Forecasting Covariance Matrices: A Mixed Frequency Approach

Roxana Halbleib and Valeri Voev

Working Paper Series
2012-30

http://www.wiwi.uni-konstanz.de/forschung/
Forecasting Covariance Matrices: A Mixed Frequency Approach*

Roxana Halbleib† University of Konstanz
Valeri Voev‡ Aarhus University

October 12, 2012

JEL classification: C32, C53
Keywords: Multivariate volatility, Volatility forecasting, High-frequency data, Realized variance, Realized covariance

*We would like to thank Luc Bauwens, Robert F. Engle, Christian Gourieroux, Asger Lunde, Andrew Patton, Jeroen Rombouts and Kevin Sheppard for comments and suggestions. All errors are ours.

†Corresponding address: University of Konstanz, Department of Economics, Universitaetsstrasse 10, 78464, Konstanz, Germany. Phone + 49 7531 88 5373, email: Roxana.Halbleib@uni-konstanz.de. The author gratefully acknowledges financial support from the European Social Fund and from the Ministry of Science, Research and Arts of Baden-Württemberg, Germany.

‡School of Economics and Management, Aarhus University, 8000 Aarhus C, Denmark. Phone +45 8942 1539, email: vvoev@creates.au.dk. Financial support by the Center for Research in Econometric Analysis of Time Series, CREATE, funded by the Danish National Research Foundation, is gratefully acknowledged.
Abstract

In this paper we introduce a new method of forecasting covariance matrices of large dimensions by exploiting the theoretical and empirical potential of using mixed-frequency sampled data. The idea is to use high-frequency (intraday) data to model and forecast daily realized volatilities combined with low-frequency (daily) data as input to the correlation model. The main theoretical contribution of the paper is to derive statistical and economic conditions, which ensure that a mixed-frequency forecast has a smaller mean squared forecast error than a similar pure low-frequency or pure high-frequency specification. The conditions are very general and do not rely on distributional assumptions of the forecasting errors or on a particular model specification. Moreover, we provide empirical evidence that, besides overcoming the computational burden of pure high-frequency specifications, the mixed-frequency forecasts are particularly useful in turbulent financial periods, such as the previous financial crisis and always outperforms the pure low-frequency specifications.
1 Introduction

Volatility modeling and forecasting have been of prime interest in financial econometrics since the seminal contributions of [Engle (1982)] and [Bollerslev (1986)]. Recently, research developments in the field have been refueled by the availability of high-frequency financial data on various financial instruments. High-frequency data has proven very useful in forecasting future volatility. Currently, there are a number of methods, mostly univariate, which propose dynamic models for realized volatility time-series, or alternatively, ways to integrate realized volatility measures into standard GARCH-type specifications. [Hansen & Lunde (2010)] provide a review of this growing literature.

In this paper, we propose a new method of forecasting covariance matrices of high dimensions by mixing volatility and correlation forecasts stemming from measures computed from data sampled at different frequencies. As a particular mixed-frequency specification, we propose to forecast daily covariance matrices by mixing forecasts computed from data sampled at daily and high frequency level.

The main theoretical contribution of the paper is the derivation, under a model-free environment, of the economic and statistical conditions under which a forecast stemming from such a mixing approach outperforms a pure single frequency (daily or high-frequency) forecast. Thus, the generality of our derivations allows a practitioner to verify and decide in advance whether the current economic conditions and/or the properties of the correlation and volatility forecasts, she has at hand, are in favor for using our approach or a standard single-frequency approach to accurately and precisely forecast her covariance matrix of daily returns. Moreover our derivations help at understanding which role the single correlation and volatility forecasts play at providing accurate forecasts for the whole covariance matrix subject to the current market conditions: high vs. low volatility/correlation states. The underlying setting of our derivations is the volatility-correlation decomposition of the covariance matrix, which enables modeling of large dimensions. Apart from that, no particular model specification or distributional assumptions are made. The loss function we consider in this study is the element-wise forecast mean squared error (MSE). Alternative loss functions (e.g., a matrix-based MSE criterion or an economically-motivated loss) can be considered in a similar framework. Clearly, if element-wise dominance of a forecast holds, dominance on a matrix loss or on a linear combination of the matrix components, such as the one implied by a deterministic portfolio choice, follows a fortiori. We express the MSE as a function of the ex-post correlation and volatility realized over the period we are forecasting, which allows us to analyze the models’ performance under various market conditions (e.g., a high-volatility, high-correlation state). We show that mixing forecasts stemming from different sampling frequencies can lead to significant forecasting gains during highly volatile periods under very general statistical properties of the mixing forecasts and at very low computational costs.

---

1This is different from the concept of conditional predictive ability analyzed in [Giacomini & White (2006)]. Our loss function is unconditional with respect to the past information. Given that we derive the results in a model-free environment, we do not explicitly define the conditioning set and thus we cannot examine conditional loss functions. With a particular model specification, conditional loss functions can be examined.
While the pure daily models for modeling and forecasting covariance matrices are very popular (see Bauwens et al. (2006) for a comprehensive survey), they suffer from curse of dimensionality, need to impose heavy parameter restrictions to assure the positive definiteness of the covariance forecasts and treat the volatility and correlation process as latent. Alternatively, high-frequency data contains information that allows for almost error-free measurement of volatility ex-post, based on the estimation of the quadratic variation of the price process, thus making it effectively observable. Early studies in the area (see e.g., Andersen, Bollerslev, Diebold & Labys (2001), Andersen, Bollerslev, Diebold & Ebens (2001), Andersen et al. (2003)) recognized that market microstructure effects can distort estimation at very high frequencies and proposed a sparse sampling approach, in which the available data is sampled every 5, 10 or 15 minutes to mitigate the impact of market microstructure noise. More recently, techniques have been developed to use the data more efficiently by designing estimators that are noise-robust (see e.g., Barndorff-Nielsen et al. (2008), Barndorff-Nielsen et al. (2009), Jacod et al. (2009), Zhang (2011), Zhang et al. (2005), Nolte & Voev (2011), etc). Most of these approaches are applicable to univariate series, i.e., for volatility, rather than for covariance estimation. While multivariate extensions of the above mentioned approaches do exist (see e.g., Voev & Lunde (2007), Barndorff-Nielsen et al. (2011), Nolte & Voev (2008), Christensen et al. (2010)) they suffer from limitations especially when applied to many assets. In most empirical work, realized covariance estimation is still carried out using the sparse-sampling approach. The problem with the sparse-sampling method is that for dimensions higher than the number of observations on the sparse subgrid (e.g., at the typical 5-minute frequency there are 78 observations for a NYSE traded stock), the realized covariance matrices are of reduced rank and, thus, singular. Generally, it can be stated that the covariance/correlation estimation with high-frequency data is much more challenging than volatility estimation due to issues of non-synchronicity of the raw multivariate series and parameter proliferation. As a result, our new specification aims at alleviating the drawbacks of the pure daily and high-frequency approaches and to "mix" their advantages in providing accurate forecasts that are easy to obtain for covariance matrices of any dimension.

For implementation purposes, we opt for the covariance matrix decomposition in terms of standard deviations and correlations. We forecast the volatility using dynamic models for the univariate series of realized volatilities, which can be estimated by any of the above-mentioned techniques. The correlation matrix forecast is conceptually identical to the DCC specification of Engle (2002), but with the important difference that we standardize (de-volatilize) returns by realized volatilities rather than by GARCH volatilities. Thus, the forecasting improvement over similar decomposed covariance matrix specifications stemming solely from low-frequency (daily) data (such as the DCC model of Engle (2002) and Tse & Tsui (2002)) is thus driven by the improvement in volatility forecasts and by using less noisy standardized residuals as an input to the correlation model. Andersen, Bollerslev, Diebold & Ebens (2001) find that the normal distribution provides a "strikingly" close match to the density of returns scaled by

\footnote{See Hautsch et al. (2010) and Lunde et al. (2011) for a recent contribution on estimation of large-dimensional realized matrices.}
realized volatilities.

Comparing to pure high-frequency data approaches, our method only requires the estimation of realized volatility series, rather than realized covariance matrices, which, as discussed above, is more problematic. This gives the advantage that the mixed-frequency framework is better suited to handle matrices of large dimensions. Furthermore, model specifications for realized covariance/correlation matrices are only recently gaining more attention (see e.g., Gourieroux et al. (2009), Bauer & Vorkink (2011), Chiriac & Voev (2011), Jin & Maheu (2012), Noureldin et al. (2012), Bauwens et al. (2012), Hautsch et al. (2011)) and there is still a lot of empirical work needed in order for these models to gain broader recognition.

The empirical study of the paper examines the validity of the theoretical conditions when forecasting the covariance matrix of 30 very liquid stocks and finds that, over the whole sample, that includes the recent financial crisis, the mixed-frequency specification outperforms decisively the daily-frequency specification and performs as good as a much more computationally intensive pure high-frequency model. However, particularly during very turbulent financial periods, such as the one following the Lehman Brother’s bankruptcy in September 2008, the mixed-frequency approach outperforms both daily-frequency and high-frequency models.

To the best of our knowledge, the papers of Bannouh et al. (2010) and Colacito et al. (2011) are the only other studies that consider mixed-frequency covariance models. However the differences of these studies to our are very stark and it suffices to mention only a few points of departure. Firstly, the model of Bannouh et al. (2010) uses a factor structure in which the factor covariance matrix is estimated with high-frequency data and the loadings on the factors are estimated with daily data. Our approach does not assume a factor structure of the covariance matrix, while it clearly does not exclude that there is one. Secondly, their model is a static one, in the sense that they focus on the issue of estimation of covariance matrices of very large dimension, rather than on forecasting. Alternatively, Colacito et al. (2011) use mixed frequency to forecast conditional covariance matrices. Differently from our approach, that mixes volatility forecasts stemming from intradaily data with correlation forecasts stemming from daily data, the DCC-MIDAS model of Colacito et al. (2011) mixes short run components computed from daily data with long run components computed from monthly or quarterly data for both volatility and correlation forecasts. In fact, the only thing that the three papers have in common is that they use in some way data at different frequencies, however under very different implementations.

The remainder of the paper is structured as follows: Section 2 introduces the mixed-frequency approach and presents the theoretical results, Section 3 contains the empirical study, and Section 4 concludes. The proofs to the two propositions in Section 2 are contained in Appendix A. Appendix B contains tables and graphs.
2 Theory

2.1 General Settings

Let \( r_t \) be a vector of daily log returns of dimension \( n \) – the number of assets considered. In this paper, we are interested in computing one-step ahead conditional covariance matrix forecasts of \( r_t \), namely \( E[\Sigma_{t+1}|\mathcal{F}_t] \), where \( \mathcal{F}_t \) is a time \( t \) information set (a discussion on multi-step ahead forecasts will follow later). \( \Sigma_{t+1} \) is the actual realized covariance over the next period, i.e., the quadratic variation of the (Itô semi-martingale) price process from \( t \) to \( t + 1 \). We propose to employ the following decomposition of the covariance matrix

\[
\Sigma_t = D_t R_t D_t, \tag{1}
\]

where \( D_t \) is a diagonal matrix given by the conditional standard deviations of each stock and \( R_t \) is the correlation matrix. This decomposition has been used in Engle (2002) and Tse & Tsui (2002) in a dynamic conditional correlation (DCC) framework. More precisely, here we propose to compute the conditional forecasts of \( \Sigma_t \) from the conditional forecasts of \( D_t \) and \( R_t \) as follows:

\[
\hat{\Sigma}_{t+1|t} = \hat{D}_{t+1|t} \hat{R}_{t+1|t} \hat{D}_{t+1|t}, \tag{2}
\]

