Modelling and Forecasting Multivariate Realized Volatility*

Roxana Chiriac† Valeri Voev‡

University of Konstanz University of Aarhus
CoFE CREATES

Abstract

This paper proposes a methodology for modelling time series of realized covariance matrices in order to forecast multivariate risks. The approach allows for flexible dynamic dependence patterns and guarantees positive definiteness of the resulting forecasts without imposing parameter restrictions. We provide an empirical application of the model, in which we show by means of stochastic dominance tests that the returns from an optimal portfolio based on the model's forecasts second-order dominate returns of portfolios optimized on the basis of traditional MGARCH models. This result implies that *any* risk-averse investor, regardless of the type of utility function, would be better-off using our model.

JEL classification: C32, C53, G11

Keywords: Forecasting, Fractional integration, Stochastic dominance, Portfolio optimization, Realized covariance

*We would like to thank Ingmar Nolte, Winfried Pohlmeier, Joachim Grammig, participants in the Financial Econometrics Seminar in Riezlern and participants in the Multivariate Volatility Models Conference in Faro for helpful discussions. All remaining errors are ours.

†Department of Economics, Box D124, University of Konstanz, 78457 Konstanz, Germany. Phone +49-7531-88-2204, Fax -4450, email: Roxana.Chiriac@uni-konstanz.de. Financial support from the German federal state of Baden-Württemberg through a Landesgraduiertenstipendium and the German Science Foundation within the project "Price-, Liquidity- and Credit Risks: Measurement and Allocation" is gratefully acknowledged.

‡School of Economics and Management, University of Aarhus, 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, CREATES, funded by the Danish National Research Foundation, is gratefully acknowledged.

1 Introduction

Multivariate volatility modelling is of particular importance to the fields of risk management, portfolio management and asset pricing. Typical approaches employed in the modelling of multivariate volatility are the multivariate GARCH (MGARCH) models (for a comprehensive review see Bauwens, Laurent, and Rombouts (2006)), stochastic volatility (SV) models (reviewed in Asai, McAleer, and Yu (2006)) and, more recently, realized covariance models (see Barndorff-Nielsen and Shephard (2004) and Andersen, Bollerslev, Diebold, and Ebens (2001), among others). While the MGARCH and SV approaches model the volatility process as latent, the realized covariance methods employ high-frequency data to enable highly precise estimation of the daily covariance of the underlying assets, thus making it observable.

A prominent feature of volatility is the presence of long memory, which led, within the GARCH framework, to the development of the integrated GARCH (Engle and Bollerslev (1986)), the fractionally integrated GARCH (Baillie, Bollerslev, and Mikkelsen (1996)) and the linear ARCH (Robinson (1991), Giraitis, Robinson, and Surgailis (2000)) models. With high frequency data, the long persistence in a series of realized volatilities is portrayed by a slow decay in the autocorrelation function (see e.g., Andersen and Bollerslev (1997), Andersen, Bollerslev, Diebold, and Ebens (2001)), and is modeled by means of fractionally integrated ARMA (ARFIMA) processes by Andersen, Bollerslev, Diebold, and Labys (2003), Oomen (2001) and Koopman, Jungbacker, and Hol (2005), among others.

Recently, the literature on MGARCH models has been advancing towards flexible model specifications applicable to higher dimensional problems. Yet there is little research on time series models for covariance matrices estimated with high frequency data. The existing literature concerning the dynamic modelling of realized covariance matrices has typically concentrated on univariate approaches for a time series of realized volatilities or a single realized covariance (correlation) series. Andersen, Bollerslev, Diebold, and Ebens (2001) model the series of log-realized volatilities and realized correlations with univariate ARFIMA models, while Corsi (2005) and Corsi and Audrino (2007) apply univariate Heterogenous Autoregressive (HAR) models to capture the high persistence of the series through an autoregressive representation of volatilities/correlations realized over different time horizons. However, the matrix constructed from the variance and correlation forecasts obtained from these univariate

models is not guaranteed to be positive definite. In order to obtain a forecast of the entire covariance matrix, Voev (2007) proposes a methodology in which the univariate variance and covariance forecasts can be combined to produce a positive definite matrix forecast. A drawback of this approach is that the dynamic linkages among the variance and covariance series (e.g., volatility spillovers) is neglected. Among the few proposed models for the dynamics of the whole realized covariance matrix are the Wishart Autoregressive (WAR) model of Gourieroux, Jasiak, and Sufana (2005), based on the distribution of the sample variance-covariance matrix, known in the literature as the Wishart distribution, and the model of Bauer and Vorkink (2006), who employ the matrix log transformation to guarantee positive definiteness of the matrix forecast. The WAR model, however, is incapable of producing long memory type dependence patterns and is built on latent processes, whose interpretation is difficult and which makes the introduction of exogenous forecasting variables hard. The study of Bauer and Vorkink (2006) differs from ours in that it introduces a latent factor model for the log-transform of the covariance matrix and investigates the forecasting power of various predictive variables, such as past returns, risk-free interest rate, dividend yield, credit spread and slope of the term structure.

The model developed in this paper has the advantages of all above-mentioned approaches while alleviating their limitations. We propose the following 3-step procedure: firstly, decompose the series of covariance matrices into their Cholesky factors, secondly forecast the Cholesky series with a well defined time series model and thirdly reconstruct the matrix forecast. The positivity of the matrix forecast is thus ensured by the "squaring" of the Cholesky factors which can be modelled as flexibly as needed without imposing any parameter restrictions. The degree of parameterization (flexibility) of the dynamic model for the Cholesky series should be guided by the dimension of the matrix we are considering, as well as by the application we have in mind; do we aim at a good in-sample fit, or are we more interested in out-of-sample forecasting? In this paper our aim will be the latter and hence we tend to favor very moderately parameterized models. In fact, our preferred specification has only three dynamic parameters regardless of the dimension of the covariance matrix – an AR-, an MA- and a parameter for the degree of fractional integration motivated by the strong persistence of the series. An additional advantage is that the inclusion of an arbitrary number of explanatory predictive variables is straightforward. The model can be seen as an application and extension of the multivariate ARFIMA model of Sowell (1989) which we estimate by conditional maximum likelihood (ML) based on the work of Beran (1995). The conditional approach is preferred over the exact ML methods pro-

posed in the univariate case by Sowell (1992) and An and Bloomfield (1993), since the exact ML approach requires the inversion of a $Tn \times Tn$ matrix, where T is the sample size, and n is the dimension of the process. For a nice review of inference on and forecasting of ARFIMA models we direct the reader to Doornik and Ooms (2004).

A minor complication of the new approach is the difficulty of interpreting the model coefficients. To overcome this problem, we derive the functional form of the marginal effects (impulse responses) which reveal the dynamic linkages among the variance and covariance series.

To assess the merits of our model we consider a risk-averse investor who faces the problem of optimal portfolio selection. A crucial input to this problem is a covariance matrix forecast. We provide him with three choices: a forecast based on our vector ARFIMA (VARFIMA) model, a DCC (Engle (2002)) forecast and a BEKK (Engle and Kroner (1995)) forecast. We then compare the ex-post realized performance of the three sets of portfolio returns. It is common for similar comparisons to be carried out by means of looking at the Sharpe ratio (see Fleming, Kirby, and Ostdiek (2001, 2003), and Şerban, Brockwell, Lehoczky, and Srivastava (2007), among others). This is unsatisfactory from the point of view that the Sharpe ratio is only sufficient if the investor has a quadratic utility and/or if the return distribution is fully described by its first two moments (e.g., a normal distribution). Both of these conditions are at best questionable. We provide a much more powerful comparison which holds for any concave utility function and any return distribution by testing whether a given return distribution stochastically dominates, in terms of second order stochastic dominance, another return distribution. The results strongly suggest that any risk-averse investor would prefer to use our forecasts.

The paper is structured as follows: Section 2 presents the model and the resulting forecasting procedure, Section 3 reports estimation and forecasting results and Section 4 concludes.

2 The Model

Let r_t be a vector of daily log returns of dimension $n \times 1$, where n represents the number of assets considered. The process r_t can be written as:

$$r_t = E[r_t|\mathcal{F}_{t-1}] + \epsilon_t, \tag{1}$$

where $\mathcal{F}_{t-1}$ is the information set consisting of all relevant information up to and including $t-1$ and

$$\epsilon_t = H_t^{1/2} z_t,$$

where H_t is a positive definite matrix of dimension $n \times n$, $H_t^{1/2}$ is its Cholesky decomposition and z_t is an $n \times 1$ vector assumed to be i.i.d. with $E[z_t] = 0$ and $V[z_t] = I_n$. The covariance matrix of the returns is given by:

$$V[r_t|\mathcal{F}_{t-1}] = V[\epsilon_t|\mathcal{F}_{t-1}] = H_t$$

Barndorff-Nielsen and Shephard (2004) and Andersen, Bollerslev, Diebold, and Ebens (2001) propose the realized covariance matrix Y_t as a consistent estimator of H_t.[1] The Cholesky decomposition of the matrix Y_t is given by the upper triangular matrix P_t, for which $P_t' P_t = Y_t$:

$$\begin{aligned} Y_t &= \begin{pmatrix} Y_{11,t} & Y_{12,t} & \dots & Y_{1n,t} \\ Y_{12,t} & Y_{22,t} & \dots & Y_{2n,t} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{1n,t} & \dots & \dots & Y_{nn,t} \end{pmatrix} \\ &= \begin{pmatrix} P_{11,t} & 0 & \dots & 0 \\ P_{12,t} & P_{22,t} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ P_{1n,t} & P_{2n,t} & \dots & P_{nn,t} \end{pmatrix} \cdot \begin{pmatrix} P_{11,t} & P_{12,t} & \dots & P_{1n,t} \\ 0 & P_{22,t} & \dots & P_{2n,t} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & P_{nn,t} \end{pmatrix} \end{aligned}$$

