UNIVERSITY OF SALERNO
Department of Economics and Statistics

DOCTORAL THESIS

Essays on the modelling and prediction of financial volatility and trading volumes

Author:
Antonio NAIMOLI

Supervisor:
Prof Giuseppe STORTI

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in the discipline of

Analisi Economica, Giuridica e Statistica delle Politiche, dei Mercati e delle Imprese

XXIX Cycle

Metodi Statistici per la Valutazione

Academic Year 2015/2016

The significant problems we have cannot be solved at the same level of thinking with which we created them.

Albert Einstein

Acknowledgements

The completion of this doctoral dissertation would not have been possible without the support of several people. I wish to express my deep thanks to everyone who offered kind assistance or support to me during the course of my PhD.

First and foremost, I would like to express my sincerest gratitude to my supervisor Prof Giuseppe Storti for his excellent guidance and patience during my PhD study. His motivation, flexibility and immense knowledge helped me in all the time of research and without his support and constant feedback this thesis could not have reached the current stage. Whatever will come in the future, I will always be grateful to him for believing in me.

I gratefully acknowledge the constructive comments and suggestions offered by the referees Prof Giampiero M. Gallo and Prof Malvina Marchese. I express my sincere gratitude for your review, which will help to improve the quality of the thesis significantly. I would also like to express my appreciation to the committee members Prof Gianluca Cubadda, Prof Giovanni De Luca and Prof Michele La Rocca for their time, interest, and helpful comments.

I would like to convey my gratitude to the Director of the Department of Economics and Statistics (DISES) Prof Cira Perna and to the PhD Coordinator Prof Sergio Destefanis for the academic support and the facilities provided to carry out the research work at the university. I am also very grateful to Prof Alessandra Amendola who supported and helped me in numerous ways during various stages of my PhD. I thank you kindly for keeping the Lab office door open and available for all the times I needed it.

Part of this research was conducted during my visiting period at the Université Catholique de Louvain in Belgium and I would like to express my gratitude to Prof Luc Bauwens for his constructive comments and valuable suggestions on my research work. I am also very thankful to all the members of the Center for Operations Research and Econometrics,

who were always so helpful and provided me their assistance during my time at Louvain la Neuve.

My deep appreciation goes to all the professors at the DISES, since they have contributed immensely to my professional and personal life. A very special thanks to Maria Rizzo, Fabio Penna and Pasquale Sabatino for being very kind and patient and always willing to lend their support at any time. I would also like to thank my colleagues and friends in the department for all the enjoyable times that we have shared.

I am particularly thankful to my dear friend Giancarlo Cesale, since I gained a lot from him through his personal and academic interactions.

During the studies a large part of the time was made enjoyable by the group of my best friends that has now become part of my life. I am very grateful to each of them for the time spent together.

Of course no acknowledgement would be complete without giving thanks to my family. I owe a lot to my parents, Vincenzo and Beatrice, who encouraged and helped me at every stage of my personal and academic life and longed to see this achievement come true. I am grateful to my brothers, Erminio and Stefano, for their encouragement and support in all my pursuits. I also wish to thank my loved ones who have left this life too soon. Their memory will be with me always.

Last, but certainly not least, I would like to thank Chiara for all her love and encouragement. She has been by my side throughout this PhD, living every single minute of it with me. Her devotion, unconditional love and support, patience, optimism and advice was more valuable than she could ever imagine.

Antonio Naimoli
06 June 2017

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The rapid growth of the financial markets and the continuous development of new and more complex financial instruments, have remarkably increased the importance of modelling and forecasting asset returns volatility as accurately as possible. Volatility plays a key role in asset pricing, risk management, asset allocation, portfolio optimization and in several other financial applications. Since the introduction of the *autoregressive conditional heteroskedastic* (ARCH) model by Engle (1982), followed by the generalization to GARCH of Bollerslev (1986), a long list of financial volatility models have been developed and extended to explain various stylized facts about financial returns and volatility, such as volatility clustering, leverage effects, leptokurtosis, and long memory. The central element within the GARCH framework is the specification of the conditional variance. However, the main problem in using this class of models is that the volatility is latent and, therefore, it is not directly observable. The standard GARCH model usually relies on daily stock returns to estimate the latent conditional volatility, using all current and past daily squared returns to provide expectations on future volatility.

More recently, the increasing availability of ultra high-frequency data for a host of different financial instruments, has given rise to the question

whether such data provides additional information when compared to the commonly used daily data. This issue has led to a new branch in financial econometrics literature, revealing that the high-frequency data are much more informative about the price process not only on an intra-daily level, but also on a daily level.

Andersen and Bollerslev (1998) show that daily latent volatility can be more accurately measured by the daily aggregated squared intraday returns, reducing the measurement error in quantifying the true latent integrated volatility. Assuming that the underlying stochastic logarithmic price is a continuous Brownian semi-martingale process, the integrated variance, which is the quadratic returns variation, can be consistently estimated by the realized variance. The daily *realized variance*, defined as the sum of all available intraday high-frequency squared returns, provides a natural ex-post return variability measure, since, under suitable conditions, it is an unbiased and highly efficient estimator of return volatility as discussed in Andersen et al. (2001) and Barndorff-Nielsen and Shephard (2002). During the last fifteen years, exploiting the superior information contained in intraday return data, several non-parametric realized volatility estimators have been proposed, such as the *Bipower* and *Tripower Variation* of Barndorff-Nielsen and Shephard (2004b), the *Subsampled RV* of Zhang et al. (2005), the *Realized Kernels* of Barndorff-Nielsen et al. (2008), the *medRV* and *minRV* of Andersen et al. (2012), among others.

In this regard, following the theoretical results of Barndorff-Nielsen and Shephard (2002), Meddahi (2002) and Andersen et al. (2003), several papers have studied the properties of these estimators. Because of a host of practical market microstructure frictions, the realized variance suffers from bias problems related to the sampling frequency. The highest

possible sampling frequency should be used for efficiency, but sampling at ultra high-frequency tends to bias the estimate of the realized measures. Aït-Sahalia et al. (2005), Zhang et al. (2005), Hansen and Lunde (2006), Bandi and Russell (2006) and Bandi and Russell (2008), among others, have discussed various solutions to account for the trade-off between bias and variance, though in most empirical application the 5 minutes realized variance is used. This choice is supported by empirical evidence showing that in practical application it is often difficult to beat this simple benchmark (Liu et al., 2015).

A second relevant research question, that is currently being actively debated in the literature, is related to the benefits, in term of forecasting accuracy of the daily volatility, that can derive from incorporating realized variance estimators in volatility forecasting models. Engle (2002) introduce the GARCH-X model including the realized variance as an exogenous variable in the volatility dynamics equation. In this framework Engle (2002) and Engle and Gallo (2006), have developed a Multiplicative Error Model (MEM) by specifying a separate dynamic equation for the realized measure, completing the GARCH-X model. Extensions of the MEM model to the multivariate case have been proposed in Cipollini et al. (2006, 2013), developing the vector MEM. Along the same track, Shephard and Sheppard (2010) proposed the HEAVY model, extending the GARCH-X model in such a way that a multi-step volatility forecast is feasible. A similar model which accounts for realized measures is the Realized GARCH model introduced by Hansen et al. (2012). This model is conceptually closely related to the standard GARCH model, but includes a measurement equation that relates the realized measures to latent conditional volatility[1]. The Realized GARCH model not only outperforms

[1]The HEAVY and the Realized GARCH models can be represented as special cases of the vector MEM. For a detailed discussion see Cipollini et al. (2013).

the standard GARCH in forecasting volatility, but it also preserves many of its attractive features. In particular, it preserves the ARMA structure which characterises the conventional GARCH framework simplifying the derivation of model properties. The research on this issue is currently ongoing and there are still many open issues deserving solution and investigation. First, the use of realized measures in volatility models gives rise to an errors in variables problem and, hence, to attenuation bias (Bollerslev et al., 2016). Parameter estimates should be corrected for this bias. The models that have been so far proposed, with a few exceptions (Hansen and Huang, 2016), require the choice of specific volatility measures while it is reasonable to expect that using a variable combination of different measures, according to market conditions, could be beneficial for volatility forecasting.

Moving from the daily horizon to a finer time scale, the growing popularity of high-frequency financial data, due to the technological progress in trading systems, has increased the attention of traders and researchers on the importance of intraday trading, optimal trade execution, order placement and liquidity dynamics. Consequently, financial time series analysis has focused on modelling volatility measure based on high-frequency data, as well as on intra-daily volumes, number of trades and durations. In this framework, being all positive-valued variables, the *autoregressive conditional mean* models play a dominant role in the literature. This class of models has been introduced by Engle and Russell (1998), developing the Autoregressive Conditional Duration (ACD) to model the dynamic behaviour of the time between trades, later generalized in Multiplicative Error Model (MEM) by Engle (2002) and Engle and Gallo (2006), as a general class of time series models for positive-valued random variables which are decomposed into the product of their conditional mean and a positive-valued i.i.d. error term with unit mean. Accordingly, more

attention has been paid to variables different from volatility, such as intra-daily trading volumes (see e.g. Manganelli (2005), Brownlees et al. (2011) and Hautsch et al. (2013), among others).

Most high-frequency variables share the features of being positive-valued, positively autocorrelated, strongly persistent and of following distinct intraday seasonal patterns. The presence of long-range dependence in financial variables is conventionally modelled by autoregressive fractionally integrated moving average (ARFIMA) process as in Andersen et al. (2003) or by using regression models mixing information at different frequencies as the Heterogeneous AR (HAR) model of Corsi (2009) and the MIDAS (Mixed Data Sampling) regression of Ghysels et al. (2006). On the other hand, the intra-daily seasonalities are usually estimated using Flexible Fourier Form mainly based on the work of Gallant (1981).

In order to capture the heavy dependence structure which characterise financial variables such as volatility and volume, dynamic models with slowly moving components are gaining importance in recent years. The Spline-GARCH of Engle and Rangel (2008) represents the starting point in capturing secular trends in financial volatility. Along this line, Engle et al. (2013) introduce a new class of models called GARCH-MIDAS to examine whether information contained in monthly, quarterly, or biannual financial or macroeconomic variables can improve the prediction of the returns conditional variance. Brownlees et al. (2011) propose a multi-component model with a dynamic specification to capture salient features of intra-daily volumes such as strong persistence, asymmetry and intra-daily periodicities. A distinctive feature of their approach is the inclusion of a daily trend component in the dynamics of the intra-daily volumes.

Technological progress together with the growing ascendancy of electronic

trading systems allows to store market activity on ultra high-frequency with very high precision. However, high-frequency financial transaction are characterised by issues concerning both the structure of the market and the statistical properties of data. The sequence and the structure of the data strongly depend on trading rules, trading forms and institutional settings. An important property of financial high-frequency data is the irregular spacing in time. A further major feature of transaction data is the discreteness of prices. Furthermore, high-frequency datasets are characterised by the presence of different types and sources of errors, such as simultaneous observations, decimal errors, transposition errors, isolated bad ticks and multiple bad ticks in succession, where the *tick* is the minimum price fluctuation. Therefore, since high-frequency datasets might be affected by errors and false information, the data cleaning has become an essential preliminary step to avoid misleading results in subsequent statistical analysis. To address this problem, several methods have been proposed to filter ultra high-frequency data (see e.g. Dacorogna et al. (2001), Falkenberry (2002), Brownlees and Gallo (2006) and Barndorff-Nielsen et al. (2009), among others).

1.1 Structure of the Thesis

Aim of this thesis is to propose and discuss novel model specifications for predicting volatility (Chapters 2 and 3) and volumes (Chapter 4) at horizons ranging from one day, for volatility, to a few minutes, for trading volumes. In the following, the most important contributions and findings of each chapter will be summarized in more detail.

The models we propose for forecasting daily volatility make use of intra-daily information, in the form of realized measures, and try to provide a solution to some open problems currently being debated in the

literature. Chapter 2 proposes a flexible generalization of the standard Realized GARCH model of Hansen et al. (2012) that allows to account for the heteroskedasticity of the error term featuring the realized volatility measure as well as to correct the attenuation bias effect in a dynamic and a fully data driven fashion. Since volatility is latent, proxies are needed in order to model its dynamics. The realized variance, under the assumption that the return generating process is a continuous semimartingale, is a consistent estimator for the integrated variance. However, it is well known that, because of microstructure noise, the realized variance is influenced by measurement error, where the impact of this error is strictly related to the market activity. In this context seems to be natural consider heteroskedasticity, therefore the variability of the measurement error is assumed to vary over time as a function of an estimator of the integrated quarticity of intra-daily returns. A further complication that arises from the noisy estimates of the integrated variance is the so-called errors-in-variables problem. This typically leads to the rise of what is usually known as attenuation bias, with the realized measure being less persistent than the latent integrated variance. In particular, the magnitude of the measurement error is negatively correlated with the persistence of the observed process. Therefore, correcting for the attenuation bias can potentially lead to improved volatility forecasts. In order to account for dynamic attenuation bias effects, we allow the volatility persistence to depend on the time-varying variance of the measurement noise, assigning more weight to the realized measure when it is more accurately measured. Accordingly this parametrization provides stronger persistence when the measurement error is relatively low, which is in line with the findings in Bollerslev et al. (2016). The resulting model is called the Time Varying Heteroskedastic Realized GARCH (TV-HRGARCH). The proposed approach is also extended to consider potential jump effects by adding the log-ratio between realized variance and a

realized volatility measure robust to the jumps as an explanatory variable in the measurement equation. This allows to separate the measurement error due to microstructure noise and discretization from the impact of jumps, since the jump correction is carried out only on days when jumps are most likely to occur, while resorting to the use of more efficient measures in jumps free periods. The empirical analysis on four stocks traded on the Xetra Market in the German Stock Exchange shows that the introduction of heteroskedasticity as well as of time-varying persistence provides positive effects on the empirical fit and on the out-of sample forecasting performance of the models. Similar results also apply to the class of models that accounts for jumps in the measurement equation.

Chapter 3 presents a further extension of the Realized GARCH model with the aim of improving the predictive ability of the model, by reducing the bias related to the sampling frequency of the intraday returns. It is well documented that the realized variance provide a consistent estimate of the integrated variance when prices are observed continuously and without measurement error. Consequently, the realized volatility should be computed using intra-daily returns sampled at the highest possible frequency. Unfortunately, market microstructure frictions introduce severe bias on daily volatility estimation as the sampling frequency of the intra-daily returns tend to increase. Thus, there is a trade-off between bias and variance related to the choice of the sampling frequency at which the realized variance is computed. This trade-off justifies the use of realized volatility measures based on intraday returns sampled at a frequency ranging from 5 to 30 minutes. In this field, the contribution of the chapter is to introduce a dynamic model for forecasting daily volatility exploiting information coming from intra-daily returns based on different sampling frequencies, in order to achieve the optimal trade-off between bias and efficiency. In particular, volatility dynamics are determined by a weighted

average of the considered realized measures, where the weights are time-varying and adaptively determined according to the estimated amount of noise and jumps. The proposed model is called the Mixed Frequency Adaptive Realized GARCH (MF-ARGARCH). As in Hansen and Huang (2016) the model includes separate measurement equations for each of the realized measures considered. However, a more parsimonious version of the MF-ARGARCH model is even considered by collapsing the two measurement equations into one single equation, resulting in the Single equation MF-ARGARCH (SMF-ARGARCH) model. Our results point out that mixing the 5 minutes realized variance with realized measures based on lower frequency information is very beneficial in terms of goodness of fit and of forecasting accuracy.

Chapter 4 moves the attention on the modelling and prediction of high-frequency trading volumes. High-frequency trading has become a widespread feature of equity markets in recent years and volume plays an important role in several intraday trading strategies. The volume-weighted average price (VWAP) strategy is used to minimize the execution costs of large trades in financial markets. The VWAP of a stock over a specified market period is simply computed as an average of intra-daily transaction prices weighted by the corresponding volume shares over the total daily volume. Since prices are practically unpredictable, the performance of the strategies are mainly based on accurate predictions of intra-daily trading volumes. In this field, as a novel approach for modeling and forecasting intra-daily volumes, we propose the Heterogeneous MIDAS Component Multiplicative Error Model (H-MIDAS-CMEM), in order to account for the main stylized facts characterising high-frequency trading volumes, such as clustering of the trading activity, strong persistence and intraday seasonality. Following the logic of the Component Multiplicative Error Model of Brownlees

et al. (2011), these empirical features are explained by components that move at different frequencies, since it is possible to distinguish a lower frequency component that moves around the overall series, a distinctive intra-daily dynamic non-periodic component and an intra-daily periodic component which allows for intraday periodicities. The periodic component is modelled employing a Fourier Flexible Form, while the short-run component follows a unit mean reverting GARCH-type process. The main innovation concerns the specification of the long-run component, since it is modelled by using a flexible MIDAS polynomial structure based on an additive cascade of linear filters adopting heterogeneous components which can move on multiple frequencies, in order to account for the strong persistent autocorrelation structure that is typical of the high-frequency volumes. Furthermore, this cascade structure reproduces the natural heterogeneity of the different categories of agents operating in the market at different frequencies. To capture sudden changes from states of low trading intensity to states characterised by a very high trading intensity, the proposed modelling approach has been further extended by introducing a time-varying intercept in the specification of the long-run component. To assess the relative merits of the proposed approach, the empirical analysis is carried out on three German stocks which are characterised by different liquidity levels according to the number of non trading intra-daily intervals. To control the proportion of zero observations the Zero-Augmented Generalized F distribution of Hautsch et al. (2013) is employed. We show that the H-MIDAS-CMEM model is able to capture the salient empirical features of high-frequency volumes both for liquid and illiquid stocks, outperforming several competitor in minimizing some common loss functions.

Finally, Appendix A illustrates the procedure used for the cleaning of the tick-by-tick dataset from which the time series used in the empirical

sections of Chapters 2 to 4 have been extracted. High-frequency data contain detailed information about financial market activity by considering tick-by-tick transactions. However, it is well known that high-frequency datasets are affected by bad observations due to problems arising from human errors, which can be unintentionally such as typing errors or intentionally for example producing dummy ticks for technical testing, and from computer errors caused by technical failures. Furthermore, the structure of the data strictly depends on the trading rules, trading forms and procedures employed by institution to produce and collect information. Further issues are related to the stochastic nature of financial transactions, since tick-by-tick data are irregularly spaced in time and changes in transaction price are discrete, because institutional settings allow prices to be only multiples of a tick. These data problems must be taken into account to avoid misleading results in subsequent statistical analysis, therefore cleaning high-frequency data becomes a necessary preliminary step. Our working dataset relies on transaction data of the German stock market indices DAX, TecDAX and MDAX, covering the period from January 2002 to December 2012. After highlighting the main features of high-frequency data and showing different types and sources of errors which produce bad ticks, we discuss on the most commonly used filtering approaches and perform some preliminary dataset manipulations before implementing the Brownlees and Gallo (2006) outliers detection procedure. Finally, to get regularly spaced time series, interpolation methods have been used to transform data from inhomogeneous to homogeneous time series.

Chapter 2

Time Varying Heteroskedastic Realized GARCH models for tracking measurement error bias in volatility forecasting

2.1 Aim and motivation

In the econometric literature it is widely acknowledged that the use of intra-daily information, in the form of realized volatility measures (Hansen and Lunde, 2011), can be beneficial for forecasting financial volatility on a daily scale. This is typically done by taking two different approaches.

First, dynamic models can be directly fitted to time series of realized measures. Examples of application of this approach are given by Heterogeneous AutoRegressive (HAR) (Corsi, 2009) and Multiplicative Error Models (MEM) (Engle, 2002; Engle and Gallo, 2006)[1]. One drawback of this approach is that the estimate is given by the expected level of the realized measure rather than by the conditional variance of returns. As it

[1] A more detailed discussion of MEM and HAR models will be given in Chapter 4.

will be made clear in the next section, realized measures are designed to consistently estimate the integrated variance that, under general regularity conditions, can be interpreted as an unbiased estimator of the conditional variance of returns, but will not be equal to it. Two main sources of bias can however arise as a consequence of microstructure noise and jumps. In practical applications an additional major source of discrepancy is due to the fact that realized measures are usually built not taking into account the contribution of overnight volatility which is indeed a relevant part of the actual conditional variance of interest for risk managers and other professionals.

The second approach makes use of GARCH type models driven by realized volatility measures. The main idea is to replace a noisy volatility proxy, the squared daily returns, with a more efficient one given by an appropriately chosen realized measures. Differently from the previous approach, in this case both low frequency, daily returns, and high frequency, realized measures, information is used for model building. Examples of models falling within this class are given by the HEAVY model of Shephard and Sheppard (2010) and the Realized GARCH model of Hansen et al. (2012). These two models are closely related but, nevertheless, they are characterised by some distinctive features. HEAVY models are designed for the generation of multi-step ahead forecasts which is guaranteed by the inclusion of a dynamic updating equation for the conditional expectation of the chosen realized measure. On the other hand, Realized GARCH models include a measurement equation which allows to gain, in a fully data driven fashion, deeper insight on the statistical properties of the realized measure for the empirical problem of interest.

A complication arising with both approaches is due to the fact that realized measures are noisy estimates of the underlying integrated variance,

generating a classical errors-in-variables problem. This typically leads to the rise of what is usually indicated as attenuation bias with the realized measure being less persistent than the latent integrated variance. Although it is evident that correcting for the attenuation bias can potentially lead to improved volatility forecasts, this issue has not received much attention in the literature. Recently, Bollerslev et al. (2016) have found that, in a HAR model, allowing the volatility persistence to depend on the estimated degree of measurement error allows to remarkably improve the model's predictive performance. Along the same track, Shephard and Xiu (2016) find evidence that the magnitude of the response coefficients associated with different realized volatility measures in a GARCH-X is related to the quality of the measure itself. Finally, Hansen and Huang (2016) observe that the response of the current conditional variance to past unexpected volatility shocks is negatively correlated with the accuracy of the associated realized volatility measure.

In this chapter, exploiting the flexibility of the Realized GARCH framework, we develop a novel modelling approach that allows to correct the attenuation bias effect in a natural and fully data driven way. To this purpose, we extend the standard Realized GARCH in three different directions. First, differently from Hansen et al. (2012) we allow the variability of the measurement error to be time-varying as a function of an estimator of the integrated quarticity of intra-daily returns. Second, we allow the volatility dynamics to depend on the accuracy of the realized measure. Namely, we allow the response coefficient of the lagged realized volatility to depend on a measure of its accuracy, given by the estimated variance of the volatility measurement error, such in a way that more weight is given to lagged volatilities when they are more accurately measured. Finally, we further extend the proposed modelling approach to account for the effect of jumps. This is done by introducing in the

measurement equation an additional component aimed at quantify the bias due to the effect of jumps. This allows to separate the measurement error due to microstructure noise and discretization from the impact of jumps. It can be shown that the proposed model can be represented as a time-varying parameter Realized GARCH, where the dynamics of the conditional variance are driven by a jumps-free volatility measure. A notable feature of the proposed model is that the jump correction occurs only on days in which jumps are most likely to happen, while resorting to the use of more efficient standard measures, such as realized variances and kernels, in jumps free periods.

The chapter is organized as follows. In Section 2.2 we review the basic theoretical framework behind the computation of realized volatility measure, while in Section 2.3 we discuss the Realized GARCH model of Hansen et al. (2012). Section 2.4 presents a time-varying parameter heteroskedastic Realized GARCH model that allows to account for attenuation bias effects. We first consider a jumps-free setting and later move to discuss a modification of the proposed model aimed at explicitly taking into account the impact of jumps. QML estimation of the proposed models is discussed in Section 2.5, while Sections 2.6 to 2.8 are dedicated to the empirical analysis. Section 2.6 presents the main features of the dataset used for the analysis. Section 2.7 focuses on the in-sample performance of the proposed models compared to the standard Realized GARCH model taken as a benchmark, whereas the out-of-sample forecasting performance is analysed in Section 2.8. Section 2.9 concludes.

2.2 Realized measures: a short review

In recent years, the availability of high-frequency financial market data has enabled researchers to build reliable measures of the latent daily volatility

based on the use of intra-daily returns. In the econometric and financial literature, these measures are widely known as *realized* volatility measures. In order to properly introduce the theoretical background supporting the use of realized volatility measures, we need to focus on the dynamic specification of the price process in continuous time. Formally, let the logarithmic price p_t of a financial asset, to be determined by the stochastic differential process

$$dp_t = \mu_t dt + \sigma_t dW_t + dJ_t \qquad 0 \leq t \leq T \tag{2.1}$$

where μ_t and σ_t are the drift and the instantaneous volatility processes, respectively, while W_t is a standard Brownian motion, with σ_t assumed to be independent of W_t and J_t is a finite activity jump process. Under assumption of jump absence ($dJ_t = 0$) and a frictionless market the logarithmic price p_t follows a semi-martingale process.

In this case, it can be easily shown that the Quadratic Variation (QV) of log-returns $r_t = p_t - p_{t-1}$ coincides with the Integrated Variance (IV) given by

$$IV_t = \int_{t-1}^{t} \sigma_s^2 ds. \tag{2.2}$$

In absence of jumps, microstructure noise and measurement error, Barndorff-Nielsen and Shephard (2002) show that the IV can be consistently estimated by the Realized Volatility (RV)

$$RV_t = \sum_{i=1}^{M} r_{t,i}^2 \tag{2.3}$$

where

$$r_{t,i} = p_{t-1+i\Delta} - p_{t-1+(i-1)\Delta}$$

is the i-th Δ-period intraday return and $M = 1/\Delta$. It is worth noting that, although IV and conditional variance of returns do not coincide, there is a precise relationship between these two quantities. Under standard integrability conditions (Andersen et al., 2001) it can be shown that

$$E(IV_t|\mathcal{F}_{t-1}) = var(r_t|\mathcal{F}_{t-1}),$$

where $\mathcal{F}_{t-1}$ denotes the information set available up to time $(t-1)$. In other words, the optimal forecast of the IV can be interpreted as the conditional variance of returns and, as a consequence, the difference between these two quantities is given by a zero mean error.

Barndorff-Nielsen and Shephard (2002) show that the RV consistently estimates the true latent volatility when $\Delta \longrightarrow 0$ but, in practice, due to data limitations, the following hold

$$RV_t = IV_t + \varepsilon_t \tag{2.4}$$

and

$$\varepsilon_t \sim N(0, 2\Delta IQ_t) \tag{2.5}$$

where $IQ_t = \int_{t-1}^{t} \sigma_s^4 ds$ is the Integrated Quarticity (IQ) which in turn can be consistently estimated as

$$RQ_t = \frac{M}{3} \sum_{i=1}^{M} r_{t,i}^4. \tag{2.6}$$

On the other hand, if jumps are present, the QV will differ from the IV of returns with the difference between the two quantities being given by the cumulated squared jumps. Namely, let

$$dJ_t = k_t dq_t$$

where $k_t = p_t - p_{t-}$ is the size of the jump in the logarithmic price p_t and q_t is a counting process with possibly time-varying intensity λ_t such that

$$P(dq_t = 1) = \lambda_t dt.$$

Then, under the assumptions in Andersen et al. (2007), we will have that

$$RV_t \underset{p}{\to} QV = IV + \sum_{t-1 \leq s \leq t} k^2(s).$$

Hence the RV estimator will be consistent for the QV, but not for the IV of the process. An alternative is to use jump-robust estimators, such as the *Bipower* and *Tripower Variation* (Barndorff-Nielsen and Shephard, 2004b), *minRV* or *medRV* (Andersen et al., 2012), that will be consistent for IV even in the presence of jumps. In the empirical applications carried out in this work, among the different proposals arisen in the literature, we will focus on the *medRV* estimator. The reasons for our choice are mainly of a theoretical nature. Specifically, Andersen et al. (2012) show that in the jump-free case *"the medRV estimator has better theoretical efficiency properties than the tripower variation measure and displays better finite-sample robustness to both jumps and the occurrence of "zero" returns in the sample"*. In addition, it is worth remarking that, unlike the *bipower variation* measure, for the *medRV* estimator an asymptotic limit theory in the presence of jumps is available.

The *medRV* estimator proposed by Andersen et al. (2012) is given by

$$medRV_t = \frac{\pi}{6 - 4\sqrt{3}\pi} \left(\frac{M}{M-2} \right) \sum_{i=2}^{M-1} med\left(|r_{t,i-1}|, |r_{t,i}|, |r_{t,i+1}| \right)^2. \quad (2.7)$$

Nevertheless it is worth remarking that, in the jumps free case, the above discussed jump robust estimators can be shown to be substantially less

efficient than the simple RV estimator. Namely, the *Bipower* and *Tripower Variation*, the *medRV* and *minRV* are found to be asymptotically normal with their asymptotic variance proportional (up to different scale factors) to the integrated quarticity (Andersen et al., 2012). Furthemore, in presence of jumps, this will be not consistently estimated by the RQ estimator, but some alternative jumps robust estimator will be needed. Among the several proposals presented in the literature, for the same reasons discussed for the jump-robust estimation of the IV, in the empirical section of this work we will focus on the $medRQ$ estimator proposed by Andersen et al. (2012)

$$medRQ_t = \frac{3\pi M}{9\pi + 72 - 52\sqrt{3}} \left(\frac{M}{M-2} \right) \sum_{i=2}^{M-1} med \left(|r_{t,i-1}|, |r_{t,i}|, |r_{t,i+1}| \right)^4 . \tag{2.8}$$

2.3 Realized GARCH models

The Realized GARCH (RGARCH) introduced by Hansen et al. (2012) has extended the class of GARCH models by replacing, in the volatility dynamics, the squared returns with a much more efficient proxy such as a realized volatility measure. The resulting model specification could be seen as a GARCH-X model where the realized measure is used as explanatory variable. Furthermore, the Realized GARCH "completes" the GARCH-X by adding a measurement equation to explicitly model the relationship between the realized measure and the model-based conditional variance of returns.

Namely, let $\{r_t\}$ be a time series of stock returns and $\{x_t\}$ be a time series of realized measures of volatility. The standard RGARCH model, in its linear

formulation, is then defined by the following equations

$$r_t = \mu_t + \sqrt{h_t}\, z_t \tag{2.9}$$

$$h_t = \omega + \beta\, h_{t-1} + \gamma\, x_{t-1} \tag{2.10}$$

$$x_t = \xi + \varphi\, h_t + \tau(z_t) + u_t \tag{2.11}$$

where $h_t = var(r_t|\mathcal{F}_{t-1})$ is the conditional variance of returns and $\mathcal{F}_{t-1}$ the historical information set available until time $t-1$. As the analysis focuses on conditional variance, without loss of generality and in order to simplify the exposition, in the reminder it is assumed that the conditional mean $\mu_t = E(r_t|\mathcal{F}_{t-1}) = 0$. The innovations z_t and u_t are assumed to be mutually independent with $z_t \overset{iid}{\sim} (0,1)$ and $u_t \overset{iid}{\sim} (0,\sigma_u^2)$. The three equations are denoted as, in order of presentation, the *return equation*, the *volatility equation* and the *measurement equation*, respectively.

The measurement equation is justified by the fact that any consistent estimator of the integrated variance can be written as the sum of the conditional variance plus a random innovation, where the latter is captured by $\tau(z_t) + u_t$. The function $\tau(z_t)$ can accommodate for leverage effects, because it captures the dependence between returns and future volatility. A common choice (see e.g. Hansen et al. (2012)), that has been found to be empirically satisfactory, is to use the specification

$$\tau(z_t) = \tau_1\, z_t + \tau_2(z_t^2 - 1).$$

Substituting the measurement equation into the volatility equation, it can be easily shown that the model implies an AR(1) representation for h_t

$$h_t = (\omega + \xi\gamma) + (\beta + \varphi\gamma)h_{t-1} + \gamma\, w_{t-1} \tag{2.12}$$

where $w_t = \tau(z_t) + u_t$ and $E(w_t) = 0$. The coefficient $(\beta + \varphi\gamma)$ reflects the persistence of volatility, whereas γ summarizes the impact of the past realized measure on future volatility.

The general conditions required to ensure that the volatility process h_t is stationary and the unconditional variance of r_t is finite and positive are given by

$$\begin{aligned} \omega + \xi\gamma &> 0 \\ 0 < \beta + \varphi\gamma &< 1. \end{aligned} \tag{2.13}$$

If the conditions in (2.13) are fulfilled, the unconditional variance of r_t, taking expectations of both sides in equation (2.12), can be easily shown to be equal to $(\omega+\xi\gamma)/[1-(\beta+\varphi\gamma)]$. Finally, as for standard GARCH models, the positivity of h_t $(\forall t)$ is achieved under the general condition that ω, γ and β are all positive.

As an alternative to the above discussed linear Realized GARCH model, Hansen et al. (2012) consider a log-linear specification whose volatility and measurement equations, in their simplest form, are given by

$$log(h_t) = \omega + \beta\, log(h_{t-1}) + \gamma\, log(x_{t-1}) \tag{2.14}$$

$$log(x_t) = \xi + \varphi\, log(h_t) + \tau(z_t) + u_t \tag{2.15}$$

where $z_t \overset{iid}{\sim} (0,1)$ and $u_t \overset{iid}{\sim} (0, \sigma_u^2)$. Compared to its linear counterpart, the log-linear specification has two main advantages. First, it is by far more flexible since no constraints on the parameters are required in order to ensure the positivity of the conditional variance that holds by construction. Second, the logarithmic transformation substantially reduces, but does not eliminate, the heteroskedasticity of the measurement equation error term.

Just as in the linear case, stationarity conditions for this model are easily obtained through the derivation of an analogous first order Markovian for $\log(h_t)$. For these reasons, in the remainder of the paper we will exclusively focus our attention on the log-linear specification of the Realized GARCH model.

2.4 Time Varying Coefficient Heteroskedastic Realized GARCH models with dynamic attenuation bias

In this section we propose a generalization of the basic Realized GARCH specification that allows to account for the natural heteroskedasticity of the measurement error u_t, as well as for dynamic attenuation bias. As previously discussed, in a jump-free world any consistent estimator of the IV can be written as the sum of the conditional variance plus a random innovation. Since the variance of this innovation term is function of the IQ, it seems natural to model the variance of the noise u_t in equation (2.15) as function of the RQ. Thus, we assume that the measurement noise variance is time-varying, i.e. $u_t \overset{iid}{\sim} (0, \sigma^2_{u,t})$. In order to model the time varying variance of the measurement noise we consider the specification

$$\sigma^2_{u,t} = exp\left\{\delta_0 + \delta_1 log\left(\sqrt{RQ_t}\right)\right\} \tag{2.16}$$

where the exponential formulation allows to guarantee the positivity of the estimated variance without imposing any constraint on the value of the parameters δ_0 and δ_1. We denote the resulting model as Heteroskedastic Realized GARCH (HRGARCH). It is easy to see that the homoskedastic Realized GARCH is nested within the class for $\delta_1 = 0$ that can be tested by means of a simple Wald statistic.

In order to account for dynamic attenuation effects in the volatility persistence, in the sense of Bollerslev et al. (2016), we further extend the basic HRGARCH specification allowing for time-varying persistence in the volatility equation. This is achieved by letting γ, the impact coefficient of the lagged realized measure, to depend on the time-varying variance of the measurement noise u_t. In line with Bollerslev et al. (2016), we expect the impact of past realized measures on current volatility to be down-weighted in periods in which the efficiency of the realized measure is low. The resulting model is called the Time Varying Heteroskedastic Realized GARCH (TV-HRGARCH).

