n}{N}$, $n = 1, 2, \ldots, N$. There are 8 knots set for each hour of bonds of Belgium, Germany, Italy, and Spain and an extra knot set for the final half-hour. Three knots at 11:00,12:00, and 13:00 are omitted for estimation reason for bonds of Austria, France and the Netherlands.



to the bottom level around 13:00 or 14:00. Then the effect of American opening is still present and lasts until 16:00. In the final half-hour, Spanish bonds complete the J-shape pattern. Overall, one common point that the seven countries share is that the volatility opens at a high level. This could be due to market makers competing less aggressively at the opening or to a greater uncertainty about the bond prices right after the overnight period.

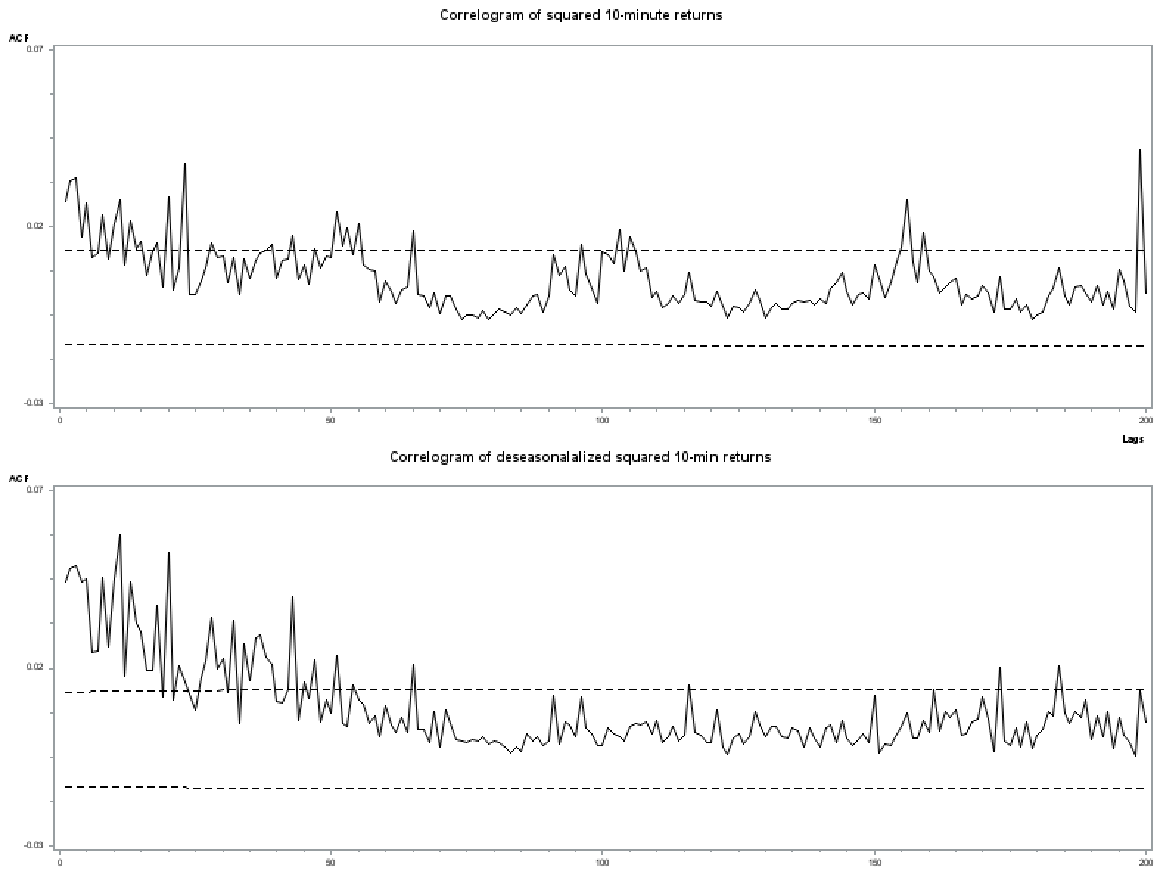
3.4.2.2 Correlogram of Squared Intraday Returns

The original goal of capturing intraday periodicity is to remove the recurring cycle of intraday volatility so that the filtered series follows a typical process as the daily or weekly returns. The effect of our estimation method of intraday seasonality is depicted by the correlograms of raw and filtered series. Here we choose the correlogram of Spanish bonds as an example.

There are a number of peaks in autocorrelations appearing at different frequencies. The wave-like pattern is very clear in the upper panel of Figure 3.7. The first peak is achieved at the 23rd lag in the upper panel of Figure 3.7, approximately corresponding to the half-day lag given that there are 52 returns each day . The correlation pertaining to interval 51 are significant, suggesting a significant daily frequency cycle. With lags advancing further away, the pattern persists and repeats the cycle at the 156th and 199th lags of Figure 3.7. The lower panel of figure 3.7 shows the correlogram of deseasonalized return $y_{t,n} = r_{t,n}/\hat{h}_t\hat{s_n}$. The notable peaks corresponding to the aforementioned lags mostly become less significant. The discernible pattern is generally destroyed by the the deseasonalization. More importantly, there is no sign of non-stationary situations. We therefore assume that the filtered returns are covariance-stationary. The stationarity is important to the following intraday volatility model

Figure 3.7: Correlogram of 10-minute returns of 10-year Spanish government bonds

Dashed lines represent the two times standard errors of autocorrelations. The upper panel depicts the correlogram of original 10-minute returns while the lower panel plots that of deseasonalized ones. The intraday data covers the period from April 02, 2012 to December 30, 2013.



and the discussion of the asymptotic distribution of the estimator. However, there are still some remaining recurring patterns of lower frequency as a possible manifestation of longer-horizon activities such as scheduled macro announcements.[31] The finding is similar to Andersen and Bollerslev (1997), who also discover some visible correlation cycles after the filtering. Andersen and Bollerslev (1997) point out that the intraday

[31] The autocorrelation of 116, 173 and 184th lags is still significant.

periodicity, being a composition of activities with different frequencies, strengthens the notion of decomposition of intraday volatility.

3.4.2.3 Intraday Volatility

Interestingly, we are modeling daily and intraday volatility in the same manner as a GARCH(1,1) model. The two GARCH(1,1) models enable us to compare the behavior of daily and intraday volatility. From Table 3.9, we can see that most of the spline parameters $\delta_1 - \delta_9$ are significant as well as other parameters. Notably, the relative magnitudes of α and β change dramatically across countries, with Spanish bonds possessing the highest persistence of volatility, probably due to the general success of capturing the periodicity of intraday volatility, whereas the β of the Netherlands drops to the bottom of the seven countries. The volatility of the 10-year bonds of Austria, Belgium, France, Germany and Italy maintains the characteristics of the daily GARCH volatility. Italy still has a relatively low β and the highest α. The overall scale of volatility is partially reflected in parameter δ_0, which is the constant in the spline equation. Still Spain has the highest δ_0, with Italy and Austria following it. None of the other countries has a constant exceeding 0.05. It seems that the estimation of intraday volatility of Spain and the Netherlands provides a different picture from daily volatility. However, the dynamics of intraday volatility still vary significantly across countries.

3.5 Forecast Evaluation

In view of the general success of GARCH(1,1) model in forecasting daily volatility of bond markets, we want to compare the forecast performance of our model against the GARCH(1,1) model estimated for daily returns. The out-of-sample period covers the

Table 3.9: Estimation of the intraday volatility model

The intraday unit GARCH is specified as $q_{t,n} = 1 - \alpha - \beta + \alpha \left(\frac{r^2_{t,n-1}}{s_{n-1}h_t} \right) + \beta q_{t,n-1}$. It is jointly estimated with the diurnal part, which is a exponential piecewise linear function of time interval.$s_n = \delta_0 * exp(\sum_{j=1}^{m} \delta_j * (\Delta_n - k_j)_+)$ where $\Delta_n - k_j > 0$ when $\Delta_n > k_j$ and $\Delta_n - k_j = 0$ $otherwise$, $\Delta_n = 1/N, 2/N, \ldots, 1$

Country	α	β	δ_0	δ_1	δ_2	δ_3	δ_4	δ_5	δ_6	δ_7	δ_8	δ_9
Austria	0.0651	0.8917	0.0717	−10.302	8.9880				9.6131	−12.195	0.9862	4.1739
	(17.2530)	(129.8230)	(17.0070)	(−21.7838)	(17.1380)				(21.6807)	(−16.1481)	(1.1660)	(2.5207)
Belgium	0.0866	0.8399	0.0311	−5.5979	1.8813	4.8921	−5.2441	7.3146	2.5396	−8.9567	−0.6500	6.5569
	(17.5083)	(96.3974)	(16.6534)	(−9.7082)	(2.0185)	(6.2388)	(−6.6089)	(8.5084)	(3.0232)	(−11.3734)	(−0.7396)	(3.8864)
France	0.0805	0.8661	0.0322	−5.4253	4.1413				10.2893	−14.333	4.1029	−0.5934
	(16.8397)	(98.0778)	(16.6469)	(−10.6934)	(7.3420)				(22.7116)	(−18.5009)	(4.5570)	(−0.3536)
Germany	0.0532	0.8994	0.0218	−4.1331	3.8132	−0.5480	−2.6307	6.8180	3.7221	−11.131	1.9272	0.3657
	(12.4611)	(100.5398)	(12.6202)	(−4.5372)	(2.3187)	(−0.4385)	(−3.1607)	(8.0356)	(3.9193)	(−12.1327)	(1.8381)	(0.1536)
Italy	0.1053	0.8484	0.054	−6.2639	5.9838	−3.1780	1.1216	3.8395	1.0415	−3.0627	−1.3971	2.4838
	(17.7867)	(93.7866)	(15.1586)	(−11.0628)	(6.4409)	(−3.9693)	(1.5069)	(4.6574)	(1.1765)	(−3.4437)	(−1.5149)	(1.5213)
the Netherlands	0.0753	0.8204	0.029	−3.3011	2.0024				10.4359	−13.924	2.7475	1.2561
	(14.5030)	(51.8162)	(17.3643)	(−6.5688)	(3.6026)				(22.5383)	(−17.8215)	(3.0339)	(0.7248)
Spain	0.0598	0.9039	0.1400	−12.429	13.6642	−7.6326	4.5944	0.5584	5.8541	−2.7180	−6.6631	17.2191
	(21.0195)	(202.3823)	(15.9266)	(−21.4210)	(14.0371)	(−8.5132)	(5.5998)	(0.7805)	(7.8211)	(−3.2709)	(−7.7462)	(10.7740)

[a] The t-statistics are shown in parentheses.

first two months of 2014. We filter the intraday observations by the most suitable thresholds derived from in-sample cleaning (see Section 3.3.5). Since the bonds of Italy and Germany require adaptive filtering, which utilizes future information, we restrict the bonds to have a percentage spread less than 50 basis point. In addition, if there is a new issue during the out-of-sample period, we switch to the new bond according to the rule described in Section 3.3.2.

Four criteria are considered to evaluate the forecast performance, namely mean square error (MSE), quasi-likelihood based error[32] (QLIKE), mean absolute error (MAE), and correlation between volatility forecast and benchmark volatility, which is approximated by the realized volatility of raw 2-hour returns. The validity of using raw 2-hour return to compute realized volatility is proven in the robustness check of our filtering MSE result(see Section 3.3.5). As Patton (2011) has shown, the "MSE" and "QLIKE" loss functions, which lead to unbiased predictors give a consistent ranking of volatility forecasts when the benchmark is a noisy volatility proxy. The "MAE" loss function, though it may not have the nice properties of "MSE" and "QLIKE", is robust to outliers. The "CORR" function generally measures the closeness between the "Patterns" of volatility forecasts and the volatility proxies. The one-day-ahead forecast of the daily GARCH(1,1) model for day t is denoted as $h^f_{1,t}$ while the forecast from the intraday model is labelled as $h^f_{2,t}$

$$
\begin{aligned}
MSE(h^f_{i,t}) &= \frac{1}{\bar{T}} \sum_{t=1}^{\bar{T}} (h^f_{i,t} - \hat{V}_t)^2 \\
MAE(h^f_{i,t}) &= \frac{1}{\bar{T}} \sum_{t=1}^{\bar{T}} |h^f_{i,t} - \hat{V}_t|
\end{aligned}
$$

[32]This is a likelihood based loss function that asymmetrically penalizes over- and under-prediction

$$QLIKE(h_{i,t}^f) = \frac{1}{\bar{T}} \sum_{t=1}^{\bar{T}} (log(h_{i,t}^f) + \hat{V}_t / h_{i,t}^f)$$

$$CORR(h_{t,1}) = \frac{1}{\bar{T}} \sum_{t=1}^{\bar{T}} (h_{i,t}^f - \bar{h}_{i,t}^f)(\hat{V}_t - \bar{\hat{V}}_t)$$

where $i = 1$ or 2

The forecasting schemes for the two models are now laid out for the purpose of fair comparison, i.e. using all the information which can be processed by each model before day t. In Section 3.4.1 and Table 3.7, we have already seen the predictive power of the daily GARCH(1,1) model for Italy and Spain. The parameters involved in forecasting are derived from a fixed-sample and all daily forecasts are generated from these parameters. In order to use new information to improve the daily model's forecast, we estimate the GARCH(1,1) model whenever a new day can be included in the fitting sample and produce the forecast for the next day. The forecasts generated by the dynamic sample approach, can be substantially different from those generated by fixed sample approach especially for Italy and Spain (see Table 3.5 for the volatile period of Italian and Spanish bonds during 2012). For intraday model, the one-day-ahead forecast $h_{2,t}^f$ is equal to $(h_{1,t}^f \sum_{n=1}^{N} s_n q_{t,n}^f)$ where $q_{t,n}^f$ is a n-step-ahead forecast generated from the intraday GARCH(1,1) model. For the first interval every day, the $q_{t,n}^f$ is initialized by $1/N \sum_{n=1}^{N} r_{t-1,n}^2 / (h_{t-1,n}^f s_n)$. Obviously, both methods exclude the information that becomes available during the forecasting day and the forecast from the GARCH model estimated on daily data is nested in $h_{2,t}^f$. It is also evident that the extra predictive power as compared to the daily model stems from the diurnal and intraday GARCH components. The accuracy of $h_{2,t}^f$ relies on the success of estimating the fixed diurnal component and an adequate specification of the GARCH component. The intraday periodicity s_n is assumed to be unchanged during the out-of-sample period. In fact, we can view the $\sum_{n=1}^{N} s_n q_{t,n}^f$ as a factor that modifies the daily GARCH(1,1) forecast according to a larger information set. If the intraday information is indeed relevant, it will improve the daily GARCH(1,1) forecast. To measure the extra information content we propose to re-estimate the intraday model with a daily dynamic-sample forecast and normalize the diurnal component so that $\sum_{n=1}^{N} s_n = 1$. Since $E(q_{t,n}) = 1$, the intraday model will provide very little information if $q_{t,n}^f$ stays close to its unconditional expectation and if the summation of s_n is 1. In other words, if this is indeed the case, then $h_{1,t}^f$ and $h_{2,t}^f$ would be identical. The normalization of the intraday volatility pattern is a common practice in fitting and forecasting intraday volatility. Taylor and Xu (1997), for example, standardize the sum of their variance seasonal pattern to be 24 for studying foreign exchange volatility. Table 3.10 presents the Diebold and Mariano (1995) test for forecast performance comparison between the two models. A negative value indicates that the component GARCH model which uses information from the intraday model produces better volatility forecasts than the daily model.

The forecast daily volatility is presented in Figure 3.8 for four major European countries. The correlation between the daily GARCH(1,1) forecast and the intraday component GARCH forecast is around 0.4 for Austria, Belgium, France and Germany while it increases to roughly 0.6 for Italy and Spain and reaches 0.8 for the Netherlands. However, the low correlation does not necessarily indicate a better forecast ability, as it is seen below that the intraday component GARCH model is more suitable for forecasting the volatility of the Dutch bonds. From Figure 3.8 we can see that the two forecasts tend to diverge when there is little variation of returns from the previous trading day. This can be explained by the nature of $q_{t,n}^f$– the multi-step-ahead forecast

Figure 3.8: Forecast plots for different countries

The blue line represents the realized volatility computed from 2-hour returns. The red line is the daily volatility forecast from the daily GARCH(1,1) model. The green line is the forecast given by the intraday GARCH model.

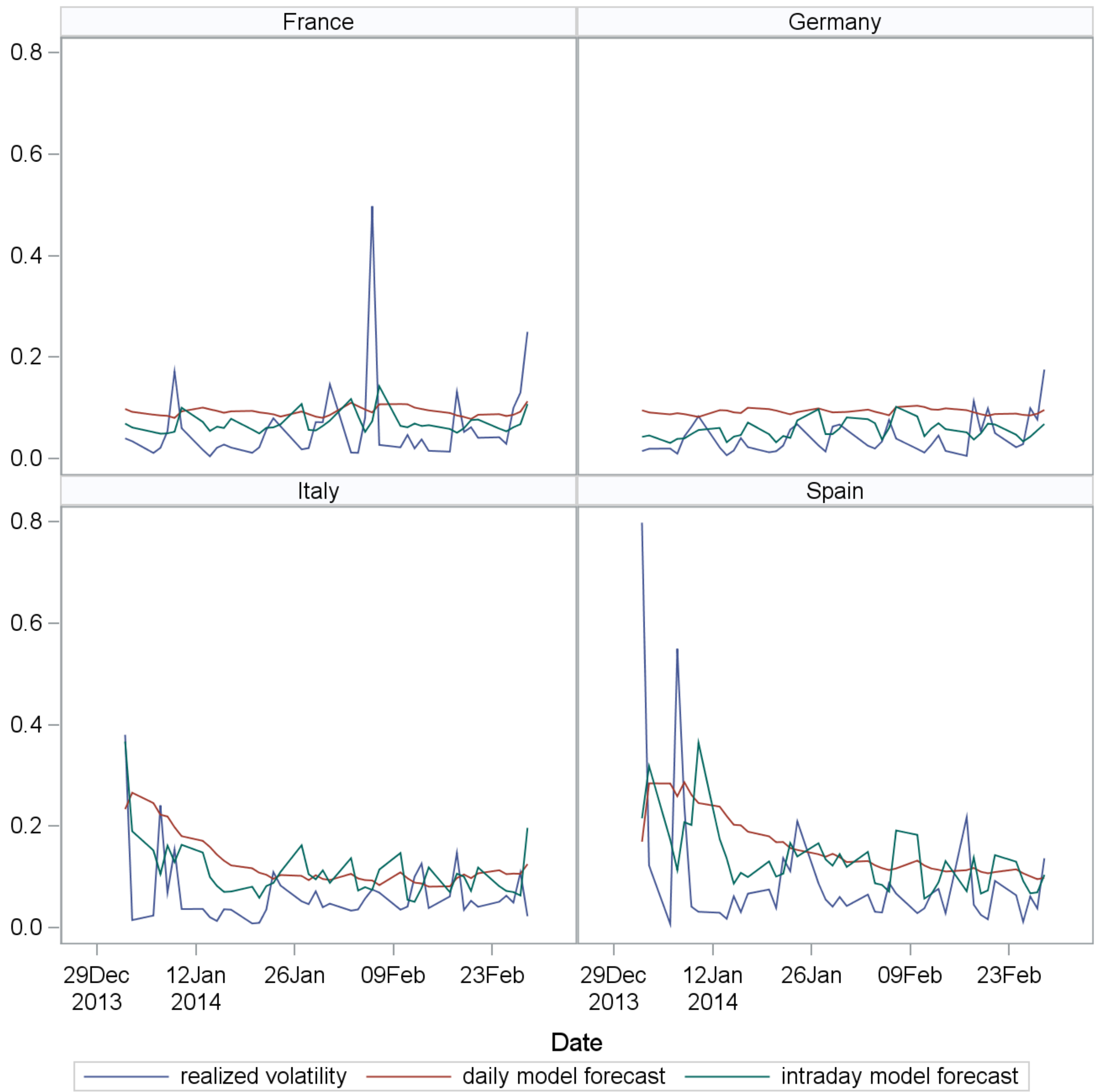


Table 3.10: The Diebold and Mariano (1995) test for a predictive ability comparison between the daily GARCH(1,1) model and the intraday multiplicative component GARCH model. Negative values show the preference to the intraday model.

***,**,* denote 1%,5%,10% significance respectively

Country	MSE	QLIKE	MAE	CORR
Austria	−0.0012**	−0.1046***	−0.0131***	−0.0001
Belgium	−0.0017***	−0.1370***	−0.0209***	−0.0000
France	−0.0008	−0.0743	−0.0130***	−0.0002
Germany	−0.0023***	−0.2260***	−0.0245***	−0.0002***
Italy	−0.0026	−0.0582	−0.0117	−0.0008
the Netherlands	−0.0004**	−0.0353***	−0.0043***	−0.0000
Spain	−0.0014	−0.0521	−0.0139	0.0001

which is a component of $h^f_{2,t}$. The half-life of $q^f_{t,n}$ is roughly 15 (or even lower for some countries) intervals, which corresponds to two-and-a-half hours whereas the half-life of $h^f_{1,t}$ is around 20 days! Therefore, when there is a shock followed by a few quiet trading days, the daily GARCH(1,1) model will generally over-predict the daily volatility but the intraday model is capable of quickly giving a low volatility forecast.

It turns out that the intraday model provides a superior forecast for most of the less volatile bonds, whereas there is no ”winner model” for Italian and Spanish bonds. The ”MSE” and ”QLIKE” measures both confirm the better forecast accuracy of the intraday model and Figure 3.8 suggests that the daily GARCH(1,1) model generally produces too high a volatility forecast for safer government bonds. For ”CORR”, which measures the synchronicity of volatility forecasts and the volatility proxy, neither of the two models seems to be better than the other. An insignificantly different forecast performance is expected for Italy and Spain, as the two models are both fitted to a high-volatility environment but the volatility is very low during the out-of-sample period. On the other hand, since the volatility of the other five bond series is always low, the intraday model does provide extra information to the daily GARCH(1,1) forecast. One exception is French bonds. Only the "MAE" loss function gives a significant result, which may be explained by the sudden spike in the middle of the forecasting period. The other measures are easily influenced by this outlier. Overall, we do see that the intraday data can be employed to improve the daily volatility forecast if volatility stays in one regime. In the robustness check, we investigate the possibility that the over-prediction generated by the daily GARCH(1,1) model is due to the omission of the overnight movements in realized volatility computation. We redo the Diebold and Mariano (1995) test, adding the square of the overnight returns[33] to the realized volatility. The test result is not changed in any significance level.

3.6 Conclusion

In this paper, we examine the daily and intraday volatility of the long-term government bonds of seven European countries during the sovereign debt crisis. A new specification of intraday periodicity, along with a unit GARCH(1,1) model, is formulated under the framework of Engle and Sokalska (2012). Several filters are presented and tested against the benchmark inspired by Bandi and Russell (2008) using the data of the MTS interdealer market. It appears that the percentile approach is most suitable for our data. The necessity of filtering suggests that only part of the information contained in

[33]The overnight return is the log of the mid-quote price at 9:00 minus the log of the mid-quote price at 17:30 the previous day.

the intraday data is relevant for longer horizon volatility. The risk of Italian and Spanish bonds is emphasized in both daily and high-frequency estimations. The standard deviations in daily and intraday returns give the first clue. The importance of the ECB interventions on the secondary sovereign debt markets is evident. The ECB's SMP considerably changed the features of bond volatility of Italy and Spain at a daily level where the effect of shocks on volatility increases during the ECB intervention. At the intraday level, periodicity is confirmed and captured successfully for some countries. The volatility transmission from US to European markets is demonstrated in all countries. The evaluation of the forecasting ability of the daily GARCH(1,1) model and the intraday multiplicative component GARCH model demonstrates that the intraday information is able to improve the forecast accuracy for less volatile bonds.

Chapter 4

Managing Portfolio Risk during Crisis Times: A Dynamic Conditional Correlation Perspective

4.1 Introduction

We study the intraday correlation in European bond markets. First, we find that during the debt crisis, there is no heightened correlation between safer and riskier bonds. For most of the bond pairs, the unconditional correlations were significantly lowered during the ECB's Security Purchase Program (SMP). Safer bonds of Austria, Belgium, France and Germany exhibited low and even negative correlation on average with bonds of Italy and Spain. Nevertheless, the conditional correlation was boosted by the program from the pre-purchase fall. The Italy-Spain correlation decreased during the purchase, suggesting that we can improve the diversification of a bond portfolio if we include both Italian and Spanish government securities during a crisis.

Second, we show that the bivariate DCC is most suitable for measuring extreme intraday VaR. Specifically, we conduct a backtesting procedure, which involves a one-step-ahead forecast of the covariance matrix and compares the four most common ways to compute portfolio VaR, i.e. a historical simulation, a Constant Conditional Correlation (CCC) model, a bivariate DCC model, and a multivariate DCC which deals information in the aggregate using a composite likelihood approach. We find that the bivariate DCC model gives the most accurate exceptions when VaR is violated equal to or below 1% of the time. The multivariate DCC model comes in the second place in terms of unconditional and conditional accuracy. Our result can be useful for portfolio managers who try to monitor their portfolios' risk closely, especially during the debt crisis period. Furthermore, large banks may want to use bivariate DCC model to achieve capital efficiency. With so few studies devoted to intraday correlation and especially to bond correlations, we can help the portfolio managers to quantify the correlation risk when they try to control for the transaction costs of executing a portfolio. It is crucial not to sell two highly correlated assets at the same time as one's selling would probably trigger a fall of the price for the other.

Our work extends previous the literature that studies contagion phenomenon. Forbes and Rigobon (2002) restrict their definition of cross market contagion to a sudden increase of a bias-corrected unconditional correlation coefficient. The author find no contagion during 1997 Asian crisis, 1994 Mexican devaluation and 1987 U.S. market crash. Using the Dynamic Conditional Correlation (DCC) model of Engle (2002a), we can easily cure the bias and produce a much richer picture by using the time-varying conditional correlation as in, e.g. Chiang et al. (2007) and Dimitriou et al. (2013). These two papers both base their analysis on the fitted conditional correlation series.

They both find that there is a increased correlation between the U.S. market and other markets during the subprime crisis. Regarding bond correlations, Dungey et al. (2006) observe contagion between emerging and developed fixed-income markets when Russian bond defaulted and the Long-Term Capital Management (LTCM) collapsed. However, we find that there is no contagion during the sovereign debt crisis and the problem remain to the peripheral European countries.

We have extensively tested the accuracy of VaR computed from DCC model for an equally weighted portfolio. The seminal paper of Engle (2002a) evaluates the performance of various correlation specifications using the dynamic quantile test of Engle and Manganelli (2004) to compute the 5% and 1% VaR. The mean reverting scalar bivariate DCC appears to be a competitive model to compute the two VaRs but the empirical analysis does not cover unconditional risk coverage and an evaluation for the efficiency of capital allocation. Billio and Caporin (2009) extend the DCC model to reflect the different dynamics of correlations between each pair of assets while keep the dimensionality of the parameter space relatively low. We show that the DCC model can provide both unconditional and conditional risk coverage.

Our goal in the second part is similar to Giot (2005), who applies four univariate intraday volatility models to examine their performances on computing VaR based on a 5-month dataset from Trade and Quote (TAQ) containing 3 stocks. He fits the various models for the first 3 months and then generating one-step-ahead forecasts for the remaining 2 months. The fitting and forecasting are done for 15-minute and 30-minute returns, respectively. In this paper, we use a longer 2012–2013 sample of seven major European countries from the MTS dataset and the backtesting is conducted on the last year of the sample. We also construct a bond portfolio during the debt crisis period and the DCC multivariate GARCH model enables us to evaluate VaR on a portfolio level.

4.2 Data

We use MTS intraday quote data from European bond markets. The MTS market contains several separate platforms dedicated to trading country specific fixed income securities. All trading and quoting runs from 8:15–17:30 Central European Time (CET). The details of the markets and the dataset are documented in Dufour and Skinner (2004). We choose on-the-run 10-year sovereign bonds of Austria, Belgium, France, Germany, Italy, the Netherlands, and Spain based on the criteria described in Chapter 3. We filter the data using the optimal methods suggested by Chapter 3. Although many filters' result are not statistically different, we still choose the best ones, which will clean most of the noise and preserve as much data as possible. In order to build a multivariate series, we need to match the data of all the 7 markets according to the common opening time. We ascertain the time window when all markets actively update their limit order book each day. Our data records start from the time when all bonds have the first quote updates and stop at the time when one market ceases to post new quotes.[34]

The logarithmic returns are computed from mid-quote prices sampled every 10 minutes. We remove the first 45 minutes due to lack of available quotes so that the intraday data commence from 9:00 or later. There are 52 10-minute returns for each day with a few days being exceptions. On occasions, the first quote update of one market appears later than 9:00, which leads to loss of two-way quotes of other markets. We denote the

[34]If the stopping time is later than 17:25, we keep the later quotes of all markets until 17:30 for that day.

10-minute returns by $r_{t,n}$, where $t = 1, 2, \ldots, 1272$ and $n = 1, 2, \ldots, 52$. In total there are 65599 observations for each country in our dataset. Next we split the dataset into two parts. The first part covers from 02 January 2009 to 30 December 2011, which we use to fit the intraday correlation model and analyze the pattern of European bond correlations during the crisis time. The second part covers from January 02, 2012 to December 30, 2013, in which the last year is used to test the ability of various methods to compute intraday VaR.