where \( \hat{D}_{t+1|t} \equiv E[D_{t+1}|\mathcal{F}_t^D] \), \( \hat{R}_{t+1|t} \equiv E[R_{t+1}|\mathcal{F}_t^R] \) and \( \mathcal{F}_t^D \) and \( \mathcal{F}_t^R \) are volatility and correlation specific information sets up to time \( t \). In our approach, we set \( F^D_t \) and \( F^R_t \) are volatility and correlation specific information sets up to time \( t \). In the sequel, we will differentiate between forecasts based on the information set containing daily measures computed on high-frequency data (e.g., time series of realized volatilities and correlations), \( \mathcal{F}_t^H \), and forecasts based on the information set containing data at the low frequency (typically daily returns), \( \mathcal{F}_t^L \). Let \( \hat{\sigma}_{i,t+1|t} = E[d_{i,t+1}|\mathcal{F}_t^{D,H}] \) and \( \hat{\rho}_{ij,t+1|t} = E[\rho_{ij,t+1}|\mathcal{F}_t^{R,H}] \), with \( \mathcal{F}_t^{D,H} \) and \( \mathcal{F}_t^{R,H} \) be volatility and correlation specific \( \mathcal{F}_t^H \), be the \( i \)-th volatility and \( ij \)-th correlation forecast from a dynamic model for the daily series of realized measures computed on high frequency data, such as the Autoregressive Fractionally Integrated Moving Average (ARFIMA) approach suggested by Andersen, Bollerslev, Diebold & Labys (2001) and Andersen, Bollerslev, Diebold & Ebens (2001) or the Heterogenous Autoregressive (HAR) model of Corsi (2009) and Corsi & Audrino (2010) estimated on series of daily realized volatilities and realized correlations. Enhanced GARCH specifications in which realized measures are included as volatility and correlation predictors (see, e.g., Engle & Gall (2006), Shephard & Shephard (2010), Hansen et al. (2012), Noureldin et al. (2012) and Hansen, Lunde & Voev
can also be considered as HF-data models. Further, let \( \hat{d}_{i,t+1}^L = E[d_{i,t+1} | \mathcal{F}_t^{D,L}] \)
and \( \hat{\rho}_{ij,t+1}^L = E[\rho_{ij,t+1} | \mathcal{F}_t^{R,L}] \), with \( \mathcal{F}_t^{D,L} \) and \( \mathcal{F}_t^{R,L} \) be volatility and correlation specific \( \mathcal{F}_t^L \), be the \( i \)-th and \( ij \)-th volatility and correlation forecasts from models using only daily data, such as the (Generalized) Autoregressive Conditional Heteroscedastic ((G)ARCH) model of Engle (1982) and Bollerslev (1986) and the Dynamic Conditional Correlation (DCC) approach of Engle (2002).

2.2 The Model

In the following we introduce the mixed-frequency (MF) approach that uses high-frequency (intradaily) data to estimate and model daily volatilities \( (D_t) \) and daily data in the model for daily correlations \((R_t)\). Based on the decomposition from Equation (2), the mixed-frequency forecast of \( \Sigma_t \) is given by:

\[
\hat{\Sigma}_{t+1|t}^{MF} = \hat{D}_{t+1|t}^H \hat{R}_{t+1|t}^L \hat{D}_{t+1|t}^H
\]

(5)

Clearly, the mixed-frequency model \( \hat{D}_{t+1|t}^L \hat{R}_{t+1|t}^L \hat{D}_{t+1|t}^L \) is also conceivable, but not of practical interest. Below we will compare the performance of the new approach against the performance of single-frequency approaches (pure low-frequency or pure high-frequency-based methods) by means of mean squared forecast error. In order to focus solely on comparing the impact of choosing a single data frequency against a mixture of frequencies to forecast covariance matrices and, thus, isolate the effect coming from comparing forecasting models, we choose to define the single-frequency approaches based on the same decomposition as given in Equation (2). More precisely, we define:

\[
\hat{\Sigma}_{t+1|t}^{LF} = \hat{D}_{t+1|t}^L \hat{R}_{t+1|t}^L \hat{D}_{t+1|t}^L
\]

(6)

\[
\hat{\Sigma}_{t+1|t}^{HF} = \hat{D}_{t+1|t}^H \hat{R}_{t+1|t}^H \hat{D}_{t+1|t}^H
\]

(7)

where Model (6) is based on daily frequency data (we will refer to this as the low-frequency (LF) model) and Model (7) is built on high-frequency data (we will refer to this as the high-frequency (HF) model). A standard implementation example for the LF model is the DCC model of Engle (2002). We note that the HF model in Equation (7) has been mentioned in Andersen et al. (2006), who also have a brief section on a version of the mixed-frequency model of Bannouh et al. (2010). However, the theoretical results that follow are not affected by the particular model specification being employed.

Before turning to the formal comparison of the three approaches above, we provide some intuition on why we believe that the mixed-frequency approach might be a valuable alternative to the single-frequency models. High-frequency data has proven to be extremely useful in the ex-post measurement of daily volatility. Nevertheless, multivariate approaches are not so well developed and suffer from difficulties associated with non-synchronous trading and dimensionality. This leads to data losses in approaches such as the multivariate kernels of Barndorff-Nielsen et al. (2011), who employ a synchronized sampling scheme or necessitates estimation of all covariances on an element-by-element basis (see, e.g., Nolte & Voev (2008) and Christensen et al.
(2010)), which does not guarantee positive-definiteness of the matrix and involves an exponentially growing number of estimations as \( n \) increases. Hautsch et al. (2010) propose a blocking approach to the kernel estimation which can, to some extent, alleviate the information loss associated with the direct estimation of a large-dimensional kernel and Lunde et al. (2011) extends this approach by means of composite kernels. Generally, it is fair to say that correlations are in practice much harder to estimate with high-frequency data compared to volatilities. Furthermore, forecasting models for realized covariance matrices, especially with many assets (more than six), are rather cumbersome and the research in this direction is still in its early stages (Bauwens et al. 2012, Hautsch et al. 2011).

The mixing approach developed in this paper is a simple alternative to the existing approaches of forecasting covariance matrices of large dimension, which extracts the informational content of HF data in the estimation of volatilities and makes use of the developed body of literature on modeling correlations with daily data. The approach is particularly attractive due to its ease of implementation compared to pure high frequency data models, since it only requires the estimation of \( n \) series of realized volatility measures (compared to a series of \( n \times n \) realized covariance/correlation matrices).

In the following, we derive and discuss the conditions under which the MF approach (5) provides smaller forecast mean squared errors (MSE) compared to the single-frequency models (6) and (7). The derivations allow a practitioner to easily verify whether a particular model of his choice satisfies the conditions for providing a superior forecast compared to another potentially simpler specification, subject to current market conditions (e.g., high-volatility) or to statistical properties of the volatility and correlation forecasts of her choice.

### 2.3 MSE Comparisons

Because we focus on one-step ahead forecasts, we are able to simplify the notation in the following manner: \( \hat{\sigma}_{ij} \equiv \sigma_{ij,t+1} \) and \( \hat{\rho}_{ij} \equiv \rho_{ij,t+1} \) for all \( i, j = 1, \ldots, n \).

We will use the representations:

\[
\hat{\sigma}_{ii} = (d_i + \epsilon_{d_i})^2 = d_i^2 + 2d_i\epsilon_{d_i} + \epsilon_{d_i}^2 \equiv \sigma_{ii} + \epsilon_{\sigma_{ii}} \quad \forall i = 1, \ldots, n
\]

\[
\hat{\sigma}_{ij} = (d_i + \epsilon_{d_i})(\rho_{ij} + \epsilon_{\rho_{ij}})(d_j + \epsilon_{d_j})
= d_i\rho_{ij}d_j + d_i\rho_{ij}\epsilon_{d_i} + d_j\rho_{ij}\epsilon_{d_i} + d_i\epsilon_{d_i}\epsilon_{\rho_{ij}} + d_j\epsilon_{d_j}\epsilon_{\rho_{ij}} + \rho_{ij}\epsilon_{d_i}\epsilon_{d_j} + \epsilon_{d_i}\epsilon_{\rho_{ij}}\epsilon_{d_j} \equiv \sigma_{ij} + \epsilon_{\sigma_{ij}} \quad \forall i \neq j, \ i, j = 1, \ldots, n
\]
with $\varepsilon_{\sigma_{ij}} \equiv 2d_i\varepsilon_{d_i} + \varepsilon_{d_i}^2$ and $\varepsilon_{\sigma_{ij}} \equiv d_i\rho_{ij}\varepsilon_{d_j} + d_j\rho_{ij}\varepsilon_{d_i} + d_i\varepsilon_{d_i}\rho_{ij} + \rho_{ij}\varepsilon_{d_i}\varepsilon_{d_j} + \varepsilon_{d_i}\varepsilon_{d_j}\rho_{ij}\varepsilon_{d_i}$. Note that the complexity of the error term $\varepsilon_{\sigma_{ij}}$ (and to a smaller extent of $\varepsilon_{\sigma_{ii}}$) arises due to the fact that the covariance (variance) forecast is a non-linear transformation of the correlation and volatility forecasts. The error term thus accounts for the noise in these original forecasts and the effect of the transformation. Note that the unbiasedness of correlation and volatility forecasts is not sufficient for unbiasedness of the covariance unless we also have that volatility and correlation forecasts are independent or at least uncorrelated.

In the following, we compare the MF, HF and LF models based on their MSE where we make use of the decomposition:

$$MSE(\hat{\sigma}_{ij}) = E[\varepsilon_{\sigma_{ij}}]^2 + V[\varepsilon_{\sigma_{ij}}].$$

(13)

The mean and variance of $\varepsilon_{\sigma_{ij}}$ are derived in Appendix [A]. The expressions in Equations (11) and (12) allow us to analyze the MSE as a function of $d_i$, $d_j$ and $\rho_{ij}$, which in turn makes it possible to draw conclusions about the models’ performance in different volatility/correlation regimes. We are now in a position to derive the conditions under which the forecast MSE of the mixing approach, $\hat{\sigma}_{ij}^{MF}$, is smaller than the forecast MSE of the single-frequency models, $\hat{\sigma}_{ij}^{MF}$ and $\hat{\sigma}_{ij}^{HF}$ for each $i,j = 1,\ldots,n$. Note that we look at elementwise MSE. Alternatively, one can define a loss based on the matrix error by means of an $L_2$ matrix norm, e.g., the Frobenius norm. A conceptually different possibility would be to consider some sort of an economic loss function, such as hedging error variance, portfolio variance, etc. An analysis of such loss functions is outside of the scope of this paper, but can be considered in a separate study.

**Proposition 1: Variance elements**

If for a given $i \in \{1,2,\ldots,n\}$, $E[(\varepsilon_{d_i}^X)^k] < \infty$, where $X \in \{H,L\}$ and $k = 2,3,4$, the difference in the MSE of the MF (HF) variance forecast $\hat{\sigma}_{ii}^{MF}$ and the LF variance forecast $\hat{\sigma}_{ii}^{LF}$ is given by

$$MSE(\hat{\sigma}_{ii}^{MF}) - MSE(\hat{\sigma}_{ii}^{LF}) = 4(E[(\varepsilon_{d_i}^H)^2]\varepsilon_{d_i}^L)^2 + 2(E[(\varepsilon_{d_i}^H)^3] - E[(\varepsilon_{d_i}^L)^3])d_i + (E[(\varepsilon_{d_i}^H)^4] - E[(\varepsilon_{d_i}^L)^4]).$$

(14)

A set of sufficient (but not necessary) conditions for $MSE(\hat{\sigma}_{ii}^{MF}) \leq MSE(\hat{\sigma}_{ii}^{LF})$ is that $E[(\varepsilon_{d_i}^H)^2] - E[(\varepsilon_{d_i}^L)^2] \leq 0$, $E[(\varepsilon_{d_i}^H)^3] - E[(\varepsilon_{d_i}^L)^3] \leq 0$ and $E[(\varepsilon_{d_i}^H)^4] - E[(\varepsilon_{d_i}^L)^4] \leq 0$. The minimal sufficient conditions, i.e., the necessary and sufficient conditions for the inequality to hold are clearly weaker and are provided in Appendix [A].

In Proposition [1](as well as for the results further below) apart from the existence of certain moments, we do not make any distribution, unbiasedness, etc. assumptions on the forecast (errors). This is theoretically appealing, since this allows us to specialize the expressions should we desire to make such assumptions (e.g., that errors have a

\[\text{Note that we do not compare variance forecast from the MF model to the HF model since they are identical by construction.}\]
zero expectation and/or are normally distributed). Furthermore, models for volatility forecasting often employ specifications for a transformation of volatility (e.g., log realized volatility). Strictly speaking, naïve forecasts via re-transforming the forecast are biased, although they might be empirically superior to bias-corrected forecasts (see e.g., Chiriac & Voev (2011)). Also, GARCH models typically model the variance, but there are specifications for the volatility (in its narrow meaning as standard deviation) as well. Generally, it is difficult to take a stand on the unbiasedness of variance/volatility forecasts and therefore we do not assume that. This, of course, does not mean that we do not allow for it; we simply let the data “tell” us.