Because the matrix Y_t is symmetric and positive definite by construction, the elements of the matrix P_t are all real (see e.g, Golub and van Loan (1996)). Let $X_t = vech(P_t)$ be the vector obtained by stacking the upper triangular components of the matrix P_t

[1] The realized covariance matrix is the sum of the products of high-frequency (e.g., 5-minute) returns within a given day t. We show how to compute Y_t in the empirical section.

in a vector. X_t is of dimension $m \times 1$, where $m = n(n+1)/2$:

$$X_t = vech(P_t) = \begin{pmatrix} P_{11,t} \\ P_{12,t} \\ P_{22,t} \\ \vdots \\ P_{nn,t} \end{pmatrix} \equiv \begin{pmatrix} X_{1,t} \\ X_{2,t} \\ X_{3,t} \\ \vdots \\ X_{m,t} \end{pmatrix}.$$

We model the dynamics of the vector X_t by using the Vector Autoregressive Fractionally Integrated Moving Average (VARFIMA(p, d, q)) model defined below:

Definition 1: The VARFIMA(p, d, q) model for the vector process X_t is defined as

$$\Phi(L)D(L)[X_t - BZ_t] = \Theta(L)\varepsilon_t, \qquad \varepsilon_t \sim N(0, \Sigma), \tag{2}$$

where Z_t is a vector of exogenous variables of dimension $k \times 1$, B is a matrix of coefficients of dimension $m \times k$, $\Phi(L) = I_m - \Phi_1 L - \Phi_2 L^2 - \ldots - \Phi_p L^p$, $\Theta(L) = I_m + \Theta_1 L + \Theta_2 L^2 + \ldots + \Theta_q L^q$ are matrix lag polynomials with $\Phi_i, i = 1, \ldots, p$ and $\Theta_j, j = 1, \ldots, q$ – the AR- and MA-coefficient matrices, and $D(L) = diag\{(1 - L)^{d_1}, \ldots, (1 - L)^{d_m}\}$, where $d_1, \ldots, d_m$ are the degrees of fractional integration of each of the m elements of the vector X_t. We assume that the roots of $\Phi(L)$ and $\Theta(L)$ lie outside the unit circle. $\diamond$

The model presented here has been studied by Sowell (1989), who shows that an element of the vector X_t, say X_{it}, is stationary if $d_i < 0.5$. Moreover, the whole vector process X_t is stationary if $d_i < 0.5$ for $i = 1, \ldots, m$. In Equation (2), one could consider including exogenous variables that are documented to have an effect on stock market volatility, such as lags of squared daily returns (Black (1976)), functions of trading volume (Lamoureaux and Lastrapes (1990)), corporate bond returns (Schwert (1989)) or short term interest rates (Glosten, Jagannathan, and Runkle (1993)).

The assumption of normally distributed error terms gives rise to a Gaussian likelihood function, which, maximized under certain regularity conditions (see Gourieroux and Monfort (1995)), and the assumption that the conditional mean function is well specified, provides consistent estimates of the parameters of the VARFIMA model defined above. Although the diagonal elements of the Cholesky decomposition are by construction positive, the Gaussianity assumption for the corresponding error terms

in Equation (4) is not problematic. The positive definiteness condition for the covariance matrix based on the forecasted Cholesky factors does not impose positivity restrictions on the elements of the predicted X_{t+s}, for some $s > 0$. Any (invertible) upper triangular matrix constructed from the elements of the forecast of X_{t+s} provides a positive definite matrix of predicted covariances. More formally, the reverse transformation from X_t to Y_t is given by:

$$Y_t = upmat(xpnd(X_t))'upmat(xpnd(X_t)),$$

where the *xpnd* operator is the inverse of the *vech* operator and the *upmat* operator creates an upper triangular matrix. The (i, j)-element of Y_t is related to X_t as follows:

$$Y_{ij,t} = \sum_{l=1+\frac{i(i-1)}{2}}^{\frac{i(i+1)}{2}} X_{l,t} X_{l+\frac{j(j-1)}{2}-\frac{i(i-1)}{2},t}, \quad i, j = 1, \ldots, n, \ j \geq i. \tag{3}$$

where $X_{l,t}$ is the l-th element of X_t. This transformation embodies and illustrates the main advantage of our model specification: it guarantees the positive definiteness and symmetry of the covariance matrix without imposing any restrictions on the parameters in the model for X_t. In terms of estimation, we face the problem that the parameters of the unrestricted VARFIMA models are not identified, which results from the non-uniqueness of VARMA models, discussed at length in Lütkepohl (2005). The problem in the multivariate case is even more severe than in the univariate ARMA case, in which root cancelation in the AR and MA-polynomials can occur. In the multivariate case, even after assuming that the AR and the MA polynomials have no common roots, one can still factor out infinitely many times a so-called unimodular lag operator without changing the structure of the process.[2] Lütkepohl (2005) discusses two forms of a general VARMA model which are unique representations of a given VARMA process: final equations form and echelon form. In our paper we consider the final equations form, for which we provide a definition below.

Definition 2: The n-dimensional VARMA(p, q) representation $\Phi(L)Y_t = \Theta(L)\varepsilon_t$ is said to be in final equations form if $\Theta_0 = I_n$ and $\Phi(L) = 1 - \phi_1 L - \ldots - \phi_p L^p$ is a scalar operator with $\phi_p \neq 0$. $\diamond$

Following this definition, we estimate the model in final equations form, restricting

[2] A unimodular lag operator is an operator whose determinant is a non-zero constant, i.e., the determinant does not involve powers of L.

Table 1: Number of parameters for the general VARFIMA(p, d, q) model and its restricted specifications, considered in this paper.

	Dimension	Number of parameters	Model 1	Model 2
$\Phi(L)$	1×1	p	1	1
$D(L)$	$m \times m$	m	m	1
B	$m \times k$	km	m^*	m^*
$\Theta(L)$	$m \times m$	qm^2	m^2	1
Total number of parameters		$qm^2 + (k+1)m + p$	$m^2 + 2m + 1$	$m + 3$

Note: $^*k = 1$ (constant). Model 1 is an unrestricted VARFIMA$(1, d, 1)$ and Model 2 is a scalar VARFIMA$(1, d, 1)$ with $d_1 = d_2 = \ldots = d_m$ and a scalar Θ.

the AR polynomial to be a scalar polynomial. Apart from guaranteeing uniqueness of the representation, this approach leads to a reduction in the number of parameters to be estimated. Table 1 gives the total number of parameters for a general VARFIMA(p, d, q) model in final equations form, as well as for two restricted model specifications considered in this paper and described below. Note, that we exclude Σ as a parameter of the model since we refrain from estimating it, as will become clear in the empirical section.

For our purposes, we employ the model in Equation (2) with AR and MA polynomials of order one and a mean vector c of dimension $m \times 1$ (Model 1):

$$\Phi(L)D(L)[X_t - c] = \Theta(L)\varepsilon_t, \qquad \varepsilon_t \sim N(0, \Sigma) \tag{4}$$

Given that $m = \frac{n(n+1)}{2}$, where n represents the number of stocks considered in the application, Model 1 has a total of $\frac{(n^2+n+2)^2}{4}$ parameters. In order to reduce the number of parameters, we assume in a restricted version of the model that all Cholesky decomposition series are fractionally integrated with the same degree of integration $d = d_1 = \ldots = d_m$, and, consequently, $D(L) = (1 - L)^d I_m$. Further reduction of the number of parameters is achieved by restricting the parameter matrix Θ to be scalar (Model 2). In practice the mean vector c can be estimated in a first step as the sample mean of X_t which leaves only three parameters for estimation in the second step. This approach is related to correlation targeting in DCC models in which the unconditional correlation matrix is set equal to the sample correlation matrix of the series in order to reduce the number of parameters to be estimated.

Forecasting

In what follows, we present the theory of forecasting with the VARFIMA model presented above. The forecasting performance of the model is assessed in the next section by using historical stock return data.

For ease of exposition and since the exogenous regressors in the model in Equation (2) are by assumption predetermined, we neglect the term BZ_t. The fractionally differenced series $D(L)X_t$ follows a stationary VARMA process, and therefore we can obtain forecasting formulas through its infinite Vector Moving Average (VMA(∞)) representation (see e.g., Lütkepohl (2005), pp. $432-434$). For each $j = 1, \ldots, m$, the fractionally differenced series $(1-L)^{d_j} X_{j,t}$ is given by:

$$(1-L)^{d_j} X_{j,t} = \sum_{h=0}^{\infty} \delta_{j,h} X_{j,t-h} = X_{j,t} + \sum_{h=1}^{\infty} \delta_{j,h} X_{j,t-h}, \tag{5}$$

where $\delta_{j,0} = 1$ and $\delta_{j,h} = \prod_{0<r\leq h} \frac{r-1-d_j}{r}$, $h = 1, 2, \ldots$. Therefore, we can rewrite Equation (2) as:

$$\Phi(L)\Lambda(L)X_t = \Theta(L)\varepsilon_t, \tag{6}$$

where $\Lambda(L) = I_m + \sum_{h=1}^{\infty} \Delta_h L^h$ and $\Delta_h = diag\{\delta_{1,h}, \ldots, \delta_{m,h}\}$. From Equation (6) we can derive the VMA(∞) representation:

$$X_t = \Phi(L)^{-1}\Lambda(L)^{-1}\Theta(L)\varepsilon_t = \sum_{i=0}^{\infty} \Psi_i \varepsilon_{t-i},$$

where $\Psi_0 = I_m$ and the optimal predictor of X_t in terms of the VMA(∞) representation is given by:

$$E_t[X_{t+s}] = \sum_{i=s}^{\infty} \Psi_i \varepsilon_{t+s-i} = \sum_{i=0}^{\infty} \Psi_{s+i} \varepsilon_{t-i}$$

The resulting forecast is unbiased (that is, the forecast errors have zero mean) and since the ε_t are assumed to be normally distributed, the forecast errors are also normally distributed as:

$$u_{t,t+s} \equiv X_{t+s} - E_t[X_{t+s}] \sim N(0, \Sigma_s),$$

where

$$\Sigma_s = E[(X_{t+s} - E_t[X_{t+s}])(X_{t+s} - E_t[X_{t+s}])'] = E[u_{t,t+s}u'_{t,t+s}] = \sum_{i=0}^{s-1} \Psi_i \Sigma \Psi'_i$$

It follows that the forecast errors of the one-step ahead forecast, $u_{t,t+1}$, are normally distributed with zero-mean and variance-covariance matrix $\Sigma_1 = \Sigma$. As seen in Equation (5), for each $j = 1, \ldots, m$, the $X_{j,t}$ process has an infinite autoregressive representation that can be truncated at, say $h = 1000$ lags for practical purposes.