4.3 Econometric Methodology

The econometric methodology closely follows Engle (2002a) and Engle et al. (2007). The entire model is estimated in separate steps where a component GARCH model is fitted first and the returns are subsequently normalized by the estimated conditional variances. In the second step, either the dynamic parameters of a bivariate model or a multivariate model including all seven European countries is estimated. The conditional mean equation for the univariate GARCH model of intraday log return is specified as:

$$r^p_{t,n} = \sqrt{s^p_n q^p_{t,n}} \varepsilon^p_{t,n} \text{ and } \varepsilon^p_{t,n} | \Phi_{t,n-1} \sim N(0, 1) \tag{4.1}$$

where s^p_n is the intraday periodicity or diurnal component

$q^p_{t,n}$ is the intraday variance component

$\varepsilon^p_{t,n}$ is an error term

$\Phi_{t,n-1}$ denote the set containing all the available information up to the preceding bin of the current time interval. Here the usual standard normal conditional distribution of the error term is assumed. The overnight return $r^p_{t,0}$ is excluded because we are interested in modelling the dynamics of sovereign bond correlation during the time interval when all markets are open. The diurnal component is modeled as an exponential linear spline and the intraday volatility follows a GARCH(1,1) process:

$$s_n^p = \delta_0^p * exp(\sum_{j=1}^{m} \delta_j^p (\Delta_n - k_{j-1})_+) \tag{4.2}$$

$$q_{t,n}^p = \alpha_0^p + \alpha_1^p \left(\frac{(r_{t,n-1}^p)^2}{s_{n-1}^p} \right) + \beta^p q_{t,n-1}^p + \gamma_1^p (pspread_{t,n-1}^p) + \gamma_2^p (pspread_{t,n-1}^p)^2 \tag{4.3}$$

$$\Delta_n - k_j = \begin{cases} (\Delta_n - k_j) & \text{if } \Delta_n > k_j \\ \\ 0 & \text{otherwise} \end{cases}$$

$$\Delta_n = \frac{n}{N} \quad n = 1, 2, \ldots, N$$

where k_j represents the knots of the linear spline and the distance between two consecutive knots is 1 hour. In contrast with Engle and Sokalska (2012), we remove the daily component from the return process. We want a good measure of intraday correlation and therefore simplify the model to focus on intraday correlation. Moreover, we do not need to assume $E(q_{t,i}) = 1$. In the original model of Engle and Sokalska (2012), the conditional variance of intraday return $r_{t,n}^p$ will finally converge to the unconditional variance $h_t^p * s_{t,n}^p$ because of the assumption $E(q_{t,i}) = 1$. Engle and Sokalska (2012) impose the assumption by forcing $\alpha_0 = 1 - \alpha_1 - \beta$ whereas we can directly estimate α_0, providing flexibility for the univariate GARCH model. We propose to include in the conditional variance equation factors that capture extreme illiquidity events. After experimenting with alternative specifications, we choose to include a quadratic function of lagged percentage spread in the conditional variance equation. In Chapter 3 we clearly see that a sudden change in the percentage spreads is associated with a big jump in returns. In order to prepare the data for the analysis, we use the same procedure presented in Chapter 3. However, we choose here to use intraday data of 2009-2011 which is not included in the evaluation of Chapter 3.[35] We find that the illiquidity still strongly influences the volatility estimation and may cause integrated conditional variance. Obviously volatility should not exhibit such high persistence due to liquidity shocks which quickly die out. It is shown that for some particular months even the 95th percentile of percentage bid-ask spreads can reach as high as several hundred basis points in Chapter 3. The liquidity effect must be controlled for and a quadratic functional form is chosen out of parsimonious concern. The parameter of the percentage spread is occasionally forced to take value zero in order to insure a positive conditional variance.[36]

The DCC part can be constructed in the following way. Vectors and matrices are denoted by bold symbols. The system of return equations is defined as:

$$\boldsymbol{r_{t,n}} = \sqrt{\boldsymbol{s_n} \circ \boldsymbol{q_{t,n}}} \circ \boldsymbol{\varepsilon_{t,n}} \tag{4.4}$$

where $\boldsymbol{r_{t,n}} = (r^1_{t,n}, r^2_{t,n}, \ldots, r^P_{t,n})'$ and $\boldsymbol{\varepsilon_{t,n}}$ contains P standardized residuals. $\boldsymbol{s_n}$, $\boldsymbol{q_{t,n}}$ contain the intraday periodicity and the univariate GARCH variance for P securities computed at time interval n. "$\circ$" denotes Hadamard product of equal size matrices. We assume a multivariate normal conditional distribution for the P securities. Specifically,

[35]Only the daily data extracted from intraday data is part of the filtering evaluation, which ignores the intraday variation in returns for this period.

[36]Bollerslev and Domowitz (1993) study the interaction between volatility and market activity using a similar specification

let $\boldsymbol{D_{t,n}} = diag(\sqrt{\boldsymbol{q_{t,n}}})$ and $\boldsymbol{S_n} = diag(\sqrt{\boldsymbol{s_n}})$. Then

$$\boldsymbol{r_{t,n}}|\Phi_{t,n-1} \sim N(0, \boldsymbol{H_{t,n}})$$

$$\boldsymbol{H_{t,n} = S_n D_{t,n} R_{t,n} D_{t,n} S_n} \tag{4.5}$$

$$\boldsymbol{\varepsilon_{t,n}} = (\boldsymbol{S_n^{-1}} \circ \boldsymbol{D_{t,n}^{-1}})\ (\boldsymbol{r_{t,n}}) \tag{4.6}$$

$$\boldsymbol{Q_{t,n}} = \boldsymbol{\Omega}(1-a-b) + a\boldsymbol{\varepsilon_{t,n-1}}\boldsymbol{\varepsilon'_{t,n-1}} + b\boldsymbol{Q_{t,n-1}} \tag{4.7}$$

$$\boldsymbol{R_{t,n}} = diag(\boldsymbol{Q_{t,n}})^{-1/2}\, \boldsymbol{Q_{t,n}}\ diag(\boldsymbol{Q_{t,n}})^{-1/2} \tag{4.8}$$

After taking out intraday periodicity and a volatility component, the conditional variance of the standardized residual $\boldsymbol{\varepsilon_{t,n}}$ is 1 for all $n = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$. Therefore, the conditional covariance matrix of $\varepsilon_{t,n}$ becomes the conditional correlation matrix of $r_{t,n}$. The quasi-correlation matrix $\boldsymbol{Q_{t,n}}$ follows an autoregressive process and then is normalized in Equation (4.8) so that the diagonal elements are exactly 1 for all $n = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$. Since there are $P(P-1)/2$ parameters in $\boldsymbol{\Omega}$ and only two dynamics parameters, $\boldsymbol{\Omega}$ is approximated by the sample correlation matrix of $\boldsymbol{\varepsilon_{t,n}}$. The approximation considerably reduces the dimensionality of the parameter space and makes the DCC model easy to estimate.

The normality assumption gives the likelihood function of the bivariate DCC model and the parameters are estimated by Gaussian quasi-maximum likelihood (QML).[37] The consistency of the estimator is proved in Engle and Sheppard (2001) given a mild regularity condition. The robust standard error is computed according to Newey and McFadden (1994) and Engle (2002a). The required gradient and Hessian matrix is

[37] We assign zero weight to the first observation every day in order to exclude effects from the previous day.

approximated by the central finite difference method.[38] Let $\boldsymbol{\lambda}, \boldsymbol{\theta}$, denote the parameter vector of GARCH part and DCC part, respectively. We have:

$$
\begin{aligned}
&\boldsymbol{r_{t,n}}|\Phi_{t,n-1} \sim N(0, \boldsymbol{H_{t,n}}) \\
&L = -\frac{1}{2}\sum_{t=1}^{T}\sum_{n=1}^{N}(\log(2\pi) + \log|\boldsymbol{H_{t,n}}| + \boldsymbol{r'_{t,n}}\boldsymbol{H_{t,n}^{-1}}\boldsymbol{r_{t,n}}) \\
&\quad = -\frac{1}{2}\sum_{t=1}^{T}\sum_{n=1}^{N}\log(2\pi) + \log|\boldsymbol{S_n D_{t,n} R_{t,n} D_{t,n} S_n}| + \boldsymbol{r'_{t,n} S_n^{-1} D_{t,n}^{-1} R_{t,n}^{-1} D_{t,n}^{-1} S_n^{-1} r_{t,n}} \\
&\quad = -\frac{1}{2}\sum_{t=1}^{T}\sum_{n=1}^{N}\log(2\pi) + 2\log|\boldsymbol{S_n D_{t,n}}| + \log|\boldsymbol{R_{t,n}}| + \boldsymbol{\varepsilon'_{t,n} R_{t,n}^{-1} \varepsilon_{t,n}} \\
&\quad = -\frac{1}{2}\sum_{t=1}^{T}\sum_{n=1}^{N}\log(2\pi) + 2\log|\boldsymbol{S_n D_{t,n}}| + \boldsymbol{r'_{t,n} S_n^{-1} D_{t,n}^{-1} D_{t,n}^{-1} S_n^{-1} r_{t,n}} \\
&\qquad \boldsymbol{-\ \varepsilon'_{t,n}\varepsilon_{t,n}} + \log|\boldsymbol{R_{t,n}}| + \boldsymbol{\varepsilon'_{t,n} R_{t,n} \varepsilon_{t,n}}
\end{aligned}
\tag{4.9}
$$

We can split the likelihood function into two parts:

$$L = L_{GARCH}(\boldsymbol{\lambda}) + L_{DCC}(\boldsymbol{\theta}, \boldsymbol{\lambda}) \tag{4.10}$$

$$L_{GARCH}(\boldsymbol{\lambda}) = -\frac{1}{2}\sum_{t=1}^{T}\sum_{n=1}^{N}\log(2\pi) + 2\log|\boldsymbol{S_n D_{t,n}}| + \boldsymbol{r'_{t,n} S_n^{-1} D_{t,n}^{-1} D_{t,n}^{-1} S_n^{-1} r_{t,n}} \tag{4.11}$$

$$L_{DCC}(\boldsymbol{\theta}, \boldsymbol{\lambda}) = -\frac{1}{2}\sum_{t=1}^{T}\sum_{n=1}^{N}\log|\boldsymbol{R_{t,n}}| + \boldsymbol{\varepsilon'_{t,n} R_{t,n} \varepsilon_{t,n}} - \boldsymbol{\varepsilon'_{t,n}\varepsilon_{t,n}} \tag{4.12}$$

Inverting the correlation matrix $\boldsymbol{R_{t,n}}$ at each step may become burdensome for a large number of securities. Engle et al. (2007) introduce the composite-likelihood (CL) method to the DCC models, which maximizes the summation of pairwise likelihood functions, instead of maximizing the full likelihood function and inverting a large matrix. They focus on the two DCC parameters and categorize others as nuisance

[38]We carry out the estimation in SAS/IML modules and the code is cross checked with the Matlab package 'MFE' written by Kevin Sheppard. The optimization is achieved by quasi-Newton algorithm.

parameters. And they show that the estimation of the target parameters is consistent and the standard error must be adjusted according to the influence of nuisance parameters. We estimate the bivariate scalar DCC model via the usual QML method and the multivariate scalar DCC ($P = 7$) via the CL method.

Specifically, we have in total $J = comb(P, 2)$ pairs of assets, where *comb* is the combinatorial function. The composite-likelihood estimator is:

$$\hat{\theta} = \arg\max \frac{1}{J} \sum_{j=1}^{J} \sum_{t=1}^{T} \sum_{n=1}^{N} \log L_{jtn}(\boldsymbol{\theta}, \hat{\boldsymbol{\lambda}}_{\boldsymbol{j}}, \widehat{\Omega_{j,12}}) \tag{4.13}$$

where $\boldsymbol{\theta}$ is the vector of the two DCC parameters and $\boldsymbol{\lambda_j}$ includes all the GARCH parameters for the jth pair.[39] $\Omega_{j,12}$ is the off-diagonal entry of the symmetric pairwise unconditional correlation matrix $\boldsymbol{\Omega_j}$ of the jth pair. $\boldsymbol{\lambda_j}$ and $\Omega_{j,12}$ are estimated based on some moment conditions: $g_{jtn}(\boldsymbol{\theta}, \boldsymbol{\lambda_j}, \Omega_{j,12})$. The moment conditions for the GARCH part are simply the first derivatives of the likelihood function with respect to all GARCH parameters:

$$E\left(\frac{\partial \log L_{jtn}(\boldsymbol{\theta}, \boldsymbol{\lambda_j}, \Omega_{j,12})}{\partial \boldsymbol{\lambda_j}}\right) = \mathbf{0} \tag{4.14}$$

and the moment condition for the unconditional correlation matrix is:

$$\Omega_{j,12} - \frac{1/(TN) \sum\limits_{t=1}^{T} \sum\limits_{n=1}^{N} \varepsilon_{1,j,t,n} \varepsilon_{2,j,t,n}}{\sqrt{1/(TN) \sum\limits_{t=1}^{T} \sum\limits_{n=1}^{N} \varepsilon^2_{1,j,t,n}} \sqrt{1/(TN) \sum\limits_{t=1}^{T} \sum\limits_{n=1}^{N} \varepsilon^2_{2,j,t,n}}} = 0 \tag{4.15}$$

where $\varepsilon_{i,j,t,n}$, ($i = 1$ *or* 2) represents the ith asset in the jth pair of assets. When modeling correlation, there is one important fact to consider, namely nonsynchronous

[39] Since we are concentrating on pairs of assets, the GARCH coefficients of both assets are in $\boldsymbol{\lambda_j}$.

trading. Lo and MacKinlay (1990) demonstrate that the difference in trading speed between two stocks can lead to spurious correlations. Lo and MacKinlay argue that the log returns are driven by the true returns, which are sometimes unobservable due to lack of trading activity. They develop a simple model where an indicator random variable is 1 with probability p_i if security i is not traded at time t and zero otherwise. They show that the cross correlation between two assets is determined by the probabilities of nontrading and the common factor that drives both returns. The induced correlation does however not vanish when the expected returns of any assets become zero. Therefore, if econometricians want to associate correlation changes with specific news or economic episodes, the contemporaneous cross correlation must be adjusted. In other words, the long-term trend needs to be inferred from the original correlation series. Burns et al. (1998) give a simple and powerful way to account for such an effect. They observe that there is a deterministic relationship in the daily opening and closing prices of one market relative to another. For example, information contained in daily closing price of the New York stock market will be impounded in Japanese stock market one day later. A first order vector moving average model is suggested to capture this lag effect. Burns et al. (1998) modify the nonsynchronous covariance matrix based on estimated vector moving average parameters.

We argue that the nonsynchronous trading effect in intraday bond data cannot be as easily controlled as when working with daily data. First, the lead and lag of opening and closing does not exist for European government bond markets: all markets open and close at the same time. All news can potentially affect all markets simultaneously. Second, quote updates for sovereign bonds arrive at irregular intervals. Differences in trading frequencies across bonds may affect the correlations. So it is unlikely that the effect of nonsynchronous trading can be captured by some deterministic parameters as in Burns et al. (1998). In order to controls for the asynchronous trading effect we suggest adopting an alternative approach in the spirit of Rangel and Engle (2012). In that paper, the authors estimate the long-term trend of correlation as a correlation between two macroeconomic-related components, which are modeled by quadratic splines. On the other hand, the cubic spline is widely used and is more flexible than the quadratic spline as it can have more than one turning point between two knots (see Wegman and Wright, 1983 for an early survey). Instead of fitting a spline along with a dynamic model, we choose to estimate a cubic spline with evenly spaced knots based on the correlation derived from a DCC model (see Table A.1 in the Appendix for the estimation of the model). The number of knots of the cubic spline is determined using the Bayesian information criterion (BIC). The approach is simpler and eliminates the needs of a multivariate spline but it still extracts the long-term trend from high-frequency data.

The *natural* cubic spline of the correlation $\rho_{t,n}^{i,j}$ between security i and j is fitted by partitioning the entire time interval.[40] The usual continuity constraints are enforced and the second derivatives at the first and last knot are zero. We choose the cubic spline with an optimal number of knots according to BIC information criterion. The number of knots for each spline varies from 2 to 30. We do not add any macroeconomic variables to the cubic spline. As suggested by Poirier (1973), the structural changes or business cycles can be inferred and tested from a cubic spline regression. However, under the crisis situation, the interpretation may be different for the change in spline parameters. There may not be a overall shift of economic environment associated when the curve changes its pattern. Rather we can conjecture that the market reacts to political or central banks announcements that lift the hope of solving the crisis.

[40] The cubic spline is estimated by using the built-in functions of SAS/IML.

4.4 Empirical Result

4.4.1 Summary Statistics

The mean of 10-minute returns in Table 4.1 is extremely close to zero and the standard deviations are qualitatively similar to those described in Chapter 3. However, the kurtosis for Spain is much larger than that of the intraday return described in Chapter 3, suggesting that there are some extreme movements in price. In fact the kurtosis of the entire series for Spain is driven by only a few observations belonging to December 2011. During this month the 95th percentile of percentage spread is above 2000 basis points and there are 6 10-minute returns that exceed 5% with an associated percentage spread rising to at least 300 basis points or even higher! Obviously none of the filters can easily identify them as outliers because most of the returns have high spreads[41] and price very often moves wildly during the month.[42] As an experiment, we remove the six extreme returns from the series and the kurtosis dramatically drops to 172 which is only one tenth of the current kurtosis.[43] Other candidates, who might have illiquidity issues are Italy and Belgium, which also have higher kurtosis. Hence it is rather crucial to control for the illiquidity effect along with efficient data cleaning procedures in GARCH estimation.

Figure 4.1: Diurnal Component $s^p_{t,n}$ for each country

The diurnal component is a linear spline: $s^p_n = \delta^p_0 * exp(\sum_{j=1}^{m} \delta^p_j (\Delta_n - k_{j-1})_+)$. The knots are set at the start of every hour (9:00, 10:00, ..., 16:00) and the last half hour from 17:00 to 17:30.

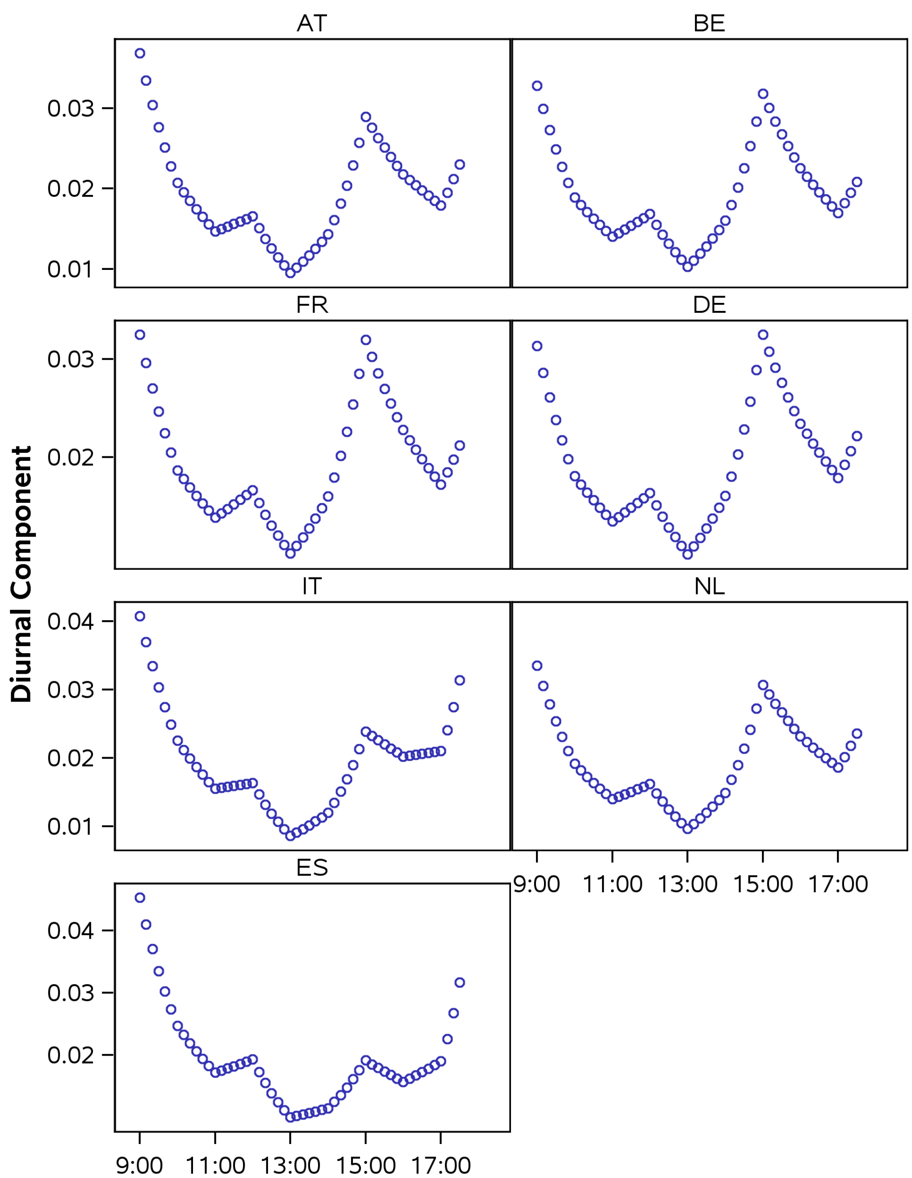


Table 4.1: Summary statistics of 10-minute intraday returns

The sample period lasts from January 02, 2009 to December 30, 2011. There are 32493 observations for each country. Spain has unusually large and positive skewness and extremely large kurtosis.

Country	Mean	St.d.	Skewness	Ex. Kurt.	Max	Min
AT	0.0003	0.074	−0.216	22.605	0.964	−1.332
BE	0.0004	0.080	−0.250	41.462	1.702	−1.751
FR	0.0005	0.057	−0.083	10.146	0.687	−0.768
DE	0.0006	0.056	−0.443	25.848	1.115	−1.320
IT	0.0004	0.078	−2.337	118.20	1.541	−2.752
NL	0.0007	0.054	−0.057	7.551	0.703	−0.736
ES	0.0003	0.132	4.130	2112.6	9.040	−9.702

4.4.2 Intraday GARCH Estimation

We present the estimation of the main GARCH parameters in Table 4.2. The robust t-statistics of quasi-maximum likelihood estimation are reported in parentheses. Although the t-test does not show any significance for the liquidity parameters, the null hypothesis that both coefficients are zero is rejected at the 1% significance level in the likelihood ratio test. The likelihood ratio test also testifies a great improvement of our conditional variance specification in fitting the data. Notice that the persistence parameter of Belgian bonds is much lower than the one reported in Table 3.9 when adding the liquidity variables. Instead of displaying the diurnal parameters, the diurnal com-

[41] The mean of the percentage spread of December 2011 is around 200 basis points and the standard deviation is around 300 basis points.

[42] The excess kurtosis of Spanish bond returns is 137 for the month

[43] If we removed the entire month from our sample, the kurtosis would drop further, to 91.

Table 4.2: Intraday GARCH estimation

The intraday 10-minute return series are constructed from bid and ask quotations of European government bonds from January 02, 2009 to December 30, 2011 obtained from the MTS dataset. The results are based on 39423 10-minute returns. 40 lags are chosen for the Newey-West standard error. *pspread* is the percentage spread computed from the best bid and ask. The models are:

$$r_{t,n}^{p} = \sqrt{s_n^p q_{t,n}^p}\varepsilon_{t,n}^p$$
$$s_n^p = \delta_0^p * exp(\sum_{j=1}^{m} \delta_j^p (\Delta_n - k_{j-1})_+)$$
$$q_{t,n}^p = \alpha_0^p + \alpha_1^p \left(\frac{(r_{t,n-1}^p)^2}{s_{n-1}^p}\right) + \beta^p q_{t,n-1}^p + \gamma_1^p (pspread_{t,n-1}^p) + \gamma_2^p (pspread_{t,n-1}^p)^2$$

The likelihood ratio (LR) test of the null hypothesis $H_0 : \gamma_1^p = 0$ and $\gamma_2^p = 0$ or $H_0 : \gamma_2^p = 0$ is shown in the last column. The 1% $\chi^2(2)$ critical value is 9.21.

Country	α_0	α_1	β	γ_1	γ_2	Log Likelihood	LR test
AT	0.0008	0.0282	0.9667		0.0012	53601.85	148.16
	(0.1174)	(5.0675)	(132.2298)		(0.1228)		
BE	0.0137	0.1316	0.6880	0.0258	0.1740	53853.81	1192.58
	(0.7708)	(5.7164)	(8.3588)	(0.7350)	(0.7769)		
FR	0.0025	0.0468	0.9210		0.0485	60095.27	219.85
	(1.1335)	(2.9668)	(24.7679)		(1.0682)		
DE	0.0009	0.0325	0.9578	0.0064	0.0158	60234.87	182.19
	(0.6526)	(4.4586)	(87.2412)	(0.5768)	(0.6212)		
IT	0.0015	0.1151	0.8087	0.0231	0.2665	60016.18	1184.28
	(0.0521)	(1.4381)	(5.9967)	(0.0482)	(0.0498)		
NL	0.0005	0.0197	0.9751	0.0014	0.0028	61137.78	69.52
	(0.2511)	(5.3725)	(184.9905)	(0.2307)	(0.2473)		
ES	0.0012	0.0298	0.9568		0.0162	51639.31	1251.08
	(2.0645)	(7.3354)	(114.4685)		(1.7080)		

ponent is plotted in Figure 4.1. The patterns of the 7 European countries are consistent with the ones shown in Figure 3.6 though the magnitude of the diurnal component of Italy and Spain are now smaller. This may well be due to the removal of daily variance component and different sample periods that the two GARCH estimation are based on. Table 3.9 and Figure 3.6 are derived from the data spanning from April 02, 2012 to December 30, 2013. The estimation of DCC parameters are shown in Appendix A.

4.5 Cubic Spline of Fitted Correlation Series

When modeling volatility, we must consider the impact of the Securities Purchase Program (SMP) initiated by the European Central Bank (ECB) in 2010 (see Chapter 3 for the details of SMP). The first round of purchases covered Greek, Irish, and Portuguese government debts, whereas the second round focused on Italian and Spanish government bonds. The ECB only published a weekly aggregate in the purchasing amounts, without giving the details of these transactions. We do not have data on specific purchases however we do need to study the SMP effect which has been documented in the paper by Ghysels et al. (2014). The authors examine the SMP's effect on daily and intraday volatility using confidential data from the ECB. They show a change in the innovation parameter in a multiplicative GARCH model.

The two gray areas in Figure 4.2 represent the periods when the ECB initiated the SMP to purchase distressed countries' government bonds. Unlike the clear identification of different phases of the U.S. subprime crisis (see Federal Reserve of St. Louis (2009) for a detailed description of the crisis timeline), there is no consensus on how to categorize the progression of the European debt crisis into different periods. According to the findings of Chapter 3, the ECB's interventions heavily influenced the volatility of

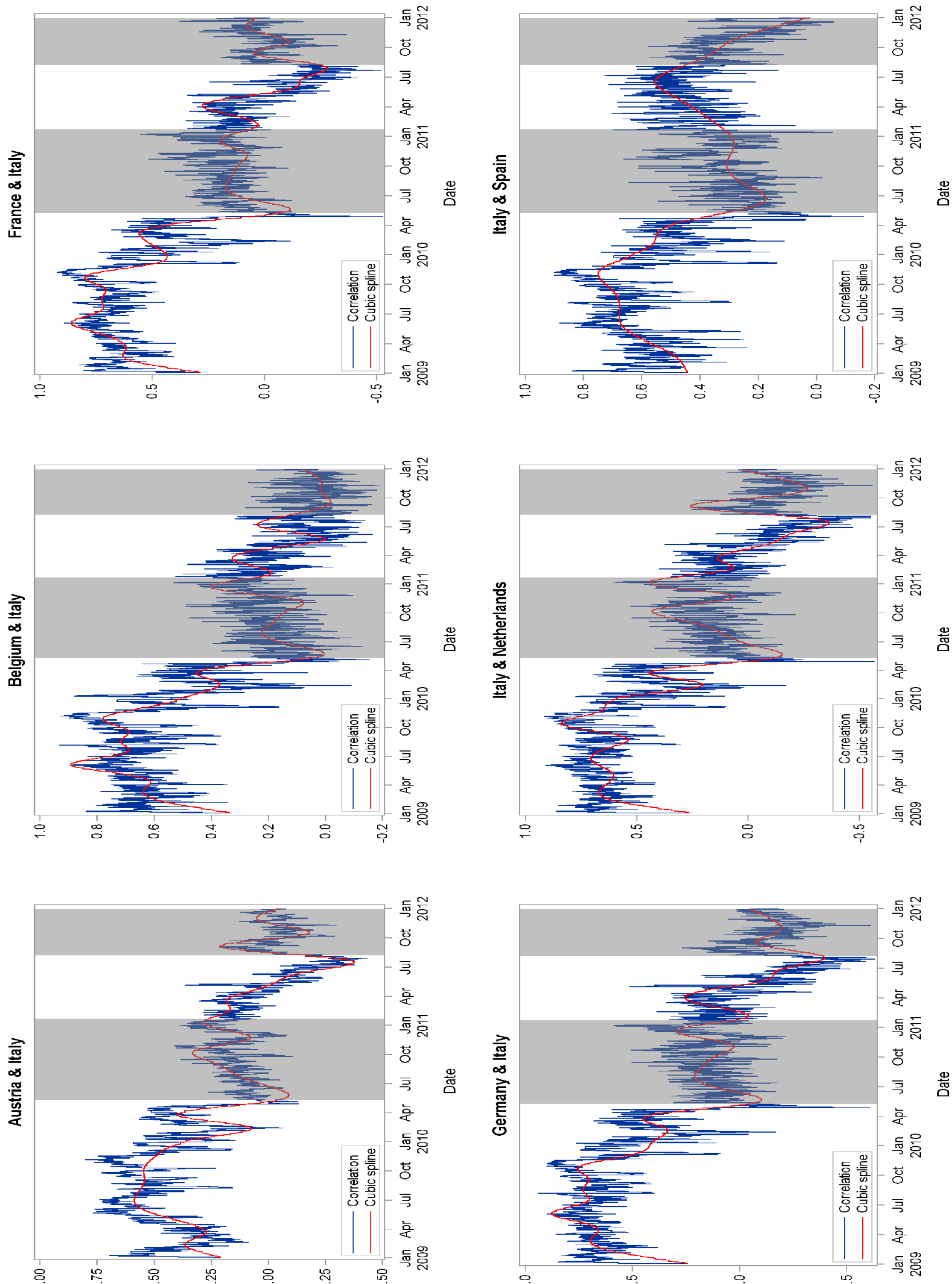


Figure 4.2: Fitted correlation and cubic spline

benchmark bonds under selling pressures. The dynamics of Italian and Spanish government bonds have completely changed–the persistence of volatility is much lower than in the no-intervention period and the shocks have much greater impacts, suggesting that investors attached more importance to macro announcements. Hence we are interested in the behavior of correlations between two countries during and outside of the ECB's bond purchasing program.