Focusing on a log-linear specification, the volatility updating equation of the TV-HRGARCH is given by

$$log(h_t) = \omega + \beta\, log(h_{t-1}) + \gamma_t\, log(x_{t-1}) \tag{2.17}$$

where

$$\gamma_t = \gamma_0 + \gamma_1\, \sigma^2_{u,t-1} \tag{2.18}$$

and $\sigma^2_{u,t}$ which follows the specification in (2.16). Accordingly, as its fixed coefficients counterpart, the TV-HRGARCH can be represented in terms of a time-varying coefficients AR(1) model for $log(h_t)$

$$log(h_t) = (\omega + \xi\gamma_t) + (\beta + \varphi\gamma_t)log(h_{t-1}) + \gamma_t\, w_{t-1}. \tag{2.19}$$

So far we have focused on a simplified setting in which we rule out the possibility of jumps. We now move to considering a variant of the proposed modelling approach featuring a jumps component as an additional source of bias in the measurement equation. This is achieved

by adding the log-ratio between RV and a realized volatility measure robust to the jumps as an explanatory variable. A variety of realized volatility estimators have been developed for considering jumps along asset price time series which provide jump robustness in estimating IV. In our empirical application, as anticipated in Section 2.2, we employ the $medRV$ estimator proposed by Andersen et al. (2012).

Generally, let $C_t = x_t/x_t^J$ be the ratio between a realized measure x_t (such as the realized variance or the realized kernel) and a realized measure robust to the jumps x_t^J. In the limit this ratio will converge in probability to the ratio between QV and IV. So we interpret values of C_t higher remarkably than 1 as providing increasing evidence of jumps occurring at time t, while the discrepancy between the two measures is expected to disappear in absence of jumps, leading to values of $C_t \approx 1$. It is worth noting that, due to sampling variability, it is however possible to observe, in a limited number of cases, values of $C_t < 1$. This is compatible with the fact that the observed C_t is given by the combination of a latent signal $\bar{C}_t \geq 1$ and a measurement error to which we can ascribe observed values of C_t below the unity threshold. A simple way of avoiding observations below unity would be to truncate the distribution of C_t at this threshold, setting all the values below the truncation point equal to 1 (see e.g. Andersen et al. (2007)). This solution would however not guarantee consistent filtering of the measurement error (the truncation on the left tail is somewhat arbitrary and the right tail would be untouched) with the potential drawback of introducing an additional source of bias in the analysis. So, being aware that this implies the rise of an error in variables problem and a potential attenuation bias effect and taking into account the limited empirical incidence of values of $C_t < 1$, we decide to work with the uncensored values of C_t.

So, after adding the bias correction variable C_t as a further explanatory variable, the modified measurement equation is given by

$$log(x_t) = \xi + \varphi log(h_t) + \eta log(C_t) + \tau(z_t) + u_t^* \tag{2.20}$$

or equivalently

$$log(x_t^*) = \xi + \varphi log(h_t) + \tau(z_t) + u_t^* \tag{2.21}$$

where $log(x_t^*) = log(x_t/C_t^\eta)$.

The specification in equation (2.21) implies that when $x_t = x_t^J$ the realized measure x_t corresponds to x_t^*, since $C_t = 1$ and no bias correction will be applied. In this way, in the jumps free case, the dynamics of the predicted volatility are still driven by the standard RV estimator, or some noise-robust variant such as the *Realized Kernel* (RK), that in this situation are known to be much more efficient than jump robust estimators. On the other hand, assuming $\eta > 0$ (as systematically confirmed by our empirical results), when x_t is greater than x_t^J, in the spirit of Barndorff-Nielsen and Shephard (2004b) and Andersen et al. (2007), it follows that there is evidence of an upward bias due to the presence of jumps meaning that the realized measure x_t needs to be lowered in order to be consistent for the latent IV. This result is reached thanks to the correction variable C_t, which, in this case, takes values higher than one and consequently makes x_t^* lower than x_t.

Considering the standard Realized GARCH and the AR(1) representation for the conditional variance, it follows that

$$log(h_t) = (\omega + \xi\gamma) + (\beta + \varphi\gamma) log(h_{t-1}) + \gamma\, w_{t-1}^* \tag{2.22}$$

where

$$w_t^* = \tau(z_t) + u_t^*$$

and

$$u_t^* = log(x_t^*) - \xi - \varphi log(h_t) - \tau(z_t). \tag{2.23}$$

By substituting equation (2.23) in (2.22), the log-conditional variance can be alternatively written as

$$log(h_t) = \omega + \beta\, log(h_{t-1}) + \gamma\, log(x_{t-1}) - \gamma\eta\, log(C_{t-1}) \tag{2.24}$$

or equivalently

$$log(h_t) = \omega + \beta\, log(h_{t-1}) + \gamma\, log(x_{t-1}^*). \tag{2.25}$$

In this modified framework, it then turns out that the log-conditional variance $log(h_t)$ is driven not only by past values of the realized measure but also, with opposite sign, by past values of the associated bias. The additional parameter η allows to adjust the contribution of C_{t-1}. From a different point of view, equation (2.21) suggests that the volatility updating equation can be rewritten in a form which is similar to that of the standard RGARCH model with the substantial difference that the volatility changes are not driven by the realized measure x_{t-1}, but by its bias-corrected version x_{t-1}/C_{t-1}^{η}, where the amount of correction is determined by the estimated scaling parameter η.

This specification, of course, can be easily extended to the HRGARCH and TV-HRGARCH models with the additional modification that, in order to account for the presence of jumps, the RQ estimator in the specification of $\sigma_{u,t}^2$ must be replaced by some jump-robust estimator as it will be more extensively discussed in the empirical section. In the reminder, in

order to distinguish models incorporating the bias correction variable C_t in the measurement equation, these will be denoted by addition of the superscript *, namely: RGARCH*, HRGARCH* and TV-HRGARCH*.

2.5 Quasi Maximum Likelihood Estimation

The model parameters can be easily estimated by standard Quasi Maximum Likelihood (QML) techniques. Let Y_t indicate any additional explanatory variable eventually included in the measurement equation. Following Hansen et al. (2012), the quasi log-likelihood function, conditionally on past information $\mathcal{F}_{t-1}$ and Y_t, is given by

$$\mathcal{L}(r, x; \boldsymbol{\theta}) = \sum_{t=1}^{T} log\, f(r_t, x_t | \mathcal{F}_{t-1}, Y_t)$$

where $\boldsymbol{\theta} = (\boldsymbol{\theta}_h, \boldsymbol{\theta}_x, \boldsymbol{\theta}_\sigma)'$ with $\boldsymbol{\theta}_h$, $\boldsymbol{\theta}_x$ and $\boldsymbol{\theta}_\sigma$ respectively being the vectors of parameters appearing in the volatility equation ($\boldsymbol{\theta}_h$), in the level of the measurement equation ($\boldsymbol{\theta}_x$) and in the noise variance specification ($\boldsymbol{\theta}_\sigma$).

An attractive feature of the Realized GARCH structure is that the conditional density $f(r_t, x_t | \mathcal{F}_{t-1}, Y_t)$ can be easily decomposed as

$$f(r_t, x_t | \mathcal{F}_{t-1}) = f(r_t | \mathcal{F}_{t-1}) f(x_t | r_t; \mathcal{F}_{t-1}, Y_t).$$

Assuming a Gaussian specification for z_t and u_t, such as $z_t \overset{iid}{\sim} N(0,1)$ and $u_t \overset{iid}{\sim} N(0, \sigma_u^2)$, the quasi log-likelihood function is given by

$$\mathcal{L}(r, x; \theta) = \underbrace{-\frac{1}{2}\sum_{t=1}^{T} log(2\pi) + log(h_t) + \frac{r_t^2}{h_t}}_{\ell(r)} + \underbrace{-\frac{1}{2}\sum_{t=1}^{T} log(2\pi) + log(\sigma_u^2) + \frac{u_t^2}{\sigma_u^2}}_{\ell(x|r)}. \tag{2.26}$$

Since the standard GARCH models do not include an equation for x_t, the overall maximized log-likelihood values is not comparable to that returned from the estimation of standard GARCH-type models. Nevertheless, the partial log-likelihood value of the returns component, $\ell(r) = \sum_{t=1}^{T} log\, f(r_t|\mathcal{F}_{t-1})$, can be still meaningfully compared to the maximized log-likelihood value achieved for a standard GARCH type model.

2.6 The Data

In order to assess the empirical performance of the proposed models we have carried out an application to four stocks traded on the Xetra Market in the German Stock Exchange. This section presents the salient features of the data that have been used in the application. In particular we have considered the following assets: Allianz (ALV), which is a financial services company dealing mainly with insurance and asset management; Bayerische Motoren Werke (BMW), that is a company engaged in vehicles, motorcycle, and engine manufacturing; Deutsche Telekom (DTE), which is a telecommunications company and RWE (RWE), which is a German company that provides electric utilities, such as electricity, gas, water and heating.

The original dataset included tick-by-tick data on transactions (trades only) taking place over the period from 02/01/2002 to 27/12/2012. The raw data have been cleaned using the procedure described in Brownlees and Gallo (2006) and then converted to an equally spaced series of five-minutes log-returns. These have been aggregated on a daily basis to compute a time series of 2791 daily open-to-close log-returns, two different realized volatility measures which are the realized variance and the realized kernel and one realized volatility measure robust to the jumps, that is the *medRV*

estimator. To this purpose only continuous trading transactions taking place during the regular market opening hours 9:00 am - 5:30 pm have been considered.

Table 2.1: Summary statistics

		Min.	1Qu.	Med.	Mean	3Qu.	Max.	S.dev.	Skew.	Kurt.
r_t	ALV	-0.147	-0.010	0.000	-0.001	0.008	0.135	0.021	-0.066	8.402
	BMW	-0.135	-0.010	0.000	0.000	0.010	0.153	0.020	-0.039	7.497
	DTE	-0.117	-0.008	-0.001	-0.001	0.006	0.103	0.016	-0.027	9.621
	RWE	-0.108	-0.009	0.000	-0.001	0.008	0.097	0.016	0.065	7.415
$RV_t \times 100$	ALV	0.002	0.011	0.020	0.049	0.046	2.318	0.097	10.055	189.059
	BMW	0.003	0.015	0.026	0.043	0.048	0.969	0.058	6.380	76.847
	DTE	0.002	0.008	0.014	0.030	0.028	0.903	0.054	6.797	79.710
	RWE	0.003	0.011	0.018	0.031	0.033	1.010	0.043	7.476	122.506
$RK_t \times 100$	ALV	0.002	0.011	0.021	0.049	0.047	1.730	0.089	7.012	94.289
	BMW	0.004	0.016	0.028	0.045	0.050	0.842	0.056	5.289	50.367
	DTE	0.003	0.010	0.015	0.032	0.030	0.835	0.051	5.347	49.067
	RWE	0.003	0.012	0.019	0.033	0.035	1.009	0.045	6.703	96.421
$medRV_t \times 100$	ALV	0.002	0.010	0.019	0.045	0.043	1.606	0.083	7.241	99.614
	BMW	0.003	0.014	0.025	0.041	0.046	0.773	0.052	4.999	44.906
	DTE	0.004	0.010	0.015	0.030	0.028	0.720	0.048	5.218	44.839
	RWE	0.002	0.011	0.018	0.031	0.032	0.876	0.042	5.991	75.030
$C_t^{RV} = \frac{RV_t}{medRV_t}$	ALV	0.760	1.000	1.100	1.134	1.219	2.655	0.195	1.536	7.596
	BMW	0.738	0.996	1.093	1.126	1.213	3.131	0.191	1.867	11.612
	DTE	0.724	0.947	1.031	1.062	1.144	2.536	0.172	1.649	9.152
	RWE	0.742	1.003	1.092	1.127	1.213	2.751	0.187	1.635	8.842
$C_t^{RK} = \frac{RK_t}{medRV_t}$	ALV	0.747	0.989	1.090	1.123	1.208	2.654	0.193	1.567	7.835
	BMW	0.736	0.984	1.080	1.114	1.196	3.126	0.190	1.889	11.761
	DTE	0.718	0.936	1.019	1.048	1.131	2.489	0.169	1.665	9.254
	RWE	0.742	0.988	1.076	1.111	1.195	2.747	0.185	1.684	9.111

Summary statistics of daily log-returns r_t, daily Realized Variance RV_t* (*: $\times 100$), daily Realized Kernel RK_t* (*: $\times 100$), daily $medRV_t*$ (*: $\times 100$), bias correction variable C_t^{RV} for RV_t and bias correction variable C_t^{RK} for RK_t . Sample period: January 2002 – December 2012. Min.: Minimum; 1Qu.: First Quartile; Med.: Median; Mean; 3Qu.: Third Quartile; Max.: Maximum; S.dev.: Standard deviation; Skew.: Skewness; Kurt.: Kurtosis.

Table 2.1 reports some descriptive statistics for daily log-returns (r_t), realized variance (RV_t), realized kernel (RK_t) and $medRV_t$, as well as for the bias correction variables related to RV_t and RK_t, denoted by C_t^{RV} and

C_t^{RK}, respectively. For ease of presentation the values of the summary statistics of the realized variance and realized kernel series have been rescaled by a factor of 100.

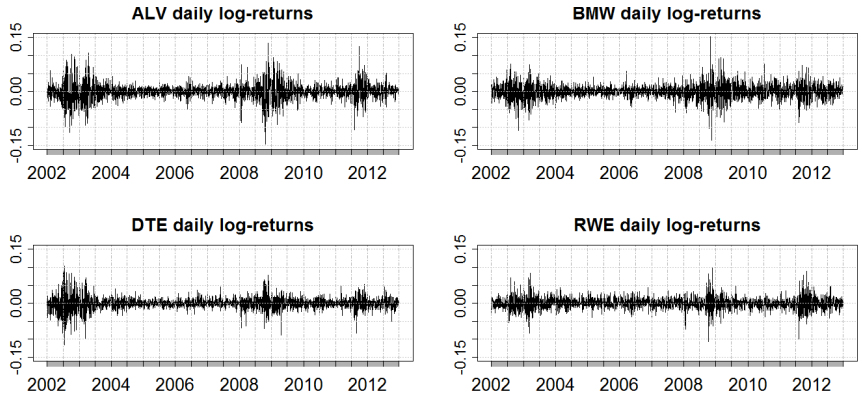
Table 2.2: C_t distribution for RV and RK

	Distribution of C_t^{RV}				Distribution of C_t^{RK}			
	ALV	BMW	DTE	RWE	ALV	BMW	DTE	RWE
0%	0.760	0.738	0.724	0.742	0.747	0.736	0.718	0.741
5%	0.894	0.893	0.845	0.895	0.887	0.883	0.836	0.883
10%	0.930	0.930	0.882	0.931	0.923	0.919	0.869	0.919
15%	0.958	0.957	0.908	0.957	0.947	0.945	0.896	0.946
20%	0.978	0.974	0.928	0.982	0.970	0.965	0.918	0.969
25%	1.000	0.996	0.947	1.003	0.989	0.984	0.936	0.988
30%	1.020	1.014	0.965	1.020	1.010	1.005	0.952	1.006
35%	1.040	1.033	0.982	1.038	1.028	1.023	0.969	1.024
40%	1.060	1.054	0.997	1.055	1.050	1.043	0.986	1.041
45%	1.083	1.074	1.015	1.074	1.072	1.061	1.003	1.061
50%	1.100	1.093	1.031	1.092	1.090	1.080	1.019	1.076
55%	1.119	1.114	1.049	1.114	1.109	1.102	1.035	1.097
60%	1.143	1.137	1.073	1.134	1.132	1.124	1.058	1.119
65%	1.168	1.159	1.093	1.158	1.156	1.146	1.079	1.141
70%	1.194	1.181	1.117	1.183	1.181	1.171	1.102	1.166
75%	1.219	1.213	1.144	1.213	1.208	1.196	1.131	1.195
80%	1.260	1.247	1.177	1.246	1.247	1.234	1.160	1.229
85%	1.306	1.296	1.216	1.295	1.293	1.282	1.199	1.281
90%	1.375	1.354	1.273	1.359	1.362	1.342	1.251	1.343
95%	1.492	1.475	1.363	1.477	1.472	1.463	1.347	1.453
100%	2.655	3.131	2.536	2.751	2.654	3.126	2.489	2.747

The daily returns range from -0.147 to 0.153, with a standard deviation typically around 0.020 and a slight skewness, which is negative for ALV, BMW and DTE, and positive for RWE. Furthermore, the high values of the kurtosis imply that the empirical distribution of the daily returns exhibits remarkably heavier tails than the normal distribution. The analysis of the rescaled daily RV_t reveals an average volatility between 0.030 and 0.049

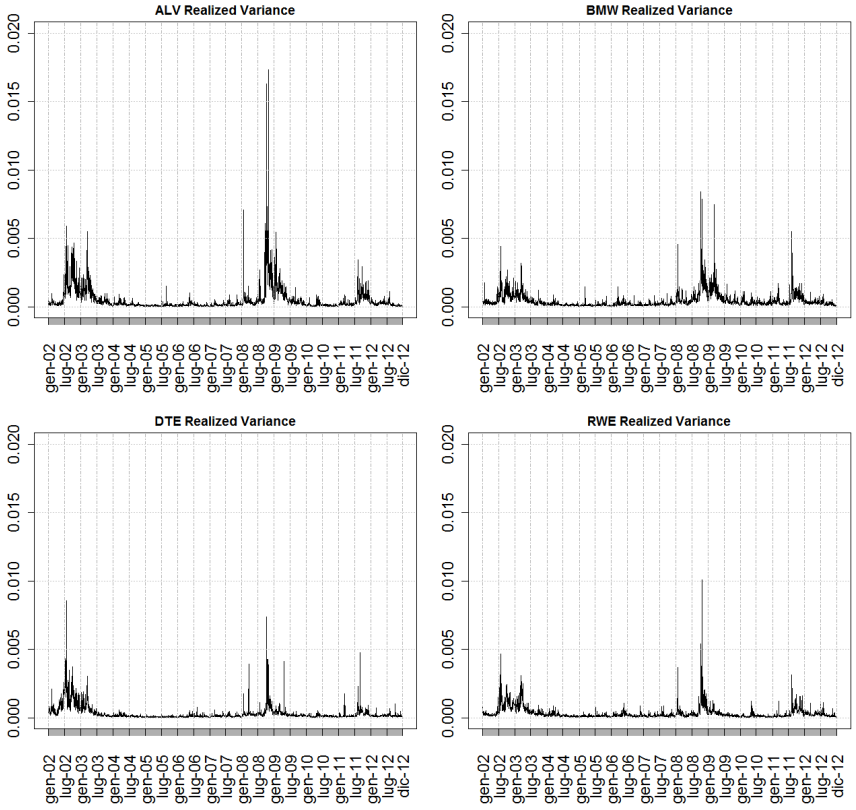
and a standard deviation ranging from 0.043 to 0.097. Finally, as expected, all the RV_t series present a very strong positive skewness. Similar considerations also hold for RK_t, even if it shows lower values of skewness and kurtosis than RV_t. The $medRV$ shows an average volatility which is in line with RV_t and RK_t, but it highlights smaller standard deviation values. Since it is a jump-robust estimator, several peaks in volatility are reduced, taking on lower maximum values than RV_t and RK_t. The bias correction variables C_t^{RV} and C_t^{RK} show an average value slightly higher than one and a positive skewness for each analysed stock. The minimum value is between 0.724 and 0.760 for C_t^{RV} and between 0.718 and 0.747 for C_t^{RK}, whereas the maximum value ranges from 2.536 to 3.131 and from 2.489 and 3.126 for C_t^{RV} and C_t^{RK}, respectively. Furthermore, looking into more detail at the distribution of C_t (Table 2.2) it is interesting to see that in three cases out of 4 only approximately the 5% of observations is below 0.90 (for DTE this values decreases to approximately 0.85) suggesting that the incidence of the measurement error in the observed C_t series is rather limited.

Figure 2.1: Time series of daily log-returns



Daily log-returns for the stocks ALV (top-left), BMW (top-right), DTE (bottom-left) and RWE (bottom-right) for the sample period 02/01/2002 – 27/12/2012.

Figure 2.2: Daily Realized Variance

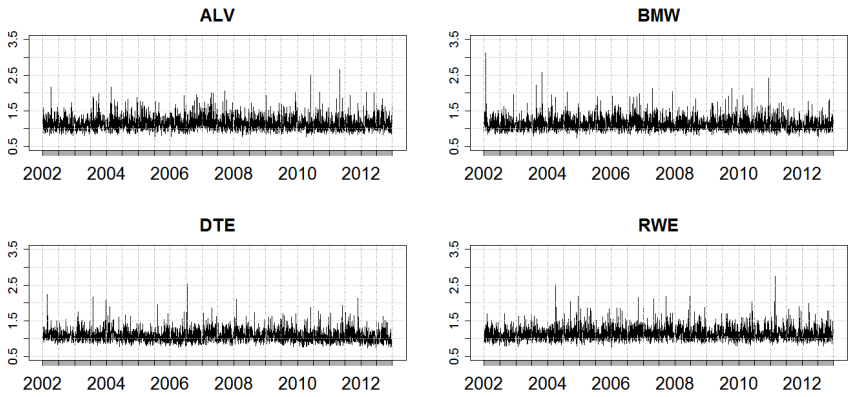


Daily Realized Variance computed using a sampling frequency of 5 minutes. ALV (top-left), BMW (top-right), DTE (bottom-left) and RWE (bottom-right). Sample period 02/01/2002 – 27/12/2012.

Figure 2.1 displays the daily returns for the four analysed stocks. A visual inspection of the plots reveals three periods of high volatility that are common to the four stocks. The first high volatility period dates back to the confidence crisis following the burst of the dot com bubble in 2002, where volatility was high through much of the latter part of 2002 and it did not go down until March of 2003. Following, the period between 2003 and 2007 was marked by a protracted unusually low volatility, interrupted by the beginning of the financial crisis in mid 2007 (mainly related to the bursting

of the United States housing bubble), peaking in September 2008 with the collapse of Lehman Brothers. The crisis in Europe then progressed from the banking system to a sovereign debt crisis, with the highest turmoil level in the late 2011, because of the critical situation of the Greek, Italian and Spanish economies and the resulting political and monetary instability of the Euro area. These are even more evident looking at Figure 2.2 reporting the time plots of the daily 5-minutes Realized Variance series. To be noted that DTE, operating in the telecommunications sector, shows the highest value of RV during the first crisis linked to the dot-com burst, while the same goes for ALV looking at the financial crisis of 2008, since it offers financial services. The most recent crisis, shows similar values across the assets with picks around 0.003 for ALV and RWE and picks close to 0.005 for BMW and DTE.

Figure 2.3: Time series of daily bias correction variable C_t^{RV}



Daily bias correction variable $C_t = RV_t/medRV_t$ for the stocks ALV (top-left), BMW (top-right), DTE (bottom-left) and RWE (bottom-right) for the sample period 02/01/2002 – 27/12/2012.

Finally, Figure 2.3 shows the bias correction variable C_t given by the ratio between RV_t and $medRV_t$. This variable fluctuates approximately around a

base level 1 with an evident positive skewness due to the the upward peaks (jumps) while downward variations due to measurement noise appear to be much less pronounced and negligible.

2.7 In-sample estimation

In this section we focus on the in-sample performance of the proposed models. The empirical analysis is carried out considering the full available sample and focusing on the log-linear specification. For ease of exposition, we separately present in sections 2.7.1 and 2.7.2 the estimation results obtained for the jump free models (RGARCH, HRGARCH and TV-HRGARCH) and for the modified models in which the impact of jumps is considered in the measurement equation (RGARCH*, HRGARCH* and TV-HRGARCH*). This choice is also due to the fact that the two categories of models are not directly comparable in terms of log-likelihood, because of the different specification of the u_t and u_t^*.

2.7.1 Estimation results for RGARCH, HRGARCH and TV-HRGARCH

In this section we report the in-sample estimation results for the HRGARCH and TV-HRGARCH models. As a benchmark for comparison we also consider the standard RGARCH model. The top panel in Table 2.3 reports parameter estimates and robust standard errors for the RGARCH, HRGARCH and TV-HRGARCH, while the second panel shows the corresponding values of the log-likelihood $\mathcal{L}(r, x)$ and partial log-likelihood $\ell(r)$, together with Bayesian Information Criterion (BIC), for the four analysed stocks.

Table 2.3: In-Sample Estimation Results for RGARCH, HRGARCH and TV-HRGARCH using 5-minutes Realized Variance

	ALV			BMW			DTE			RWE		
	RGARCH	HRGARCH	TV-HRGARCH	RGARCH	HRGARCH	TV-HRGARCH	RGARCH	HRGARCH	TV-HRGARCH	RGARCH	HRGARCH	TV-HRGARCH
ω	-0.132 $\star$ 0.257	0.011 $\star$ 0.094	1.999 0.386	-0.127 $\star$ 0.111	0.036 $\star$ 0.082	1.611 0.323	-0.067 $\star$ 0.133	0.124 $\star$ 0.095	0.878 0.444	-0.180 0.085	-0.368 0.132	0.940 0.325
γ	0.401 0.037	0.381 0.031	-	0.286 0.027	0.280 0.024	-	0.337 0.032	0.330 0.028	-	0.338 0.028	0.293 0.027	-
γ_0	-		0.490 0.051	-	-	0.410 0.041	-	-	0.415 0.058	-	-	0.389 0.037
γ_1	-		1.059 0.238	-	-	0.616 0.151	-	-	0.189 0.086	-	-	0.616 0.204
β	0.583 0.033	0.620 0.028	0.565 0.035	0.701 0.023	0.727 0.023	0.692 0.026	0.662 0.027	0.691 0.024	0.669 0.030	0.646 0.026	0.670 0.024	0.634 0.028
ξ	-0.094 $\star$ 0.610	-0.478 0.219	-0.468 0.211	0.113 $\star$ 0.353	-0.589 0.254	-0.574 $\star$ 0.375	-0.140 $\star$ 0.373	-0.952 0.246	-0.959 0.229	0.088 $\star$ 0.230	0.608 $\star$ 0.474	0.644 $\star$ 0.355
φ	0.989 0.073	0.948 0.027	0.949 0.026	1.004 0.043	0.924 0.031	0.926 0.046	0.965 0.043	0.879 0.028	0.878 0.026	0.993 0.027	1.058 0.056	1.061 0.042
τ_1	-0.069 0.008	-0.069 0.008	-0.069 0.007	-0.023 0.008	-0.032 0.008	-0.032 0.007	-0.019 0.007	-0.026 0.007	-0.027 0.007	-0.038 0.008	-0.040 0.007	-0.040 0.007
τ_2	0.108 0.006	0.111 0.006	0.111 0.006	0.082 0.006	0.086 0.005	0.085 0.005	0.091 0.005	0.091 0.005	0.091 0.005	0.082 0.005	0.083 0.005	0.082 0.005
σ_u^2	0.185 0.006	-	-	0.171 0.006	-	-	0.142 0.005	-	-	0.164 0.006	-	-
δ_0	-	-0.403 $\star$ 0.258	-0.393 0.195	-	0.322 $\star$ 0.320	0.148 $\star$ 0.237	-	1.589 0.326	1.188 0.517	-	0.283 $\star$ 0.351	-0.133 $\star$ 0.283
δ_1	-	0.163 0.032	0.166 0.024	-	0.270 0.041	0.251 0.030	-	0.436 0.039	0.389 0.061	-	0.261 0.043	0.212 0.034
$\ell(r)$	7605.335	7606.093	7608.968	7466.828	7468.052	7466.639	8234.436	8236.882	8236.071	7986.446	7988.292	7990.691
$\mathcal{L}(r,x)$	6001.795	6030.413	6061.639	5970.296	6024.041	6050.956	7020.422	7161.425	7166.077	6567.031	6614.276	6632.093
BIC	-11940.116	-11989.418	-12043.935	-11877.120	-11976.674	-12022.571	-13977.370	-14251.443	-14252.813	-13070.588	-13157.144	-13184.844

In-sample parameter estimates for the full sample period 02 January 2002 - 27 December 2012 using 5-minutes Realized Variance. $\star$: parameter not significant at 5%. $\ell(r)$: partial log-likelihood. $\mathcal{L}(r,x)$: log-likelihood. BIC: Bayesian Information Criterion. Robust standard errors are reported in small font under the parameter values.

Table 2.4: In-Sample Estimation Results for RGARCH, HRGARCH and TV-HRGARCH using 5-minutes Realized Kernel

	ALV			BMW			DTE			RWE		
	RGARCH	HRGARCH	TV-HRGARCH	RGARCH	HRGARCH	TV-HRGARCH	RGARCH	HRGARCH	TV-HRGARCH	RGARCH	HRGARCH	TV-HRGARCH
ω	0.175 $\star$ 0.096	0.016 $\star$ 0.100	1.933 0.366	-0.051 $\star$ 0.039	0.034 $\star$ 0.099	1.507 0.315	-0.077 $\star$ 0.436	0.127 $\star$ 0.101	0.728 $\star$ 0.447	-0.163 0.048	-0.354 0.113	0.817 0.344
γ	0.439 0.035	0.380 0.031	-	0.297 0.023	0.277 0.025	-	0.366 0.064	0.329 0.028	-	0.332 0.024	0.293 0.025	-
γ_0	-	-	0.485 0.047	-	-	0.397 0.040	-	-	0.397 0.060	-	-	0.381 0.040
γ_1	-	-	1.004 0.238	-	-	0.570 0.145	-	-	0.142 0.068	-	-	0.531 0.185
β	0.581 0.031	0.622 0.028	0.568 0.034	0.699 0.024	0.729 0.022	0.698 0.026	0.633 0.030	0.692 0.024	0.676 0.030	0.655 0.025	0.671 0.024	0.639 0.028
ξ	-0.775 0.187	-0.495 0.234	-0.482 0.223	-0.290 0.140	-0.583 $\star$ 0.318	-0.569 $\star$ 0.364	-0.071 $\star$ 0.238	-0.968 0.264	-0.980 0.267	-0.032 $\star$ 0.075	0.560 $\star$ 0.392	0.585 $\star$ 0.470
φ	0.910 0.023	0.947 0.029	0.948 0.027	0.957 0.018	0.926 0.039	0.928 0.044	0.972 0.030	0.879 0.030	0.877 0.030	0.981 0.017	1.054 0.046	1.056 0.055
τ_1	-0.067 0.008	-0.068 0.008	-0.068 0.008	-0.036 0.008	-0.032 0.008	-0.032 0.008	-0.022 0.007	-0.025 0.007	-0.025 0.008	-0.041 0.007	-0.039 0.007	-0.040 0.008
τ_2	0.108 0.006	0.112 0.006	0.112 0.006	0.081 0.006	0.086 0.005	0.086 0.005	0.090 0.006	0.091 0.005	0.091 0.005	0.081 0.005	0.083 0.005	0.082 0.005
σ_u^2	0.187 0.006	-	-	0.172 0.006	-	-	0.138 0.005	-	-	0.164 0.005	-	-
δ_0	-	-0.373 $\star$ 0.249	-0.376 $\star$ 0.192	-	0.312 $\star$ 0.304	0.157$\star$ 0.247	-	1.630 0.343	1.331 0.530	-	0.298 $\star$ 0.329	-0.092 $\star$ 0.299
δ_1	-	0.165 0.031	0.167 0.024	-	0.268 0.039	0.250 0.031	-	0.440 0.041	0.405 0.063	-	0.262 0.040	0.215 0.036
$\ell(r)$	7604.307	7605.255	7607.555	7467.287	7467.758	7466.359	8233.687	8236.207	8235.456	7986.291	7988.366	7990.536
$\mathcal{L}(r,x)$	5986.784	6017.651	6046.408	5958.409	6009.207	6033.109	7005.854	7151.095	7154.516	6552.303	6599.838	6613.798
BIC	-11910.095	-11963.895	-12013.475	-11853.344	-11947.007	-11986.876	-13948.235	-14230.782	-14229.690	-13041.133	-13128.268	-13148.254

In-sample parameter estimates for the full sample period 02 January 2002 - 27 December 2012 using 5-minutes Realized Kernel. $\star$: parameter not significant at 5%. $\ell(r)$: partial log-likelihood. $\mathcal{L}(r,x)$: log-likelihood. BIC: Bayesian Information Criterion. Robust standard errors are reported in small font under the parameter values.

The use of the log-linear specification ensures that the positivity of the conditional variance is always fulfilled even when the intercept ω is negative. However ω is in most cases not significant, except when the TV-HRGARCH specification is used. For this class of models the value of ω is considerably greater than the usual intercept of the standard Realized GARCH model. For each stock the parameter β is always slightly higher for HRGARCH than RGARCH and TV-HRGARCH, whereas the parameter φ takes values closer to one both for the standard RGARCH and for the RGARCH models which account for heteroskedasticity in the variance of the noise u_t. These results are in line with the findings in Hansen et al. (2012), since $\varphi \simeq 1$ suggesting that the log-transformed realized measure x_t, is roughly proportional to the conditional variance. The parameters of the leverage function $\tau(z)$ are always significant, with τ_1 which takes on negative values, while τ_2 is positive.

The empirical results also highlight that the parameter δ_1 is always positive and significant at the usual 0.05 level. This means that, as expected, RQ_t positively affects the dynamics of the variance of the error term u_t in the measurement equation. Furthermore, this implies that $\sigma^2_{u,t}$ tends to take on higher values in periods of turmoil and lower values when volatility tends to stay low. For the standard RGARCH the constant variance σ^2_u is quite stable across assets, assuming a value between 0.142 and 0.185. The coefficient γ, which summarizes the impact of the realized measure on future volatility, ranges from 0.286 to 0.401 for RGARCH, while for HRGARCH these values are slightly lower as ranging from 0.280 to 0.381 for BMW and ALV, respectively. For the TV-HRGARCH this effect is explained, in an adaptive fashion, by the time varying γ_t, which is function of the past noise variance $\sigma^2_{u,t-1}$. Since γ_1 is always positive and $log(x_t)$ (as well as $log(h_t)$) is negative, when the lagged variance of the error

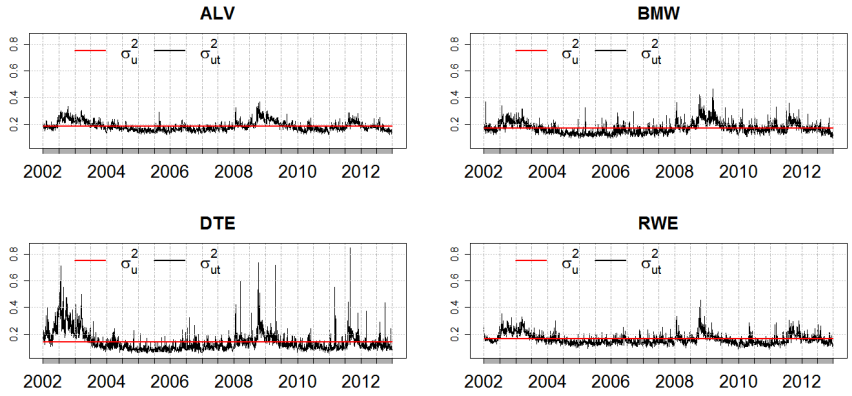
term of the realized measure $\sigma^2_{u,t-1}$ is high the impact of the lagged log-transformed realized measure $log(x_{t-1})$ on $log(h_t)$ will be negative and lower than that would have been implied by the same value of $log(x_{t-1})$ in correspondence of a lower value of $\sigma^2_{u,t-1}$. Said differently, the impact of x_{t-1} on h_t will be downscaled towards zero when $\sigma^2_{u,t-1}$ increases. Equivalently, variations in h_t ($\nabla h_t = h_t - h_{t-1}$) will be negatively correlated with the values of γ_t and $\sigma^2_{u,t-1}$. These results are in line with the recent findings of Bollerslev et al. (2016).

