Discussion Paper Series

A Cholesky-MIDAS model for predicting stock portfolio volatility

By

Ralf Becker, Adam Clements, Robert O'Neill

Centre for Growth and Business Cycle Research, Economic Studies, University of Manchester, Manchester, M13 9PL, UK

September 2010
Number 149

A Cholesky-MIDAS model for predicting stock portfolio volatility*

Ralf Becker[†], Adam Clements[‡] and Robert O'Neill[†]
[†]Economics, School of Social Sciences, University of Manchester
[‡]School of Economics and Finance, Queensland University of Technology

August 26, 2010

Abstract

This paper presents a simple forecasting technique for variance covariance matrices. It relies significantly on the contribution of Chiriac and Voev (2010) who propose to forecast elements of the Cholesky decomposition which recombine to form a positive definite forecast for the variance covariance matrix. The method proposed here combines this methodology with advances made in the MIDAS literature to produce a forecasting methodology that is flexible, scales easily with the size of the portfolio and produces superior forecasts in simulation experiments and an empirical application.

Corresponding author
Ralf Becker
Economics, School of Social Sciences
University of Manchester

email `ralf.becker@manchester.ac.uk`
Ph +44 (0)161 275 4807.

1 Introduction

In recent decades modelling and forecasting the volatility of financial assets has been a fertile ground for research. For a long period this research focused on modelling and estimating volatility in a univariate context (see Poon and Granger, 2003). More recently, an increasing amount of research has been concerned with modelling entire variance-covariance matrices, an issue of importance in finance (see Andersen, Bollerslev, Christoffersen and Diebold, 2006) when considering portfolios of financial assets. It is this literature this paper seeks to contribute to.

The volatility forecasting literature has seen exciting developments based on the availability of intraday data and more precise volatility measurements. Two research strands, in the context of univariate volatility forecasting, have emerged. One based on traditional univariate time-series

*The authors gratefully acknowledge comments from Denise Osborn, Arthur Sinko and seminar participants at Lancaster University and The University of Manchester. The responsibility for errors or omissions lies entirely with the authors.

models, using measurements of realized volatilities (Andersen *et al.*, 2006), the second making direct use of high-frequency data (Ghysels, Santa-Clara and Valkanov 2004, 2006) labeled the Mixed Data Sampling (MIDAS) approach.

The literature on multivariate variance-covariance matrix (VCM) modelling and forecasting has been somewhat slower in utilising intraday data as two additional issues arise in this context. While the dynamic conditional correlation (DCC) model of Engle and Sheppard (2001) has been used extensively for low-dimensional portfolios, it has proven difficult to estimate highly parameterised models for high-dimensional systems. Optimal parameter estimates are difficult to find without reducing the parameter space which leads to constraints on the correlation dynamics. Engle (2008) and Engle, Shephard and Sheppard (2008) demonstrate how such systems can be estimated efficiently, ensuring positive definiteness of VCMs, although these models do not make use of intraday data.

While using intraday data allows econometricians to model variances and covariances as observed variables, the issue of positive definitness and ensuring a reasonably sized parameter space remain very much relevant. A promising approach to address the former issue is proposed by Chiriac and Voev (2010) who model the elements of the Cholesky decomposition of realized VCMs, calculated on the basis of intraday data. Modelling the behaviour of the elements of the decomposition requires no restrictions while ensuring positive definiteness of the corresponding VCMs. However, the dynamics of the elements of the Cholesky decomposition remain restricted to ensure a reasonable parameter space.

In this paper we combine this latter framework with the MIDAS type approach for utilising data observed at different frequencies. The resulting model offers a number of advantages. *First*, as in Chiriac and Voev (2010), it uses high frequency data to obtain relatively precise measurements of the latent variance-covariance matrices, an approach which has delivered significant improvements in forecast performance in the univariate context (see Anderson, Bollerslev, Diebold and Labys 2003 and Koopman, Jungbacker and Hol 2005). *Second*, despite allowing for complicated lag structures, the resulting nonlinear optimisation problem is low dimensional and therefore feasible even for large portfolios. *Third*, the modelling framework can, in a straightforward manner, be extended to include any weakly exogenous variables.

Simulation evidence illustrates that the proposed estimation methodology can deliver significantly improved multi-step ahead VCM forecasts when compared to the most popular method based solely on daily returns, the dynamic conditional correlation (DCC) approach of Engle and Sheppard (2001). This superiority is evident even when the data follows a DCC process.

The remainder of this paper is structured as follows, The next section reviews the literature in several areas which are important in developing our model, while Section 3 introduces our approach to modelling and forecasting variance covariance matrices. Section 4 describes how we compare competing forecasts. This is followed in Section 5 by a simulation study, which compares the forecasting performance of various methods for small portfolios. Empirical evidence for the method's performance in the context of a large scale portfolio is presented in Section 6. This section includes simulation and empirical evidence. The final section concludes and notes future areas of interest.

2 Literature Review

This paper draws closely on the contributions made in a number of different areas of the vast volatility and correlation modelling literature. In particular it relies on research that outlines how to use high-frequency data to obtain proxies for latent variance-covariance matrices, the mixed data sampling (MIDAS) literature, that has demonstrated how to efficiently use information observed at varying sampling frequencies and lastly the literature that illuminates how to ensure that forecasts of variance-covariance matrices are positive definite.

Developments in trading technology have allowed the recording of incredibly detailed financial market data. However, only fairly recently, has it been demonstrated how to harness this data effectively in order to facilitate the measurement, modelling and forecasting of volatility. The contributions in this area are too numerous to be reviewed here in any great detail (see Andersen *et al.* 2006, for an overview). It should, however, be mentioned that despite the conceptual beauty of realized volatilities and covariances, there are numerous practical issues such as missing observations (Hansen and Lunde, 2005), microstructure noise (Zhang, Mykland and Aït-Sahalia 2005), measurement error (Hansen and Lunde, 2006) and the presence of discontinuous jump processes (Barndorff-Nielsen and Sheppard, 2006) to be potentially dealt with. Each of these issues complicates the computation of realized volatilities although, in principle, technologies have been developed to deal with them. This does not, however, apply to the computation of realized covariances, for which the above issues remain largely unexplored. In addition, when computing realized covariances, and subsequently realized correlations, new issues arise such as non-synchronicity in trading and the Epps effect. Hayashi and Yoshida (2005) suggest a solution to deal with non-synchronous trading, though this issue is not a focus of the present study. The Epps effect, as discussed in Epps (1979), reflects the fact that correlations (and covariances) between financial assets decrease as we use an increasingly fine sampling interval to compute realized covariances. This paper relies on these contributions as it introduces a forecasting model which relies on realized variance-covariance matrices constructed from high-frequency return data. When the forecasting method presented here is applied in practice, all the above issues will have to be dealt with prior to the application of the model presented here.

The second area of recent developments this paper is indebted to is the MIDAS literature, that began with Ghysels, Santa-Clara and Valkanov (2004, 2006). This literature has revived old techniques of lag parameterizations that allow for a flexible, yet parsimonious, utilization of many lags of lagged explanatory variables. MIDAS has been employed to investigate several economic relationships[1]. In the context of univariate volatility modelling the approach has proven to be useful when applied to multi-step ahead volatility forecasting (see Ghysels, Santa-Clara and Valkanov, 2004). Ghysels and Sinko (2006) and Ghysels, Sinko and Valkanov (2008) further investigate the use of daily realized data, corrected for microstructure noise, as regressors in volatility forecasting models and find that unadjusted measures of realized volatility provide superior forecasts than their noise adjusted counterparts. Lastly, Ghysels *et al.* (2008) also discuss the extension of the MIDAS approach to a multivariate setting and emphasize the flexibility of the method, which makes it attractive when forecasting the variance-covariance matrices of stock portfolios over a trading month.

The MIDAS technology has also been utilised in combination with more established approaches to multivariate volatility modelling. Colacito, Engle and Ghysels (2007) introduce the DCC-MIDAS model in which the correlation matrix of stocks is allowed to have a long-run time-varying component, modelled using MIDAS technology. Although the same paper discusses the conditions required for ensuring that the resulting correlation matrices are positive semi-definite, these restrictions become complex even for a relatively small number of assets, limiting the practical appeal of the model for high dimensions. This helps to highlight the challenge of ensuring the positive semi-definiteness of variance-covariance matrices that has been tackled in a number of different ways, most prominently in the finance literature by conditional correlation models, led by the DCC model of Engle and Sheppard (2001)[2]. This typically employs daily return observations and uses restricted parameter values in order to ensure that forecast correlation matrices are positive definite.

[1] For example Ghysels, Santa-Clara and Valkanov (2005) use MIDAS to provide evidence of the existence of a risk-return tradeoff, Clements & Galvão (2008) use it to improve forecasts of US quarterly output growth and it has been used in the forecasting of aggregate output and employment by Armesto, Hernández-Murillo, Owyang & Piger (2007).

[2] Other variants of the conditional correlation approach include the CCC model of Bollerslev (1990), the RSDC model of Pelletier (2006), the block-DCC of Billio, Caporin and Gobbo (2003) and the asymmetric-DCC of Cappiello, Engle and Sheppard (2006).