The correlations of Figure 4.2 indicate a divergence of European governments from 2009 to 2011. From January to October 2009, we observe a strong, positive (above 0.5) correlation between German and Italian government bond returns. However, during the last quarter of 2009 and the first quarter of 2010, as the confidence in the Italian government bonds deteriorates, the correlation starts decreasing. In April 2010, we observe a sharp drop in the estimated conditional correlation reaching highly negative values. Just before the beginning of the first ECB's securities market programme, the correlation recovers and stays just above zero for the duration of the programme. Once the ECB's support ceases the correlation decreases sharply to negative values. With the second phase of the SMP, the correlation increases again towards zero. The empirical evidence seems to support the effectiveness of both ECB interventions in restoring confidence in the Italian sovereign market. Additionally, when the ECB decided to buy Italian and Spanish bonds, their correlation fell all the way down to almost zero from their pre-purchase level. Clearly the purchase decoupled the Italian bonds from the Spanish ones.

To test whether there is any significant change in correlation size, we run a regression of computed dynamic correlations on the two dummies corresponding to the two periods of bond acquisition, and an additional dummy to control for the temporary break of

Table 4.3: Regression estimation of equation (4.16)

The robust t-values are reported in parenthesis. 40 lags are included in Newey-West standard error correction. The regression is $\rho^{i,j}_{t,n} = \rho_0 + \beta_1 I_1 + \beta_2 I_2 + \beta_3 I_{break} + \eta^{i,j}_{t,n}$

Country	ρ_0	β_1	β_2	β_3
Austria & Italy	0.4420	−0.3115	−0.4795	−0.4565
	(53.5750)	(−30.4882)	(−44.6048)	(−29.6097)
Belgium & Italy	0.5853	−0.3889	−0.5445	−0.4387
	(65.7422)	(−35.6034)	(−48.5850)	(−34.5753)
France & Italy	0.5820	−0.4150	−0.5775	−0.5921
	(58.3539)	(−34.5178)	(−41.7462)	(−34.6105)
Germany & Italy	0.5795	−0.4423	−0.7091	−0.6454
	(56.6586)	(−33.5257)	(−47.9383)	(−33.1022)
Italy & Netherlands	0.5764	−0.4162	−0.6799	−0.6161
	(58.8070)	(−32.8644)	(−47.3701)	(−34.9603)
Italy & Spain	0.5795	−0.2804	−0.2931	−0.1212
	(74.8356)	(−26.2311)	(−22.0266)	(−10.9794)

the program[44]:

$$\rho^{i,j}_{t,n} = \rho_0 + \beta_1 I_1 + \beta_2 I_2 + \beta_3 I_{break} + \eta^{i,j}_{t,n} \tag{4.16}$$

Table 4.3 describes the unconditional changes in correlations during the SMP in contrast with the conditional variation in Figure 4.2. Although the trend was moving upward, the overall level of correlation fell significantly. The correlation reduction was

[44]The two dummies take value 1 whenever the weekly report of purchasing amount is positive and take value 0 otherwise.

stronger when it resumed after a long inactivity. The average correlation between other countries and Italy became very small and even negative towards the end of 2011. This phenomenon is in contrast with the definition of contagion in Forbes and Rigobon (2002). There was no apparent elevated correlation during 2010 and 2011, when the crisis was most severe. Furthermore, it is valuable to include riskier bonds such as Italian and Spanish bonds in order to provide stronger diversification effect.

4.6 VaR Analysis

4.6.1 VaR Methodology

In this section we evaluate the performance of the bivariate DCC model for forecasting Value-at-Risk against three alternative approaches. In particular, we conduct a backtesting procedure (see Campbell (2005) for a review of backtesting methods) which involves forecasting the covariance of a portfolio. The methods from the simplest to the most sophisticated are: historical VaR, the constant conditional correlation (CCC) model, the bivariate DCC model, and the multivariate (in our case, we have 7 variables of interest) DCC model with composite likelihood. We construct an equally weighted portfolio containing 7 benchmark bonds with rebalancing at the end of each month. The portfolio variance is $\omega' H_{t,n} \omega$. We agree that the DCC model can be used to minimize portfolio variance by choosing weights based on the model generated correlation, but our objective here is to test the DCC model's ability to compute an adequate VaR measure for a given portfolio. We generate a one-step-ahead forecast of volatility and correlation based on different processes assumed by the models and then we compare the VaR performance based on a series of statistical tests. We adopt a dynamic sampling scheme: the models are re-estimated at the end of every month in 2013 using the sample of the last 12 months.[45] The one-step ahead forecasts for the subsequent month assume that the parameters are fixed at the estimated values. For example, let $q_{t,n}^{f,p}$ denote the forecast for $q_{t,n}^{p}$ in Equation 4.3. Then:

$$q_{t,n}^{f,p} = \hat{\alpha_0}^p + \hat{\alpha_1}^p \left(\frac{(r_{t,n}^p)^2}{\hat{s_{n-1}}^p} \right) + \hat{\beta}^p q_{t,n-1}^{f,p} + \hat{\gamma}_1^p (pspread_{t,n-1}^p) + \hat{\gamma}_2^p (pspread_{t,n-1}^p)^2 \quad (4.17)$$

where $r_{t,n-1}^p$ and $pspread_{t,n-1}^p$ are the 10-minute return and percentage bid-ask spreads from the previous 10-minute interval, respectively. $\hat{\alpha_0}^p$, $\hat{\alpha_1}^p$, $\hat{s_{n-1}^p}$, $\hat{\beta}^p$, $\hat{\gamma_1}^p$, $\hat{\gamma_2}^p$ are the estimated coefficients of the univariate GARCH model.[46]

The 2-year data has 13120 observations for 2012 and 13056 observations for 2013. The forecasted VaRs for intraday return $r_{t,n}$ from three statistical models are computed as:

$$VaR_{t,n}(\alpha) = -F_{t,n}^{-1}(\alpha) * Vol_{portfolio}^{f} \quad (4.18)$$

where $F_{t,n}^{-1}(\alpha)$ is the $1-\alpha$ percentile of the normal cumulative distribution function of volatility standardized return. $Vol_{portfolio}^{f}$ stands for the forecast of portfolio variance from different models. Therefore we tend to update $F_{t,n}^{-1}(\alpha)$ every 10 minutes by looking back exactly 1 year in order to maximize the information incorporated in the VaR. The historical simulation generates VaR by computing the corresponding percentiles using the sample in the same fashion as the realized distribution function.

Kupiec (1995) suggests a test for evaluating the adequacy of the VaR measures: the Proportion of Failure (PF) test on the hypothesis that the required violation frequency

[45] The rolling estimation frequency is chosen due to limited computing power.

[46] The first forecast of each day is generated from the last observation of the previous day.

is achieved. The test assumes that the violation of VaR follows a binomial distribution and the test statistics asymptotically follow the chi-square distribution with 1 degree of freedom. The violation of VaR can be expressed by a "hit" function, i.e.:

$$I_{t,n}(\alpha) = \begin{cases} 1 & \text{if } r_{t,n} < -VaR_{t,n,\alpha} \\ 0 & \text{otherwise} \end{cases}$$

$$I(\alpha) = \sum_{t=1}^{T} \sum_{n=1}^{N} I_{t,n}(\alpha) \tag{4.19}$$

$$\hat{\alpha} = I(\alpha)/TN \tag{4.20}$$

The PF likelihood ratio test statistic is:

$$LR_{PF} = 2[\log(\hat{\alpha}^{I(\alpha)}(1-\hat{\alpha})^{TN-I(\alpha)}) - \log(\alpha^{I(\alpha)}(1-\alpha)^{TN-I(\alpha)})] \tag{4.21}$$

The PF test has the property that if $\hat{\alpha} = \alpha$ then the test statistic is zero. However, Christoffersen (1998) argues that counting exceptions and performing unconditional coverage tests cannot fully validate a VaR measure. The independence of VaR exceptions is also important, in the sense that the persistence of VaR exceptions, if any, indicates that the VaR measure does not cover market risk exposure properly. A VaR measure that gives a correct coverage on average may fail to do so in any particular period, thereby reporting excessive losses. If volatility is clustered, then a non-adjusting VaR will give too few exceptions in tranquil times and too many exceptions in volatile periods. Correspondingly, banks using such VaR measures would set aside too much capital when there are not many losses and too little when losses happen very often.

This not only proves that one VaR measure is conditionally inadequate, but also implies the inefficiency of using capitals and large opportunity cost. To gauge the conditional coverage we use the Dynamic Quantile (DQ) test of Engle and Manganelli (2004), which has been proved to perform well in many cases by Berkowitz et al. (2009). Specifically, it is an F test conducted in a regression of $I_{t,n}(\alpha) - \alpha$ on multiple lags of the dependent variable, the current VaR and other explanatory variables to test the hypothesis that all coefficients including the intercept are zero. Here we use 6 lags and the current VaR as regressors. To measure losses, we propose the following function which evaluates the distance between a VaR measure and the realized return at exceptions:

$$Lf_L = \sum_{t=1}^{T} \sum_{n=1}^{N} (r_{t,n} - VaR_{t,n}) * I_{t,n}(\alpha) \tag{4.22}$$

Compared to the loss function of Lopez (1998), who studies exceptions and the magnitude of losses at the same time, we do not include the former in our loss function. Also there seems no reason to choose quadratic form as it will put more emphasis on larger losses. All returns and VaRs are measured in basis points in Equation (4.22).

4.6.2 Interpretation of VaR Backtesting

Bivariate and multivariate DCC tend to report higher than theoretical exception times according to Panel A of Table 4.4 at almost all percentage levels, whereas historical simulation is a more conservative measure for lower than 1% (inclusive) VaR. Billio and Caporin (2009) find that the CCC model gives almost identical results to DCC model in a daily VaR exercise. However, the CCC model completely breaks down for forecasting intraday VaR in Table 4.4, which provides strong evidence that time-varying correlation is a crucial factor in portfolio risk management, especially in intraday operations. The

Table 4.4: Evaluation of VaR measures

The bold text indicates non-rejection of the null hypothesis.

Model	0.1%	0.25%	0.5%	0.75%	1%	2.5%	5%	7.5%	10%
Panel A: Exception times									
Theoretical	13	32.64	65.28	97.92	130.56	326.4	652.8	979.2	1305.6
Historical	9	31	56	89	116	346	697	1060	1435
CCC	22	63	115	172	216	514	967	1401	1777
Bivariate DCC	11	34	73	111	152	363	748	1110	1478
Multivariate DCC	11	33	69	118	156	379	761	1132	1493
Panel B: Kupiec (1995) PF likelihood ratio test. The $\chi^2(1)$ critical value of 5% significance level is 3.84									
Historical	**1.417**	**0.084**	**1.393**	**0.845**	**1.704**	**1.184**	**3.085**	7.033	13.851
CCC	5.077	22.208	30.987	46.055	47.174	94.393	139.550	174.999	171.939
Bivariate DCC	**0.343**	**0.056**	**0.884**	**1.687**	**3.378**	4.064	13.986	18.164	24.366
Multivariate DCC	**0.343**	**0.004**	**0.209**	3.893	4.713	8.272	17.964	24.634	28.701
Panel C: Engle and Manganelli (2004) DQ test. The $F_{8,13048}$ critical value of 5% significance level is 1.94.									
Historical	20.753	9.750	15.862	13.139	12.292	39.883	46.285	45.012	42.822
CCC	3.621	3.919	5.762	7.289	10.122	20.188	28.648	32.171	29.276
Bivariate DCC	**0.074**	**1.248**	**1.479**	**1.061**	**1.572**	6.647	9.582	14.784	11.991
Multivariate DCC	**0.115**	**1.346**	**0.779**	**1.575**	2.075	4.975	9.772	12.305	11.666
Panel D: Loss function in basis points									
Historical	43.268	101.812	173.627	263.051	339.431	812.665	1557.47	2288.49	3020.95
CCC	65.625	135.839	247.711	349.884	442.637	994.502	1800.26	2547.98	3258.94
Bivariate DCC	40.504	92.540	164.762	232.679	300.335	725.912	1376.60	1990.25	2637.92
Multivariate DCC	41.527	92.321	161.516	238.883	306.067	736.161	1386.20	2006.25	2657.50

multivariate DCC model gives qualitatively similar results as the bivariate DCC model except for 1% VaR . The slightly inferior performance of the multivariate scalar DCC model may be due to the restriction imposed on the correlation process. It assumes that all correlation has the same dynamic patterns and each correlation is only affected by its own shocks. The bivariate scalar DCC, however, does not impose the former assumption and allows different parameter values for each correlation. It may be advisable for researchers to allow a more complex structure of the DCC model, such as in Billio and Caporin (2009) when dealing with a large number of assets in aggregate. For all of the percentage levels at which DCC can provide adequate unconditional coverage, it can also provide correct conditional coverage as reported by Panel C. Its VaR violations are evenly spread out during the test sample period which can be seen in Figure 4.3. Although historical VaR generally performs quite well in terms of PF test, the ”hit” of the VaR is clustered based on DQ test and Figure 4.4. It is apparent that the violations of the historical VaR concentrate in the middle of the year when volatility is high. Panel D of Table 4.4 indicates that the difference of losses between the historical VaR and the bivariate DCC VaR could be up to 30 basis points for 1% VaR. Thinking of a portfolio worth €1million, the efficiency loss is 3000. Combining with the fact that historical VaR has fewer exceptions than DCC VaR, the inefficiency is quite severe.

4.7 Conclusion

We fit a bivariate DCC model in order to study the correlation between European government bonds in the face of sovereign debt crisis. The correlations decreased for the most of the time when the ECB started buying distressed countries’ debt. It should be noted that the decreased correlation cannot be seen as caused by the ECB.

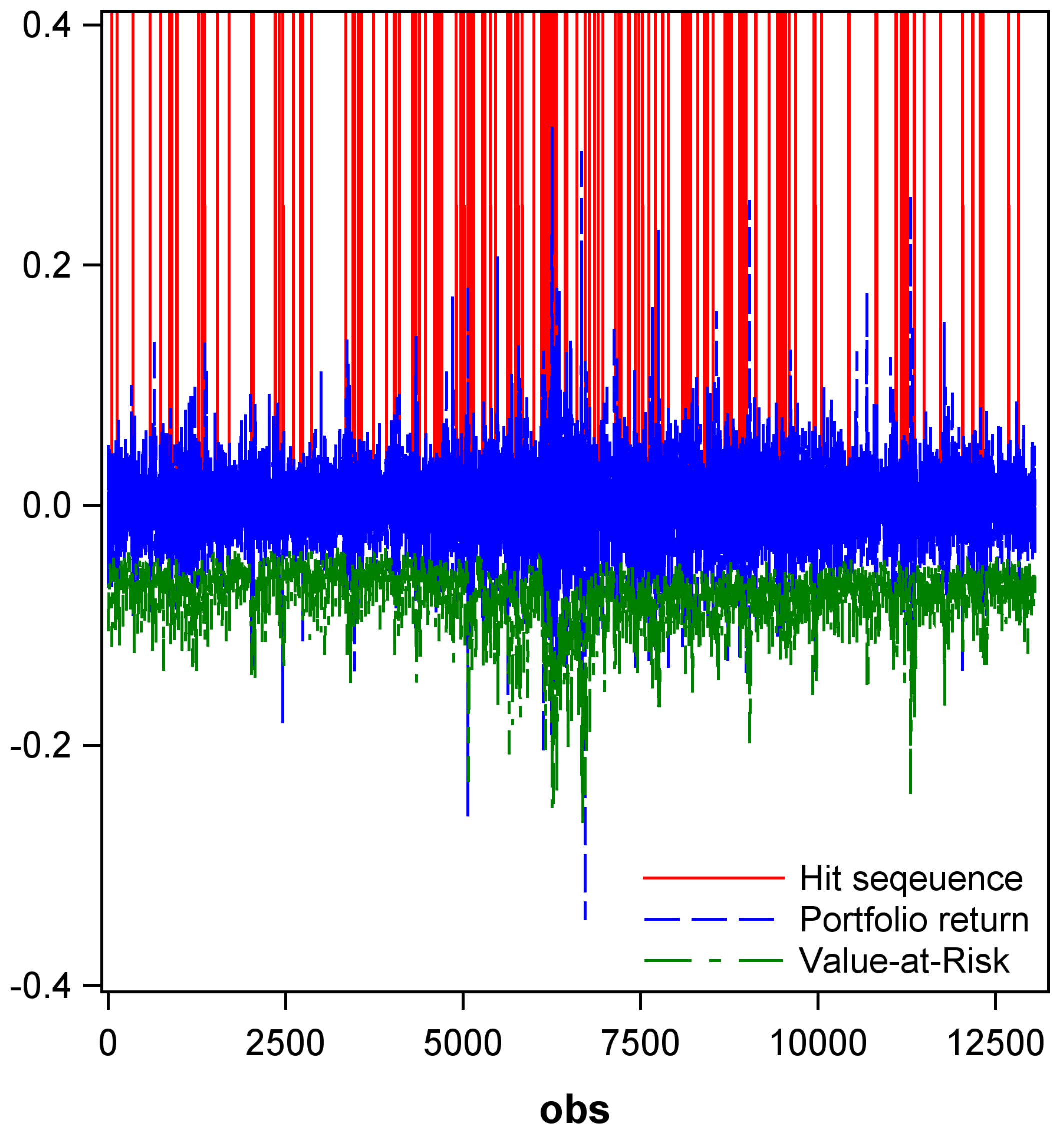


Figure 4.3: Value-at-Risk plot generated by the bivariate DCC model

Hit sequence $I_{t,n}(\alpha)$ takes values 1 when $r_{t,n} < VaR_{t,n}(\alpha)$. The y axis stops at value 0.4 to save space.

The drop of the correlation happened before the ECB intervened the markets. The lowered correlations suggest there was no contagion happened during the crisis and bond portfolios can enhance the diversification effect by including Italian and Spanish

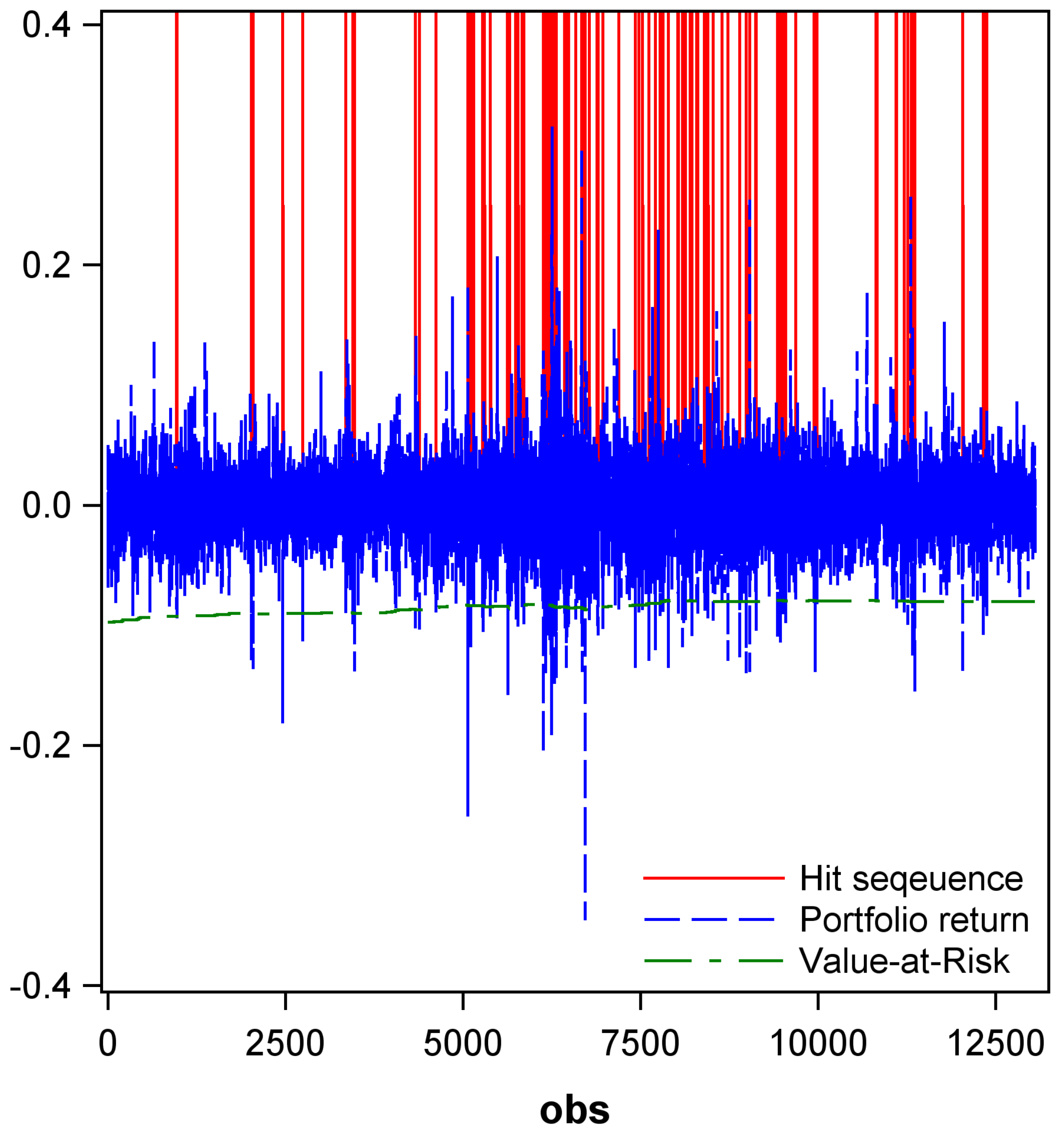


Figure 4.4: Value-at-Risk plot generated by historical simulation

Hit sequence $I_{t,n}(\alpha)$ takes values 1 when $r_{t,n} < VaR_{t,n}(\alpha)$. The y axis stops at value 0.4 to save space.

government bonds. To further control for the variance risk of a bond portfolio, we compare the adequacy of various VaR measures derived from correlation models and simple historical simulation. Although historical simulation sometimes performs well on average, the bivariate scalar DCC model provides accurate and serially independent VaR measures for including and lower than 1% percentage levels. In addition, the bivariate DCC model yields lower losses at exceptions and release more capitals for investment.

Chapter 5

From a Quote-Driven to an Order-Driven market: The Case of the EuroMTS Government Bond Trading Platform

5.1 Introduction

Over the last 8 years, starting from 2008, the European sovereign debt markets have experienced a series of shocks and regulatory interventions which have severely affected the dynamics of liquidity (see Fender and Lewrick, 2015 for an overview of the development of European government bond liquidity). In order to defend the creditworthiness of the GIIPS countries (Greece, Ireland, Italy, Portugal, Spain), the European Central Bank has implemented several policies to facilitate the functioning of secondary bond markets. Hence, a variety of market designs have been adopted for electronic trading platforms to promote liquidity and we have seen the rise of a hybrid market structures, where ordinary market participants compete with professional dealers to provide liquidity. There are designated market makers for some, if not all, stocks in the LSE, NASDAQ, NYSE, Euronext-Paris and Xetra, which are actually order-driven markets (see e.g. Aitken et al., 2009). Dating back to Glosten (1989), ample empirical evidence and theoretical arguments have shown the value of the designated market makers in auction type markets. However, the popularity of the mechanism in bond market has not been examined adequately, which bears great significance when bond market liquidity is in distress. For example, the U.S. Treasury market exhibited some very extreme price movements on October 15, 2014. The Federal Reserve Report (2015) reveals that algorithmic traders and large banks withdrew from the market during that time, resulting in insufficient liquidity for 10-year US Treasury bonds. To avoid such issues happening in Europe, the European Securities and Market Authorities (ESMA) is working to implement more rigid rules on the behaviour of algorithmic trading firms in MiFiD II.[47]

MTS is a leading electronic trading platform for European fixed-income securities with average daily volumes above 100 billion Euros. There are many local (domestic) segments of MTS and an international market called EuroMTS, where European benchmark government bonds are traded. On November 15, 2012 there was an important change to the EuroMTS rule book. From this date onwards, all market participants (including those previously defined as price takers) were allowed to submit one-sided limit orders to the market. Hence, this date signed the transformation of EuroMTS from a quote-driven market to an order-driven market in which all participants compete

[47] See `https://www.esma.europa.eu/policy-rules/mifid-ii-and-mifir`

to provide liquidity to more aggressive traders.[48] Before the rule change, only market makers known as primary dealers could submit limit orders (see Dufour and Skinner, 2004 for an overall introduction of the EuroMTS market prior to the change). To clarify the terminology involved, we will describe some of the institutional details of the EuroMTS market.

5.2 Institutional Details

There are two types of dealers in the MTS inter-dealer markets: ”primary” dealers and ”ordinary” dealers. Primary dealers act as designated market makers with obligations for the bonds allocated to them, whereas ordinary dealers were price takers before November 15, 2012. Trading in the EuroMTS market starts from 8:15 CET and ends at 17:30 CET. Market makers for some bonds can also quote for other bonds, for which they have no obligations. Moreover, primary dealers have the discretion to reject market orders whose quantities are below the minimum tradable quantity set by the MTS company. All quotes are centrally managed in the order book and the visible quantity for each quote is specified by the primary dealer who usually quotes on both EuroMTS and the local MTS platform at the same time. There are many local MTS platforms, each of which corresponds to the debts of one European country. The trading of one benchmark bond, thus, can be implemented in the EuroMTS or one of the local MTS platform. The parallel quotes, which have the same price but possibly different sizes, can only be hit once. Generally three types of orders exist in the MTS markets: two-sided limit orders (buy and sell at the same time), one-sided limit orders (only

[48] The decision is confirmed in the data as there was a trade of a short-term German bond (ISIN Code: DE0001137362) initiated by a trader using an one-sided limit order in the EuroMTS on November 19, 2012.

buy or sell) and market orders.[49] The submission of two-sided limit orders is still a privilege to primary dealers. The local MTS markets are not affected by this rule change. Primary dealers and ordinary dealers post anonymous limit orders. Hence, we cannot see whether the bid or ask is contributed via a two-sided or a one-sided limit order from the data. In order to be consistent with the literature, from now on we will only use the word ”dealers” to refer to the primary dealers. Others might be referred as ”traders” and ”market participants”. The old quote-driven market resembles the traditional dealership markets, which have been proven to provide less liquidity and higher transaction costs by many earlier studies.[50]

5.3 Related Literature

Our paper is related to other studies examining the effects of micro-structure and institutional changes on the quality of markets. Early papers focus on assessing the impact of the introduction of electronic trading platforms with limit order books. Amihud et al. (1997) investigate the process of Tel Aviv Stock Exchange have gradually moved from a call auction system to a continuous trading system. The authors present a significant increase of trading volume and liquidity ratio for more than 70% of the transferring stocks, which results in price appreciation. Muscarella and Piwowar (2001) derive similar results for the Paris Bourse, transferring the frequently traded stocks from call trading to continuous trading and moving the less traded stocks in the reverse direction. They demonstrate that while continuous trading increases the Cumulative Abnormal Return (CAR) and liquidity for the stocks, the call market does the opposite for stocks

[49] Other types of limit orders are also available, e.g. iceberg orders, fill-or-kill orders.

