Universität St.Gallen

Forecasting correlations during the late-2000s financial crisis: short-run component, long-run component, and structural breaks

Francesco Audrino

April 2011 Discussion Paper no. 2011-12

School of Economics and Political Science, Department of Economics

University of St. Gallen

Editor:	Martina Flockerzi University of St. Gallen School of Economics and Political Science Department of Economics Varnbüelstrasse 19 CH-9000 St. Gallen Phone +41 71 224 23 25 Fax +41 71 224 31 35 Email seps@unisg.ch
Publisher:	School of Economics and Political Science Department of Economics University of St. Gallen Varnbüelstrasse 19 CH-9000 St. Gallen Phone +41 71 224 23 25
Electronic Publication:	Fax +41 71 224 31 35 http://www.seps.unisg.ch

Forecasting correlations during the late-2000s financial crisis: short-run component, long-run component, and structural breaks

Francesco Audrino

Author's address:	Prof. Dr. Francesco Audrino Institute of Mathematics and Statistics Bodanstrasse 6 9000 St. Gallen Tel. +41 71 224 2431 Fax +41 71 224 2894 Email francesco.audrino@unisg.ch Website www. mathstat.unisg.ch

Abstract

We empirically investigate the predictive power of the various components affecting correlations that have been recently introduced in the literature. We focus on models allowing for a flexible specification of the short-run component of correlations as well as the long-run component. Moreover, we also allow the correlation dynamics to be subjected to regime-shift caused by threshold-based structural breaks of a different nature. Our results indicate that in some cases there may be a superimposition of the long- and short-term movements in correlations. Therefore, care is called for in interpretations when estimating the two components. Testing the forecasting accuracy of correlations during the late-2000s financial crisis yields mixed results. In general component models allowing for a richer correlation specification possess a (marginally) increased predictive accuracy. Economically speaking, no relevant gains are found by allowing for more flexibility in the correlation dynamics.

Keywords

Correlation forecasting; Component models; Threshold regime-switching models; Mixed data sampling; Performance evaluation.

JEL Classification

C32; C52; C53.

1 Introduction

Nowadays there is a widespread empirical evidence that correlations vary over time and across assets. Starting with the seminal dynamic conditional correlation (DCC) model proposed by Engle (2002) as a generalization of Bollerslev's (1990) constant conditional correlation (CCC) model, researchers have tried to take this time-varying behavior of the correlation dynamics into account. Most such models focused on the idea that there are different short- and long-run sources that affect correlations, which may be naturally interpreted according to certain economic principles. The approaches proposed in the literature mainly estimated only one such component (see, for example, the DCC model proposed by Engle (2002) and almost all its generalizations, and Ledoit et al. (2003) for the short-run component, and Pelletier (2006) for the long-run component). More recently some studies that place more emphasis on the estimation of both sources simultaneously have been proposed, including, among others, the models proposed by Colacito et al. (2010) and Audrino and Trojani (2011).

Following a recent stream in the literature on volatility, Colacito et al. (2010) apply the idea of a component model to the estimation of dynamic correlations. The model proposed in that study uses an Engle's DCC-type approach to capture the short-term correlation dynamics in connection with a long-run component. This component is extracted via the mixed data sampling (MIDAS) technique recently discussed by Ghysels et al. (2006) and Forsberg and Ghysels (2007), among others, and applied to the problem of volatility forecasting. In two applications, Colacito et al. (2010) found that extending the standard autoregressive short-lived DCC dynamics to take into account a (possibly) economically interpretable long-term component significantly improves the fit of the model and allows one to devise more profitable portfolio allocation strategies.

Audrino and Trojani (2011) propose a unified, generalized framework for dynamic correlations in which the choice of the type of short-run dynamics (i.e. constant, CCC, or dynamically changing, DCC) is completely data driven. In their model, the long-run component is allowed to change smoothly over time via the idea of some rolling, equally weighted unconditional correlation averages in the spirit of the MIDAS approach. Moreover, the long-term correlation component is subjected to drastic structural breaks and regime shifts according to the multivariate dynamic behavior of some relevant (exogenous or endogenous) underlying variables. Fitting their model on three different data sets, Audrino and Trojani found that both components are not always relevant for improving the accuracy of the correlations estimates and predictions: in some cases, a simple CCC structure for the short-term behavior of correlations and a slowly moving long-run component subjected to jumps caused by structural breaks of a different nature yield superior performances over, for example, DCC-type models.

The main purpose of this empirical study is to contribute to the correlation literature by better understanding and disentangling the role of the short- and long-run components outlined above for correlation prediction in periods of market turbulence. Moreover, we allow one or both components to be subjected to regime shifts and we evaluate the additional forecasting power we may gain from this increased flexibility in modeling correlation dynamics.

To this end we run a horse race considering the most relevant models applied to two (updated) data sets already used in the past literature by Colacito et al. (2010) and Audrino and Trojani (2011): namely, Fama-French industry portfolios and the long-term bond and a nine-dimensional US stock example. Performance is evaluated both in statistical and economic terms and, at least for the US stock data set, comparisons are based on very accurate correlation measures computed using high-frequency data. In particular, we unify the settings considered in Colacito et al. (2010) and Audrino and Trojani (2011) and we focus more intensively on the differences in the forecasting power of the three effects that have been identified as main drivers of correlation dynamics: short-run and long-run components, and regime-switches.