In contrast Chiriac and Voev (2010) propose to tackle the issue of positive semi-definiteness using measurements of variance covariance matrices derived from high-frequency data. Their key contribution is the use of Cholesky decompositions to ensure that forecasts of the VCMs are positive semi-definite. This is achieved by modelling and forecasting the elements of the Cholesky decomposition, which are not subject to any restrictions. When transforming such forecasts back into variance-covariance matrices, positive semi-definiteness is guaranteed via the properties of the decomposition.

This paper draws on a number of the contributions described above to devise a modelling and forecasting strategy for multi-step ahead variance-covariance matrices based on high-frequency data. As in Chiriac and Voev (2010) we focus on the elements of the Cholesky decomposition, but in contrast to that paper utilise a generalized MIDAS approach in modelling their behaviour.

3 Cholesky-MIDAS Methodology

3.1 Assumptions & Background

The model we present is used to forecast the VCM of stock returns in an n stock portfolio. For any given day, t, the $n \times 1$ vector of returns is denoted by $\mathbf{r}_t = (r_{1t}, ..., r_{nt})'$, where r_{it} is the return on stock i on day t, and we assume that given all information available at time $t-1$, $\mathcal{F}_{t-1}$, the mean is unforecastable, i.e. $E(\mathbf{r}_t|\mathcal{F}_{t-1}) = 0$. The object of interest is the $n \times n$ conditional variance-covariance matrix of returns, $Var(\mathbf{r}_t|\mathcal{F}_{t-1}) = \mathbf{\Sigma}_t$, which we assume to be time-varying, predictable, and although unobserved, can be consistently estimated by a realized variance-covariance matrix $\mathbf{V}_t$.[3]

3.2 Realized Variance-Covariance Matrix Calculation

The model presented below relies on the calculation of realized variance covariance matrices over both single and multi day frequencies, and this section provides a brief introduction to how these realizations are obtained. The approach is covered in more depth in appendix A and a more detailed introduction to the subject can be found in Andersen et al (2006).

Within trading day t we obtain an $n \times 1$ vector of stock returns over each x minute trading period, $\mathbf{r}_{q,t}$, where $q = 1, .., Q$ in a day containing Q trading periods. The realized covariance matrix relating to the trading portion on day t is the sum of the products of these vectors, i.e. $\mathbf{V}_{TR,t} = \sum_{q=1}^{Q} \mathbf{r}_{q,t}\mathbf{r}'_{q,t}$, where $\mathbf{V}_{TR,t}$ is the realized variance covariance matrix of the trading period

[3] Generally in this paper $\mathbf{\Sigma}$ represents the actual VCM (unobserved except in simulations), $\mathbf{V}$ is an observed realized value of the VCM, calculated from intraday data and $\mathbf{H}$ is used to denote a forecast of the matrix.

of day t.

As we forecast the VCM over a period in excess of a single trading day, we are also interested in the additional volatility attributable to the time over which stocks are not traded, thus we follow one of the methods proposed in Hansen and Lunde (2005) for consistent estimation of the 24-hour realized volatility matrix. Specifically the close to open period is treated as a separate return period so that the total 24-hour realized covariance matrix for day t, $\mathbf{V}_t$, is computed as $\mathbf{V}_t = \mathbf{r}_{co,t}\mathbf{r}'_{co,t} + \sum_{q=1}^{Q} \mathbf{r}_{q,t}\mathbf{r}'_{q,t}$ where $\mathbf{r}_{co,t}$ is the vector of close to open returns on day t.

Once we have obtained the daily values of the realized variance-covariance matrix we can find the matrix for an m day long period by summing the m daily realized matrices. In this paper we denote this matrix by $\mathbf{V}_t^{(m)}$, which is the matrix covering the period $t-m+1$ to t hence $\mathbf{V}_t^{(m)} = \sum_{i=t-m+1}^{t} \mathbf{V}_i$.

3.3 General CD-MIDAS Procedure

This section introduces the Cholesky–decomposition MIDAS methodology (CD-MIDAS), which we use to model and forecast variance-covariance matrices. For an $(n \times n)$ dimensional VCM this involves the following steps:

1. Use high frequency return data to calculate daily realized VCMs, $\mathbf{V}_t$, and realized VCMs over periods corresponding to the forecasting horizon of m days, $\mathbf{V}_t^{(m)}$. The notation $\mathbf{V}_t^{(m)}$ refers to a time period that is m days long ending at time t, hence $\mathbf{V}_t^{(m)}$ is the realized VCM for the time period $t-m+1$ to t.

2. Compute the Cholesky decomposition of realized VCMs at horizons of both 1 and m days, such that

$$\mathbf{V}_t = \mathbf{C}_t\mathbf{C}'_t \tag{1}$$

$$\mathbf{V}_t^{(m)} = \mathbf{C}_t^{(m)}\mathbf{C}_t^{(m)\prime} \tag{2}$$

and the related $(\widetilde{n} \times 1)$, $\widetilde{n} = n(n+1)/2$, vectors of unique lower diagonal elements are

$$\mathbf{P}_t = vech(\mathbf{C}_t) \tag{3}$$

$$\mathbf{P}_t^{(m)} = vech\left(\mathbf{C}_t^{(m)}\right). \tag{4}$$

3. Construct MIDAS models in which each of the $\widetilde{n}$ elements of $\mathbf{P}_{t+m}^{(m)}$ is modelled as a function of lags of (potentially) all elements in $\mathbf{P}_t, ..., \mathbf{P}_{t-K+1}$, where K is the maximum number of daily lags in the MIDAS specification (A more detailed discussion of the methodology in this step is provided in Section 3.4.).

4. Use the MIDAS models estimated in step 3 to forecast m periods ahead, $\widehat{\mathbf{P}}_{T+m}^{(m)}$, using the observed realizations of $\mathbf{P}_T, ..., \mathbf{P}_{T-K+1}$.

5. Populate $\widehat{\mathbf{C}}_{T+m}^{(m)}$ with the appropriate elements from $\widehat{P}_{T+m}^{(m)}$ and produce a VCM forecast $\mathbf{H}_{T+m}^{(m)}$ according to
$$\mathbf{H}_{T+m}^{(m)} = \widehat{\mathbf{C}}_{T+m}^{(m)} \widehat{\mathbf{C}}_{T+m}^{(m)\prime}. \tag{5}$$
The output, $\mathbf{H}_{T+m}^{(m)}$, is a forecast of the VCM over the period from $t+1$ to $t+m$.

Steps 1, 2 and 4 rely on calculations of realized variance-covariance matrices and transformations of the Cholesky decomposition. As noted in Chiriac and Voev (2010) the decomposition is particularly useful in this context as any operation of the type shown in equations (1) and (5) yields a positive definite variance covariance matrix as long as $\mathbf{C}$ is a lower diagonal matrix of real numbers, no other restrictions are required, making it possible to model each unique element in $\mathbf{C}$ individually.

3.4 MIDAS Specification

This section provides more detail on the forecasting model used in step 3. The CD-MIDAS specification allows us to model each of the $\widetilde{n}$ unique elements of the Cholesky decomposition via a single equation. As discussed above, the only constraint that is to be imposed on the elements in $\widehat{P}_{T+m}^{(m)}$ is that they ought to be real. As this is sufficient to achieve positive definitness in the VCM forecats, it allows the econometricion to contemplate a wide range of models for the individual elements of the Cholesky decomposition. One example is the CD-VARFIMA model proposed in Chiriac and Voev (2010). In the form in which the authors propose to apply it to large dimensional systems that model turns into $\widetilde{n}$ ARFIMA models with a common fractional integration parameter.

Here we propose a model that is somewhat simpler to estimate as it will capture strong persistance by a potentially long lag structure rather than a fractionally integrated process. In particular we propose to emply the MIDAS methodology of Ghysels *et al.* (2004, 2006) in using daily realized observations of the VCM to forecast the same matrix at a lower frequency. The use of tightly parameterised lag functions allows us to be conservative in the number of parameters used while allowing for flexible lag distributions.

In our notation below we denote an individual element of the decomposition as P_{it} in the daily case and $P_{it}^{(m)}$ in the m-period case where $i = 1, ..., \widetilde{n}$. We propose to use a weighted average of past values of P_{it} in a forecasting model for $P_{it}^{(m)}$

$$P_{i,t+m}^{(m)} = \beta_{i0} + \beta_{i1} \sum_{k=1}^{K} B\left(k, 1, \theta_i\right) P_{i,t-k+1} + v_t. \tag{6}$$

where we use a beta lag structure to determine the weights, such that,

$$\begin{aligned} \text{Beta Lag - } B(k,1,\theta_i) &= \frac{f(\frac{k}{K},1,\theta_i)}{\sum_{k=1}^{K} f(\frac{k}{K},1,\theta_i)} &(7) \\ f(z,a,b) &= \frac{z^{a-1}(1-z)^{b-1}\Gamma(a+b)}{\Gamma(a)\Gamma(b)}. &(8) \end{aligned}$$

Although two parameters determine the beta function the first of these is set to unity, such that weights applied to lags are decreasing in k[4]. The maximum lag considered is K. When considering this forecasting model it is important to keep in mind that the elements of P_{it} are nonlinear combinations of elements in the VCM $\mathbf{V}_t^{(m)}$. This implies that the history of P_{it} will also contain the history of all the elements of the VCM that are relevant for the calculation of P_{it}.