In terms of goodness of fit, from the second panel of Table 2.3 it clearly emerges that the TV-HRGARCH features the lowest value of the BIC for the four examined stocks. Looking at the value of $\mathcal{L}(r, x)$, the TV-HRGARCH model gives an improvement of 59.844 for ALV, of 80.660 for BMW, of 145.655 for DTE and of 65.062 for RWE, than the standard RGARCH[2]. This improvement is less pronounced if the comparison is done between RGARCH and HRGARCH. Nevertheless, it is easy to check, by means a simple likelihood ratio test, that both HRGARCH and TV-HRGARCH give rise to a significant likelihood improvement over the benchmark RGARCH. Moving to the analysis of the partial returns log-likelihood component $\ell(r)$, the differences across the analysed models are less striking, even if, in comparison with the benchmark, both HRGARCH and TV-HRGARCH show slightly greater values of $\ell(r)$.

As a robustness check, all the models have been re-estimated using the 5-minutes Realized Kernel as a volatility proxy. The results (Table 2.4) are very similar to those obtained using the 5-minutes Realized Variance.

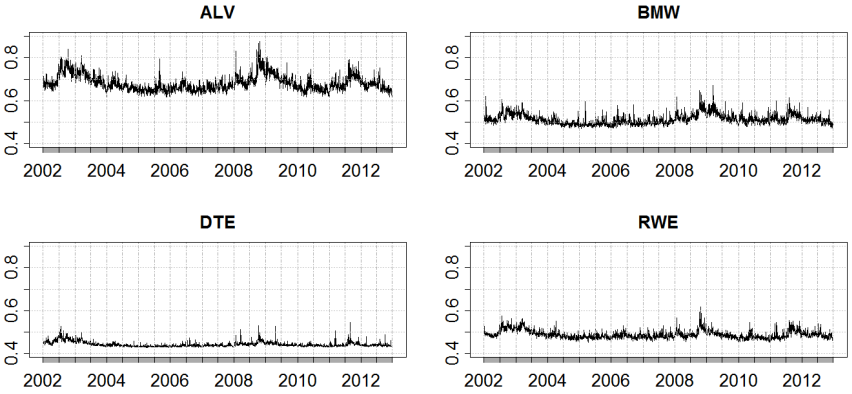
[2]Differently from Hansen et al. (2012) we obtain positive values for the log-likelihood. This is mainly due to the fact that they use percentage log-returns, which approximately fall in the range (-30, 30). It follows that the conditional variances are often above 1, returning positive log-variances that multiplied by -1 in the log-likelihood, explaining the large negative log-likelihoods that they typically get.

Figure 2.4: Constant versus time-varying variance of the noise u_t of the HRGARCH fitted using the 5 minute RV



The Figure shows the constant variance σ_u^2 (red-line) estimated with RGARCH together with the time-varying variance $\sigma_{u,t}^2$ (black-line) estimated with HRGARCH. Both models have been fitted taking the 5-minutes RV as volatility proxy. Sample period 02 January 2002 - 27 December 2012.

Figure 2.5: Time-varying coefficient γ_t given by the TV-HRGARCH model



The Figure shows the time-varying coefficient $\gamma_t = \gamma_0 + \gamma_1 \sigma_{u,t-1}^2$ for the sample period 02 January 2002 - 27 December 2012.

Figure 2.4 compares the constant variance σ_u^2 estimated by RGARCH with the time varying variance $\sigma_{u,t}^2$ given by the HRGARCH model estimated using the 5 minute RV[3]. For the four analysed stocks the trend of $\sigma_{u,t}^2$ follows the dynamics of the realized measure, being higher in period of storm and lower in periods of calm, while the constant variance σ_u^2 estimated within the RGARCH (red line in the plot) is approximately equal to the average level of the time-varying variance of the measurement noise. Figure 2.5 displays the time plot of the γ_t coefficient for the four considered stocks. It is evident that when the variance of the measurement error is high, γ_t is high leading to a less substantial increase of h_t compared to days in which, coeteris paribus, $\sigma_{u,t}^2$ is low and the realized measure provides a stronger more reliable signal. Furthermore, it is worth noting that the value of the γ_t coefficient tends to be higher than the value of the time invariant γ estimated within the RGARCH and HRGARCH models.

2.7.2 Estimation results for RGARCH*, HRGARCH* and TV-HRGARCH*

The estimation results for RGARCH*, HRGARCH* and TV-HRGARCH* are reported in Table 2.5. Since in this modified framework we are not excluding the presence of jumps, it is required to replace in equation (2.16) RQ_t with a jump-robust estimator of the IQ. In particular, for the same reasons leading use to choose $medRV_t$ as volatility robust estimator, our choice falls on $medRQ_t$.

For all the models the estimates of the ω parameter are in line with the results discussed in the previous section, being not significant in most cases. The values of the parameter β range between 0.579 and 0.720 and therefore are lower than the usual values encountered in a standard

[3]We do not report results for models using the RK as a volatility proxy since these are virtually identical to those reported here for the 5 minute RV.

Table 2.5: In-Sample Estimation Results for RGARCH*, HRGARCH* and TV-HRGARCH* using 5-minutes Realized Variance

	ALV			BMW			DTE			RWE		
	RGARCH*	HRGARCH*	TV-HRGARCH*	RGARCH*	HRGARCH*	TV-HRGARCH*	RGARCH*	HRGARCH*	TV-HRGARCH*	RGARCH*	HRGARCH*	TV-HRGARCH*
ω	-0.026 ⋆ 0.098	0.012 ⋆ 0.106	1.537 0.448	0.104 ⋆ 0.072	0.041 ⋆ 0.101	1.066 0.295	-0.016 ⋆ 0.018	0.173 ⋆ 0.093	0.523 ⋆ 0.302	-0.289 0.129	-0.375 0.125	0.294 0.129
γ	0.406 0.032	0.394 0.032	-	0.342 0.028	0.287 0.026	-	0.342 0.027	0.339 0.030	-	0.362 0.032	0.296 0.026	-
γ_0	-	-	0.449 0.053	-	-	0.349 0.037	-	-	0.378 0.046	-	-	0.339 0.031
γ_1	-	-	0.884 0.232	-	-	0.468 0.139	-	-	0.079 0.036	-	-	0.308 0.101
β	0.588 0.031	0.605 0.030	0.579 0.033	0.673 0.027	0.720 0.023	0.707 0.025	0.662 0.027	0.687 0.026	0.679 0.028	0.610 0.030	0.665 0.024	0.654 0.026
ξ	-0.324 ⋆ 0.212	-0.472 ⋆ 0.242	-0.474 0.202	-0.701 0.196	-0.574 ⋆ 0.317	-0.521 ⋆ 0.383	-0.312 0.096	-1.112 0.228	-1.133 0.231	0.247 ⋆ 0.354	0.633 ⋆ 0.439	0.662 ⋆ 0.339
φ	0.967 0.026	0.953 0.030	0.953 0.025	0.909 0.024	0.928 0.039	0.935 0.047	0.948 0.011	0.864 0.026	0.861 0.026	1.014 0.041	1.064 0.051	1.067 0.040
τ_1	-0.068 0.008	-0.069 0.007	-0.069 0.007	-0.030 0.008	-0.031 0.008	-0.031 0.008	-0.029 0.007	-0.024 0.007	-0.025 0.007	-0.041 0.008	-0.040 0.007	-0.041 0.007
τ_2	0.108 0.006	0.110 0.006	0.110 0.006	0.080 0.005	0.084 0.005	0.083 0.005	0.087 0.005	0.084 0.005	0.084 0.005	0.079 0.005	0.081 0.005	0.081 0.005
η	0.399 0.050	0.402 0.049	0.465 0.051	0.278 0.054	0.243 0.052	0.309 0.053	0.542 0.047	0.560 0.043	0.563 0.043	0.279 0.050	0.272 0.049	0.309 0.051
$\sigma^2_{u^*}$	0.180 0.006	-	-	0.169 0.006	-	-	0.132 0.005	-	-	0.160 0.005	-	-
δ_0	-	-0.551 0.255	-0.511 0.256	-	-0.115 ⋆ 0.336	-0.135 ⋆ 0.298	-	1.643 0.320	1.512 0.410	-	0.081 ⋆ 0.333	-0.151 ⋆ 0.142
δ_1	-	0.145 0.031	0.151 0.031	-	0.212 0.042	0.211 0.038	-	0.448 0.038	0.432 0.049	-	0.235 0.040	0.207 0.042
$\ell(r)$	7606.606	7606.523	7610.151	7467.651	7468.990	7467.012	8235.128	8237.538	8238.099	7989.563	7988.720	7988.765
$\mathcal{L}(r,x)$	6039.755	6062.608	6081.077	5989.603	6022.243	6033.203	7099.296	7236.314	7237.838	6586.725	6622.973	6627.373
BIC	-12008.103	-12045.874	-12074.879	-11907.799	-11965.144	-11979.131	-14127.184	-14393.286	-14388.401	-13102.043	-13166.604	-13167.470

In-sample parameter estimates for the full sample period 02 January 2002 - 27 December 2012 using 5-minutes Realized Variance. ⋆: parameter not significant at 5%. $\ell(r)$: partial log-likelihood. $\mathcal{L}(r,x)$: log-likelihood. BIC: Bayesian Information Criterion. Robust standard errors are reported in small font under the parameter values.

Table 2.6: In-Sample Estimation Results for RGARCH*, HRGARCH* and TV-HRGARCH* using 5-minutes Realized Kernel

	ALV			BMW			DTE			RWE		
	RGARCH*	HRGARCH*	TV-HRGARCH*	RGARCH*	HRGARCH*	TV-HRGARCH*	RGARCH*	HRGARCH*	TV-HRGARCH*	RGARCH*	HRGARCH*	TV-HRGARCH*
ω	-0.011 $\star$ 0.009	0.015 $\star$ 0.093	1.576 0.383	0.050 $\star$ 0.084	0.040 $\star$ 0.086	1.068 0.308	0.071 $\star$ 0.102	0.173 $\star$ 0.106	0.501 $\star$ 0.299	-0.295 0.110	-0.365 0.105	0.279 $\star$ 0.278
γ	0.401 0.030	0.393 0.032	-	0.316 0.026	0.285 0.025	-	0.361 0.030	0.339 0.030	-	0.327 0.027	0.296 0.025	-
γ_0	-	-	0.452 0.049	-	-	0.346 0.037	-	-	0.374 0.046	-	-	0.337 0.029
γ_1	-	-	0.891 0.218	-	-	0.474 0.155	-	-	0.072 $\star$ 0.038	-	-	0.293 0.095
β	0.595 0.030	0.606 0.030	0.578 0.032	0.692 0.024	0.721 0.023	0.708 0.025	0.653 0.028	0.687 0.026	0.681 0.028	0.643 0.027	0.666 0.024	0.655 0.026
ξ	-0.353 0.078	-0.481 0.209	-0.482 0.234	-0.539 0.241	-0.570 0.261	-0.517 $\star$ 0.330	-0.560 0.251	-1.112 0.267	-1.145 0.220	0.380 $\star$ 0.298	0.597 $\star$ 0.357	0.625 $\star$ 0.351
φ	0.965 0.020	0.953 0.026	0.954 0.029	0.929 0.030	0.930 0.032	0.937 0.040	0.920 0.029	0.864 0.030	0.861 0.025	1.032 0.038	1.061 0.042	1.064 0.041
τ_1	-0.067 0.008	-0.068 0.007	-0.068 0.007	-0.029 0.008	-0.031 0.008	-0.031 0.008	-0.029 0.007	-0.024 0.007	-0.024 0.007	-0.039 0.008	-0.040 0.007	-0.041 0.007
τ_2	0.108 0.006	0.111 0.006	0.110 0.006	0.081 0.006	0.084 0.005	0.084 0.005	0.087 0.006	0.084 0.005	0.084 0.005	0.080 0.005	0.081 0.005	0.080 0.005
η	0.428 0.049	0.431 0.049	0.493 0.048	0.309 0.055	0.276 0.052	0.341 0.054	0.572 0.047	0.560 0.044	0.582 0.043	0.318 0.052	0.303 0.050	0.338 0.051
$\sigma^2_{u^*}$	0.181 0.006	-	-	0.170 0.006	-	-	0.133 0.005	-	-	0.161 0.005	-	-
δ_0	-	-0.537 0.265	-0.496 0.241	-	-0.145 $\star$ 0.342	-0.157 $\star$ 0.303	-	1.643 0.319	1.548 0.385	-	0.087 $\star$ 0.352	-0.140 $\star$ 0.368
δ_1	-	0.146 0.032	0.153 0.029	-	0.208 0.043	0.207 0.038	-	0.448 0.038	0.436 0.046	-	0.235 0.042	0.208 0.044
$\ell(r)$	7606.038	7605.978	7609.572	7468.509	7468.878	7466.875	8236.628	8237.538	8237.725	7989.168	7988.760	7988.868
$\mathcal{L}(r,x)$	6031.986	6055.256	6073.888	5980.886	6010.990	6021.962	7095.484	7236.314	7232.245	6578.436	6612.497	6616.411
BIC	-11992.565	-12031.170	-12060.500	-11890.365	-11942.638	-11956.648	-14119.560	-14393.286	-14377.214	-13085.465	-13145.653	-13145.546

In-sample parameter estimates for the full sample period 02 January 2002 - 27 December 2012 using 5-minutes Realized Kernel. $\star$: parameter not significant at 5%. $\ell(r)$: partial log-likelihood. $\mathcal{L}(r,x)$: log-likelihood. BIC: Bayesian Information Criterion. Standard errors are reported in small font under the parameter values.

GARCH(1,1) model, taking on greater values for HRGARCH* than RGARCH* and TV-HRGARCH*, as in Table 2.3. The use of the realized measure x_t^* makes the coefficient φ slightly higher for the stocks ALV, BMW and RWE, but slightly lower for DTE according to the HRGARCH* and the TV-HRGARCH* models, while for the RGARCH* this parameter is lowered with the exception of RWE. The leverage function, as in Table 2.3, shows negative values for τ_1 and positive values for τ_2. The coefficient η related to the bias correction variable is always positive and statistically significant and its values ranges from 0.278 to 0.542 for RGARCH*, from 0.243 to 0.560 for HRGARCH* and from 0.309 to 0.563 for TV-HRGARCH*. Given that $0 < \eta < 1$, the impact of the rescaled realized measure $log(x_t/C_t^{\eta})$ is determined along the same lines as for $log(x_t)$. Therefore, the amount of smoothing is not arbitrarily chosen, but data driven through the estimated parameter η.

A very interesting result is that for RGARCH* the variance $\sigma_{u^*}^2$ of the measurement noise u_t^* is reduced than what found for the standard Realized GARCH model for each analysed asset, providing evidence of an improved goodness of fit in the modified measurement equation. For HRGARCH* and TV-HRGARCH* the parameter δ_1 is always statistically significant and positive, giving empirical confirmation to the intuition that the variance of the measurement error is time varying and in accordance with the asymptotic theory suggesting that this is positively related to the integrated quarticity. The impact of the past realized measure on future volatility is increased by the introduction of the bias correction variable C_t, as the coefficient γ always takes on higher values confirming the idea that accounting for jumps further reduces the attenuation bias effect on γ. Also, within the same framework, for the TV-HRGARCH* model the coefficient γ_1 is positive (even if lower than the ones showed for the TV-HRGARCH), confirming that more weight is given to the realized measure

when it is more accurately measured. Thus, this class of models provides stronger persistence when the measurement error is relatively low. Very similar results are obtained when using the 5-minutes Realized Kernel as a volatility proxy (Table 2.6). The main difference is that, in this case, for DTE the parameter γ_1, which drives the dynamics of the time-varying persistence, is statistically not significant, returning, in this way, to the HRGARCH* model which also shows the lowest BIC.

Table 2.7: In-sample partial log-likelihood comparison using 5-min RV

	ALV	BMW	DTE	RWE
RGARCH	7605.335	7466.828	8234.436	7986.446
HRGARCH	7606.093	7468.052	8236.882	7988.292
TV-HRGARCH	7608.968	7466.639	8236.071	**7990.691**
RGARCH*	7606.606	7467.651	8235.128	7989.563
HRGARCH*	7606.523	**7468.990**	8237.538	7988.720
TV-HRGARCH*	**7610.151**	7467.012	**8238.099**	7988.765

Table 2.8: In-sample partial log-likelihood comparison using 5-min RK

	ALV	BMW	DTE	RWE
RGARCH	7604.307	7467.287	8233.687	7986.291
HRGARCH	7605.255	7467.758	8236.207	7988.366
TV-HRGARCH	7607.555	7466.359	8235.456	**7990.536**
RGARCH*	7606.038	7468.509	8236.628	7989.168
HRGARCH*	7605.978	**7468.878**	8237.538	7988.760
TV-HRGARCH*	**7609.572**	7466.875	**8237.725**	7988.868

The second panel of Table 2.5 shows that, even in this framework, the model with time-varying persistence provides the lowest BIC values,

except for the stock DTE where the HRGARCH* is preferred. Under this respect, we remind that, because of the different specification of the error term, the two sets of models, including or less the bias correction variable C_t in the measurement equation, cannot be compared in terms of the overall log-likelihood, but only in terms partial log-likelihood. In this regard, the results in Table 2.7 show that the introduction of heteroskedasticity, as well as of time-varying persistence, has positive effects. However, the best results are achieved by the models that allow for bias effects in the measurement equation, namely HRGARCH* and TV-HRGARCH*, and only in the case of RWE the TV-HRGARCH prevails. Therefore, in general, with the exception of TV-HRGARCH for the RWE assets, RGARCH*, HRGARCH* and TV-HRGARCH* outperform their counterparts in terms of partial returns log-likelihood. Again, for completeness we also report in Table 2.8 estimation results obtained using the Realized Kernel as a volatility proxy that, as in the previous case, are virtually identical to those based on 5-minutes Realized Variance.

2.8 Out-of-sample Analysis

In this section the out-of-sample predictive ability of the models estimated in sections 2.7.1 and 2.7.2 has been assessed by means of a rolling window forecasting exercise using a window of 1500 days. The out-of sample period goes from 02 January 2008 to the end of the sample and includes 1270 daily observations covering the credit crisis and the turbulence period running from November 2011 to the beginning of 2012.

As in Hansen et al. (2012), the predictive log-likelihood has been used as a first criterion for assessing the predictive accuracy of the fitted models.

The one-day ahead predictive log-likelihood is given by

$$\hat{\ell}(r,x)_{t+1} = -\frac{1}{2}\left[log(2\pi) + log(\hat{h}_{t+1}) + \frac{r^2_{t+1}}{\hat{h}_{t+1}}\right] - \frac{1}{2}\left[log(2\pi) + log(\hat{\sigma}^2_u) + \frac{u^2_{t+1}}{\hat{\sigma}^2_u}\right]. \tag{2.27}$$

Therefore, the forecast $\hat{h}_{t+1}$ is plugged into the one-step-ahead log-density, getting the one-step-ahead predictive density estimate. Consequently, the aggregated predictive log-likelihood is computed by summing the density estimates for each day in the forecast period.

Table 2.9: Predictive log-likelihood and predictive partial log-likelihood using 5-minutes RV.

	ALV	BMW	DTE	RWE
RGARCH	2568.420 3334.337	2430.166 3131.385	3111.784 3698.522	2887.809 3562.323
HRGARCH	2583.630 3334.156	2451.359 3132.354	3189.043 3698.924	2925.647 3561.528
TV-HRGARCH	**2602.690** 3336.313	**2466.983** 3130.976	**3192.937** 3699.402	**2932.639** 3565.076
RGARCH*	2583.520 3334.529	2439.665 3131.750	3150.693 3698.896	2907.078 3562.610
HRGARCH*	2599.042 3334.508	2457.695 3132.582	3227.635 3699.068	2938.611 3561.881
TV-HRGARCH*	**2607.499** 3335.925	**2468.806** 3130.976	**3228.019** 3699.246	**2940.683** 3564.577

The table reports the predictive log-likelihood and the predictive partial log-likelihood in smaller font underneath. In **bold** the preferred model according to predictive log-likelihood. In blue the best model in terms of predictive partial log-likelihood. Out-of-sample period 02 January 2008 – 27 December 2012.

Table 2.9 shows the values of the predictive log-likelihood and of the predictive partial log-likelihood, corresponding to the different models considered in our analysis. In terms of predictive log-likelihood the

out-of-sample performance of the models is to be evaluated separately for RGARCH, HRGARCH and TV-HRGARCH and for RGARCH*, HRGARCH* and TV-HRGARCH*, because of the different specification of the error term in the measurement equation. Both in the first and in the second class of models the ones that allow for heteroskedasticity and time-varying persistence are the preferred models as they feature the greatest value of the predictive log-likelihood. This result also holds for the predictive partial log-likelihood, except for the stock BMW where the HRGARCH* is the favourite model. Note that the predictive partial log-likelihood, as its in-sample counterpart, allows to compare the predictive ability of the whole set of models considered.

Table 2.10: Average values of QLIKE loss using 5-min RV as volatility proxy (top) and MCS p-values (bottom). For each stock: **bold**: minimum loss; red: model $\in$ 75% MCS; blue: model $\in$ 90% MCS

QLIKE				
	ALV	BMW	DTE	RWE
RGARCH	-6.964	-6.666	-7.342	-7.254
HRGARCH	-6.963	-6.665	-7.344	-7.253
TV-HRGARCH	**-6.967**	**-6.669**	**-7.346**	**-7.258**
RGARCH*	-6.964	-6.667	-7.343	-7.254
HRGARCH*	-6.964	-6.666	-7.344	-7.254
TV-HRGARCH*	-6.966	**-6.669**	-7.345	-7.257
MCS p-values				
	ALV	BMW	DTE	RWE
RGARCH	0.086	0.080	0.070	0.022
HRGARCH	0.086	0.044	0.186	0.022
TV-HRGARCH	1.000	1.000	1.000	1.000
RGARCH*	0.086	0.393	0.098	0.026
HRGARCH*	0.086	0.072	0.186	0.022
TV-HRGARCH*	0.209	0.593	0.728	0.103

As a further criterion for assessing and comparing the forecasting accuracy of the fitted models, we have considered the QLIKE loss function. Our choice is motivated by the consideration that, compared to other robust alternatives, such as the Mean Squared Error (MSE), this loss function has revealed to be more powerful in rejecting poorly performing predictors (see Liu et al. (2015)). The QLIKE loss has been computed according to the formula

$$QLIKE = \frac{1}{T}\sum_{t=1}^{T} log(\hat{h}_t) + \frac{RV_t}{\hat{h}_t} \tag{2.28}$$

where the 5-minutes realized variance has been chosen as volatility proxy. The significance of differences in the QLIKE values across different models is tested by means of the Model Confidence Set (MCS) approach of Hansen et al. (2011).

Looking at average values of the QLIKE loss function for each considered model, reported in the top panel of Table 2.10, it turns out that for all the assets the lowest value is obtained for the TV-HRGARCH model (together with the TV-HRGARCH* for the stock BMW). The bottom panel of the table displays the p-values given by the MCS employing the test statistic T_{max}, discussed in Hansen et al. (2011). In order to estimate the optimal block-bootstrap length we use the method described in Patton et al. (2009), with results based on 3000 bootstrap resamples.

Assessing the significance of differences across different models, it turns out that the TV-HRGARCH is the only model always entering into the MCS at both confidence levels considered (0.75 and 0.90) and for ALV and RWE no other model is entering in the MCS at the 0.75 level. The standard RGARCH never comes into the *set of superior models* and the HRGARCH falls into the MCS only for the asset DTE at a confidence level of 0.90, but it is excluded as we move to the more restrictive 0.75 level.

Within the class of models which allow for upward and downward bias effects in the measurement equation, the RGARCH* is included in the set of superior models at the level of 0.75 for the stock BMW, whereas the HRGARCH* comes into the 90% MCS only for DTE. On the other hand, the TV-HRGARCH* shows better predictive ability, since it comes into the set of superior models at a 0.90 confidence level for ALV and RWE and at a 0.75 level for BMW and DTE. Overall, it is interesting to see, for both predictive partial likelihood and QLIKE loss, that the models including jump-correction register the best performance in the case of the BMW stock which is the asset characterised by the most extreme jumps (see Table 2.1).

Table 2.11: Predictive log-likelihood and predictive partial log-likelihood using 5-minutes RK.

	ALV	BMW	DTE	RWE
RGARCH	2564.496 3333.812	2426.601 3130.976	3108.492 3698.528	2882.355 3562.848
HRGARCH	2580.540 3333.797	2448.082 3132.242	3186.439 3698.978	2919.863 3562.009
TV-HRGARCH	**2598.487** 3335.651	**2462.938** 3130.853	**3189.973** 3699.536	**2925.268** 3565.264
RGARCH*	2581.083 3334.374	2438.195 3131.644	3148.365 3699.064	2904.114 3562.773
HRGARCH*	2597.039 3334.295	2455.614 3132.529	3225.735 3699.092	2935.192 3562.243
TV-HRGARCH*	**2604.231** 3335.374	**2465.788** 3130.951	**3225.906** 3699.247	**2936.449** 3564.387

The table reports the predictive log-likelihood and the predictive partial log-likelihood in smaller font underneath using five-minutes RK as realized measure. In **bold** the preferred model according to predictive log-likelihood. In blue the best model in terms of predictive partial log-likelihood. Out-of-sample period 02 January 2008 – 27 December 2012.

As a robustness check we have repeated the out-of-sample forecast evaluation using the 5-minutes Realized Kernel. The results, reported in Table 2.11 and in Table 2.12 are practically identical to those obtained using the 5-minute realized variance, confirming the robustness of our findings.

Table 2.12: Average values of QLIKE loss using 5-min RK as volatility proxy (top) and MCS p-values (bottom). For each stock: **bold**: minimum loss; red: model $\in$ 75% MCS; blue: model $\in$ 90% MCS

QLIKE				
	ALV	BMW	DTE	RWE
RGARCH	-6.972	-6.675	-7.359	-7.272
HRGARCH	-6.972	-6.674	-7.361	-7.270
TV-HRGARCH	**-6.975**	**-6.678**	**-7.363**	**-7.274**
RGARCH*	-6.973	-6.676	-7.360	-7.271
HRGARCH*	-6.973	-6.675	-7.362	-7.271
TV-HRGARCH*	-6.974	**-6.678**	**-7.363**	**-7.274**
MCS p-values				
	ALV	BMW	DTE	RWE
RGARCH	0.096	0.048	0.052	0.036
HRGARCH	0.096	0.029	0.196	0.020
TV-HRGARCH	1.000	1.000	1.000	1.000
RGARCH*	0.096	0.262	0.077	0.036
HRGARCH*	0.096	0.078	0.196	0.030
TV-HRGARCH*	0.123	0.438	0.734	0.126

2.9 Conclusion

The empirical results show that the introduction of heteroskedasticity in modelling volatility has a remarkable effect on the estimated parameters

and on the out-of sample forecasting performance of the model. In particular, the model that allows for heteroskedasticity and time-varying persistence, shows the lowest values of the QLIKE loss function, entering in the MCS always with the highest p-value at the considered confidence levels. Furthermore, this specification maximize the predictive log-likelihood and the predictive partial log-likelihood values.

The class of models that accounts for jumps in the measurement equation has the advantage to remarkably improve the partial log-likelihood related to the returns in fitting models and also in terms of predictive ability they show an improvement over the standard Realized GARCH approach.

Chapter 3

Mixed Frequency Realized GARCH models

3.1 Aim and motivation

High frequency financial data are becoming increasingly available and this has inspired researchers in developing new and more complex econometric models making use of this information. The introduction of ex-post volatility estimators to measure the quadratic variation of asset prices is a prominent example in this direction. Furthermore, it is widely accepted that incorporating intra-daily data in dynamic models for predicting volatility allows to generate more accurate forecasts. The estimation of volatility based on high-frequency data is not, however, devoid of problems. These are mainly related to inhomogeneity of the intraday returns series across the trading day, diurnal patterns and other market microstructure frictions which largely influence the modelling of realized volatility measures.

In absence of microstructure noise and measurement error Andersen et al. (2003) showed, using a seminal result in semimartingale process theory, that the Realized Variance (RV), using all data available, yields a

consistent estimate of volatility. This finding suggests that the RV should be computed by using tick-by-tick data or intra-daily returns sampled at the highest possible frequency. However, because of a host of practical market microstructure frictions, the RV suffers from a bias problem as the sampling frequency of the intraday returns tends to increase. Thus, as discussed in Aït-Sahalia et al. (2005), Zhang et al. (2005), Hansen and Lunde (2006), Bandi and Russell (2006) and Bandi and Russell (2008), among others, there is a trade-off between bias and variance related to the choice of the working sampling frequency. This trade-off is the reason why realized volatility measures are often based on intraday returns sampled at a frequency ranging from 5 to 30 minutes.

A graphical tool that helps to identify the optimal compromise between bias and variance, is the *volatility signature plot*, which shows the sample average of the RV over a long time span as a function of the sampling frequency. In this field, several empirical applications on the use of realized volatility have pointed out that the sampling frequencies tend to stabilize within the 5 – 30 minutes range providing empirical justification to the common practice of using volatility estimates based on a 5 minutes grid (Liu et al., 2015). Nevertheless, a 5 minutes aggregation interval is in most settings sufficient to approximately cancel out the effects of microstructure biases, but it is usually not wide enough to filter out the impact of jumps. Effective solutions require the use of a jump-robust estimator, such as the *medRV* estimator (used in Chapter 2) or, alternatively, a lower aggregation frequency for computing realized variance estimates. Both solutions require a substantial efficiency loss in days where jumps are negligible or not present at all.

The above discussion makes clear and evident the duality between variability, maximized at high sampling frequencies, and bias due

microstructure noise and jumps, vanishing at low frequencies. In empirical analysis researchers are usually called to identify the sampling frequency that allows to attain the optimal bias-variance trade-off for the data at hand. An alternative approach would be to use an adaptive frequency varying according to market conditions. Ideally, one would tend to use the highest possible frequency in noise and jumps free periods, while reverting to a lower frequency when these are more likely to take place.

This chapter enhances the theoretical discussion and the empirical analysis performed in Chapter 2, proposing a dynamic model for forecasting daily volatility based on the optimal combination, in an adaptive fashion, of realized variances computed at different frequencies in order to achieve the optimal trade-off between bias and efficiency. Namely, in our approach future volatility forecasts are driven by the variation of a weighted average of the involved realized measures, where the weight is time-varying on the daily scale depending on the estimated amount of noise and jumps. The weight dynamics are driven by an appropriately chosen state variable whose level is expected to be related to the impact of microstructure noise and jumps on the higher frequency measure.

Again, as in Chapter 2, the proposed model is developed building on the Realized GARCH framework of Hansen et al. (2012). Although in this work we focus on models including two realized variances based on different discretization grids, the proposed approach could be easily generalized to incorporate a higher number of frequencies. The proposed model is called the Mixed Frequency Adaptive Realized GARCH (abbreviated MF-ARGARCH). The model includes separate measurement equations for each of the realized measures considered. For the sake of parsimony, we also develop a variant of the MF-ARGARCH characterised by a single measurement equation, called the Single equation

MF-ARGARCH (abbreviated SMF-ARGARCH) model, in which the dependent variable is a time varying weighted average of the raw realized measures.

The chapter is structured as follows. Section 3.2 introduces the novel MF-ARGARCH modelling approach while details about QML estimation of model parameters are provided in Section 3.3. Sections 3.4 to 3.6 are dedicated to the empirical applications. Section 3.4 describes the data used for the analysis, while the in-sample estimation and out-of-sample forecasting exercise are illustrated in Section 3.5 and 3.6, respectively. The results of the out-of-sample forecasting comparison provide evidence that mixed frequency models allow to improve over the predictive performance of the standard single frequency Realized GARCH model. In particular, the best results are obtained when the 5 minutes realized variance is combined with a lower frequency estimator with a sampling interval ranging from 10 to 30 minutes. Section 3.7 concludes.

3.2 Mixed Frequency Realized GARCH models

This section presents an extension of the standard Realized GARCH model that allows to exploit information coming from realized volatility measures based on different sampling frequencies in order to achieve the optimal trade-off between bias and variability and thus to improve the predictive ability of the model.

The proposed model evolves over the standard RGARCH specification under two different aspects. First, volatility dynamics are given by a weighted average of the raw measures considered, where the weights are time-varying and adaptively determined. Second, as in Hansen and Huang (2016), the model includes a different measurement equation for

each of the realized measures considered. Due to its inherent features, the resulting model specification is called the Mixed Frequency Adaptive Realized GARCH (MF-ARGARCH) model.

In particular, the MF-ARGARCH is defined by the following equations

$$log(h_t) = \omega + \beta\, log(h_{t-1}) + \gamma\, log(\tilde{x}_{t-1}) \tag{3.1}$$

$$log(\tilde{x}_t) = \alpha_t\, log\left(x_t^{(H)}\right) + (1-\alpha_t)\, log\left(x_t^{(L)}\right) \tag{3.2}$$

$$log\left(x_t^{(j)}\right) = \xi_j + \varphi_j\, log(h_t) + \tau_j(z_t) + \tilde{u}_{j,t} \qquad j = H, L \tag{3.3}$$

where $x_t^{(H)}$ and $x_t^{(L)}$ denote realized measures computed at a high and a low frequency, respectively.

Therefore, $\log(\tilde{x}_t)$ is a weighted average of two (log-transformed) realized measures based on different sampling frequencies, where the weight α_t is of the form

$$\alpha_t = \alpha_0 + \alpha_1 R_{t-1} \quad \text{and} \quad R_t = \sqrt{\frac{RQ_t^{(L)}}{RQ_t^{(H)}}}. \tag{3.4}$$

So, α_t is given by a linear function of the ratio between the Realized Quarticity (RQ) measures computed at time $(t-1)$ over the two chosen frequencies, with

$$RQ_t^{(j)} = \frac{M}{3}\sum_{i=1}^{M} r_{t,i}^4, \qquad j = H, L$$

where $r_{t,i}$ are the intra-daily returns and M the sampling frequency.

After some simple algebra, by replacing the measurement equation in the GARCH equation, the log-conditional variance can be specified as

$$log(h_t) = \bar{\mu} + \bar{\pi} log(h_{t-1}) + \bar{w}_{t-1} \tag{3.5}$$

where

$$\begin{aligned}
\bar{\mu} &= \omega + \gamma(\alpha_t \, \xi_H + (1 - \alpha_t) \, \xi_L) \\
\bar{\pi} &= \beta + \gamma(\alpha_t \, \varphi_H + (1 - \alpha_t) \, \varphi_L) \\
\bar{w}_t &= \gamma(\alpha_t \, \tilde{w}_{H,t} + (1 - \alpha_t) \, \tilde{w}_{L,t})
\end{aligned}$$

and

$$\begin{aligned}
\tilde{w}_{j,t} &= \tilde{u}_{j,t} + \tau_j(z_t), \qquad j = H, L \\
\tilde{u}_{j,t} &= log\left(\tilde{x}_t^{(j)}\right) - \xi_j - \varphi_j \, log(h_t) - \tau_j(z_t), \qquad j = H, L
\end{aligned}$$

where α_t is of the form in (3.4).