[50] There is a series of papers evaluating the liquidity for stocks switching from the NASDAQ to the NYSE (see Christie and Schultz, 1994, Barclay, 1997, Bennett and Wei, 2006). All the existing research concludes that stocks have higher liquidity when trading on the NYSE.

with low trading frequency. Henke and Lauterbach (2005) explore the same issue and proves the existence of liquidity growth when, instead of exchange, listed firms propose switching the trading mechanism to ensure continuous trading in the Warsaw stock exchange.

The introduction of the SETS orderbook and the SETSmm trading in the London Stock Exchange (LSE) has attracted much academic attention. Gajewski and Gresse (2007) illustrate that bid-ask spreads are higher and that depth is also larger in the SETS market than in the Euronext Paris market. This seems to indicate that retail trading is cheaper in Paris with tighter spreads but institutional trading is more convenient in London with larger depth. Chelley-Steeley and Skvortsov (2010) indicate the existence of volume enlargement and illiquidity ratio reduction after the SETSmm trading mechanism was introduced. Gregoriou (2015) also finds that transaction costs of illiquid stocks are reduced with the introduction of the AlM electronic platform to the LSE.

The findings of other markets also seem to favor auction type markets in promoting liquidity. Barclay et al. (1999) demonstrate that quoted and effective spread were both lower without depressing the depth, after the SEC required a limit to orders of normal customers to be displayed in the NASDAQ. Nimalendran and Petrella (2003) suggest that a hybrid order-driven market with the introduction of specialists for illiquid stocks offers greater liquidity by testing the data of the ISE. More recently, Anand et al. (2009) highlight that professional liquidity providers paid by companies can reduce the stocks' quoted bid-ask spreads in the Stockholm Stock Exchange.[51]

[51] Other important topics include: Chordia et al. (2014) studying the effect of decimalization of the NYSE and NASDAQ on liquidity; Jain et al. (2008) comparing of the market quality before and after the implementation of the Sarbanes-Oxley Act; and Boehmer et al. (2005) testing the effect of the increased pre-trade transparency on the quality of the market after NYSE started providing order book information to traders.

In general, we show that the rule change for the EuroMTS market leads to an increase in the level of market liquidity. We consider several dimensions of liquidity: bid-ask spread, depth and immediacy (measured as the proportion of time with sufficiently tight bid-ask spreads). An event study approach is implemented by the nonparametric Wilcoxon signed-rank test and a full regression with several control variables. The daily time-weighted percentage bid-ask spread shows the first sign of promotion in liquidity both in the nonparametric test and in the regression, i.e. the spread has decreased since the introduction of the new rule. Regarding the depth of the markets, the increment is not universal, as dealership markets may provide larger depth (see above and Gajewski and Gresse, 2007). Due to the fact that market makers and limit order users exercise more of the options that they can specify a very small amount, the depth at the best quotes of long-term bonds has been reduced significantly for some of the European countries as it happened in the conversion of the NASDAQ, found by Barclay et al. (1999). Here, we will go deeper into the further tiers of the order book and examine the percentage spread associated with 10 million bonds, which is shown to decrease on average. We develop a further measure of liquidity in order to study the immediacy offered by the market for large sizes. Similar to Hodrick and Moulton (2009)'s emphasis on the importance of looking at price, quantity and timing of the trade simultaneously, we define the immediacy of the market as the time length during which people can trade at a percentage spread, less than the maximum daily average time-weighted percentage spread, plus the 3 maximum standard deviation of the daily spread.

In addition to the operation we have just performed, we run a series of robustness checks. First, we restrict the sample to on-the-run bonds, while the conclusion stays unchanged. We also test the main empirical findings using a difference-in-difference approach where the liquidity of the domestic MTS platform is used as the control group. When differencing the liquidity measures between the EuroMTS platform and the domestic platforms, the results are somewhat mixed. The assumptions (independence of the two platforms) which the methodology relies upon are not satisfied by our data because of the parallel quoting for primary dealers (Darbha and Dufour, 2013). Therefore, the difference-in-difference approach may not properly reflect the liquidity conditions before and after the event.

Our research makes three main contributions which are important to academics and practitioners. First, we study and quantify the effect on several liquidity dimensions of an important change in the microstructure of the MTS markets. The structural change we consider marks the transformations of one of the MTS platforms from a dealer-driven system to an order-driven system. Hence, this study is a further contribution to the vast literature that compares and contrasts these two alternative market structures. Second, we consider the effect of the structural change on a large sample of government bonds with varying degrees of liquidity. Prior analysis has focused on a small number of very liquid securities. However, our objective is similar to Albanesi and Rindi (2000) who examine a series of mechanism changes in the MTS markets prior to 2000. Here, the authors have focused on two events related to the MTS market: the introduction of the primary dealers in the MTS markets in 1994 and the introduction of the anonymity of trades in 1997. Albanesi and Rindi (2000) choose 4 Italian benchmark bonds associated with three separate months in 1993, 1995 and 1997, respectively, to gauge the effect on the two events. The long-term, market-wide influence may not be seen from the results. On the other hand, we explore the market aggregate liquidity affected by a rule change for 7 European countries using a longer and continuous sample period.

Apparently, our approach is able to provide a systematic description of the liquidity development. Thirdly, our results are relevant for portfolio managers, regulators and owners of market systems who are all concerned by the liquidity of government bonds. When portfolio managers want to execute a large number of fixed income securities, transaction costs pose a great threat to the overall profits. Regulators such as the ESMA are trying to prevent liquidity-led market crash. The fierce competition to attract order flows between exchanges requires the market system owners to be ever vigilant when promoting liquidity. The order-driven feature with the presence of obligated dealers for the EuroMTS market may be the ideal place for portfolio managers to reduce transaction costs, for regulators to monitor the market and for market owners to review how to promote liquidity.

The rest of the paper is organized as follows: Section 5.4 explains our testing hypothesis followed by Section 5.5 describing our dataset and bond portfolio constructions. Section 5.6 explain how we compute the liquidity proxies and Section 5.7 outlines the econometric methodology, while Section 5.8 interprets the final empirical results whose robustness is tested in Section 5.9. Finally , Section 5.10 concludes the entire analysis. Full regression estimations are displayed in Appendix B.

5.4 Hypothesis Design

Following Darbha and Dufour (2013), we rely on a time-weighting scheme to construct the liquidity measures for a single bond. Since the intensity of quoting varies over time and differs across bonds, the simple average of bid-ask spreads over a trading day for a single bond cannot properly reflect the liquidity conditions. Instead, the weight of intraday bid-ask spreads and depth is calculated as the proportion of the trading day in which the quote is available. The testing statistics are based on the log of average daily measures of all bonds belonging to the same category (see Section 5.5 for building the categories and forming portfolios, and 5.6 for the computation of the liquidity measures). Since there are few papers comparing liquidity between different forms of fixed-income markets, we draw the following hypothesis mainly from the literature on stock markets.

H1 The average daily percentage spread for the European government bonds has decreased since the privilege of inserting limit orders is introduced.

H2 The average daily depth for the European government bonds at the best 3 price levels has decreased because of the change.

H3 The average daily percentage spread for the exact 10 million European government bonds in the market has decreased

Many studies suggest that auction-type markets tend to have lower spreads (Lee, 1993; Schmidt and Iversen, 1993; Petersen and Fialkowski, 1994; Christie and Huang, 1994) but also lower depth (De Jong et al., 1995; Gajewski and Gresse, 2007) compared to quote-driven markets. **H2** might need more elaboration as the evidence is not consistent. While Barclay et al. (1999) and Nimalendran and Petrella (2003) both document nondecreasing or even greater depth in the same form of the market as the EuroMTS market now, some papers argue that quoting size is superior in a market where dealers have monopolistic powers to manage their inventories (see e.g. Vijh, 1990). Obviously, the dealers on EuroMTS no longer have an oligopolistic market making the power, as all traders can now provide liquidity by submitting limit orders. In addition, we are not able to assess the liquidity of the whole order book since we only have the data for the top 3 quotes. The depth for each observation is calculated as the average of total bid depth and ask depth, which we can see from the dataset. However, the effects of the rule change on depth are somewhat mixed based on our analysis. We further combine the first two measures by computing the spread for a larger trade size than what can be traded at the best prices. Portfolio manager often need to execute trades with sizes larger than what is available at the best quotes (Hodrick and Moulton, 2009). The 10-million threshold is chosen because the minimum quoting quantity in the local MTS markets is usually 5 million bonds.[52]

H4 The average daily immediacy, measured as the proportion of time during the trading day where the percentage spread is less than the maximum daily average time-weighted percentage spread, plus 3 maximum standard deviation of the time-weighted percentage spread, has increased.

H4 captures the market quality when traders demand a quick execution. This way of measuring immediacy is not normally used for stock market in which transactions are frequent. On the other hand, bond transactions in the EuroMTS constitute a large amount and are very sparse over time. We have seen in Chapter 3 that there are times in which the market is essentially closed because of the wide spread. It is thus important for customers to know when they can trade in quality and the associated costs they are likely to pay. Analyzing transaction data alone cannot properly reflect the liquidity changes. According to the literature, the immediacy is a concept combining transaction costs with the length of the trading horizon. Grossman and Miller (1988) builds a theoretical model measuring immediacy as the total transaction size in period

[52]The minimum is 2.5 million for some of the Italian bonds . Since we only have the limited levels of the order book, enlarging the quantity of bonds would create more missing values, although the actual order book may be large enough .

1 and 2 of a 3-period model. Perold (1988) emphasizes the importance of the time spent to execute a portfolio when considering transaction costs. Chacko et al. (2008) models transaction costs from the perspective of a monopolistic market maker who benefits from investors who do not want to wait for long time. We thus propose to measure the immediacy as the relative time related to low transaction costs.

5.5 Data and Bond Portfolio Construction

We choose a sample that lasts from January 09, 2012 to September 30, 2013, which contains 220 days before and after the event date to capture the long-term effect on liquidity.[53] The coupon bearing government bonds of 7 European countries are covered, i.e. Austria (AT), Belgium (BE), France (FR), Germany (DE), Italy (IT), the Netherlands (NL), and Spain (ES). We use the same procedures presented in Chapter 1 and 2 for optimally filtering the quote data. Further criteria for choosing a bond are implemented as follows:

1. We construct portfolios of bonds with similar residual duration. We categorize the bonds into 3 maturity buckets following Dunne et al. (2007): 1.25–3.5 years (short term), 3.5–6.5 years (medium term) and 6.5–13.5 years (long term). Here we will not study the very long-term bonds as we observe that a large number of very long-maturity bonds issued 20 years ago have been listed on the EuroMTS since April 2013. Through a direct communication with the company, we confirm that the company introduced more securities to the platform other than the tra-

[53]The length of the window can be varied to be shorter or longer, while the empirical results will not be modified materially. However, as documented in Chapter 3, some Italian and Spanish government bonds experienced exceptional liquidity shocks in December 2011, a fact that may compromise the analysis. Thus it is better not to extend the sample to include the 2011 data.

ditional benchmark securities. It is very hard to distinguish the very old bonds with the existing benchmark securities. Once a bond's residual maturity is below the lower bound of a range, the bond is deleted from the corresponding portfolio and immediately added to the portfolio for the lower-maturity bucket, e.g. once the time to maturity of a bond in the long-term portfolio falls below the 6.5-year threshold, then the bond is moved to the medium-term portfolio. The categorization implies a bond-portfolio approach where each bond takes equal weight in the residual maturity portfolio.[54]

2. If the bond is newly issued and meets the criterion 1, then its introduction to the corresponding residual maturity portfolio is delayed by at least one month. This criterion is consistent with the approach we used to identify the on-the-run benchmark bonds in Chapter 3. Pasquariello and Vega (2009) points out that the bid-ask spread of the newly issued bond is higher than the just off-the-run bond for at least 10 days in the US Treasury market. Avoiding the first month after the issuance can alleviate the shock to the liquidity of the portfolio when the newly issued bond is added. In addition, the delay can be up to five months since the first auction as some of the bonds are not listed on the EuroMTS market until then.[55]

5.6 Liquidity Proxies

In view of the fact that we have a different number of observations each day, we let t denote the day, n_t ($n_t = 1, 2, \ldots, N_t$) denote the nth observation of the day t , T_t

[54]Alternatively, we would choose to remove the bond when its residual maturity is exactly equal to the lower bound. Our results have not been affected by this choice.

[55]The late listing does not appear to be a problem for long-term bonds.

represent the duration of the trading day t and T_{n_t} stand for the time stamp of the n_tth observation. $Bid_{n_t,i}$ and $Ask_{n_t,i}$ are the ith ($i = 1, 2, 3$) tier of the bid and ask prices for the n_tth observation, respectively. $AskSize_{n_t,i}$ and $BidSize_{n_t,i}$ are the corresponding sizes. We formulate the following liquidity proxies for individual bonds:

1. For **H1**, the daily average Time-Weighted Percentage Bid-Ask Spread (TWP-BAS), which is calculated in basis points.

$$\mathrm{TWPBAS}_t = 10000 * \frac{1}{T_t} \sum_{n_t=1}^{N_t} \frac{(Ask_{n_t,1} - Bid_{n_t,1})}{(Ask_{n_t,1} + Bid_{n_t,1})/2} * (T_{n_t+1} - T_{n_t}) \tag{5.1}$$

2. For **H2**, the depth of each intraday observation is the average of the top 3 bid sizes and ask sizes The daily average Time-Weighted Depth (TWDEP) is measured as

$$\mathrm{TWDEP}_t = \frac{1}{T_t} \sum_{n_t=1}^{N_t} (\frac{1}{2} * \sum_{i=1}^{3} (AskSize_{n_t,i} + BidSize_{n_t,i})) * (T_{n_t+1} - T_{n_t}) \tag{5.2}$$

3. For **H3**, we first calculate the average size-weighted bid (SWBid) and ask price (SWAsk) within a 10 trade :

$$\begin{aligned}
\mathrm{SWBid}_{n_t} = {} & Bid_{n_t,1} * \frac{\min\{10mil, BidSize_{n_t,1}\}}{10mil} \\
& + Bid_{n_t,2} * \frac{\max\{10mil - Bidsize_{n_t,1}, 0\}}{10mil} \\
& + Bid_{n_t,3} * \frac{\max\{10mil - BidSize_{n_t,1} - BidSize_{n_t,2}, 0\}}{10mil}
\end{aligned}$$

$$\begin{aligned}
\mathrm{SWAsk}_{n_t} = {} & Ask_{n_t,1} * \frac{\min\{10mil, AskSize_{n_t,1}\}}{10mil} \\
& + Ask_{n_t,2} * \frac{\max\{10mil - Asksize_{n_t,1}, 0\}}{10mil}
\end{aligned}$$

$$+ Ask_{n_t,3} * \frac{\max\{10mil - AskSize_{n_t,1} - AskSize_{n_t,2}, 0\}}{10mil}$$

Then we compute the daily average Time-Weighted Percentage Bid-Ask Spread for the 10 million bond trade in basis points as:

$$\text{TWPBAS}_{t|\text{DEP}=10mil} = 10000 * \frac{1}{T_{t|\text{DEP}>=10mil}} \sum_{n_t=1}^{N_t} \frac{(\text{SWAsk}_{n_t} - \text{SWBid}_{n_t})}{(\text{SWAsk}_{n_t} + \text{SWBid}_{n_t})/2}$$
$$* (T_{n_t+1} - T_{n_t}) * \mathbb{1}_{\text{DEP}_{n_t}>=10mil} \tag{5.3}$$

Notice that the $\mathbb{1}_{\text{DEP}_{n_t}>=10mil}$ indicator function suggests that we only measure the spread when the average depth is larger than 10 million. If either side of the order book does not have the sufficient amount, then we do not include the observation in the calculation.

4. For **H4**, the maximum daily average time weighted percentage spread of all bonds for one portfolio[56] and the maximum of the corresponding time-weighted standard deviation is computed first. Then the daily immediacy (IMMED) is

$$\text{IMMED}_t = \frac{1}{T_t} \sum_{n_t=1}^{N_t} (T_{n_t+1} - T_{n_t}) * \mathbb{1}_{t|\text{PBAS} <= \max(\text{TWPBAS}_t) + 3*\max(\text{TWSTD}_t)} \tag{5.4}$$

Immediacy is a variable ranging from 0 to 1, which can be interpreted as the probability of the spread being below the threshold during the day. Using this limited dependent variable as the explanatory variable in an OLS regression leads to a heteroskedasticity problems, which can be controlled with robust standard errors (see below).

[56]The portfolio comes from the one maturity category and contains a single country's debts.

The liquidity proxies for the subsequent statistical analysis are the log of equally weighted average liquidity proxies of individual bonds that belong to the same portfolio.

5.7 Econometric Methodology

We intend to compare the unconditional and conditional liquidity difference before and after the event. The unconditional difference is examined in the Wilcoxon signed-rank test following Bessembinder et al. (2009) with a de-trending operation. The conditional difference is reflected in an OLS regression with various control variables.[57] The Wilcoxon signed-rank test is a nonparametric test, which does not require normality of the liquidity proxies. The absence of normality in the assumptions of the signed-rank test is most convenient for the immediacy – a nonnegative number between 0 and 1. Hence, we detrend the variables for the nonparametric test because we observe a time trend that drives the liquidity proxies used to test **H1**, **H2**, and **H3** due to the ending phase of the sovereign debt crisis. The immediacy measure is only demeaned. Specifically, the detrending regression for the entire sample is:

$$y_t = \alpha + \beta t + \varepsilon_t. \tag{5.5}$$

The log of average daily liquidity proxies is used as the regressands. The null hypothesis for the Wilcoxon signed-rank test is that the median of the difference in the residuals before and after the event is zero, in comparison to the two-sided alternative that the median is nonzero. The Wilcoxon signed-rank test requires that the pairwise difference is serially independent and symmetrically distributed around its mean or median. It

[57] Newey-West (1987) standard error with 7 lags is computed

also assumes that there is no tie rank and no zero difference. We randomly permute the observations of the two sample windows in order to eliminate any autocorrelations of the differenced series. The Ljung-Box test is conducted on the permuted series, while no autocorrelation is eventually detected. The median number does not vary significantly across different permutations of the residuals. The tie rank is given by the average rank and the small-sample P-value is provided when the sample size is less than 20 because of too many discarded zero differences.[58] The symmetry of distribution is checked via histogram, while the absolute value of skewness occasionally exceeds 2 for the liquidity variables. The sign test is used in this case.[59]

The OLS regression is specified as:

$$y_t = \alpha + \beta_{EuroMTS} * Dummy_{t,EuroMTS} + \sum_{k=1}^{K} \beta_k * Control\ Variable_{t,k} + \varepsilon_t \tag{5.6}$$

The $Dummy_{t,EuroMTS}$ takes value 1 since November 15, 2012 (inclusive) and value 0 before that date. As in the detrending regression, the regressand is still the log of average liquidity proxies. There are several control variables we use to capture the liquidity effects attributed to events other than the rule change, including:

1. A dummy for the last day of the year to account for the apparent slow-down of quoting, which is a holiday effect for the MTS dataset. We also experimented with other holiday dates retrieved from the following Bloomberg terminal Chordia et al. (2005). However, we do not find any significant effects in our regression here. Through a visual inspection of the liquidity variables, we cannot tie low liquidity levels with holidays. Monthly and weekly seasonality are not found in the t-test

[58] The small sample is unlikely after the permutation.

[59] Other thresholds are checked for robustness, which does not alter the conclusion. The biggest skewness is 4.7 for the skewed distribution of immediacy for French long-term bonds in Table 5.2.

compared to the mean of the liquidity variables.

2. A dummy for the dates when introducing newly-issued bonds to the portfolio; a dummy for the dates when introducing old dated bonds whose liquidity are usually lower than newly-issued bonds to the portfolio; a dummy for the dates when newly-issued bonds are removed from the portfolio; and a dummy for the dates when dated bonds are excluded from the portfolio. These four dummies intend to capture the jumps of liquidity on the dates when the portfolio construction changed. Potentially, the newly-issued benchmark bonds enjoy greater liquidity than the old dated bonds in the portfolio. The addition of such bonds may lead to sudden increases in the average liquidity and vice versa. Since we construct the portfolios according to residual maturities, the newly issued bonds are naturally defined as the bonds issued for the specific maturity category. Hence, we shall refer them as the standard bonds in the OLS regression. For example, a bond issued with a maturity of 7.5 years is a standard bond for the medium-term category. It becomes a dated bond for the short-term category when its remaining time to maturity has dropped below 3.5 years. The inclusion and exclusion of the dated bonds have the opposite effects of adding and deleting the standard bonds, respectively.

3. A time trend variable t which represents the improving liquidity conditions due to the end of the sovereign debt crisis. Chordia et al. (2005) add t^2 into their OLS regression. We do the same, but find no effect of t^2 on the liquidity variables. t is not included in the regression for immediacy.

4. One lag of the volatility of the market index to control for market-wide uncertainty

according to Engle et al. (2012) and Chordia et al. (2005). The volatility is defined as the absolute value of the log return of the corresponding maturity bond index. The intuition behind the above is that the higher the uncertainty of the market, the lower of the liquidity. The indices downloaded from Data Stream are the JP Morgan GBI Europe Index for bonds with maturities of 1-3 years, 3-5 years and 5-10 years, respectively.

5. The log of the average trading imbalance over all eligible bonds for the last 5 days to check for trading impacts. The variable is computed by taking the log of the absolute value of the trading imbalance and then multiplying it with the sign of the original imbalance. Before taking logs, we add 1 to the imbalance figure, which leads to zero log imbalance if there are no transactions for the last 5 days or the imbalance is indeed zero. We separate positive from negative trading imbalances (which is defined as buyer-initiated volume minus seller-initiated volume) in Equation (5.6) as we expect some asymmetrical effects of purchasing and selling pressure. For example, a large net sale of Italian bonds may signal the deepening of the crisis, which may increase spreads more than an equally large net purchase. Therefore, we will not normalize the trading imbalance by total volume as there is usually only one transaction per bond (if any) in a trading day.

6. The absolute standardized surprise to the macroeconomic announcements of US and Europe $|\frac{E-A}{\sigma}|$, where A represents the actual value and E is the expected value of the news statistics retrieved from the Bloomberg terminal. The Long Term Refinance Operation (LTRO) data is obtained from the ECB website and there is no surprise associated with this news. Instead, a dummy variable for the LTRO is used. The dummy variable takes value 1 on the days when the LTRO

is announced by the ECB and 0 otherwise.[60] Beetsma et al. (2013) document ”good” and ”bad” news effects on interest rates in the European bond markets during the crisis. Taking absolute values helps us to avoid differentiating between the effects of positive and negative surprises on liquidity and helps us focus on the effect of uncertainty on liquidity. The US news is never released after the 17:30 closing time of the MTS market as we covert the US EDT time to the CET time. Table B.1 presents the selected news, which is only comprised of a small part of the public information pool, but the importance of these news is emphasized in the literature (see Balduzzi et al., 2001 and Andersen et al., 2003b for reference).

We have also tested other possible candidates for control variables, such as the number of market participants, the number of market makers, bond duration, and bond age. but most of these variables do not vary enough over time, causing multicollinearity problems.

5.8 Empirical Results

5.8.1 Summary Statistics

It is apparent in Table 5.1 that most of the liquidity measures improved greatly after the EuroMTS gave market participants access to using limit orders. After the rule change, the TWPBAS is roughly one-third of the past level for Austria, Spain and even for France. Although the liquidity for German and Dutch bonds has always been good, there is still a visible improvement. Generally, the longer the maturity, the greater the reduction in the TWPBAS. For example, there is a 15 basis points decrease for

[60] The LTRO is settled 3 days after the announcement.

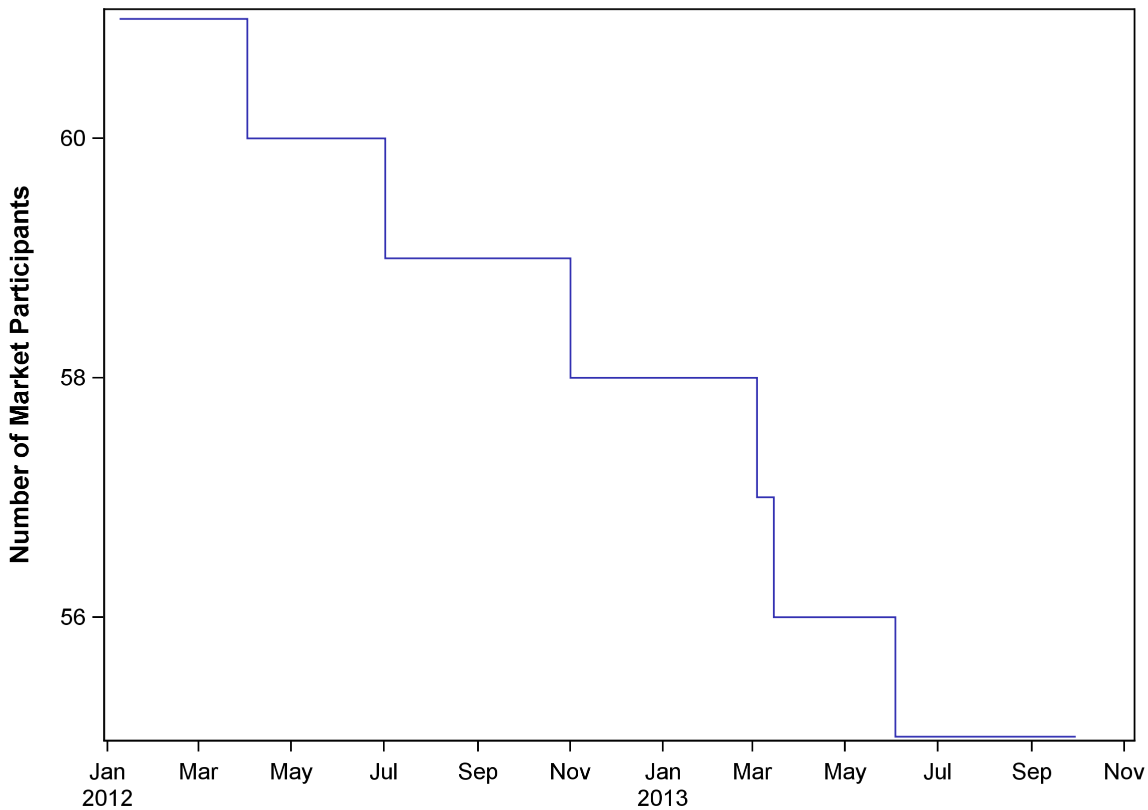
Italian short-term bonds, while the decline of long-term TWPBAS becomes more than 30 basis points. The increment for the TWDEP is less striking but also evident. It seems that the average total depth increases by 11 million, which concentrates on short-term and medium-term bonds excluding German bonds. Similar to the TWPBAS, the enhancement of depth is weakest for Germany and we observe even a small deterioration for the medium and long-term portfolio probably because of losing customers for this market, shown in Figure 5.1. By checking the information on bond characteristics, we observe that the number of market participants for all German bonds gradually moved from 61 to 55 during the sample period. The 10-million-bond TWPBAS follows similar pattern to the one observed for the top-of-the-book TWPBAS, suggesting a uniform reduction in spread for lower tiers in the order book. The immediacy measure tells the same story of a higher liquidity after the event. Not is only the bid-ask spread lower, but the probability of encountering low spread is also higher, except for Germany; meanwhile, a 3% increase is observed for Austria.

The downward sloping time trend of the bid-ask spread is obvious in Figure 5.2 for all countries' debts, including the German ones. The large single jump in the plots of France, Germany and the Netherland is due to the end-of-year trading effect and it is controlled in Equation (5.6) with the end-of-year dummy. One may observe that there is a structural shift of the TWPBAS of Italian bonds around August 2012 in Figure 5.3. Note that, around the same time, there was a concurrent decline in the spread for the Spanish bonds. The systematic decline of bid-ask spread was suddenly driven by lower market volatility. This may well be attributed to some news from the Eurozone or ECB interventions. The plots in Figure B.1 for depth do not show a clear trend. Nonetheless, for the most part the trend variable for TWDEP is significantly different

Table 5.1: The mean of the average liquidity measures before and after the event

TWPBAS (bps)	Sample Period	AT	BE	FR	DE	IT	NL	ES
short term	before	43.15	16.19	17.48	6.27	27.37	6.51	51.56
	after	13.72	5.37	6.22	5.17	12.75	3.47	21.18
medium term	before	48.16	28.27	22.13	6.92	34.14	12.46	89.05
	after	17.84	10.65	9.91	5.88	14.70	7.54	30.36
long term	before	45.67	29.38	23.63	9.93	51.69	16.16	116.23
	after	21.51	13.61	14.46	7.40	19.83	10.33	41.62
TWDEP (mil.)								
short term	before	39.82	50.51	53.16	31.33	56.67	57.66	40.64
	after	45.35	58.31	78.92	31.40	69.04	68.07	50.94
medium term	before	39.89	59.49	53.71	37.49	54.43	59.43	40.64
	after	45.82	78.74	69.34	36.22	69.91	69.38	43.48
long term	before	36.77	55.12	46.50	37.01	51.01	51.87	36.06
	after	37.03	68.08	51.62	33.44	56.07	57.15	36.98
$\text{TWPBAS}_{\text{DEP}=10mil}$								
short term	before	46.68	16.79	18.29	6.82	27.60	6.56	57.17
	after	13.89	5.46	6.42	5.66	12.79	3.48	22.61
medium term	before	51.40	29.68	23.06	7.24	33.64	12.67	96.96
	after	18.14	10.82	10.09	6.58	14.28	7.62	32.92
long term	before	50.01	30.99	25.64	11.44	50.76	16.79	127.32
	after	24.38	14.04	14.75	8.63	19.79	10.72	45.83
IMMED								
short term	before	0.9438	0.9787	0.9893	0.9981	0.9605	0.9865	0.9411
	after	0.9950	0.9957	0.9970	0.9972	0.9747	0.9942	0.9694
medium term	before	0.9433	0.9633	0.9842	0.9974	0.9813	0.9792	0.9407
	after	0.9927	0.9884	0.9962	0.9944	0.9938	0.9875	0.9500
long term	before	0.9370	0.9572	0.9718	0.9966	0.9798	0.9771	0.9447
	after	0.9845	0.9870	0.9918	0.9995	0.9932	0.9845	0.9489

Figure 5.1: The average number of market participants for short-, medium- and long-term German bonds in the EuroMTS market.



from zero in the detrending regression. Hence, with depth we also control for the time trend (see Appendix B for the estimation of the de-trending regression).