The conditions we provide, reveal that it is sufficient for the MF (HF) approach to outperform the LF model if the second, third and fourth (uncentered) moments of the HF volatility errors are smaller than their LF counterparts. The conditions on the second and fourth moment are likely to be satisfied, since the basic motivation of using HF data is that it helps in measuring and forecasting volatility more precisely than with daily data.

In the following proposition, we derive and compare MSE’s of covariance elements for the MF, LF and HF models. As before, we make minimal assumptions on the moments and dependence structure of the forecast errors.

**Proposition 2: Covariance elements**

If, for a given $i \neq j$, $i, j \in \{1, 2, \ldots, n\}$ and $k, l, m = 0, 1, 2$, it holds that

(i) $E[(\varepsilon^H_{d_i})^k(\varepsilon^H_{d_j})^l(\varepsilon^L_{\rho_{ij}})^m] < \infty$

(ii) $E[(\varepsilon^L_{d_i})^k(\varepsilon^L_{d_j})^l(\varepsilon^L_{\rho_{ij}})^m] < \infty$

(iii) $E[(\varepsilon^H_{d_i})^k(\varepsilon^H_{d_j})^l(\varepsilon^H_{\rho_{ij}})^m] < \infty$

then it follows that

1. (mixed frequency to low frequency model comparison)

$$MSE(\hat{\sigma}_{ij}^{MF}) - MSE(\hat{\sigma}_{ij}^{LF}) = F_1(d_i, d_j, \rho_{ij}),$$

where $F_1(d_i, d_j, \rho_{ij})$ is a fourth-order polynomial in $d_i$, $d_j$ and $\rho_{ij}$ as given in Equation (40) in Appendix A. If $\rho_{ij} \geq 0$, sufficient (but not necessary) conditions that $MSE(\hat{\sigma}_{ij}^{MF}) \leq MSE(\hat{\sigma}_{ij}^{LF})$ are that all parameters of $F_1(d_i, d_j, \rho_{ij})$ are non-positive (see Proof A.2 in Appendix A).

2. (mixed frequency to high frequency model comparison)

$$MSE(\hat{\sigma}_{ij}^{MF}) - MSE(\hat{\sigma}_{ij}^{HF}) = F_2(d_i, d_j, \rho_{ij}),$$

where $F_2(d_i, d_j, \rho_{ij})$ is a fourth-order polynomial in $d_i$, $d_j$ and $\rho_{ij}$ as given in Equation (41) in Appendix A. If $\rho_{ij} \geq 0$, sufficient (but not necessary) conditions that $MSE(\hat{\sigma}_{ij}^{MF}) \leq MSE(\hat{\sigma}_{ij}^{HF})$ are that all parameters of the $F_2(d_i, d_j, \rho_{ij})$ polynomial are non-positive (see Proof A.2 in Appendix A).
In the proposition above the MSE’s of the covariance forecasts are expressed as functions of the true ex-post values of the volatility and correlation. Given the minimal assumptions we make, the expressions are rather lengthy and defy straightforward interpretation. Nevertheless, under the empirically relevant assumption that \( \rho_{ij} \geq 0 \), some intuition of the sufficient conditions stated above can be provided. For the MF to outperform the LF model, it is sufficient but not necessary, that

1. Second moments of volatility forecast errors are smaller for the HF forecast compared to the LF forecast;

2. Cross moments up to order six (see the expressions for \( B_{ij} \) in Appendix A) of forecast errors for the MF model (using HF volatility and LF correlation forecasts) are smaller than those of forecast errors for the LF model (using LF volatility and correlation forecasts).

The first condition is intuitive and likely to hold. We let the data reveal whether the second condition is satisfied. Intuitively this condition will hold, if HF volatility and LF correlation errors are in some sense less dependent than LF volatility errors and LF correlation errors. In any case, it is important to emphasize that these are not necessary conditions; the MF model can outperform the LF model even when they fail.

Similar arguments can be used to analyze the MF to HF model comparison. In particular, for the MF to outperform the HF model, it is sufficient but not necessary, that

1. Second moments of correlation forecast errors are smaller for the LF forecast compared to the HF forecast;

2. Cross moments up to order six (see the expressions for \( C_{ij} \) in Appendix A) of forecast errors for the MF model (using HF volatility and LF correlation forecasts) are smaller than those of forecast errors for the HF model (using HF volatility and correlation forecasts).

Note that while the first condition is likely to fail empirically (at least in small dimensions when the HF model is feasible), the conditions above are not necessary, i.e., we can still have \( MSE(\hat{\sigma}_{ij}^{MF}) \leq MSE(\hat{\sigma}_{ij}^{HF}) \) even if they are violated. Thus the MF method can still provide in practice better or equally good forecasts as the HF method as long as the effect of having better HF than LF correlation forecasts is smaller compared to the effect of reducing the dependency among the volatility and correlation forecast errors by considering mixed-frequency over single-frequency (HF) data. The empirical results presented in Section 3 prove this fact: the performances of MF and HF approaches in forecasting covariances are statistically equal (not significantly

Note that this assumption is needed just for the ease of interpreting the sufficient conditions. However, although not impossible, the proportion of negative correlations in the stock market is very small; in our empirical exercise, negative correlation days are about 3% of the total sample. Moreover, the positive correlation assumption is particularly relevant for samples including high volatility periods, due to the empirically observed volatility-in-correlation effect (high volatility induces high correlation) typical to stock markets.
different MSE’s) during the whole sample and MF outperforms HF during very volatile financial times.

The usefulness of Proposition 2 is that it gives us the exact form of the MSE difference as a function of the variables \(d_i, d_j\) and \(\rho_{ij}\). Since the parameters of the polynomials \(F_1\) and \(F_2\) are easily estimated from data (simply as sample counterparts of the population moments), we can use the results in Proposition 2 to analyze how the relative model performance behaves in various volatility/correlation environments. To approach this issue more formally, we compute partial derivatives of \(F\) with respect to \(d_i\) and \(d_j\) which are reported below:

\[
\frac{\partial F_1(d_i, d_j, \rho_{ij})}{\partial d_i} = 2B_{1,i}d_i \rho_{ij}^2 + 2B_{3,i}d_i + B_{6,i}d_i \rho_{ij}^2 + 2B_{7,i}d_i d_j \rho_{ij} + 2B_{18,i}d_i \rho_{ij} + B_{19,i}
\]

\[
\frac{\partial F_1(d_i, d_j, \rho_{ij})}{\partial \rho_{ij}} = 2B_{1,i}d_i^2 \rho_{ij} + 2B_{2,i}d_j \rho_{ij} + 2B_{5,i}d_i \rho_{ij} + 2B_{6,i}d_i d_j \rho_{ij} + B_{7,i}d_i^2 d_j
\]

\[
\frac{\partial F_2(d_i, d_j, \rho_{ij})}{\partial d_i} = 2C_{1,i}d_i \rho_{ij} + 2C_{2,i}d_i d_j + 2C_{4,i}d_i d_j \rho_{ij} + 2C_{5,i}d_i \rho_{ij} + C_{6,i}d_j \rho_{ij}
\]

\[
\frac{\partial F_2(d_i, d_j, \rho_{ij})}{\partial \rho_{ij}} = 2C_{4,i}d_i^2 \rho_{ij} + 2C_{5,i}d_i^2 \rho_{ij} + 2C_{7,i}d_i d_j + C_{8,i}d_i^2 d_j + C_{10,i}d_j + C_{16,i}
\]

where the expressions for B’s and C’s are given in the appendix.

The implications are as follows:

1. If all B’s in expression (17) are non-positive and \(\rho_{ij} \geq 0\), then \(\frac{\partial F_1(d_i, d_j, \rho_{ij})}{\partial d_i} \leq 0\), i.e., the larger \(d_i\) (\(d_j\)), the better the MF model compared to the LF model. If \(\rho_{ij} < 0\), the sign of the derivative is undetermined.

2. If all B’s in expression (18) are non-positive and \(\rho_{ij} \geq 0\), then \(\frac{\partial F_1(d_i, d_j, \rho_{ij})}{\partial \rho_{ij}} \leq 0\), i.e., the larger \(\rho_{ij}\), the better the MF model compared to the LF model. If \(\rho_{ij} < 0\), the sign of the derivative is undetermined.

3. If all C’s in expression (19) are non-positive and \(\rho_{ij} \geq 0\), then \(\frac{\partial F_2(d_i, d_j, \rho_{ij})}{\partial d_i} \leq 0\), i.e., the larger \(d_i\) (\(d_j\)), the better the MF model compared to the HF model. If \(\rho_{ij} < 0\), the sign of the derivative is undetermined.

4. If all C’s in expression (20) are non-positive, then \(\frac{\partial F_2(d_i, d_j, \rho_{ij})}{\partial \rho_{ij}} \leq 0\), i.e., the larger \(\rho_{ij}\), the better the MF model compared to the HF model for all \(\rho_{ij}\).
In order to give an intuitive description of the sufficient conditions above, one should remember that it is a widely-observed empirical fact that high volatility induces high correlations (i.e., volatility-in-correlation effect) and that when $d_i$ is large, $\rho_{ij}$ is also large. Our theoretical expressions from above confirm this empirical fact by the large proportion of common B’s in the expressions (17) and (18) and common C’s in the expressions (19) and (20). Moreover, by summing up the non-negativity conditions stated above, we observe that they are identical (with the exception of three B’s and two C’s) to the sufficient conditions of Proposition 2. Consequently, we can expect that the MF method outperforms the LF and HF approaches during high volatility/correlation periods. In a matter of fact, the empirical results presented in Section 3 show that MF outperforms LF and HF methods during very turbulent financial times (high volatility and correlation).

As we mentioned already, our focus in this paper is to provide conditional one-period ahead covariance matrix forecasts. Within our framework, a multi-step forecast can be easily obtained from mixing forecasts of volatility and correlation similar to Equation (2) as follows:

$$\hat{\Sigma}_{t:t+h|t} = \hat{D}_{t:t+h|t} \hat{R}_{t:t+h|t} \hat{D}_{t:t+h|t}. $$

However the theoretical properties derived above are not directly applied. If the multi-step forecasts are obtained from a “direct” procedure, by aggregating data over $h$ periods, then the theoretical results derived so far will hold as long as the finite moment assumptions from propositions 1 and 2 hold for the aggregated forecast errors. However, if the volatility and correlation multi-step forecasts stem from an iterated procedure, where the $h$-step ahead forecasts are computed from aggregating $h$ one-period forecasts, in order to derive the theoretical results, one has to account for the joint distribution of the $h$ one-step forecast errors. To illustrate this issue, we provide an example with $h = 2$ (cf. Equation (3) above):

$$\hat{\sigma}_{ii,t:t+2|t} = \hat{d}_{i,t:t+2|t}^2 $$

$$\hat{d}_{i,t:t+2|t} = \hat{d}_{i,t:t+1|t} + \hat{d}_{i,t+1:t+2|t} = d_{i,t:t+1} + \varepsilon_{d_{i,t:t+2|t}} + \varepsilon_{d_{i,t+1:t+2|t}} + \varepsilon_{d_{i,t:t+2|t}}. $$

Now we would need to account for the properties of the aggregated forecast error, $\varepsilon_{d_{i,t:t+2|t}}$, which imply conditions on the time-variation in the moments and cross moments of the daily forecast errors as well as for their serial correlation and cross-correlation. We refrain from deepening this analysis in the current study, by letting it for further research.

3 Empirical Application

In this section, we present results on the mixing and single-frequency multivariate volatility forecasting approaches presented in Section 2. As volatility is not observable, we use a realized covariance proxy in our evaluation. We note that the MSE is a loss
function which satisfies the conditions in Patton (2011) of being robust to the noise in the volatility proxy. To be consistent with the theoretical framework and results described in Section 2, we compare the forecast performance of the MF approach against LF and HF approaches solely by means of their mean squared forecast errors. It would be also interesting to evaluate their performance by some economic criteria, such as within an optimal portfolio application. However, because such an evaluation is not of interest to the present paper, we leave it for further research.