The ratio R_t plays the role of a state variable whose level is related to the level of bias due to noise and jumps. Although, in principle other choices could have been considered, we decided to use the ratio of RQs for their ability to amplify the impact of these biases.

As discussed in Barndorff-Nielsen and Shephard (2004a) and Bandi and Russell (2008), among others, the RQ highlights the sensitivity of fourth moment returns to outliers and market microstructure noise. They found that these effects tend to be reduced as the returns sampling frequency shrinks. Being the ratio between RQ computed at different frequencies, R_t has the function of correcting for the upward and downward bias eventually affecting the high frequency realized measures. As the

frequencies in the numerator and in the denominator of R_t decrease, the expected ratio tends to unity. Otherwise, its level can substantially deviate from this value due to noise and outliers. Overall, in the ideal situation in which the two frequencies are well separated, since $RQ_t^{(H)}$ will be more sensitive to jumps and microstructure noise than $RQ_t^{(L)}$, we expect that values of $R_t < 1$ will be prevalently due to both jumps and microstructure noise, while, values of $R_t > 1$ will be more likely ascribed to microstructure noise. In both cases, the observed value of R_t will be also affected by an observation error due to sampling variability. On the other hand, when the numerator and denominator frequencies are very close, we expect that variations in the observed R_t will be spurious being dominated by sampling variability. In this case, in the limit situation in which $R_t = 1, \forall t$, an identification issue arises in the estimation of α_0 and α_1.

When α_1 is not statistically significant the bias correction is not applied and the log-transformed realized measure $log(\tilde{x}_t)$ is given by the constant weights convex combination

$$log(\tilde{x}_t) = \alpha_0 \, log\left(x_t^{(H)}\right) + (1 - \alpha_0) \, log\left(x_t^{(L)}\right). \tag{3.6}$$

Therefore, α_0 and $(1-\alpha_0)$ determine, respectively, the importance assigned to $x_t^{(H)}$ and $x_t^{(L)}$ for the overall realized measure $\tilde{x}_t$. For $\alpha_1 = 0$ and α_0 equal to either 0 or 1 the standard Realized GARCH model is obtained as a special case.

A more parsimonious version of the MF-ARGARCH model can be obtained by collapsing the two measurement equations into one single equation in which the dependent variable is given by the weighted average $log(\tilde{x}_t)$. An appealing feature of this solution is that it allows to directly

obtain insight on the statistical properties of the model based on the realized measure $\tilde{x}_t$, through the parameters of the dedicated measurement equation. The resulting specification is called the Single equation MF-ARGARCH (abbreviated SMF-ARGARCH) model and it is defined by the following equations

$$log(h_t) = \omega + \beta\, log(h_{t-1}) + \gamma\, log(\tilde{x}_{t-1}) \tag{3.7}$$

$$log(\tilde{x}_t) = \alpha_t\, log\left(x_t^{(H)}\right) + (1-\alpha_t)\, log\left(x_t^{(L)}\right) \tag{3.8}$$

$$log(\tilde{x}_t) = \xi + \varphi\, log(h_t) + \tau(z_t) + \tilde{u}_t. \tag{3.9}$$

Similarly to the standard RGARCH model, the SMF-ARGARCH can be rewritten as a Markovian process by replacing the measurement equation in the volatility equation

$$log(h_t) = (\omega + \xi\gamma) + (\beta + \varphi\gamma) log(h_{t-1}) + \gamma\tilde{w}_{t-1} \tag{3.10}$$

where

$$\tilde{w}_t = \tau(z_t) + \tilde{u}_t$$
$$\tilde{u}_t = log(\tilde{x}_t) - \xi - \varphi\, log(h_t) - \tau(z_t)$$

and $log(\tilde{x}_t)$ which follows the specification in (3.8).

As previously discussed, the main appeal of both the MF-ARGARCH and the Single equation MF-ARGARCH is that these models are able to incorporate information from different realized volatility measures based on intra-daily returns observed at different frequencies, allowing to reach a time-varying optimal bias-variance trade-off between the two measures. This feature has two main implications. First, in an ex ante perspective,

we expect the Mixed Frequency models to be able to generate improved volatility forecasts since, at each time point, the forecast is based on an adaptively optimized volatility measure. Second, ex post, the estimated model can be used to generate optimized volatility measures to be used for more accurate volatility estimation. In this work we mainly focus on the first point, while leaving the second one for future research.

3.3 Quasi-Maximum Likelihood estimation procedure

In this section we discuss the quasi-maximum likelihood estimation of the Mixed Frequency Realized GARCH models presented in the previous section. For the sake of simplicity, we start considering the SMF-ARGARCH model. In this case the likelihood derivation closely follows that of the standard Realized GARCH model.

The log-likelihood function of the SMF-ARGARCH model has the following structure

$$\mathcal{L}(r, \tilde{x}; \boldsymbol{\theta}) = \sum_{t=1}^{T} log\, f(r_t, \tilde{x}_t | \mathcal{F}_{t-1})$$

where $\boldsymbol{\theta} = (\boldsymbol{\theta}_h, \boldsymbol{\theta}_{\tilde{x}})'$ with $\boldsymbol{\theta}_h$ and $\boldsymbol{\theta}_{\tilde{x}}$ the vectors of parameters characterising the volatility equation ($\boldsymbol{\theta}_h$) and the measurement equation ($\boldsymbol{\theta}_{\tilde{x}}$), respectively.

By standard probability theory results, this can be rewritten by factorizing the joint conditional density $f(r_t, \tilde{x}_t | \mathcal{F}_{t-1})$ as

$$f(r_t, \tilde{x}_t | \mathcal{F}_{t-1}) = f(r_t | \mathcal{F}_{t-1}) f(\tilde{x}_t | r_t; \mathcal{F}_{t-1}).$$

Finally, the quasi log-likelihood function of the SMF-ARGARCH model, where a Gaussian specification is assumed for z_t and $\tilde{u}_t$, such that $z_t \overset{iid}{\sim} N(0,1)$ and $u_t \overset{iid}{\sim} N(0,\sigma_{\tilde{u}}^2)$, is given by

$$\mathcal{L}(r,\tilde{x};\boldsymbol{\theta}) = \underbrace{-\frac{1}{2}\sum_{t=1}^{T} log(2\pi) + log(h_t) + \frac{r_t^2}{h_t}}_{\ell(r)} + \underbrace{-\frac{1}{2}\sum_{t=1}^{T} log(2\pi) + log(\sigma_{\tilde{u}}^2) + \frac{\tilde{u}_t^2}{\sigma_{\tilde{u}}^2}}_{\ell(\tilde{x}|r)}. \tag{3.11}$$

For the double measurement equation model, the MF-ARGARCH, some modifications arise. Due to the presence of two measurement equations and, hence, two measurement errors, in this case it is indeed necessary to consider an additional term in the conditional density $f(r_t, x_t^{(H)}, x_t^{(L)}|\mathcal{F}_{t-1})$, namely

$$f(r_t, x_t^{(H)}, x_t^{(L)}|\mathcal{F}_{t-1}) = f(r_t|\mathcal{F}_{t-1}) f(x_t^{(H)}|r_t;\mathcal{F}_{t-1}) f(x_t^{(L)}|r_t;\mathcal{F}_{t-1}).$$

As a consequence, the quasi log-likelihood function is given by

$$\begin{aligned}\mathcal{L}(r,x^{(H)},x^{(L)};\boldsymbol{\theta}) = &\underbrace{-\frac{1}{2}\sum_{t=1}^{T} log(2\pi) + log(h_t) + \frac{r_t^2}{h_t}}_{\ell(r)} + \\ &\underbrace{-\frac{1}{2}\sum_{t=1}^{T} log(2\pi) + log\left(\sigma_{\tilde{u}_{H,t}}^2\right) + \frac{\tilde{u}_{H,t}^2}{\sigma_{\tilde{u}_{H,t}}^2}}_{\ell(r|x^{(H)})} + \\ &\underbrace{-\frac{1}{2}\sum_{t=1}^{T} log(2\pi) + log\left(\sigma_{\tilde{u}_{L,t}}^2\right) + \frac{\tilde{u}_{L,t}^2}{\sigma_{\tilde{u}_{L,t}}^2}}_{\ell(r|x^{(L)})}.\end{aligned} \tag{3.12}$$

It is worth noting that the different specification of the error terms for the RGARCH, MF-ARGARCH and SMF-ARGARCH, makes the models not

directly comparable on the basis of the overall maximized log-likelihood values, but they are comparable only by the partial log-likelihood values of the returns component $\ell(r)$. Under this respect, the factorizations in (3.11) and (3.12) are particularly useful in what they can be used to compare the simple GARCH and RGARCH models to the more sophisticated Mixed Frequency Realized GARCH models by the partial log-likelihood value of the returns component, $\ell(r) = \sum_{t=1}^{T} log\, f(r_t|\mathcal{F}_{t-1})$.

3.4 Data description

In order to assess the relative merits and drawbacks of the proposed modelling approach in comparison with the standard RGARCH model we perform an empirical application on a set of 4 different German stocks. Namely, the variables we consider in our application are daily open-to-close returns and daily realized variances and quarticities based on 30 seconds, 1, 5, 10, 15 and 30 minutes frequencies, while the sample period spans a total of 2791 trading days ranging from 02/01/2002 to 27/12/2012. Since the analysis focuses on the same stocks used for the empirical application performed in Chapter 2, Allianz (ALV), Bayerische Motoren Werke (BMW), Deutsche Telekom (DTE) and RWE (RWE), in order to avoid tedious repetitions, we refer the reader to consult *The Data* section (2.6) of the previous chapter to get more information on data characteristics.

In this section we focus on the features of the ratio $R_t = \sqrt{RQ_t^{(L)}/RQ_t^{(H)}}$. This has been computed matching the usual 5 minutes frequency, kept fixed, with other frequencies ranging from 30 seconds to 30 minutes. In the following, we refer to $m = minutes$ and $s = seconds$ to denote the sampling frequency of the realized volatility measures and of the realized quarticity measures used in our analysis.

Table 3.1: Summary statistics of the ratio $\sqrt{RQ_t^{(L)}/RQ_t^{(H)}}$

		Min.	1Qu.	Med.	Mean	3Qu.	Max.	S.dev	Skew.	Kurt.
$R_t = \sqrt{\frac{RQ_t^{(5m)}}{RQ_t^{(30s)}}}$	ALV	0.027	0.496	0.624	0.713	0.789	2.941	0.687	2.235	7.395
	BMW	0.080	0.469	0.607	0.698	0.778	2.474	0.687	2.078	5.899
	DTE	0.070	0.339	0.438	0.523	0.566	2.265	0.540	2.254	7.198
	RWE	0.039	0.455	0.574	0.664	0.723	2.772	0.656	2.278	7.610
$R_t = \sqrt{\frac{RQ_t^{(5m)}}{RQ_t^{(1m)}}}$	ALV	0.035	0.593	0.736	0.807	0.894	2.684	0.687	1.750	4.899
	BMW	0.090	0.575	0.722	0.795	0.888	2.384	0.687	1.697	4.391
	DTE	0.099	0.478	0.586	0.660	0.720	1.982	0.595	1.902	5.019
	RWE	0.112	0.564	0.693	0.770	0.851	2.414	0.669	1.688	4.236
$R_t = \sqrt{\frac{RQ_t^{(10m)}}{RQ_t^{(5m)}}}$	ALV	0.071	0.740	0.894	0.947	1.067	1.858	0.708	1.128	2.287
	BMW	0.144	0.733	0.892	0.946	1.066	1.881	0.712	1.123	2.297
	DTE	0.209	0.699	0.850	0.905	1.019	1.922	0.686	1.226	2.528
	RWE	0.230	0.727	0.876	0.937	1.055	1.920	0.714	1.220	2.496
$R_t = \sqrt{\frac{RQ_t^{(15m)}}{RQ_t^{(5m)}}}$	ALV	0.020	0.636	0.828	0.922	1.040	2.585	0.817	1.545	3.551
	BMW	0.082	0.630	0.823	0.923	1.052	2.540	0.818	1.470	3.277
	DTE	0.026	0.604	0.762	0.853	0.960	2.284	0.750	1.532	3.452
	RWE	0.133	0.619	0.804	0.893	1.002	2.429	0.784	1.484	3.340
$R_t = \sqrt{\frac{RQ_t^{(30m)}}{RQ_t^{(5m)}}}$	ALV	0.015	0.517	0.718	0.888	0.996	3.004	0.937	1.802	4.349
	BMW	0.086	0.508	0.719	0.911	1.018	4.274	1.006	2.165	6.970
	DTE	0.032	0.469	0.657	0.803	0.894	3.850	0.864	2.227	7.732
	RWE	0.043	0.481	0.684	0.856	0.948	3.187	0.921	1.901	4.932

Summary statistics of the ratio $\sqrt{RQ_t^{(L)}/RQ_t^{(H)}}$, where $RQ_t^{(H)}$ and $RQ_t^{(L)}$ are realized quarticities based on a high and a low sampling frequency, respectively. Sample period: January 2002 – December 2012. Min.: Minimum; 1Qu.: First Quartile; Med.: Median; Mean; 3Qu.: Third Quartile; Max.: Maximum; S.dev.: Standard deviation; Skew.: Skewness; Kurt.: Kurtosis.

The summary statistics reported in Table 3.1 highlight some empirical regularities. In particular, the highest maximum value of R_t is always given by the ratio between 30 minutes RQ and 5 minutes RQ, while the lowest value is related to $RQ_t^{(10_m)}/RQ_t^{(5_m)}$, for each examined stocks. This is reasonable since the $RQ_t^{(30_m)}$ is smoother than the $RQ_t^{(10_m)}$ and thus the differences with the $RQ_t^{(5_m)}$ are more pronounced. In other words, the 30 minutes RQ is less sensitive to the shocks than the 5 and 10 minutes RQ.

Table 3.2: R_t distribution

		0%	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%	55%	60%	65%	70%	75%	80%	85%	90%	95%	100%
$R_t = \sqrt{\frac{RQ_t^{(5m)}}{RQ_t^{(30s)}}}$	ALV	0.027	0.339	0.388	0.433	0.464	0.496	0.520	0.543	0.570	0.597	0.624	0.651	0.680	0.714	0.751	0.789	0.833	0.892	0.979	1.119	2.941
	BMW	0.080	0.304	0.363	0.410	0.442	0.470	0.495	0.523	0.550	0.576	0.607	0.636	0.669	0.699	0.739	0.778	0.824	0.880	0.962	1.079	2.474
	DTE	0.070	0.242	0.273	0.296	0.319	0.339	0.358	0.377	0.399	0.417	0.438	0.460	0.483	0.510	0.536	0.566	0.607	0.659	0.745	0.865	2.265
	RWE	0.039	0.304	0.359	0.397	0.427	0.455	0.481	0.504	0.527	0.552	0.574	0.601	0.626	0.655	0.686	0.723	0.773	0.833	0.916	1.078	2.772
$R_t = \sqrt{\frac{RQ_t^{(5m)}}{RQ_t^{(1m)}}}$	ALV	0.035	0.415	0.477	0.523	0.564	0.593	0.624	0.651	0.681	0.710	0.736	0.763	0.796	0.826	0.857	0.894	0.938	0.996	1.081	1.220	2.684
	BMW	0.090	0.391	0.462	0.508	0.543	0.575	0.606	0.637	0.665	0.692	0.722	0.749	0.778	0.816	0.849	0.888	0.939	0.991	1.078	1.201	2.384
	DTE	0.099	0.350	0.393	0.428	0.452	0.478	0.503	0.524	0.545	0.567	0.586	0.612	0.638	0.663	0.689	0.720	0.767	0.812	0.888	1.022	1.982
	RWE	0.112	0.388	0.455	0.502	0.537	0.564	0.594	0.617	0.641	0.667	0.693	0.719	0.746	0.778	0.809	0.851	0.899	0.960	1.043	1.178	2.414
$R_t = \sqrt{\frac{RQ_t^{(10m)}}{RQ_t^{(5m)}}}$	ALV	0.071	0.526	0.597	0.652	0.703	0.740	0.769	0.801	0.829	0.861	0.894	0.926	0.961	0.994	1.029	1.067	1.109	1.169	1.249	1.375	1.859
	BMW	0.144	0.518	0.596	0.656	0.693	0.733	0.767	0.799	0.831	0.861	0.892	0.921	0.952	0.989	1.023	1.066	1.115	1.175	1.263	1.368	1.881
	DTE	0.209	0.518	0.582	0.625	0.670	0.699	0.729	0.759	0.795	0.822	0.850	0.877	0.908	0.941	0.981	1.020	1.059	1.103	1.188	1.313	1.922
	RWE	0.230	0.525	0.598	0.650	0.690	0.727	0.758	0.792	0.821	0.845	0.876	0.907	0.937	0.975	1.009	1.055	1.097	1.149	1.237	1.382	1.920
$R_t = \sqrt{\frac{RQ_t^{(15m)}}{RQ_t^{(5m)}}}$	ALV	0.020	0.422	0.498	0.551	0.591	0.636	0.677	0.716	0.751	0.790	0.828	0.868	0.903	0.944	0.990	1.040	1.104	1.177	1.283	1.445	2.585
	BMW	0.082	0.408	0.482	0.541	0.581	0.630	0.673	0.715	0.753	0.787	0.823	0.864	0.905	0.953	0.994	1.052	1.109	1.187	1.290	1.448	2.541
	DTE	0.026	0.402	0.477	0.526	0.564	0.604	0.634	0.666	0.695	0.728	0.762	0.793	0.828	0.867	0.913	0.960	1.021	1.089	1.180	1.339	2.284
	RWE	0.133	0.409	0.479	0.535	0.576	0.619	0.658	0.695	0.733	0.770	0.804	0.834	0.869	0.910	0.949	1.002	1.067	1.142	1.254	1.399	2.429
$R_t = \sqrt{\frac{RQ_t^{(30m)}}{RQ_t^{(5m)}}}$	ALV	0.015	0.290	0.365	0.419	0.472	0.517	0.551	0.588	0.632	0.675	0.718	0.768	0.807	0.857	0.921	0.996	1.077	1.181	1.327	1.540	3.005
	BMW	0.086	0.290	0.358	0.413	0.463	0.508	0.549	0.591	0.630	0.673	0.719	0.768	0.820	0.877	0.938	1.018	1.099	1.211	1.340	1.625	4.275
	DTE	0.032	0.267	0.328	0.379	0.424	0.469	0.505	0.543	0.580	0.620	0.657	0.691	0.730	0.777	0.828	0.894	0.969	1.051	1.182	1.375	3.850
	RWE	0.043	0.283	0.349	0.405	0.442	0.481	0.527	0.562	0.605	0.645	0.684	0.729	0.775	0.827	0.880	0.948	1.030	1.121	1.269	1.488	3.188

A further evidence is that R_t takes on the smallest average values for $RQ_t^{(5_m)}$/ $RQ_t^{(30_s)}$ and the biggest ones for $RQ_t^{(10_m)}/RQ_t^{(5_m)}$, ranging from 0.713 to 0.947 for ALV, from 0.698 to 0.946 for BMW, from 0.523 to 0.905 for DTE and from 0.664 to 0.937 for RWE. The RQ based on a 30 seconds frequency is more heavily affected by jumps than the 5 minutes RQ and consequently the ratio R_t tends to be closer to zero when the bias triggered by jumps is heavier. On the other hand, the value of the ratio tends to move toward unity when jumps and microstructure biases are not present or the two measures involved in the ratio are computed at very close frequencies. The standard deviation of the ratio R_t, for all the analysed assets, tends to increase passing from $RQ_t^{(5_m)}/RQ_t^{(30_s)}$ to $RQ_t^{(30_m)}$/ $RQ_t^{(5_m)}$, since the higher the difference in frequency the higher the variability.

The distribution of R_t in Table 3.2 points out that for $RQ_t^{(5_m)}/RQ_t^{(30_s)}$ approximately the 5% of observations is greater than 1.1, while this value rises to 1.2 for $RQ_t^{(5_m)}/RQ_t^{(1_m)}$. In the other cases, R_t is below 1.2 in a number of cases falling within the range 85-90%. These results seem to suggest that the bias due to jumps widely prevails over the opposite bias effects.

3.5 Empirical analysis

In this section the two variants of the Mixed Frequency Realized GARCH models are fitted to the observed data, considering the full sample period 02 January 2002 – 27 December 2012. As a benchmark we also consider a standard RGARCH model taking the 5 minutes RV as volatility measure. This choice is due to two main reasons. First, as already mentioned, the 5 minutes frequency is a standard rule of the thumb among practitioners and applied researchers. Second, it has been provided empirical evidence that in most settings it ensures better performance in terms of fitting and

forecasting than realized measures computed at different frequencies (see e.g. Liu et al. (2015)). Along the same line, the 5 minutes RV is also taken as basis frequency in the fitted (S)MF-ARGARCH models, but in this case it is "melted" with realized measures which are computed at higher frequencies such as 30 seconds and 1 minute, or at lower frequencies such as 10, 15 and 30 minutes.

The estimation results for RGARCH and SMF-ARGARCH models are presented together in section 3.5.1, while the results for MF-ARGARCH models are reported in section 3.5.2. It is worth reminding that, since each model is using different couples of realized variances, the MF-ARGARCH and SMF-ARGARCH models are not comparable on the basis of the maximized full log-likelihood value, even within the same category of models.

3.5.1 Estimation results for the Single equation Mixed Frequency Adaptive Realized GARCH models

The in-sample estimation results for the standard RGARCH and the SMF-ARGARCH models are reported in Table 3.3, showing the parameter estimates and robust standard errors in small font under the parameter values, as well as the returns partial log-likelihood $\ell(r)$, the full log-likelihood $\mathcal{L}(r, x)$ and Bayesian Information Criterion (BIC), for the four analysed stocks.

The superscript $(5_m, j_k)$, where $j = 1, 10, 15, 30$ and $k = m, s$ with $m =$ *minutes* and $s =$ *seconds*, related to the SMF-ARGARCH model, denotes the sampling frequency of the realized measures which are combined in the measurement equation for $log(\tilde{x}_t)$. It follows that the coefficient α_t, which is a linear function of the ratio $\sqrt{RQ_t^{(L)}/RQ_t^{(H)}}$, is always referred to

the realized variance computed at a 5 minutes frequency, whereas $(1 - \alpha_t)$ provides the weight of the realized measure taking a higher or a lower frequency within the the SMF-ARGARCH$^{(5_m, j_k)}$ model.

The log-linear specification has the advantage of giving more flexibility to the models, since the positivity of the conditional variance is, by construction, always fulfilled without imposing constraints on the parameters. As a result the intercept ω takes on even negative values, though in most cases it is not significant for all the examined assets. The parameter γ, that explains the impact of the past realized measure on future volatility, presents quite different values among the proposed models, since it ranges from 0.385 to 0.476 for ALV, from 0.286 to 0.406 for BMW, from 0.313 to 0.450 for DTE and from 0.315 to 0.375 for RWE. Interestingly, the upper limit always corresponds to models given by the combination of the 5 minutes RV with realized measures computed at higher frequencies. In particular, for ALV, BMW and DTE the upper bound is provided by the SMF-ARGARCH$^{(5_m, 30_s)}$, while for RWE it comes out from the SMF-ARGARCH$^{(5_m, 1_m)}$. The values of the parameter β of the RGARCH and SMF-ARGARCH models are lower than those usually estimated in a GARCH(1,1) model, with an average of 0.632.

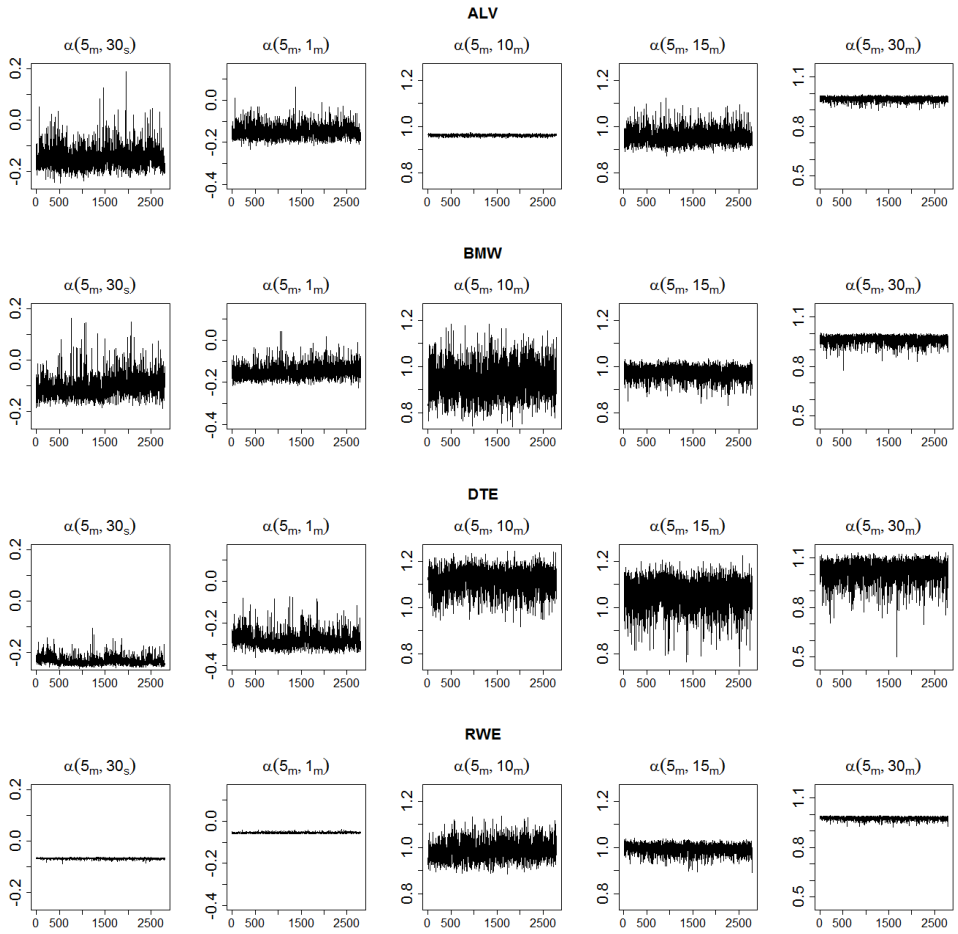
Looking at the parameters that characterise the measurement equation, ξ mainly takes negative values, but as for the intercept of the log-conditional variance it is often not significant at the usual 5% level. The parameters of the leverage function $\tau(z)$ are always significant, where τ_1 is negative and τ_2 is positive. The coefficient φ is close to 1 for all the proposed models. This suggests that the log-transformed realized measure (modelled by the measurement equation) is approximately unbiased for the conditional variance, confirming the findings in Hansen et al. (2012).

Table 3.3: In-Sample Estimation Results for RGARCH and SMF-ARGARCH

		ω	γ	β	ξ	φ	τ_1	τ_2	σ_u^2	α_0	α_1	$\ell(r)$	$\mathcal{L}(r,x)$	BIC
ALV	RGARCH	-0.132⋆ 0.257	0.401 0.037	0.583 0.033	-0.094⋆ 0.610	0.989 0.073	-0.069 0.008	0.108 0.006	0.185 0.006	-	-	7605.335	6001.795	-11940.116
	SMF-ARGARCH$^{(5_m,30_s)}$	-0.103⋆ 0.131	0.476 0.035	0.530 0.033	-0.114⋆ 0.254	0.948 0.031	-0.064 0.006	0.061 0.005	0.107 0.004	-0.249 0.053	0.149 0.060	7589.607	6753.510	-13427.679
	SMF-ARGARCH$^{(5_m,1_m)}$	-0.156⋆ 0.136	0.385 0.030	0.605 0.028	0.086⋆ 0.339	0.986 0.042	-0.068 0.007	0.072 0.005	0.123 0.005	-0.239 0.076	0.113⋆ 0.086	7593.173	6555.217	-13031.092
	SMF-ARGARCH$^{(5_m,10_m)}$	0.000⋆ 0.108	0.400 0.032	0.600 0.030	-0.375⋆ 0.245	0.956 0.030	-0.069 0.008	0.109 0.006	0.185 0.006	0.974 0.140	-0.013⋆ 0.144	7606.465	6004.173	-11929.005
	SMF-ARGARCH$^{(5_m,15_m)}$	-0.040⋆ 0.106	0.389 0.030	0.606 0.029	-0.288⋆ 0.249	0.967 0.030	-0.068 0.008	0.110 0.006	0.184 0.006	0.871 0.111	0.098⋆ 0.107	7606.384	6005.540	-11931.739
	SMF-ARGARCH$^{(5_m,30_m)}$	-0.028⋆ 0.108	0.392 0.032	0.604 0.029	-0.289⋆ 0.253	0.967 0.031	-0.066 0.008	0.110 0.006	0.182 0.006	0.992 0.046	-0.032⋆ 0.047	7606.367	6005.136	-11930.930
BMW	RGARCH	-0.127⋆ 0.111	0.286 0.027	0.701 0.023	0.113⋆ 0.353	1.004 0.043	-0.023 0.008	0.082 0.006	0.171 0.006	-	-	7466.828	5970.296	-11877.120
	SMF-ARGARCH$^{(5_m,30_s)}$	0.040⋆ 0.187	0.406 0.035	0.614 0.028	-0.429⋆ 0.428	0.910 0.052	-0.040 0.006	0.054 0.004	0.101 0.004	-0.201 0.055	0.147 0.071	7454.505	6692.501	-13305.661
	SMF-ARGARCH$^{(5_m,1_m)}$	0.103⋆ 0.124	0.376 0.030	0.648 0.025	-0.587 0.294	0.898 0.036	-0.044 0.006	0.060 0.004	0.111 0.004	-0.237 0.074	0.117⋆ 0.084	7461.441	6566.673	-13054.004
	SMF-ARGARCH$^{(5_m,10_m)}$	-0.019⋆ 0.135	0.299 0.027	0.701 0.024	-0.333⋆ 0.421	0.950 0.051	-0.028 0.008	0.082 0.006	0.171 0.006	0.701 0.147	0.257⋆ 0.152	7468.550	5976.876	-11874.411
	SMF-ARGARCH$^{(5_m,15_m)}$	0.097⋆ 0.095	0.314 0.026	0.701 0.024	-0.701 0.262	0.905 0.032	-0.029 0.008	0.082 0.006	0.171 0.006	1.045 0.086	-0.084⋆ 0.094	7467.428	5974.616	-11869.890
	SMF-ARGARCH$^{(5_m,30_m)}$	0.111⋆ 0.092	0.316 0.027	0.701 0.024	-0.710 0.248	0.905 0.031	-0.028 0.008	0.084 0.006	0.171 0.006	1.007 0.040	-0.054⋆ 0.043	7467.200	5976.559	-11873.777
DTE	RGARCH	-0.067⋆ 0.133	0.337 0.032	0.662 0.027	-0.140⋆ 0.373	0.965 0.043	-0.019 0.007	0.091 0.005	0.142 0.005	-	-	8234.436	7020.422	-13977.370
	SMF-ARGARCH$^{(5_m,30_s)}$	-0.358⋆ 0.374	0.450 0.040	0.564 0.040	0.450⋆ 0.807	0.930 0.094	-0.028 0.004	0.033 0.003	0.045 0.002	-0.269 0.021	0.072 0.026	8156.010	8535.395	-16991.449
	SMF-ARGARCH$^{(5_m,1_m)}$	0.392 0.102	0.450 0.034	0.637 0.037	-1.243 0.219	0.766 0.025	-0.031 0.005	0.047 0.004	0.063 0.003	-0.379 0.035	0.153 0.042	8215.869	8127.557	-16175.773
	SMF-ARGARCH$^{(5_m,10_m)}$	0.184⋆ 0.102	0.378 0.032	0.652 0.028	-0.820 0.228	0.885 0.026	-0.031 0.007	0.089 0.005	0.139 0.005	1.284 0.109	-0.192⋆ 0.121	8235.996	7032.858	-13986.374
	SMF-ARGARCH$^{(5_m,15_m)}$	0.121⋆ 0.105	0.371 0.031	0.651 0.028	-0.684 0.248	0.901 0.028	-0.031 0.007	0.089 0.005	0.139 0.005	1.231 0.071	-0.212 0.077	8236.939	7033.326	-13987.311
	SMF-ARGARCH$^{(5_m,30_m)}$	-0.107⋆ 0.102	0.313 0.026	0.681 0.025	-0.010⋆ 0.303	0.978 0.035	-0.029 0.007	0.089 0.006	0.138 0.005	1.141 0.036	-0.166 0.046	8231.196	7029.375	-13979.409
RWE	RGARCH	-0.180 0.085	0.338 0.028	0.646 0.026	0.088⋆ 0.230	0.993 0.027	-0.038 0.008	0.082 0.005	0.164 0.006	-	-	7986.446	6567.031	-13070.588
	SMF-ARGARCH$^{(5_m,30_s)}$	-0.586⋆ 0.171	0.374 0.035	0.576 0.029	1.167 0.511	1.086 0.059	-0.044 0.006	0.053 0.004	0.099 0.004	-0.062⋆ 0.050	-0.009⋆ 0.062	7970.711	7235.797	-14392.253
	SMF-ARGARCH$^{(5_m,1_m)}$	-0.557 0.133	0.375 0.031	0.576 0.030	1.031 0.371	1.079 0.043	-0.043 0.006	0.057 0.004	0.112 0.004	-0.063⋆ 0.070	0.010⋆ 0.083	7973.697	7070.429	-14061.516
	SMF-ARGARCH$^{(5_m,10_m)}$	-0.394 0.152	0.316 0.030	0.644 0.026	0.695⋆ 0.511	1.063 0.060	-0.040 0.008	0.082 0.005	0.162 0.005	0.851 0.145	0.147⋆ 0.158	7990.092	6570.804	-13062.266
	SMF-ARGARCH$^{(5_m,15_m)}$	-0.354 0.132	0.315 0.028	0.650 0.026	0.587⋆ 0.430	1.050 0.050	-0.041 0.008	0.082 0.005	0.162 0.005	1.049 0.071	-0.065⋆ 0.078	7988.912	6570.503	-13061.665
	SMF-ARGARCH$^{(5_m,30_m)}$	-0.378 0.118	0.319 0.027	0.642 0.027	0.636⋆ 0.373	1.056 0.044	-0.040 0.008	0.083 0.005	0.162 0.005	0.994 0.042	-0.023⋆ 0.051	7990.151	6570.892	-13062.443

In-sample parameter estimates for RGARCH and SMF-ARGARCH$^{(5_m,j_k)}$, where $j = 1, 10, 15, 30$ and $k = m, s$ with $m = \textit{minutes}$ and $s = \textit{seconds}$. ⋆: parameter not significant at 5%. $\ell(r)$: partial log-likelihood. $\mathcal{L}(r,x)$: log-likelihood. BIC: Bayesian Information Criterion. Robust standard errors are reported in small font under the parameter values. Sample period 02 January 2002 - 27 December 2012.

Figure 3.1: Time-varying coefficient α_t of SMF-ARGARCH models



The figure shows the dynamics of the coefficient $\alpha_t = \alpha_0 + \alpha_1 R_{t-1}$ which is function of the different frequencies of the realized quarticities involved in the ratio $R_t = \sqrt{RQ_t^{(L)}/RQ_t^{(H)}}$.