Having forecasted X_{t+s}, we construct the forecast of the daily volatility matrix Y_{t+s} by applying the transformation in Equation (3). As a brief aside, note that we are in fact forecasting the series Y_t, while ideally we would like to forecast H_t. H_t, however, is not observable, implying that the quality of the forecast does not fully depend on the dynamic specification of Y_t but also on the quality of the realized covariance estimator. It is beyond the scope of this paper to address the latter issue; the search for better and better multivariate volatility measures using high frequency data is currently a very active area of research. In this paper we use an estimator which has been shown to be reliable and much more precise than any estimator based on daily data.

Let us define the forecast errors for the individual elements of Y_{t+s} as $e_{ij,t+s} = E_t[Y_{ij,t+s}] - Y_{ij,t+s}$. Since the forecast of Y_{t+s} is a quadratic transformation of the forecast of X_{t+s}, the mean of $e_{ij,t+s}$ is generally no longer zero, and depends on the variance Σ_s of the forecast error $u_{t,t+s}$. Thus, in order to obtain unbiased predictions, each component of the covariance matrix forecast, $E_t[Y_{ij,t+s}]$, should be corrected for the bias given by $E_t[e_{ij,t+s}] \equiv \sigma^*_{s,ij} \neq 0$. From Equation (3), it follows that the bias correction for $E_t[Y_{ij,t+s}]$ can be obtained from the elements of the matrix Σ_s using the following formula:

$$\sigma^*_{s,ij} = \sum_{l=1+\frac{i(i-1)}{2}}^{\frac{i(i+1)}{2}} \sigma_{s(l,l+\frac{j(j-1)}{2}-\frac{i(i-1)}{2})}, \tag{7}$$

where $j \geq i$, $i = 1, \ldots, n$ and $\sigma_{s(u,v)}$ is the (u, v)-element of Σ_s.

Since the model is applied to a transformation of the realized covariance matrix, namely the series of Cholesky factors, the parameters in Equation (2) are not directly interpretable. However, one can derive the dynamic linkages among the variance and covariance series as functions of these parameters. The elements of the predicted covariance matrix $E_t[Y_{t+s}]$ are (nonlinear) functions of elements of the forecast $E_t[X_{t+s}] \equiv E[X_{t+s}|\mathcal{F}_t]$, and, therefore, functions of the estimated parameter vector and the variables included in $\mathcal{F}_t$, which in our case is the history of the process X_t up to time t, denoted by $\boldsymbol{X}_t$. We can write the (i, j)-element of the predicted covariance

matrix as (see Equation 3):

$$E_t[Y_{ij,t+s}] = \sum_{l=1+\frac{i(i-1)}{2}}^{\frac{i(i+1)}{2}} E_t\left[X_{l,t+s}X_{l+\frac{j(j-1)}{2}-\frac{i(i-1)}{2},t+s}\right] \equiv G_{i,j,s}(\boldsymbol{X}_t, \vartheta), \tag{8}$$

where $i, j = 1, \ldots, n$, $j \geq i$ and $E_t\left[X_{l,t+s}\right]$ is the l-th element of the vector $E_t\left[X_{t+s}\right]$. $G_{i,j,s}(\cdot)$ is a scalar function of $\boldsymbol{X}_t$ and ϑ, corresponding to the (i,j)-element of the matrix $E_t[Y_{t+s}]$, and ϑ is the vector of all model parameters. For example, the impact of a shock in the covariance $Y_{ij,t}$ on the predicted variance $E_t[Y_{ii,t+s}]$ can be computed as follows:

$$\frac{\partial E_t[Y_{ii,t+s}]}{\partial Y_{ij,t}} = \frac{\partial G_{i,i,s}}{\partial G_{i,j,0}} = \sum_{r=1}^{m} \frac{\partial G_{i,i,s}}{\partial X_{r,t}} \frac{\partial X_{r,t}}{\partial G_{i,j,0}} = F^{s,t}_{ii,ij}(\boldsymbol{X}_t, \vartheta), \tag{9}$$

where $G_{i,j,0} \equiv Y_{ij,t} = \sum_{l=1+\frac{i(i-1)}{2}}^{\frac{i(i+1)}{2}} X_{l,t}X_{l+\frac{j(j-1)}{2}-\frac{i(i-1)}{2},t}$, $j \geq i$ and $F^{s,t}_{ii,ij}(\cdot)$ is a scalar function. In a similar way, one can derive the the impact of the variance $Y_{ii,t}$ on the predicted covariance $E_t[Y_{ij,t+s}]$:

$$\frac{\partial E_t[Y_{ij,t+s}]}{\partial Y_{ii,t}} = \frac{\partial G_{i,j,s}}{\partial G_{i,i,0}} = \sum_{r=1}^{m} \frac{\partial G_{i,j,s}}{\partial X_{r,t}} \frac{\partial X_{r,t}}{\partial G_{i,i,0}} = F^{s,t}_{ij,ii}(\boldsymbol{X}_t, \vartheta), \tag{10}$$

where $G_{i,i,0} \equiv Y_{ii,t} = \sum_{l=1+\frac{i(i-1)}{2}}^{\frac{i(i+1)}{2}} X^2_{l,t}$. The expressions for $G_{i,j,s}$, $G_{i,i,s}$, $F^{s,t}_{ii,ij}$ and $F^{s,t}_{ij,ii}$ are derived in Appendix A. In Section 3 we report estimated values of such marginal effects for our empirical example.

3 Empirical Application

In this section we present results from estimating and forecasting the VARFIMA model using historical return data for $n = 6$ stocks traded at the New York Stock Exchange (NYSE). For the estimation, we use a multivariate extension of the conditional maximum likelihood approach of Beran (1995). It is important to state that in this empirical paper, we focus mainly on evaluating the out-of-sample performance rather than on in-sample fit of the model: while in-sample evaluation methods are in general limited and cumbersome when applied to highly dimensional models (Engle and Sheppard (2007), Bauwens, Laurent, and Rombouts (2006)) and less relevant for practical purposes, the out-of-sample assessment of covariance models is of key impor-

tance for the evaluation of our ability to precisely predict financial risks. We compare our model to two very popular MGARCH models – the DCC of Engle (2002) and the BEKK of Engle and Kroner (1995) – by applying statistical and economic criteria, such as ex-post performance of mean-variance optimal portfolios.

3.1 Data

We use tick-by-tick bid and ask quotes from the NYSE Trade and Quotations (TAQ) database sampled from 9:45 until 16:00 over the period January 1, 2001 to June 30, 2006 (1381 trading days). Although the NYSE market opens at 9:30, we filter out the quotes recorded in the first 15 minutes in order to eliminate the opening auction effect on the price process. For the current analysis, we select the following six stocks: American Express Inc. (AXP), Citigroup (C), Home Depot Inc. (HD), Hewlett-Packard (HWP), International Business Machines (IBM) and JPMorgan Chase & Co (JPM). All stocks trade on the NYSE and are highly liquid, which motivated the choice. In order to obtain a regularly spaced sequence of midquotes, we use the previous-tick interpolation method, described in Dacorogna, Gençay, Müller, Olsen, and Pictet (2001). The mid-quotes are thus sampled at the 5-minute and daily frequency, from which 5-minute and daily log returns are computed. Thus we obtain 75 intraday observations which are used to compute the realized variance-covariance matrices for each day. Table B.1 in Appendix B reports summary statistics of both 5-minute and daily returns. We observe typical stylized facts such as overkurtosis and tendency for negative skewness of intradaily and daily returns (across all six stocks, the average kurtosis of 5-minute return series is about 269.2, while of daily returns is about 10.9). For estimation, we scale up the daily and intradaily returns by 100, i.e., we consider percentage returns.

For each $t = 1, \ldots, 1381$, we construct series of daily realized covariance matrices, Y_t, from with 5-minute returns as:

$$Y_t = \sum_{j=1}^{M} r_{j,t} r'_{j,t} \tag{11}$$

where $M = 75$ and $r_{j,t}$ is the $n \times 1$ vector of 5-minute returns 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 midquote price at time $j\Delta$ in day t. By construction, the realized covariance matrices are symmetric and, for $n < M$, they are positive definite almost surely. Since by sampling sparsely we disregard a lot of data, we refine the estimator by considering subsamples. With $\Delta = 300$ sec, we construct 30 regularly Δ-spaced subgrids starting at seconds $1, 11, 21, \ldots, 291$, compute the realized covariance matrix for each subgrid and average over the subgrids. The resulting subsampled realized covariance is much more robust to noise and non-synchronicity than the simple 5-minute based one. As we are interested in the covariance matrix of the whole day (close-to-close), and Y_t estimates only its open-to-close portion, we use the scaling method introduced by Hansen and Lunde (2005) adapted to the multivariate case: we scale each (co)variance estimate corresponding to the trading period by an average scaling factor, which incorporates the overnight information over all series. This procedure preserves the positive-definiteness of the resulting covariance matrix. Table B.2 in Appendix B reports summary statistics of realized variances and covariances of the six stocks considered in the study. As already documented by Andersen, Bollerslev, Diebold, and Ebens (2001), both realized variance and covariance distributions are extremely right skewed and leptokurtic.