5.8.2 Wilcoxon Signed-Rank Test

The Wilcoxon test in Table 5.2 corroborates our earlier description of the mean of liquidity proxies. Almost all the bonds enjoy a smaller bid-ask spread, including German bonds, even after controlling for the better liquidity conditions due to the fading of the European sovereign bond crisis. Notice the rule change translates into a reduction of nearly 10 (($\approx (1-\exp(-0.20))*51$) basis points in the TWPBAS (equivalent to €5000 saving for a transaction of 5 million euros) for short-term and long-term Spanish debts, while a comparable drop happened to long-term Dutch debts. The TWDEP turns out to be augmented strongly for Dutch bonds whose depth has expanded by an average of at least 9% ($\approx \exp(0.09)-1$) . A similar result holds for the $\text{TWPBAS}_{\text{DEP}=10mil}$. The higher probability of having tradable spread is observed for most of the bonds, with Austrian debts receiving the greatest liquidity enhancement. Some of the significant results are too small to be economically relevant as the liquidity is large throughout the sample; see, for example, the results for the short-term and medium-term bonds of Germany – the decrease in the immediacy corresponds to 3.33 ($= 33300 * 0.0001$) seconds reduction in time.[61] Even long-term French debts have greater immediacy. **H1**, **H3**, **H4** where they are generally accepted at 1% and 5% significance level whereas **H2** is occasionally rejected for long-term bonds.

[61] Trading lasts for 9 hours and 15 minutes, which is equivalent to 33300 seconds.

Figure 5.2: Time series plots of the TWPBAS of short-term bonds of non-distressed countries

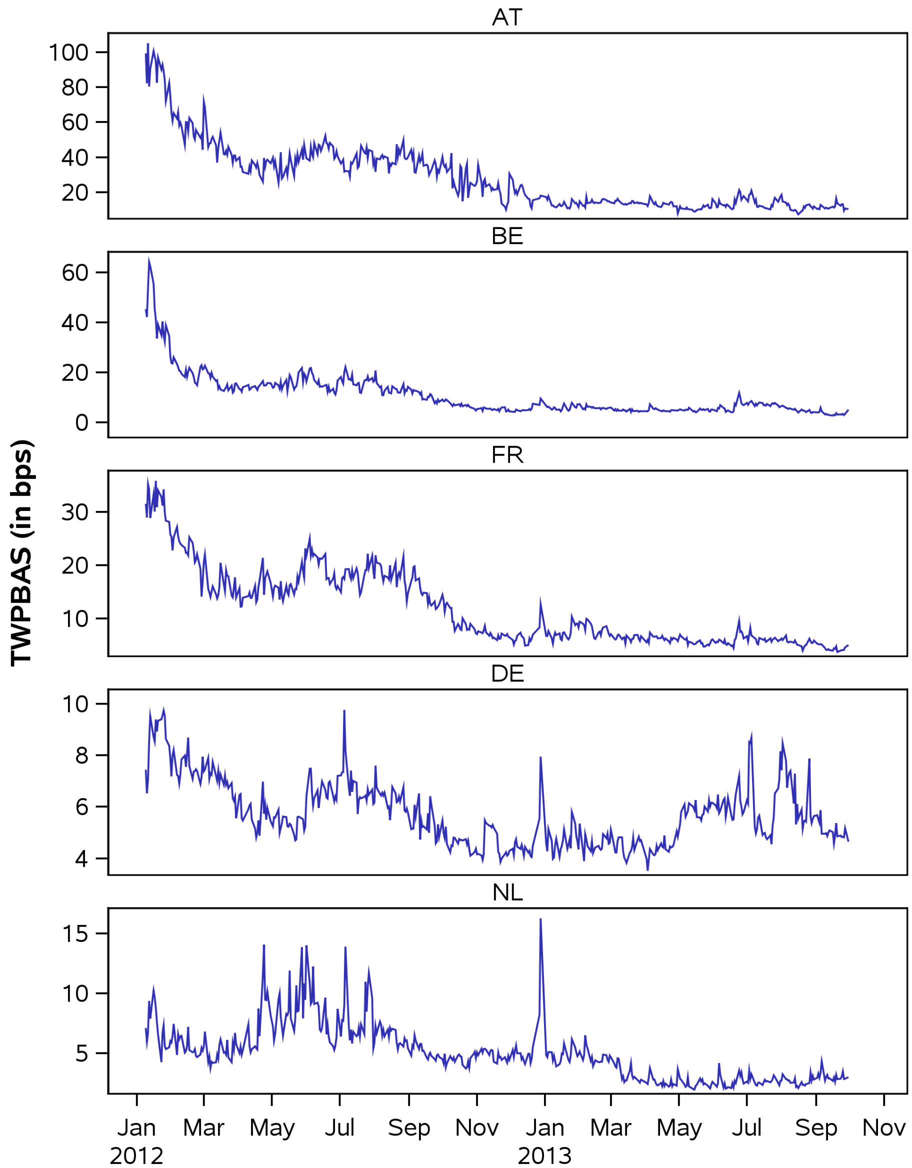


Figure 5.3: Time series plots of the TWPBAS of the Italian and Spanish short-term bonds and market volatility.

Market volatility is computed as the absolute return in percentage points of the JP Morgan GBI Europe index.

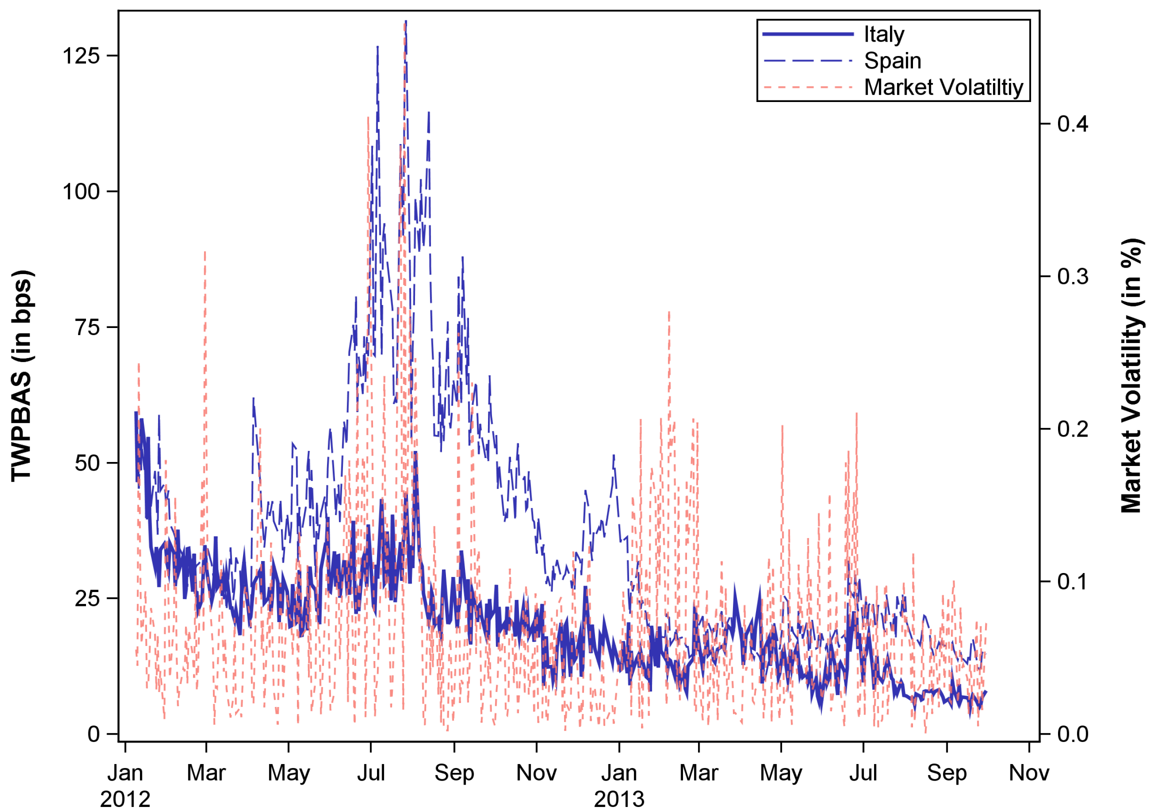


Table 5.2: Median of the difference in liquidity for detrended and demeaned series

The null hypothesis of the Wilcoxon signed-rank test is that the median of the difference in the residuals of the liquidity variables before and after the event is zero. For the test on the spread and depth proxies we use the residuals of the detrending regression Equation (5.5). For the test on the immediacy proxy we simply use the demeaned series. ***,**,* denote 1%, 5%, 10% significance respectively. The superscript s indicates that the sign test is used.

TWPBAS	AT	BE	FR	DE	IT	NL	ES
short term	−0.1263***	−0.0659**	−0.1034***	−0.0086	−0.0176	−0.0400	−0.1975***
medium term	−0.0838***	−0.0655*	−0.0680***	0.0083	−0.0022	−0.0600**	−0.1194***
long term	−0.0221**	−0.0135	−0.1195***	−0.0518**	−0.0680**	−0.1171***	−0.1727***
TWDEP							
short term	−0.0410	0.0005	0.0064	0.0421***	0.0146	0.0940***	0.0021
medium term	−0.0058	0.0321***	−0.0011	0.0314**	0.0072	0.1227***	−0.0544**
long term	−0.0607**	0.0260**	0.0097	−0.0233*	−0.0521***	0.0999***	−0.0151
$\text{TWPBAS}_{\text{DEP}=10mil}$							
short term	−0.1360***	−0.0484	−0.0761***	−0.0832***	−0.0155	0.0077	−0.2608***
medium term	−0.0718***	−0.0509**	−0.0754***	−0.0000	0.0145	−0.0584**	−0.1398***
long term	−0.0438**	−0.0005	−0.0844***	−0.0525**	−0.0388**	−0.0770**	−0.1765***
IMMED							
short term	0.0345***	0.0034^{***s}	0.0000^{s}	−0.0001***	0.0128***	0.0000***	0.0275***
medium term	0.0329***	0.0113***	0.0000^{s}	−0.0001***	0.0072***	0.0000*	0.0114***
long term	0.0269***	0.0107^{***s}	0.0080^{***s}	0.0012***	0.0106^{***s}	0.0027^{**s}	0.0025

5.8.2.1 Why We Have Some Smaller Depth for Long-Term Bonds?

A notable deterioration of the depth pertains to long-term Austrian and Italian bonds. One explanation could be that market participants and market makers now submit more frequently minimum-size orders,[62] whereas a bigger minimum quote size for market makers exists in the domestic platforms.[63] We compute the average daily proportion of order updates with quoting quantity equal to the 2 million minimum amount for a bond portfolio for all 3 levels of the quoted prices in the EuroMTS platform. The daily proportion takes account of the fact that, sometimes, not all tiers are populated with quoted price and size and the quote updates are only counted when a quote exists for the tier.[64] Table 5.3 presents the median of the difference in the demeaned relative proportion of the small orders, before and after the event. For the countries with the most severe depth deterioration, dealers and traders indeed tend to use small orders more frequently after the event. Austrian long-term bonds have witnessed a 19% percent surge in the usage of such orders in the top bid tier and a total 5% increase and 10% happened for Italian bonds on the bid and ask side, respectively. The decline in depth for Spanish bonds though insignificant may also be explained by this phenomenon. Investors supply a very limited amount for Spanish debts, almost 3% more often on average for bid prices on a daily basis. The results are even more evident for Spain on the ask side with a 13% expansion of minimum-size orders! Besides losing customers in the German market, the willingness to provide a deep market has also been contracted on this front. The small order frequency is too small to have any real effects for the rest

[62] The minimum quantity for limit orders is 1 or 2 million in the EuroMTS market.

[63] Usually the minimum quoting quantity is 5 million or more for the benchmark bonds in the local MTS platforms.

[64] A new record is generated whenever there is an update for price or size in any of the three tiers on either side of the order book.

of the countries, which is consistent with a non-decreasing or larger depth available in the market. In fact, the overall proliferation of limited quantity suggests that the rule change has really worked. Non-dealers are now competing with dealers in providing liquidity. They provide tighter spreads with relatively smaller depth, which benefits traders with smaller transaction sizes. Our finding is consistent with Barclay et al. (1999), who make the conjecture that an insignificant reduction in daily average time-weighted depth is caused by the NASDAQ cutting the minimum quote size from 1000 shares to 100 shares and dealers utilized their opportunities to post 100 shares during the transition to an order-driven market.

Table 5.3: Median of the difference in relative update frequency of minimum-size limit orders for a demeaned series of long-term bonds

The null hypothesis of the Wilcoxon signed-rank test is that the median of the difference in the residuals of the relative small-order frequency before and after the event is zero. ***,**,* denote 1%, 5%, 10% significance respectively. The superscript s indicates that the sign test is used.

Bid Tiers	AT	BE	FR	DE	IT	NL	ES
$BidSize_{1,n_t}$	0.1850***	0.0000***	0.0000***	0.1569***	0.0465***	0.0000^{***s}	0.0261**
$BidSize_{2,n_t}$	0.0087^{***s}	0.0000***	0.0000^{***s}	0.0320***	0.0033***	0.0000^{***s}	0.0064^{***s}
$BidSize_{3,n_t}$	0.0050^{***s}	0.0000***	0.0000^{***s}	0.0001**	0.0016***	0.0000***	0.0045***
Ask Tiers							
$AskSize_{1,n_t}$	0.0003***	0.0000***	0.0000***	0.1697***	0.0939***	0.0000^{**s}	0.1129***
$AskSize_{2,n_t}$	0.0026^{***s}	0.0000***	0.0000^{***s}	0.0384***	0.0046***	0.0000^{***s}	0.0127***
$AskSize_{3,n_t}$	0.0037***	0.0000***	0.0000^{***s}	0.0007	0.0020***	0.0000***	0.0078***

5.8.3 The OLS Regression

The OLS regression Equation (5.6) measures more accurately the conditional effect of the market rule change by controlling for a series of other factors which may have affected the change in liquidity. The figure is much more strengthened in Table 5.4 than the one suggested by the Wilcoxon signed-rank test, proving the necessity of controlling different variables and providing more confidence to the acceptance of hypothesis **H1**, **H3**, **H4**. Notably, the TWPBAS of short-term Spanish bonds has fallen to $20 \approx \exp(3.8056 - 0.8024)$ basis points on average from its pre-event level of $44 \approx \exp(3.8056)$ basis points, due to the rule change. A similarly large reduction in the proportional spread is found for short-term Austrian bonds, with a width of bid-ask spread reduced by 21 ($\approx \exp(3.9851) * (1 - \exp(-0.5087))$) basis points. Even stronger results are found in the reduction of the $\text{TWPBAS}_{\text{DEP}=10mil}$, which also implies a decreased spread for lower tiers. The magnitude of changes is stronger for TWDEP with a more significant smaller depth for medium-term Spanish bonds. The increase for the Netherlands is an impressive 20 million bonds ($\approx \exp(17.9) * (\exp(0.27) - 1)$) and even more for medium-term Dutch bonds! The increase in the immediacy is also enhanced when isolating other effects. All the Austrian bonds are 4% more likely to trade with a percentage spread less than the maximum daily average, as a time-weighted percentage spread plus 3-standard deviation. Smaller immediacy was also found for short-term and medium-term German bonds, but the significance is only marginal.

Table 5.4: Estimation of the constant α and the $\beta_{EuroMTS}$ in Equation (5.6)

The regression (5.6) is estimated via OLS with Newey-West (1987) standard error (7 lags included). ***,**,* denote 1%, 5%, 10% significance respectively.

TWPBAS	Parameter	AT	BE	FR	DE	IT	NL	ES
short term	α	3.9851***	2.9139***	3.0338***	1.8285***	3.5180***	2.0269***	3.8056***
	$\beta_{EuroMTS}$	−0.5087***	−0.2571*	−0.3820***	−0.1706*	−0.0395	−0.1428	−0.8024***
medium term	α	4.0953***	3.4617***	3.1859***	1.9839***	3.6505***	2.6369***	4.5473***
	$\beta_{EuroMTS}$	−0.2884***	−0.2401*	−0.3420***	0.0172	−0.0858	−0.2265**	−0.6784***
long term	α	3.9735***	3.5224***	3.2336***	2.3109***	4.1257***	2.8144***	4.9374***
	$\beta_{EuroMTS}$	−0.2981***	−0.1203	−0.2978***	−0.1193**	−0.2362**	−0.2261**	−0.5308***
TWDEP								
short term	α	17.4421***	17.6677***	17.5844***	17.3525***	17.8127***	17.9297***	17.4107***
	$\beta_{EuroMTS}$	−0.0587	−0.0340	0.0091	0.1835***	−0.0128	0.2740***	−0.0169
medium term	α	17.4044***	17.8379***	17.7230***	17.5395***	17.7310***	18.0430***	17.4165***
	$\beta_{EuroMTS}$	−0.0280	0.1090**	0.0803**	0.1250***	−0.0089	0.3361***	−0.1337**
long term	α	17.3492***	17.7921***	17.6457***	17.4503***	17.6435***	17.8953***	17.3580***
	$\beta_{EuroMTS}$	−0.1485***	0.1085**	0.0568	−0.0737	−0.1964***	0.2721***	−0.0748
$\text{TWPBAS}_{\text{DEP}=10mil}$								
short term	α	4.0602***	2.9346***	3.0906***	1.8438***	3.5549***	2.0337***	3.8722***
	$\beta_{EuroMTS}$	−0.5966***	−0.2555	−0.3626***	−0.3310***	0.0262	−0.1402	−0.8367***
medium term	α	4.1578***	3.5002***	3.2315***	1.9768***	3.6320***	2.6528***	4.5935***
	$\beta_{EuroMTS}$	−0.3313***	−0.2400*	−0.3264***	−0.0144	−0.0926	−0.2211**	−0.7128***
long term	α	4.0566***	3.5711***	3.3227***	2.4778***	4.1125***	2.8346***	4.9811***
	$\beta_{EuroMTS}$	−0.3133***	−0.1185	−0.2888***	−0.0863	−0.2238*	−0.2356**	−0.5715***
IMMED								
short term	α	0.9488***	0.9904***	0.9957***	0.9982***	0.9621***	0.9906***	0.9635***
	$\beta_{EuroMTS}$	0.0485***	0.0139***	0.0121*	−0.0009*	0.0144***	0.0098***	0.0246***
medium term	α	0.9471***	0.9759***	0.9924***	0.9992***	0.9894***	0.9870***	0.9560***
	$\beta_{EuroMTS}$	0.0478***	0.0237***	0.0113**	−0.0032*	0.0109***	0.0087**	0.0084
long term	α	0.9414***	0.9665***	0.9804***	0.9975***	0.9826***	0.9837***	0.9590***
	$\beta_{EuroMTS}$	0.0465***	0.0292***	0.0180***	0.0026***	0.0118***	0.0074*	0.0016

5.9 Robustness Test

We now run a series of robustness tests. First, we apply the same econometric methodology to on-the-run bonds only to assure that the liquidity progression is not driven by portfolio construction. Indeed, most of the results stay qualitatively unchanged in Table 5.5 and 5.6. Second, we conduct a difference-in-difference test between the domestic MTS platforms and the EuroMTS platform. We difference the liquidity proxies between the two platforms first and then perform the signed-rank test on the already differenced series. The signed-rank test is implemented in the same fashion as described in Section 5.7 except that there is no detrending or demeaning for the series. Due to the similarity between the two platforms, any time-trend and/or mean is obviously removed by the first difference operation. The presence of a trend is tested on the differenced series but the test produces insignificant t results. It would appear that the empirical finding of the TWPBAS is less uniform in Table 5.7 than in Table 5.2. While Belgium, France and Spain still maintain the decline in the TWPBAS, Germany, Italy and the Netherlands have flipped signs. Not surprisingly, the $\text{TWPBAS}_{\text{DEP}=10mil}$ follows the same pattern as the TWPBAS. The TWDEP has more negative outcomes supported by most of the countries, with the exception of Germany, which actually validates our reasoning for lower depth. The larger minimum quantity requirement for the local platforms prevents the overuse of insufficient size, while the more depressing results suggest the superiority of the local platforms to the EuroMTS in terms of depth in Table 5.7. The immediacy is essentially indistinguishable from the pre-event level, though an observable deterioration happened to France and a slightly higher probability has developed in the Dutch market. The difference-in-difference results rely heavily upon the assumption that the liquidity in the domestic MTS platforms is independent of liquidity in the EuroMTS platform. However, the assumption is unlikely to be fulfilled because most of the dealers tend to quote on both platforms.

Third, we have rerun the analysis by allowing the time trend parameter to change after the event and the results stay qualitatively the same. The detrending regression and the OLS regression become the following:

$$y_t = \alpha + \beta_1 t + \beta_2 * Dummy_{t,EuroMTS} * t + \varepsilon_t. \tag{5.7}$$

$$\begin{aligned} y_t = {} & \alpha + \beta_{EuroMTS} * Dummy_{t,EuroMTS} + \sum_{k=1}^{K} \beta_k * Control\ Variable_{t,k} \\ & + \beta_{k+1} * Dummy_{t,EuroMTS} * t + \varepsilon_t \end{aligned} \tag{5.8}$$

Table 5.8 below shows the results of the Wilcoxon signed rank test on the detrended series (or demeaned series for the immediacy proxy). The detrending operation allows a change in the coefficient of the time variable t. A test for the time variation of the intercept will be conducted on the detrended series. The regression estimation in Table 5.9 also accommodates the time trend change and is consistent with the results of Table 5.8. We observe that the results for Dutch and Spanish bonds now show a surprising increase in bid-ask spreads for the period after the event. However, the overall results still hold with a decrease in bid-ask spread and an increase in depth for other markets. The immediacy measure stays the same as we do not apply the detrending to this series. The first differencing approach would be more appropriate if we compared the changes in liquidity before and after the event but we are primarily interested in comparing the level of liquidity before and after the event.

Finally, we decide to remove the data the week before and the week after the event to account for any transitory effect. There is no uncertainty with respect to both the announcement date and the implementation date for the rule change. However, it is reasonable to assume that market participants started anticipating the event and adjusted their trading behaviour gradually before the event. Table 5.10 and Table 5.11 present the results for this robustness test. Overall, the estimation results do not change qualitatively. We also allow for the coefficient of the t variable to change in this test but the combined test does not yield further insights.

5.10 Concluding Remarks

We have studied the development of liquidity after the EuroMTS market, granting every investor access to posting one-sided limit orders. We have observed a significant decline in the daily average time-weighted percentage bid-ask spread on the top level and deeper levels of the order book. According to our observations, the daily average time-weighted depth has mostly become greater, with some deterioration due to dealers and market participants using more 2-million-size orders and losing customers for German markets. Although it can be argued that the willingness to trade long term bonds has diminished because of the insolvency risk and long-term uncertainty and thus the depth was lower, the low minimum size makes the deterioration more likely. The immediacy, which can be interpreted as the probability of encountering a low spread, has also increased, most significantly for Austrian bonds. When controlling for various variables that might have impacts on liquidity, our empirical discovery is enhanced. Robustness test of having a controlled group from the local MTS platforms turns out to be less supportive, but the liquidity of local MTS and the EuroMTS depend on each other. Overall, the transition from a quote-driven market to an order-driven market has helped

Table 5.5: Median of the difference in liquidity for on-the-run bonds

The null hypothesis of the Wilcoxon signed-rank test is that the median of the difference in the residuals of the liquidity variables before and after the event is zero. For the test on the spread and depth proxies, we use the residuals of the detrending regression Equation (5.5). For the test on the immediacy proxy we simply use the demeaned series."***,**,* denote 1%, 5%, 10% significance respectively. The superscript s indicates that the sign test is used.

TWPBAS	AT	BE	FR	DE	IT	NL	ES
short term	−0.1903***	0.0398	−0.0161	−0.0018	−0.0167	−0.0060	−0.1993***
medium term	−0.0816***	−0.0058	−0.1018***	0.0402	−0.0357	−0.1256***	−0.1643***
long term	−0.1029***	−0.0647**	−0.0444*	−0.0480**	−0.0646**	−0.0339**	−0.0678***
TWDEP							
short term	−0.0350	0.0049	−0.0251*	0.1177***	−0.0036	0.1275***	−0.0425**
medium term	0.0288	0.0539***	0.0349**	0.0676***	−0.0379**	0.1132***	−0.0444**
long term	−0.0143	0.0462***	0.0533***	−0.0100	−0.0004	0.1421***	0.0060
$\text{TWPBAS}_{\text{DEP}=10mil}$							
short term	−0.2146***	0.0147	−0.0244	−0.0556*	−0.0767	−0.0247	−0.2336***
medium term	−0.0806***	−0.0641	−0.0815***	0.0420	−0.0452	−0.1640***	−0.1548***
long term	−0.1340***	−0.0618**	−0.0745***	−0.0386*	−0.0898***	−0.0597**	−0.1091***
IMMED							
short term	0.0260^{***s}	0.0042^{***s}	0.0000^{s}	0.0000***	0.0062***	0.0000***	0.0196***
medium term	0.0333***	0.0084^{***s}	0.0000^{**s}	0.0000***	0.0232***	0.0000***	0.0000^{s}
long term	0.0409^{***s}	0.0097^{***s}	0.0220^{***s}	0.0038***	0.0083***	0.0000^{s}	0.0129**

Table 5.6: Estimation of the constant α and the $\beta_{EuroMTS}$ in Equation (5.6) for on-the-run bonds

The regression (5.6) is estimated via OLS with Newey-West (1987) standard error (7 lags included). ***,**,* denote 1%, 5%, 10% significance respectively.

TWPBAS	Parameter	AT	BE	FR	DE	IT	NL	ES
short term	α	3.9992***	3.3037***	3.4621***	1.7353***	3.3961***	2.1454***	3.8576***
	$\beta_{EuroMTS}$	−0.5335***	0.1025	−0.0734	0.0412	−0.0598	−0.1541	−0.8886***
medium term	α	4.0568***	3.4826***	3.2830***	2.0210***	3.6444***	2.6078***	4.7045***
	$\beta_{EuroMTS}$	−0.2512**	−0.1303	−0.3036***	0.0810	−0.0242	−0.3470***	−0.6695***
long term	α	3.9743***	3.3686***	3.1526***	2.1286***	3.4608***	2.5895***	4.3473***
	$\beta_{EuroMTS}$	−0.3317***	−0.2416	−0.1581*	−0.1678***	−0.2462*	−0.2059**	−0.3445***
TWDEP(mil.)								
short term	α	17.4569***	17.7253***	17.4274***	17.3174***	17.7168***	17.9394***	17.3915***
	$\beta_{EuroMTS}$	−0.0373	0.0220	−0.0907*	0.1503**	−0.0356	0.3432***	−0.1551***
medium term	α	17.4055***	17.8976***	17.6614***	17.4050***	17.5571***	17.9927***	17.4495***
	$\beta_{EuroMTS}$	0.0808	0.1913***	0.1443***	0.3249***	−0.2253***	0.2836***	−0.1334**
long term	α	17.3303***	17.8030***	17.7193***	17.1235***	17.2853***	17.8499***	17.2627***
	$\beta_{EuroMTS}$	−0.0504	0.1585**	0.1964***	−0.0820	−0.0879	0.3410***	−0.0444
$\text{TWPBAS}_{\text{DEP}=10mil}$								
short term	α	4.0513***	3.3376***	3.5272***	1.8193***	3.4449***	2.1504***	3.9301***
	$\beta_{EuroMTS}$	−0.6342***	0.1069	−0.0410	−0.1468*	−0.1297	−0.1504	−0.9105***
medium term	α	4.1224***	3.5231***	3.3113***	2.0641***	3.6519***	2.6325***	4.7290***
	$\beta_{EuroMTS}$	−0.3091***	−0.1364	−0.2978***	0.0185	−0.0952	−0.3382**	−0.7171***
long term	α	4.0500***	3.4130***	3.2279***	2.1663***	3.4898***	2.6199***	4.4142***
	$\beta_{EuroMTS}$	−0.4111***	−0.2496	−0.2214**	−0.0972	−0.2752**	−0.2269**	−0.4054***
IMMED								
short term	α	0.9526***	0.9765***	0.9872***	0.9896***	0.9674***	0.9787***	0.9669***
	$\beta_{EuroMTS}$	0.0445***	0.0234***	0.0127	0.0054**	0.0094**	0.0201***	0.0193***
medium term	α	0.9461***	0.9735***	0.9881***	0.9918***	0.9614***	0.9817***	0.9603***
	$\beta_{EuroMTS}$	0.0509***	0.0266***	0.0147*	0.0068*	0.0243***	0.0117**	−0.0020
long term	α	0.9373***	0.9663***	0.9623***	0.9806***	0.9673***	0.9829***	0.9550***
	$\beta_{EuroMTS}$	0.0560***	0.0341***	0.0414***	0.0160***	0.0166***	0.0097*	0.0107

Table 5.7: Median of the difference in liquidity for the difference-in-difference series

The null hypothesis of the Wilcoxon signed-rank test is that the median of the difference in residual before and after the event is zero. The residual is obtained by differencing the two series generated from the EuroMTS market and the corresponding local MTS markets. ***,**,* denote 1%, 5%, 10% significance respectively. The superscript s indicates that the sign test is used.