However, for the SMF-ARGARCH models, φ takes smaller values than the ones of the standard RGARCH according to the stocks ALV, BMW and DTE. Exceptions are made for SMF-ARGARCH$^{(5m,30m)}$, for DTE, while for the stock RWE the different specifications of the SMF-ARGARCH always provide higher values of φ than RGARCH. Furthermore, the SMF-ARGARCH$^{(5m,30s)}$ and the SMF-ARGARCH$^{(5m,1m)}$ always show the greatest discrepancy from 1 of the φ coefficient. On the other hand, the variance of the measurement error $\tilde{u}_t$ of these models takes on the lowest values for each asset, confirming that the higher the frequency, the lower the error variability, but of course a greater bias occurs. Finally the SMF-ARGARCH models yield values of $\sigma^2_{\tilde{u}}$ slightly lower or at most equal to those given by the RGARCH.

Regarding the coefficients of the weighting function α_t, which is always linked to the 5 minutes realized measure, it can be noted that α_0 is significant for each specification of the SMF-ARGARCH model, except for SMF-ARGARCH$^{(5m,30s)}$ and SMF-ARGARCH$^{(5m,1m)}$ for the stock RWE, assuming negative values when the 5 minutes realized measure is associated with higher frequency realized measures and positive values when it is combined with realized measures computed according to a lower sampling frequency. The parameter α_1, which is related to the dynamics of the ratio R_t that allows to adjust bias effects, is significant just in a few cases and in particular for the SMF-ARGARCH$^{(5m,30s)}$, for the stocks ALV and BMW, and for the SMF-ARGARCH where the 5 minute RV is combined with the realized measures based on 30 seconds, 1 minute, 15 and 30 minutes, for the stock DTE, while it is never significant for RWE.

Graphically, the magnitude of the weight α_t is displayed in Figure 3.1, where it can be easily seen that when α_1 is not statistically significant the variation range of α_t is very small consistently with a constant weights

model. However, in cases in which the coefficient α_1 is significant the sign of the estimated parameter is such that, when $R_t < 1$, namely when $RQ_t^{(H)} > RQ_t^{(L)}$, the weighting function α_t is designed to balance the realized measures in $\tilde{x}_t$, shrinking the level of $x_t^{(H)}$ (which is up-weighted) and raising the one of $x_t^{(L)}$ (which is down-weighted). On the other hand, when $R_t > 1$, that is $RQ_t^{(H)} < RQ_t^{(L)}$, the impact of $x_t^{(L)}$ (which is up-weighted) is lowered, whereas that of $x_t^{(H)}$ (which is down-weighted) is increased. Accordingly, the combination of the realized measures sampled at different frequencies by means of the weighting function α_t yields an optimized measure $\tilde{x}_t$ in terms of trade-off between bias and variability.

In terms of goodness of fit the estimated models can only be compared in terms of partial log-likelihood and, namely, of the returns component reported in the $\ell(r)$ column of Table 3.3. The results show that, while SMF-ARGARCH models combining the 5 minutes frequency with a lower frequency ($10m$, $15m$, $30m$) lead to a partial log-likelihood improvement with respect to the basic Realized GARCH, this is not case when the basis 5 minutes frequency is combined with realized measures based on higher frequency returns (30 seconds and 1 minute).

3.5.2 Estimation results for the Mixed Frequency Adaptive Realized GARCH models

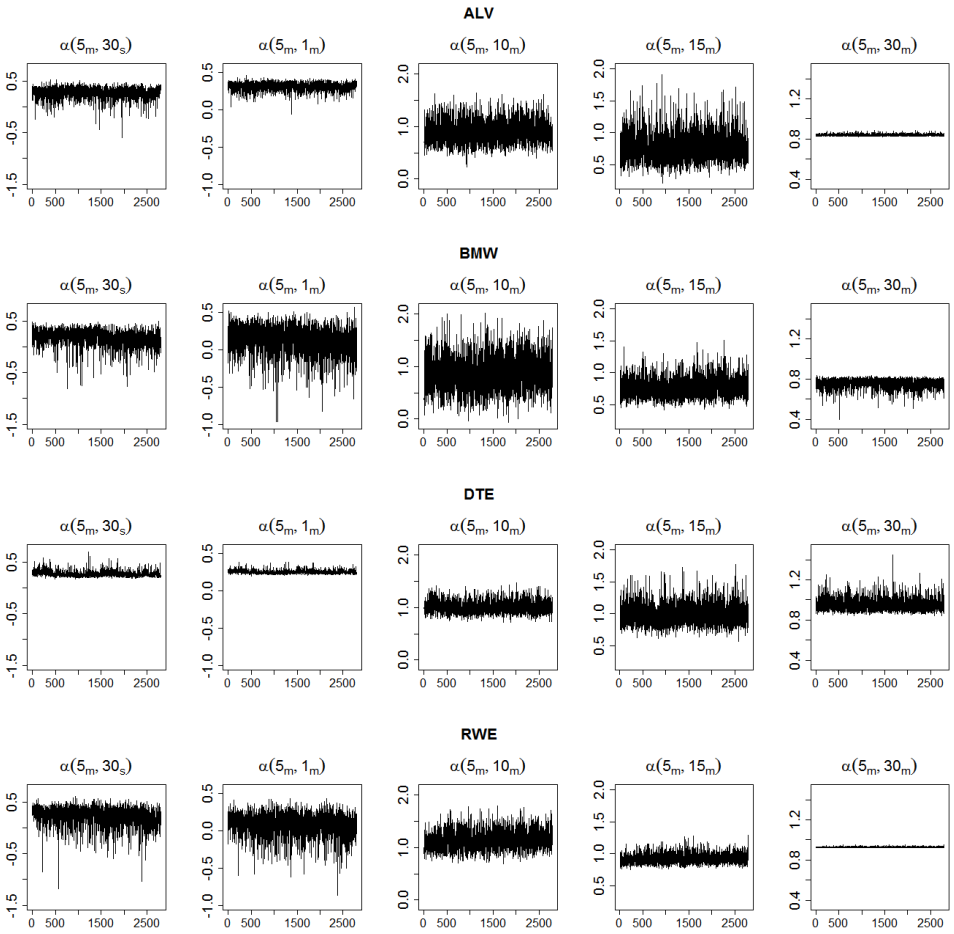
Table 3.4 shows the parameter estimates, with the associated robust standard errors in small font underneath, for the MF-ARGARCH models. Unlike the results in Table 3.3, the intercept ω of the GARCH-equation takes mainly positive values and it is significant in half of the cases. The parameter γ, that relates the realized measure to the future conditional variance of returns, lies in the range between 0.326 and 0.476 for ALV, between 0.276 and 0.383 for BMW, between 0.312 and 0.448 for DTE

Table 3.4: In-Sample Estimation Results for MF-ARGARCH

		ω	γ	β	ξ_H	φ_H	$\tau_{H,1}$	$\tau_{H,2}$	$\sigma^2_{u_H}$	ξ_L	φ_L	$\tau_{L,1}$	$\tau_{L,2}$	$\sigma^2_{u_L}$	α_0	α_1	$\ell(r)$	$\mathcal{L}(r,x)$	BIC
ALV	MF-ARGARCH$^{(5_m,30_s)}$	0.241⋆	0.476	0.563	-1.002	0.845	-0.068	0.063	0.114	-0.458⋆	0.944	-0.066	0.110	0.186	0.543	-0.388	7603.889	5124.153	-10129.293
		0.142	0.039	0.032	0.235	0.029	0.006	0.005	0.004	0.270	0.033	0.008	0.007	0.006	0.098	0.166			
	MF-ARGARCH$^{(5_m,1_m)}$	0.311	0.440	0.603	-1.038	0.855	-0.068	0.076	0.126	-0.822	0.904	-0.064	0.108	0.184	0.462	-0.193⋆	7601.811	4947.795	-9776.578
		0.127	0.038	0.030	0.221	0.027	0.007	0.006	0.005	0.232	0.029	0.008	0.007	0.006	0.167	0.252			
	MF-ARGARCH$^{(5_m,10_m)}$	0.064⋆	0.407	0.598	-0.526	0.938	-0.069	0.109	0.184	-0.445⋆	0.956	-0.059	0.121	0.218	0.172⋆	0.789	7605.230	4170.814	-8222.617
		0.112	0.033	0.030	0.236	0.029	0.008	0.006	0.006	0.239	0.029	0.008	0.006	0.007	0.276	0.327			
	MF-ARGARCH$^{(5_m,15_m)}$	0.137⋆	0.326	0.686	-0.632	0.931	-0.064	0.119	0.186	-0.643	0.943	-0.053	0.139	0.253	0.200⋆	0.659	7600.141	3955.544	-7792.076
		0.097	0.029	0.023	0.248	0.031	0.008	0.007	0.006	0.244	0.030	0.009	0.007	0.007	0.169	0.205			
	MF-ARGARCH$^{(5_m,30_m)}$	0.126⋆	0.409	0.604	-0.605	0.929	-0.068	0.109	0.185	-0.639	0.947	-0.052	0.166	0.334	0.823	0.022⋆	7605.056	3580.308	-7041.603
		0.128	0.036	0.030	0.272	0.033	0.008	0.006	0.006	0.259	0.032	0.010	0.008	0.010	0.087	0.105			
BMW	MF-ARGARCH$^{(5_m,30_s)}$	0.029⋆	0.329	0.683	-0.414⋆	0.914	-0.037	0.056	0.104	-0.301⋆	0.954	-0.028	0.082	0.170	0.550	-0.551	7464.138	5183.008	-10247.004
		0.103	0.025	0.022	0.281	0.034	0.006	0.004	0.004	0.284	0.035	0.008	0.006	0.006	0.113	0.195			
	MF-ARGARCH$^{(5_m,1_m)}$	0.489	0.383	0.684	-1.448	0.796	-0.040	0.061	0.113	-1.320	0.832	-0.030	0.082	0.169	0.633	-0.672	7456.337	5063.977	-10008.942
		0.179	0.044	0.025	0.318	0.039	0.006	0.004	0.004	0.328	0.040	0.008	0.006	0.006	0.173	0.259			
	MF-ARGARCH$^{(5_m,10_m)}$	0.232	0.307	0.722	-1.028	0.866	-0.028	0.082	0.170	-0.896	0.891	-0.028	0.097	0.204	-0.232⋆	1.192	7464.908	4233.351	-8347.690
		0.104	0.027	0.021	0.275	0.034	0.008	0.006	0.006	0.278	0.034	0.009	0.007	0.007	0.245	0.302			
	MF-ARGARCH$^{(5_m,15_m)}$	-0.003⋆	0.276	0.724	-0.392⋆	0.942	-0.027	0.083	0.171	-0.273⋆	0.972	-0.017⋆	0.112	0.246	0.381	0.444	7470.578	3978.700	-7838.388
		0.114	0.022	0.021	0.365	0.044	0.008	0.006	0.006	0.387	0.047	0.009	0.007	0.008	0.163	0.193			
	MF-ARGARCH$^{(5_m,30_m)}$	0.401	0.327	0.724	-1.542	0.803	-0.028	0.081	0.172	-1.464	0.836	-0.012⋆	0.139	0.327	0.841	-0.103⋆	7457.208	3577.301	-7035.590
		0.089	0.030	0.024	0.194	0.024	0.008	0.006	0.006	0.200	0.024	0.011	0.009	0.010	0.098	0.109			
DTE	MF-ARGARCH$^{(5_m,30_s)}$	0.675	0.426	0.688	-1.863	0.678	-0.029	0.042	0.052	-0.952	0.869	-0.026	0.089	0.142	0.159	0.241	8228.113	7160.183	-14201.354
		0.161	0.041	0.024	0.226	0.025	0.004	0.003	0.003	0.271	0.030	0.007	0.005	0.005	0.035	0.109			
	MF-ARGARCH$^{(5_m,1_m)}$	0.628	0.448	0.651	-1.737	0.726	-0.031	0.056	0.068	-1.098	0.856	-0.028	0.091	0.139	0.188	0.098⋆	8233.133	6827.930	-13536.849
		0.127	0.039	0.027	0.179	0.021	0.005	0.004	0.003	0.202	0.023	0.007	0.005	0.005	0.065	0.169			
	MF-ARGARCH$^{(5_m,10_m)}$	0.304	0.387	0.655	-1.154	0.848	-0.030	0.090	0.140	-0.868	0.897	-0.023	0.105	0.187	0.622	0.443⋆	8236.338	5407.409	-10695.805
		0.098	0.034	0.030	0.195	0.023	0.007	0.005	0.005	0.205	0.024	0.008	0.007	0.006	0.232	0.302			
	MF-ARGARCH$^{(5_m,15_m)}$	0.374	0.406	0.643	-1.194	0.844	-0.031	0.090	0.139	-1.037	0.887	-0.023	0.118	0.221	0.553	0.534	8234.955	5175.955	-10232.897
		0.116	0.037	0.029	0.216	0.025	0.007	0.005	0.005	0.219	0.025	0.009	0.006	0.007	0.134	0.203			
	MF-ARGARCH$^{(5_m,30_m)}$	0.251⋆	0.312	0.720	-0.948	0.876	-0.038	0.094	0.138	-0.556⋆	0.956	-0.022	0.149	0.304	0.835	0.160⋆	8234.256	4687.264	-9255.516
		0.188	0.048	0.028	0.452	0.052	0.007	0.006	0.005	0.461	0.053	0.010	0.008	0.009	0.075	0.121			
RWE	MF-ARGARCH$^{(5_m,30_s)}$	0.273	0.315	0.728	-1.092	0.828	-0.043	0.053	0.105	-1.183	0.848	-0.043	0.084	0.164	0.644	-0.660	7949.103	5718.930	-11318.848
		0.073	0.021	0.018	0.200	0.024	0.006	0.004	0.004	0.205	0.024	0.008	0.005	0.005	0.106	0.194			
	MF-ARGARCH$^{(5_m,1_m)}$	-0.149⋆	0.397	0.600	0.066⋆	0.966	-0.041	0.057	0.111	-0.079⋆	0.971	-0.035	0.083	0.159	0.498	-0.567	7971.570	5673.857	-11228.701
		0.144	0.040	0.031	0.365	0.043	0.006	0.004	0.004	0.355	0.042	0.008	0.005	0.005	0.137	0.215			
	MF-ARGARCH$^{(5_m,10_m)}$	-0.130⋆	0.314	0.676	-0.004⋆	0.982	-0.039	0.082	0.161	-0.342⋆	0.953	-0.032	0.097	0.198	0.547	0.645	7981.403	4870.348	-9621.684
		0.102	0.024	0.024	0.300	0.035	0.008	0.005	0.005	0.291	0.034	0.009	0.006	0.006	0.251	0.317			
	MF-ARGARCH$^{(5_m,15_m)}$	0.147⋆	0.417	0.606	-0.761	0.894	-0.038	0.080	0.163	-1.115	0.871	-0.037	0.108	0.234	0.727	0.234⋆	7978.065	4625.599	-9132.186
		0.120	0.036	0.031	0.257	0.030	0.008	0.005	0.005	0.240	0.028	0.009	0.006	0.007	0.134	0.184			
	MF-ARGARCH$^{(5_m,30_m)}$	0.386	0.357	0.692	-1.246	0.841	-0.040	0.081	0.165	-1.910	0.790	-0.028	0.139	0.319	0.919	0.011⋆	7957.930	4184.810	-8250.607
		0.126	0.036	0.025	0.261	0.031	0.008	0.006	0.006	0.229	0.027	0.011	0.007	0.009	0.082	0.109			

In-sample parameter estimates for MF-ARGARCH$^{(5_m,j_k)}$, where $j = 1, 10, 15, 30$ and $k = m, s$ with $m = \textit{minutes}$ and $s = \textit{seconds}$. ⋆: parameter not significant at 5%. $\ell(r)$: partial log-likelihood. $\mathcal{L}(r,x)$: log-likelihood. BIC: Bayesian Information Criterion. Robust standard errors are reported in small font under the parameter values. Sample period 02 January 2002 - 27 December 2012.

Figure 3.2: Time-varying coefficient α_t of MF-ARGARCH models



The figure shows the dynamics of the coefficient $\alpha_t = \alpha_0 + \alpha_1 R_{t-1}$ which is function of the different frequencies of the realized quarticities employed in the ratio $R_t = \sqrt{RQ_t^{(L)}/RQ_t^{(H)}}$.

and between 0.314 and 0.417 for RWE. These ranges always include the value of γ given by the standard RGARCH model. Furthermore, as in the previous analysis, the upper bound of the range for γ is reached by models mixing the 5 minutes RV with realized measures computed at the higher frequencies of 30 seconds and 1 minute, except for the stock RWE where it corresponds to the MF-ARGARCH$^{(5_m,15_m)}$. Also, as for the SMF-ARGARCH model, the parameter β is much smaller than the customary values encountered in a typical GARCH(1,1) model estimated on daily returns, since it takes an average value of 0.662.

The intercepts of the measurement equations for $x_t^{(H)}$ and $x_t^{(L)}$, ξ_H and ξ_L respectively, take on negative values (almost always significant) except for the MF-ARGARCH$^{(5_m,1_m)}$ even if, in this case, ξ_H is not statistically significant. The parameters τ_1 and τ_2 of the leverage function $\tau(z)$ take negative and positive values respectively, both for the high frequency measure and for the low frequency measure, confirming the results displayed in Table 3.3. The coefficient φ_H is always lower than its counterpart φ_L for the stocks ALV, BMW and DTE, while for the asset RWE this holds only for the MF-ARGARCH$^{(5_m,30_s)}$ and the MF-ARGARCH$^{(5_m,1_m)}$. Furthermore, φ_H and φ_L are smaller than the φ coefficient given by the standard RGARCH model for all the examined cases.

The variance of the measurement noise of the high frequency realized measure $\sigma^2_{\tilde{u}_H}$ is, as expected, always smaller than the variance related to the realized measure based on a lower frequency $\sigma^2_{\tilde{u}_L}$. This result is in line with what is shown in Table 3.3, since the variability of the error decreases as the sampling frequency increases, even if the bias goes up. On the other hand, when the sampling frequency decreases the realized measure is less variable, but less biased. Furthermore, the variance of the measurement

noise associated to the 5 minutes RV provided by the MF-ARGARCH models is perfectly in line with the values of σ_u^2 of the standard RGARCH model for all the analysed stocks.

It is particularly interesting to see that, unlike the SMF-ARGARCH models in Table 3.3, the parameter α_1, which drives the impact of the upward and downward bias in the weighting function α_t, is significant for several specifications of the MF-ARGARCH model. This is probably due to the greater flexibility of the MF-ARGARCH model that, incorporating two different measurement equations, allows to separately characterise the bias and variability properties of the realized measures involved in the model. The coefficient α_0 mainly takes positive values, except for the MF-ARGARCH$^{(5_m,10_m)}$ for BMW, even though it is not statistically significant. Similarly, the estimated α_0 is not significant at the 5% level for MF-ARGARCH$^{(5_m,10_m)}$ and MF-ARGARCH$^{(5_m,15_m)}$ models fitted to the asset ALV. Figure 3.2 depicts the dynamics of the time-varying coefficient α_t for the MF-ARGARCH models. The variation range showed in each plot is larger than those displayed in Figure 3.1 for the SMF-ARGARCH models.

Similarly to what observed for the SMF-ARGARCH model, the analysis of the returns partial log-likelihood reveals that the combination of the 5 minutes RV with a realized measure based on a lower frequency, in general, provides a better performance than models combining the basis 5 minutes frequency with realized measures based on 30 seconds or 1 minute returns.

An overall comparison of the analysed models in terms of partial log-likelihood is reported in Table 3.5, where the values in boldface refer to the preferred models. The results show that the values of the $\ell(r)$ are always maximized by the SMF-ARGARCH and the MF-ARGARCH

models. In particular, the best performance for the stock ALV is achieved by the SMF-ARGARCH$^{(5_m,10_m)}$, while for BMW it is given by the MF-ARGARCH$^{(5_m,15_m)}$. For DTE and RWE the single equation model prevails through the combinations $(5_m, 15_m)$ and $(5_m, 30_m)$, respectively.

Table 3.5: In-sample partial log-likelihood comparison

	ALV	BMW	DTE	RWE
RGARCH	7605.335	7466.828	8234.436	7986.446
SMF-ARGARCH$^{(5_m,30_s)}$	7589.607	7454.505	8156.010	7970.711
SMF-ARGARCH$^{(5_m,1_m)}$	7593.173	7461.441	8215.869	7973.697
SMF-ARGARCH$^{(5_m,10_m)}$	**7606.465**	7468.550	8235.996	7990.092
SMF-ARGARCH$^{(5_m,15_m)}$	7606.384	7467.428	**8236.939**	7988.912
SMF-ARGARCH$^{(5_m,30_m)}$	7606.367	7467.200	8231.196	**7990.151**
MF-ARGARCH$^{(5_m,30_s)}$	7603.889	7464.138	8228.113	7949.103
MF-ARGARCH$^{(5_m,1_m)}$	7601.811	7456.337	8233.133	7971.570
MF-ARGARCH$^{(5_m,10_m)}$	7605.230	7464.908	8236.338	7981.403
MF-ARGARCH$^{(5_m,15_m)}$	7600.141	**7470.578**	8234.955	7978.065
MF-ARGARCH$^{(5_m,30_m)}$	7605.056	7457.208	8234.256	7957.930

In **bold** the preferred model. Sample period 02 January 2002 - 27 December 2012.

3.6 Out-of-sample forecasting

The out-of-sample predictive ability of the fitted models has been assessed by means of a rolling window forecasting exercise using a window of 1500 days, where the model parameters are recursively re-estimated every 22 days. The forecasting exercise covers the 2008 credit crisis and the turmoil period related to the instability of the Euro area in the late 2011, ranging from 15 November 2007 to 27 December 2012 for a total of 1298 days. The forecast accuracy has been evaluated by the predictive partial log-likelihood and by the QLIKE loss function, while the significance of differences in forecasting performance across different models has been

tested by the Model Confidence Set (MCS) approach of Hansen et al. (2011) considering the confidence levels of 75% and 90%.

3.6.1 Forecast evaluation

Following Hansen et al. (2012), for each model the out-of-sample performance of the proposed models has been first assessed by computing the predictive log-likelihood over the out-of-sample forecasting period

$$\hat{\ell}(r,x)_{t+1} = -\frac{1}{2}\left[log(2\pi) + log(\hat{h}_{t+1}) + \frac{r^2_{t+1}}{\hat{h}_{t+1}}\right] - \frac{1}{2}\left[log(2\pi) + log(\hat{\sigma}^2_u) + \frac{u^2_{t+1}}{\hat{\sigma}^2_u}\right] \tag{3.13}$$

where $\hat{\ell}(r,x)_{t+1}$ is the one-day ahead predictive log-likelihood and $\hat{h}_{t+1}$ denotes the forecast conditional variance. These log-density estimates are computed for each day in the forecast period and consequently summed to get the overall predictive log-likelihood for each model. However, it is worth reminding once more that, because of the different specification of the error term in the measurement equation, the RGARCH, SMF-ARGARCH and MF-ARGARCH are not comparable in terms of overall log-likelihood, while it makes sense to compare the models looking at the returns partial log-likelihood.

In this regard, Table 3.6 reports the values of the predictive partial log-likelihood of the returns component for the different models considered in our analysis. It clearly emerges that the MF-ARGARCH model always outperforms the competitors, since it gives the greatest value of the predictive partial log-likelihood for all the examined stocks. In particular, for ALV the best performer is the MF-ARGARCH$^{(5_m,30_m)}$, whereas for BMW and DTE the MF-ARGARCH$^{(5_m,15_m)}$ outperforms the other models. Finally, for RWE the partial predictive log-likelihood is maximized by the MF-ARGARCH$^{(5_m,10_m)}$.

Table 3.6: Predictive partial log-likelihood

	ALV	BMW	DTE	RWE
RGARCH	3410.194	3207.235	3791.972	3642.903
SMF-ARGARCH$^{(5_m,30_s)}$	3398.878	3191.235	3769.119	3635.468
SMF-ARGARCH$^{(5_m,1_m)}$	3402.977	3200.283	3778.495	3632.168
SMF-ARGARCH$^{(5_m,10_m)}$	3408.670	3208.552	3789.531	3642.539
SMF-ARGARCH$^{(5_m,15_m)}$	3411.735	3208.200	3788.943	3644.480
SMF-ARGARCH$^{(5_m,30_m)}$	3412.000	3208.497	3787.941	3640.509
MF-ARGARCH$^{(5_m,30_s)}$	3405.494	3202.308	3786.371	3644.076
MF-ARGARCH$^{(5_m,1_m)}$	3411.271	3201.788	3790.531	3647.868
MF-ARGARCH$^{(5_m,10_m)}$	3409.017	3208.683	3794.819	**3648.373**
MF-ARGARCH$^{(5_m,15_m)}$	3411.641	**3212.213**	**3796.071**	3648.090
MF-ARGARCH$^{(5_m,30_m)}$	**3413.778**	3210.893	3794.701	3648.176

In **bold** the best model in terms of predictive partial log-likelihood of the returns component. Out-of-sample period 15 November 2007 - 27 December 2012.

In order to assess the predictive ability of the models, we also consider the QLIKE loss function, which is defined as

$$QLIKE = \frac{1}{T}\sum_{t=1}^{T} log(\hat{h}_t) + \frac{RV_t}{\hat{h}_t}. \tag{3.14}$$

Among the class of robust loss functions for volatility forecast evaluation, the QLIKE loss function has been found to be powerful in rejecting poorly performing predictors (see e.g. Patton (2011) and Liu et al. (2015) among others). The top panel of Table 3.7 shows that the QLIKE loss function, using the 5 minutes realized variance as volatility proxy, is minimized by the MF-ARGARCH models mixing the 5 minutes RV with realized measures based on lower frequencies such as 10, 15 and 30 minutes.

In order to test the significance of differences in the QLIKE values across different models the MCS approach of Hansen et al. (2011) is used. The

bottom panel of Table 3.7 reports the MCS p-values corresponding the *Tmax* statistic, based on 3000 bootstrap resamples. The optimal block-length used for the implementation of the bootstrap algorithm has been estimated following the procedure proposed in Patton et al. (2009).

Table 3.7: Average values of QLIKE loss using 5-min RV as volatility proxy (top) and MCS p-values (bottom). For each stock: **bold**: minimum loss; red: model $\in$ 75% MCS; blue: model $\in$ 90% MCS

QLIKE				
	ALV	BMW	DTE	RWE
RGARCH	-6.973	-6.675	-7.347	-7.259
SMF-ARGARCH$^{(5_m,30_s)}$	-6.954	-6.643	-7.269	-7.248
SMF-ARGARCH$^{(5_m,1_m)}$	-6.961	-6.662	-7.310	-7.238
SMF-ARGARCH$^{(5_m,10_m)}$	-6.972	-6.676	-7.346	-7.261
SMF-ARGARCH$^{(5_m,15_m)}$	-6.975	-6.676	-7.348	-7.264
SMF-ARGARCH$^{(5_m,30_m)}$	-6.975	-6.675	-7.352	-7.258
MF-ARGARCH$^{(5_m,30_s)}$	-6.967	-6.670	-7.331	-7.268
MF-ARGARCH$^{(5_m,1_m)}$	-6.976	-6.670	-7.343	-7.273
MF-ARGARCH$^{(5_m,10_m)}$	-6.976	-6.684	**-7.366**	**-7.281**
MF-ARGARCH$^{(5_m,15_m)}$	**-6.978**	**-6.685**	-7.360	-7.279
MF-ARGARCH$^{(5_m,30_m)}$	**-6.978**	**-6.685**	**-7.366**	-7.271

MCS p-values				
	ALV	BMW	DTE	RWE
RGARCH	0.469	0.004	0.065	0.124
SMF-ARGARCH$^{(5_m,30_s)}$	0.039	0.000	0.000	0.036
SMF-ARGARCH$^{(5_m,1_m)}$	0.106	0.000	0.001	0.048
SMF-ARGARCH$^{(5_m,10_m)}$	0.458	0.012	0.054	0.116
SMF-ARGARCH$^{(5_m,15_m)}$	0.516	0.004	0.074	0.124
SMF-ARGARCH$^{(5_m,30_m)}$	0.516	0.005	0.141	0.116
MF-ARGARCH$^{(5_m,30_s)}$	0.178	0.000	0.004	0.124
MF-ARGARCH$^{(5_m,1_m)}$	0.633	0.002	0.020	0.131
MF-ARGARCH$^{(5_m,10_m)}$	0.516	0.551	0.908	1.000
MF-ARGARCH$^{(5_m,15_m)}$	0.775	0.551	0.242	0.374
MF-ARGARCH$^{(5_m,30_m)}$	1.000	1.000	1.000	0.131

The results point out that the MF-ARGARCH$^{(5_m,10_m)}$ is the only model which always falls into the MCS at the confidence level of 75%, followed by the MF-ARGARCH models based on the frequency combinations $(5_m, 15_m)$ and $(5_m, 30_m)$ entering the 75% MCS in all cases except one in which they are still included in the more conservative 90% set. In addition, for the BMW the only three models entering the set of superior models are of the MF-ARGARCH type, thus confirming the good predictive power of this class of models. Differently, the MF-ARGARCH models combining the 5 minutes realized volatility with higher frequencies are overall characterised by a lower predictive accuracy since they are only included in the 90% MCS for RWE and in 75% MCS for ALV. However, the analysis of this asset highlights a fairly similar predictive ability within of the whole set of models considered, failing to indicate a clear winner, even if the class of the MF-ARGARCH model (except for the $(5_m, 30_s)$ specification) shows the highest p-values.

On the other hand, the predictive performance of models including a single measurement equation, the SMF-ARGARCH, has not been found to be equally satisfactory. It can be noted that the SMF-ARGARCH$^{(5_m,30_s)}$ never enters the set of the superior models and the SMF-ARGARCH$^{(5_m,1_m)}$ enters just once at the 90% level for the stock ALV. The other specifications within this class of models fall into the MCS at the confidence level of 0.75 for ALV and at the confidence level of 0.90 for RWE, while for the asset DTE, only the SMF-ARGARCH$^{(5_m,30_m)}$ comes into the 90% MCS.

Finally, the standard RGARCH enters in the set of superior models just for the stock ALV (for which almost all models show a very similar forecast accuracy) at 75% level and for RWE at 90% level while, in the case of BMW and DTE, it is clearly outperformed by its mixed frequency counterparts.

3.7 Conclusion

In this paper we propose a flexible generalization of the Realized GARCH model of Hansen et al. (2012) where the conditional variance dynamics are driven by a weighted average of realized variances computed using intra-daily information observed at different frequencies. The coefficients of the weighted average are time varying and adaptively estimated in order to guarantee, in a fully data driven fashion, an optimal bias-variance trade-off. The proposed models can be used to generate improved conditional variance forecasts and, in an ex-post framework, to compute an optimized volatility measure that follows in the spirit the two time scales estimator of Zhang et al. (2005). Investigation of the properties of this measure and an empirical comparison with other commonly used realized measures is left for investigation in future research activity.

The in-sample empirical analysis shows that the SMF-ARGARCH and the MF-ARGARCH models, incorporating information from multiple realized volatility measures computed at different frequencies, lead to substantial improvements over the standard Realized GARCH model, in terms of a goodness of fit, measured according to the returns partial log-likelihood. The best results are achieved by models mixing the 5 minutes RV with realized measures based on lower frequency information.

An out-of-sample forecasting comparison shows that, compared to the standard Realized GARCH model, the mixed-frequency models can allow for substantial improvements in terms of forecasting accuracy as measured by the partial predictive log-likelihood and the QLIKE loss functions.

Chapter 4

Heterogeneous Component MEM models for forecasting trading volumes

4.1 Recent developments in the analysis of high frequency financial time series

Thanks to the rapid growth of information technology, in recent years high frequency data analysis has been continuously gaining importance. The availability of financial data recorded at very high frequencies has inspired the development of new types of econometric models to capture the specific properties of these observations. The most important characteristics of high-frequency data are strong serial dependencies, irregular spacing in time, price discreteness and intraday seasonal patterns. To model the dynamic behaviour of irregularly spaced transaction data, Engle and Russell (1998) proposed the Autoregressive Conditional Duration (ACD) model, later generalized in Multiplicative Error Model (MEM), by Engle (2002), as a general class of time series models for positive-valued random variables which are decomposed into the product of their conditional mean and a positive-valued i.i.d. error

term with unit mean. Discussions and extensions on the properties of this model class can be found in Chou (2005), Manganelli (2005), Cipollini et al. (2006), Lanne (2006), Brunetti and Lildholdt (2007) and Brownlees et al. (2011), among others.

The continuous development of new and more complex financial instruments, has considerably increased the necessity to understand the theoretical and empirical behaviour of non-negative valued processes, such as number of trades and volumes, high-low range, absolute returns, financial durations and volatility measures derived from ultra high-frequency data. It is well established that all these variables share the feature of clustering and high-persistence. The recurrent feature of long-range dependence is conventionally modelled as an autoregressive fractionally integrated moving average (ARFIMA) process as in Andersen et al. (2003), or using regression models mixing information at different frequencies such as the Heterogeneous AR (HAR) model of Corsi (2009). This has been extended by Andersen et al. (2007) inserting a volatility jump component for capturing the abrupt changes that characterise the realized volatility and by Ghysels et al. (2006) in a MIDAS framework.

The HAR model, inspired by the Heterogeneous Market Hypothesis of Müller et al. (1993) and the asymmetric propagation of information between long and short horizons, is formulated as a multicomponent volatility model with an additive hierarchical structure by specifying daily volatility as the sum of volatility components over different time horizons. Deo et al. (2006) discuss the specification and estimation of stochastic volatility models with long memory. McAleer and Medeiros (2008) extend the HAR-RV by proposing a flexible multiple regime smooth transition model to capture non-linearities and long-range dependence in the time series dynamics. Scharth and Medeiros (2009) in order to describe

the dynamics of realized volatilities of several DJIA stocks consider a regression tree model, incorporating past cumulated daily returns to account for regime switches. Raggi and Bordignon (2012) propose a single model which combines, both non-linearity effects, through a Markov switching process, and high persistence, through fractionally integrated dynamics. Groß-KlußMann and Hautsch (2013) introduce a long memory autoregressive conditional Poisson model to forecast bid-ask spreads, capturing salient features such as strong autocorrelation and discreteness of observations. Recently, Gallo and Otranto (2015) have employed a regime switching model with Markovian dynamics and a smooth transition non-linearity approach, within the class of Multiplicative Error Models, to capture the slow-moving long-run average level in realized volatility series.

Another class of models has recently received increasing attention from researchers and practitioners. This includes component models featuring two or more components moving at different frequencies. These models are able to parsimoniously characterise the rich dependence structure of financial variables such as volatility and volume, including their highly persistent dependences structures. It is worth noting that the HAR model itself can be represented as a component model aggregating volatility components moving at different frequencies.