It is important to understand that, as in Ghysels *et al.* (2006), the specification in (6) is not a model representing the dynamic process of the elements in $P_t^{(m)}$ but merely a potentially useful forecasting device. In order to use (6) for forcasting three parameters, β_{i0}, β_{i1} and θ_i, require estimation. These parameters are estimated equationwise for all $\widetilde{n}$ elements of $P_t^{(m)}$, allowing for varying degrees of dependence. Each such estimation is a straightforward nonlinear least squares (NLS) estimation. For ease of notation we collect all $3\widetilde{n}$ parameters in the $(3\widetilde{n} \times 1)$ parameter vector $\mathbf{\Theta}$.

After estimation of the CD-MIDAS parameters, we can obtain forecasts for each element of the Cholesky decomposition. We use the K lags of daily data prior to the start of the forecast period to do so. For example if we wish to forecast the VCM over $t+1$ to $t+m$ we do so using the realizations $P_t, P_{t-1}, ..., P_{t-K+1}$ and the set of estimated parameters, $\widehat{\mathbf{\Theta}}$. Thus $\hat{P}_{t+m}^{(m)} = f(P_t, ..., P_{t-K+1}, \widehat{\mathbf{\Theta}})$. The forecast VCM, $\mathbf{H}_{t+m}^{(m)}$, is then obtained using the relationship in (5). Note that we do not need to iterate a forecasting procedure in order to obtain a monthly forecast, as is customary for the DCC and proposed in the CD-VARFIMA model of Chiriac and Voev (2010). This, however, comes at the price of having to re-estimate the parameters for different forecast horizons.

It is a nice feature of modelling elements of the Cholesky-Decomposition that any variable X_t deemed important in forecasting the elements in $P_{t+m}^{(m)}$ can easily be incorporated into the MIDAS

[4] This restriction can easily be relaxed.

model in (6):

$$P_{i,t+m}^{(m)} = \beta_{i0} + \beta_{i1} B\left(k,1,\theta_i\right) P_{i,t-k+1} + \beta_{ix} B\left(k,1,\theta_{ix}\right) X_{t-k+1} + v_t. \tag{9}$$

As the only requirement for any $\widehat{P}_{i,t+m}^{(m)}$ from such a model is that it ought to be real, no restrictions on any of the parameters in (6) are required. Indeed one may consider the inclusion of $P_{j,t}$, where $j \neq i$, into the forecasting model. This would turn the forecasting strategy into a CD-MIDAS-VAR model, which would seem a natuaral extension from the CD-VARFIMA of Chiriac and Voev (2010) or the discussion in Andersen *et al.* (2006).

In the end whether these terms, or any other explanatory variable, will significantly contribute to the forecast performance of the model is an empirical issue. But beyond this there is a potential econometric issue. NLS estimation of the parameters in (9) will be compromised if $\beta_{ix} = 0$ as, in this case, θ_{ix} will be unidentified. This will be particularly problematic if one was to consider a large number of $P_{j,t}$ as additional explanatory variables in which case it is increasingly likely that one would have to deal with this problem. In this case one would have to eliminate insignificant explanatory variables in a step prior to the NLS estimation[5].

We experimented with such a procedure, potentially allowing for several $P_{j,t}$, $j \neq i$, to be included into the MIDAS forecast model. However, no improvements of the forecasts performance were achieved by allowing extra $P_{j,t}$s. For this reason this approach will not be pursued further.

3.5 Multiple Stock Orderings

If we have n stocks then there are $n!$ possible permutations of these. Due to the nature of the calculation of Cholesky decompositions, each ordering of stocks, and its associated VCM, will result in a different Cholesky decomposition with no linear relationship to that obtained from an alternate ordering. As a result the procedure above yields different forecasts of the covariance matrix for each ordering.

In the light of the previous research on combining forecasts, we propose to use the available forecasts to effectively generate one superior VCM forecast by averaging across forecasts obtained from different orderings[6]. Beginning with Bates and Granger (1969), who show that combining two forecasts may outperform either of the constituent predictions, there is an established literature on the benefits of combining forecasts. The intuitive rationale for such gains in performance, proposed by Newbold and Harvey (2004), is that by averaging the forecaster is reducing the risk of relying on

[5] One procedure achieving this would be to fix all beta weight parameters θ_{ix} at a reasonable value and then eliminating terms according to some "pseudo-significance" criterion for the β_{ix}.

[6] As each individual forecast is positive semi-definite, the average will be as well.

one model, in an analogous way to an investor decreasing their risk by diversifying their portfolio. For reviews of the literature around the subject we refer the reader to Clemen (1989) and Newbold and Harvey (2004), both of whom note that forecast accuracy can often be improved by simple averaging of multiple individual forecasts. Newbold and Harvey (2004) also note that the literature shows several examples in which gains are made even when the individual forecasts come from a similar source, which seems applicable to the CD-MIDAS model with different stock orderings.

We therefore use model averaging to evaluate the impact of changing the orderings of stocks in the CD-MIDAS model. We will examine whether forecast accuracy changes when averaging forecasts obtained under different orderings and evaluate whether these predictions are more or less accurate than the values obtained when using a single ordering of the stocks. The fact that our model proposes different, yet equally valid, forecasts also illustrates that this model should merely be understood as a forecast tool and not as an estimated representation of an underlying structural relationship.

4 Forecast Evaluation

Below we compare several sets of 22-day VCM forecasts from different models: DCC, CD-MIDAS, CD-VARFIMA (allowing for models with multiple orderings for both CD models), Riskmetrics[7] and a rolling average of the realized variance-covariance matrices. The method used to determine which provides the most accurate predictions of portfolio volatility is the model confidence set (MCS) approach of Hansen, Lunde and Nason (2004).

The MCS takes a set of models and obtains a reduced group that contains the most accurate forecasts, with a given confidence level. The statistical process by which MCS results are obtained relies on evaluating the forecasting performance of each model relative to ex-post observations of the variable of interest via a loss function. The models that remain in the MCS at the end of the process are judged to have equal predictive power.

We begin the process of forming the MCS with a set of forecasting models Γ_0. The first stage of the process tests the null hypothesis that all of the these models have equal predictive accuracy (EPA) when their performance is measured against a set of ex-post observations. If $h_{kt}^{(m)}$ is the k^{th} forecast of the (scalar) variance over time $t-m+1$ to t and $\sigma_t^{(m)}$ is the observed value of the variance (or a consistent estimate) for the same period then the value of a loss function based on comparison of these is denoted $L(h_{kt}^{(m)}, \sigma_t^{(m)})$. The evaluation of the EPA hypothesis is based on

[7] An exponentially weighted moving average approach introduced in J.P. Morgan (1996).

loss differentials between forecasting models k and j at time t, $d_{kj,t}$,

$$d_{kj,t} = L(h_{kt}^{(m)}, \sigma_t^{(m)}) - L(h_{jt}^{(m)}, \sigma_t^{(m)}). \tag{10}$$

If all of the forecasts are equally accurate then the loss differentials between all pairs of forecasts should not be significantly different from zero. The null hypothesis of EPA is then

$$H_0 : E\left(d_{kj,t}\right) = 0 \ \forall k > j \in \Gamma \tag{11}$$

We test (11) using the semi-quadratic test statistic described in Hansen and Lunde (2007). If the null hypothesis is rejected at an $\alpha\%$ confidence level, we remove the worst performing model, in terms of the loss functions and begin the process again with the reduced set of forecasts Γ_1. This process is iterated until the test of equal predictive accuracy cannot be rejected, or a single model remains. The model(s) which survive form the $\alpha\%$ confidence MCS.