After computing the series of realized covariance matrices, we construct the series of Cholesky factors, which inherit the long memory property of realized (co)variances documented by Andersen and Bollerslev (1997) and Andersen, Bollerslev, Diebold, and Ebens (2001). To get an idea about the degree of fractional integration, we run OLS regressions of log-autocorrelations on log-lags (see Beran (1998), pp. 89-92) and obtain a cross-sectionally averaged estimate of 0.24.

3.2 MGARCH Models

For our comparative study we consider two popular MGARCH approaches for the conditional covariance matrix: the DCC model (Engle (2002)) and the diagonal BEKK model (Engle and Kroner (1995)). We assume here that the conditional mean of daily returns is constant, $E[r_t|\mathcal{F}_{t-1}] = \mu$ (see Equation (1)) and we estimate it along with the MGARCH parameters.

DCC GARCH

Engle (2002) proposed a multivariate GARCH model with univariate GARCH(1,1)

conditional variances, $h_{ii,t}$, and dynamic conditional correlations:

$$H_t = D_t R_t D_t,$$

where $D_t = diag(h_{11,t}^{1/2} \dots h_{nn,t}^{1/2})$ and

$$h_{ii,t} = w_i + \alpha_i \epsilon_{i,t-1}^2 + \beta_i h_{ii,t-1},$$

where $w_i, \alpha_i, \beta_i \geq 0$ and $\alpha_i + \beta_i < 1$, $\forall i = 1, \dots, n$.

$$R_t = (diag(Q_t))^{-\frac{1}{2}} Q_t (diag(Q_t))^{-\frac{1}{2}},$$

where Q_t is an $n \times n$ symmetric and positive definite matrix given by:

$$Q_t = (1 - \theta_1 - \theta_2)\bar{Q} + \theta_1 u_{t-1} u_{t-1}' + \theta_2 Q_{t-1},$$

where u_t is the vector of standardized residuals with elements

$$u_{i,t} = \frac{\epsilon_{i,t}}{\sqrt{h_{ii,t}}}, \quad i = 1, \dots, n$$

and $\bar{Q}$ is the unconditional covariance of u_t. For $n = 6$ assets, the DCC model implies a total of 26 parameters, which are estimated by maximizing the normal pseudo-likelihood.

Diagonal BEKK

Engle and Kroner (1995) suggested a multivariate GARCH model, where the conditional return covariance matrix H_t is parameterized as a function of lags and lagged squared innovations:

$$H_t = C'C + A'\epsilon_{t-1}\epsilon_{t-1}' A + B' H_{t-1} B,$$

where C is an upper $n \times n$ triangular matrix and A and B are $n \times n$ parameter matrices. Under certain restrictions, described in Engle and Kroner (1995), the resulting covariance matrices are assured to be positive definite and stationary. In the present paper, we estimate the diagonal specification of the model, where A and B are diagonal matrices. The model includes 39 parameters, which are estimated by maximum likelihood assuming conditional normality.

3.3 Estimation Results

Before turning to the forecasting evaluation, we present here estimation results for the full sample of data. The results of the DCC and diagonal BEKK models are reported in Table B.3. Due to the "curse of dimensionality", the unrestricted VARFIMA model (Model 1) with $n = 6$ ($m = 21$) implies the estimation of 484 parameters, and therefore we present here the empirical results only for the scalar model (Model 2). As discussed earlier, we set c equal to the sample mean $\bar{X}$ of X_t and estimate the model for $X_t - \bar{X}$ with just three parameters: d, ϕ and θ. In order to avoid estimating the $m \times m$ matrix Σ, we set it equal to the unconditional covariance matrix for a stationary VARFIMA$(1, d, 1)$ process for which we use an approximation based on the VARMA(∞,1) representation truncated at 1000 lags. Thus Σ is a function of the parameters d, ϕ and θ, and of the sample covariance matrix of X_t. The long memory of the series is reflected in an estimated degree of fractional integration d of approximately 0.210; the autoregressive parameter is significantly positive ($\hat{\phi} = 0.025$), while the moving average parameter is significantly negative ($\hat{\theta} = -0.105$). Similar results are obtained by Oomen (2001) who estimates ARFIMA models for log realized volatilities. As a result of pre-estimating c and using the sample covariance of X_t in the computation of Σ, the resulting "second-step" QML standard errors of the estimated parameters are incorrect. Therefore, to assure a robust inference of the model parameters, we derive the standard errors by employing the subsampling bootstrap method developed by Politis and Romano (1994a) and Politis, Romano, and Wolf (1999) for dependent and cross-correlated time series. The values of the standard deviations of $\hat{d}$, $\hat{\phi}$ and $\hat{\theta}$ are approximately 0.0063, 0.0052 and 0.0065, respectively, indicating that the estimated parameters are significant at the 5% level.

As already mentioned in Section 2, although the estimated coefficients are not directly interpretable, we can measure the dynamic marginal effects among the series (see Appendix A). For example, we find that the variance of the AXP stock returns has a positive effect on the one-step ahead correlation between AXP and HWP returns; the estimated value of the marginal effect of AXP variance on the conditional expectation of the covariance between AXP and HWP returns is around 0.044 (based on Equation (A6)). This well documented phenomenon is referred to in the literature as the "volatility in correlation effect" (see Andersen, Bollerslev, Diebold, and Ebens (2001)) and indicates the strong linkage between volatilities and correlations on the stock market. The interested reader can obtain the values of all marginal effects from the authors on request.

Figure B.1 in Appendix B plots the autocorrelograms of the standardized residual series of the estimated VARFIMA model. The residual autocorrelogram occasionally reveals some remaining autocorrelation, which might be the result of truncating the infinite AR polynomial in Equation (5) at $h = 1000$ for practical reasons. We emphasize again, however, that in the present study, we are more interested in analyzing the out-of-sample performance of the new model, rather than concentrating on its in-sample fit. Moreover, it is clear that the tight parametrization leads to a low in-sample performance but might very well improve the quality of the out-of-sample forecasts (see e.g., Engle and Sheppard (2007), among others).

3.4 Forecasting Results

In order to compare the forecasting performance of the three models, we divide the overall sample of 1381 days into two subsamples: an in-sample period on which we estimate the model, and an out-of-sample period which serves to evaluate the forecasting performance. The in-sample period contains initially the first 1181 observations. In each forecasting step, we increase the in-sample period by one observation, re-estimate the models and make a new one-step ahead forecast. This procedure is carried out 200 times, and as a result we obtain a total of 200 one-step ahead forecasts for each covariance estimator. Given that there are no significant differences in the VARFIMA forecasts with and without bias correction, we ignore the bias correction for computational reasons. As a quick comparison we mention that the cross-sectional average root mean squared prediction errors for the VARFIMA, DCC and BEKK forecasts are 0.508, 0.580 and 0.564, respectively. These are very crude "goodness" measures and we do not claim that there are statistically significant differences among them. They should, therefore, be interpreted as simply indicative. In this paper, we refrain from further statistical forecast comparisons.

In order to asses the economic value of the three model forecasts, we construct portfolios which are supposed to maximize the utility of a risk-averse investor. If the utility function is second degree polynomial or logarithmic and/or the return distribution is completely characterized by its first two moments (e.g. normal distribution), the portfolio optimization reduces to finding the asset weights which minimize the portfolio volatility while aiming for a target expected return or maximize the portfolio return while targeting a certain volatility (Markowitz (1952)).

We assume that an investor minimizes his portfolio volatility subject to a fixed expected return (10% p.a.). He is allowed (Scenario 1) or prohibited (Scenario 2) to sell assets he does not own (short selling). In this context, the optimal portfolio is given by the solution to the following quadratic problem:

$$\min_{w_{t+1|t}} \quad w'_{t+1|t}\hat{H}_{t+1|t}w_{t+1|t}$$

subject to:
Scenario 1: $w'_{t+1|t}E_t[R_{t+1}] + (1 - w'_{t+1|t}\iota)R_f = R^*$
Scenario 2: $w'_{t+1|t}E_t[R_{t+1}] + (1 - w'_{t+1|t}\iota)R_f = R^*$, $w_{t+1|t} \geq 0$,
where $\hat{H}_{t+1|t}$ is the covariance forecast at day t for day $t+1$, $w_{t+1|t}$ is the $n \times 1$ vector of portfolio weights chosen at day t for day $t+1$, ι is an $n \times 1$ vector of ones, R_f is the risk free rate (3% p.a.) and R^* is the target expected return (10% p.a.).

Given that there is hardly any predictable return variation at the daily level, we assume that the expected returns are constant as in Fleming, Kirby, and Ostdiek (2001, 2003). Having solved for the optimal weights based on the three different conditional covariance forecasts, we compute the ex-post daily portfolio returns and the corresponding Sharpe ratios, given by:

$$SR = \frac{\bar{R}_p - R_f}{\hat{\sigma}_{R_p}},$$

where $\bar{R}_p$ is the sample mean and $\hat{\sigma}_{R_p}$ – the sample standard deviation of the ex-post realized portfolio return series.

Table 2: Annualized Sharpe ratios and standard deviations of out-of-sample realized portfolio returns

Portfolio	VARFIMA	DCC	BEKK
	Sharpe Ratios		
Scenario 1	0.976	0.615	0.491
Scenario 2	0.531	0.455	0.242
	Standard Deviations		
Scenario 1	12.71	12.97	13.16
Scenario 2	15.38	17.25	16.72

Table 2 reports the annualized realized Sharpe ratios and standard deviations of the

three sets of minimum-covariance portfolios. The numbers in this table should be interpreted simply as indicative that for the considered sample, the VARFIMA-based portfolio delivers a smaller standard deviation and a higher Sharpe ratio than the GARCH-based ones. We relegate the formal comparison of these results by means of significance tests to the following discussion on stochastic dominance which is a much more general way of assessing whether a given return distribution is "better" than another one.