Starting from the Spline-GARCH of Engle and Rangel (2008), where returns volatility is specified to be the product of a slow-moving component, represented by an exponential spline, and a short-run component which follows an unit GARCH process, several contributions have extended this idea. Brownlees and Gallo (2010) propose a dynamic model using a long-run component, based on some linear basis expansion of time, bounded with a penalized maximum likelihood estimation

strategy. Engle et al. (2013) introduce a new class of models called GARCH-MIDAS, since it is structured as a mean reverting unit daily GARCH process and a MIDAS polynomial filter which applies to monthly, quarterly, or biannual financial or macroeconomic variables. Amado and Teräsvirta (2013) decompose the variance into a conditional and an unconditional component such that the latter evolves smoothly over time through a linear combination of logistic transition functions with time as the transition variable. Moving to the analysis of intra-daily data, Engle and Sokalska (2012) develop the Multiplicative Component GARCH, decomposing the volatility of high-frequency asset returns into product of three components, namely the conditional variance is a product of daily, diurnal, and stochastic intraday components. Brownlees et al. (2011) propose a component model with a dynamic specification to capture salient features of intra-daily volumes such as high-persistence, asymmetry and intra-daily periodicity.

4.2 A new model for forecasting high-frequency trading volumes

This paper proposes and investigates a dynamic component model for high-frequency trading volumes. These are typically characterised by two prominent features: a slowly moving long-run level and a highly persistent autocorrelation structure that has often been found to be long memory. So, as a novel approach for modelling and forecasting intra-daily volumes, we propose the Heterogeneous MIDAS Component Multiplicative Error Model (H-MIDAS-CMEM), which is an extension of the class of Multiplicative Error Models for non-negative time series by Engle (2002) and Engle and Gallo (2006) and dynamic models with slowly moving components in the spirit of Engle et al. (2013) and Corsi (2009).

The empirical characteristics of intra-daily volumes led us to adopt a model that is based on components that move at different frequencies, since it is possible to distinguish a lower frequency component that moves around the overall series, an intra-daily periodic component which has a regular U-shape pattern and a distinctive intra-daily dynamic non-periodic component. These components are differently modelled, since the periodic component is treated employing a Fourier Flexible Form, the short-run component follows a unit mean reverting GARCH-type process and the long-run component is based on a MIDAS polynomial structure through an additive cascade of linear filters adopting heterogeneous components which can take on multiple frequencies, even different from the usual ones applied in the HAR framework. From an economic point of view, this cascade structure reproduces the natural heterogeneity of the different categories of agents operating in the market at different frequencies. This results in a variety of sources separately affecting the variation of the average volume at different speeds. On a statistical ground, the cascade structure has the advantage of increasing model's flexibility since it allows to separately parametrize the dynamic contribution of each of these sources. Namely, we consider a linear combination of MIDAS filters of volumes aggregated over different time intervals ranging from one hour to one day. Doing so we also allow to account for the apparent long memory behaviour of high-frequency volumes.

This parametrization is flexible enough to reproduce smooth changes in the long level of volumes. Nevertheless, it fails to fully capture another empirical regularity of high-frequency trading volumes which is the sudden change from states of low trading intensity to states characterised by a very high average trading intensity. To this purpose we introduce in the trend specification a time-varying intercept modelled as a weighted average of two different values associated to low and high trading intensity

states, respectively. The weights of the combination can be interpreted as related to the probability of staying in each of these states. They are assumed to be time-varying as a logistic function of past observed state variables.

4.3 Relations with the existing literature

Relatively to the link with previous research in this field, the H-MIDAS-CMEM extend the logic of the Component Multiplicative Error Model proposed by Brownlees et al. (2011), by considering a more flexible parametrization of the long-run component. This work is also related to the literature on algorithmic trading (see Madhavan (2002), Białkowski et al. (2008) and Brownlees et al. (2011) among others), which is an area of increasing importance in financial exchanges. Volumes are a crucial ingredient for the volume-weighted average price (VWAP) trading strategy, which is one of the most common benchmarks used by institutional investors for judging the execution quality of individual stocks. The VWAP of a stock over a particular time horizon (usually one day) is simply given by the total traded value divided by the total traded volume during that period, i.e. the price of each transaction is weighted by the corresponding traded volume. The aim of using a VWAP trading target is to minimize the price impact of a given order by slicing it into smaller transaction sizes, reducing, in this way, the difference between expected price of a trade and its actual traded price. Investors, spreading the timing of transactions throughout the day, seek to achieve an average execution price as close as possible to the VWAP in order to lower market impact costs. Therefore, in this context, the key for a good strategy relies on accurate predictions of intra-daily volumes, since prices are substantially unpredictable.

4.4 Empirical analysis

In order to assess the relative merits of the proposed approach we present the results of an empirical analysis. This is carried out on high-frequency trading volume data from January 2003 to December 2007 for three stocks traded on the Xetra Market in the German Stock Exchange, which are characterised by different liquidity levels according to the number of non trading intra-daily intervals. We filter the raw data using the procedure proposed in Brownlees and Gallo (2006), deleting transactions that occurred outside regular trading hours from 9:00 am to 5:30 pm. The data are aggregated by computing cumulated trading volumes over 10-minutes intervals, resulting in 51 observations per day. Estimation is performed by the method of maximum likelihood under the assumption that the innovation are distributed according to a Zero-Augmented Generalized F distribution by Hautsch et al. (2013). The reason for this choice is twofold. First, it delivers a flexible probabilistic model for the conditional distribution of volumes. Second, it allows to control for the relevant proportion of zeros present in our data. We show that the H-MIDAS-CMEM model captures the salient empirical features of high-frequency volumes such as strong serial dependencies, clusters of trading activity and intra-daily periodic patterns. The evaluation of forecasting performance, provided by the Model Confidence Set (MCS) procedure proposed by Hansen et al. (2011), shows that the H-MIDAS-CMEM outperforms several competing models in minimizing the Slicing Loss function developed by Brownlees et al. (2011) and the usual Mean Squared Prediction Error, since it is the only model always included in the set of superior models at different confidence levels.

The remainder of this chapter is organized as follows. Section 4.5 describes the H-MIDAS-CMEM model. In Section 4.6 we present the estimation

procedure. Section 4.7 describes the data characteristics (4.7.1) and reports in-sample estimation results and diagnostics (4.7.2). Section 4.8 compares the predictive performance of competing models employing the MCS procedure and different loss functions. Section 4.9 concludes.

4.5 Model formulation

Let $\{x_{t,i}\}$ be a time series of intra-daily trading volumes. We denote days by $t \in \{1, \ldots, T\}$, where each day is divided into I equally spaced intervals indexed by $i \in \{1, \ldots, I\}$, then the total number of observations is given by $N = T \times I$.

The empirical regularities of high persistence and clustering of trading activity about intra-daily volumes lead us to build a Multiplicative Error Model consisting of multiple components that move at different frequencies. Extending the logic of the Component Multiplicative Error Model (CMEM) by Brownlees et al. (2011) and MIDAS regression, we propose the H-MIDAS-CMEM which is formulated as

$$x_{t,i} = \tau_t \, g_{t,i} \, \phi_i \, \varepsilon_{t,i}. \tag{4.1}$$

The multiplicative innovation term $\varepsilon_{t,i}$ is assumed to be i.i.d., non-negative and to have unit mean and constant variance σ^2, that is

$$\varepsilon_{t,i} | \mathcal{F}_{t,i-1} \sim \mathcal{D}^{+}(1, \sigma^2) \tag{4.2}$$

where $\mathcal{F}_{t,i-1}$ is the sigma-field generated by the available information until interval $i-1$ of day t. Then, the expectation of $x_{t,i}$, given the information set $\mathcal{F}_{t,i-1}$, is the product of:

ϕ_i is an *intra-daily periodic component* parametrized by a Fourier Flexible Form, which reproduces the approximately U-shaped intra-daily seasonal pattern typically characterising trading activity;

$g_{t,i}$ is an *intra-daily dynamic non-periodic component*, based on a unit mean reverting GARCH-type process, that reproduces autocorrelated and persistent movements around the current long-run level;

τ_t is a *lower frequency component* given by the sum of MIDAS filters defined at different frequencies. This component is designed to track the dynamics of the long-run level of trading volumes. A time-varying intercept allows to reproduce sudden switches from very low to high trading intensity periods.

4.5.1 Intra-daily periodic component

Intra-daily volumes usually exhibit a U-shaped daily seasonal pattern, i.e. the trading activity is higher at the beginning and at the end of the day than around lunch time. To account for the periodic intraday factor we divide volumes $x_{t,i}$ by a seasonal component ϕ_i that is specified via a Fourier Flexible Form as proposed by Gallant (1981)

$$\phi_i = \sum_{q=0}^{Q} a_{0,q}\, \iota^q + \sum_{p=1}^{P} [a_{c,p}\, cos(2\pi p\, \iota) + a_{s,p}\, sin(2\pi p\, \iota)] \tag{4.3}$$

where $\iota = i/I \in (0, 1]$ is a normalized intraday time trend.

Andersen et al. (2000) suggest that the Fourier terms in (4.3) do not add any significant information for $Q > 2$ and $P > 6$, so the model precision by using $Q = 2$ and $P = 6$ is enough to capture the behaviour of the intra-day

periodicities[1].

Assuming a multiplicative impact of intra-day periodicity effects, diurnally adjusted trading volumes are computed as

$$y_{t,i} = \frac{x_{t,i}}{\phi_i}. \tag{4.4}$$

4.5.2 Intra-daily dynamic non-periodic component

The intra-daily non-periodic component, unlike the seasonal component, takes a distinctive and non-regular dynamic. In order to make the model identifiable the intra-daily dynamic component follows a unit mean reverting GARCH-type process, namely $g_{t,i}$ has unconditional expectation equal to 1.

The short-run component, in its simplest form, is formulated as

$$g_{t,i} = \omega^* + \alpha_1 \frac{y_{t,i-1}}{\tau_t} + \alpha_0 I(y_{t,i-1} = 0) + \beta_1 g_{t,i-1}, \tag{4.5}$$

where $\omega^* = (1 - \alpha_1 - (1 - \pi)\alpha_0 - \beta_1)$, π is the probability that $y_{t,i} > 0$ and $I(\bullet)$ denotes an indicator function which is equal to 1 if the argument is true and to 0 otherwise. The dummy variable in equation (4.5) allows us to better capture the liquidity dynamics, since we have the following two different processes:

$$\begin{cases} \forall\, y_{t,i-1} > 0, \quad g_{t,i} = \omega^* + \alpha_1 \frac{y_{t,i-1}}{\tau_t} + \beta_1 g_{t,i-1} \\ \\ \forall\, y_{t,i-1} = 0, \quad g_{t,i} = \omega^* + \alpha_0 + \beta_1 g_{t,i-1} = \omega_0^* + \beta_1 g_{t,i-1} \quad \text{where } \omega_0^* = \omega^* + \alpha_0. \end{cases}$$

[1]This result is confirmed by computing the Bayesian Information Criterion (BIC), not reported in the paper, for the estimation lags.

The additional intercept α_0 aims at making the specifications of the non-periodic short-run component more flexible since it allows to differently characterise the dynamics of the trading volumes depending on the presence of trading activity at time $t-1$. It is immediate to see that the hypothesis of homogeneous dynamics corresponding to $\alpha_0 = 0$ can be easily tested by means a simple Wald statistic. A similar formulation is adopted by Hautsch et al. (2013). Consequently, the task of the indicator function $I(y_{t,i-1} = 0)$ is to adjust the intra-daily dynamic non-periodic component intensity when volume switches from trading to non-trading intervals.

4.5.3 The low frequency component

The low frequency component is modelled as a linear combination of MIDAS filters of past volumes aggregated at different frequencies. A relevant issue is related to the identification of the frequency of the information to be used by the filters, that notoriously acts a smoothing parameter. The simplest choice would be to use information observed at a daily scale. Then, considering volumes aggregated at daily frequency, the trend component τ_t is given by

$$log\,\tau_t = m_d + \theta_d \sum_{k=1}^{K_d} \varphi_k(\omega_{1,d}, \omega_{2,d}) YD_{t-k} \tag{4.6}$$

where $YD_t = \sum_{i=1}^{I} y_{t,i}$, denotes the daily cumulative volume, whereas the subscript d means that the parameters refer to a daily frequency. The long-run component is considered in terms of logarithmic specification since it does not require parameter constraints to ensure the positivity of τ_t.

A common choice for determining $\varphi_k(\boldsymbol{\omega})$ is the Beta weighting scheme

$$\varphi_k(\boldsymbol{\omega}) = \frac{(k/K)^{\omega_1-1}(1-k/K)^{\omega_2-1}}{\sum_{j=1}^{K}(j/K)^{\omega_1-1}(1-j/K)^{\omega_2-1}} \tag{4.7}$$

where the weights in equation (4.7) sum up to 1. As discussed in Ghysels et al. (2007), this Beta-specification is very flexible, being able to accommodate increasing, decreasing or hump-shaped weighting schemes. The Beta lag structure in (4.7) includes two parameters, but in our empirical applications ω_1 is always setted equal to 1 such that the weights are monotonically decreasing over the lags. As in Engle and Rangel (2008), the number of lags K is properly chosen by information criteria to avoid overfitting problems.

An alternative specification is to use higher frequency volumes aggregated over intervals of length equal to $1/H$ days

$$log\,\tau_{t/H} = m_h + \theta_h \sum_{k=1}^{K_d}\sum_{j=1}^{H} \varphi_{[j+(k-1)H]}(\omega_{1,h},\omega_{2,h}) Y H_{t-k}^{(H-j+1)} \tag{4.8}$$

where $t/H \in \{1,\ldots,H\times T\}$ denotes the hourly frequency, with H the number of intervals in which the day is divided, while the subscript h refers to the parameters corresponding to the hourly frequency. The variable $YH_t^{(j)}$ corresponds to the (j)-th hourly cumulative volume of the day t, where

$$YH_t^{(j)} = \sum_{i=I\frac{(j-1)}{H}+1}^{I\frac{j}{H}} y_{t,i} \qquad j=1,\ldots,H. \tag{4.9}$$

Accordingly, for example, by setting $H=3$ we have a long-run component

that changes every 2:50 hours, while for $H = 17$ the trend moves with a 30 minutes frequency.

A more general formulation of the long-run component encompassing the previous two is then given by

$$\begin{aligned} log\,\tau_t = m &+ \theta_d \sum_{k=1}^{K_d} \varphi_k(\omega_{1,d}, \omega_{2,d}) Y D_{t-k} \\ &+ \theta_h \sum_{k=1}^{K_d} \sum_{j=1}^{H} \varphi_{[j+(k-1)H]}(\omega_{1,h}, \omega_{2,h}) Y H_{t-k}^{(H-j+1)}. \end{aligned} \tag{4.10}$$

This multiple frequencies specification appear to be preferable to the previous single-frequency models in (4.6) and (4.8) for three different reasons. First, the modeller is not bound to choose a specific frequency for trend estimation, but can determine the optimal blend of low and high frequency information in a fully data driven fashion. Second, it is compatible with the heterogeneous market assumption of Müller et al. (1993), enforcing the idea that market agents can be divided in a different groups characterised by different interest and strategies. Third, as pointed out in Corsi (2009), an additive cascade of linear filters moving at different frequencies allows to reproduce very persistent dynamics such as those typically observed for high-frequency trading volumes. These have indeed been frequently found to be observationally equivalent to long-memory processes.

Remark:
A further generalization of the proposed approach could be to augment the model in (4.10) with volumes cumulated over intervals of length greater than one day. For example, the trend component might follow

a specification that moves on a monthly basis, namely

$$\begin{aligned} log\,\tau_{t\times M} = m &+ \theta_m \sum_{k=1}^{K_m} \varphi_k(\omega_{1,m},\omega_{2,m}) YM_{(t\times M)-k} \\ &+ \theta_w \sum_{k=1}^{K_m}\sum_{j=1}^{W} \varphi_{[j+(k-1)W]}(\omega_{1,w},\omega_{2,w}) YW^{(W-j+1)}_{(t\times M)-k} \\ &+ \theta_d \sum_{k=1}^{K_m}\sum_{j=1}^{WD} \varphi_{[j+(k-1)(WD)]}(\omega_{1,d},\omega_{2,d}) YD^{((WD)-j+1)}_{(t\times M)-k} \end{aligned} \tag{4.11}$$

where $t \times M$ denotes the monthly frequency and $YM_{t\times M}$ is the monthly volume, while $YW^{(j)}_{t\times M}$ and $YD^{(j)}_{t\times M}$ are the (j)-th weekly and daily volume of the month $t \times M$ respectively. This specification involves a multiple weighting scheme, since the weight $\omega_{2,\cdot}$ is not the same for each component, as it can be noted from the different subscript. Furthermore, K_m denotes the number of MIDAS lags based on the monthly frequency, whereas the index W represents the number of weeks in a month and D the number of trading days in a week. Therefore, the long-run component can potentially include multiple heterogeneous components, employing variables with a lower or a higher frequency than the ones used in equation (4.11).

The clustering of the trading activity involves a continuous variation of the average volume level, since periods in which the stocks are very traded are alternated with periods in which the stocks are less traded. This is even more evident when there is a sequence of zero observations which precedes or follows high trading periods. In other words, the dynamics of trading volumes are typically characterised by sudden transitions from states of very low trading activity to state of intense trading. In order to account for this switching-state behaviour we further extend the proposed modelling approach introducing a time-varying intercept in the specification of the

long-run component. This is specified as a convex combination of two different unknown parameters m_1 and m_2, that is

$$m_t = \lambda_t \, m_1 + (1 - \lambda_t) \, m_2. \tag{4.12}$$

The combination weights are time-varying, since they change as a function of observable state-variables. The weight function λ_t follows a logistic specification of the type

$$\lambda_t = \frac{1}{1 + exp(\gamma(\delta - s_{t-1}))}, \qquad (\gamma, \delta) > 0 \tag{4.13}$$

where γ and δ are unknown coefficients and s_{t-1} is an appropriately chosen state-variable. The γ parameter determines the transition speed from one state to the other one, since the higher γ, the faster the switching. The constraint $\gamma > 0$ allows us to associate the parameter m_1 with high volume levels. In fact, when s_{t-1} goes to infinity, the weight λ_t tends to 1, excluding the second component m_2, since the corresponding weight $(1 - \lambda_t)$ tends to zero. Similarly, when s_{t-1} goes to zero, λ_t tends to its minimum value $(1 + exp(\gamma\delta))^{-1}$.

A suitable choice for the state-variable is the daily average of intra-daily volumes $\bar{y}_t$. Assuming that $E(g_{t,i}|\mathcal{F}_{t-1})$ is equal to its unconditional expectation ($E(g_{t,i}|\mathcal{F}_{t-1}) = 1$), it represents an unbiased proxy of the daily component τ_t. Since $\bar{y}_t$ is always different from zero, as we exclude non-trading days, we introduce the further restriction $\delta > 0$, given that δ can be interpreted as the threshold around which the level of the volume switches from one regime to the other one.

4.6 Inference

Multiplicative Error Models are usually estimated by QMLE, assuming that the density of the innovation term follows a Gamma distribution or one of its generalizations or special cases. However, in the case of non-liquid assets these distributions cannot account for a point mass at zero, as the corresponding log-likelihood functions exclude zero realizations, with the exception of the Exponential distribution that is obtained as particular case of the Gamma.

The specification of an Exponential distribution for $\varepsilon_{t,i}$ is a natural choice in a MEM context, as the Exponential distribution can be seen as the counterpart of the Normal distribution for positive-valued random variables. Under the assumption of correct specification of the conditional mean function, the maximization of the Exponential quasi log-likelihood function leads to consistent and asymptotically normal estimates of the conditional mean parameters that can be interpreted as QML estimates. This result can be formalized from Engle and Russell (1998) and Engle (2002) exploiting the results of Lee and Hansen (1994) for GARCH models.

Engle and Gallo (2006) show that the consistency of the Exponential QML-estimator can be alternatively achieved by choosing a Gamma distribution, since the first order conditions for the conditional mean parameters are the same for the two estimators. Consequently, the estimator of the conditional mean parameters obtained employing the Gamma distribution is identical to the one reached from the exponential QMLE.

Assuming a *Gamma(a,b)* distribution for the innovation term $\varepsilon_{t,i}$, it follows that

$$f(\varepsilon_{t,i}|\mathcal{F}_{t,i-1}) = \frac{1}{\Gamma(a)\, b^a} \varepsilon_{t,i}^{a-1}\, exp\left(-\frac{\varepsilon_{t,i}}{b}\right)$$

where $a, b > 0$. However, to ensure that $\varepsilon_{t,i}$ has unit mean the restriction $b = \frac{1}{a}$ holds

$$f(\varepsilon_{t,i}|\mathcal{F}_{t,i-1}) = \frac{1}{\Gamma(a)}\, a^a\, \varepsilon_{t,i}^{a-1}\, exp\left(-a\, \varepsilon_{t,i}\right).$$

Therefore, by the properties of the Gamma distribution we have

$$f(y_{t,i}|\mathcal{F}_{t,i-1}) = \frac{1}{\Gamma(a)}\, a^a\, y_{t,i}^{a-1} \mu_{t,i}^{-a}\, exp\left(-a\frac{y_{t,i}}{\mu_{t,i}}\right), \quad \text{where } \mu_{t,i} = \tau_t\, g_{t,i}.$$

Accordingly, the process $y_{t,i}$ have conditional expectation $E[y_{t,i}|\mathcal{F}_{t,i-1}] = \mu_{t,i}$ and conditional variance $var[y_{t,i}|\mathcal{F}_{t,i-1}] = \mu_{t,i}^2/a$.

However, if the data are affected by exact zero values, some problems arise. The Gamma distribution cannot be defined for $a < 1$ when zero observations occur and, on the other hand, the first condition order for a is not fulfilled for $a \geq 1$ for exact zero values[2]. Therefore, the Maximum Likelihood estimation of a is not feasible in the case in which the data are characterised by zero outcomes. To avoid this problem Engle (2002) assume that $\varepsilon_{t,i}|\mathcal{F}_{t,i-1} \sim Exponential(1)$, which is equivalent to the Gamma density by setting $a = 1$.

The continuous nature of the Exponential distribution implies that the proportion of zeros must be trivial to avoid misspecification at the lower boundary of the support. In presence of zero observations the Generalized Method of Moments (GMM) can be a valid alternative estimation strategy, as discussed in Brownlees et al. (2011), since it does not need the adoption of a specific density function for the innovation term. Exponential-QML and GMM provide consistent estimates of conditional mean parameters, but these become quite inefficient in the presence of a high proportion of zeros, in both cases. To address this problem Hautsch et al. (2013) propose

[2]See Cipollini et al. (2006) for further details.

an alternative estimation strategy based on the introduction of what they call a Zero-Augmented Generalized F (ZAF) distribution. Their results provide evidence that in the presence of a non-trivial proportion of zero outcomes MLE by the ZAF distribution allows to overcome the potential inconsistency of the standard QMLE and, in any case, to obtain substantial efficiency gains over the latter.

As proposed in Hautsch et al. (2013), a ZAF distribution is applied to a non-negative random variable Z with a high proportion of zero outcomes, assigning a discrete probability mass at the exact zero value as follows

$$\pi = P(Z > 0), \quad (1 - \pi) = P(Z = 0). \tag{4.14}$$

Suppose that when $Z > 0$, Z follows a Generalized F distribution with density function

$$g(z; \boldsymbol{\zeta}) = \frac{a z^{ab-1} [c + (z/\nu)^a]^{(-c-b)} \, c^c}{\nu^{ab} \, \mathcal{B}(b, c)} \tag{4.15}$$

where $\boldsymbol{\zeta} = (a, b, c, \nu)$, $a > 0$, $b > 0$, $c > 0$ and $\nu > 0$. $\mathcal{B}(\cdot, \cdot)$ is the Beta function with $\mathcal{B}(b, c) = [\Gamma(b)\Gamma(c)]/\Gamma(b + c)$. The Generalized F distribution is based on a scale parameter ν and three shape parameters a, b and c, thus it is very flexible, nesting different error distributions, such as Weibull for $b = 1$ and $c \to \infty$, Generalized Gamma for $c \to \infty$ and Log-Logistic for $b = 1$ and $c = 1$.

Then the ZAF distribution is semi-continuous with the density function

$$f_Z(z) = (1 - \pi)\eta(z) + \pi g(z) I_{(z>0)} \tag{4.16}$$

where $0 \leq \pi \leq 1$, $\eta(z)$ is a point probability mass at $z = 0$ and $I_{(z>0)}$ denotes an indicator function taking the value 1 for $z > 0$ and 0 else. Can

be easily noted that the ZAF density reduces to the Generalized F for $\pi = 1$.

The moments of the ZAF distribution are given by

$$\begin{aligned} E[Z^r] &= \pi E[Z^r|Z>0] + (1-\pi)E[Z^r|Z=0] \\ &= \pi \nu^r c^{r/a} \frac{\Gamma(b+r/a)\Gamma(c-r/a)}{\Gamma(b)\Gamma(c)}, \quad b < ac. \end{aligned} \tag{4.17}$$

The MEM structure in (4.1) implies that $\nu = (\pi\xi)^{-1}$ to ensure that the unit mean assumption for ε_t is fulfilled and

$$\xi = c^{1/a} \left[\Gamma(b+1/a)\Gamma(c-1/a)\right] \left[\Gamma(b)\Gamma(c)\right]^{-1}. \tag{4.18}$$

Considering seasonally adjusted volume $y_{t,i}$, the log-likelihood function based on the density in (4.16) is given by

$$\begin{aligned} \mathcal{L}(\boldsymbol{y};\boldsymbol{\psi},\pi) = n_z\, log(1-\pi) + n_{nz}\, log\,\pi + \sum_{i\in\mathcal{J}_{n_{nz}}} \Bigg\{ log\, a + (ab-1)\, log\left(\frac{y_{t,i}}{\tau_t\, g_{t,i}}\right) + c\, log\, c \\ -(c+b)\, log\left[c + \left(\frac{\pi\xi y_{t,i}}{\tau_t\, g_{t,i}}\right)^a\right] - log(\tau_t\, g_{t,i}) - log\,\mathcal{B}(b,c) + ab\, log(\pi\xi)\Bigg\} \end{aligned} \tag{4.19}$$

where $\mathcal{J}_{n_{nz}}$ denotes the set of all observations different from zero, while n_z and n_{nz} denote the number of zero and non-zero observations respectively.

4.7 Empirical application

4.7.1 Data description

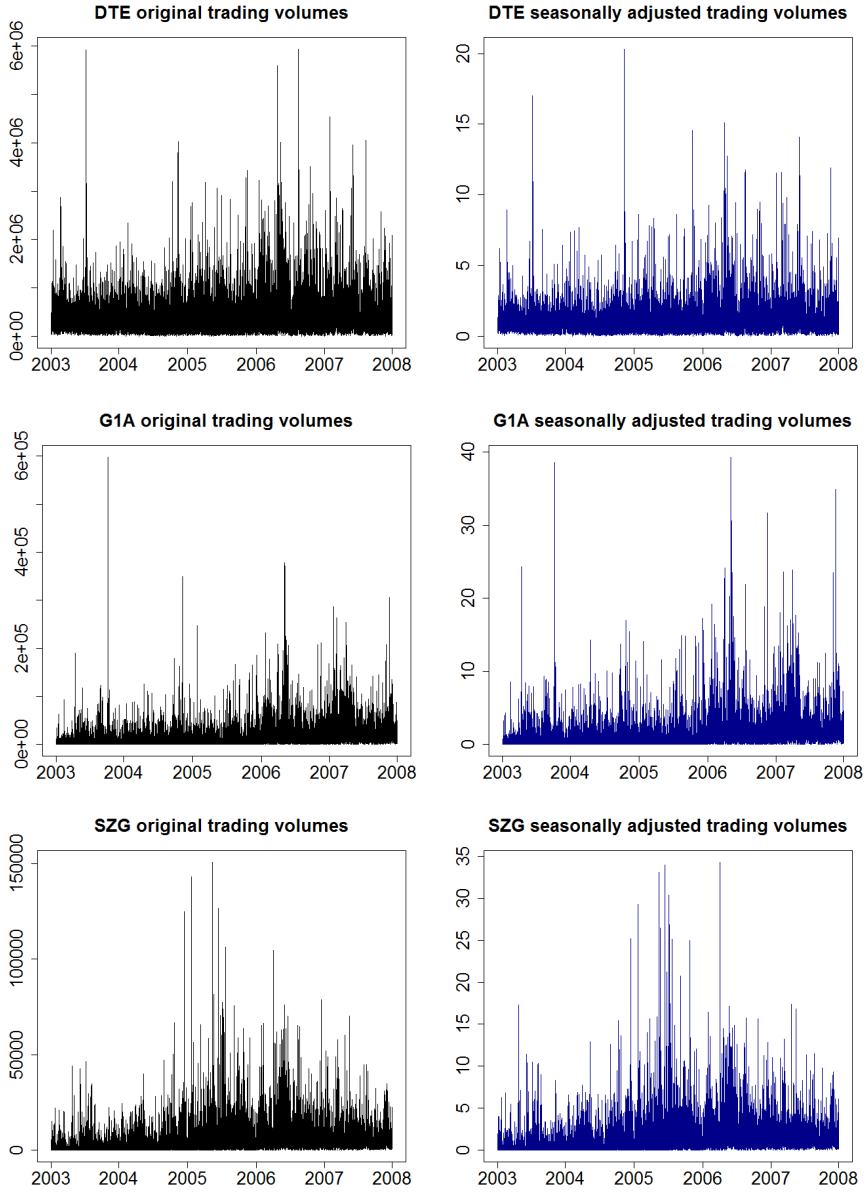
High-frequency trading volume data used in our analysis refer to the stocks Deutsche Telekom (DTE), GEA Group (G1A) and Salzgitter (SZG) traded on the Xetra Market in the German Stock Exchange. The raw data have been filtered employing the procedure proposed by Brownlees

and Gallo (2006), only considering regular trading hours from 9:00 am to 5:30 pm. Tick-by-tick data are aggregated computing intra-daily volumes over 10-minutes intervals, which means 51 observations per day. The empirical analysis covers the period between January 2003 and December 2007 which includes 1269 trading days, corresponding to 64719 intra-daily observations for each stock.

In order to capture some potential trading behaviour, the data have been seasonally adjusted using the Fourier Flexible Form described in equation (4.3). Figure 4.1 depicts the overall volume time series (left panel) and the seasonally adjusted trading volumes (right panel) of the three analysed assets in our empirical application for the period that goes from January 2003 to December 2007, whereas Figure 4.2 shows the intraday seasonal pattern estimated via the Fourier Flexible Form defined in equation (4.3). For DTE and SZG the trading intensity is higher near the open and just prior to the close, since time between trades, or durations, tend to be shortest near the open and the close of the trading day than the middle of the day, as documented in Engle and Russell (1998). The stock G1A, unlike the others, exhibits a lower trading intensity not only around lunch time, but also in the first part of the trading day.

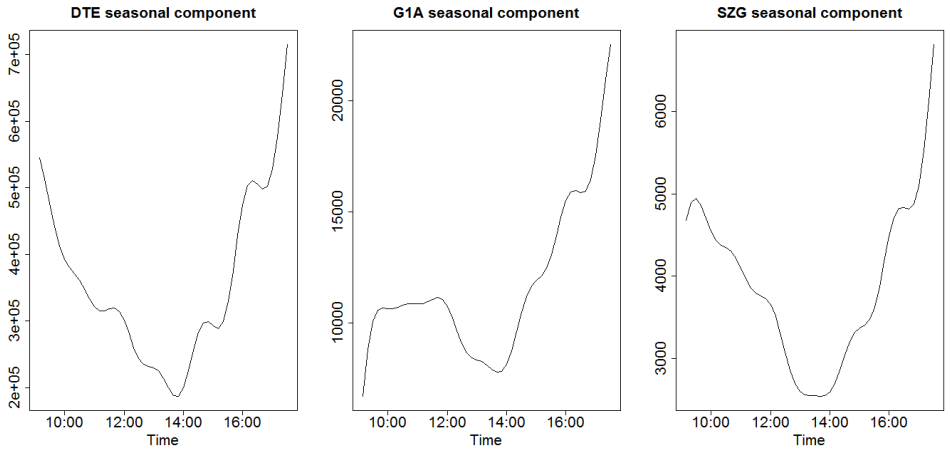
Descriptive statistics of seasonally adjusted 10-minutes trading volumes and of original 10-minutes trading volumes, whose values are reported in parentheses, of the three analysed stocks are shown in Table 4.1. An important feature of the data is the different number of zeros induced by non-trading intervals, i.e. the summary statistics point out that the proportion of zero observations are ranging from 0.03% to 15.78%, for DTE and SZG respectively. Being a relatively calm market period, the share of zeros is largely related to the regular trading activity which characterises each stock, rather than to turbulence caused by external shocks.

Figure 4.1: Original trading volume and seasonally adjusted trading volume for the three analysed stocks



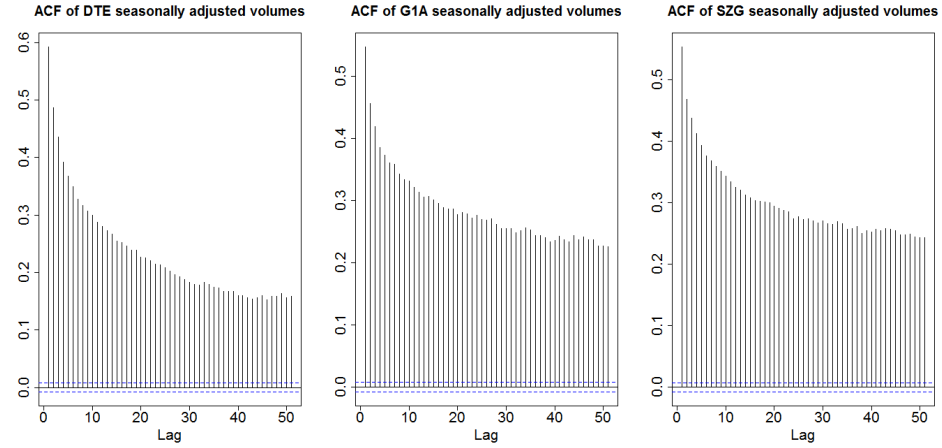
Left panel: original trading volumes for the period January 2003 – December 2007 of the stocks DTE, G1A and SZG. **Right panel:** seasonally adjusted trading volumes by using the Fourier Flexible Form specified in equation (4.3).

Figure 4.2: Intraday seasonal pattern



Intra-daily periodic component for the trading hour 9:00 am - 5:30 pm, estimated employing the Fourier Flexible Form specified in equation (4.3), for the stocks DTE, G1A and SZG.

Figure 4.3: Autocorrelation Function of seasonally adjusted volumes



Autocorrelation Function of seasonally adjusted 10-minutes trading volumes. The charts are sorted from the most liquid to the less liquid stock.

Table 4.1: Summary statistics for trading volumes

Ticker	Zero%	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.	$\hat{\rho}_{(1)}$	$\hat{\rho}_{(51)}$
DTE	0.031	0.000	0.505	0.792	1.000	1.229	20.310	0.829	0.592	0.159
		(0.000)	(159200)	(275300)	(358900)	(458200)	(5940000)	(310842.40)	(0.662)	(0.318)
G1A	7.046	0.000	0.180	0.544	1.000	1.274	39.300	1.443	0.547	0.226
		(0.000)	(2000)	(6113)	(11680)	(14770)	(598200)	(16960.19)	(0.573)	(0.280)
SZG	15.779	0.000	0.124	0.586	1.000	1.320	34.320	1.420	0.554	0.243
		(0.000)	(488)	(2162)	(3917)	(5110)	(150700)	(5686.33)	(0.585)	(0.288)

Summary statistics of seasonally adjusted 10 minutes trading volumes and of original trading volumes, where the latter are reported in parentheses. Sample period: January 2003 – December 2007. Zero%: Percentage of zero observations; Min.: Minimum; 1st Qu.: First Quartile; Median; Mean; 3rd Qu.: Third Quartile; Max.: Maximum; Std. Dev.: Standard Deviation; $\hat{\rho}_{(1)}$: Autocorrelation at the lag 1; $\hat{\rho}_{(51)}$: Autocorrelation at the lag 51 (1 day).