3.1 Data

The data consists of tick-by-tick transaction prices from the Trade and Quotations (TAQ) database sampled from 9:30 until 16:00 for the period 11.10.2001 – 01.06.2009 ($T = 1915$ trading days). For the current analysis, we select a total of 30 liquid assets, from which 29 stocks were in the past or are currently constituents of the Dow Jones Industrial Average (DJIA), and the Spider Trust Series I (SPY), which is an exchange traded fund that tracks the performance of the S& P 500 index. We employ the previous-tick interpolation method, described in Dacorogna et al. (2001) and obtain 78 intraday returns by sampling every 5 minutes and one daily return. For estimation, we multiply the returns by 100, i.e., we consider percentage returns. For each $t = 1, \ldots, 1915$, a series of daily realized covariance matrices can be constructed as:

$$RCov_t = \sum_{j=1}^{M} r_{j,t} r_{j,t}^{\prime}$$  

where $M = 78$. The 5-minute returns, $r_{j,t}$, are computed as

$$r_{j,t} = p_{j\Delta,t} - p_{(j-1)\Delta,t}, \quad j = 1, \ldots, M$$

where $\Delta = 1/M$ and $p_{j\Delta,t}$ is the log transaction price at time $j\Delta$ on day $t$. The realized covariance matrices are symmetric by construction and, for $n < M$, positive definite almost surely. In order to enhance the precision, we refine the estimator by subsampling. As in Chiriac & Voev (2011), we construct 30 regularly $\Delta$-spaced subgrids ($\Delta = 300$ seconds) starting at seconds 1, 11, 21, $\ldots$, 291, compute the realized covariance matrix on each subgrid and take the average. The resulting subsampled realized covariance is much more robust to the so called market microstructure noise than the simple 5-minute based one and given the high liquidity of all stocks, we are confident that the effect

---

5 We are grateful to Asger Lunde for providing us with the data.
6 The stocks are: Alcoa Inc. (AA), American International Group Inc. (AIG), American Express Company (AXP), Boeing Corporation (BA), Bank of America Corporation (BAC), Citigroup Inc. (C), Caterpillar Inc. (CAT), Chevron Corporation (CVX), Dupont (DD), Walt Disney Company (DIS), General Electric Company (GE), General Motors Corporation (GM), Home Depot Inc. (HD), International Business Machines (IBM), Intel (INTC), Johnson & Johnson (JNJ), J.P. Morgan Chase & Company (JPM), Coca-Cola Company Kraft Foods Inc. (KO), McDonald’s Corporation (MCD), 3M Company (MMM), Merck & Company Inc. (MRK), Microsoft (MSFT), Pfizer Inc. (PFE), Procter & Gamble Company (PG), AT&T (T), United Technologies Corporation (UTX), Verizon Communications Inc. (VZ), Wal-Mart Stores Inc. (WMT), Exxon Mobil Corporation (XOM)
of non-synchronicity is rather mild at the chosen frequency. Moreover, we avoid the noise induced by measuring the overnight volatility as the squared overnight return, by applying the open-to-close data and measuring the volatility over the trading session. Consequently, the daily return is also computed as from the open-to-close prices. Table B.1 in Appendix B reports summary statistics for the realized variances and covariances averaged over all 30 stocks and, respectively, over all 435 pairs of stocks as well as the summary statistics of daily returns averaged over all stocks. For both realized variance and covariance measures, we observe typical stylized facts such as extremely right skewness and leptokurtosis (Andersen, Bollerslev, Diebold & Ebens, 2001). The daily returns exhibit skewness close to zero and overkurtosis.

The 30 daily realized variance series are given by the diagonal elements of the realized covariance (RCov) matrix defined above. Please note the distinction we make here between realized variance (RV) and its square root, for which we use the term realized volatilities (RVol). The series of daily realized correlation matrices RCorr are computed from RCov in the usual way.

3.2 Forecasting Models

In this section we elaborate on the implementation of the three forecasting approaches introduced in Section 2. The choice of the models implemented here is broad and, in general, motivated by various theoretical or empirical results reported by the related literature on forecasting daily volatilities and correlations based on daily and high frequency data. For both volatility and correlation forecasts we opt not just for one model choice, but for a series of models in order to let data “tell” which is the best model and mixing strategy and to make our results robust. As already mentioned, the choice of the models is ultimately an empirical question. Therefore, the present empirical exercise should be regarded as a simple illustrative example of the theoretical results presented in Section 2.

The mixed-frequency (MF) model

The covariance matrix forecast from the MF approach is given by:

\[ \hat{\Sigma}_{t+1|t}^{MF} = \hat{D}_{t+1|t}^H \hat{R}_t^L \hat{D}_{t+1|t}^H, \]

where \( \hat{D}_{t+1|t}^H = \text{diag}(RVol_{1,t+1|t}, \ldots, RVol_{30,t+1|t}) \) and \( RVol_{i,t+1|t}, i = 1, \ldots, 30 \) are one-day ahead realized volatility forecasts. We choose to follow the main body of literature (e.g., Andersen, Bollerslev, Diebold & Labys, 2001, Andersen et al., 2000, among others) and to apply an autoregressive framework that is able to capture the strong persistence (long memory) of daily log-realized volatility series evident in the slowly decaying autocorrelation functions in Figure B.1 in Appendix B. Thus we apply for each \( i = 1, \ldots, 30 \) the following ARFIMA(\( p,d,q \)) model:

\[ \Phi_i(L)(1 - L)^d X^*_i,t = \Theta_i(L) \varepsilon_{i,t}, \quad \varepsilon_{i,t} \sim N(0, \omega_i), \]

One may also choose the HAR model of Corsi (2009). However, Chiriac & Voev (2011) show that there are no significant forecasting gains from choosing the HAR models over the ARFIMA models.
where $X^*_{i,t}$ are the demeaned series of daily log-$RVol_i$, $\Phi_i$ and $\Theta_i$ are the AR and MA lag polynomials and $d_i$ is the parameter of fractional integration. We implement for each demeaned log-$RVol_i$ series the following fractional integrated model specifications: ARFIMA$(1,d,1)$, ARFIMA$(1,d,0)$ and ARFIMA$(0,d,0)$. Specific to the model choice, the $X^*_{i,t}$ series are stationary and invertible as long as $d_i < 0.5$ and the roots of $\Phi_i(L)$ and $\Theta_i(L)$ lie outside the unit circle. We estimate all models by the maximum likelihood approach described in Beran (1995).

$\hat{R}_{t+1|t}$ is the one-day ahead correlation matrix forecast, which we derive from the dynamic correlation (DCC) approach of Engle (2002) estimated on daily data as follows:

$$
\begin{align*}
R_t^L &= (\text{diag}(Q_t^L))^{-\frac{1}{2}}Q_t^L(\text{diag}(Q_t^L))^{-\frac{1}{2}} \\
Q_t^L &= (1 - \theta_1 - \theta_2)\bar{Q}L + \theta_1 u_{t-1}'u_{t-1} + \theta_2 Q_{t-1}^L,
\end{align*}
$$

(27)

where $u_t$ is the vector of de-volatilized residuals with elements

$$
u_{i,t} = \frac{\epsilon_{i,t}}{RVol_{i,t}}, \quad i = 1, \ldots, 30,$$

and $\bar{Q}$ is the unconditional covariance of $u_t$. Furthermore, we assume that the conditional mean of daily returns is constant, $r_{i,t} = E[r_{i,t} | \mathcal{F}_{t-1}] + \epsilon_{i,t} = \mu_i + \epsilon_{i,t}$ and estimate the model in Equation (27) on the demeaned series of daily returns. Note that we standardize the daily returns here by realized volatilities, rather than by GARCH volatilities as in the standard implementation of DCC. In the theoretical section of the paper, we treat correlation errors in the MF and the LF model as identical. In fact, the standardization by $RVol$ is likely to improve upon the correlation model and is a secondary channel through which HF data leads to improvements. In this sense, the theoretical results on the conditions for the MF model to outperform the LF model are too conservative since they do not take into account these additional gains. We estimate Equation (27) by the maximum likelihood approach described in Engle (2002). However, given that there is empirical evidence (Engle et al. (2009)) that the maximum likelihood estimator is severely biased when estimated on correlation matrices of high dimensions ($n$ larger than 20), we also apply different techniques proposed in the literature to reduce the bias, such as: the consistent DCC (cDCC) approach of Aielli (2011) together with the three shrinkage approaches of Hafner & Reznikova (2012), for which the targets are: the identity matrix (cDCC-I), the single factor model (cDCC-M) and the equicorrelation matrix (cDCC-E) as well as with the composite maximum likelihood method (cDCC-CL) proposed by Engle et al. (2009).

The low frequency (LF) model

The covariance matrix forecasts with daily data are obtained in this empirical application with the DCC model of Engle (2002):

$$
\Sigma_{t+1|t}^{LF} = \hat{D}_{t+1|t}^L\hat{R}_{t+1|t}^L\hat{D}_{t+1|t}^L
$$

(28)
where $\hat{D}_{t+1|t} = \text{diag}(h_{1,t+1|t}^{1/2}, \ldots, h_{30,t+1|t}^{1/2})$ and $h_{i,t+1|t}$ are one-day ahead forecasts computed from univariate GARCH(1,1) models:

$$h_{i,t} = w_i + \alpha_i \epsilon_{i,t-1}^2 + \beta_i h_{i,t-1} \quad \forall i = 1, \ldots, 30$$

(29)

with $w_i, \alpha_i, \beta_i \geq 0$ and $\alpha_i + \beta_i < 1, \forall i = 1, \ldots, 30$. The one-day ahead correlation forecast $\hat{R}_{t+1|t}$ is given in Equation (27), however, importantly, the standardized (devolatilized) residuals are now given by

$$u_{i,t} = \frac{\epsilon_{i,t}}{\sqrt{h_{i,t}}}, \quad i = 1, \ldots, 30.$$

We estimate the DCC model by the two-step maximum likelihood approach described in [Engle (2002)]. Similar to the mixed-frequency approach presented above, we implement further enhancements of the DCC model which aim at correcting for the bias induced by the maximum likelihood estimation when applied on matrices of large dimensions.

**The high frequency (HF) model**

The HF covariance matrix forecasts are given by:

$$\hat{\Sigma}_{t+1|t}^{HF} = \hat{D}_{t+1|t}^{HF} \hat{R}_{t+1|t}^{HF \ H} \hat{D}_{t+1|t}^{HF}$$

(30)

where $\hat{D}_{t+1|t}^{HF}$ are obtained in the same manner as in the MF model and $\hat{R}_{t+1|t}^{HF}$ is obtained here from the following autoregressive framework:

$$\hat{R}_{t+1|t}^{HF} = \left(1 - \sum_{l=1}^{t} \lambda_l\right) \bar{RCorr} + \sum_{l=1}^{t} \lambda_l \tilde{RCorr}_{t-l+1},$$

(31)

where $\bar{RCorr} = \frac{1}{t} \sum_{i=1}^{t} RCorr_i$, $\tilde{RCorr}_t = RCorr_t - \bar{RCorr}$ and $\lambda_l$ is the sequence of coefficients of a pure AR-representation of the following fractional integrated vector ARFIMA($0,d,0$) process:

$$D(L)Y_t = \zeta_t, \quad \zeta_t \sim N(0, \Omega),$$

(32)

where $Y_t$ is the vector obtained by stacking the lower triangular portion of $\tilde{RCorr}_t$ without the main diagonal and $D(L) = (1 - L)^d I_m$, where $m$ is the number of correlation series $m = n(n-1)/2$. The choice of the long memory process is due to the persistence of the 435 series of daily realized correlations (Figure B.2 in Appendix B depicts the autocorrelation function of 30 randomly chosen realized correlation series). Due to the large dimension of the $Y_t$ vector ($m = 435$), we also consider a simple short memory vector AR process, as follows:

$$Y_t = \psi Y_{t-1} + \xi_t, \quad \xi_t \sim N(0, \Upsilon),$$

(33)

One can apply different specifications of GARCH. However, as [Hansen & Lunde (2005)] show, the GARCH(1,1) is hard to beat within the GARCH class of models.
where \(-1 < \psi < 1\) in order to assure stationarity of the \(Y_t\) process. It remains an empirical question whether a fractional autoregressive or a simple autoregressive model is better for forecasting realized correlations. We present empirical results for both vector ARFIMA\((0,d,0)\) and vector ARFIMA\((1,0,0)\) models. We note that the model specifications in equations (32) and (33) are clearly not inferior (most likely even superior) in terms of flexibility to the DCC model.