The assumption of a "mean-variance" investor is rather restrictive from an economic point of view. A more meaningful evaluation of the optimality of the portfolios can be achieved by comparing the whole distribution of the portfolio returns as opposed to just the first two moments. For example, the skewness and the shape of the tails of the return distribution are of significant relevance in the investment decision process. Therefore, in what follows, we compare the VARFIMA-, DCC- and BEKK-based portfolio return distributions by means of stochastic dominance tests. To this end we need an additional definition.

Definition 3: Let X_1 and X_2 be two real random variables. It is said that X_1 s-th order stochastically dominates X_2 ($X_1 \succeq_s X_2$, $s > 0$) if and only if $F^s_{X_1}(x) \leq F^s_{X_2}(x)$ for all x with strict inequality for some x, where $F^s_{X_i}(x) = \int_{-\infty}^{x} F^{s-1}_{X_i}(t)dt$ for $s \geq 2$, $F^1_{X_i}(x) = F_{X_i}(x)$ and $F_{X_i}(x)$ is the cumulative distribution function (CDF) of X_i, $i = 1, 2$. $\diamond$

Fishburn (1980) and Bawa (1975), among others, show that X_1 s-th order stochastically dominates X_2 if and only if $E[u(X_1)] \geq E[u(X_2)]$ (with strict inequality for some x from the common support of X_1 and X_2) for every function u with $(-1)^{j+1}u^{(j)}(x) \geq 0$ for all $j \in 1, \ldots, s$ where $u^{(j)}(x)$ stands for the j-th derivative of $u(x)$. The implications of this for our analysis are as follows: Let us have two optimal portfolio strategies (forecasting models), A and B and $R_{p,A}$ and $R_{p,B}$ be the realized returns of the two minimum-variance portfolios with CDF's $F_A(x)$ and $F_B(x)$. A risk-averse investor with an increasing utility function $u(x)$, translating into $u^{(1)}(x) \geq 0$ and $u^{(2)}(x) \leq 0$, chooses portfolio A over portfolio B if and only if portfolio A second order stochastically dominates portfolio B, i.e., $\int_{-\infty}^{r} F_A(x)dx \leq \int_{-\infty}^{r} F_B(x)dx$ for $r \in \Pi$, where Π is the common support of $R_{p,A}$ and $R_{p,B}$, with strict inequality for at least one $r \in \Pi$. In this case the investor has a larger expected utility from portfolio A than from portfolio B, $E[u(R_{p,A})] \geq E[u(R_{p,B})]$.

Comparing the integrated cumulative distributions (i.e., $F^2(\cdot)$) of the VARFIMA-based portfolio pairwise against the DCC- and BEKK-based ones, we find that the former is strictly smaller for each value of the common return support, which is a first indication that the VARFIMA-based portfolio second order stochastically dominates the other two portfolios. To check the robustness of these results, we apply a number of stochastic dominance tests on the estimated distributions.

The literature on stochastic dominance tests is separated into two groups: one group (McFadden (1989), Klekan, McFadden, and McFadden (1991), Barett and Donald (2003), Linton, Maasoumi, and Whang (2005)) tests the null hypothesis of dominance ($H_0 : A \succeq_2 B$) against the alternative of non-dominance ($H_1 : A \not\succeq_2 B$), while the other group (Kaur, Rao, and Singh (1994), Davidson and Duclos (2000)) tests the null hypothesis of non-dominance, against the alternative hypothesis of dominance. Most of these tests are developed on the assumptions of i.i.d. and cross-independent observations. Due to the fact that we deal with serially (due to GARCH effects) and cross-dependent portfolio returns, we apply here two tests which account for these features: the Linton, Maasoumi, and Whang (2005) (LMW) test and Kaur, Rao, and Singh (1994) (KRS) test. We use the LMW test with the subsampling procedure (Sub) of Politis and Romano (1994a) and Politis, Romano, and Wolf (1999) and the stationary bootstrap (SB) procedure of Politis and Romano (1994b) to obtain consistent critical values for the test.

Table 3 reports the p-values of the LMW and KRS tests for various null hypotheses described in the first column. Regardless of the investment strategy, all tests with the null hypothesis of stochastic dominance of the VARFIMA portfolio against the other two portfolios have a p-value well in excess of 60% indicating a strong support for the null. Changing the testing direction, we strongly reject the null hypothesis of dominance of the MGARCH portfolios against the VARFIMA for Scenario 2 and with the SB-based LMW test for Scenario 1. Similar results are obtained from the KRS test with null hypotheses on non-dominance. Generally, for Scenario 2 we find ample evidence for the dominance of the VARFIMA-based portfolio, while for Scenario 1 the data is inconclusive, but still delivers some support for our model. Referring again to Table 2, it is evident that for Scenario 2, the differences in the variance of the portfolio distributions are substantial, which is the reason for the much more clear-cut test results compared to Scenario 1. The relevance of the constrained portfolio optimization problem in Scenario 2 is supported by the fact that many institutional

investors are forbidden by law from short selling. Furthermore, a recent study of Boehmer, Jones, and Zhang (2008) reveals that on the NYSE only up to 2% of short sales are undertaken by individual traders. Thus, we conclude that combining the precision of high-frequency data to measure realized volatility with a sensible time-series model to forecast it, is a worthwhile strategy to pursue, as it has the potential of providing added economic value.

Table 3: P-values of the LMW and KRS tests for 2^{nd} order stochastic dominance.

		Scenario 1		*Scenario 2*	
Test/Portfolio B		DCC	BEKK	DCC	BEKK
LMW Test					
Sub	$H_0 : A \succeq_2 B$	0.803	0.625	0.812	0.785
	$H_0 : B \succeq_2 A$	0.633	0.160	0.000	0.000
SB	$H_0 : A \succeq_2 B$	0.930	0.960	0.871	0.792
	$H_0 : B \succeq_2 A$	0.019	0.009	0.000	0.000
KRS Test					
SB	$H_0 : A \not\succeq_2 B$	0.254	0.107	0.098	0.009
	$H_0 : B \not\succeq_2 A$	0.960	0.950	0.990	0.990

Note: Portfolio A denotes the minimum covariance portfolio based on the VARFIMA forecasts. The critical values of the tests are derived from bootstrap procedures which account for serial and cross dependence of the observations: subsampling bootstrap ("Sub") and stationary bootstrap ("SB"). The subsampling size is $b = 90$ observations. The "block" length of the stationary bootstrap is driven by the average value of the first order serial correlation of the series.

4 Conclusion

In this paper, we present an approach for modelling the dynamics of realized covariance matrices. The model we propose explicitly accounts for the empirically observed long memory of financial volatility and allows for inclusion of predictive variables (e.g., traded volume, interest rates, etc.) which have been found to influence volatility. The main feature of our specification is the decomposition of the realized covariance matrices into their Cholesky factors. The dynamics of the elements of the Cholesky decompositions are modelled with a multivariate vector fractionally integrated ARMA (VARFIMA) model without imposing restrictions on the admissible parameter space. By subsequent "squaring" of the forecasted Cholesky elements, we automatically obtain positive definite covariance forecasts.

The model is estimated on six and a half years of daily realized covariances of six stocks traded on the NYSE and shows a reasonable in-sample fit. More importantly, we assess its forecasting performance by applying it to an optimal portfolio selection problem. We compare the resulting optimal portfolio returns to the returns generated by using forecasts of two well established multivariate GARCH models, the DCC of Engle (2002) and the BEKK of Engle and Kroner (1995). By employing tests for stochastic dominance, we show that among these three alternatives, *any* risk-averse investor would achieve the highest expected utility by using our model's forecasts to optimize his portfolio.

The methodology presented in this study can be extended in a number of ways: one interesting direction is to consider alternative estimation techniques such as non-linear least squares and minimum distance estimation to overcome some of the difficulties associated with the maximum likelihood approach used in this study. In order to fully realize the potential of our methodology as well as to further test the performance of the model, we believe it to be worthwhile to increase the number of assets under consideration as well as to test the model on different time periods.

References

An, S., and P. Bloomfield (1993): "Cox and Reid's Modification in Regression Models with Correlated Errors," Discussion Paper, Department of Statistics, North Carolina State University.

Andersen, T. G., and T. Bollerslev (1997): "Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns," *Journal of Finance*, 52, 975–1005.

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

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003): "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579–625.

Asai, M., M. McAleer, and J. Yu (2006): "Multivariate Stochastic Volatility: A Review," *Econometric Reviews*, 25, 145–175.

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

Barett, G., and S. Donald (2003): "Consistent Tests for Stochastic Dominance," *Econometrica*, 71, 71–104.

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

Bauer, G. H., and K. Vorkink (2006): "Multivariate Realized Volatility," Working Paper, Bank of Canada.

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

Bawa, V. S. (1975): "Optimal Rules for Ordering Uncertain Prospects," *Journal of Financial Economics*, 2, 95–121.

Beran, J. (1995): "Maximum Likelihood Estimation of the Differencing Parameter for Invertible Short and Long Memory Autoregressive Integrated Moving Average Models," *Journal of the Royal Statistical Society*, 57, 659–672.

——— (1998): *Statistics for Long-Memory Processes*. CRC Press.

Black, F. (1976): "Studies of Stock Price Volatility Changes," in *Proceedings of the 1976 Business Meeting of the Business and Economic Statistics Section, American Statitistical Association*, pp. 177–181.

Boehmer, E., C. Jones, and X. Zhang (2008): "Which Shorts are Informed?," *Journal of Finance*, 63, 491–527.

Corsi, F. (2005): "Measuring and Modelling Realized Volatility: from Tick-by-tick to Long Memory," Dissertation, Univerity of Lugano.

Corsi, F., and F. Audrino (2007): "Realized Correlation Tick-By-Tick," Working Paper, University of St. Gallen.

Şerban, M., A. Brockwell, J. Lehoczky, and S. Srivastava (2007): "Modelling the Dynamic Dependence Structure in Multivariate Financial Time Series," *Journal of Time Series Analysis*, 28, 763–782.