A further feature of high-frequency volumes is the strong serial dependence. As shown in Figure 4.3, the autocorrelation function of the three series appears to be very persistent. Also the decay pattern at high lags seems to be much lower than what implied by the assumption of exponential decay. Interestingly, the autocorrelations of the most liquid stocks decay more rapidly than the autocorrelations of the less liquid stocks, that is the latter are apparently more persistent. This higher persistence is probably driven by the presence of long sequence of zero values in the time series of volumes observed for less liquid stocks. The magnitude of the autocorrelations for the different series is slightly reduced dividing the original volumes by the intraday seasonal component, as documented in the last two columns of Table 4.1, reporting the value of the autocorrelation at the lag 1 and lag 51 (1 day) of seasonally adjusted and of original trading volumes.

Looking at Figure 4.1 we observe that high-frequency volumes, as most financial time series, are characterised by the presence of clusters trading activity, i.e. large (small) change in volumes following on large (small) change in volumes. A quantitative manifestation of this stylized fact is that the autocorrelation between $x_{t,i}$ and $x_{t,i-j}$ is greater than zero for j which ranges from a few minutes to several weeks. This suggests that the mean level around which the high-frequency volumes fluctuate is likely to be not constant, but smoothly time-varying.

4.7.2 In sample estimation results and model diagnostics

In this section we compare several models to analyse the liquidity dynamics of the stocks DTE, G1A and SZG. In particular, we investigate the features of intra-daily volumes presenting estimation results and diagnostics about the full sample, comparing the H-MIDAS-CMEM model against a number of alternative specifications. For each model we chose the best specification selecting the number of lags involved in the trend filter and in the short-run component by BIC and changing the starting parameters recursively in order to maximize the log-likelihood value. To achieve the optimal long-run frequency for trend estimation, we tested a lot of specifications by using filters corresponding to different time-horizons. The number of filters has been selected excluding the component for which the slope parameter θ was not significant and then the best trend structure was chosen according to BIC.

The examined models follow the GARCH-type process specified in (4.5) for the intra-daily dynamic component, while the long-run component is defined according to five specifications. Of these three make use of MIDAS filters and are given by the MIDAS-MEM in (4.6), the H-MIDAS-CMEM in (4.10) and a variant of the latter characterised by a time-varying intercept

(H-MIDAS-CMEM-TVI). The remaining two trend models are given by the CMEM of Brownlees et al. (2011), with

$$\tau_t = m + \alpha_{1,d}\,\bar{y}_{t-1} + \beta_{1,d}\,\tau_{t-1}$$

and by the HAR-MEM, with

$$\tau_t = m + \beta_{1,d}\,\bar{y}_{t-1} + \beta_{1,w}\,\bar{y}_{t-1:t-5} + \beta_{1,m}\,\bar{y}_{t-1:t-22}$$

where $\bar{y}_t = I^{-1}\sum_{i=1}^{I} y_{t,i}$, $\bar{y}_{t-1:t-5} = \sum_{j=1}^{5}\bar{y}_{t-j}$ and $\bar{y}_{t-1:t-22} = \sum_{j=1}^{22}\bar{y}_{t-j}$ are the daily, weekly and monthly average volumes, respectively.

Table 4.2 reports parameter estimates for the five models analysed, distinguishing the parameters of the intra-daily component, of the trend component and of the ZAF distribution, where all the parameters are statistically significant at 5%.

The panel of the intra-daily component parameters shows that the coefficient α_1 takes always the lowest value for the H-MIDAS-CMEM and in particular for the H-MIDAS-CMEM-TVI, whereas it presents more similar values among other models. The parameter α_0 is negative for the most liquid stock DTE, and positive for the stocks G1A and SZG, which are less liquid. In particular α_0 gets the minimum value for DTE and the maximum value for G1A and SZG, with reference to H-MIDAS-CMEM-TVI. For each stock the intra-daily dynamic non-periodic component is specified employing two GARCH lags related to the coefficient β_1 and β_2, except for the stock SZG which needs only one lag for the H-MIDAS-CMEM-TVI. Interestingly, the H-MIDAS-CMEM and its time-varying specification provide the lowest values in terms of persistence, which is given by $\varrho_g = \sum_{i=1}^{p}\alpha_i + \sum_{j=1}^{q}\beta_j$, and the highest values of the intercept ω^*.

This also holds for the HAR-MEM when it is compared to MEM, CMEM and MIDAS-MEM.

The second panel of Table 4.2 displays the estimated values of the trend parameters. Regarding the models with a long-run component which follow a MIDAS specification, the number of optimal lags K is equal to 300 days for all stocks analysed[3]. The MIDAS-MEM model with daily filter shows some regularity, since the slope parameter θ_d is always positive and takes approximately the same intensity for the different stocks. Also the weight ω_d seems to be very similar among the assets. The models that imply a Heterogeneous-MIDAS specification of the trend component exhibit, for each analysed asset, slope parameters θ_d and θ_h with opposite sign. In the H-MIDAS-CMEM the weights ω_d and ω_h present values not so far from each other, while in its time-varying specification there is a greater discrepancy, with respect to individual stocks. Regarding the coefficient of the time-varying intercept, γ rises as the liquidity decreases. Then, the higher the number of zeros, the higher the speed of transition from a trading to a non-trading state.

Finally, the panel of distribution parameters points out that the coefficient π is very close to 1 (or equal to 1 by rounding to three decimal places) for DTE, which means that the ZAF distribution seems to be reduced to the Generalized F. If we consider the empirical frequency $\hat{\pi} = N^{-1} \sum_{i=1}^{N} 1I_{(y_{t,i}>0)}$ as an estimate for the probability that $y_{t,i} > 0$, it is equal to 0.9997 for the most liquid stock, to 0.9295 and to 0.8422 for G1A and SGZ respectively. Therefore, the estimated parameter π of the ZAF distribution, is very close to the empirical frequency of non-zero outcomes[4].

[3]The parameters ω_1 of the Beta weighting scheme is always setted equal to 1 in order to have a monotonically decreasing weights.

[4]For further details about the possibility of using the empirical frequency $\hat{\pi}$ rather than to estimate the parameter π of the ZAF distribution see Hautsch et al. (2013).

Table 4.2: In sample parameter estimates

		Intra-daily parameters						Trend parameters												Distribution parameters			
Ticker	Model	ω^*	α_1	α_0	β_1	β_2	ϱ_g	m	$\alpha_{1,d}$	$\beta_{1,d}$	$\beta_{1,w}$	$\beta_{1,m}$	θ_d	θ_h	ω_d	ω_h	m_1	m_2	γ	δ	a	b	c	π
DTE	MEM	0.088	0.381	-0.079	0.355	0.176	0.912	-	-	-	-	-	-	-	-	-	-	-	-	-	3.911	0.661	0.830	1.000
			(0.005)	(0.010)	(0.014)	(0.012)															(0.087)	(0.020)	(0.025)	(0.000)
	CMEM	0.094	0.367	-0.091	0.251	0.288	0.906	0.336	0.270	0.441	-	-	-	-	-	-	-	-	-	-	3.644	0.729	0.794	1.000
			(0.005)	(0.010)	(0.012)	(0.012)		(0.086)	(0.013)	(0.091)											(0.080)	(0.022)	(0.029)	(0.000)
	HAR-MEM	0.109	0.358	-0.097	0.323	0.210	0.892	0.285	-	0.291	0.173	0.281	-	-	-	-	-	-	-	-	3.482	0.773	0.973	1.000
			(0.005)	(0.011)	(0.015)	(0.013)		(0.040)		(0.022)	(0.039)	(0.053)									(0.089)	(0.027)	(0.036)	(0.000)
	MIDAS-MEM	0.069	0.371	-0.075	0.249	0.312	0.931	-0.417	-	-	-	-	0.011	-	48.902	-	-	-	-	-	4.495	0.554	0.664	0.999
			(0.005)	(0.007)	(0.012)	(0.012)		(0.053)					(0.001)		(0.514)						(0.096)	(0.015)	(0.019)	(0.000)
	H-MIDAS-CMEM	0.210	0.346	-0.199	0.185	0.260	0.790	-1.054	-	-	-	-	-0.172	0.192	55.576	52.587	-	-	-	-	5.202	0.471	0.579	1.000
			(0.005)	(0.012)	(0.013)	(0.013)		(0.026)					(0.003)	(0.003)	(0.190)	(0.102)					(0.265)	(0.029)	(0.036)	(0.000)
	H-MIDAS-CMEM-TVI	0.252	0.339	-0.245	0.267	0.142	0.748	-	-	-	-	-	-0.065	0.081	15.611	17.175	-0.529	-1.742	1.961	0.398	2.803	1.023	1.257	1.000
			(0.005)	(0.002)	(0.016)	(0.013)							(0.010)	(0.010)	(0.046)	(0.293)	(0.017)	(0.053)	(0.048)	(0.023)	(0.027)	(0.017)	(0.019)	(0.000)
G1A	MEM	0.042	0.285	0.026	0.325	0.346	0.956	-	-	-	-	-	-	-	-	-	-	-	-	-	3.073	0.386	0.818	0.929
			(0.005)	(0.006)	(0.016)	(0.015)															(0.088)	(0.013)	(0.032)	(0.001)
	CMEM	0.084	0.293	0.049	0.295	0.324	0.912	0.051	0.165	0.786	-	-	-	-	-	-	-	-	-	-	3.049	0.390	0.794	0.931
			(0.006)	(0.009)	(0.016)	(0.015)		(0.012)	(0.022)	(0.031)											(0.063)	(0.010)	(0.024)	(0.001)
	HAR-MEM	0.085	0.281	0.045	0.371	0.259	0.912	0.237	-	0.182	0.259	0.292	-	-	-	-	-	-	-	-	2.741	0.448	0.936	0.930
			(0.006)	(0.009)	(0.019)	(0.017)		(0.026)		(0.023)	(0.043)	(0.049)									(0.091)	(0.018)	(0.047)	(0.001)
	MIDAS-MEM	0.064	0.287	0.053	0.367	0.277	0.932	-0.692	-	-	-	-	0.014	-	49.659	-	-	-	-	-	3.157	0.382	0.731	0.927
			(0.006)	(0.008)	(0.019)	(0.017)		(0.043)					(0.000)		(0.376)						(0.073)	(0.011)	(0.024)	(0.001)
	H-MIDAS-CMEM	0.151	0.260	0.032	0.389	0.197	0.847	-0.922	-	-	-	-	-0.155	0.169	46.325	46.209	-	-	-	-	2.809	0.433	0.939	0.930
			(0.006)	(0.011)	(0.021)	(0.019)		(0.024)					(0.004)	(0.004)	(0.073)	(0.047)					(0.078)	(0.014)	(0.042)	(0.001)
	H-MIDAS-CMEM-TVI	0.199	0.242	0.091	0.419	0.134	0.794	-	-	-	-	-	0.041	-0.042	56.698	62.378	1.540	-1.330	1.687	1.134	4.847	0.237	0.453	0.930
			(0.006)	(0.015)	(0.025)	(0.021)							(0.002)	(0.001)	(0.093)	(0.054)	(0.075)	(0.069)	(0.097)	(0.036)	(0.233)	(0.012)	(0.028)	(0.001)
SZG	MEM	0.055	0.294	0.035	0.310	0.336	0.939	-	-	-	-	-	-	-	-	-	-	-	-	-	3.066	0.463	0.772	0.846
			(0.005)	(0.005)	(0.017)	(0.015)															(0.079)	(0.015)	(0.027)	(0.001)
	CMEM	0.103	0.266	0.086	0.369	0.248	0.883	0.104	0.130	0.705	-	-	-	-	-	-	-	-	-	-	3.320	0.420	0.752	0.841
			(0.006)	(0.010)	(0.020)	(0.017)		(0.014)	(0.014)	(0.030)											(0.105)	(0.016)	(0.033)	(0.001)
	HAR-MEM	0.109	0.261	0.110	0.384	0.229	0.874	0.261	-	0.150	0.181	0.157	-	-	-	-	-	-	-	-	2.934	0.477	0.938	0.842
			(0.006)	(0.011)	(0.020)	(0.017)		(0.019)		(0.018)	(0.029)	(0.031)									(0.106)	(0.021)	(0.052)	(0.001)
	MIDAS-MEM	0.093	0.262	0.072	0.358	0.276	0.896	-0.870	-	-	-	-	0.011	-	45.901	-	-	-	-	-	3.557	0.385	0.689	0.843
			(0.005)	(0.009)	(0.019)	(0.017)		(0.033)					(0.001)		(0.750)						(0.008)	(0.003)	(0.007)	(0.001)
	H-MIDAS-CMEM	0.157	0.238	0.093	0.340	0.251	0.830	-1.282	-	-	-	-	-0.115	0.132	51.257	50.554	-	-	-	-	4.232	0.313	0.588	0.857
			(0.005)	(0.011)	(0.019)	(0.018)		(0.028)					(0.003)	(0.003)	(0.283)	(0.221)					(0.114)	(0.010)	(0.021)	(0.001)
	H-MIDAS-CMEM-TVI	0.248	0.205	0.168	0.521	-	0.726	-	-	-	-	-	-0.075	0.078	38.967	42.086	0.257	-2.083	2.225	0.430	4.499	0.302	0.531	0.843
			(0.006)	(0.017)	(0.013)								(0.003)	(0.003)	(0.293)	(0.370)	(0.048)	(0.133)	(0.082)	(0.056)	(0.252)	(0.019)	(0.038)	(0.001)

In sample January 2003 – December 2007 parameter estimates for the models analysed on 10 minutes trading volumes: Standard MEM (MEM), Component MEM (CMEM), Heterogeneous AR-MEM (HAR-MEM), MIDAS-MEM with daily filter (MIDAS-MEM), Heterogeneous MIDAS Component MEM with constant intercept (H-MIDAS-CMEM) and time-varying intercept (H-MIDAS-CMEM-TVI). ϱ_g: estimated persistence of the dynamic intra-daily non-periodic component. Standard errors are reported in parentheses. All coefficients are statistically significant at 5%.

Table 4.3: In-sample diagnostics for the fitted models

		Residuals $\hat{\varepsilon}_{t,i}$				Squared Residuals $\hat{\varepsilon}^2_{t,i}$					
Ticker	Model	$\hat{\rho}_{(1)}$	Q_1	$\hat{\rho}_{(51)}$	Q_{51}	$\hat{\rho}_{(1)}$	Q_1	$\hat{\rho}_{(51)}$	Q_{51}	$\mathcal{L}(y)$	BIC
DTE	MEM	-0.003	0.550 (0.458)	0.025	327.342 (0.000)	0.002	0.123 (0.726)	0.000	0.431 (1.000)	-31198.14	62482.74
	CMEM	0.005	1.078 (0.299)	0.020	233.619 (0.000)	0.002	0.289 (0.591)	0.000	0.800 (1.000)	-30977.80	62074.49
	HAR-MEM	0.009	3.643 (0.056)	0.019	227.216 (0.000)	0.004	0.752 (0.386)	0.000	1.827 (1.000)	-30911.36	61952.41
	MIDAS-MEM	0.005	1.291 (0.256)	0.019	220.691 (0.000)	0.001	0.067 (0.796)	0.000	0.189 (1.000)	-31122.52	62363.93
	H-MIDAS-CMEM	-0.008	3.421 (0.064)	0.024	168.488 (0.000)	0.002	0.240 (0.624)	0.000	1.328 (1.000)	-30456.97	61054.44
	H-MIDAS-CMEM-TVI	0.003	0.324 (0.569)	0.020	136.020 (0.000)	0.001	0.014 (0.905)	0.000	0.074 (1.000)	-30292.07	60757.08
G1A	MEM	0.008	2.811 (0.094)	0.007	254.768 (0.000)	0.000	0.000 (0.984)	-0.001	404.931 (0.000)	-60351.41	120789.30
	CMEM	0.001	0.021 (0.885)	0.002	185.817 (0.000)	0.000	0.000 (0.987)	-0.002	334.736 (0.000)	-60159.79	120438.50
	HAR-MEM	0.006	1.650 (0.199)	0.001	173.427 (0.000)	0.001	0.028 (0.867)	-0.002	394.680 (0.000)	-60135.86	120401.40
	MIDAS-MEM	0.006	1.908 (0.167)	0.003	193.863 (0.000)	0.001	0.021 (0.885)	-0.001	457.766 (0.000)	-60161.05	120441.00
	H-MIDAS-CMEM	0.001	0.024 (0.877)	0.009	136.363 (0.000)	-0.003	0.503 (0.478)	0.000	297.982 (0.000)	-59632.03	119404.60
	H-MIDAS-CMEM-TVI	0.005	1.084 (0.298)	0.008	95.844 (0.000)	-0.005	1.406 (0.236)	0.001	109.699 (0.000)	-59478.36	119129.60
SZG	MEM	0.004	0.976 (0.323)	0.018	515.965 (0.000)	0.002	0.131 (0.718)	0.004	68.463 (0.052)	-76550.46	153187.40
	CMEM	0.008	3.338 (0.068)	0.013	431.739 (0.000)	0.005	1.447 (0.229)	0.004	93.349 (0.000)	-76348.06	152815.00
	HAR-MEM	0.007	2.434 (0.119)	0.011	396.676 (0.000)	0.005	1.380 (0.240)	0.004	96.885 (0.000)	-76356.64	152843.00
	MIDAS-MEM	0.007	2.614 (0.106)	0.014	390.341 (0.000)	0.005	1.196 (0.274)	0.005	108.711 (0.000)	-76376.39	152871.70
	H-MIDAS-CMEM	-0.004	0.708 (0.400)	0.016	319.100 (0.000)	-0.008	3.014 (0.083)	0.009	111.967 (0.000)	-76035.26	152211.00
	H-MIDAS-CMEM-TVI	0.006	1.589 (0.208)	0.020	189.382 (0.000)	-0.010	5.034 (0.025)	0.010	105.176 (0.000)	-75255.22	150672.60

In sample diagnostics for the models analysed. $\hat{\rho}_{(l)}$: Autocorrelation at the l-th lag; Q_l: Ljung-Box statistics at the l-th lag with the corresponding p-values in parenthesis; $\mathcal{L}(y)$: Log-likelihood value; BIC: Bayesian Information Criterion.

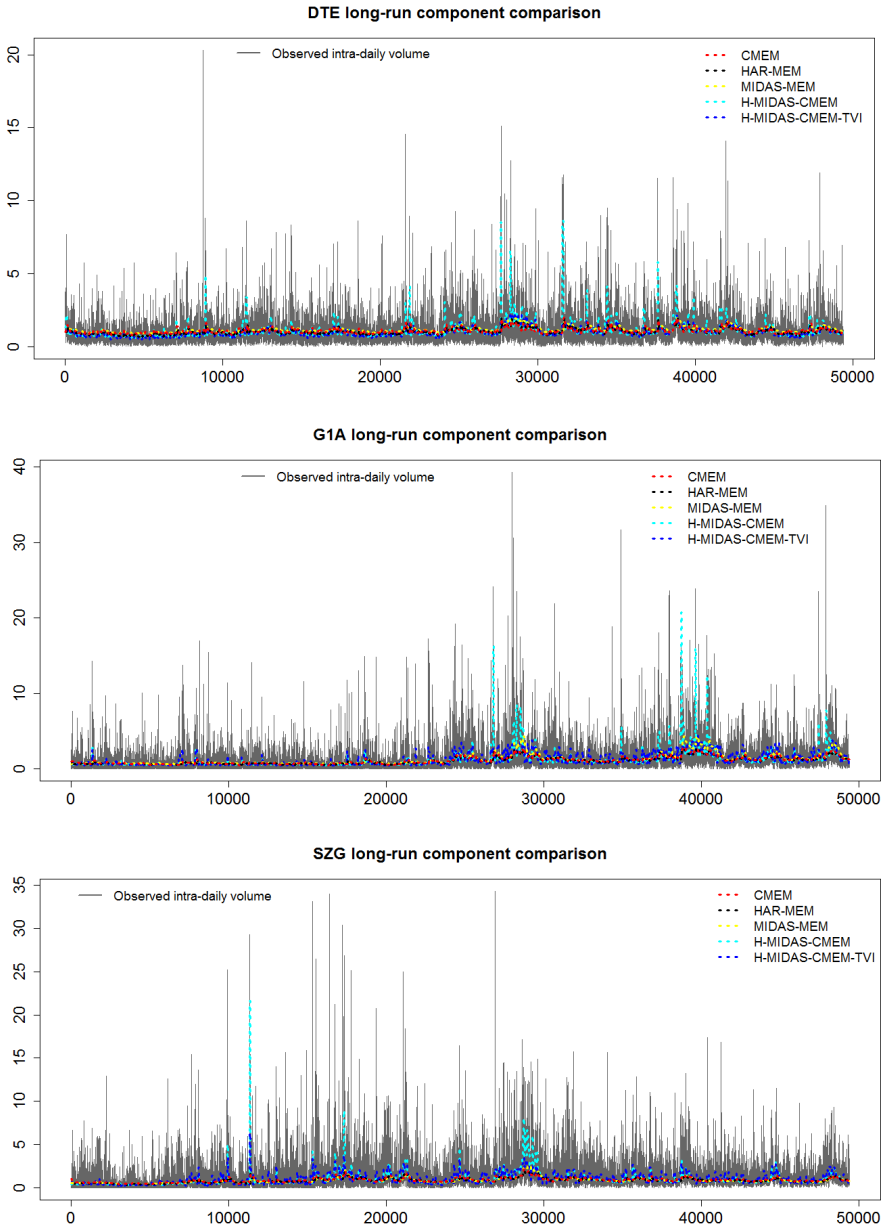
Diagnostics on estimated residuals in Table 4.3 show that $\hat{\rho}_{(1)}$ is statistically equal to zero for every model examined, while the white noise hypothesis is rejected at lag 51, though Ljung-Box statistics at lag 1-day (Q_{51}) reveal that H-MIDAS-CMEM models, both with fixed and time-varying intercept, exhibit the lowest serial autocorrelation for the three analysed stocks (as measured by the value of the Q_{51} statistic). However, considering the huge number of intra-daily observations $(N = 64719)$ it is very likely that the

white noise hypothesis is violated, since the confidence bands of residuals ACF are bounded approximately by ± 0.009. Focusing on the squared residuals $\hat{\varepsilon}_{t,i}^2$ there are no problems arising at lags 1 and 1-day for DTE, whereas for the less liquid stocks G1A and SGZ the white noise hypothesis is rejected, except for the standard MEM in the analysis of SZG. However, the use of the standard MEM shows the highest Q_{51} for $\hat{\varepsilon}_{t,i}$ in all observed cases.

From a comparison of the values in Table 4.3 it clearly emerges that the log-likelihood values recorded for the H-MIDAS-CMEM and H-MIDAS-CMEM-TVI are much larger than those of competing models. In particular, the BIC values show that there is a big improvement coming from the inclusion of a time-varying intercept in the H-MIDAS-CMEM, confirming the hypothesis that having an intercept which easily adapts to changes in the average volume level is very realistic. Component models different from H-MIDAS-CMEM exhibit results very close to each other, while the standard MEM seems to be the weakest competitor looking at the BIC values, thereby enhancing the empirical evidence that the overall volume series clusters around a lower frequency component and thus suggesting the use of component models.

In order to investigate the different behaviour of the long-run component provided by component models used in our empirical application, Figure 4.4 shows the comparison among the trend estimated by CMEM (red dashed line), HAR-MEM (black dashed line), MIDAS-MEM (yellow dashed line), H-MIDAS-CMEM (cyan dashed line) and H-MIDAS-CMEM-TVI (blue dashed line), for the three analysed stocks, whose observed volumes are drawn in grey. The long-run component given by CMEM, HAR-MEM and MIDAS-MEM display very similar dynamics to each

Figure 4.4: Long-run component comparison



Comparison of long-run component estimated using CMEM (red), HAR-MEM (black), MIDAS-MEM (yellow), H-MIDAS-CMEM (cyan) and H-MIDAS-CMEM-TVI (blue). Observed intra-daily volumes are drawn in grey.

other, but they are smoother than the trend estimated using the H-MIDAS-CMEM and the H-MIDAS-CMEM-TVI, especially in periods characterised by high trading intensity. In particular, the long-run component of the H-MIDAS-CMEM takes values much higher than the competing models when volumes shows high peaks.

4.8 Forecasting

To evaluate the predictive ability of the H-MIDAS-CMEM models we consider a forecasting exercise for the period January-December 2007, which includes 251 days. In order to capture the salient features of the data, the model parameters are recursively estimated every day starting from January 2006 with a 1-year rolling window. Therefore at each step we predict 51 intra-daily volumes before re-estimating the models, for a total of 251 days and 12801 intra-daily observations. The out-of-sample performance of the models examined is evaluated by computing some widely used forecasting loss functions. The significance of differences in forecasting performance is assessed by the Model Confidence Set (MCS) approach.

4.8.1 Out-of-sample evaluation

To compare the out-of-sample predictive performances we use the following loss functions:

$$L^{MSE} = \sum_{t=1}^{T} \sum_{i=1}^{I} (y_{t,i} - \hat{y}_{t,i})^2$$

$$L^{Slicing} = - \sum_{t=1}^{T} \sum_{i=1}^{I} (w_{t,i} \, log \, \hat{w}_{t,i})$$

where L^{MSE} is the Mean Squared prediction Error (MSE) of the volumes, while $L^{Slicing}$ is the Slicing Loss function developed by Brownlees et al. (2011) to evaluate VWAP trading strategies. The slicing weights $\hat{w}_{t,i}$ are computed under the static and dynamic VWAP replication strategies, with weights $w_{t,i} = y_{t,i}/\sum_{i=1}^{I} y_{t,i}$ which represents the i-th intra-daily volume proportion of day t. In the static VWAP replication strategy the weight $\hat{w}_{t,i}$ of the day t is given by the i-th one-step ahead intra-daily volume forecast relative to the sum of all the predicted intra-daily volumes for the same day. The dynamic strategy, instead, employs the static weights to upload the predicted volume proportions through new intra-daily information[5].

Table 4.4: Out-of-sample loss functions comparison

	DTE			G1A			SGZ		
	L^{MSE}	L^{SL}_{stc}	L^{SL}_{dyn}	L^{MSE}	L^{SL}_{stc}	L^{SL}_{dyn}	L^{MSE}	L^{SL}_{stc}	L^{SL}_{dyn}
MEM	0.481	3.920	2.767	1.997	3.921	2.765	0.869	3.921	2.764
CMEM	0.477	3.918	2.766	1.966	3.916	2.762	0.858	3.918	2.762
HAR-MEM	0.476	3.918	2.766	1.963	3.916	2.762	0.857	3.918	2.762
MIDAS-MEM	0.477	3.918	2.766	1.977	3.917	2.762	0.858	3.918	2.762
H-MIDAS-CMEM	0.465	3.915	2.764	1.958	3.909	2.757	0.850	3.912	2.758
H-MIDAS-CMEM-TVI	**0.455**	**3.914**	**2.763**	**1.850**	**3.907**	**2.756**	**0.799**	**3.911**	**2.757**

Loss functions values for Mean Squared Error (L^{MSE}) and Slicing Loss with weights computed under the static (L^{SL}_{stc}) and dynamic (L^{SL}_{dyn}) VWAP replication strategy. In **bold** the best model.

The loss functions for single model shown in Table 4.4 point out that the H-MIDAS-CMEM with fixed and mainly with time-varying intercept returns the lowest values for both the Mean Squared prediction Error (L^{MSE}) and the Slicing Loss using weights computed under the static (L^{SL}_{stc}) and dynamic (L^{SL}_{dyn}) VWAP replication strategy. A lower value of

[5]See Brownlees et al. (2011) for further details.

L^{MSE} provides evidence of a greater ability to capture the continuous variation from calms to storms periods, since intra-daily volume series are highly volatile, whereas minimizing the Slicing Loss function increases the chances to achieve the VWAP target for a given trading strategy.

4.8.2 Model comparison

The availability of several competing models to address the same problem raises the question of selecting the ones that more effectively reach the goal set. In most cases it is not trivial to establish which is the model that outperforms the others, since, as observed by Hansen and Lunde (2005) and Hansen et al. (2011), they might be statistically equivalent or because there could not be enough information in the data that allows us to select, doubtless, the best models. Starting from Diebold and Mariano (1995), who provide a test for the null hypothesis that two prediction models have equal accuracy, a lot of testing procedures have been developed to check whether a particular benchmark is significantly outperformed by any of the alternatives used in the forecasting comparison, see e.g. the Reality Check for data snooping of White (2000), the Stepwise Multiple Testing procedure of Romano and Wolf (2005), the Superior Predictive Ability test of Hansen (2005) and the Conditional Predictive Ability test of Giacomini and White (2006). A further generalisation is proposed by Hansen et al. (2011) in the form of the Model Confidence Set (MCS) procedure, which unlike the tests mentioned above does not require an a priori benchmark, providing greater flexibility in applications without an obvious benchmark.

Then, to evaluate the predictive power of H-MIDAS-CMEM models, we employ the MCS procedure of Hansen et al. (2011), which relies on a sequence of statistic tests to construct a set of superior models in terms

of predictive ability at certain confidence level $(1 - \alpha)$. In particular, given an arbitrary loss function, the MCS procedure sequentially tests the null hypothesis that all models have equal predictive ability and if it is accepted, the procedure stops and the set of superior models is created, otherwise, the equal predictive ability is tested again after the elimination of worst model.

Table 4.5 reports MCS p-values employing the T_{max} statistic, discussed in Hansen et al. (2011), at a confidence level $(1 - \alpha)$ of 0.75 and 0.95[6]. We use the block-bootstrap, estimating the optimal block length by the method described in Patton et al. (2009), with results based on 3000 bootstrap resamples. The highlighted values confirm the strength of the H-MIDAS-CMEM, since the model with time-varying intercept is always included in the MCS with p-value equal to 1 referring to the set of loss functions employed to measure the predictive ability of the models. For what concerns the H-MIDAS-CMEM with fixed intercept, it falls in the set of the superior models at both confidence levels considered for the static and dynamic Slicing Loss (except for the stock G1A for the static Slicing Loss at the 0.75 level), whereas looking at the Mean Squared Prediction Error this holds only for the less liquid stocks G1A and SZG at a confidence level of 95%, even if it provides p-values always higher then the competitors. Furthermore, the specifications different from the H-MIDAS-CMEM models never fall into the MCS according to the loss functions and the confidence levels considered.

[6]We report p-values only for T_{max} because the results corresponding to the T_{Range} statistic are practically equivalent.

Table 4.5: MCS p-values for different loss functions and confidence levels

	Mean Squared Error						Static Slicing Loss						Dynamic Slicing Loss					
	$\alpha = 0.05$			$\alpha = 0.25$			$\alpha = 0.05$			$\alpha = 0.25$			$\alpha = 0.05$			$\alpha = 0.25$		
	DTE	G1A	SZG	DTE	G1A	SZG	DTE	G1A	SZG	DTE	G1A	SZG	DTE	G1A	SZG	DTE	G1A	SZG
MEM	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.001	0.000	0.000	0.004	0.000	0.000	0.004	0.000	0.002
CMEM	0.001	0.002	0.000	0.000	0.000	0.000	0.010	0.000	0.002	0.008	0.002	0.003	0.019	0.000	0.000	0.020	0.001	0.002
HAR-MEM	0.001	0.000	0.000	0.000	0.000	0.000	0.014	0.000	0.004	0.015	0.002	0.005	0.025	0.000	0.003	0.026	0.000	0.004
MIDAS-MEM	0.000	0.000	0.000	0.000	0.000	0.000	0.006	0.000	0.002	0.008	0.002	0.003	0.016	0.000	0.000	0.020	0.000	0.002
H-MIDAS-CMEM	0.027	0.056	0.086	0.025	0.059	0.074	0.263	0.232	0.334	0.268	0.224	0.330	0.450	0.378	0.592	0.462	0.370	0.616
H-MIDAS-CMEM-TVI	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

MCS p-values corresponding to Mean Squared prediction Error and Slicing Loss with static and dynamic weights for the three examined stocks considering the whole set of models. Highlighted values suggest the models included in the MCS at a confidence level of 75% and 95%.

4.9 Conclusions

This paper introduces the Heterogeneous MIDAS Component MEM model (H-MIDAS-CMEM) as a novel approach to fit and forecast high-frequency volumes. Extending the logic of the CMEM developed in Brownlees et al. (2011) by the use of an Heterogeneous-MIDAS component specified as an additive cascade of linear filters which take on different frequencies, we are able to better capture the main empirical properties of intra-daily trading volumes, such as memory persistence and clustering of the trading activity. An important role is also played by the dummy variable included in the intra-daily dynamic non-periodic component, since it makes the model more flexible and adaptive to variations in liquidity dynamics when the volume series switches from trading to non trading intervals and vice versa.

From the analysis of three stocks (characterised by different levels of liquidity) traded in the German Stock Exchange it arises that the H-MIDAS-CMEM provides a very good fit both for liquid and illiquid stocks, estimating the model parameters by the ZAF distribution of Hautsch et al. (2013) in order to account for zero outcomes. The out-of-sample analysis confirms the strength of the H-MIDAS-CMEM, since it outperforms the competitors as it is clear from the MCS procedure proposed by Hansen et al. (2011). In fact, the H-MIDAS-CMEM is the only model always included in the set of superior models (at different confidence levels), minimizing the Slicing Loss with weights computed under the static and dynamic VWAP replication strategy and providing the highest p-values even for the MSE loss function. In addition, the inclusion of a time-varying intercept in the long-run component greatly improves in sample results and the predictive power of the model, making it the dominant one.

The model introduced in this paper is motivated by the empirical properties of high-frequency volumes, but our study also contributes to the literature dedicated to the analysis of other financial variables which share the same features, such as durations, bid-ask spread and volatility measures by employing a suitable distribution. A natural extension is to construct multivariate specification which can jointly model financial and/or economic variables as usually it occurs in a MIDAS framework, but, in addition, using filters that move at different frequencies. Furthermore, it is possible to differently model the function which drives the trend of the time-varying intercept or use a more flexible form for the seasonality component.

Appendix A

Filtering ultra high-frequency stock prices: an application to German stock market data

A.1 Introduction

Thanks to the rapid growth of information technology, in recent years it has become increasingly easy to store, transfer and process huge volumes of ultra high-frequency data. Demand for high-frequency data is growing at the same time, both for decision making and research purpose. Traders analyse tick-by-tick data to take decisions and implement trading strategies, whereas researchers use transaction data to observe market microstructure, test hypothesis and develop new econometric models. The *tick* is the minimum upward or downward movement in price of a security, but it also represents the information set of each transaction. The tick is usually recorded with information on the *time stamp* of a trade, the price at which a trade is executed and the volume in number of shares. The time stamp denotes the date and the time of a specific market activity.