### 3.3 Forecast Evaluation

We split the whole sample of data into an in-sample period from 11.10.2000 to 06.06.2007 (1414 days) and an out-of-sample period from 07.06.2007 to 01.06.2009 (500 days). The forecasts are carried out in a recursive manner, i.e., at each step the models are re-estimated with all of the available data. Due to the large heterogeneity in the variances of stocks and the presence of extreme outliers in the evaluation period, we opt for reporting median rather than mean values across stocks and stock pairs.

<table>
<thead>
<tr>
<th>Model for forecasting</th>
<th>Model for forecasting correlations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>cDCC</td>
</tr>
<tr>
<td>MF</td>
<td>ARFIMA((0,d,0))</td>
</tr>
<tr>
<td>MF</td>
<td>ARFIMA((1,d,0))</td>
</tr>
<tr>
<td>MF</td>
<td>ARFIMA((1,d,1))</td>
</tr>
<tr>
<td>LF</td>
<td>GARCH</td>
</tr>
<tr>
<td>HF</td>
<td>ARFIMA((0,d,0))</td>
</tr>
<tr>
<td>HF</td>
<td>ARFIMA((1,d,0))</td>
</tr>
<tr>
<td>HF</td>
<td>ARFIMA((1,d,1))</td>
</tr>
</tbody>
</table>

Table 1: Median value of the MSE’s over all (465) elements of the covariance matrix.

For a fair comparison, we consider additionally the model where the cross product of standardized daily returns \(u_{t-1}u'_{t-1}\) in Equation (27) is replaced by the observed matrices of daily realized correlations:

\[
R^M_t = (1 - \delta_1 - \delta_2)\hat{R}_{Corr} + \delta_1 R_{corr_{t-1}} + \delta_2 R^M_{t-1},
\]

This approach has the same decomposition structure as MF and LF with the difference of replacing the outer product of daily standardized returns by the \(R_{Corr}\). This is likely to improve upon the correlation model as a whole, given that the \(R_{Corr}\) is a better estimator of daily correlations than the outer product of daily standardized returns. This model is very much in the spirit of the multivariate High-Frequency-Based Volatility (HEAVY) model proposed by Noureldin et al. (2012), which instead apply a BEKK-type representation to obtain the correlation matrix forecasts. Although the theoretical results on the conditions of the MF model cannot be directly applied to this approach due to the nonlinear way of mixing daily (in the estimation likelihood) and intraday information (in \(R_{Corr}\)), one still may consider it as a natural extension of the newly introduced MF method. However, given that the empirical results do not significantly differ from the HF approach, we refrain from reporting them here. They may be obtained by request from the authors.
variance and 435 covariance forecasts. Table 1 shows that the MF and HF approaches provide comparable MSE’s, both smaller than the MSE of the LF model. For both MF and HF models, among the approaches considered in the present empirical exercise, the ARFIMA(1,d,1) model for forecasting volatilities seems to play an important role in providing qualitatively good forecasts (all smaller than in the LF approach). Among the models for forecasting correlations based on low frequency data, the cDCC-CL approach is the method that improves the forecast quality of both MF and LF methods. The VARFIMA(0,d,0) for forecasting realized correlations exhibits in general a good performance within the HF framework. For comparing the MSE results, we implement the model confidence set (MCS) of Hansen, Lunde & Nason (2011) which selects the set of models containing the best one with a certain degree of confidence. Across all elements of the covariance matrices, the MF models are in the 95% MCS between 6.4% (ARFIMA(0,d,0) for volatility and cDCC-I for correlation forecasts) and 54% (ARFIMA(1,d,1) for volatility and cDCC-CL for correlation forecasts) of the cases, the LF models between 2% (cDCC forecasts) and 31% (cDCC-CL forecasts) of the cases and the HF models between 4% (ARFIMA(0,d,0) for volatility and VARFIMA(1,0,0) for correlation forecasts) and 69% (ARFIMA(1,d,1) for volatility and VARFIMA(0,d,0) for correlation forecasts) of the cases.

An interesting analysis for our purposes is to see how the variance and covariance forecasts differ among the three approaches. Table B.2 in Appendix B provides the median values of MSE’s of the variance forecast errors for the MF (HF) and LF approaches. Due to the precise measurement of volatilities based on high frequency data and the long memory specification, all MF (HF) approaches provide a clear improvement in the variance forecast when compared to the LF model, which uses daily data and the GARCH variance specification. The MF (HF) models are contained in the 95% MCS for 80% of the stocks, while the LF models are in the set in only 26% of the cases. Among the HF-based models, the ARFIMA(1,d,0) and ARFIMA(1,d,1) variance forecasts are in the 95% MCS in 90% of the cases, while the ARFIMA(0,d,0) is in the 95% MCS for 70% of the 30 stocks.

The MF variance forecasts are subject to bias when obtained naïvely from log-volatility forecasts. Based on the results of Chiriac & Voev (2011), Lütkepohl & Xu (2012) and our empirical findings, that show that the bias correction increases the MSE of variance forecasts regardless of the model choice, we opt for no bias correction.

The MSE results for the covariances are given in Table B.3. Similar to the results presented in Table 1, the MF and HF methods perform equally well: the ARFIMA(1,d,1) combined with cDCC-CL in the MF framework and with VARFIMA(0,d,0) in the HF framework provides the smallest MSE’s among all models. These approaches are in the 95% MCS for 51% and 67% of the stock pairs, respectively. The LF models are in the MCS in only 8% of the cases.

The fact that the MF and HF models provide comparatively equal performances is very encouraging, especially if we consider larger dimensions. Besides the good forecasting power, the MF models are much easier to implement, given their straightforward correlation specification. The difficulty in the implementation of the HF models

---

10We implement the Ox package MulCom v2.00 provided by the authors.
increases with the dimension of the system due to issues of non-synchronicity and parameter proliferation. Thus, the results we provide should be viewed in the light of the considerable simplicity of the MF model compared to the HF models.\footnote{Although deviating from the decomposition structure given in Equation (2), we also consider for comparison purposes the HF models of Chiriac & Voev (2011) and Hautsch et al. (2011). However, given that the empirical results do not significantly differ from the HF approach considered here, we refrain from reporting them. They may be obtained by request from the authors.}

Moreover, if we have a detailed look at the results and consider in particular the most turbulent period on the stock market following the Lehman Brothers’ bankruptcy (16.09.2008-23.10.2008\footnote{In this period the largest realized volatility within the whole sample (around 80% on 16.09.2008) as well as the largest realized correlation (0.95 on 23.10.2008) registered. During this period, one can also observe the volatility-in-correlation effect: on 16.09.2009 the largest realized correlation among the stocks was around 0.90, while on 23.10.2008 the largest realized volatility was around 67%.}), our results show that in fact the MF provides the smallest median MSE across all series – of around 262.41 given by the combination ARFIMA(1,d,1) and cDCC-I), while the smallest median MSE stemming from HF approaches is about 270.80 (from the combination ARFIMA(1,d,1) - VARFIMA(0,d,0)). In what regards each series, as expected, the variance is better forecasted by a HF- than by a LF-based model (ARFIMA(1,d,1) with MSE of around 657.66 vs. GARCH(1,1) with MSE of 795.92). However, the covariances are better forecasted by the MF approach (the smallest median MSE is about 257.89 and stems from the combination ARFIMA(1,d,1) - VARFIMA(0,d,0)). Thus the MF approach, besides having a significant computational advantage, also outperforms the HF method during this extremely turbulent financial period.

In the following, we report some empirical results on the necessary and sufficient conditions in propositions 1 and 2. For this purpose, we choose from each of the mixing frequency approaches the model with the smallest MSE in Table 1 (the bold entries). The median estimate of the parameters $A_{1,i} = E[(\varepsilon_{H,i}^d)^2] - E[(\varepsilon_{L,i}^d)^2]$, $A_{2,i} = E[(\varepsilon_{H,i}^d)^3] - E[(\varepsilon_{L,i}^d)^3]$ and $A_{3,i} = E[(\varepsilon_{H,i}^d)^4] - E[(\varepsilon_{L,i}^d)^4]$ from Equation (14) are $-0.2338$, $0.2600$, and $-3.0429$, respectively. Although the median values of the three parameters do not fulfill the sufficient conditions of Proposition 1 (to be negative), they fulfill the necessary conditions presented in Appendix A (see case 3.1 in Proof A.1). Stock-wise, parameter $A_2$ is either negative (8 of 30 stocks) or very close to zero (22 of 30 stocks). Thus, the sufficient conditions of Proposition 1 are fulfilled in 8 cases, while the necessary conditions (case 3) are valid in the remaining 22 cases. The empirically valid necessary conditions are very intuitive: on average, one expects a gain in the variance (efficiency) and in the kurtosis of volatility forecasts stemming from HF data compared to LF data, while there are no particular reasons for asymmetric forecast errors (skewness almost zero). The sign and relative magnitude of the estimated parameters confirm these expectations.

In Figure 1 we plot the median value of the polynomial in Equation (14) for a wide range of values for $d_i$, i.e, $d_i = [0,10]$, where the upper limit is chosen to correspond approximately to an extreme value of 160% annualized volatility along with the median 95% confidence intervals. In order to avoid any distributional assumptions...
on the forecast errors, we opt for approximating the confidence intervals by the bootstrapped percentile intervals as described in [Efron & Tibshirani (1993), pp. 170-173]. We implement the stationary bootstrap procedure of [Politis & Romano (1994b)], which accounts for some degree of serial correlation of the forecast errors. To account for cross correlation among different forecast errors, we bootstrap them in groups.

One may observe that for all values of $d$, the parabola as well as the confidence bounds lie below the zero line, which indicates that, for most of the stocks, the MF (HF) models provide significantly smaller mean squared forecast errors than the LF models. Moreover, one may observe the left asymmetry of the confidence intervals around the median value of the estimated polynomial, especially for the small values of the true volatility. This indicates that we may have a larger probability of a negative MSE difference for values of $d_i$ close to zero. Thus, while in a few cases the difference between the MSE of MF (HF) and LF variance forecast errors is minimal, in most of the cases, for the forecast models chosen in this empirical application, the MF (HF) models will provide significantly smaller forecast errors than the LF approach.

![Figure 1](image.png)

**Figure 1:** Median value of $\hat{MSE}(\hat{\sigma}_{MF}^{d_i}) - \hat{MSE}(\hat{\sigma}_{LF}^{d_i}) = 4\hat{A}_i d_i^2 + 4\hat{B}_i d_i + \hat{C}_i$. The median value of the polynomial is given by the thick solid line. The dotted lines represent the 95% confidence intervals and the simple solid line represents the zero bound. On X-axis we plot the true standard deviation, $d_i$ that ranges from 0 to 10.

Next, we verify empirically the conditions of Proposition $\hat{2}$ on the MSE inequalities of MF and LF covariance forecasts. Both first and second sufficient conditions of the first part of Proposition $\hat{2}$ are confirmed empirically for most of the stock pairs, which indicates that the chosen volatility and correlation forecast models improve the forecasting power of the MF model compared to LF model. More precisely, the first sufficient condition that states that the second moments of volatility forecast errors is smaller for the HF forecast compared to the LF forecast is mirrored in the fact that the median values of the related parameters in Equation (40), namely $B_1$, $B_2$, $B_5$, $B_6$, and $B_7$,
\( B_6, B_{10} \) and \( B_{14} \), are negative. The second sufficient condition that states that the dependencies (cross-moments) among HF volatility and LF correlation forecast errors are smaller than the ones among LF volatility and LF correlation forecast errors is empirically confirmed – the median value of the associated parameters in Equation (40) is on average also negative.

For a detailed analysis with respect to the market conditions, we plot in Figure 2 the median estimate of the polynomial from Proposition 2, \( \hat{F}_1(d_i, d_j, \rho_{ij}) \) along with the median 95\% percentile-bootstrapped confidence intervals for \( d_i = [0, 10] \) and \( d_j = [0, 10] \) and for different values of \( \rho_{ij} \). The choice of \( \rho_{ij} \) is motivated by the descriptive statistics of the realized correlations over the whole window: on average, the daily realized correlations among all stocks is around 0.32, the maximum is 0.95 and the minimum is –0.65. As one may notice, compared to the LF model, the MF model provides significantly smaller MSE in particular for large values of volatilities and all values of correlation (the surfaces lie below the zero bound). This indicates that the MF approach is especially attractive in a highly volatile environment. During volatile periods, the correlations also tend to be large. In this scenario, the MSE gains of the MF model are even stronger (see the right panel of Figure 2). Generally, we observe that for positive and negative correlations the polynomial and confidence bound surfaces lie below the zero surface for the entire range of volatility values.