The literature on MCS typically assumes scalar forecasts and hence scalar loss functions are straightforward. As the forecasts considered here are for VCMs, we need to transform them to scalar loss functions. Two general approaches are considered here. *First* we form an equally weighted portfolio of all stocks considered. Two loss functions commonly employed in MCS evaluations, the mean squared error (MSE) and mean absolute deviation (MAD) can then be employed to the scalar portfolio variance. Let $h_{kt}^{(m)}$ be the variance forecast of the equally weighted portfolio, which is a function of the elements in the VCM forecast from model k, $\mathbf{H}_{kt}^{(m)}$. This will be compared to the realised portfolio variance, $\sigma_t^{(m)}$, being a linear combination of the elements of the actual VCM $\mathbf{\Sigma}_t^{(m)}$[8]. The loss functions are then applied to evaluate the accuracy of the portfolio variance forecasts:

$$\text{Mean Squared Error } L(h_{kt}^{(m)}, \sigma_t^{(m)}) = (h_{kt}^{(m)} - \sigma_t^{(m)})^2 \tag{12}$$

$$\text{Mean Absolute Deviation } L(h_{kt}^{(m)}, \sigma_t^{(m)}) = |h_{kt}^{(m)} - \sigma_t^{(m)}| \tag{13}$$

The *second* approach to constructing a loss function from a VCM forecast, $\mathbf{H}_{kt}^{(m)}$, and the actual VCM, $\mathbf{\Sigma}_t^{(m)}$, is to apply the multivariate QLIKE loss function

$$\text{MV QLIKE } L(\mathbf{H}_{kt}^{(m)}, \mathbf{\Sigma}_t^{(m)}) = tr(\mathbf{H}_{kt}^{(m)-1}\mathbf{\Sigma}_t^{(m)}) - \log\left|\mathbf{H}_{kt}^{(m)-1}\mathbf{\Sigma}_t^{(m)}\right| - n \tag{14}$$

[8] In the simulation experiments reported below we have access to the true VCMs and so can use $\Sigma_t^{(m)}$, where we discuss real stocks this is a latent matrix and we replace it with the realized VCM, $\mathbf{V}_t^{(m)}$, which is a consistent estimator of $\Sigma_t^{(m)}$.

which Patton and Sheppard (2009) and Laurent, Rombouts and Violante (2009) show to be a robust loss function for evaluating covariance matrices[9]. Becker, Clements, Hurn and Doolan (2009) find in simulations that QLIKE is most likely to select the forecast model which shares the specification of the data generating process with fewer remaining forecasting models in the MCS. Hence it is a more discerning loss function, and we therefore employ it in our simulations (see also Laurent, Rombouts and Violante, 2010).

By using the results from these three loss functions we aim to form a more robust picture of the performance of the CD-MIDAS compared to other forecasting methods[10].

5 Simulation Evidence

In this section we investigate the forecasting power of the CD-MIDAS through Monte Carlo simulations for dimensions of 3 and 20 stocks. The CD-MIDAS forecasts will be compared to forecasts from the DCC, CD-VARFIMA, Riskmetrics and the simple, but popular, rolling average. The results show that the CD-MIDAS model holds significant promise.

5.1 Data Generating Process

We first provide a brief overview of the method used to generate data for the simulation study; a more detailed description, including parameter values, can be found in Appendix B. The simulation assumes that the variance of returns for each of the n stocks follows a GARCH(1,1). The conditional correlation matrix, governing the strength of the relationships between the stock returns, is assumed to be generated from a scalar DCC(1,1) model as specified in Engle and Sheppard (2001). Given starting values we use the DCC and GARCH equations to determine the path of the variances and the correlation matrix at a daily level. The GARCH/DCC approach thus provides us with a variance-covariance matrix for each day.

In order to generate intraday data we use the daily variance-covariance matrix generated by the DCC. Using this, $\mathbf{\Sigma}_t$, and denoting the number of required intraday periods by Q, we obtain intraday returns for day t by taking Q draws from a $N(0_n, I_n)$[11] distribution and premultiplying the resulting vectors by the lower diagonal Cholesky decomposition of $\frac{1}{Q}\mathbf{\Sigma}_t$. The result is Q vectors of simulated returns, $\mathbf{r}_{qt}$ where

[9] Consistent, here, refers to a loss function that identifies the best forecast model even if the latent $\Sigma_t^{(m)}$ is replace with the realized VCM, $\mathbf{V}_t^{(m)}$.

[10] It can be shown by simulations that for a given dimension of VCM the MVQLIKE puts relatively more emphasis on the fit of the variances. As the dimension of the VCM increases the covariances become more important, but that increase in importance is somewhat slower for the MVQLIKE than it is for the MSE (evaluated on an equally weighted portfolio). These results are available on request.

[11] 0_n represents an $n \times 1$ vector of zeros and I_n is an $n \times n$ identity matrix.

$$
\begin{aligned}
\mathbf{r}_{qt} &= chol\left(\frac{1}{Q}\boldsymbol{\Sigma}_t\right) e_{qt} \\
e_{qt} &\sim N(0_n, I_n)
\end{aligned}
$$

and, as the vectors are independent within the day, the sum of the intraday return vectors, $\mathbf{r}_t = \sum_{q=1}^{Q} \mathbf{r}_{qt}$, has a realized VCM equal to $\mathbf{H}_t$. Hence we have a set of intraday returns with a DCC correlation structure at the daily level, whose component volatilities conform to a GARCH process. This process is conducted for $t = 1, ..., T$ so that we obtain the following:

1. Actual VCMs, $\boldsymbol{\Sigma}_t$ $t = 1, ..., T$, for use in forecast evaluation and calculation of a rolling average forecast.

2. Intraday returns data with 25 return periods per trading day, $\mathbf{r}_{qt}$ $q = 1, ..., 25$ and $t = 1, ..., T$. This is used in the calculation of realized covariance matrices which are inputs in the CD-MIDAS model as explained above.

3. Daily returns data, $\mathbf{r}_t$ $t = 1, ..., T$, for use in estimation and forecasting in the DCC model.

Data generation requires the values of the GARCH and DCC parameters, in order to ensure these are realistic we calibrate the values by estimating a DCC model for daily observations of the highly liquid Coca-Cola, American Express and Disney stocks, over the period 3/4/97-31/08/06. In the first of our simulations we use these estimates and allow the DCC model of the DGP to remain unchanged throughout. In the second simulation we consider an environment in which the DGP parameters change over time. We allow all of the GARCH and DCC parameters to change every 1,000th trading day in order to introduce a structural break in the data[12]. At these breaks we vary the GARCH and DCC parameters by small amounts to reflect typical values for these models, with larger changes made in long term correlations between the stocks. This is consistent with the nature of realized correlations over twenty-two day periods between the three stocks used for calibration, as these range between -0.11 and 0.91 and all of the correlations we use are in this range, the actual values used can be found in Table 6 in Appendix B. In this context the constant parameter DCC forecasts come from a misspecified model.

[12] Given that the estimation period was set to 1,000 observations, this break placement implies a varying placement of the break period in the estimation period. Further, on a small number of occasions the break is placed in the forecasting period.

5.2 Forecasting Models

In order to gauge the value of averaging forecasts across different permutations of stocks we include versions of the CD-MIDAS model based on one, two or three different orderings of the three stocks[13]. These three models are represented by *CDM1*, *CDM2* and *CDM3* respectively in the tables below.

In addition to the DCC and CD-MIDAS model we include a simple rolling average forecast (RA) to ensure that the two model approaches are able to significantly improve on a simple non-modelled approach. Rolling averages are computed from the actual VCMs for the last twenty non-overlapping 22 day periods prior to the start of the forecast period, hence in total the average is taken over 440 trading days. We also include the Riskmetrics model (*RISKM*) introduced in J.P. Morgan (1996), this is an exponentially weighted moving average of the cross products of monthly return vectors and is included here due to its popularity in the risk management industry.

The CD-VARFIMA model introduced in Chiriac and Voev (2010)[14] which uses a fractionally integrated VAR model to forecast the elements of the Cholesky decomposition is also included in our study. We follow the recomendations of Chiriac and Voev (2010) in including a single MA and AR term in the modelling of the behaviour of the decomposition and we restrict the fractional integration, MA and AR paramaters to be the same for all elements in the decomposition. In order to allow a fair comparison with our model we also obtain forecasts from this model for up to three orderings of the elements of the original VCMs and average the resulting forecasts. The results are labelled *CDV1*, *CDV2* and *CDV3* in the results presented below.

We simulate daily data series of length 3,200. The first 1,000 observations form the first estimation sample (for *CDM* and *CDV* models) used to forecast the VCM for days 1,001 to 1,022. The estimation sample is then moved forward 22 days to forecast the VCM over days 1,023 to 1,044 and so forth producing 100 non-overlapping forecasts. We then perform an MCS analysis on the set of 100 forecasts. For each DGP we obtain 1,000 such replications of the data, resulting in 1,000 model confidence sets. The results will report characteristics of these 1,000 model confidence sets we obtain for each DGP.

5.3 Results

We focus the analysis of our results on the number of times that a forecasting method is included in the 1,000 different MCS we obtain for each DGP. The tables below report the percentage of model

[13] The first ordering of the stocks is (1,2,3) in which stock 1's variance is the first diagonal element in the covariance matrix, stock 2's variance is the second diagonal element and stock 3 in the third diagonal position. The other orderings are (3,2,1) and (2,3,1).

[14] For our VARFIMA estimations we used code gratefully provided by Roxanna Chiriac, translated into OX.

confidence sets that contain a given model, and how often a specific model is the only remaining member of the set. Hence if we report that the *CDM1* model is included 88.4% of the time it means it was included in the model confidence set in 884 of the 1,000 replications.