The growing dominance of electronic trading systems together with the rapid technological progress, allows to record market activity on ultra high-frequency with high precision. However, high-frequency data show several issues arising not only from the structure of the market, but also from the statistical properties of the data. The structure of the data is strongly related to the trading rules, trading forms and procedures which the institutions employ to produce and collect information. For example, might be some differences between electronic and hybrid markets. Furthermore, regulatory changes and technological advances could make the data structurally different into the sample period. The stochastic nature of financial transactions highlights an important property of high-frequency data which is the irregular spacing in time. A further major feature of transaction data is the discreteness of prices. Institutional settings allow prices to be only multiples of a tick, consequently transaction prices take on discrete values. Furthermore, high-frequency datasets are characterised by a myriad of types and sources of errors, such as simultaneous observations, decimal errors, transposition errors, isolated and multiple bad ticks in succession. These errors are mainly linked to the trading intensity, since the higher the velocity in trading, the higher the probability of reporting incorrect information about transaction data (Brownlees and Gallo, 2006). As a consequence, there could be potential errors affecting high-frequency datasets, so before carrying out further analysis it is crucial to clean data from "bad ticks".

The remainder of this paper is structured as follows. Section A.2 describes the structure of the working dataset. Section A.3 highlights the main features of high-frequency data and shows different types and sources of errors which produce bad ticks. Section A.4 discusses different approaches for cleaning data. In section A.5 preliminary dataset manipulations are performed, whereas in section A.6 the Brownlees and Gallo (2006) outliers

detection procedure is implemented. In section A.7 interpolation methods are used to construct homogeneous time series. Section A.8 concludes.

A.2 Dataset structure

The analysed high-frequency dataset relies on transaction data of the German stock market indices DAX, TecDAX and MDAX covering the period from January 2002 to December 2012. The DAX consists of the 30 major German companies in terms of market capitalization. The TecDAX index tracks the performance of the 30 largest German companies from the technology sector. In terms of market capitalization the companies within the TecDAX rank below those included in the DAX. Finally, the MDAX includes the 50 largest companies from classic sectors ranking immediately below the companies included in the DAX index. The overall structure of our dataset is showed in Table A.1.

Table A.1: Dataset structure

Column	Content	Format
1	ISIN	Character
2	Date	YYYYMMDD
3	Time	HH:MM:SS
4	Price	Numeric
5	Price flag	Character
6	Volume	Numeric

The International Securities Identification Number (ISIN) uniquely identifies a security. The ISIN code is a 12-character alpha-numerical code that does not contain information characterising financial instruments, but

serves to uniformly identify an asset. The Price flag[1] describes different price categories, such as opening price, closing price and continuous trading price, which is the price that we take into account in our analysis. In continuous trading the price is determined by an order-driven mechanism which allows the trading of highly liquid securities.

Figure A.1: Xetra Continuous Trading and Auctions schedule

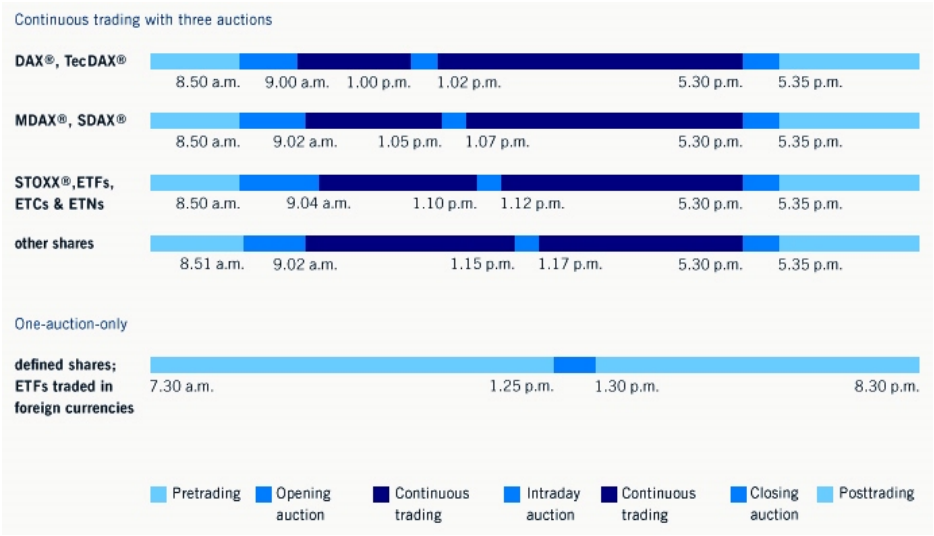


Figure A.1 shows the Xetra continuous trading and auctions schedule. In continuous trading, the orders are executed immediately as soon as corresponding orders are available. Auctions are conducted at the beginning and at the end of the trading day with an intraday auction around lunch time[2].

[1]Price Flags: B = Quote of the issuer; DA = End-of-day auction; E = Opening price; EA = Opening auction; IA = Intraday auction; L = Closing price from the day before; V = Continuous trading; VA = Auction after volatility circuit breaker; S = Closing price; SA = Closing auction; XB = Xetra best.

[2]More information are available on www.xetra.com

A.3 Properties and issues of transaction data

High-frequency financial data show features which are generally not observable in data measured at a lower intensity. One of the most important characteristics of transaction data is the irregular spacing in time. Since market transactions occur irregularly, tick-by-tick data are, consequently, irregularly spaced in time, giving rise to inhomogeneous time series. Furthermore, this leads to significant differences in tick frequency among securities. A further prominent feature is that high-frequency data are, by nature, discrete. Trading rules allow prices to be only multiples of a tick, where the minimum tick size varies from asset to asset and also across exchanges. Therefore prices fall into a set of discrete values. High-frequency data often show strong periodic patterns in tick frequency. A prominent example in this field is the well known U-shaped pattern in asset return volatility and trading volume, since the trading intensity is higher near the open and just prior to the close of the trading day than the middle of the day.

The considerable advances in computer science make the huge volume of high-frequency data easier to manage. Nevertheless, the amount of erroneous observations grows with the increasing of the volume and sampling frequency of the data. Simultaneous observations, decimal errors, transposition errors, isolated bad ticks and multiple bad ticks in succession, provide strong evidence of errors affecting ultra high-frequency datasets. Therefore, there are different types and sources of errors that produce bad ticks[3]. These "wrong" ticks could greatly affect the outcome of the analysis if they are not properly treated.

[3]See Dacorogna et al. (2001) and Falkenberry (2002) for an exhaustive overview about data error types.

Generally, errors can be classified into three different groups:

1. Human errors
    - unintentional errors, caused by the huge volume of observations and by the fast frequency of the transactions, such as decimal errors, typing errors or transposition errors. These errors are detected both in fully automated and in partially automated trading systems;
    - intentional errors, such as dummy ticks to test network connection.
2. System errors
    - technical errors caused by computer system, software and network connection failures.
3. Market bad data
    - errors given by processes inherent to trading. For example trade cancelled, replaced or corrected;
    - issues related to the simultaneous trading of the same asset on multiple markets.

Therefore, raw data need to be filtered in order to remove implausible ticks with market activity.

A.4 Data filtering approaches

Recently, several approaches have been proposed for filtering high-frequency data, but generally the cleaning procedures can be classified in *Search and Delete*, when bad ticks are identified and deleted, and in *Search and Replace*, whenever outliers are detected and replaced.

The latter, as discussed in Falkenberry (2002), uses a threshold and a moving average (which is function of the tick frequency) to determine if a certain observation is an outlier. Practically, when the absolute difference between price and moving filter exceed the user-defined threshold the tick is deemed bad. Ticks that are considered bad are replaced with the value of the moving filter, but without changing the information about volume. Therefore, this approach has the advantage of maintaining the volume of a trade even if the associated price has not actually been recorded.

Barndorff-Nielsen et al. (2009) proposed a *Search and Delete* outliers detection procedure both for trades and quotes, in order to investigate the performance of the Realised Kernel volatility estimator, applied simultaneously to trade and quote data. Since this cleaning procedure requires the use of bid-ask spread in different steps and our dataset does not provide information about bid and ask prices, we prefer to apply the filtering procedure introduced by Brownlees and Gallo (2006). It is an heuristic procedure easy to implement for removing "false" ticks which are relevant for the user, without changing the real-time properties of the raw data.

A.5 Preliminary filtering procedure

The filtering procedure has been carried out by using the statistical software *R*, because it provides a wide range of packages to manage, clean and match ultra high-frequency data and because it is a programming language easily extensible through functions written by the users. In order to get a more homogeneous dataset, some preliminary manipulations have been performed before applying the outliers detection procedure of Brownlees and Gallo (2006).

Table A.2: Dataset structure before preliminary filtering procedure

ISIN	Date	Time	Price	Price Flag	Volume
DE0007236101	20020731	12:59:25	53.32	V	2000
DE0007236101	20020731	12:59:45	53.25	V	400
DE0007236101	20020731	13:02:31	53.35	IA	10147
DE0007236101	20020731	13:02:35	53.35	V	200
DE0007236101	20020731	13:02:56	NA	.	NA

Table A.3: Dataset structure after preliminary filtering procedure

Stock	Time stamp	Price	Volume
Siemens	2002-07-31 12:59:25	53.32	2000
	2002-07-31 12:59:45	53.25	400
	2002-07-31 13:02:35	53.35	200

In particular, the time-date format of the data has been changed using the *timeDate* package, then data have been converted in *.xts* format to exploit the different tools for high-frequency data analysis given by the *highfrequency* package. In the next step zero or not available (NA) prices have been removed, selecting only continuous trading prices (V) within the regular trading hour 9:00 am - 5:30 pm. Finally, the ISIN code has been replaced with the corresponding company's name. Furthermore, the last trading day of each year is systematically excluded from the dataset, since different trading times apply. These preliminary steps allow to better manage the dataset in order to detect and remove wrong ticks and to construct homogeneous time series. Table A.2 and Table A.3 show, respectively, an example of the dataset structure before and after the preliminary filtering procedure.

A further type of error that usually affects high-frequency datasets is related to the so called simultaneous observations. The simultaneous observations problem occurs when several transactions are reported at the same time, but executed at different price levels. This phenomenon can be explained by the possibility that securities can be traded on different exchanges or because the execution of a market order, in some cases, could produce more than one transaction report. This issue can also be related to approximation errors, reporting non simultaneous prices as simultaneous.

Since most of the time series analysis methods require one observation per time stamp, to overcome the problem of simultaneous prices, observations with the same time stamp are merged and the corresponding values are aggregated by taking the median price. The discrete nature of the tick-by-tick data together with the high probability of detecting outliers, explain the choice of taking the median price. An example of the simultaneous observations is displayed in Table A.4.

Table A.4: Simultaneous prices

Stock	Time stamp	Price
	2008-01-02 09:09:39	23.34
DEUTSCHE POST AG	2008-01-02 09:09:39	23.35
	2008-01-02 09:09:39	23.39

A.6 Outliers detection procedure

It is not trivial to manage a huge dataset, since it is impossible to make assumptions on the dynamic properties of the data. Trading activity differs greatly across securities and this represents a significant problem

in developing a filter which is able to remove bad ticks. Therefore, the filtering process can not be based on complex parametric models, but it must follow empirical procedures simple to implement without changing the dynamic properties of the raw data.

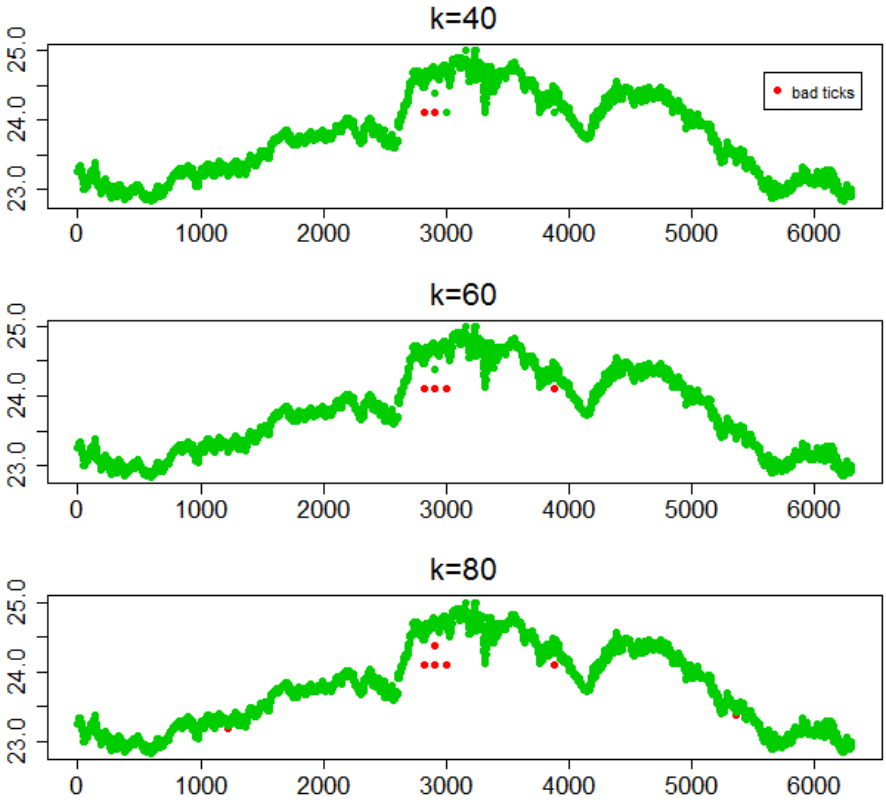
Brownlees and Gallo (2006) proposed the following procedure to detect outliers

$$(|p_i - \bar{p}_i(k)| < 3s_i(k) + \gamma) = \begin{cases} \text{True} \Longrightarrow \text{ observation } i \text{ is kept} \\ \text{False} \Longrightarrow \text{ observation } i \text{ is removed} \end{cases}$$

where $\{p_i\}_{i=1}^N$ is an ordered tick-by-tick price series, $\bar{p}_i(k)$ is the δ-trimmed sample mean of a neighborhood of k observations around i, $s_i(k)$ is the sample standard deviation of a neighborhood of k observations around i and γ is a *granularity* parameter.

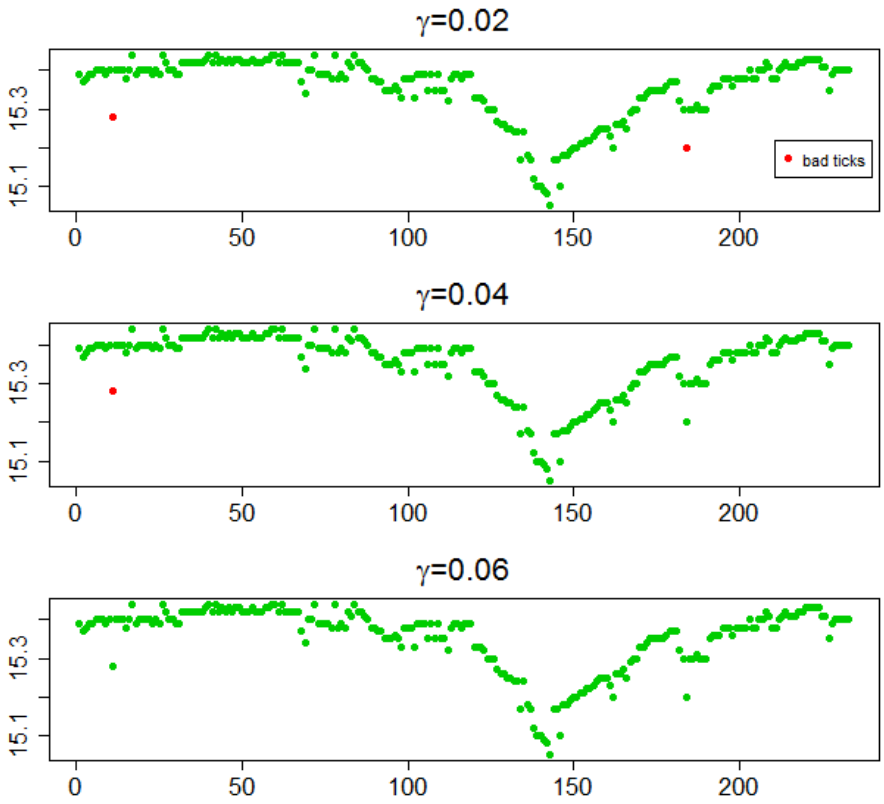
As in Falkenberry (2002), the filtering algorithm is based on a threshold, though, in this case it is not fixed, but it takes different values according to the neighborhood of the k observations around the i-th tick. The parameter γ avoids that zero variances are produced by sequences of k equal prices. The parameter δ should be chosen on the basis of the outliers frequency, that is the higher the frequency, the higher the percentage of trimming. The value of k should be fixed according to the level of the trading intensity. Therefore, if the trading activity is not so high, k should be reasonably small, in order to avoid that the window of observations includes too distant prices. The neighborhood of observations is always picked such that a given tick is compared with the prices of the same trading day. To get the optimal parameters, the cleaning algorithm was run several times for a grid of different values of k and γ, keeping fixed δ at 10%.

Figure A.2: Sensitivity analysis for k



The figure shows the sensitivity analysis for the parameter k, considering the transaction price series of the day 08 October 2008 for the stock BMW within the trading hour 12:00-14:30. The green points represent the clean price series, while the red points are the bad ticks. Top panel: clean time series with $k = 40$ and $\gamma = 0.02$. Center panel: clean time series with $k = 60$ and $\gamma = 0.02$. Bottom panel: clean time series with $k = 80$ and $\gamma = 0.02$.

Figure A.3: Sensitivity analysis for γ



The figure shows the sensitivity analysis for the parameter γ, considering the transaction price series of the day 14 January 2003 for the stock CON. The green points represent the clean price series, while the red points are the bad ticks. Top panel: clean time series with $k = 30$ and $\gamma = 0.02$. Center panel: clean time series with $k = 30$ and $\gamma = 0.04$. Bottom panel: clean time series with $k = 30$ and $\gamma = 0.06$.

A graphical representation of the sensitivity analysis of the parameters k and γ is showed in Figure A.2 and Figure A.3, respectively. In order to consider multiple scenarios, we focus on the transaction price series of the stocks Bayerische Motoren Werke (BMW) and Continental (CON), since they show a different trading intensity level. In particular, BMW (which is a German luxury vehicle, motorcycle, and engine manufacturing company) is a very liquid stock, as it shows a high trading intensity, while CON (which is a leading German automotive manufacturing company specialising in tyres, brake systems and automotive safety) is a less liquid stock, since it has a low trading intensity.

Figure A.2 provides an example of the sensitivity of the parameter k, showing the transaction price series of the day 08 October 2008 for the asset BMW within the trading hour 12:00-14:30. The analysis is addressed by changing the values of k and by keeping the values of γ fixed at 0.02. It can be noted that the number of bad ticks detected by the filtering algorithm rises with the increasing of k. In particular, for $k = 40$ the outliers detection procedure captures only two bad ticks, for $k = 60$ the outliers become four, while by setting $k = 80$ the algorithm removes seven anomalous observations.

Figure A.3 displays the sensitivity analysis for the parameter γ, plotting the price series of the stock CON of the day 14 January 2003 by considering three different scenarios. The sensitivity is assessed by allowing for different values of γ and setting $k = 30$. In this case, the higher the γ, the lower the number of anomalous prices highlighted by the cleaning algorithm. It clearly emerges that, when $\gamma = 0.02$ the number of outliers is equal to two, choosing $\gamma = 0.04$ the filtering algorithm detects only one bad tick, whereas for $\gamma = 0.06$ there are no false ticks.

As discussed in Falkenberry (2002), the main problem in developing cleaning algorithms is the handling of marginal errors, since it involves a trade off between *overscrubbing* and *underscrubbing* of tick data. If data are filtered too loosely we still have unusable dirty data and, on the other hand, if data are filtered too tightly we might lose relevant information, changing the statistical properties of the raw data. Furthermore, the quality of the data might have a different relevance among users according to their interests and market strategies. What is considered a bad tick for a tick-based trader may be considered negligible for a trader using lower frequency strategy, such as hourly or daily. Therefore, marginal errors are function of the time unit (tick-by-tick, 1-minute, 60-minute, etc.) used by the traders to achieve their targets. In describing the issues associated with maintaining and cleaning a high-frequency financial database, Falkenberry (2002) states that

"the primary objective in developing a set of tick filters is to manage the overscrub/underscrub trade off in such a fashion as to produce a time series that removes false outliers in the trader's base unit of analysis that can support historical backtesting without removing real-time properties of the data".

Following these considerations, we found empirical evidence that the optimal overscrub/ underscrub trade off that complies with our research interests is achieved by setting $\gamma = 0.02$ and by using $k = 30$ for illiquid assets, $k = 40$ for liquid assets and $k = 60$ for stocks showing a very high trading intensity.

A.7 Data aggregation

Tick-by-tick data are, by nature, irregularly spaced in time. Furthermore, transactions on multiple assets rarely occur at the same time, since the

trading activity varies considerably across assets. As high-frequency models are typically based on homogeneous time series, it is necessary to ensure that each trading session has the same number of observations. Several data aggregation procedures have been proposed to get regularly spaced time series. Interpolation methods allow to transform raw data from inhomogeneous to homogeneous time series. Interpolation is a method for constructing artificial data points using the available information about existing data. To transform data in equally spaced time series Dacorogna et al. (2001) proposed some interpolation methods, where the most commonly used in empirical applications are the *linear interpolation* and the *previous-tick interpolation*.

Let $\{(x_i, t_i)\}_{i=1}^{N}$ be an irregular time series where t_i and x_i denote, respectively, the time and value of the i-th observation and let $\{(x_j^*, t_j^*)\}_{j=1}^{N^*}$ be the lower frequency time series that we wish to build according to a suitable aggregation function.

Using the previous-tick interpolation it follows that

$$x_j^* = x_p,$$

where $t_p = max\{t_i | t_i < t_j^*\}$.

On the other hand, employing the linear interpolation we get

$$x_j^* = \left(1 - \frac{t_j^* - t_p}{t_n - t_p}\right) x_p + \left(\frac{t_j^* - t_p}{t_n - t_p}\right) x_n,$$

where $x_j^* = x_n$ and $t_n = min\{t_i | t_i > t_j^*\}$.

The differences between the two interpolation methods, showed in Figure A.4, are quite negligible when ultra high-frequency data are used (Dacorogna et al., 2001).

Using the linear interpolation, irregularly spaced price series have been aggregated into equally spaced time grids, thus forming homogeneous time series. Figure A.5 and Figure A.6 show examples of equally-spaced time series of stock prices and log-returns computed at different frequencies. In particular, for the most liquid stock BMW, Figure A.5 displays intra-daily prices and log-returns aggregated at the frequencies of 30 seconds and 1 minute of the day 08 October 2008, whereas for the stock CON, which shows a lower trading activity, Figure A.6 depicts time series of intra-daily 5 minutes and 10 minutes transaction prices and log-returns of the day 14 January 2003.

Figure A.4: Different interpolation methods

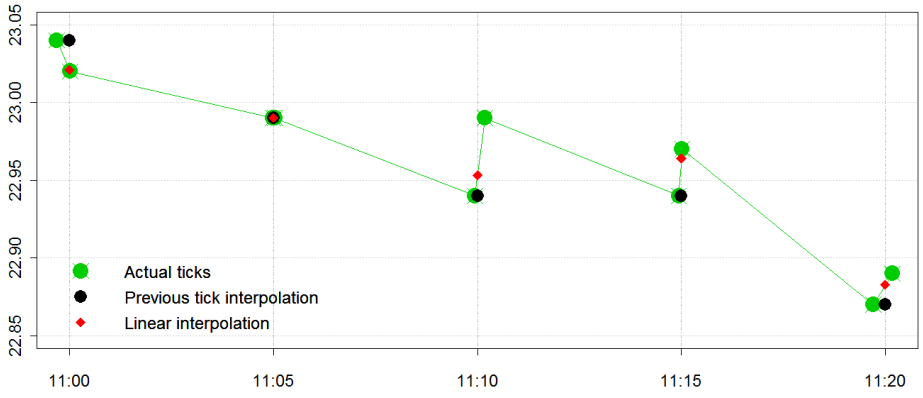
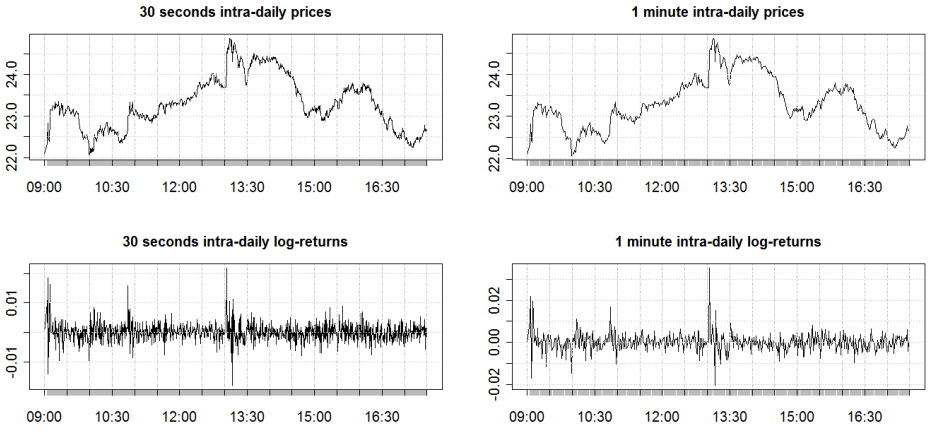
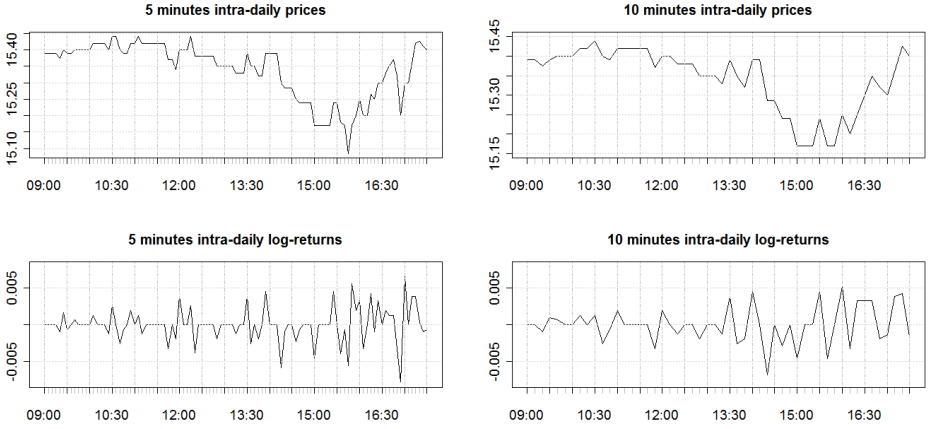


Figure A.5: BMW intra-daily 30 seconds and 1 minute stock prices and log-returns



Time series plots of intra-daily 30 seconds and 1 minute transaction prices and log-returns for the stock BMW of the day 08 October 2008.

Figure A.6: CON intra-daily 5 minutes and 10 minutes stock prices and log-returns



Time series plots of intra-daily 5 minutes and 10 minutes transaction prices and log-returns for the stock CON of the day 14 January 2003.

A.8 Conclusion

The availability of low-cost financial transaction data has given rise to new challenges in econometrics and empirical finance. In this paper we have pointed out the main features and discussed on relevant issues concerning high-frequency data. High frequency analysis is playing an important role in market microstructure theory, option pricing, risk management and decision making process, therefore filtering tick-by-tick data accurately becomes a key requirement in these fields. The cleaning procedure by Brownlees and Gallo (2006) proves a very useful tool in detecting and removing bad ticks. To achieve satisfactory results, the outlier detection procedure must be performed for each stock in order to remove observations which are not plausible with the market activity. However, the assessment on the data quality is highly related to the interests and aims of the users.

Bibliography

Aït-Sahalia, Y., P. A. Mykland, and L. Zhang (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial studies 18*(2), 351–416.

Amado, C. and T. Teräsvirta (2013). Modelling volatility by variance decomposition. *Journal of Econometrics 175*(2), 142–153.

Andersen, T. G. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International economic review*, 885–905.

Andersen, T. G., T. Bollerslev, and J. Cai (2000). Intraday and interday volatility in the japanese stock market. *Journal of International Financial Markets, Institutions and Money 10*(2), 107–130.

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*(4), 701–720.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American statistical association 96*(453), 42–55.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*(2), 579–625.

Andersen, T. G., D. Dobrev, and E. Schaumburg (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics 169*(1), 75–93.

Bandi, F. M. and J. R. Russell (2006). Separating microstructure noise from volatility. *Journal of Financial Economics 79*(3), 655–692.

Bandi, F. M. and J. R. Russell (2008). Microstructure noise, realized variance, and optimal sampling. *The Review of Economic Studies 75*(2), 339–369.

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*(6), 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*(3), C1–C32.

Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 64*(2), 253–280.

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*(3), 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), 1–37.

Białkowski, J., S. Darolles, and G. Le Fol (2008). Improving vwap strategies: A dynamic volume approach. *Journal of Banking & Finance 32*(9), 1709–1722.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics 31*(3), 307–327.

Bollerslev, T., A. J. Patton, and R. Quaedvlieg (2016). Exploiting the errors: A simple approach for improved volatility forecasting. *Journal of Econometrics 192*(1), 1–18.

Brownlees, C. T., F. Cipollini, and G. M. Gallo (2011). Intra-daily volume modeling and prediction for algorithmic trading. *Journal of Financial Econometrics 9*(3), 489–518.

Brownlees, C. T. and G. M. Gallo (2006). Financial econometric analysis at ultra-high frequency: Data handling concerns. *Computational Statistics & Data Analysis 51*(4), 2232–2245.

Brownlees, C. T. and G. M. Gallo (2010). Comparison of volatility measures: a risk management perspective. *Journal of Financial Econometrics 8*(1), 29–56.

Brunetti, C. and P. M. Lildholdt (2007). Time series modeling of daily log-price ranges for chf/usd and usd/gbp. *The Journal of Derivatives 15*(2), 39–59.

Chou, R. Y. (2005). Forecasting financial volatilities with extreme values: the conditional autoregressive range (carr) model. *Journal of Money, Credit and Banking*, 561–582.

Cipollini, F., R. F. Engle, and G. M. Gallo (2006). Vector multiplicative error models: representation and inference. Technical report, National Bureau of Economic Research.

Cipollini, F., R. F. Engle, and G. M. Gallo (2013). Semiparametric vector mem. *Journal of Applied Econometrics 28*(7), 1067–1086.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics*, 174–196.

Dacorogna, M., R. Gencay, U. Muller, R. Olsen, and O. Pictet (2001). *An introduction to high frequency finance*. Academic Press, New York.

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*(1), 29–58.

Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics 13*(3).

Engle, R. (2002). New frontiers for arch models. *Journal of Applied Econometrics 17*(5), 425–446.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987–1007.

Engle, R. F. and G. M. Gallo (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics 131*(1), 3–27.

Engle, R. F., E. Ghysels, and B. Sohn (2013). Stock market volatility and macroeconomic fundamentals. *Review of Economics and Statistics 95*(3), 776–797.

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*(3), 1187–1222.

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

Engle, R. F. and M. E. Sokalska (2012). Forecasting intraday volatility in the us equity market. multiplicative component garch. *Journal of Financial Econometrics 10*(1), 54–83.

Falkenberry, T. N. (2002). High frequency data filtering. *Tick Data, Technical*.

Gallant, A. R. (1981). On the bias in flexible functional forms and an essentially unbiased form: the fourier flexible form. *Journal of Econometrics 15*(2), 211–245.

Gallo, G. M. and E. Otranto (2015). Forecasting realized volatility with changing average levels. *International Journal of Forecasting 31*(3), 620–634.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics 131*(1), 59–95.

Ghysels, E., A. Sinko, and R. Valkanov (2007). Midas regressions: Further results and new directions. *Econometric Reviews 26*(1), 53–90.

Giacomini, R. and H. White (2006). Tests of conditional predictive ability. *Econometrica 74*(6), 1545–1578.

Groß-KlußMann, A. and N. Hautsch (2013). Predicting bid–ask spreads using long-memory autoregressive conditional poisson models. *Journal of Forecasting 32*(8), 724–742.

Hansen, P. and A. Lunde (2011). Forecasting volatility using high frequency data. *The Oxford Handbook of Economic Forecasting, Oxford: Blackwell*, 525–556.

Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business & Economic Statistics 23*.

Hansen, P. R. and Z. Huang (2016). Exponential garch modeling with realized measures of volatility. *Journal of Business & Economic Statistics 34*(2), 269–287.

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*(6), 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*(7), 873–889.

Hansen, P. R. and A. Lunde (2006). Realized variance and market microstructure noise. *Journal of Business & Economic Statistics 24*(2), 127–161.

Hansen, P. R., A. Lunde, and J. M. Nason (2011). The model confidence set. *Econometrica 79*(2), 453–497.

Hautsch, N., P. Malec, and M. Schienle (2013). Capturing the zero: a new class of zero-augmented distributions and multiplicative error processes. *Journal of Financial Econometrics*, 1–33.

Lanne, M. (2006). A mixture multiplicative error model for realized volatility. *Journal of Financial Econometrics 4*(4), 594–616.

Lee, S.-W. and B. E. Hansen (1994). Asymptotic theory for the garch (1, 1) quasi-maximum likelihood estimator. *Econometric theory 10*(01), 29–52.

Liu, L. Y., A. J. Patton, and K. Sheppard (2015). Does anything beat 5-minute rv? a comparison of realized measures across multiple asset classes. *Journal of Econometrics 187*(1), 293–311.

Madhavan, A. N. (2002). Vwap strategies. *Trading 2002*(1), 32–39.

Manganelli, S. (2005). Duration, volume and volatility impact of trades. *Journal of Financial markets 8*(4), 377–399.

McAleer, M. and M. C. Medeiros (2008). A multiple regime smooth transition heterogeneous autoregressive model for long memory and asymmetries. *Journal of Econometrics 147*(1), 104–119.

Meddahi, N. (2002). A theoretical comparison between integrated and realized volatility. *Journal of Applied Econometrics 17*(5), 479–508.

Müller, U. A., M. M. Dacorogna, R. D. Davé, O. V. Pictet, R. B. Olsen, and J. R. Ward (1993). Fractals and intrinsic time: A challenge to econometricians. *Unpublished manuscript, Olsen & Associates, Zürich*.

Patton, A., D. N. Politis, and H. White (2009). Correction to "automatic block-length selection for the dependent bootstrap" by d. politis and h. white. *Econometric Reviews 28*(4), 372–375.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics 160*(1), 246–256.

Raggi, D. and S. Bordignon (2012). Long memory and nonlinearities in realized volatility: a markov switching approach. *Computational statistics & data analysis 56*(11), 3730–3742.

Romano, J. P. and M. Wolf (2005). Stepwise multiple testing as formalized data snooping. *Econometrica 73*(4), 1237–1282.

Scharth, M. and M. C. Medeiros (2009). Asymmetric effects and long memory in the volatility of dow jones stocks. *International Journal of Forecasting 25*(2), 304–327.

Shephard, N. and K. Sheppard (2010). Realising the future: forecasting with high-frequency-based volatility (heavy) models. *Journal of Applied Econometrics 25*(2), 197–231.

Shephard, N. and D. Xiu (2016). Econometric analysis of multivariate realised qml: efficient positive semi-definite estimators of the covariation of equity prices.

White, H. (2000). A reality check for data snooping. *Econometrica 68*(5), 1097–1126.

Zhang, L., P. A. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association 100*(472), 1394–1411.