TWPBAS	AT	BE	FR	DE	IT	NL	ES
short term	0.0007^{***s}	−0.0127***	−0.0013^{***s}	0.1082***	0.0134^{s}	0.0010^{s}	−0.0122***
medium term	0.0011^{***s}	−0.0064***	−0.0001	0.0279***	−0.0004^{s}	0.0033**	−0.0074^{***s}
long term	0.0003	−0.0067***	0.0621	0.0233***	0.0017^{***s}	0.0023**	−0.0163***
TWDEP							
short term	−0.0862***	−0.0189***	−0.0104***	0.0437***	−0.0443***	−0.0141	−0.0170
medium term	−0.0651***	−0.0055**	0.0212**	0.0064*	−0.0014	0.0205***	−0.0041
long term	−0.0438***	−0.0024	0.0134**	0.0121***	0.0197***	−0.0396***	0.0699***
TWPBAS$_{\text{DEP}=10mil}$							
short term	0.0001^{**s}	−0.0132***	−0.0052^{***s}	0.0530***	0.0178^{s}	0.0055	−0.0073***
medium term	0.0014***	−0.0057***	−0.0005^{s}	0.0143***	0.0031^{***s}	0.0037^{*s}	−0.0067^{***s}
long term	0.0052^{***s}	−0.0066***	−0.0103^{***s}	0.0103***	0.0014^{***s}	0.0001	−0.0090^{***s}
IMMED							
short term	0.0000^{***s}	0.0000^{***s}	0.0000^{s}	0.0000	−0.0009^{s}	0.0000^{s}	0.0000^{s}
medium term	0.0000^{s}	0.0000^{**s}	0.0000^{s}	0.0000*	0.0000^{s}	0.0000^{s}	−0.0003^{s}
long term	0.0000	0.0000	−0.0008^{***s}	0.0000***	−0.0000	0.0002***	−0.0006

the EuroMTS, improving its liquidity in the dimension of bid-ask spread, quoted depth and immediacy. Our results show how market participants can now execute the trade of bond portfolios with lower costs and greater immediacy.

Table 5.8: Median of the difference in liquidity after allowing the time trend to change after the event

The null hypothesis of the Wilcoxon signed-rank test is that the median of the difference in the residuals of the liquidity variables before and after the event is zero. For the test on the spread and depth proxies, the Wilcoxon signed rank test allowing the coefficient of the time variable t to change after the rule change in the detrending regression equation (5.7). For the test on the immediacy proxy we simply use the demeaned series. ***,**,* denote 1%, 5%, 10% significance respectively. The superscript s indicates that the sign test is used.

TWPBAS	AT	BE	FR	DE	IT	NL	ES
short term	-0.1053***	-0.1053***	-0.0413**	-0.0664***	0.0227	0.0261	-0.0106
medium term	-0.0524**	-0.0641***	-0.0871***	-0.0105*	-0.0327	0.0530**	0.0219
long term	-0.0312	-0.0586**	-0.1382***	-0.0199*	-0.0272*	0.0347*	0.0439
TWDEP							
short term	-0.0568***	0.0175	0.0147	0.0223***	0.0249	0.0417**	0.0172
medium term	-0.0057**	0.0149	0.0205**	0.0384***	0.0134	0.0738***	0.0011
long term	-0.0414***	0.0113	0.0171***	0.0291***	-0.0291***	0.0513***	0.0235
$\text{TWPBAS}_{\text{DEP}=10mil}$							
short term	-0.0961***	-0.0937***	-0.0433**	-0.0873***	-0.0325	0.0153	-0.0041
medium term	-0.0563***	-0.1014***	-0.0773***	-0.0147**	-0.0073	0.0569**	0.0204
long term	-0.0418***	-0.0525**	-0.1000***	-0.0431***	-0.0616*	0.0046	0.0490
IMMED							
short term	0.0345***	0.0034^{***s}	0.0000^{s}	-0.0001^{***}	0.0128***	0.0000***	0.0275***
medium term	0.0329***	0.0113***	0.0000^{s}	-0.0001^{***}	0.0072***	0.0000*	0.0114***
long term	0.0269***	0.0107^{***s}	0.0080^{***s}	0.0012***	0.0106^{***s}	0.0027^{**s}	0.0025

Table 5.9: Estimation of the constant α and the $\beta_{EuroMTS}$ in Equation (5.8) after allowing the time trend to change after the event.

The regression (5.8) is estimated via OLS with Newey-West (1987) standard error (7 lags included). ***,**,* denote 1%, 5%, 10% significance respectively.

TWPBAS	Parameter	AT	BE	FR	DE	IT	NL	ES
short term	α	4.0194***	2.9478***	3.0540***	1.8333***	3.5239***	2.0227***	3.8260***
	$\beta_{EuroMTS}$	-0.6282***	-0.2748*	-0.3998***	-0.1793*	-0.0495	-0.1444	-0.9016***
medium term	α	4.1593***	3.5108***	3.2048***	1.9923***	3.6772***	2.6245***	4.5784***
	$\beta_{EuroMTS}$	-0.3356***	-0.2659*	-0.3542***	0.0288	-0.0840	-0.2466**	-0.7495***
long term	α	4.0272***	3.5836***	3.2752***	2.3114***	4.1363***	2.8076***	4.9538***
	$\beta_{EuroMTS}$	-0.3505***	-0.1284	-0.3111***	-0.1324**	-0.2601**	-0.2767**	-0.6059***
TWDEP								
short term	α	17.4871***	17.6772***	17.5942***	17.3556***	17.8119***	17.9478***	17.4497***
	$\beta_{EuroMTS}$	-0.1312***	-0.0392	0.0098	0.2023***	-0.0169	0.3047***	-0.0408
medium term	α	17.4554***	17.8702***	17.7406***	17.5395***	17.7218***	18.0593***	17.4547***
	$\beta_{EuroMTS}$	-0.0913**	0.1105**	0.0974***	0.1354***	-0.0162	0.3659***	-0.1508***
long term	α	17.4025***	17.8344***	17.6740***	17.4427***	17.6450***	17.9103***	17.3803***
	$\beta_{EuroMTS}$	-0.2118***	0.1127**	0.0619	-0.0850	-0.2233***	0.2771***	-0.1040**
$\text{TWPBAS}_{\text{DEP}=10mil}$								
short term	α	4.0484***	2.9531***	3.0975***	1.8445***	3.5611***	2.0265***	3.8526***
	$\beta_{EuroMTS}$	-0.6472***	-0.2689*	-0.3834***	-0.3601***	0.0255	-0.1445	-0.9177***
medium term	α	4.1719***	3.5152***	3.2324***	1.9851***	3.6579***	2.6323***	4.5832***
	$\beta_{EuroMTS}$	-0.3263***	-0.2624*	-0.3486***	-0.0072	-0.0923	-0.2449**	-0.7783***
long term	α	4.0571***	3.5866***	3.3301***	2.4788***	4.1237***	2.8158***	4.9658***
	$\beta_{EuroMTS}$	-0.3149***	-0.1263	-0.3077***	-0.1017*	-0.2435**	-0.2751**	-0.6330***
IMMED								
short term	α	0.9488***	0.9904***	0.9957***	0.9982***	0.9621***	0.9906***	0.9635***
	$\beta_{EuroMTS}$	0.0485***	0.0139***	0.0121*	−0.0009*	0.0144***	0.0098***	0.0246***
medium term	α	0.9471***	0.9759***	0.9924***	0.9992***	0.9894***	0.9870***	0.9560***
	$\beta_{EuroMTS}$	0.0478***	0.0237***	0.0113**	−0.0032*	0.0109***	0.0087**	0.0084
long term	α	0.9414***	0.9665***	0.9804***	0.9975***	0.9826***	0.9837***	0.9590***
	$\beta_{EuroMTS}$	0.0465***	0.0292***	0.0180***	0.0026***	0.0118***	0.0074*	0.0016

Table 5.10: Median of the difference in liquidity after dropping one week before and one week after the rule change

The null hypothesis of the Wilcoxon signed-rank test is that the median of the difference in the residuals of the liquidity variables before and after the event is zero. For the test on the spread and depth proxies, we use the residuals of the detrending regression Equation (5.5). For the test on the immediacy proxy we simply use the demeaned series. ***,**,* denote 1%, 5%, 10% significance respectively. The superscript s indicates that the sign test is used.

TWPBAS	AT	BE	FR	DE	IT	NL	ES
short term	-0.1532***	-0.0513*	-0.0951***	-0.0213	-0.0200	-0.0113	-0.2267***
medium term	-0.0681***	-0.0888**	-0.0510***	0.0050	-0.0110	-0.0357*	-0.1669***
long term	-0.0672***	-0.0136	-0.1199***	-0.0328**	-0.0933***	-0.0864***	-0.1231***
TWDEP							
short term	-0.0681**	-0.0075	0.0059	0.0483***	-0.0105	0.1020***	-0.0061
medium term	0.0022	0.0319**	0.0144*	0.0319**	0.0115	0.1222***	-0.0681**
long term	-0.0722***	0.0135**	0.0223	-0.0190*	-0.0632***	0.0920***	-0.0303*
$\text{TWPBAS}_{\text{DEP}=10mil}$							
short term	-0.1446***	-0.0519*	-0.0871***	-0.0188***	-0.0025	-0.0219	-0.1841***
medium term	-0.0586***	-0.0573**	-0.0695***	-0.0024	-0.0328	-0.0573**	-0.1704***
long term	-0.0660***	-0.0020	-0.1262***	-0.0442**	-0.0949**	-0.1233***	-0.1923***
IMMED							
short term	0.0350***	0.0048^{***s}	0.0000^{s}	-0.0001**	0.0111***	0.0000***	0.0252***
medium term	0.0347***	0.0130***	0.0000^{s}	0.0000	0.0070***	0.0000	0.0056*
long term	0.0307***	0.0092^{***s}	0.0077^{***s}	0.0012***	0.0094^{***s}	0.0003^{s}	0.0004

Table 5.11: Estimation of the constant α and the $\beta_{EuroMTS}$ in Equation (5.6) after dropping one week before and one week after the rule change

The regression (5.6) is estimated via OLS with Newey-West (1987) standard error (7 lags included). ***,**,* denote 1%, 5%, 10% significance respectively.

TWPBAS	Parameter	AT	BE	FR	DE	IT	NL	ES
short term	α	4.0194***	2.9478***	3.0540***	1.8333***	3.5239***	2.0227***	3.8260***
	$\beta EuroMTS$	-0.6282***	-0.2748*	-0.3998***	-0.1793*	-0.0495	-0.1444	-0.9016***
medium term	α	4.1593***	3.5108***	3.2048***	1.9923***	3.6772***	2.6245***	4.5784***
	$\beta EuroMTS$	-0.3356***	-0.2659*	-0.3542***	0.0288	-0.0840	-0.2466**	-0.7495***
long term	α	4.0272***	3.5836***	3.2752***	2.3114***	4.1363***	2.8076***	4.9538***
	$\beta EuroMTS$	-0.3505***	-0.1284	-0.3111***	-0.1324**	-0.2601**	-0.2767**	-0.6059***
TWDEP								
short term	α	17.4871***	17.6772***	17.5942***	17.3556***	17.8119***	17.9478***	17.4497***
	$\beta EuroMTS$	-0.1312***	-0.0392	0.0098	0.2023***	-0.0169	0.3047***	-0.0408
medium term	α	17.4554***	17.8702***	17.7406***	17.5395***	17.7218***	18.0593***	17.4547***
	$\beta EuroMTS$	-0.0913**	0.1105**	0.0974***	0.1354***	-0.0162	0.3659***	-0.1508***
long term	α	17.4025***	17.8344***	17.6740***	17.4427***	17.6450***	17.9103***	17.3803***
	$\beta EuroMTS$	-0.2118***	0.1127**	0.0619	-0.0850	-0.2233***	0.2771***	-0.1040**
$\text{TWPBAS}_{\text{DEP}=10mil}$								
short term	α	4.0484***	2.9531***	3.0975***	1.8445***	3.5611***	2.0265***	3.8526***
	$\beta EuroMTS$	-0.6472***	-0.2689*	-0.3834***	-0.3601***	0.0255	-0.1445	-0.9177***
medium term	α	4.1719***	3.5152***	3.2324***	1.9851***	3.6579***	2.6323***	4.5832***
	$\beta EuroMTS$	-0.3263***	-0.2624*	-0.3486***	-0.0072	-0.0923	-0.2449**	-0.7783***
long term	α	4.0571***	3.5866***	3.3301***	2.4788***	4.1237***	2.8158***	4.9658***
	$\beta EuroMTS$	-0.3149***	-0.1263	-0.3077***	-0.1017*	-0.2435**	-0.2751**	-0.6330***
IMMED								
short term	α	0.9511***	0.9900***	0.9954***	0.9982***	0.9617***	0.9902***	0.9645***
	$\beta EuroMTS$	0.0473***	0.0140***	0.0122*	-0.0008	0.0150***	0.0101***	0.0240***
medium term	α	0.9466***	0.9749***	0.9919***	0.9991***	0.9889***	0.9863***	0.9563***
	$\beta EuroMTS$	0.0486***	0.0241***	0.0115**	-0.0033*	0.0112***	0.0088**	0.0076
long term	α	0.9419***	0.9651***	0.9796***	0.9975***	0.9820***	0.9835***	0.9601***
	$\beta EuroMTS$	0.0466***	0.0297***	0.0186***	0.0026***	0.0124***	0.0072*	0.0008

Chapter 6

Conclusion

6.1 Summary of the Thesis

With the recent sovereign debt crisis, we have observed a substantial change of intraday volatility and a sovereign bond flash crash in the US. The orderliness of markets is often hindered by sudden liquidity shortages and greater temporary volatility. This is a great concern for both regulators and practitioners. The wide usage of algorithmic trading and automated market making emphasized the importance of studying the high-frequency data and intraday volatility. The field of financial econometrics has been greatly expanded on this front with some exciting advancements and a better understanding of fast moving markets.

In Chapter 2, we review a whole range of topics closely related to intraday volatility. We commence the survey from a fundamental issue not seen in low-frequency literature, i.e. how to sample the tick-by-tick data. Three approaches are proposed, and each has its own merits. The calendar time sampling is arguably the most suitable for the MTS dataset. The realized variance are intimately related to intraday volatility as both use high-frequency data. We describe some stylized features about intraday volatility, starting from the common ones shared with daily volatility. The unique intraday periodicity is captured by various functional forms and it is associated mainly with scheduled macroeconomic announcements. Three papers have endeavored to accommodate all the stylized facts of intraday volatility and two of them lay the foundation of our empirical works, namely Andersen and Bollerslev (1998), and Engle and Sokalska (2012). Finally, we review the studies that have addressed the link between volatility and liquidity, stemming from the mixture of distribution hypothesis, where the relationship between daily price changes (a measure of volatility) and trading volumes (a measure of liquidity) has been one focus of this strand of literature.

We make three contributions to the literature of intraday volatility and liquidity. In Chapter 3, we evaluate different types of filters based on the multiplicative component GARCH model of Engle and Sokalska (2012), which we modify by replacing the simple average of intraday periodicity with a linear spline. We devise a benchmark for the evaluation in the spirit of Bandi and Russell (2008), which minimizes the distance between the summation of intraday return variance and daily realized variance. We find that the percentile approach with monthly recalculation is the best method to clean the market microstructure noise and help to estimate volatility dynamics. In the estimation section, we present the unique periodicity of intraday volatility of Italian and Spanish bonds, pointing to a U-shape or a J-shape pattern. We also compare the daily volatility forecast of the intraday GARCH with the daily GARCH model, and we find a superior performance of the intraday GARCH for the safer debts. The reason stems from the fact that the intraday GARCH has a much shorter half-life.

In Chapter 4, we first use a bivariate DCC model to examine the contagion effect across European government bonds during the sovereign debt crisis. We do not find any sustained high correlation between the bonds of Italy/Spain and the rest of the less risky bonds. The ECB intervention has been successful in restoring the confidence of bond markets in distressed European bond issuers. Second, we show how our model can be used for risk management purposes and for computing adequate VaRs. We demonstrate that European treasury bond portfolios are better diversified when containing Italian and Spanish bonds. We prove that the bivariate DCC model is capable of giving an adequate VaR measure for lower than 1% (inclusive) confidence level. In addition, the flexibility of the DCC model gives portfolio managers more space to allocate risk capital dynamically.

In Chapter 5, we change the course to study the importance of the liquidity and compare the merits of the newly installed auction market with the old dealership market. The inter-dealer MTS market is a major electronic trading platform for European fixed-income securities. There are many domestic MTS trading platforms and the EuroMTS, where European government bonds are traded. There are designated market makers who post two-sided quotes to provide liquidity and other market participants were price takers on the EuroMTS. On November 15, 2012, the EuroMTS market started allowing every market participant to submit one-sided limit orders. We construct several measures of liquidity, trying to study the impacts of this measure. And we find a significant improvement in liquidity for most of the European benchmark bonds. The daily average time-weighted percentage bid-ask spread became lower with a 20-basis-point reduction for short-term Spanish bonds. The depth generally became larger but smaller depth is found for Austrian and Italian long-term bonds after the event probably because the EuroMTS market has a lower minimum size (2 million bonds) requirement than the local MTS platforms and liquidity providers tend to supply more often the minimum amount. The finding is drawn from the Wilcoxon signed-rank test and a fully controlled regression including market volatility, macroeconomic announcements, and other carefully chosen liquidity-related variables. We test the robustness of our results by applying the same methodology to on-the-run bonds and comparing the EuroMTS with the local MTS markets in a difference-in-difference framework. The outcomes of the on-the-run bonds reinforce the conclusion for liquidity enhancement, which we describe for the bond portfolios consisting of all the benchmark bonds. Nonetheless, the difference-in-difference comparison has more mixed results and is not exempt from endogeneity problem as most of the market makers for one country's debts quote on both platforms.

6.2 Limitation of the Thesis and Future Research Suggestions

One limitation for the whole thesis would be that we did not formally test our model specifications. There are a few reasons that are worth to discuss here. In Chapter 3, we primarily want to evaluate the filters and the volatility forecasts. There is ample evidence already mentioned and suggesting that GARCH(1,1) is the best model for estimating and forecasting volatility in bond markets. The asymmetric effects of negative and positive returns in the intraday intervals has been tested, which yields insignificant parameters. More importantly, the more complicated structure adds little value to our comparison of filters and the exercise of generating better daily volatility forecasts. It will make the model estimation harder to converge and more sensitive to the extra assumptions. In Chapter 3, Engle (2002a) has shown some evidence that the DCC model is robust to misspecification of the univariate and multivariate models when the true correlation is simulated from different processes. The univariate model in Engle (2002a) is indeed a GARCH(1,1) model as we do for the intraday volatility. In Chapter 5, the Wilcoxon signed-rank test does not have the misspecification problem and the control variables in the regression are the best to our knowledge. We have consulted a wide range of literature to find the control variables, but a systematic model of liquidity remains few. The controlled regression examines the conditional change of the liquidity proxies as compared to the unconditional change suggested by the nonparametric test, and no contradictive conclusion is found between the two. In an unreported Ramsey RESET test, the specification is actually not adequate for describing the liquidity dynamics despite the fact that a high R^2 has often been produced by our regression. We mainly follow Chordia et al. (2005) for the specification of the regression.

For future research directions, we think that the interaction between liquidity and volatility is a burgeoning field, where many research projects are possible. As the volatility and liquidity are both intrinsically unobservable, applying statistical filters is also a promising subject. The unobservability points to a Bayesian approach, where auxiliary particle filters have been used to study intraday volatility. Stroud and Johannes (2014) have presented a very sophisticated research methodology, which is shown to possess some remarkable forecasting power for the realized variance.

Appendix A

In-Sample Estimation of DCC Models

Table A.1: The quasi-maximum likelihood estimation of the bivariate scalar DCC model in Equation (4.9) and the composite likelihood estimation of the multivariate scalar DCC model.

The robust t-values are reported in parentheses. 40 lags are chosen for the Newey-West (1987) standard error.

Country	a	b	Log Likelihood	Country	a	b	Log Likelihood
AT & BE	0.0267	0.9686	9489.10	FR & DE	0.0502	0.9399	27984.93
	(4.6852)	(132.9523)			(6.9371)	(87.2513)	
AT & FR	0.0407	0.9488	14676.23	FR & IT	0.0134	0.9857	5171.57
	(5.5017)	(92.5779)			(4.2463)	(292.0438)	
AT & DE	0.0383	0.9515	16827.35	FR & NL	0.0429	0.9493	25832.16
	(6.5325)	(116.5077)			(6.3687)	(107.3721)	
AT & IT	0.0091	0.9901	2637.46	FR & ES	0.0140	0.9856	9526.20
	(5.0840)	(500.0889)			(2.6450)	(176.7081)	
AT & NL	0.0454	0.9396	15700.37	DE & IT	0.0177	0.9809	5380.63
	(6.3683)	(87.3028)			(1.5106)	(73.6995)	
AT & ES	0.0178	0.9809	5010.90	DE & NL	0.0594	0.8863	34471.77
	(2.2930)	(113.7665)			(5.4595)	(29.8098)	
BE & FR	0.0279	0.9692	16430.05	DE & ES	0.0163	0.9832	9551.62
	(4.7679)	(132.3200)			(4.6493)	(268.4222)	
BE & DE	0.0217	0.9768	17768.15	IT & NL	0.0182	0.9802	5155.46
	(4.0894)	(165.3250)			(3.8818)	(181.2062)	
BE & IT	0.0168	0.9816	5134.70	IT & ES	0.0146	0.9834	5815.56
	(0.9876)	(49.7846)			(5.1845)	(292.5506)	
BE & NL	0.0262	0.9710	17130.48	NL & ES	0.0211	0.9779	9472.15
	(5.4008)	(170.7037)			(3.8573)	(166.3348)	
BE & ES	0.0130	0.9864	8562.18	Multivariate DCC	0.0219	0.9768	264761.15
	(3.0364)	(216.1870)			(9.1912)	(370.1861)	

Appendix B

Appendix to Chapter 5

Table B.3, B.4, B.5, and B.6 present the OLS estimation of Equation (5.6) for the TWPBAS, TWDEP, $\mathrm{TWPBAS}_{\mathrm{DEP}=10mil}$, and the immediacy respectively. The end-of-year dummy, which represents the holiday effect of the MTS dataset, is significantly positive, suggesting a widening of the spread as we have seen in Figure 5.2 and Figure 5.3. The four dummies, checked for the change of portfolio constitution, shows the sign and significance as we expect in most cases. To our surprise, the short-term bond portfolio of Italy has a slightly lower bid-ask spread than the rest of the time when deleting a standard bond, probably because the dated bonds take a relatively small part of the portfolio. If the liquidity of the portfolio is largely driven by standard bonds, then deleting a standard bond due to small residual maturity will improve the liquidity. The time trend t is no longer significant for portfolios of Germany and Spain compared to Table B.2, where only time trend variable is included. The explanatory power may be absorbed by other variables, such as macro news. The lagged market volatility is consistent with the findings of Chordia et al. (2005) who show that higher market uncertainty leads to lower liquidity. While market volatility and macro news represent the effect of informed trading, the log order imbalances essentially capture the uninformed trading. For short term debts, a net purchase would reduce the spread whereas a net sale would increase the spread. Combining the two variable would reduce the significance and magnitude of their coefficients. The ECB Main Refinancing Rate and the quarterly report of GDP growth are the most important news in determining the bid-ask spread. Apart from Austria and Germany, the shock to the ECB's decision on interest rate has significantly increased the average daily time-weighted bid-ask spread for the European countries. On the other hand, a positive shock to GDP has reduced the spread. We conjecture that the negative signs of GDP are driven by speculators' trading on the news. More importantly, the unexpected growth of GDP signals a recovery of the European economy; it thus relieves the concerns of market participants about debt insolvency and the funding liquidity of the governments.

Table B.1: Macroeconomic announcements considered in the OLS regression

US Economic News	Label in the regression	Frequency	CET time
Chicago Purchasing Manager	CPM_{US}	Monthly	15:45
Consumer Confidence	CC_{US}	Monthly	16:00
CPI	CPI_{US}	Monthly	14:30
Durable Goods Orders	DGO_{US}	Monthly	14:30
Initial Jobless Claims	IJC_{US}	Weekly	14:30
New Home Sales	NHS_{US}	Monthly	16:00
Change in Nonfarm Payrolls	NP_{US}	Monthly	14:30
Uni. of Mich. Sentiment	$Senti_{US}$	Monthly	15:55
Unemployment Rate	UR_{US}	Monthly	14:30
Euro-Area Economic News			
Eurozone Manufacturing PMI	PMI_{EU}	Monthly	10:00
CPI Estimate	CPI_{EU}	Monthly	11:00
GDP	GDP_{EU}	Quarterly	11:00
Economic Confidence	EC_{EU}	Monthly	11:00
Consumer Confidence	CC_{EU}	Monthly	11:00
PPI	PPI_{EU}	Monthly	11:00
Retail Sales	RS_{EU}	Monthly	11:00
Unemployment Rate	UR_{EU}	Monthly	11:00
ECB Main Refinancing Rate	ECB_{EU}	Daily	13:45
Long Term Refinance Operation	$LTRO_{EU}$	Semi-monthly	11:00

Table B.2: The estimation of the detrending regression

A detrending regression is applied to the time series of log average liquidity proxies computed for all bond portfolios. This Table presents the estimated coefficients of the detrending regression: $y_t = \alpha + \beta t + \varepsilon_t$ Newey-West (1987) standard errors (with 7 lags) are used. The log of average liquidity proxies are used as the regressands.

TWPBAS	Parameter	AT	BE	FR	DE	IT	NL	ES
short term	α	4.1398***	3.1093***	3.2111***	1.8821***	3.6484***	2.1202***	4.0925***
	β	−0.0045***	−0.0043***	−0.0041***	−0.0007***	−0.0035***	−0.0028***	−0.0030***
medium term	α	4.2063***	3.6249***	3.3440***	2.0097***	3.8700***	2.6969***	4.7723***
	β	−0.0039***	−0.0038***	−0.0031***	−0.0007***	−0.0037***	−0.0021***	−0.0040***
long term	α	4.0665***	3.6182***	3.3005***	2.3893***	4.2772***	2.9264***	5.1061***
	β	−0.0030***	−0.0030***	−0.0019***	−0.0011***	−0.0039***	−0.0018***	−0.0042***
TWDEP								
short term	α	17.4206***	17.6464***	17.5696***	17.3019***	17.7484***	17.8546***	17.3765***
	β	0.0006***	0.0007***	0.0018***	−0.0002***	0.0009***	0.0004**	0.0011***
medium term	α	17.4262***	17.7751***	17.6729***	17.4823***	17.6775***	17.9037***	17.4203***
	β	0.0006***	0.0012***	0.0011***	−0.0003***	0.0011***	0.0003	0.0005***
long term	α	17.3795***	17.7381***	17.6059***	17.4548***	17.6515***	17.7802***	17.3476***
	β	0.0002	0.0008***	0.0004***	−0.0004***	0.0006***	0.0001	0.0002*
$\text{TWPBAS}_{\text{DEP}=10mil}$								
short term	α	4.2384***	3.1419***	3.2630***	1.9346***	3.6633***	2.1274***	4.1940***
	β	−0.0047***	−0.0044***	−0.0042***	−0.0006***	−0.0036***	−0.0028***	−0.0031***
medium term	α	4.2831***	3.6754***	3.3907***	2.0139***	3.8553***	2.7149***	4.8435***
	β	−0.0041***	−0.0039***	−0.0032***	−0.0004***	−0.0037***	−0.0022***	−0.0039***
long term	α	4.1612***	3.6728***	3.4013***	2.5308***	4.2469***	2.9581***	5.1747***
	β	−0.0030***	−0.0031***	−0.0021***	−0.0011***	−0.0038***	−0.0018***	−0.0041***

Figure B.1: The plots of average daily depth for all bonds included in the short-term bond portfolio

The details about the construction of daily depth from intraday data is well-explained in the text. This is the time-series plot of short-term bond portfolio daily depth.