We first consider the case where the data generating process is the constant coefficient scalar *DCC* of Engle and Sheppard (2001). The top panel of Table 1 reports the proportion of times that each of the considered models was included in the MCS for this DGP. The results are reported for two loss functions based on the total portfolio variance of an equally weighted portfolio (MSE, MAD) and one based on an elementwise analysis of the variance covariance matrix (MVQLIKE). Under all three loss functions the *CDM2* and *CDM3* versions of the CD-MIDAS model outperfrom the correctly specified DCC model, however under both MAD and MSE the difference in the inclusion rates is small and the *DCC* is included in over 90% of the model confidence sets. Under the MVQLIKE loss function the difference between the performance of the CD-MIDAS models and the *DCC* is much bigger. The difference in inclusion rates between *CDM2* and *DCC* in is 25.8% in this case, implying that under this robust loss function even when the *DCC* is correctly specified it is outperformed by the CD-MIDAS model.

It is also apparent that increasing the orderings of stocks in the CD-MIDAS models impacts positively on forecast accuracy. Under all three loss functions the performance of *CDM3* and *CDM2* are broadly similar, however, under all loss functions both perform significantly better than the single ordering *CDM1* version. This implies that although increasing the number of orderings does generally improve forecast performance, the vast majority of improvements may be achieved with relatively few orderings (here 2 orderings appear to suffice to achieve the majority of the gain).

Generally the performance of the CD-VARFIMA models is not encouraging. Although they outperform the rolling average and Riskmetrics forecasts, overall their forecasting performance is inferior to the CD-MIDAS and the DCC models in the context of a DCC GDP[15].

The bottom section of Table 1 reports the percentage of occasions on which one of the models was the only element remaining in the model confidence set. No model appears to dominate all others more than 2.6% of the time, in accordance with the results in the upper section of Table 1 in which *CDM2* and *CDM3* are included together in the vast majority of the cases. Under the MVQLIKE loss function 19.5% of the confidence sets are made up of only the *CDM2* and *CDM3* models which reinforces that these were the two best performing forecastors.

[15] It should be noted here that Chiriac and Voev (2010) find the CV-VARFIMA to outperform the DCC in the context of a 6 stock portfolio using real data. In that situation it is unlikly that the DCC is a good representation of the unobserved DGP and hence the results presented here are complementary to those presented in Chiriac and Voev. They do not present any simulation results.

	Inclusion In MCS								
	CDM1	CDM2	CDM3	DCC	RA	CDV1	CDV2	CDV3	RISKM
MAD	77.4%	98.7%	97.7%	92.2%	8.4%	3.1%	4.6%	11%	5.8%
MSE	87.3%	97.7%	97.7%	93.7%	7%	8.8%	7.9%	17.9%	13.3%
MVQLIKE	54%	98.4%	96.1%	72.6%	0.3%	2%	0.1%	0.3%	0.3%

	MCS consisting of only one model								
	CDM1	CDM2	CDM3	DCC	RA	CDV1	CDV2	CDV3	RISKM
MAD	0.2%	0.9%	0.2%	0.6%	0%	0%	0%	0%	0%
MSE	0.2%	0.4%	0.1%	1%	0%	0%	0%	0%	0%
MVQLIKE	0.2%	2.6%	0.6%	0.5%	0%	0%	0%	0%	0%

Table 1: Forecast evaluation for DCC DGP. The top panel of this table reports the percentage of times that models are included in the MCS, the bottom panel reports the percentage of times that a given model is included in the MCS on its own. The results are based on data procured from a DCC DGP process

Given that the *DCC* model is specified to match the DGP the fact that our CD-MIDAS model outperforms the *DCC* model seems counterintuative. We believe that the problem with the *DCC* forecasts is caused by estimation error. In order to provide more evidence for this conjecture we report results in which we replace the DCC model with *DCCact*, a model in which we use the true paramater values to forecast the VCM based on returns data obtained from our simulation process. The *CDV* and *CDM* models still require parameter estimation as above. The results are reported in Table 2. The results show that, when the true paramaters are employed, the DCC is the best performing model, this is demonstrated most emphatically under the MVQLIKE loss function, when in 96% of cases the MCS consists of only the DCC model. These results support our hypothesis that the significant deterioration in performance of the estimated DCC is due to estimation uncertainty.

We now consider results from the case in which the DCC does not share the specification of the DGP, a situation believed to be more in keeping with a real world environment. Table 3 reports the proportion of occasions on which models were included in the MCS when the data generating process was a DCC process with structural breaks every $1{,}000^{th}$ trading day. This allows us to compare the performance of the models when none of the estimated models accurately identifies the DGP of stock returns, a sure certainty in practice. Under all loss functions the two best performing models are the *CDM2* and *CDM3* models, with only a marginal difference in their performance. Under all loss functions the *CDM1* model is the third best performing model. As in the no breaks case the rolling average, CD-VARFIMA and Riskmetrics models are inferior to the DCC and CD-MIDAS models. As the DGP is still based on a DCC process (albeit with changing parameters) it is not surprising that the DCC remains superior to some other forecasting models that do not make use of any information on the structure of the DGP. We see that under all loss functions the CD-MIDAS model improves its performance when averaged across more than one ordering of the stocks, also consistent with Table 1 we see that in this case the biggest increase in performance is associated with the addition of the first alternative ordering of the stocks.

The bottom section of Table 3 reports the percentage of times that a particular model considered is the only remaining model. The results reveal that no single model makes up the MCS on its own more than 2.6% of the time. These results seem insignificant and are concurrent with the result that more than one CD-MIDAS model is commonly included in the MCS. An additional column, headed CDMIDAS is included in the bottom section of Table 3 reporting the proportion of times the MCS was made up only of a combination of CD-MIDAS models[16]. We see that under all three

[16] In these cases the MCS can be made up of one, two or three of the CD-MIDAS models included in the investigation.

	Inclusion In MCS								
	CDM1	CDM2	CDM3	DCCact	RA	CDV1	CDV2	CDV3	RISKM
MAD	73.3%	83.7%	83%	100%	7.7%	3.4%	4%	11%	5.8%
MSE	79.2%	86.7%	87.1%	99.8%	5.8%	6.1%	7%	14.1%	10.6%
MVQLIKE	3.9%	4%	4%	100%	0.2%	0.2%	0%	0.1%	0%

	MCS consisting of only one model								
	CDM1	CDM2	CDM3	DCCact	RA	CDV1	CDV2	CDV3	RISKM
MAD	0%	0%	0%	16%	0%	0%	0%	0%	0%
MSE	0%	0%	0%	10.7%	0%	0%	0%	0%	0%
MVQLIKE	0%	0%	0%	96%	0%	0%	0%	0%	0%

Table 2: Forecast evaluation for DCC DGP (including the DGP as one forecasting model). Percentage of times models are included in the MCS are reported in the top panel. The bottom panel reports the percentage of times that each model was the only member of the MCS. These results are based on a DCC DGP process

	Inclusion in MCS									
	CDM1	CDM2	CDM3	DCC	RA	CDV1	CDV2	CDV3	RISKM	
MAD	87.1%	96.8%	97.1%	59.8%	15.4%	0.1%	0%	0%	1%	
MSE	92.8%	95.4%	96.6%	68.3%	6.9%	0%	0.1%	0%	3.3%	
MVQLIKE	82.4%	97.6%	97.7%	21.9%	0.5%	0%	0%	0%	0%	
	MCS consisting of only one model*									
	CDM1	CDM2	CDM3	DCC	RA	CDV1	CDV2	CDV3	RISKM	CDMIDAS
MAD	1.2%	1.3%	1.7%	0.1%	0%	0%	0%	0%	0%	38.5%
MSE	2.6%	0.5%	0.8%	0.1%	0%	0%	0%	0%	0%	31%
MVQLIKE	0.8%	1.4%	1.5%	0%	0%	0%	0%	0%	0%	77.9%

Table 3: Forecast evaluation for DCC Break DGP. The top section of this table reports the percentage of times a model is included in the MCS. The second panel reports the percentage of occasions on which a model is included in the MCS without any of the other models. * The column CDmidas reports the percentage of MCS analysis in which the reuslt was made up only of some combination of CDM1, CDM2 and CDM3. These MCS results are based on simulations of a DGP process of a DCC model with breaks every 1,000th tradingday

loss functions a substantial amount of model confidence sets are made up only from the CD-MIDAS models, especially in the MVQLIKE case where 77.9% of the sets are composed exclusively of this type of model, this further underlines the usefulness of the forecasts from this approach.

Overall the CD-MIDAS outperforms all other models considered regardless of the data generating process, a positive reflection on the value of the proposed forecasting tool. We have also seen evidence that the practice of averaging over several VCM orderings in the CD-MIDAS model can significantly improve forecasts, albeit with diminishing returns to the number of orderings used.

6 Large Scale Portfolios

In multivariate modelling of covariance matrices there are two problems that require attention. The first is the problem of ensuring positive definite VCM forecasts, solved here by the use of the Cholesky decomposition as proposed by Chiriac and Voev (2010), and secondly there is the problem of parameter estimation for portfolios of large dimensions. The analysis above considered the case where n, the dimension of the VCM, was equal to three. However, traditionally it has been thought that around 20 stocks are necessary for the full effects of diversification to be obtained (see Bloomfield, Leftwich and Long, 1977). The purpose of this section is to show that the CD-MIDAS model can be adapted in a straightforward manner for use in forecasting the VCM in settings with higher dimensions; here we set $n = 20$ as a basis for examining large scale matrices.