We now focus on the conditions of the second part Proposition 2 which concerns the comparison between the MF and HF covariance MSE’s. The first condition, that the second moment of LF correlation forecasts is smaller than the HF forecasts is empirically violated, however minimally: the differences between the two moments is for most of the pairs very small (the median value of the \( C_1 \) parameter in Equation (41) is about 0.008).

**Figure 2:** Median value of \( \hat{MSE}(\hat{\sigma}^{MF}_{ij}) - \hat{MSE}(\hat{\sigma}^{LF}_{ij}) \). The median value of the polynomial is given by the rough surface. The dotted surfaces represent the 95\% confidence intervals and the surface with rectangular pattern represents the zero bound. On X-axis and Y-axis we plot the true standard deviations, \( d_i \) and \( d_j \), respectively, that range from 0 to 10.
Figure 3: Median value of $\hat{\text{MSE}}(\hat{\sigma}_{ij}^{MF}) - \hat{\text{MSE}}(\hat{\sigma}_{ij}^{HF})$. The median value of the polynomial is given by the rough surface. The dotted surfaces represent the 95% confidence intervals and the surface with rectangular pattern represents the zero bound. On X-axis and Y-axis we plot the true standard deviations, $d_i$ and $d_j$, respectively, that range from 0 to 10.

Moreover, the difference between the dependency among HF volatility forecasts and LF correlation forecasts and the dependency among HF volatility and correlation forecasts, which describes the second sufficient condition, is also very small: on average the median value of the associated parameters in Equation (41) is about $-9.699 \times 10^{-5}$. Given that both effects described in the sufficient conditions are very small, we expect that the MF and HF approaches have a comparative forecasting ability. This result is indeed confirmed by the small difference between the smallest median value of the MSE’s given in Table B.3 (bold entries). The corresponding MF and HF approaches are simultaneously part of the 95% MCS in 33% of the 435 stock pair combinations.

To analyze in more detail the covariance MSE comparison between the MF and HF approaches with respect to market conditions, we plot in Figure 3 the median value of $\hat{F}_2(d_i, d_j, \rho_{ij})$ and the median 95% percentile-bootstrapped confidence intervals over all pairs of stocks for the minimum, average and maximum value of the empirical correlation of the sample and for $d_i, d_j \in [0, 10]$.

One can observe almost no difference between graphs with respect to the choice of $\rho_{ij}$. Similar to the argument stated above, this is due to the fact that the associated parameters in Equation (20) are very small. Moreover, one can observe that half of the polynomial and confidence surfaces lie under the zero bound, which indicates better or equal performance of MF compared to the HF. Thus, when the stocks face very large volatilities (one extreme, of around 160% annually and the other one of around 80% annually), the MF approach is providing better forecasts than the HF. The graph further shows that, in the case of extreme volatility for both stocks, the HF outperforms the MF, however at a lower magnitude than MF outperforms LF (see Figure 2). Thus, in such a situation, one investor has to think twice before going through the computational burden specific to the HF approach in order to have only a small MSE gain compared to the MF approach. Moreover, in order to be able to profit from this gain, one should be in the case where almost all stocks face extreme volatility. This scenario is however not always realistic, as during crises not all stocks are equally affected by extreme events. In a matter of fact, on 16.09.2008, which is the
day with the largest realized volatility in our sample (AIG with 80%), the median value of the realized volatilities among the other 29 stocks was about 3.4%. This explains the previous results that, across all stocks, during the turbulent period starting with 16.09.2008, the MF provides better covariance MSE’s than the HF approach.

Summing up the empirical results, we can conclude that the MF approach is a better choice in forecasting covariance matrices than single-frequency (pure LF or pure HF) methods for three reasons: 1) it has the advantage of using HF data over LF data to estimate and forecast volatilities, which is already widely recognized in literature; 2) it exploits the advantage that mixing frequencies empirically reduces the dependency among volatility and correlation forecasts compared to single-frequency approaches and 3) although it uses a less accurate correlation forecast (based on the outer product of daily returns, instead of daily realized correlation matrices), it has a substantial computational advantage over HF approaches by avoiding the estimation issue and parameter proliferation in modeling the dynamics of realized correlation matrices of large dimensions.

4 Conclusion

In this paper we introduce a new method to forecast covariance matrices of large dimensions by exploring theoretically and empirically the potential of using mixed-frequency data. The method is based on a decomposition of the covariance matrix into a diagonal matrix of volatilities and a correlation matrix. With this decomposition in hand, it becomes easy to employ high-frequency data in the modeling volatilities while still using a standard low frequency model for the correlations in the spirit of the DCC model of Engle (2002).

In the theoretical section of the paper, we derive necessary and sufficient conditions for a mixed frequency model of this kind to provide a smaller element-wise mean squared error relative to a pure low-frequency or a pure high-frequency model (in which correlations are also modeled with the aid of high-frequency data). The conditions we provide do not depend on distributional assumptions on the forecast error or a particular model specifications and are easily checked empirically. This allows practitioners to verify whether in a particular empirical implementation, a mixed frequency model is worth considering or not. We also provide some intuition into the results as well as some general guidance in which environments a MF model is likely to have the largest gains against a single-frequency model, e.g., economic states with high/low volatility and correlation. Furthermore, we discuss how a bootstrap procedure can be implemented to assess the statistical significance of the results.

In the empirical section, we show how the theoretical results can be implemented in practice within an application to 30 highly liquid assets. The analysis reveals that MF approaches decisively outperform standard LF models and perform comparable well with pure HF models which are more cumbersome to implement and are considerably more computationally costly. Moreover, a more detailed analysis of the results show that, particularly in periods of of high volatility, such as the month following the Lehman Brother’s bankruptcy, the MF model outperforms the pure HF approach.
References


\section*{A Appendix}

In this appendix, in order to keep notational burden at a minimum, whenever we do not index variables by superscripts for models (MF, LF, HF) or data frequency (H, L), it means that the equation holds for any model/data frequency.

\subsection*{Preliminaries}

We have that:

\begin{alignat}{2}
    E[\varepsilon_{\sigma_{ii}}] &= 2d_i E[\varepsilon_{d_i}] + E[\varepsilon_{d_i}^2] \\
    E[\varepsilon_{\sigma_{ij}}] &= d_i \rho_{ij} E[\varepsilon_{d_i}] + d_j \rho_{ij} E[\varepsilon_{d_j}] + d_i E[\varepsilon_{d_i} \varepsilon_{\rho_{ij}}] + d_j E[\varepsilon_{d_j} \varepsilon_{\rho_{ij}}] + \\
    &\quad + \rho_{ij} E[\varepsilon_{d_i} \varepsilon_{\rho_{ij}}] + E[\varepsilon_{d_i} \varepsilon_{\rho_{ij}} \varepsilon_{d_j}] 
\end{alignat}

\begin{alignat}{2}
    V[\varepsilon_{\sigma_{ii}}] &= V[2d_i \varepsilon_{d_i} + \varepsilon_{d_i}^2] = 4d_i^2 V[\varepsilon_{d_i}] + V[\varepsilon_{d_i}^2] + 4d_i \text{Cov}[\varepsilon_{d_i}, \varepsilon_{d_i}^2] \\
    &= 4d_i^2 V[\varepsilon_{d_i}] + (E[\varepsilon_{d_i}^3] - E[\varepsilon_{d_i}^2]^2) + 4d_i (E[\varepsilon_{d_i}^2] - E[\varepsilon_{d_i}] E[\varepsilon_{d_i}^2]) \\
    V[\varepsilon_{\sigma_{ij}}] &= d_i^2 \rho_{ij}^2 V[\varepsilon_{d_i}] + d_j^2 \rho_{ij}^2 V[\varepsilon_{d_j}] + d_i^2 d_j^2 \rho_{ij}^2 V[\varepsilon_{d_i} \varepsilon_{d_j}] + d_i^2 V[\varepsilon_{d_i} \varepsilon_{\rho_{ij}}] + d_j^2 V[\varepsilon_{d_j} \varepsilon_{\rho_{ij}}] + \\
    &\quad + \rho_{ij}^2 V[\varepsilon_{d_i} \varepsilon_{\rho_{ij}}] + V[\varepsilon_{d_i} \varepsilon_{\rho_{ij}} \varepsilon_{d_j}] + \\
    &\quad + 2d_i d_j \rho_{ij} \text{Cov}[\varepsilon_{d_i}, \varepsilon_{d_j}] + 2d_i d_j \rho_{ij} \text{Cov}[\varepsilon_{d_i}, \varepsilon_{\rho_{ij}}] + 2d_i d_j \rho_{ij} \text{Cov}[\varepsilon_{d_i}, \varepsilon_{d_i} \varepsilon_{d_j}] + \\
    &\quad + 2d_i d_j \rho_{ij} \text{Cov}[\varepsilon_{d_i}, \varepsilon_{d_j} \varepsilon_{\rho_{ij}}] + 2d_i d_j \rho_{ij} \text{Cov}[\varepsilon_{d_i}, \varepsilon_{d_j} \varepsilon_{d_j}] + 2d_i \rho_{ij} \text{Cov}[\varepsilon_{d_i}, \varepsilon_{d_i} \varepsilon_{d_j}] + \\
    &\quad + 2d_i \rho_{ij} \text{Cov}[\varepsilon_{d_i}, \varepsilon_{d_i} \varepsilon_{\rho_{ij}}] + 2d_i \rho_{ij} \text{Cov}[\varepsilon_{d_i}, \varepsilon_{d_j} \varepsilon_{\rho_{ij}}] + 2d_j \rho_{ij} \text{Cov}[\varepsilon_{d_j}, \varepsilon_{d_i} \varepsilon_{d_j}] + \\
    &\quad + 2d_j \rho_{ij} \text{Cov}[\varepsilon_{d_j}, \varepsilon_{d_j} \varepsilon_{d_j}] + 2d_i \rho_{ij} \text{Cov}[\varepsilon_{d_i}, \varepsilon_{\rho_{ij}} \varepsilon_{d_j}] + 2d_j \rho_{ij} \text{Cov}[\varepsilon_{d_j}, \varepsilon_{\rho_{ij}} \varepsilon_{d_j}] + \\
    &\quad + 2d_i \rho_{ij} \text{Cov}[\varepsilon_{d_i}, \varepsilon_{\rho_{ij}} \varepsilon_{d_j}] + 2d_j \rho_{ij} \text{Cov}[\varepsilon_{d_j}, \varepsilon_{\rho_{ij}} \varepsilon_{d_j}] + \\
    &\quad + 2d_i \rho_{ij} \text{Cov}[\varepsilon_{d_i}, \varepsilon_{d_i} \varepsilon_{d_j}] + 2d_j \rho_{ij} \text{Cov}[\varepsilon_{d_j}, \varepsilon_{d_j} \varepsilon_{d_j}] + 2d_i \rho_{ij} \text{Cov}[\varepsilon_{d_i}, \varepsilon_{d_i} \varepsilon_{\rho_{ij}}] + \\
    &\quad + 2d_j \rho_{ij} \text{Cov}[\varepsilon_{d_j}, \varepsilon_{d_j} \varepsilon_{\rho_{ij}}] + 2d_i \rho_{ij} \text{Cov}[\varepsilon_{d_i}, \varepsilon_{\rho_{ij}} \varepsilon_{d_j}] + 2d_j \rho_{ij} \text{Cov}[\varepsilon_{d_j}, \varepsilon_{\rho_{ij}} \varepsilon_{d_j}] + \\
    &\quad + 2d_i \rho_{ij} \text{Cov}[\varepsilon_{d_i}, \varepsilon_{\rho_{ij}} \varepsilon_{d_j}] + 2d_j \rho_{ij} \text{Cov}[\varepsilon_{d_j}, \varepsilon_{\rho_{ij}} \varepsilon_{d_j}] + 2d_i \rho_{ij} \text{Cov}[\varepsilon_{d_i}, \varepsilon_{\rho_{ij}} \varepsilon_{d_j}] + \\
    &\quad + 2d_j \rho_{ij} \text{Cov}[\varepsilon_{d_j}, \varepsilon_{\rho_{ij}} \varepsilon_{d_j}] + 2d_i \rho_{ij} \text{Cov}[\varepsilon_{d_i}, \varepsilon_{\rho_{ij}} \varepsilon_{d_j}] + 2d_j \rho_{ij} \text{Cov}[\varepsilon_{d_j}, \varepsilon_{\rho_{ij}} \varepsilon_{d_j}] + \\
    &\quad + 2d_i \rho_{ij} \text{Cov}[\varepsilon_{d_i}, \varepsilon_{\rho_{ij}} \varepsilon_{d_j}] + 2d_j \rho_{ij} \text{Cov}[\varepsilon_{d_j}, \varepsilon_{\rho_{ij}} \varepsilon_{d_j}]. 
\end{alignat}