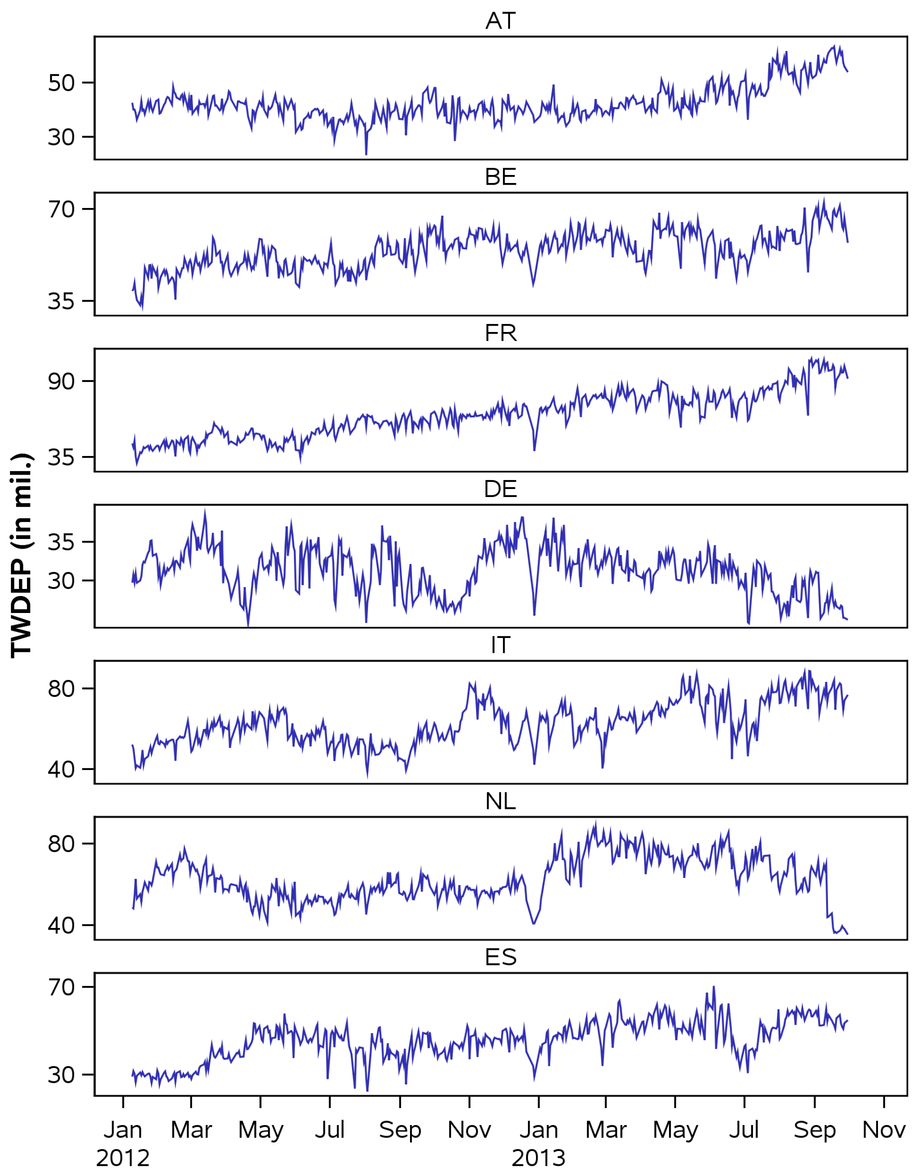


Figure B.2: The plots of average daily immediacy for all bonds included in the short-term bond portfolio

The details about the construction of daily immediacy from intraday data is well-explained in the text. This is the time-series plot of short-term bond portfolio daily immediacy.

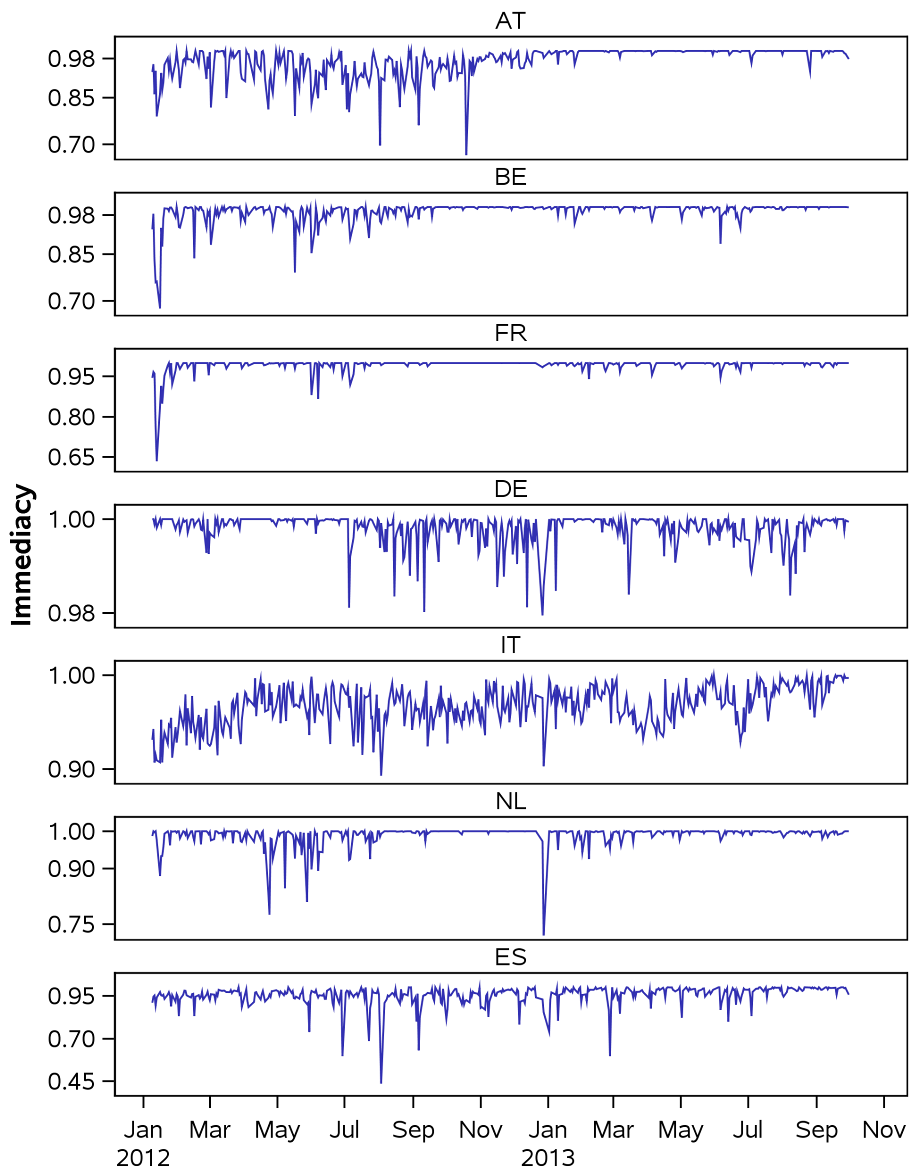


Table B.3: The estimation of the regression equation (5.6) for the TWPBAS of short-term bonds

The regression

$y_t = \alpha + \beta_{EuroMTS} * Dummy_{t,EuroMTS} + \sum_{k=1}^{K} \beta_{control,k} * Control\ Variable_{t,k} + \varepsilon_t$ is estimated by OLS with Newey-West (1987) standard error (with 7 lags). The $Dummy_{t,EuroMTS}$ takes value 1 since November 15, 2012 (inclusive) and value 0 before that date. ***,**,* denote 1%, 5%, 10% significance respectively. The labels for various macroeconomic announcements are listed in Table B.1. All the control dummy variables are described in Section 5.7. $Dummy_{holiday}$ captures the end-of-year effect and $Dummy_{ent\_standard}$ takes value 1 when a newly issued bond is introduced to the portfolio. $Dummy_{del\_standard}$ manifests liquidity change when deleting a standard bond. The same labelling logic applies to the dated bonds.

Regressors	AT	BE	FR	DE	IT	NL	ES
α	3.9851***	2.9139***	3.0338***	1.8285***	3.5180***	2.0269***	3.8056***
$Dummy_{EuroMTS}$	−0.5087***	−0.2571*	−0.3820***	−0.1706*	−0.0395	−0.1428	−0.8024***
$Dummy_{holiday}$	0.1020**	0.4312***	0.5533***	0.4432***	0.1358**	1.5062***	0.9990***
$Dummy_{ent\_standard}$			−0.0993***	−0.0623	−0.0060	−0.1143*	
$Dummy_{del\_standard}$				−0.0436	−0.1029*		
$Dummy_{ent\_dated}$	0.2456	0.0038	−0.0917	0.1583*	−0.1466*	0.1904***	0.2924***
$Dummy_{del\_dated}$	0.0937	−0.0621	0.1399			−0.1777***	−0.3008***
t	−0.0028***	−0.0034***	−0.0028***	−0.0001	−0.0033***	−0.0022***	−0.0002
Market Volatility$_{t-1}$	0.2011	1.0027***	0.7779***	0.2401*	0.7009***	0.8642***	0.9408**
Negative Log Order Imbalance	−0.0128	−0.0206***	−0.0023	−0.0061	−0.0031	0.0096	−0.0034
Positive Log Order Imbalance	−0.0233***	0.0170*	0.0064	−0.0228**	0.0128**	−0.0078	−0.0059
CCI_{US}	0.0431	0.0462	0.0329	0.0231	0.1133**	−0.0105	0.0079
CC_{EU}	0.0133	0.0283	−0.0245	−0.0108	0.1061**	−0.0599	0.0747**
CPI_{EU}	−0.0565	−0.0879	−0.0270	−0.1714***	−0.0731	0.1543***	0.0767
CPI_{US}	0.0117	−0.0160	0.0601	−0.0332	0.0702*	−0.0549	−0.0288
CPM_{US}	−0.0157	0.0469	0.0405	0.0332	0.0483	−0.0373	−0.0981

DGO_{US}	−0.0064	−0.0052	−0.0166	0.0092	0.0359	−0.0097	0.0253
ECB_{EU}	−0.2061**	0.4369**	1.0057***	0.1370	1.0400***	0.5718***	0.7929***
EC_{EU}	−0.0253	0.0177	0.0328	−0.0537**	−0.0103	0.1326***	−0.0280
GDP_{EU}	−0.2157***	−0.0689	−0.0911**	−0.0688	−0.2759***	−0.0760	−0.1501
IJC_{US}	0.0248	−0.0142	−0.0190	0.0110	0.0404*	0.0172	0.0318
$LTRO_{EU}$	0.0086	−0.0372	−0.0093	−0.0196	0.0343	0.0098	0.0161
NHS_{US}	0.0465	0.0464	0.0487	0.0244	0.0453	0.0337	0.0157
NP_{US}	0.0833*	0.0805	0.0589	0.0392	0.1517**	0.2080**	0.0717
PMI_{EU}	−0.0282	0.0151	0.0451*	0.0209	0.0465*	−0.0221	−0.0811**
PPI_{EU}	0.0142	0.0134	0.0572	0.0434	0.0226	0.0076	0.0524
RS_{EU}	0.0555	0.0591	0.0106	0.0829*	−0.0297	0.0206	−0.0289
$Senti_{US}$	−0.0048	−0.0585	−0.0176	0.0027	0.0508	−0.0265	0.0360
UR_{EU}	0.1749*	0.0929	−0.0312	0.1576**	−0.0657	−0.1144	−0.2696***
UR_{US}	−0.0039	0.0148	0.0533	−0.0256	0.0865	−0.0084	0.1240
Adj.R^2	0.8677	0.7731	0.8601	0.2337	0.7441	0.6697	0.6197

Table B.4: The estimation of the regression equation (5.6) for the TWDEP of short-term bonds

The regression $y_t = \alpha + \beta_{EuroMTS} * Dummy_{t,EuroMTS} + \sum_{k=1}^{K} \beta_{control,k} * Control\ Variable_{t,k} + \varepsilon_t$ is estimated via OLS with Newey-West (1987) standard error (with 7 lags). ***,**,* denote 1%, 5%, 10% significance respectively. The labels for various macroeconomic announcements are listed in Table B.1. All the dummy variables are described in Section 5.7. $Dummy_{holiday}$ captures the holiday effect and $Dummy_{ent\_standard}$ takes value 1 when a new bond is introduced to the portfolio. $Dummy_{del\_standard}$ manifests liquidity change for deleting a standard bond. The same labelling logic applies to the dated bonds.

Regressor	AT	BE	FR	DE	IT	NL	ES
α	17.4421***	17.6677***	17.5844***	17.3525***	17.8127***	17.9297***	17.4107***
$Dummy_{EuroMTS}$	−0.0587	−0.0340	0.0091	0.1835***	−0.0128	0.2740***	−0.0169

$Dummy_{holiday}$	−0.1905***	−0.2595***	−0.5340***	−0.2739***	−0.4614***	−0.5698***	−0.4892***
$Dummy_{ent\_standard}$			0.1442***	−0.0237	0.0278	0.1238***	
$Dummy_{del\_standard}$				0.0851***	−0.0872***		
$Dummy_{ent\_dated}$	0.0354	−0.0386	−0.0334	−0.0752***	0.0148	0.0061	−0.1024***
$Dummy_{del\_dated}$	−0.1244***	0.0028	0.2000***			−0.1019***	0.1644***
t	0.0008***	0.0008***	0.0018***	−0.0008***	0.0009***	−0.0005	0.0011***
Market Volatility$_{t-1}$	−0.3543***	−0.2021**	0.0016	−0.0387	−0.2776***	−0.0407	−0.3036**
Negative Log Order Imbalance	−0.0101*	0.0012	0.0024	−0.0070*	−0.0008	−0.0048	−0.0016
Positive Log Order Imbalance	−0.0101***	−0.0041	−0.0033*	−0.0032	−0.0118***	−0.0019	−0.0035
CCI_{US}	0.0165	−0.0268	0.0027	0.0093	−0.0483	0.0339	−0.0640*
CC_{EU}	−0.0021	0.0091	−0.0217	0.0204	−0.0336	0.0203	0.0001
CPI_{EU}	−0.0227	0.0075	0.0567**	0.0349*	−0.0046	−0.0817**	0.0619
CPI_{US}	−0.0169	0.0305**	0.0376**	0.0062	−0.0226	0.0247	0.0045
CPM_{US}	0.0132	0.0236	−0.0334	0.0101	0.0248	0.0178	0.0347
DGO_{US}	−0.0011	0.0146	−0.0427	0.0025	0.0138	−0.0048	0.0557**
ECB_{EU}	−0.8126***	−0.2729***	−0.1405***	−0.2760***	−0.8388***	−0.1337*	−1.2090***
EC_{EU}	−0.0297*	0.0308**	0.0243	−0.0002	−0.0127	−0.0137	0.0163
GDP_{EU}	−0.0086	0.0345	−0.0520	−0.0041	0.0823***	0.1149***	0.1068*
IJC_{US}	−0.0053	−0.0184*	−0.0240**	−0.0118*	−0.0498***	−0.0279*	−0.0351**
$LTRO_{EU}$	0.0073	0.0021	0.0072	0.0193*	−0.0047	0.0039	0.0232
NHS_{US}	−0.0108	0.0053	−0.0046	0.0065	−0.0095	−0.0122	−0.0532*
NP_{US}	−0.0607**	−0.0778***	−0.0410	−0.0328*	0.0154	−0.0482	0.0072
PMI_{EU}	−0.0224	0.0118	−0.0164	−0.0135**	0.0133	0.0413	0.0405*
PPI_{EU}	−0.0079	0.0127	0.0082	0.0076	−0.0256	0.0029	0.0403
RS_{EU}	0.0040	0.0017	−0.0405	0.0092	−0.0280	0.0122	−0.0390
$Senti_{US}$	0.0071	−0.0030	−0.0076	−0.0065	−0.0214*	−0.0266	−0.0101

UR_{EU}	0.0880	0.0146	−0.0089	0.0233	−0.0572	0.0662	−0.1047**
UR_{US}	−0.0163	0.0111	−0.0064	−0.0175	−0.0933***	−0.0319	−0.0684***
Adj. R^2	0.3389	0.4695	0.8511	0.3508	0.4799	0.2433	0.4699

Table B.5: The estimation of the regression equation (5.6) for the $\text{TWPBAS}_{\text{DEP}=10mil}$ of short-term bonds

The regression $y_t = \alpha + \beta_{EuroMTS} * Dummy_{t,EuroMTS} + \sum_{k=1}^{K} \beta_{control,k} * Control\ Variable_{t,k} + \varepsilon_t$ is estimated via OLS with Newey-West (1987) standard error (with 7 lags). ***,**,* denote 1%, 5%, 10% significance respectively. The labels for various macroeconomic announcements are listed in Table B.1. All the dummy variables are described in Section 5.7. $Dummy_{holiday}$ captures the holiday effect and $Dummy_{ent\_standard}$ takes value 1 when a standard bond is introduced to the portfolio. $Dummy_{del\_standard}$ manifests liquidity change when deleting a standard bond. The same labelling logic applies to the dated bonds.

Regressor	AT	BE	FR	DE	IT	NL	ES
α	4.0602***	2.9346***	3.0906***	1.8438***	3.5549***	2.0337***	3.8722***
$Dummy_{EuroMTS}$	−0.5966***	−0.2555	−0.3626***	−0.3310***	0.0262	−0.1402	−0.8367***
$Dummy_{holiday}$	0.1091**	0.4798***	0.5825***	0.4606***	0.2543***	1.5167***	1.1150***
$Dummy_{ent\_standard}$			−0.1049***	−0.0667	−0.0094	−0.1116	
$Dummy_{del\_standard}$				−0.0365	−0.0498		
$Dummy_{ent\_dated}$	0.2454	0.0094	−0.1025	0.1289	−0.1193*	0.1937***	0.2667***
$Dummy_{del\_dated}$	0.1444**	−0.0458	0.1513*			−0.1822***	−0.3295***
t	−0.0027***	−0.0034***	−0.0029***	0.0005	−0.0036***	−0.0023***	−0.0002
Market Volatility$_{t-1}$	0.2063	1.0548***	0.7500***	0.1898	0.7787***	0.8631***	1.1790***
Negative Log Order Imbalance	−0.0094	−0.0217***	−0.0022	−0.0079	−0.0011	0.0083	−0.0051
Positive Log Order Imbalance	−0.0249***	0.0197*	0.0080	−0.0220**	0.0089	−0.0079	−0.0057
CCI_{US}	0.0174	0.0477	0.0375	0.0274	0.1369**	−0.0110	0.0193
CC_{EU}	0.0111	0.0338	−0.0294	−0.0102	0.0599*	−0.0614	0.0799*
CPI_{EU}	−0.0490	−0.0937	−0.0472	−0.1524***	−0.0904	0.1524***	0.0814

CPI_{US}	0.0037	−0.0185	0.0621*	−0.0231	0.1205***	−0.0583	−0.0393
CPM_{US}	−0.0213	0.0455	0.0486	0.0365	0.0144	−0.0397	−0.1110
DGO_{US}	−0.0079	−0.0128	−0.0300	−0.0002	−0.0053	−0.0104	0.0087
ECB_{EU}	0.3234***	0.5286***	1.0601***	0.1895*	1.2796***	0.5512***	1.8065***
EC_{EU}	−0.0178	0.0194	0.0313	−0.0539*	0.0217	0.1315***	−0.0241
GDP_{EU}	−0.2195***	−0.0774	−0.0896**	−0.0720	−0.2139***	−0.0782	−0.2025**
IJC_{US}	0.0274	−0.0078	−0.0179	0.0119	0.0384	0.0167	0.0459
$LTRO_{EU}$	−0.0022	−0.0427	−0.0115	−0.0205	0.0308	0.0089	0.0058
NHS_{US}	0.0548	0.0440	0.0368	0.0282	0.0183	0.0372	0.0746
NP_{US}	0.1021**	0.1259	0.0944	0.0156	0.1660***	0.2064**	0.1141
PMI_{EU}	−0.0334	0.0215	0.0500*	0.0250	0.0448*	−0.0210	−0.0726**
PPI_{EU}	0.0329	0.0157	0.0561	0.0467	0.0557	0.0054	0.0696
RS_{EU}	0.0642*	0.0516	0.0013	0.0872*	−0.0148	0.0201	−0.0304
$Senti_{US}$	−0.0026	−0.0564	−0.0132	0.0047	0.0359	−0.0278	0.0299
UR_{EU}	0.1482	0.0847	−0.0147	0.1336**	−0.0683	−0.1107	−0.3020***
UR_{US}	0.0053	0.0024	0.0374	−0.0038	0.0489	−0.0100	0.1456
Adj. R^2	0.8875	0.7684	0.8637	0.2290	0.7707	0.6647	0.6231

Table B.6: The estimation of the regression equation (5.6) for the immediacy of short-term bonds

The regression $y_t = \alpha + \beta_{EuroMTS} * Dummy_{t,EuroMTS} + \sum_{k=1}^{K} \beta_{control,k} * Control\ Variable_{t,k} + \varepsilon_t$ is estimated via OLS with Newey-West (1987) standard error (with 7 lags). ***,**,* denote 1%, 5%, 10% significance respectively. The labels for various macroeconomic announcements are listed in Table B.1. All the dummy variables are described in Section 5.7. $Dummy_{holiday}$ captures the holiday effect and $Dummy_{ent\_standard}$ takes value 1 when a standard bond is introduced to the portfolio. $Dummy_{del\_standard}$ manifests liquidity change when deleting a standard bond. The same labelling logic applies to the dated bonds.

Regressor	AT	BE	FR	DE	IT	NL	ES

α	0.9488***	0.9904***	0.9957***	0.9982***	0.9621***	0.9906***	0.9635***
$Dummy_{EuroMTS}$	0.0485***	0.0139***	0.0121*	−0.0009*	0.0144***	0.0098***	0.0246***
$Dummy_{holiday}$	−0.0200***	−0.0107***	−0.0160***	−0.0153***	−0.0734***	−0.2730***	−0.1350***
$Dummy_{ent\_standard}$			0.0029	0.0014***	0.0090	−0.0101***	
$Dummy_{del\_standard}$				−0.0013	−0.0318***		
$Dummy_{ent\_dated}$	−0.0247	−0.0030	0.0068	−0.0029	0.0039	−0.0378	0.0096
$Dummy_{del\_dated}$	−0.0518***	0.0011	0.0212**			−0.0025	0.0296***
Market Volatility$_{t-1}$	−0.0471*	−0.0785**	−0.0300*	0.0036	−0.0150	−0.0143	−0.2173***
Negative Log Order Imbalance	−0.0029**	0.0002	0.0005	−0.0001	−0.0004	0.0017	−0.0006
Positive Log Order Imbalance	0.0005	−0.0032**	−0.0026	−0.0002	0.0003	−0.0004	−0.0002
CCI_{US}	0.0183***	0.0034	0.0018	−0.0005	−0.0101***	0.0028	−0.0094
CC_{EU}	0.0008	−0.0006	0.0052	−0.0011	−0.0034	0.0041*	−0.0046
CPI_{EU}	−0.0086*	0.0037	0.0060	0.0012	0.0042	−0.0036	0.0030
CPI_{US}	−0.0004	0.0063	0.0058	−0.0036*	−0.0050	0.0058	0.0064
CPM_{US}	0.0068	0.0017	0.0010	0.0003	−0.0029	0.0016	0.0092
DGO_{US}	0.0014	0.0053	0.0067	0.0007	−0.0007	0.0040**	0.0094
ECB_{EU}	−0.4518***	−0.0591***	0.0192**	0.0013	0.0026	0.0424***	−0.7787***
EC_{EU}	−0.0050	−0.0021	0.0037	0.0002	−0.0019	−0.0013	−0.0043
GDP_{EU}	0.0203**	0.0097**	0.0085*	0.0015***	0.0177***	0.0042	0.0275**
IJC_{US}	−0.0046	−0.0058	−0.0070*	−0.0003	−0.0016	−0.0004	−0.0116**
$LTRO_{EU}$	0.0092**	0.0064**	0.0007	−0.0000	−0.0012	−0.0026	0.0040
NHS_{US}	−0.0030	0.0007	0.0025	0.0011***	0.0034	−0.0079	−0.0451
NP_{US}	−0.0181	−0.0377**	−0.0286**	0.0007	−0.0063	−0.0164*	−0.0455
PMI_{EU}	0.0001	−0.0056*	−0.0052*	0.0005	−0.0022	−0.0047	−0.0123
PPI_{EU}	−0.0201	−0.0024	0.0025	−0.0000	−0.0041	0.0039*	−0.0165
RS_{EU}	−0.0017	0.0063	0.0056*	−0.0018**	0.0001	0.0035	0.0008

$Senti_{US}$	−0.0047	0.0004	−0.0084	−0.0005	−0.0055*	0.0032*	0.0006
UR_{EU}	0.0327***	0.0061	0.0023	−0.0021*	−0.0103	−0.0006	0.0219
UR_{US}	−0.0116*	0.0105	0.0059	0.0003	−0.0036	0.0026	−0.0133
Adj. R^2	0.3487	0.1837	0.1301	0.0482	0.1373	0.2839	0.2071

Bibliography

Admati, A. R., and P. Pfleiderer. 1988. A Theory of Intraday Patterns: Volume and Price Variability. *Review of Financial Studies* 1: 3–40.

Aielli, G. P. 2013. Dynamic Conditional Correlation: On Properties and Estimation. *Journal of Business & Economic Statistics* 31: 282–299.

Aitken, M. J., R. M. Cook, F. Harris, and T. H. McInish. 2009. Market Design and Execution Cost for Matched Securities Worldwide. *The Journal of Trading* 2009: 38–76.

Albanesi, S., and B. Rindi. 2000. The Quality of the Italian Treasury Bond Market, Asymmetric Information and Transaction Costs. *Annales Economie et de Statistique* 1–19.

Almgren, R., and N. Chriss. 2001. Optimal Execution of Portfolio Transactions. *Journal of Risk* 3: 5–39.

Amihud, Y., H. Mendelson, and B. Lauterbach. 1997. Market Microstructure and Securities Values: Evidence From the Tel Aviv Stock Exchange. *Journal of Financial Economics* 45: 365–390.

Bibliography

Anand, A., C. Tanggaard, and D. G. Weaver. 2009. Paying for Market Quality. *The Journal of Financial and Quantitative Analysis* 44: 1427–1457.

Andersen, T. G., and T. Bollerslev. 1997. Intraday Periodicity and Volatility Persistence in Financial Markets. *Journal of Empirical Finance* 4: 115–158.

Andersen, T. G., and T. Bollerslev. 1998. Deutsche Mark-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies. *The Journal of Finance* 53: 219–265.

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

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens. 2001. 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. 2003a. Modeling and Forecasting Realized Volatility. *Econometrica* 71: 579–625.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and C. Vega. 2003b. Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange. *The American Economic Review* 93: 38–62.

Andersson, M., L. J. Overby, and S. Sebestian. 2009. Which News Moves the Euro Area Bond Market? *German Economic Review* 10: 1–31.

Babecka Kucharcukova, O., P. Claeys, and B. Vasicek. 2016. Spillover of the ECB's monetary policy outside the euro area: How different is conventional from unconventional policy? *Journal of Policy Modeling* 38: 199–225.

Bibliography

Bae, K.-H., G. A. Karolyi, and R. M. Stulz. 2003. A New Approach to Measuring Financial Contagion. *Review of Financial Studies* 16: 717–763.

Baillie, R. T., T. Bollerslev, and H. O. Mikkelsen. 1996. Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* 74: 3–30.

Balduzzi, P., E. J. Elton, and T. C. Green. 2001. Economic News and Bond Prices: Evidence From the US Treasury Market. *Journal of Financial and Quantitative Analysis* 36: 523–543.

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

Barclay, M. J. 1997. Bid-Ask Spreads and the Avoidance of Odd-Eighth Quotes On Nasdaq: An Examination of Exchange Listings. *Journal of Financial Economics* 45: 35–60.

Barclay, M. J., W. G. Christie, J. H. Harris, E. Kandel, and P. H. Schultz. 1999. Effects of Market Reform On the Trading Costs and Depths of Nasdaq Stocks. *The Journal of Finance* 54: 1–34.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard. 2008. Designing Realized Kernels to Measure the Ex Post Variation of Equity Prices in the Presence of Noise. *Econometrica* 76: 1481–1536.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard. 2009. Realized Kernels in Practice: Trades and Quotes. *The Econometrics Journal* 12: C1–C32.

Bibliography

Barndorff-Nielsen, O. E., and N. Shephard. 2002a. Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)* 64: 253–280.

Barndorff-Nielsen, O. E., and N. Shephard. 2002b. Estimating Quadratic Variation Using Realized Variance. *Journal of Applied Econometrics* 17: 457–477.

Barndorff-Nielsen, O. E., and N. Shephard. 2004a. Econometric Analysis of Realized Covariation: High Frequency Based Covariance, Regression, and Correlation in Financial Economics. *Econometrica* 72: 885–925.

Barndorff-Nielsen, O. E., and N. Shephard. 2004b. Power and Bipower Variation with Stochastic Volatility and Jumps. *Journal of Financial Econometrics* 2: 1–37.