With 20 assets we have $\tilde{n} = 210$ and we estimate the general CD-MIDAS specification using the own lag specification of (6). It is worth noting that although we have increased the number of stocks we are not required to alter our estimation procedure, other than increasing the number of univariate MIDAS relationships estimated.

In order to establish a benchmark we initially present a Monte Carlo simulation for $n = 20$. The data generating process is the DCC with no structural breaks and we use the same method to obtain the data as described previously (see Appendix B for details and Table 7 for parameters). We perform 100 replications, for each of these we obtain 100 forecasts of non-overlapping 22 day periods and, as before, use the MCS to compare the results from several models: DCC, CD-MIDAS and CD-VARFIMA with 1,2 & 3 orderings, Riskmetrics and a rolling average of actual 20×20 variance-covariance matrices. Estimating a DCC for 20 stocks is realistically only feasible using the composite likelihood approach of Engle *et al* (2008). The MCS results for this analysis using MSE and MAD loss functions for an equally weighted portfolio, and the MVQLIKE loss functions are reported in Table 4.

The results reveal that our CD-MIDAS model performs worse than than the DCC model when

	Inclusion in MCS								
	CDM1	CDM2	CDM3	DCC	RA	CDV1	CDV2	CDV3	RISKM
MAD	73%	82%	83%	90%	9%	0%	0%	0%	1%
MSE	78%	81%	82%	96%	6%	0%	0%	0%	3%
MVQLIKE	0%	25%	97%	26%	0%	0%	0%	0%	0%

	MCS Consisting of one model								
	CDM1	CDM2	CDM3	DCC	RA	CDV1	CDV2	CDV3	RISKM
MAD	1%	1%	1%	11%	0%	0%	0%	0%	0%
MSE	1%	0%	0%	14%	0%	0%	0%	0%	0%
MVQLIKE	0%	1%	61%	2%	0%	0%	0%	0%	0%

Table 4: The top section of this table reports the percentage of times model forecasts were retained in the MCS. The bottom panel reports the percentage of occasions on which the model consisted of a single model. All results apply to the case of forecasting a 20x20 VCM from data generated by a scalar DCC process..

the two are compared on the basis of total portfolio volatility, which is a weighted average of the 210 unique VCM elements. Only on a few occasions does the MCS consists of a single forecast only, in the vast majority of cases it would contain the DCC and a CD-MIDAS forecast. The nature of the result changes when we employ the MVQLIKE loss function. In this case the CD-MIDAS model with three orderings is performing much better than the DCC model with a difference in inclusion rates of 71%. Consequently the CDM3 forms a single forecast MCS in 61% of cases. It is notable that as in the small stock case the CD-VARFIMA, rolling average and Riskmetrics models perform much worse than the DCC and CD-MIDAS models.

Interestingly, adding a third ordering now has a clear, positive impact when evaluating the CD-MIDAS forecasts with the MVQLIKE. In the $n = 3$ case the addition of a third ordering only offered negligable improvements. This allows the conjecture that the marginal benefit of additional orderings is increasing in n[17].

It is interesting to see the substantially different implications for the different type of loss functions. As mentioned earlier, the contribution of the variance elements to the overall measure is relatively larger in the MVQLIKE measure than in the MAD or MSE applied to an equally weighted portfolio. It is therefore plausible to obtain different results.

6.1 Applied Forecasting Experiment

In this section we apply the CD-MIDAS model to twenty real stocks and compare its performance to CD-VARFIMA, DCC, Riskmetrics and rolling forecast approaches using an MCS evaluation. The data used in the estimation and forecasting covers the dates 1/12/1997-31/8/2006, the full list of stocks included in the experiment can be found in Table 8 in Appendix C.

As in the Monte Carlo simulations above the estimation period for the CD-MIDAS and DCC models is 1,000 observations long while the rolling average forecast is taken over the last twenty non-overlapping 22 day periods. We again employ a CD-MIDAS model which uses only own lags as explanatory variables for individual elements of the decomposition. As previously we restrict the results to up to three orderings. We investigated up to five orderings but beyond three orderings no forecast improvements were available. This supports our earlier conjecture that the bulk of improvements is to be had from averaging over a very small number of orderings.

In this experiment we obtain forecasts for 49 non-overlapping 22 day periods, hence the forecasts cover a period of approximately four years in length. The MCS results using MVQLIKE, MSE and MAD, where the latter two are calculated for equally weighted portfolios, are presented in Table 5,

[17] The issue of the optimal number of orderings is worthy of additional investigation but goes beyond the scope of this paper. In the empirical application we offer some additional information on this issue.

MSE	$\overline{Loss}$	p_{MCS}	MAD	$\overline{Loss}$	p_{MCS}	MVQLIKE	$\overline{Loss}$	p_{MCS}
CDM3	2.88	1.00	CDM2	0.88	1.00	CDM3	3.67	1.00
CDM2	2.90	0.47	CDM3	0.89	0.50	CDM2	3.69	0.21
CDM1	2.97	0.16	CDM1	0.90	0.28	CDM1	3.77	0.06
DCC	3.50	0.16	CDV2	1.00	0.28	CDV2	4.82	0.06
CDV2	4.54	0.16	CDV1	1.01	0.28	CDV3	4.86	0.06
CDV1	4.56	0.16	CDV3	1.01	0.28	CDV1	5.57	0.06
CDV3	4.56	0.16	DCC	1.02	0.28	RA	5.77	0.04
RA	6.02	0.15	RA	1.56	0.24	DCC	6.25	0.02
RISKM	9.67	0.10	RISKM	2.79	0.16	RISKM	18.34	0.02

Table 5: MCS for real data. MCS results for forecasts of the variance-covariance matrix of 20 stocks. For each loss function the entries are ordered with the best forecast model listed first. $\overline{Loss}$ is the average loss. $p - MCS$ is the MCS p-value. CDMi (CDVi) represents the forecasts from the CD-MIDAS (CV-VARFIMA) model with i orderings. DCC represents the forecasts from the DCC model estimated with the component likelihood method. RA is the rolling average forecast and RISKM the Riskmetrics forecast.

these include both the mean values of the loss functions, $\overline{Loss}$, and the MCS p-values, p_{MCS}. The p-values indicate the confidence level at which a model would be removed from the MCS, hence at a 10% confidence level all models with a p-value in excess of 0.1 would be included in the MCS.

We find that in the cases of all three loss functions the CD-MIDAS model has the lowest loss function values and the highest associated p-values. Under the MAD and MSE loss function only the Riskmetrics model is excluded from the MCS at standard confidence levels, however as discussed in Becker *et. al* (2009) and Laurent *et. al* (2010) these loss functions exhibit less power relative to the MVQLIKE. Under the MVQLIKE the only models included in the MCS are models based on the CD-MIDAS approach, notably the version of the model based on a single ordering of stocks is removed from the MCS at a 10% significance level, the same level at which the forecasts based on the CD-VARFIMA model are rejected from the MCS. This is further evidence that the use of additional orderings can be used to obtain superior forecasts in the model.

In general this applied experiment further shows the ability of the CD-MIDAS to outperform the forecasts of the DCC and CD-VARFIMA models and that within the model increasing the number of permutations over which we average forecasts can increase the accuracy of forecasts of the VCM.

7 Conclusions

This paper presents a new model for forecasting the variance covariance matrix (VCM) of a stock portfolio. The model, referred to as Cholesky Decomposition-MIDAS (CD-MIDAS) is estimated using the properties of the Cholesky decomposition to ensure positive definiteness of forecasts. It also employs the MIDAS framework to provide a specification which allows realizations of daily covariance matrices to be used in forecasting monthly covariance matrices, in contrast to many existing multivariate models which ignore such information.

A Monte Carlo investigation shows that the model presented here is able to significantly improve on forecasts of return volatility, for an equally weighted portfolio, compared to the DCC, CD-VARFIMA and other more simple forecasting techniques. This is evidence that the Cholesky-MIDAS model is a potentially useful technique whenever forecasts of variance covariance matrices are used in a decision making process. This result is maintained even when the DCC's form is correctly specified but parameter estimation is required. It is important to understand that the CD-MIDAS forecasting tool (as well as the CD-VARFIMA model) cannot lay claim to represent any structural model of return (co)variances. It is merely a parsimoniously parameterised forecasting tool potentially able to capture important stylised facts in the the dynamics of variances and covariances. As such its usefulness can only be judged in a forecasting context. It is therefore encouraging that it proves to produce superior forecasts compared to a well understood forecasting model (the DCC), even when the latter is the correct representation of the DGP.

As different orderings of stocks will produce different Cholesky decompositions, the forecasts from the CD-MIDAS model are conditional on the particular ordering chosen. This opens the opportunity to produce multiple, equally valid forecasts for the same VCM. It is demonstrated here that the resulting possibility of producing forecast averages can significantly add to the quality of forecasts. It appears as if a small number of alternative orderings suffice to achieve the majority of available improvements. A deeper investigation into this issue and in particular in whether optimal ordeings can be identified ex-ante is left for future research.