\textbf{Proof A.1} (Proposition 1):

We have that

\[ E[\varepsilon_{\sigma_{ii}}]^2 = 4d_i^2 E[\varepsilon_{d_i}]^2 + E[\varepsilon_{d_i}^2]^2 + 4d_i E[\varepsilon_{d_i}] E[\varepsilon_{d_i}^2] \]

Thus the MSE of $\hat{\sigma}_{ii}$ is given by

\begin{alignat}{2}
    \text{MSE}(\hat{\sigma}_{ii}) &= E[\varepsilon_{\sigma_{ii}}]^2 + V[\varepsilon_{\sigma_{ii}}] = 4d_i^2 E[\varepsilon_{d_i}]^2 + E[\varepsilon_{d_i}^2]^2 + 4d_i E[\varepsilon_{d_i}] E[\varepsilon_{d_i}^2] + \\
    &\quad + 4d_i^2 (E[\varepsilon_{d_i}^3] - E[\varepsilon_{d_i}^2]^2) + (E[\varepsilon_{d_i}^3] - E[\varepsilon_{d_i}^2]^2)^2 + 4d_i (E[\varepsilon_{d_i}^2] - E[\varepsilon_{d_i}] E[\varepsilon_{d_i}^2]) + \\
    &= 4d_i^2 E[\varepsilon_{d_i}^3] + 4d_i E[\varepsilon_{d_i}^3] + E[\varepsilon_{d_i}^4] 
\end{alignat}

It follows that

\[ \text{MSE}(\hat{\sigma}_{ii}^{MF}) - \text{MSE}(\hat{\sigma}_{ii}^{LH}) = 4d_i^2 (E[(\varepsilon_{d_i}^H)^2] - E[(\varepsilon_{d_i}^L)^2]) + 4d_i (E[(\varepsilon_{d_i}^H)^3] - E[(\varepsilon_{d_i}^L)^3]) + \]
(E[(\epsilon_H^{H})^4] - E[(\epsilon_d^L)^4]).

Let

\[ A_{1,i} = E[(\epsilon_H^{H})^2] - E[(\epsilon_d^L)^2], \quad A_{2,i} = E[(\epsilon_H^{H})^3] - E[(\epsilon_d^L)^3], \quad A_{3,i} = E[(\epsilon_H^{H})^4] - E[(\epsilon_d^L)^4] \]

Then we have that

\[ \text{MSE}(\hat{\sigma}^{MF}_H) - \text{MSE}(\hat{\sigma}^{LF}_H) = 4A_{1,i}d_i^2 + 4A_{2,i}d_i + A_{3,i}. \]

Clearly, \( E[(\epsilon_H^{H})^2] - E[(\epsilon_d^L)^2] \leq 0 \), \( E[(\epsilon_H^{H})^3] - E[(\epsilon_d^L)^3] \leq 0 \) and \( E[(\epsilon_H^{H})^4] - E[(\epsilon_d^L)^4] \leq 0 \) are sufficient conditions for \( \text{MSE}(\hat{\sigma}^{MF}_H) \leq \text{MSE}(\hat{\sigma}^{LF}_H) \) since \( d_i \geq 0 \). To verify the necessary and sufficient conditions, consider that \( 4A_{1,i}d_i^2 + 4d_iA_{2,i} + A_{3,i} \leq 0 \) if and only if

1. \( A_{1,i} = 0 \), and \( d_i \leq -\frac{A_{3,i}}{4A_{2,i}} \) or

2. \( A_{1,i} > 0 \), \( A_{2,i} - A_{1,i}A_{3,i} > 0 \) and \( d_i \in [d_{1,ij}, d_{2,ij}] \) where \( d_{1,ij} = \frac{-A_{2,i} - \sqrt{A_{2,i}^2 - A_{1,i}A_{3,i}}}{2A_{1,i}} \) or

3. \( A_{1,i} < 0 \) and either

   3.1 \( A_{2,i} - A_{1,i}A_{3,i} \leq 0 \) or

   3.2 \( A_{2,i} - A_{1,i}A_{3,i} > 0 \) and \( d_i \notin (d_{1,ij}, d_{2,ij}) \) where \( d_{1,ij} = \frac{-A_{2,i} + \sqrt{A_{2,i}^2 - A_{1,i}A_{3,i}}}{2A_{1,i}} \).

The following table summarizes the necessary and sufficient conditions for \( \text{MSE}(\hat{\sigma}^{MF}_H) \leq \text{MSE}(\hat{\sigma}^{LF}_H) \) taking into account that \( d_i \geq 0 \):

<table>
<thead>
<tr>
<th>( A_{1,i} )</th>
<th>( A_{2,i} )</th>
<th>( A_{3,i} )</th>
<th>( d_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \leq 0 )</td>
<td>( \leq 0 )</td>
<td>( \leq 0 )</td>
<td>( \forall d_i )</td>
</tr>
<tr>
<td>( &gt; 0 )</td>
<td>( \leq 0 )</td>
<td>( \leq 0 )</td>
<td>( d_i \leq \frac{-A_{2,i} + \sqrt{A_{2,i}^2 - A_{1,i}A_{3,i}}}{2A_{1,i}} )</td>
</tr>
<tr>
<td>( \leq \frac{A_{2,i}}{A_{3,i}} )</td>
<td>( &gt; 0 )</td>
<td>( \leq 0 )</td>
<td>( \forall d_i )</td>
</tr>
<tr>
<td>( 0 &gt; A_{1,i} &gt; \frac{A_{2,i}}{A_{3,i}} )</td>
<td>( &gt; 0 )</td>
<td>( \leq 0 )</td>
<td>( 0 \leq d_i \leq \frac{-A_{2,i} - \sqrt{A_{2,i}^2 - A_{1,i}A_{3,i}}}{2A_{1,i}} )</td>
</tr>
<tr>
<td>( = 0 )</td>
<td>( &gt; 0 )</td>
<td>( \leq 0 )</td>
<td>( 0 \leq d_i \leq \frac{-A_{2,i}}{4A_{2,i}} )</td>
</tr>
<tr>
<td>( &gt; 0 )</td>
<td>( &gt; 0 )</td>
<td>( \leq 0 )</td>
<td>( d_i \geq \frac{-A_{2,i} - \sqrt{A_{2,i}^2 - A_{1,i}A_{3,i}}}{2A_{1,i}} )</td>
</tr>
<tr>
<td>( \geq 0 )</td>
<td>( \geq 0 )</td>
<td>( = 0 )</td>
<td>( d_i = 0 )</td>
</tr>
<tr>
<td>( &lt; 0 )</td>
<td>( &lt; 0 )</td>
<td>( &gt; 0 )</td>
<td>( d_i \geq \frac{-A_{2,i} + \sqrt{A_{2,i}^2 - A_{1,i}A_{3,i}}}{2A_{1,i}} )</td>
</tr>
<tr>
<td>( = 0 )</td>
<td>( &lt; 0 )</td>
<td>( &gt; 0 )</td>
<td>( d_i \geq \frac{-A_{2,i}}{4A_{2,i}} )</td>
</tr>
<tr>
<td>( 0 &lt; A_{1,i} \leq \frac{A_{2,i}}{A_{3,i}} )</td>
<td>( &lt; 0 )</td>
<td>( &gt; 0 )</td>
<td>( d_i \geq \frac{-A_{2,i} - \sqrt{A_{2,i}^2 - A_{1,i}A_{3,i}}}{2A_{1,i}} )</td>
</tr>
<tr>
<td>( &lt; 0 )</td>
<td>( \geq 0 )</td>
<td>( \geq 0 )</td>
<td>( d_i \geq \frac{-A_{2,i} + \sqrt{A_{2,i}^2 - A_{1,i}A_{3,i}}}{2A_{1,i}} )</td>
</tr>
</tbody>
</table>

\( \Box \)

**Proof A.2** (Proposition 2): Squaring the bias term in Equation (35), adding it to the variance expression in Equation
and simplifying (alternatively one can obtain the MSE directly as the expectation of the squared forecast error), one obtains

\[
MSE[\hat{\sigma}_{ij}] = \rho_j^2 E[\varepsilon_{d,i}^2] + \rho_j^2 E[\varepsilon_{d,j}^2] + d_i^2 d_j^2 E[\varepsilon_{\rho_{ij}}^2]
\]

\[
+ d_i^2 E[\varepsilon_{d,i}^2 \varepsilon_{\rho_{ij}}] + \rho_j^2 E[\varepsilon_{d,j}^2 \varepsilon_{\rho_{ij}}] + 2d_i d_j \rho_j E[\varepsilon_{d,i} \varepsilon_{d,j}]
\]

\[
+ 2d_i^2 \rho_j E[\varepsilon_{d,i}^2 \varepsilon_{\rho_{ij}}] + 6d_i d_j \rho_j E[\varepsilon_{d,i} \varepsilon_{d,j} \varepsilon_{\rho_{ij}}] + 2d_i \rho_j E[\varepsilon_{d,i} \varepsilon_{\rho_{ij}}^2] + 4d_i \rho_j E[\varepsilon_{d,i} \varepsilon_{d,j} \varepsilon_{\rho_{ij}}]
\]

\[
+ 2d_i d_j \rho_j E[\varepsilon_{d,i} \varepsilon_{d,j} \varepsilon_{\rho_{ij}}] + 2d_i \rho_j E[\varepsilon_{d,i} \varepsilon_{d,j} \varepsilon_{\rho_{ij}}] + 4d_j \rho_j E[\varepsilon_{d,j} \varepsilon_{\rho_{ij}}^2] + 4d_i d_j E[\varepsilon_{d,i} \varepsilon_{d,j} \varepsilon_{\rho_{ij}}]
\]

\[
+ 2d_j E[\varepsilon_{d,j}^2 \varepsilon_{\rho_{ij}}^2] + 2d_i^2 \rho_j E[\varepsilon_{d,i}^2 \varepsilon_{\rho_{ij}}^2] + 2d_i d_j \rho_j E[\varepsilon_{d,i} \varepsilon_{d,j} \varepsilon_{\rho_{ij}}]
\]

\[
+ 2d_j E[\varepsilon_{d,j}^2 \varepsilon_{\rho_{ij}}^2] + 2d_i \rho_j E[\varepsilon_{d,i} \varepsilon_{d,j} \varepsilon_{\rho_{ij}}] + 2d_i d_j E[\varepsilon_{d,i} \varepsilon_{d,j} \varepsilon_{\rho_{ij}}]
\]

\[
+ 2d_j E[\varepsilon_{d,j}^2 \varepsilon_{\rho_{ij}}^2] + 2d_i \rho_j E[\varepsilon_{d,i} \varepsilon_{d,j} \varepsilon_{\rho_{ij}}] + E[\varepsilon_{d,i}^2 \varepsilon_{d,j}^2 \varepsilon_{\rho_{ij}}^2]
\]