Bauwens, L., W. Ben Omrane, and P. Giot. 2005. News Announcements, Market Activity and Volatility in the Euro/Dollar Foreign Exchange Market. *Journal of International Money and Finance* 24: 1108–1125.

Bauwens, L., and P. Giot. 2000. The Logarithmic ACD Model: An Application to the Bid-Ask Quote Process of Three NYSE Stocks. *Annals of Economics and Statistics* 117–149.

Bauwens, L., and P. Giot. 2003. Asymmetric ACD Models: Introducing Price Information in ACD Models. *Empirical Economics* 28: 709–731.

Bauwens, L., S. Laurent, and J. V. Rombouts. 2006. Multivariate GARCH Models: A Survey. *Journal of Applied Econometrics* 21: 79–109.

Bibliography

Beber, A., M. W. Brandt, and K. A. Kavajecz. 2009. Flight-to-Quality or Flight-to-Liquidity? Evidence From the Euro-Area Bond Market. *Review of Financial Studies* 22: 925–957.

Beetsma, R., M. Giuliodori, F. de Jong, and D. Widijanto. 2013. Spread the News: The Impact of News on the European Sovereign Bond Markets during the Crisis. *Journal of International Money and Finance* 34: 83–101.

Bennett, P., and L. Wei. 2006. Market Structure, Fragmentation, and Market Quality. *Journal of Financial Markets* 9: 49–78.

Berkowitz, J., P. Christoffersen, and D. Pelletier. 2009. Evaluating Value-at-Risk Models with Desk-Level Data. *Management Science* 57: 2213–2227.

Bessembinder, H., K. M. Kahle, W. F. Maxwell, and D. Xu. 2009. Measuring Abnormal Bond Performance. *Review of Financial Studies* 22: 4219–4258.

Billio, M., and M. Caporin. 2009. A Generalized Dynamic Conditional Correlation Model for Portfolio Risk Evaluation. *Mathematics and Computers in Simulation* 79: 2566–2578.

Blume, L., D. Easley, and M. O'Hara. 1994. Market Statistics and Technical Analysis: The Role of Volume. *The Journal of Finance* 49: 153–181.

Blume, M. E., D. B. Keim, and S. A. Patel. 1991. Returns and Volatility of Low-Grade Bonds 1977-1989. *The Journal of Finance* 46: 49–74.

Boehmer, E., G. Saar, and L. Yu. 2005. Lifting the Veil: An Analysis of Pre-Trade Transparency at the NYSE. *The Journal of Finance* 60: 783–815.

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

Bollerslev, T. 1990. Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Model. *The Review of Economics and Statistics* 498–505.

Bollerslev, T., J. Cai, and F. M. Song. 2000. Intraday Periodicity, Long Memory Volatility, and Macroeconomic Announcement Effects in the US Treasury Bond Market. *Journal of Empirical Finance* 7: 37–55.

Bollerslev, T., R. Y. Chou, and K. F. Kroner. 1992. ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence. *Journal of Econometrics* 52: 5–59.

Bollerslev, T., and I. Domowitz. 1993. Trading Patterns and Prices in the Interbank Foreign Exchange Market. *The Journal of Finance* 48: 1421–1443.

Bollerslev, T., J. Litvinova, and G. Tauchen. 2006. Leverage and Volatility Feedback Effects in High-Frequency Data. *Journal of Financial Econometrics* 4: 353–384.

Bollerslev, T., and V. Todorov. 2011. Tails, Fears, and Risk Premia. *The Journal of Finance* 66: 2165–2211.

Boudt, K., C. Croux, and S. Laurent. 2011. Robust Estimation of Intraweek Periodicity in Volatility and Jump Detection. *Journal of Empirical Finance* 18: 353–367.

Brogaard, J., T. Hendershott, and R. Riordan. 2014. High-Frequency Trading and Price Discovery. *Review of Financial Studies* 27: 2267–2306.

Brownlees, C. T., and G. M. Gallo. 2006. Financial Econometric Analysis at Ultra-High Frequency: Data Handling Concerns. *Computational Statistics & Data Analysis* 51: 2232–2245.

Burns, P., R. F. Engle, and J. J. Mezrich. 1998. Correlations and Volatilities of Asynchronous Data. *The Journal of Derivatives* 5: 7–18.

Busch, T., B. J. Christensen, and M. A. Nielsen. 2011. The Role of Implied Volatility in Forecasting Future Realized Volatility and Jumps in Foreign Exchange, Stock, and Bond Markets. *Journal of Econometrics* 160: 48–57.

Campbell, J. Y., A. W.-C. Lo, and A. C. MacKinlay. 1997. *The Econometrics of Financial Markets*, vol. 2. Princeton, NJ: Princeton University Press.

Campbell, S. D. 2005. A Review of Backtesting and Backtesting Procedures. *Finance and Economics Discussion Series, Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington D.C.*, 2005-21.

Cappiello, L., R. F. Engle, and K. Sheppard. 2006. Asymmetric Dynamics in the Correlations of Global Equity and Bond Returns. *Journal of Financial Econometrics* 4: 537–572.

Cespa, G., and T. Foucault. 2014. Illiquidity Contagion and Liquidity Crashes. *Review of Financial Studies* 27: 1615–1660.

Chacko, G. C., J. W. Jurek, and E. Stafford. 2008. The Price of Immediacy. *The Journal of Finance* 63: 1253–1290.

Chelley-Steeley, P. L., and L. Skvortsov. 2010. Efficiency and the Trading System:

The Case of SETSmm. *Journal of International Financial Markets, Institutions and Money* 20: 509–518.

Cheung, Y. C., F. de Jong, and B. Rindi. 2005. Trading European Sovereign Bonds: The Microstructure of the MTS Trading Platforms. *ECB Working Paper*, 432.

Chiang, T. C., B. N. Jeon, and H. Li. 2007. Dynamic Correlation Analysis of Financial Contagion: Evidence From Asian Markets. *Journal of International Money and Finance* 26: 1206–1228.

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

Chordia, T., A. Subrahmanyam, and Q. Tong. 2014. Have Capital Market Anomalies Attenuated in the Recent Era of High Liquidity and Trading Activity? *Journal of Accounting and Economics* 58: 41–58.

Chou, R. Y. 1988. Volatility Persistence and Stock Valuations: Some Empirical Evidence Using GARCH. *Journal of Applied Econometrics* 3: 279–294.

Christiansen, C. 2007. Volatility-Spillover Effects in European Bond Markets. *European Financial Management* 13: 923–948.

Christie, W. G., and R. D. Huang. 1994. Market Structures and Liquidity: A Transactions Data Study of Exchange Listings. *Journal of Financial Intermediation* 3: 300–326.

Christie, W. G., and P. H. Schultz. 1994. Why Do NASDAQ Market Makers Avoid Odd-Eighth Quotes? *The Journal of Finance* 49: 1813–1840.

Christoffersen, P. F. 1998. Evaluating Interval Forecasts. *International Economic Review* 39: 841–862.

Clark, P. K. 1973. A Subordinated Stochastic Process Model with Finite Variance for Speculative Prices. *Econometrica* 41: 135–155.

Clements, M. P., and A. B. Galvao. 2008. Macroeconomic Forecasting With Mixed-Frequency Data. *Journal of Business & Economic Statistics* 26: 546–554.

Clyde, P., P. Schultz, and M. Zaman. 1997. Trading Costs and Exchange Delisting: The Case of Firms that Voluntarily Move from the American Stock Exchange to the Nasdaq. *The Journal of Finance* 52: 2103–2112.

Colacito, R., R. F. Engle, and E. Ghysels. 2011. A Component Model for Dynamic Correlations. *Journal of Econometrics* 164: 45–59.

Corsi, F. 2009. A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics* 7: 174–196.

Corsi, F., and F. Audrino. 2012. Realized Covariance Tick-by-Tick in Presence of Rounded Time Stamps and General Microstructure Effects. *Journal of Financial Econometrics* 10: 591–616.

Corsi, F., S. Mittnik, C. Pigorsch, and U. Pigorsch. 2008. The Volatility of Realized Volatility. *Econometric Reviews* 27: 46–78.

Dacorogna, M. M., U. A. Müller, R. J. Nagler, R. B. Olsen, and O. V. Pictet. 1993. A Geographical Model for the Daily and Weekly Seasonal Volatility in the Foreign Exchange Market. *Journal of International Money and Finance* 12: 413–438.

Dacorogna, M. M., U. A. Müller, O. V. Pictet, and R. B. Olsen. 1998. Modelling Short-Term Volatility with GARCH and HARCH Models. In D. Christian and B. Zhou (eds.), *Nonlinear Modelling of High Frequency Financial Time Series*, John Wiley & Sons Inc.

Darbha, M., and A. Dufour. 2013. Microstructure of Euro-Area Government Bond Market. In H. K. Baker and H. Kiymaz (eds.), *Market Microstructure in Emerging and Developed Markes*, Hoboken, NJ: John Wiley.

De Goeij, P., and W. Marquering. 2004. Modeling the Conditional Covariance between Stock and Bond Returns: A Multivariate GARCH Approach. *Journal of Financial Econometrics* 2: 531–564.

De Jong, F., T. Nijman, and A. Roell. 1995. A Comparison of the Cost of Trading French Shares on the Paris Bourse and on SEAQ International. *European Economic Review* 39: 1277–1301.

Deo, R., C. Hurvich, and Y. Lu. 2006. Forecasting Realized Volatility Using a Long-Memory Stochastic Volatility Model: Estimation, Prediction and Seasonal Adjustment. *Journal of Econometrics* 131: 29–58.

Diaz, A., J. J. Merrick Jr., and E. Navarro. 2006. Spanish Treasury Bond Market Liquidity and Volatility Pre- and Post-European Monetary Union. *Journal of Banking & Finance* 30: 1309–1332.

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

Dimitriou, D., D. Kenourgios, and T. Simos. 2013. Global Financial Crisis and Emerging Stock Market Contagion: A Multivariate FIAPARCH-DCC Approach. *International Review of Financial Analysis* 30: 46–56.

Ding, Z., and R. F. Engle. 2001. Large Scale Conditional Covariance Matrix Modeling, Estimation and Testing. *NYU Working Paper*, S-DRP-01-07.

Dionne, G., P. Duchesne, and M. Pacurar. 2009. Intraday Value at Risk (IVAR) Using Tick-by-Tick Data with Application to the Toronto Stock Exchange. *Journal of Empirical Finance* 16: 777–792.

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

Dufour, A., and F. Skinner. 2004. MTS Time Series: Market and Data Description for the European Bond and Repo Database. *ICMA Centre Discussion Paper, Henley Business School*, 2004-07.

Dufour, A., A. Stancu, and S. Varotto. 2016. The Equity-Like Features of Sovereign Bonds. *ICMA Centre Discussion Paper, Henley Business School*, ICM-2014-16.

Dungey, M., R. Fry, B. Gonzalez-Hermosillo, and V. Martin. 2006. Contagion in International Bond Markets during the Russian and the LTCM Crises. *Journal of Financial Stability* 2: 1–27.

Dungey, M., R. Fry, B. Gonzalez-Hermosillo, and V. L. Martin. 2005. Empirical Modelling of Contagion: A Review of Methodologies. *Quantitative Finance* 5: 9–24.

Dunne, P. G., M. J. Moore, and R. Portes. 2007. Benchmark Status in Fixed-Income Asset Markets. *Journal of Business Finance & Accounting* 34: 1615–1634.

Ederington, L. H., and J. H. Lee. 1993. How Markets Process Information: News Releases and Volatility. *The Journal of Finance* 48: 1161–1191.

Engle, R. 2002a. Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business & Economic Statistics* 20: 339–350.

Engle, R. 2002b. New Frontiers for ARCH Models. *Journal of Applied Econometrics* 17: 425–446.

Engle, R., and R. Colacito. 2006. Testing and Valuing Dynamic Correlations for Asset Allocation. *Journal of Business & Economic Statistics* 24: 238–253.

Engle, R., and R. Ferstenberg. 2007. Execution Risk. *Journal of Portfolio Management* 33: No.2:34–44.

Engle, R. F. 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica* 50: 987–1007.

Engle, R. F. 2000. The Econometrics of Ultra-High-Frequency Data. *Econometrica* 68: 1–22.

Engle, R. F., M. J. Fleming, E. Ghysels, and G. Nguyen. 2012. Liquidity, Volatility, and Flights to Safety in the US Treasury Market: Evidence From A New Class of Dynamic Order Book Models. *FRB of New York Staff Report*, 590.

Engle, R. F., E. Ghysels, and B. Sohn. 2008. On the Economic Sources of Stock Market Volatility. In *AFA 2008 New Orleans Meetings Paper*.

Engle, R. F., and G. Lee. 1999. A Long-Run and Short-Run Component Model of Stock Return Volatility. In *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive WJ Granger*, 475–497, Oxford: Oxford University Press.

Engle, R. F., and S. Manganelli. 2004. CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. *Journal of Business & Economic Statistics* 22: 367–381.

Engle, R. F., and A. J. Patton. 2001. What Good Is a Volatility Model? *Quantitative Finance* 1: 237–245.

Engle, R. F., and J. G. Rangel. 2008. The Spline-GARCH Model for Low-Frequency Volatility and Its Global Macroeconomic Causes. *Review of Financial Studies* 21: 1187–1222.

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

Engle, R. F., N. Shephard, and K. Sheppard. 2007. Fitting and Testing Vast Dimensional Time-Varying Covariance Models. *NYU Working Paper*, FIN-07-046.

Engle, R. F., and K. Sheppard. 2001. Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH. *NBER Working Paper*, 8554.

Engle, R. F., and M. E. Sokalska. 2012. Forecasting Intraday Volatility in the US Equity Market. Multiplicative Component GARCH. *Journal of Financial Econometrics* 10: 54–83.

Eser, F., and B. Schwaab. 2016. Evaluating the impact of unconventional monetary policy measures: Empirical evidence from the ECB's Securities Markets Programme. *Journal of Financial Economics* 119: 147–167.

Federal Reserve Bank of St Louis. 2009. The Financial Crisis. A Timeline of Events and Policy Actions, https://www.stlouisfed.org/financial-crisis/full-timeline.

Federal Reserve, Joint Staff Report. 2015. *The U.S. Treasury Market On October 15, 2014*.

Fender, I., and U. Lewrick. 2015. Shifting Tides – Market Liquidity and Market-Making in Fixed Income Instruments. *BIS Quarterly Review* March: 97–109.

Fleming, M. 2001. Measuring Treasury Market Liquidity. *FRB of New York Staff Report*, 133.

Fleming, M., and J. Lopez. 1999. Heat Waves, Meteor Showers, and Trading Volume: An Analysis of Volatility Spillovers in the US Treasury Market. *FRB of New York Staff Report*, 82.

Forbes, K. J., and R. Rigobon. 2002. No Contagion, Only Interdependence: Measuring Stock Market Comovements. *The Journal of Finance* 57: 2223–2261.

Foster, A. J. 1995. Volume-Volatility Relationships for Crude Oil Futures Markets. *Journal of Futures Markets* 15: 929–951.

French, K. R., G. Schwert, and R. F. Stambaugh. 1987. Expected Stock Returns and Volatility. *Journal of Financial Economics* 19: 3–29.

Gajewski, J.-F., and C. Gresse. 2007. Centralised Order Books versus Hybrid Order

Books: A Paired Comparison of Trading Costs on NSC (Euronext Paris) and SETS (London Stock Exchange). *Journal of Banking & Finance* 31: 2906–2924.

Gallant, A. 1981. On the Bias in Flexible Functional Forms and an Essentially Unbiased Form. *Journal of Econometrics* 15: 211–245.

Gençay, R., M. Dacorogna, U. A. Müller, O. Pictet, and R. Olsen. 2001a. *An Introduction to High-Frequency Finance*. Cambridge: Academic Press.

Gençay, R., F. Selcuk, and B. J. Whitcher. 2001b. *An Introduction to Wavelets and Other Filtering Methods in Finance and Economics*. Cambridge: Academic Press.

Ghysels, E., J. Idier, S. Manganelli, and O. Vergote. 2014. A High Frequency Assessment of the ECB Securities Markets Programme. *ECB Working Paper*, 1642.

Ghysels, E., P. Santa-Clara, and R. Valkanov. 2005. There is a Risk-Return Trade-Off After All. *Journal of Financial Economics* 76: 509–548.

Giot, P. 2000. Time Transformations, Intraday Data, and Volatility Models. *Journal of Computational Finance* 4: 31–62.

Giot, P. 2005. Market Risk Models for Intraday Data. *The European Journal of Finance* 11: 309–324.

Glosten, L. R. 1989. Insider Trading, Liquidity, and the Role of the Monopolist Specialist. *The Journal of Business* 62: 211–235.

Glosten, L. R., R. Jagannathan, and D. E. Runkle. 1993. On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *The Journal of Finance* 48: 1779–1801.

Gregoriou, A. 2015. Market Quality of Dealer versus Hybrid Markets for Illiquid Securities: New Evidence from the FTSE AIM Index. *The European Journal of Finance* 21: 466–485.

Griffin, J. E., and R. C. Oomen. 2011. Covariance Measurement in the Presence of Non-Synchronous Trading and Market Microstructure Noise. *Journal of Econometrics* 160: 58–68.

Grossman, S. J., and M. H. Miller. 1988. Liquidity and Market Structure. *The Journal of Finance* 43: 617–633.

Hansen, P. R., Z. Huang, and H. H. Shek. 2012. Realized GARCH: A Joint Model for Returns and Realized Measures of Volatility. *Journal of Applied Econometrics* 27: 877–906.

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.

Hansen, P. R., and A. Lunde. 2006. Realized Variance and Market Microstructure Noise. *Journal of Business & Economic Statistics* 24: 127–161.

Hardle, W. K., N. Hautsch, and A. Mihoci. 2012. Modelling and Forecasting Liquidity Supply Using Semiparametric Factor Dynamics. *Journal of Empirical Finance* 19: 610–625.

Harju, K., and S. M. Hussain. 2011. Intraday Seasonalities and Macroeconomic News Announcements. *European Financial Management* 17: 367–390.

Harris, L. 1987. Transaction Data Tests of the Mixture of Distributions Hypothesis. *The Journal of Financial and Quantitative Analysis* 22: 127–141.

Harris, L. 2002. *Trading and Exchanges: Market Microstructure for Practitioners*. Oxford: Oxford University Press.

Hasbrouck, J. 1991. Measuring the Information Content of Stock Trades. *The Journal of Finance* 46: 179–207.

Hasbrouck, J. 1993. Assessing the Quality of a Security Market: A New Approach to Transaction-Cost Measurement. *Review of Financial Studies* 6: 191–212.

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

Hasbrouck, J. 2012. *High Frequency Quoting: Short-Term Volatility in Bids and Offers*. Stern School of Business, New York University.

Hayashi, T., and S. Kusuoka. 2008. Consistent Estimation of Covariation under Nonsynchronicity. *Statistical Inference for Stochastic Processes* 11: 93–106.

Hayashi, T., and N. Yoshida. 2005. On Covariance Estimation of Non-Synchronously Observed Diffusion Processes. *Bernoulli* 11: 359–379.

Henke, H., and B. Lauterbach. 2005. Firm-Initiated and Exchange-Initiated Transfers to Continuous Trading: Evidence From the Warsaw Stock Exchange. *Journal of Financial Markets* 8: 309–323.

Hodrick, L. S., and P. C. Moulton. 2009. Liquidity: Considerations of a Portfolio Manager. *Financial Management* 38: 59–74.

Hsieh, D. A., and A. W. Kleidon. 1996. Bid-Ask Spreads in Foreign Exchange Markets: Implications for Models of Asymmetric Information. In J. A. Frankel, G. Galli, and A. Giovannini (eds.), *The Microstructure of Foreign Exchange Markets*, 41–72, Chicago: University of Chicago Press.

Huang, R. D., J. Cai, and X. Wang. 2002. Information-Based Trading in the Treasury Note Interdealer Broker Market. *Journal of Financial Intermediation* 11: 269–296.

Jain, P. K., J.-C. Kim, and Z. Rezaee. 2008. The Sarbanes-Oxley Act of 2002 and Market Liquidity. *Financial Review* 43: 361–382.

Jones, C. M., O. Lamont, and R. L. Lumsdaine. 1998. Macroeconomic News and Bond Market Volatility. *Journal of Financial Economics* 47: 315–337.

JP Morgan. 1996. *Riskmetrics: Technical Document*. Morgan Guaranty Trust Company of New York.

Karpoff, J. M. 1987. The Relation Between Price Changes and Trading Volume: A Survey. *The Journal of Financial and Quantitative Analysis* 22: 109–126.

Kim, O., and R. E. Verrecchia. 1991. Market Reaction to Anticipated Announcements. *Journal of Financial Economics* 30: 273–309.

Kirilenko, A., A. S. Kyle, M. Samadi, and T. Tuzun. 2016. The Flash Crash: The Impact of High Frequency Trading on an Electronic Market. *The Journal of Finance* forthcoming.

Koike, Y. 2016. Quadratic Covariation Estimation of an Irregularly Observed Semimartingale with Jumps and Noise. *Bernoulli* 22: 1894–1936.

Kupiec, P. H. 1995. Techniques for Verifying the Accuracy of Risk Measurement Models. *The Journal of Derivatives* 3: 73–84.

Kyle, A. S. 1985. Continuous Auctions and Insider Trading. *Econometrica* 53: 1315–1335.

Lamoureux, C. G., and W. D. Lastrapes. 1990. Heteroskedasticity in Stock Return Data: Volume versus GARCH Effects. *The Journal of Finance* 45: 221–229.

Lee, C. M. C. 1993. Market Integration and Price Execution for NYSE-Listed Securities. *The Journal of Finance* 48: 1009–1038.

Liu, C., and J. M. Maheu. 2012. Intraday Dynamics of Volatility and Duration: Evidence from Chinese Stocks. *Pacific-Basin Finance Journal* 20: 329–348.

Lo, A. W., and A. C. MacKinlay. 1990. An Econometric Analysis of Nonsynchronous Trading. *Journal of Econometrics* 45: 181–211.

Longstaff, F. A., and E. S. Schwartz. 1992. Interest Rate Volatility and the Term Structure: A Two Factor General Equilibrium Model. *The Journal of Finance* 47: 1259–1282.

Lopez, J. A. 1998. Methods for Evaluating Value-at-Risk Estimates. *Federal Reserve Bank of New York Economic Policy Review* 4.

Maheu, J. M., and T. H. McCurdy. 2011. Do High-Frequency Measures of Volatility Improve Forecasts of Return Distributions? *Journal of Econometrics* 160: 69–76.

Manganelli, S. 2005. Duration, Volume and Volatility Impact of Trades. *Journal of Financial Markets* 8: 377–399.

Martens, M., Y.-C. Chang, and S. J. Taylor. 2002. A Comparison of Seasonal Adjustment Methods When Forecasting Intraday Volatility. *Journal of Financial Research* 25: 283–299.

McAleer, M., F. Chan, S. Hoti, and O. Lieberman. 2008. Generalized Autoregressive Conditional Correlation. *Econometric Theory* 24: 1554–1583.

Merton, R. C. 1973. An Intertemporal Capital Asset Pricing Model. *Econometrica* 41: 867–887.

Muscarella, C. J., and M. S. Piwowar. 2001. Market Microstructure and Securities Values: Evidence from the Paris Bourse. *Journal of Financial Markets* 4: 209–229.

Müller, U. A., M. M. Dacorogna, R. D. Dave, R. B. Olsen, O. V. Pictet, and J. E. von Weizsacker. 1997. Volatilities of Different Time Resolutions – Analyzing the Dynamics of Market Components. *Journal of Empirical Finance* 4: 213–239.

Müller, U. A., M. M. Dacorogna, R. B. Olsen, O. V. Pictet, M. Schwarz, and C. Morgenegg. 1990. Statistical Study of Foreign Exchange Rates, Empirical Evidence of a Price Change Scaling Law, and Intraday Analysis. *Journal of Banking & Finance* 14: 1189–1208.

Nelson, D. B. 1990. ARCH Models as Diffusion Approximations. *Journal of Econometrics* 45: 7–38.

Nelson, D. B. 1991. Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica* 59: 347–370.

Newey, W. K., and D. McFadden. 1994. Large Sample Estimation and Hypothesis Testing. In R. F. Engle and D. L. McFadden (eds.), *Handbook of Econometrics*, vol. 4, 2111–2245, North-Holland: Elsevier.

Newey, W. K., and K. D. West. 1987. Hypothesis Testing with Efficient Method of Moments Estimation. *International Economic Review* 28: 777–787.

Nimalendran, M., and G. Petrella. 2003. Do "Thinly-Traded" Stocks Benefit from Specialist Intervention? *Journal of Banking & Finance* 27: 1823–1854.

O'Hara, M., and M. Ye. 2011. Is Market Fragmentation Harming Market Quality? *Journal of Financial Economics* 100: 459–474.

Oomen, R. C. A. 2005. Properties of Bias-Corrected Realized Variance under Alternative Sampling Schemes. *Journal of Financial Econometrics* 3: 555–577.

Oomen, R. C. A. 2006. Properties of Realized Variance under Alternative Sampling Schemes. *Journal of Business & Economic Statistics* 24: 219–237.

Pasquariello, P., and C. Vega. 2009. The On-the-Run Liquidity Phenomenon. *Journal of Financial Economics* 92: 1–24.

Pattipeilohy, C., J. W. van den End, M. Tabbae, J. Frost, and J. de Haan. 2013. Unconventional Monetary Policy of the ECB during the Financial Crisis: An Assessment and New Evidence. *DNB Working Paper*, 381.

Patton, A. J. 2011. Volatility Forecast Comparison Using Imperfect Volatility Proxies. *Journal of Econometrics* 160: 246–256.

Perold, A. F. 1988. The Implementation Shortfall: Paper versus Reality. *The Journal of Portfolio Management* 14: 4–9.

Petersen, M. A., and D. Fialkowski. 1994. Posted versus Effective Spreads: Good Prices or Bad Quotes? *Journal of Financial Economics* 35: 269–292.

Poirier, D. J. 1973. Piecewise Regression Using Cubic Splines. *Journal of the American Statistical Association* 68: 515–524.

Poon, S.-H., and C. W. Granger. 2003. Forecasting Volatility in Financial Markets: A Review. *Journal of Economic Literature* 41: 478–539.

Ranaldo, A. 2009. Segmentation and Time-of-Day Patterns in Foreign Exchange Markets. *Journal of Banking & Finance* 33: 2199–2206.

Rangel, J. G., and R. F. Engle. 2012. The Factor-Spline-GARCH Model for High and Low Frequency Correlations. *Journal of Business & Economic Statistics* 30: 109–124.

Roll, R. 1984. A Simple Implicit Measure of the Effective Bid-Ask Spread in an Efficient Market. *The Journal of Finance* 39: 1127–1139.

Schmidt, H., and P. Iversen. 1993. Automating German Equity Trading: Bid-Ask Spreads on Competing Systems. *Journal of Financial Services Research* 6: 373–397.

Schwert, G. W. 1989. Why Does Stock Market Volatility Change over Time? *The Journal of Finance* 44: 1115–1153.

Schwert, G. W. 1990. Stock Volatility and the Crash of '87. *Review of Financial Studies* 3: 77–102.

Silvennoinen, A., and T. Terasvirta. 2009. Multivariate GARCH Models. In T. Mikosch, J.-P. Kreiß, A. R. Davis, and T. G. Andersen (eds.), *Handbook of Financial Time Series*, 201–229, Berlin, Heidelberg: Springer.

Stroud, J. R., and M. S. Johannes. 2014. Bayesian Modeling and Forecasting of 24-

Hour High-Frequency Volatility. *Journal of the American Statistical Association* 109: 1368–1384.

Taylor, S. J., and X. Xu. 1997. The Incremental Volatility Information in One Million Foreign Exchange Quotations. *Journal of Empirical Finance* 4: 317–340.

Tsay, R. S. 1987. Conditional Heteroscedastic Time Series Models. *Journal of The American Statistical Association* 82: 590–604.

Tse, Y. K., and A. K. C. Tsui. 2002. A Multivariate Generalized Autoregressive Conditional Heteroscedasticity Model with Time-Varying Correlations. *Journal of Business & Economic Statistics* 20: 351–362.

Vijh, A. M. 1990. Liquidity of the CBOE Equity Options. *The Journal of Finance* 45: 1157–1179.

Voev, V., and A. Lunde. 2007. Integrated Covariance Estimation Using High-Frequency Data in the Presence of Noise. *Journal of Financial Econometrics* 5: 68–104.

Wang, G. H. K., and J. Yau. 2000. Trading Volume, Bid-ask Spread, and Price Volatility in Futures Markets. *Journal of Futures Markets* 20: 943–970.

Wegman, E. J., and I. W. Wright. 1983. Splines in Statistics. *Journal of the American Statistical Association* 78: 351–365.

Wood, R. A., T. H. McInish, and J. K. Ord. 1985. An Investigation of Transactions Data for NYSE Stocks. *The Journal of Finance* 40: 723–739.

Zhang, L., P. A. Mykland, and Y. Ai-Sahalia. 2005. A Tale of Two Time Scales. *Journal of the American Statistical Association* 100: 1394–1411.