Dacorogna, M. M., R. Gençay, U. A. Müller, R. B. Olsen, and O. V. Pictet (2001): *An Introduction to High-Frequency Finance*. Academic Press, London.

Davidson, R., and J.-Y. Duclos (2000): "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality," *Econometrica*, 68, 1435–1464.

Doornik, J. A., and M. Ooms (2004): "Inference and Forecasting for ARFIMA Models with an Application to US and UK Inflation," *Studies in Nonlinear Dynamics & Econometrics*, 8, Article 14.

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

Engle, R., and T. Bollerslev (1986): "Modelling the Persistence of Conditional Variances," *Econometric Reviews*, 5, 1–50.

Engle, R., and K. Kroner (1995): "Multivariate Simultaneous Generalized ARCH," *Econometric Theory*, 11, 122–150.

Engle, R., and K. Sheppard (2007): "Evaluating the Specification of Covariance Models for Large Portfolios," Working Paper, University of Oxford.

Fishburn, P. C. (1980): "Continua of Stochastic Dominance Relations for Unbounded Probability Distributions," *Journal of Mathematical Economics*, 7(3), 271–285.

Fleming, J., C. Kirby, and B. Ostdiek (2001): "The Economic Value of Volatility Timing," *Journal of Finance*, 56, 329–352.

Giraitis, L., P. M. Robinson, and D. Surgailis (2000): "A Model for Long Memory Conditional Heteroskedasticiity," *Annals of Applied Probability*, 10, 1002–1024.

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

Golub, G. H., and C. F. van Loan (1996): *Matrix computations*. Johns Hopkins University Press.

Gourieroux, C., J. Jasiak, and R. Sufana (2005): "The Wishart Autoregressive Process of Multivariate Stochastic Volatility," Working Paper, University of Toronto.

Gourieroux, C., and A. Monfort (1995): *Statistics and Econometric Models, Vol.1,*. Cambridge University Press, Cambridge.

Hansen, P. R., and A. Lunde (2005): "A Realized Variance for the Whole Day Based on Intermittent High-Frequency Data," *Journal of Financial Econometrics*, 3, 525–554.

Kaur, A., B. P. Rao, and H. Singh (1994): "Testing for Second Order Stochastic Dominance of Two Distributions," *Econometric Theory*, 10, 849–866.

Klekan, L., R. McFadden, and D. McFadden (1991): "A Robust Test for Stochastic Dominance," Working Paper, Department of Economics, MIT.

Koopman, S. J., B. Jungbacker, and E. Hol (2005): "Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements," *Journal of Empirical Finance*, 12, 445–475.

Lamoureaux, C., and W. Lastrapes (1990): "Heteroskedasticity in Stock Return Data: Volume versus GARCH effects," *Journal of Finance*, 45(1), 221–229.

Linton, O., E. Maasoumi, and Y. Whang (2005): "Consistent Testing for Stochastic Dominance under General Sampling Schemes," *Review of Economic Studies*, 72, 735–765.

Lütkepohl, H. (2005): *New Introduction to Multiple Time Series Analysis*. Springer, Berlin.

Markowitz, H. (1952): "Portfolio Selection," *Journal of Finance*, 7, 77–91.

McFadden, D. (1989): "Testing for Stochastic Dominance," in *Studies in the Economics of Uncertainity*, ed. by T. Fomby, and T. Seo, pp. 113–134. Springer-Verlag.

Oomen, R. (2001): "Using High Frequency Stock Market Index Data to Calculate, Model and Forecast Realized Return Variance," Economics Discussion Paper No. 2001/6, European University Institute.

Politis, D., and J. Romano (1994a): "Large Sample Confidence Regions based on Subsamples under Minimal Assumptions," *Annals of Statistics*, 22, 2031–2050.

——— (1994b): "The Stationary Bootstrap," *Journal of the American Statistical Assosiation*, 89, 1303–1313.

Politis, D., J. Romano, and M. Wolf (1999): *Subsampling*. Springer-Verlag, New York.

Robinson, P. M. (1991): "Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression," *Journal of Econometrics*, 47, 67–84.

Schwert, G. W. (1989): "Why does Stock Market Volatility Change over Time," *Journal of Finance*, 44, 1115–1153.

Sowell, F. (1989): "Maximum Likelihood Estimation of Fractionally Integrated Time Series Models," Working Paper, Carnegie Mellon University.

——— (1992): "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models," *Journal of Econometrics*, 53, 165–188.

Voev, V. (2007): "Dynamic Modelling of Large-Dimensional Covariance Matrices," in *Recent Developments in High Frequency Financial Econometrics*, ed. by L. Bauwens, W. Pohlmeier, and D. Veredas. Springer, Berlin.

A Appendix: Proofs

Derivation of the marginal effects in Equations (9) and (10)

We derive the expressions of $G_{i,j,s}$, $G_{i,i,s}$, $F^{s,t}_{ii,ij}$ and $F^{s,t}_{ij,ij}$ from Section 3, in the case of the restricted version of the VARFIMA model (Model 2), estimated in the present study. Given Equation (4), Model 2 can be written as follows:

$$(1-\phi L)D(L)[X_t - c] = \varepsilon_t + \theta\varepsilon_{t-1}, \qquad \varepsilon_t \sim N(0,\Sigma) \tag{A1}$$

where ϕ and θ are scalars, $D(L) = diag\{(1-L)^d, \ldots, (1-L)^d\}$ is of dimension $m \times m$, and c is a vector of dimension $m \times 1$. In this context, we can write Equation (A1) as:

$$(1-\phi L)(1-L)^d[X_{l,t} - c_l] = \varepsilon_{l,t} + \theta\varepsilon_{l,t-1}, \quad \varepsilon_{l,t} \sim N(0,\sigma_{ll}) \quad l = 1,\ldots,m, \tag{A2}$$

where σ_{ll} is the (l,l)-element of Σ. Given the representation in Equation (5), we can write Equation (A2) as follows:

$$(1-\phi L)[X_{l,t} - c_l + \sum_{h=1}^{\infty}\delta_h(X_{l,t-h} - c_l)] = \varepsilon_{l,t} + \theta\varepsilon_{l,t-1},$$

where $\delta_h = \prod_{0<r\leq h}\frac{r-1-d}{r}, h = 1,2,\ldots$. Thus for each $l = 1,\ldots,m$:

$$X_{l,t} = c_l + (\phi - \delta_1)(X_{l,t-1} - c_l) + \sum_{h=2}^{\infty}(\phi\delta_{h-1} - \delta_h)(X_{l,t-h} - c_l) + \varepsilon_{l,t} + \theta\varepsilon_{l,t-1},$$

From the expression above, we can derive the conditional expectation of $Y_{ij,t+s}$ from Equation (8) for any $s \geq 1$ and $i,j = 1,\ldots,n$ with $j \geq i$. We focus here on $s = 1$ (generalization to $s > 1$ is straightforward):

$$\begin{aligned} E_t[Y_{ij,t+1}] &= \sum_{l=1+\frac{i(i-1)}{2}}^{\frac{i(i+1)}{2}} E_t\left[X_{l,t+1}X_{l+\frac{j(j-1)}{2}-\frac{i(i-1)}{2},t+1}\right] \\ &\equiv \sum_{l=1+\frac{i(i-1)}{2}}^{\frac{i(i+1)}{2}} E_t\left[X_{l,t+1}X_{p,t+1}\right] \end{aligned}$$

$$
\begin{aligned}
&= \sum_{l=1+\frac{i(i-1)}{2}}^{\frac{i(i+1)}{2}} E_t[c_l c_p + c_l(\phi-\delta_1)(X_{p,t}-c_p) + c_l\sum_{h=2}^{\infty}(\phi\delta_{h-1}-\delta_h)(X_{p,t-h+1}-c_p) \\
&+ c_l\varepsilon_{p,t+1} + c_l\theta\varepsilon_{p,t} + c_p(\phi-\delta_1)(X_{l,t}-c_l) + (\phi-\delta_1)^2(X_{l,t}-c_l)(X_{p,t}-c_p) \\
&+ \sum_{h=2}^{\infty}(\phi-\delta_1)(\phi\delta_{h-1}-\delta_h)(X_{l,t}-c_l)(X_{p,t-h+1}-c_p) + (\phi-\delta_1)(X_{l,t}-c_l)\varepsilon_{p,t+1} \\
&+ (\phi-\delta_1)\theta(X_{l,t}-c_l)\varepsilon_{p,t} + c_p\sum_{h=2}^{\infty}(\phi\delta_{h-1}-\delta_h)(X_{l,t-h+1}-c_l) + \sum_{h=2}^{\infty}(\phi-\delta_1)(\phi\delta_{h-1}-\delta_h) \\
&\times (X_{p,t}-c_p)(X_{l,t-h+1}-c_l) + \sum_{h=2}^{\infty}(\phi\delta_{h-1}-\delta_h)(X_{l,t-h+1}-c_l)\sum_{h=2}^{\infty}(\phi\delta_{h-1}-\delta_h)(X_{p,t-h+1}-c_p) \\
&+ \sum_{h=2}^{\infty}(\phi\delta_{h-1}-\delta_h)(X_{l,t-h+1}-c_l)\varepsilon_{p,t+1} + \sum_{h=2}^{\infty}(\phi\delta_{h-1}-\delta_h)\theta(X_{l,t-h+1}-c_l)\varepsilon_{p,t} \\
&+ c_p\varepsilon_{l,t+1} + (\phi-\delta_1)(X_{p,t}-c_p)\varepsilon_{l,t+1} + \sum_{h=2}^{\infty}(\phi\delta_{h-1}-\delta_h)(X_{p,t-h+1}-c_p)\varepsilon_{l,t+1} \\
&+ \varepsilon_{p,t+1}\varepsilon_{l,t+1} + \theta\varepsilon_{l,t+1}\varepsilon_{p,t} + c_p\theta\varepsilon_{l,t} + \theta(\phi-\delta_1)(X_{p,t}-c_p)\varepsilon_{l,t} \\
&+ \sum_{h=2}^{\infty}(\phi\delta_{h-1}-\delta_h)\theta(X_{p,t-h+1}-c_p)\varepsilon_{l,t} + \theta\varepsilon_{l,t}\varepsilon_{p,t+1} + \theta^2\varepsilon_{l,t}\varepsilon_{p,t}] \\
&\equiv G_{i,j,1}(\boldsymbol{X}_t,\vartheta), \qquad \text{(A3)}
\end{aligned}
$$

where the index p is defined (implicitly as a function of i,j) as $p = l + \frac{j(j-1)}{2} - \frac{i(i-1)}{2}$ and $\vartheta = (c', \phi, \theta, d, \ldots, vech(\Sigma))'$.