Another area for further investigation is the possible augmentation of the MIDAS specification with additional weakly exogenous information, a property shared with the model of Chiriac and Voev (2010) who were the first to devise models for the elements of the Cholesky decomposition. Hence one could use information such as realized volatilities, information on news releases, macro-economic variables, implied volatility data from VIX and anything else that might be believed to influence the VCM. This is not commonly the case in other multivariate models and we note that the capacity of this model to incorporate additional variables could be a significant benefit.

8 Bibliography

References

Andersen, T. G., Bollerselv, T., Christoffersen, P. F. and Diebold, F. X. [2006], Volatility and correlation forecasting, in G. Elliott, C. Granger and A. Timmermann, eds, *'Handbook of Economic Forecasting'*, North Holland Press, Amsterdam.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003), "Modeling and forecasting realized volatility", *Econometrica,* vol. 71, no. 2, pp. 579-625.

Armesto, M.T., Hernández-Murillo, R., Owyang, M.T. & Piger, J.M. (2007), "Identifying asymmetry in the language of the beige book: a mixed data sampling approach" Federal Reserve bank of St. Lous Working Paper Series No. 2007-010B.

Barndorff-Nielsen, O. E. and N. Shephard (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics* 4, 1-30.

Bates, J.M. & Granger, C.W.J. (1969), "The combination of forecasts", *Operations Research Quarterly*, Vol. 20, pp465-74.

Becker, R., Clements, A., Hurn, S. & Doolan, M. (2009) "On the efficacy of techniques for evaluating multivariate volatility forecasts" NCER Working Paper Series, No. 41.

Billio, M., Caporin, M. & Gobbo, M. (2003), "Block dynamic conditional correlation multivariate GARCH models." GRETA working paper n.03.03.

Bloomfield, T., Leftwich, R. & Long, J.B. (1977), "Portfolio strategies and performance", *Journal of Financial Economics,* Vol. 5 pp201-218.

Bollerslev, T. 1990, "Modelling the coherence in short-run nominal exchange rates: A multivariate generalised ARCH model", *The Review of Economics and Statistics*, vol. 72, no. 3, pp. 498-505.

Cappiello, L., Engle, R. F. & Sheppard, K. (2006), "Asymmetric dynamics in the correlations of global equity and bond returns", *Journal of Financial Econometrics*, vol. 4, no. 4, pp. 537-572.

Chiriac, R. & Voev, V. (2010), "Modelling and forecasting multivariate realised volatility.", *Journal of Applied Econometrics*, to appear

Clemen, R.T. (1989), "Combining forecasts: A review and annotated bibliography", *International Journal of Forecasting*, Vol. 5, pp559-83

Clements, M.P. & Galvão, A. B. (2008), "Macroeconomic forecasting with mixed-frequency data: forecasting output growth in the United States", *Journal of Business & Economic Statistics* Vol. 35 No.4 pp546-554

Colacito, R., Engle, R.F. & Ghysels, E. (2007), "A component model for dynamic correlations", Unpublished

Engle R.F. (2008) “High dimensional dynamic correlations” in: Castle J.L. and Shephard N. (Eds.), *The Methodology and Practise of Econometrics: Papers in Honour of David F Hendry.* Oxford University Press, Forthcoming.

Engle, R.F., Ghysels, E. & Sohn, B. (2008), "On the economic sources of stock market volatility", Unpublished

Engle, R. F. & Sheppard, K. (2001), "Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH", NBER Working Paper No. 8554.

Engle R.F. Shephard N. and Sheppard K. (2008) “Fitting vast dimensional time-varying covariance models”, unpublished mimeo, http://www.oxford-man.ox.ac.uk/~nshephard/.

Epps, T.W. (1979), "Comovements in stock prices in the very short run", *Journal of the American Statistical Associatin,* Vol. 74, pp291-298.

Ghysels, E., Santa-Clara, P. & Valkanov, R. (2004), "The MIDAS touch: Mixed data sampling regression models", CIRANO Working Papers 2004s-20, CIRANO.

Ghysels, E., Santa-Clara, P. & Valkanov, R. (2005), "There is a risk-return tradeoff after all", *Journal of Financial Economics*, Vol. 76, pp509-548.

Ghysels, E., Santa-Clara, P. & Valkanov, R. (2006), "Predicting volatility: getting the most out of return data sampled at different frequencies", *Journal of Econometrics*, Vol. 131, pp59-95.

Ghysels, E., Sinko, A. & Valkanov, R. (2008), "MIDAS regressions: further results and new directions", *Econometric Reviews*, Forthcoming

Ghysels, E. & Sinko, A. (2006), "Volatility prediction and microstructure noise", Work in progress

Hansen, P.R. & Lunde, A. (2005) "A realized variance for the whole day based on intermittent data", *Journal of Financial Econometrics*, Vol 3 (4), pp525-554.

Hansen, R.P. & Lunde A. (2006) “Realised Variance and Market Microstructure Noise” *Journal of Business and Economic Statistics*. Vol. 24, pp. 127-218.

Hansen, R.P. & Lunde, A. (2007) "MULCOM 1.00, Econometric toolkit for multiple comparisons" (Packaged with Mulcom package)

Hansen, P.R., Lunde, A. & Nason, J.M. (2004), 'Model confidence sets for forecasting models',Federal Reserve Bank of Atlanta Working Paper No. 2005-7 http:\\ssrn.com\\paper=522382

Hayashi, T. & Yoshida, N. (2005) "On covariance estimation of non-synchronously observed diffusion processes", Bernoulli, Vol 11 (2), pp359-379.

Koopman, S. J., Jungbacker, B., & Hol, E. 2005, "Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measures", *Journal of Empirical Finance*, vol. 12, pp. 445-475.

J.P. Morgan (1996) *Riskmetrics Technical Document 4th Edition*, J.P. Morgan, New York.

Laurent, S., Rombouts, J.V.K. & Violante, F. (2009) 'On loss functions and ranking forecasting performances of multivariate volatility models', CIRPEE Working Paper 09-48.

Laurent, S., Rombouts, J.V.K. & Violante, F. (2010) 'On the Forecasting Accuracy of Multivariate GARCH Models', CIRPEE Working Paper 10-21.

Newbold, P. & Harvey, D.J. (2004)"Forecasting Combination and Encompassing" in Clements, M.P. and Hendry, D.F. (Eds), *A Companion to Economic Forecasting*, pp 268-83, Blackwells, Oxford

Patton, A. & Sheppard, K. "Evaluating Volatility anfd Correlation Forecasts" in Anderson, T.G., Davis, J.P., White, H. and Mikosch, T. (eds) *Handbook of Financial Econometrics*, Springer Verlag.

Pelletier, D. (2006) "Regime switching for dynamic correlations", *Journal of Econometrics*, vol. 131, no. 1-2, pp. 445-473.

Poon S-H. and Granger C.W.J. (2003) “Forecasting volatility in financial markets: a review” *Journal of Economic Literature*, **41**, 478-539.

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

A Realised Covariance Matrix Calculation

In the paper realized variance-covariance matrices are central to the implementation of the Cholesky-MIDAS model. As mentioned in the text there are several ways in which the matrix may be calculated, however this appendix focuses on describing only the method utilised in this paper.

We utilise one of the methods proposed in Hansen and Lunde (2005), designed to capture the volatility across an entire 24-hour period rather than just during the trading segment of the day. It is chosen in part because it makes the simulation of returns across a 24-hour period more straightforward in the Monte Carlo simulation. This makes the calculation of monthly covariance matrices easier as they will simply be the sum of 22 single daily covariance matrices.

For a given set of stocks open and closing prices are recorded as well as prices at x minute intervals within the trading day. The vector of returns over the period when the market is closed is denoted as $\mathbf{r}_{co}$ while the intraday returns over x minute periods are denoted $\mathbf{r}_{q,t}$ where $q = 1, ...Q$ and Q is the number of x minute periods in the trading day. The realized variance-covariance matrix for the day, $\mathbf{V}_t$, is calculated as

$$\mathbf{V}_t = \mathbf{r}_{co,t}\mathbf{r}'_{co,t} + \sum_{q=1}^{Q} \mathbf{r}_{q,t}\mathbf{r}'_{q,t} \tag{A.1}$$

When $\mathbf{V}_t^{(m)}$, a variance-covariance matrix for an m day period ending at time t, is required we simply sum each of the constituent matrices as in (A.2) below:

$$\mathbf{V}_t^{(m)} = \sum_{i=t-m+1}^{t} \mathbf{V}_i \tag{A.2}$$

where $\mathbf{V}_i$ is the realized covariance matrix for the i^{th} 24 hour period.

In the simulation experiment in Section 5 we allow each of the 24 hour periods to contain 25 intraday trading periods. If we attempt to interpret this in terms of a real world trading environment this is equivalent to having one period representing over-night returns and 24 returns observed within the trading day.