Given this general expression for the MSE, we obtain that

\[
MSE(\hat{\sigma}_{ij}^{MF}) - MSE(\hat{\sigma}_{ij}^{LF}) = B_{1,ij} d_i^2 \rho_j^2 + B_{2,ij} d_j^2 \rho_j^2 + B_{3,ij} d_i^2 + B_{4,ij} d_j^2 + B_{5,ij} \rho_j^2
\]

\[+ B_{6,ij} d_i d_j \rho_j^2 + B_{7,ij} d_i d_j \rho_j^2 + B_{8,ij} d_i d_j \rho_j^2 + B_{9,ij} d_i d_j \rho_j^2 + B_{10,ij} d_i \rho_j^2 + B_{11,ij} d_i \rho_j^2
\]

\[+ B_{12,ij} d_i d_j \rho_j^2 + B_{13,ij} d_i d_j \rho_j^2 + B_{14,ij} d_i d_j \rho_j^2 + B_{15,ij} d_i d_j \rho_j^2 + B_{16,ij} d_i d_j \rho_j^2
\]

\[+ B_{17,ij} d_i d_j \rho_j^2 + B_{18,ij} d_i d_j \rho_j^2 + B_{19,ij} d_i d_j \rho_j^2 + B_{20,ij} d_i d_j \rho_j^2 + B_{21,ij} d_i d_j \rho_j^2 + B_{22,ij}
\]

\[= F_1(d_i, d_j, \rho_j)
\]

where

\[
B_{1,ij} = E[(\varepsilon_{d,i}^H)^2] - E[(\varepsilon_{d,j}^L)^2]
\]

\[
B_{2,ij} = E[(\varepsilon_{d,j}^H)^2] - E[(\varepsilon_{d,j}^L)^2]
\]

\[
B_{3,ij} = E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2] - E[(\varepsilon_{d,j}^H \varepsilon_{\rho_{ij}}^L)^2]
\]

\[
B_{4,ij} = E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2] - E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2]
\]

\[
B_{5,ij} = E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2] - E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2]
\]

\[
B_{6,ij} = 2E[(\varepsilon_{d,i}^H \varepsilon_{d,j}^H)] - 2E[(\varepsilon_{d,j}^H \varepsilon_{d,j}^L)]
\]

\[
B_{7,ij} = 2E[(\varepsilon_{d,j}^H \varepsilon_{\rho_{ij}}^L)] - 2E[(\varepsilon_{d,j}^H \varepsilon_{\rho_{ij}}^L)]
\]

\[
B_{8,ij} = 2E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2] - E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2]
\]

\[
B_{9,ij} = 2E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2] - E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2]
\]

\[
B_{10,ij} = 2E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2] - E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2]
\]

\[
B_{11,ij} = 4E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2] - E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2]
\]

\[
B_{12,ij} = 2E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2] - E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2]
\]

\[
B_{13,ij} = 2E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2] - E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2]
\]

\[
B_{14,ij} = 2E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2] - E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2]
\]

\[
B_{15,ij} = 2E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2] - E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2]
\]

\[
B_{16,ij} = 2E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2] - E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2]
\]

\[
B_{17,ij} = 2E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2] - E[(\varepsilon_{d,i}^H \varepsilon_{\rho_{ij}}^L)^2]
\]
Furthermore, the MSE difference of the MF model with respect to the HF model is given by

\[ \text{MSE}(\hat{\sigma}_{ij}^{MF}) - \text{MSE}(\hat{\sigma}_{ij}^{HF}) = C_{1,ij} d_i^2 d_j^2 + C_{2,ij} d_i^2 + C_{3,ij} d_j^2 + C_{4,ij} d_i^2 d_j \rho_{ij} + C_{5,ij} d_j^2 \rho_{ij} \\
+ C_{6,ij} d_j \rho_{ij} + C_{7,ij} d_i \rho_{ij} + C_{8,ij} d_i^2 \rho_{ij} + C_{9,ij} d_j^2 \rho_{ij} + C_{10,ij} d_j \rho_{ij} + C_{11,ij} d_i^2 d_j \\
+ C_{12,ij} d_i^2 d_j^2 + C_{13,ij} d_i d_j + C_{14,ij} d_i + C_{15,ij} d_j + C_{16,ij} \rho_{ij} + C_{17,ij} \]

\[ \equiv F_2(d_i, d_j, \rho_{ij}) \]  

(41)

where

\[ C_{1,ij} = E[(\varepsilon_{\rho_{ij}}^L)^2] - E[(\varepsilon_{\rho_{ij}}^H)^2] \]
\[ C_{2,ij} = E[(\varepsilon_{d_i}^L)^2] - E[(\varepsilon_{d_i}^H)^2] \]
\[ C_{3,ij} = E[(\varepsilon_{d_j}^L)^2] - E[(\varepsilon_{d_j}^H)^2] \]
\[ C_{4,ij} = 2E[(\varepsilon_{d_i}^L \varepsilon_{d_j}^L)] - E[(\varepsilon_{d_i}^L)^2] \]
\[ C_{5,ij} = 2E[(\varepsilon_{d_i}^L)^2 - E[(\varepsilon_{d_i}^H)^2] \]
\[ C_{6,ij} = 6E[(\varepsilon_{d_i}^L \varepsilon_{d_j}^H)] - E[(\varepsilon_{d_i}^H)^2] \]
\[ C_{7,ij} = 4E[\varepsilon_{d_i}^H(\varepsilon_{d_j}^L)^2] - E[\varepsilon_{d_i}^H(\varepsilon_{d_j}^H)^2] \]
\[ C_{8,ij} = 2E[\varepsilon_{d_i}^H \varepsilon_{d_j}^L] - E[\varepsilon_{d_i}^H \varepsilon_{d_j}^L] \]
\[ C_{9,ij} = 2E[(\varepsilon_{d_i}^L)^2 \varepsilon_{d_j}^H] - E[(\varepsilon_{d_i}^L)^2 \varepsilon_{d_j}^H] \]
\[ C_{10,ij} = 4E[(\varepsilon_{d_i}^L)^2 \varepsilon_{d_j}^L] - E[(\varepsilon_{d_i}^L)^2 \varepsilon_{d_j}^H] \]
\[ C_{11,ij} = 2E[\varepsilon_{d_i}^H(\varepsilon_{d_j}^L)^2] - E[\varepsilon_{d_i}^H(\varepsilon_{d_j}^H)^2] \]
\[ C_{12,ij} = 2E[\varepsilon_{d_i}^H(\varepsilon_{d_j}^L)^2] - E[\varepsilon_{d_i}^H(\varepsilon_{d_j}^L)^2] \]
\[ C_{13,ij} = 4E[\varepsilon_{d_i}^H \varepsilon_{d_j}^H(\varepsilon_{d_j}^L)^2] - E[\varepsilon_{d_i}^H \varepsilon_{d_j}^H(\varepsilon_{d_j}^H)^2] \]
\[ C_{14,ij} = 2E[\varepsilon_{d_i}^H \varepsilon_{d_j}^H(\varepsilon_{d_j}^L)^2] - E[\varepsilon_{d_i}^H \varepsilon_{d_j}^H(\varepsilon_{d_j}^L)^2] \]
\[ C_{15,ij} = 2E[(\varepsilon_{d_i}^H)^2 \varepsilon_{d_j}^L] - E[(\varepsilon_{d_i}^H)^2 \varepsilon_{d_j}^L] \]
\[ C_{16,ij} = 2E[(\varepsilon_{d_i}^H)^2 \varepsilon_{d_j}^L] - E[(\varepsilon_{d_i}^H)^2 \varepsilon_{d_j}^L] \]
\[ C_{17,ij} = E[(\varepsilon_{d_i}^H \varepsilon_{d_j}^L)^2] - E[(\varepsilon_{d_i}^H \varepsilon_{d_j}^L)^2] \]
## Appendix

<table>
<thead>
<tr>
<th>Series</th>
<th>Mean</th>
<th>Max</th>
<th>Min</th>
<th>Std. dev.</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>daily realized variances</td>
<td>4.262</td>
<td>440.151</td>
<td>0.156</td>
<td>15.179</td>
<td>11.625</td>
<td>250.210</td>
</tr>
<tr>
<td>daily realized covariances</td>
<td>1.317</td>
<td>90.115</td>
<td>-1.732</td>
<td>3.682</td>
<td>12.266</td>
<td>257.934</td>
</tr>
<tr>
<td>daily returns</td>
<td>-0.040</td>
<td>14.637</td>
<td>-13.284</td>
<td>1.892</td>
<td>0.059</td>
<td>13.631</td>
</tr>
</tbody>
</table>

Table B.1: Averaged descriptive statistics of daily realized variances, daily realized covariances and daily returns over the period 11.10.2001 – 01.06.2009. The realized variance and covariance series are scaled by 10^4, while the daily returns series are scaled by 10^2.

<table>
<thead>
<tr>
<th>Model for forecasting standard deviations</th>
<th>Model for forecasting variance MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>MF</td>
<td>ARFIMA(0, d, 0)</td>
</tr>
<tr>
<td></td>
<td>49.34</td>
</tr>
<tr>
<td></td>
<td>ARFIMA(1, d, 0)</td>
</tr>
<tr>
<td></td>
<td><strong>49.11</strong></td>
</tr>
<tr>
<td></td>
<td>ARFIMA(1, d, 1)</td>
</tr>
<tr>
<td></td>
<td>49.14</td>
</tr>
<tr>
<td>LF</td>
<td>GARCH</td>
</tr>
<tr>
<td></td>
<td>59.85</td>
</tr>
</tbody>
</table>

Table B.2: Median value of MSE’s of all (30) variance forecasts.

<table>
<thead>
<tr>
<th>Model for forecasting standard deviations</th>
<th>Model for forecasting correlations cDCC cDCC-E cDCC-I cDCC-M cDCC-CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>MF</td>
<td>ARFIMA(0, d, 0)</td>
</tr>
<tr>
<td></td>
<td>23.09</td>
</tr>
<tr>
<td></td>
<td>23.73</td>
</tr>
<tr>
<td></td>
<td>23.40</td>
</tr>
<tr>
<td></td>
<td>23.23</td>
</tr>
<tr>
<td></td>
<td>22.08</td>
</tr>
<tr>
<td></td>
<td>ARFIMA(1, d, 0)</td>
</tr>
<tr>
<td></td>
<td>22.98</td>
</tr>
<tr>
<td></td>
<td>23.56</td>
</tr>
<tr>
<td></td>
<td>23.27</td>
</tr>
<tr>
<td></td>
<td>23.21</td>
</tr>
<tr>
<td></td>
<td>21.96</td>
</tr>
<tr>
<td></td>
<td>ARFIMA(1, d, 1)</td>
</tr>
<tr>
<td></td>
<td>23.09</td>
</tr>
<tr>
<td></td>
<td>23.41</td>
</tr>
<tr>
<td></td>
<td>23.08</td>
</tr>
<tr>
<td></td>
<td>22.98</td>
</tr>
<tr>
<td></td>
<td>21.95</td>
</tr>
<tr>
<td>LF</td>
<td>GARCH</td>
</tr>
<tr>
<td></td>
<td>24.51</td>
</tr>
<tr>
<td></td>
<td>24.62</td>
</tr>
<tr>
<td></td>
<td>24.36</td>
</tr>
<tr>
<td></td>
<td>24.46</td>
</tr>
<tr>
<td></td>
<td>22.97</td>
</tr>
<tr>
<td>HF</td>
<td>ARFIMA(0, d, 0)</td>
</tr>
<tr>
<td></td>
<td>22.68</td>
</tr>
<tr>
<td></td>
<td>21.97</td>
</tr>
<tr>
<td></td>
<td>VARCHIMA(1,0,0)</td>
</tr>
<tr>
<td></td>
<td>22.56</td>
</tr>
<tr>
<td></td>
<td>21.88</td>
</tr>
<tr>
<td></td>
<td>ARFIMA(1, d, 1)</td>
</tr>
<tr>
<td></td>
<td>22.40</td>
</tr>
<tr>
<td></td>
<td><strong>21.57</strong></td>
</tr>
</tbody>
</table>

Table B.3: Median values of MSE’s of all (435) covariance forecasts.
Figure B.1: Autocorrelation function of daily log-realized volatilities. The first graph corresponds to SPY, the rest of the 29 graphs correspond to the stocks alphabetically ordered as in the main text. The dotted lines represent the 95% confidence intervals. The X-axis depicts lags from 1 to 30 and the Y-axis depicts the values of the autocorrelation function.

Figure B.2: Autocorrelation function of the first 30 pairs of daily realized correlations: i.e., the first 29 graphs correspond to the realized correlations of SPY with the other stocks and the last graph corresponds to the realized correlation between AA and AIG. The dotted lines represent the 95% confidence intervals. The X-axis depicts lags from 1 to 30 and Y-axis depicts the values of the autocorrelation function.