From Equation (A3) we derive the expression of $G_{i,j,1}(\boldsymbol{X}_t,\vartheta)$ to be:

$$
\begin{aligned}
G_{i,j,1}(\boldsymbol{X}_t,\vartheta) &= \sum_{l=1+\frac{i(i-1)}{2}}^{\frac{i(i+1)}{2}} (c_l c_p + (\phi-\delta_1)\left[c_l(X_{p,t}-c_p) + c_p(X_{l,t}-c_l) + (\phi-\delta_1)(X_{l,t}-c_l)(X_{p,t}-c_p)\right] \\
&+ \sum_{h=2}^{\infty}(\phi-\delta_1)(\phi\delta_{h-1}-\delta_h)\left[(X_{l,t}-c_l)(X_{p,t-h+1}-c_p) + (X_{p,t}-c_p)(X_{l,t-h+1}-c_l)\right] \\
&+ \sum_{h=2}^{\infty}(\phi\delta_{h-1}-\delta_h)\left[c_l(X_{p,t-h+1}-c_p) + c_p(X_{l,t-h+1}-c_l)\right] \\
&+ \sum_{h=2}^{\infty}(\phi\delta_{h-1}-\delta_h)(X_{l,t-h+1}-c_l)\sum_{h=2}^{\infty}(\phi\delta_{h-1}-\delta_h)(X_{p,t-h+1}-c_p) \\
&+ \sum_{h=2}^{\infty}\theta(\phi\delta_{h-1}-\delta_h)\left[\varepsilon_{p,t}(X_{l,t-h+1}-c_l) + \varepsilon_{l,t}(X_{p,t-h+1}-c_p)\right] \\
&+ \theta\left[c_l\varepsilon_{p,t} + c_p\varepsilon_{l,t}\right] + (\phi-\delta_1)\theta\left[(X_{l,t}-c_l)\varepsilon_{p,t} + \varepsilon_{l,t}(X_{p,t}-c_p)\right] + \Sigma_{l,p} + \theta^2\varepsilon_{l,t}\varepsilon_{p,t}\big). \qquad \text{(A4)}
\end{aligned}
$$

In a similar manner, we derive the $G_{i,i,1}(X_t,\vartheta)$ to be:

$$\begin{aligned}
G_{i,i,1}(\boldsymbol{X}_t,\vartheta) &= \sum_{l=1+\frac{i(i-1)}{2}}^{\frac{i(i+1)}{2}} \left(c_l^2+(\phi-\delta_1)^2(X_{l,t}-c_l)^2+2c_l(\phi-\delta_1)(X_{l,t}-c_l)\right.\\
&+ \sum_{h=2}^{\infty} 2(\phi-\delta_1)(\phi\delta_{h-1}-\delta_h)(X_{l,t}-c_l)(X_{l,t-h+1}-c_l)\\
&+ \sum_{h=2}^{\infty} 2c_l(\phi\delta_{h-1}-\delta_h)(X_{l,t-h+1}-c_l)\\
&+ \left[\sum_{h=2}^{\infty}(\phi\delta_{h-1}-\delta_h)(X_{l,t-h+1}-c_l)\right]^2+2(\phi-\delta_1)\theta(X_{l,t}-c_l)\varepsilon_{l,t}\\
&+ \sum_{h=2}^{\infty} 2\theta(\phi\delta_{h-1}-\delta_h)\varepsilon_{l,t}(X_{l,t-h+1}-c_l)\\
&+ \left.2\theta c_l\varepsilon_{l,t}+\Sigma_{l,l}+\theta^2\varepsilon_{l,t}^2\right), \qquad \text{(A5)}
\end{aligned}$$

Given that $G_{i,j,0}(\boldsymbol{X}_t,\vartheta)=\sum_{l=1+\frac{i(i-1)}{2}}^{\frac{i(i+1)}{2}} X_{l,t}X_{p,t}$, we can derive $F_{ii,ij}^{s,t}$ and $F_{ij,ii}^{s,t}$ from Equations (9) and (10) for any (i,j) combination.
For example, the marginal effect of the volatility $Y_{11,t}$ at time t on the conditional expectation of the covariance $Y_{12,t+1}$ at time $t+1$ is given by:

$$F_{12,11}^{1,t}(\boldsymbol{X}_t,\vartheta)=\frac{\partial E_t[Y_{12,t+1}]}{\partial Y_{11,t}}=\frac{\partial G_{1,2,1}(\boldsymbol{X}_t,\vartheta)}{\partial G_{1,1,0}(\boldsymbol{X}_t,\vartheta)}=\frac{\partial G_{1,2,1}(\boldsymbol{X}_t,\vartheta)}{\partial X_{1,t}^2},$$

where $G_{1,2,1}(\boldsymbol{X}_t,\hat{\vartheta})$ is obtained from Equation (A4). Thus

$$F_{12,11}^{1,t}(\boldsymbol{X}_t,\vartheta) = \frac{(\phi-\delta_1+\theta)\left(c_2+(\phi-\delta_1)(X_{2,t}-c_2)+\theta\varepsilon_{2,t}+\sum_{h=2}^{\infty}(\phi\delta_{h-1}-\delta_h)(X_{2,t-h+1}-c_2)\right)}{2X_{1,t}}.$$

Similarly we can derive the marginal effect of the covariance $Y_{12,t}$ on the conditional expectation of the volatility $Y_{11,t+1}$ denoted by $F_{11,12}^{1,t}$:

$$\begin{aligned}
F_{11,12}^{1,t}(\boldsymbol{X}_t,\vartheta) &= \frac{\partial E_t[Y_{11,t+1}]}{\partial Y_{12,t}}=\frac{\partial G_{1,1,1}(\boldsymbol{X}_t,\vartheta)}{\partial G_{1,2,0}(X_t,\vartheta)}=\frac{\partial G_{1,1,1}(\boldsymbol{X}_t,\vartheta)}{\partial X_{1,t}X_{2,t}}\\
&= \frac{\partial G_{1,1,1}(\boldsymbol{X}_t,\vartheta)}{\partial X_{1,t}}\frac{\partial X_{1,t}}{\partial X_{1,t}X_{2,t}}+\frac{\partial G_{1,1,1}(\boldsymbol{X}_t,\vartheta)}{\partial X_{2,t}}\frac{\partial X_{2,t}}{\partial X_{1,t}X_{2,t}},
\end{aligned}$$

where $G_{1,1,1}(\boldsymbol{X}_t,\vartheta)$ is obtained from Equation (A5). Thus

$$F_{11,12}^{1,t}(\boldsymbol{X}_t,\vartheta) = \frac{2(\phi-\delta_1+\theta)[c_1+(\phi-\delta_1)(X_{1,t}-c_1)+\theta\varepsilon_{1,t}+\sum_{h=2}^{\infty}(\phi\delta_{h-1}-\delta_h)(X_{1,t-h+1}-c_1)]}{X_{2,t}}.$$

B Appendix: Tables and Figures

Table B.1: Summary statistics of 5-minute and daily stock returns

Stock	Mean	Max	Min	Std. dev	Skew	Kurt
5-minute returns						
AXP	0.0113	0.0703	-0.1843	0.0022	-4.7063	485.0690
HWP	-0.0016	0.1112	-0.1597	0.0031	-1.0157	256.3915
JPM	-0.0037	0.0774	-0.1186	0.0025	-0.9637	137.7105
HD	-0.0241	0.1082	-0.1271	0.0024	-2.2291	270.6422
C	0.0006	0.0845	-0.1035	0.0022	-0.4016	157.5951
IBM	-0.0119	0.1086	-0.1071	0.0019	1.5253	307.8203
Daily returns						
AXP	1.1391	0.1034	-0.1464	0.0193	-0.2277	8.5927
HWP	0.3494	0.1567	-0.2066	0.0267	-0.0234	10.7708
JPM	-0.2844	0.1578	-0.2019	0.0218	0.0683	13.7154
HD	-1.7161	0.1228	-0.1509	0.0210	-0.2066	9.2915
C	0.1761	0.1178	-0.1726	0.0184	-0.4100	13.2778
IBM	-0.7115	0.1173	-0.1106	0.0177	0.4465	10.2498

Note: This table reports descriptive statistics of the 5-minute and daily returns for the stocks AXP, C, HWP, JPM, HD and IBM over the period from 1^{st} January 2001 to 30^{th} June 2006. The means are scaled by 10^4.