The simulation experiment allows us to side-step several issues in the calculation of realized covariance matrices, we need not worry about jumps in the data and non-synchronous trading creating bias in our measure of the covariance matrix as they are not a feature of our data. Although we accept that this will affect real, data within our simulation we assume that we have been able to solve these problems and that the method borrowed from Hansen & Lunde (2005) allows us to consistently estimate the actual variance-covariance matrix for a given 24 hour period.

B Data Simulation

This appendix describes how the data used in the simulation experiment is generated. The key outputs of the process are; a) observed variance covariance matrices, b) simulated daily returns data and c) simulated intraday data which can be used to generate simulated realized variance covariance matrices.

We make several simplifying assumptions. Firstly we assume that trading occurs over a whole day so that a day can be split into Q periods of equal length. We also assume by construction that there are no jumps and further that trading activity occurs throughout the day.

The first step is to define the full set of equations which govern the dynamic behaviour of the variance covariance matrix (VCM). We use a GARCH (1,1) process to determine the volatility of each of the n stocks, as in (B.1)-(B.3). Hence each stock has 3 parameters which govern the development of its volatility process.

$$r_{it} = \eta_{it} \quad \eta_{it} \sim N(0, \sigma_{it}^2) \tag{B.1}$$

$$\sigma_{it}^2 = \alpha_{i0} + a_{i1}\eta_{it-1}^2 + \beta_i \sigma_{it-1}^2 \tag{B.2}$$

$$\eta_{it} = \sqrt{\sigma_{it}^2}\varepsilon_{it} \quad \varepsilon_{it} \sim IID(0,1) \tag{B.3}$$

We allow the movement of the variance-covariance matrix to be governed by a simple DCC process, as presented in (B.4)-(B.7).

$$\mathbf{\Sigma}_t = \mathbf{D}_t \mathbf{\Gamma}_t \mathbf{D}_t \tag{B.4}$$

$$\mathbf{\Gamma}_t = \mathbf{Q}_t^* \mathbf{Q}_t \mathbf{Q}_t^* \tag{B.5}$$

$$\mathbf{Q}_t = (1 - \gamma - \varphi)\bar{\mathbf{Q}} + \gamma(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') + \varphi \mathbf{Q}_{t-1} \tag{B.6}$$

$$\mathbf{Q}_t^* = I_3 \odot \mathbf{Q}_t^{1/2} \tag{B.7}$$

In effect this models the correlation matrix for the stocks which we then combine with the variances generated by our n GARCH models to yield the VCM. In (B.4) $\mathbf{D}_t$ is a diagonal matrix with the diagonal elements equal to the square root of the variances, σ_{it}^2, generated by the GARCH equation

(B.2), $\mathbf{\Gamma}_t$ is the correlation matrix of the stocks and it is this that is modelled by the DCC model. In (B.6) $\boldsymbol{\varepsilon}_t$ represents a $n \times 1$ vector of standardized residuals from the n univariate GARCH models. In order for the model to be complete in our simulation we need to specify the values of the DCC parameters γ and φ as well as the unique elements of the long run correlation matrix $\bar{\mathbf{Q}}$. The $n \times 1$ vector of stock returns is distributed as $\mathbf{r}_t \sim N(0_{n\times 1}, \mathbf{\Sigma}_t)$ where the dynamics of $\mathbf{\Sigma}_t$ is described by (B.1)-(B.7).

We now describe how to initialize and iterate the data in order to obtain the required simulation data. For time $t = 0$ we set $\sigma^2_{i0} = \frac{\alpha_{i0}}{1-a_{i1}-\beta_i}$ for each of the n stocks, that is we set the initial variance of each stock equal to its long run value, similarly we set $\mathbf{Q}_0 = \bar{\mathbf{Q}}$ and from (B.5) we can see that this translates to setting the initial correlation matrix equal to its long run value. From this we obtain Σ_0 from (B.4). Hence for the initial day of the simulation we obtain a $n \times 1$ vector of returns such that $\mathbf{r}_0 \sim (0, \mathbf{\Sigma}_0)$. In order to obtain realized VCM we need Q intraday observations of this data. We obtain these using the following steps.

1. Multiply Σ_0 by $\frac{1}{Q}$, which gives the matrix $\tilde{\Sigma}_0 = \frac{1}{Q}\Sigma_0$ which is the variance-covariance matrix for each of the intraday periods.

2. Obtain Q random vectors drawn from a normal distribution, each random vector, $\lambda_{q,0}$ ($q = 1, ..., Q$) is such that $\lambda_{q,0} \sim N(0_{n\times 1}, I_n)$.

3. Pre-multiply each $\lambda_{q,0}$ by the Cholesky decomposition of $\tilde{\Sigma}_0$, $\mathbf{C}_0$, to obtain the q^{th} vector of intraday returns. Each of the intraday return vectors $\mathbf{r}_{q,0} = \mathbf{C}_0\lambda_{q,0}$ and so $\mathbf{r}_{q,0} \sim N(0_{n\times 1}, \tilde{\Sigma}_0)$.

4. The daily return is then equal to the sum of the intraday periods, that is $\mathbf{r}_0 = \sum_{q=1}^{Q} \mathbf{r}_{q,0}$, and from the rules of linear combinations of independent vectors $\mathbf{r}_0 \sim N(0, \Sigma_0)$

We can now generate σ^2_{i1} for each stock using the daily returns r_{i0} where t_{i0} is the ith element of $\mathbf{r}_0$. In combination with the recursion in equation (B.6) and the relations in (B.7), (B.5) and (B.4) this delivers $\mathbf{\Sigma}_1$ which is then used to generate intraday returns for $t = 1$ as described in steps 1 to 4 above.

Data are simulated for $n = 3$ (Section 5.1) and for $n = 20$ (Section 6). The parameter values used in the case of the simulations for $n = 3$ are shown in Table 6. The parameter values used in the $n = 20$ case are provided in Table 7 below.

`DCC without break, n=3`							
`DCC with breaks (Regime 1), n=3`							
GARCH	$\alpha_0 \cdot 10^4$	α_1	β				
stock1	0.017635	0.07228	0.9177	$\bar{Q}$	1.000		
stock2	0.005927	0.045517	0.943804		0.365	1.000	
stock3	0.05444	0.09182	0.905986		0.434	0.295	1.000
					$\gamma = 0.01$		$\varphi = 0.98$
`DCC with breaks (Regime 2), n=3`							
GARCH	$\alpha_0 \cdot 10^4$	α_1	β				
stock1	0.002	0.05228	0.9377	$\bar{Q}$	1.000		
stock2	0.01927	0.075517	0.903804		0.050	1.000	
stock3	0.03444	0.03182	0.945986		0.650	0.400	1.000
					$\gamma = 0.05$		$\varphi = 0.94$
`DCC with breaks (Regime 3), n=3`							
GARCH	$\alpha_0 \cdot 10^4$	α_1	β				
stock1	0.0015	0.03228	0.9577	$\bar{Q}$	1.000		
stock2	0.01127	0.045517	0.933804		0.150	1.000	
stock3	0.0444	0.0218	0.925986		0.650	0.250	1.000
					$\gamma = 0.03$		$\varphi = 0.93$
`DCC with breaks (Regime 4), n=3`							
GARCH	$\alpha_0 \cdot 10^4$	α_1	β				
stock1	0.004	0.06228	0.9277	$\bar{Q}$	1.000		
stock2	0.023927	0.035517	0.963804		0.250	1.000	
stock3	0.047	0.01182	0.975986		0.500	0.460	1.000
					$\gamma = 0.02$		$\varphi = 0.97$

Table 6: DCC parameter values (n=3). Paramaters used for the simulation of GARCH and DCC data in the simulation described in appendix B

`DCC without break, n=20`			
Correlation Dynamics	γ	ψ	
	0.01	0.98	
Unconditional Correlations	Min	Max	Avg
	0.03	0.58	0.31
Univariate GARCH Parameters	Min	Max	Avg
α	0.0005	0.0948	0.0459
β	0.8176	0.9930	0.9054
$\alpha + \beta$	0.9013	0.9999	0.9513

Table 7: DCC parameter values (n=20). Paramaters used for the simulation of GARCH and DCC data in the simulation described in appendix B

	Ticker symbol	Company Name
1	AA	Alcoa Inc
2	AXP	American Express Inc
3	BA	Boeing Co.
4	BAC	Bank of America Corp.
5	BMY	Bristol Myers Squibb Co.
6	CL	Colgate Palmolive
7	DD	El DuPont de Nemours co.
8	DIS	Walt Disney Corp.
9	GD	General Dynamics Corp
10	GE	General Electric Co.
11	IBM	IBM
12	JNJ	Johnson & Johnson
13	JPM	JP Morgan Chase Co.
14	KO	Coca Cola Corp.
15	MCD	Mcdonald's Corp
16	MER	Merril Lynch Co. Inc
17	MMM	3M co.
18	PEP	Pepsico Inc.
19	PFE	Pfizer Inc
20	TYC	Tyco International ltd.

Table 8: List of Stocks. Stocks included in the forecasting experiment in section 6.1.

C Stocks Used In Forecasting Experiment

Table 8 provides a full list of the stocks used in the analysis in section 6.1.