Table B.2: Summary statistics of realized variances and realized covariances of the stocks AXP, C, HWP, JPM HD and IBM

Stock	Mean	Max	Min	Std. dev	Skew	Kurt
Realized Variance						
AXP	0.0390	0.8339	0.0011	0.0635	5.4105	46.2969
HWP	0.0656	1.4397	0.0028	0.0961	6.6996	75.0095
JPM	0.0490	2.8130	0.0017	0.1083	14.9024	334.1691
HD	0.0413	0.7317	0.0012	0.0533	4.9629	41.7344
C	0.0386	1.4113	0.0013	0.0738	10.0771	151.6528
IBM	0.0267	0.8111	0.0013	0.0387	8.1390	131.8510
Realized Covariance						
AXP-HWP	0.0154	0.4085	-0.0145	0.0290	6.3709	61.8298
AXP-JPM	0.0169	0.6035	-0.0791	0.0325	8.3144	117.355
AXP-HD	0.0143	0.3223	-0.0060	0.0256	5.5456	48.0592
AXP-C	0.0171	0.4900	-0.0130	0.0312	5.8290	50.4219
AXP-IBM	0.0128	0.3288	-0.0185	0.0226	5.0769	43.5852
HWP-JPM	0.0170	0.4047	-0.0054	0.0294	6.2420	59.3477
HWP-HD	0.0150	0.3183	-0.1175	0.0249	7.0555	82.6114
HWP-C	0.0171	0.2913	-0.0473	0.0270	4.7059	34.1161
HWP-IBM	0.0150	0.3334	-0.0026	0.0233	14.3477	317.0420
JPM-HD	0.0152	0.3637	-0.0345	0.0268	5.8616	56.1417
JPM-C	0.0221	1.2769	-0.0552	0.0498	6.2820	61.3870
JPM-IBM	0.0141	0.4329	-0.0098	0.0253	5.3664	48.0463
HD-C	0.0156	0.4063	-0.0051	0.0269	7.4878	92.5154
HD-IBM	0.0127	0.2234	-0.0037	0.0195	4.8518	35.7024
C-IBM	0.0142	0.4839	-0.0151	0.0252	7.8865	110.4219

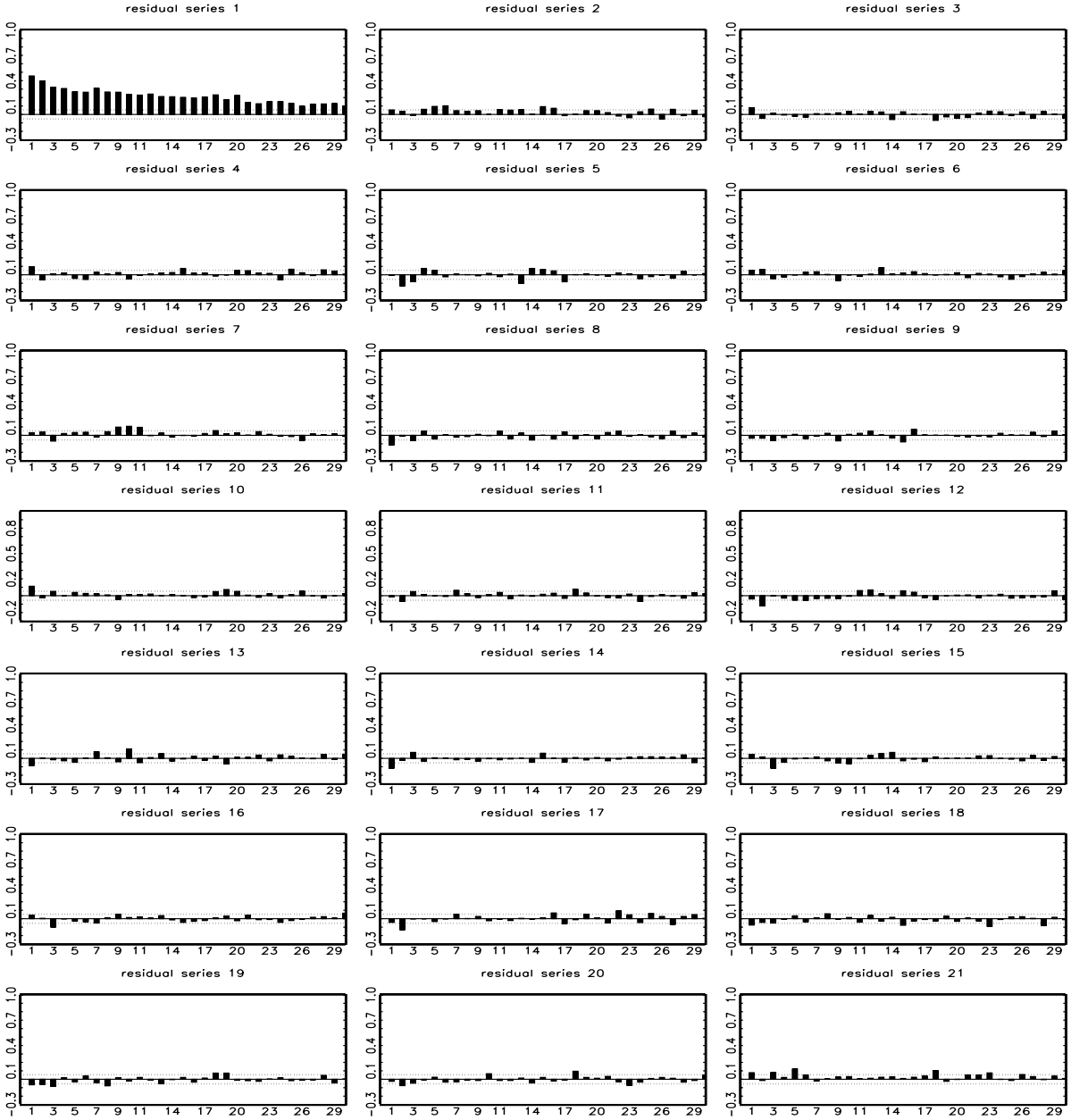
Note: This table reports the descriptive statistics of realized covariances and variances of the six stocks. The realized variances and covariances are calculated from 5-minute intraday returns, as described in the main text. The realized variances and covariances are scaled by 10^2.

Table B.3: Estimation results of the diagonal BEKK(1,1,1) and DCC model

Parameter/Stock	AXP	HWP	JPM	HD	C	IBM
diagonal BEKK(1,1,1)						
μ_i	0.1039	0.0578	-0.1042	0.0556	-0.0349	-0.0886
	(0.0689)	(0.0523)	(0.0542)	(0.0417)	(0.0425)	(0.0290)
C	0.0420	-0.0221	-0.0301	-0.0892	0.0485	-0.0139
	(0.0219)	(0.0386)	(0.0246)	(0.0441)	(0.0223)	(0.0278)
		-0.0538	-0.0060	0.0706	0.0394	-0.0393
		(0.0235)	(0.0128)	(0.0173)	(0.0253)	(0.0456)
			-0.0042	-0.0171	0.0032	-0.1054
			(0.0288)	(0.0207)	(0.0136)	(0.0389)
				0.1607	0.0850	0.1946
				(0.0851)	(0.0117)	(0.0485)
					0.1499	0.1777
					(0.0304)	(0.0292)
						0.1412
						(0.0340)
$diag(A)$	0.9845	0.9947	0.9788	0.9870	0.9814	0.9857
	(0.0153)	(0.0014)	(0.0097)	(0.0056)	(0.0054)	(0.0068)
$diag(B)$	0.0617	0.0620	0.0353	0.0336	0.0203	0.0180
	(0.0410)	(0.0611)	(0.0358)	(0.0440)	(0.0309)	(0.0350)
DCC of Engle (2002)						
μ_i	0.0717	0.0490	0.0313	0.0182	0.0264	0.0262
	(0.0354)	(0.0589)	(0.0340)	(0.0418)	(0.0320)	(0.0534)
w_i	0.0236	0.0144	0.0117	0.0155	0.0167	0.0273
	(0.0157)	(0.0163)	(0.0079)	(0.0139)	(0.0137)	(0.0550)
α_i	0.0867	0.0097	0.0658	0.0403	0.0670	0.0714
	(0.0341)	(0.0046)	(0.0270)	(0.0138)	(0.0359)	(0.1191)
β_i	0.9087	0.9871	0.9315	0.9549	0.9266	0.9194
	(0.0314)	(0.0065)	(0.0252)	(0.0165)	(0.0373)	(0.1259)
	θ_1	0.0067		θ_2	0.9776	
		(0.0031)			(0.0139)	

Note: QML standard errors are reported in parenthesis.

Figure B.1: ACF of the standardized residuals of the restricted VARFIMA model (Model 2)



Research Papers 2008

2008-25: Mark Podolskij and Mathias Vetter: Bipower-type estimation in a noisy diffusion setting

2008-26: Martin Møller Andreasen: Ensuring the Validity of the Micro Foundation in DSGE Models

2008-27: Tom Engsted and Thomas Q. Pedersen: Return predictability and intertemporal asset allocation: Evidence from a bias-adjusted VAR model

2008-28: Frank S. Nielsen: Local polynomial Whittle estimation covering non-stationary fractional processes

2008-29: Per Frederiksen, Frank S. Nielsen and Morten Ørregaard Nielsen: Local polynomial Whittle estimation of perturbed fractional processes

2008-30: Mika Meitz and Pentti Saikkonen: Parameter estimation in nonlinear AR-GARCH models

2008-31: Ingmar Nolte and Valeri Voev: Estimating High-Frequency Based (Co-) Variances: A Unified Approach

2008-32: Martin Møller Andreasen: How to Maximize the Likelihood Function for a DSGE Model

2008-33: Martin Møller Andreasen: Non-linear DSGE Models, The Central Difference Kalman Filter, and The Mean Shifted Particle Filter

2008-34: Mark Podolskij and Daniel Ziggel: New tests for jumps: a threshold-based approach

2008-35: Per Frederiksen and Morten Ørregaard Nielsen: Bias-reduced estimation of long memory stochastic volatility

2008-36: Morten Ørregaard Nielsen: A Powerful Test of the Autoregressive Unit Root Hypothesis Based on a Tuning Parameter Free Statistic

2008-37: Dennis Kristensen: Uniform Convergence Rates of Kernel Estimators with Heterogenous, Dependent Data

2008-38: Christian M. Dahl and Emma M. Iglesias: The limiting properties of the QMLE in a general class of asymmetric volatility models

2008-39: Roxana Chiriac and Valeri Voev: Modelling and Forecasting Multivariate Realized Volatility