For some applications, disentangling the long-run movements from the short-term behavior of correlations may be very difficult. In fact, with certain data sets, analysis can not yield any clear conclusion about which type of component has been estimated. This superimposition leads to enormous difficulties in interpreting the results on the one hand and does not allow clarification of the predictive power of the different effects on the other hand. One such example is shown in Figure 1. The data under investigation are the short- and long-run components of the daily correlations between two Fama-French industry portfolios, i.e. the energy and the hi-tech portfolios, and the 10 year bond. This data set is an updated version of the one investigated in Colacito et al. (2010).

FIGURE 1 ABOUT HERE.

As can be seen in the figure, the ex-post estimated correlations using a model focusing only on short-lived dynamics (DCC) and those from a model capturing the long-term correlation behavior (CCC-MIDAS) are almost indistinguishable. In fact, absolute differences between the correlation estimates are never larger than 8%. As this empirical study demonstrates, this fact will be reflected in a comparable statistical and economic performance of the models. Interpretation of such correlation estimates as short- or long-run dynamics must therefore be made with caution.

Clearly this may not be the case for other data sets, as shown in Figure 2. The data considered in this case are part of a nine-dimensional US stock example. As an illustration we report the correlation estimates between Nike and Microsoft. Once again this is an updated version of the data set investigated in Audrino and Trojani (2011).

FIGURE 2 ABOUT HERE.

Differences between the short-term DCC movements and those stemming from a CCC-MIDAS model focusing on long-run dynamics are particularly evident. In particular, the long-run component obtained by the CCC-MIDAS estimation show some trend movements during the time period under investigation whereas, as expected, the DCC short-run dynamics are less disperse and mean-reverting around a constant long-term correlation value of about 0.3. Such different behavior of the two correlation components allows for a clear distinction and (possibly) a natural interpretation of the two effects also using certain economic principles. This impression is even more evident when superimposing the correlation estimates obtained using CCC-MIDAS and DCC-MIDAS approaches. Results are shown in Figure 3.

FIGURE 3 ABOUT HERE.

The plot at the top of Figure 3 shows that the MIDAS long-run components estimated in the CCC and DCC settings match particularly well and are characterized by a general downward trend during the period under investigation. The bottom panel clearly shows that for this data the estimation of both short- and long-run components produce different results than when considering only one of these components alone. Both effects seem to be relevant to the production of accurate estimates and forecasts of the correlation dynamics. As will be demonstrated, our performance results bear out this first impression.

We perform two out-of-sample tests. In both cases the out-of-sample period corresponds to the most recent years available (2006-2009 in one case and 2007-2010 in the second case). Since the late-2000s financial crisis is included in those periods, our forecasting exercise may be interpreted as applying a sort of stress test of the different approaches under very difficult conditions. We evaluate the statistical performance of the correlation predictions according to tests for superior predictive ability (SPA) introduced by Hansen (2005), and we construct model confidence sets (MCS, see Hansen et al., 2003 and 2011) at different significance levels based on several goodness-of-fit measures such as out-of-sample likelihood, out-of-sample mean square errors (MSE), and out-of-sample mean absolute errors (MAE). Similarly to Colacito et al. (2010), we use the portfolio allocation problem first introduced in Engle and Colacito (2006) to evaluate the economic value of the different models for correlation dynamics.

In general, our correlation forecasting performance results are mixed. Investigating the role of short- and long-run components and structural breaks for predicting correlations among Fama-French industry portfolios and the long-term bond, we do not find any significant difference in statistical or economic terms among models that consider only one or both components. We interpret this result as a consequence of the superimposition of short- and long-run components when estimated using the different models, as discussed above. In contrast, allowing for structural breaks of a different nature yields to a marginal improvement of correlation prediction accuracy. In our second empirical example on US stock data, results are more in favor of component models where both the short- and long-run movements are explicitly taken into account, although the very simple DCC model is never significantly outperformed. Correctly predicting the short-term movements of correlations seems to be the crucial part of the job in this case. Allowing for structural breaks does not increase the forecasting power of the models. In testing differences in economic terms, we are unable to find any significant gain achievable by an investor using more sophisticated approaches over the very simple CCC model: in fact, 15 basis points is the maximal gain obtained overall. This result highlights how widely interpretations about the goodness of a model may vary depending on whether statistical or economic gains are considered.

The remainder of the paper is organized as follows. The competing models, data sets, and performance measures in our horse race are introduced in Section 2. Results are presented and discussed in Section 3. Section 4 summarizes and concludes.

2 The horse race

This section describes the main ingredients of our race in full detail, namely the alternative approaches considered, the data sets investigated, and the performance measures used to rank the predictive power of the models.

2.1 The competing models

Let us consider a d-dimensional multivariate stochastic process $(\mathbf{r}_t)_{t\in\mathbb{Z}}$ with values in $\mathbb{R}^d$ denoting the returns of a set of d assets: $\mathbf{r}_t = [r_{1,t}, \ldots, r_{d,t}]'$. For simplicity, let us assume that the vector of returns $\mathbf{r}_t = [r_{1,t}, \ldots, r_{d,t}]'$ follows the process:

$$\mathbf{r}_t = D_t \epsilon_t, \tag{2.1}$$

where $D_t := \text{diag}[\sigma_{1,t}, \ldots, \sigma_{d,t}]$ and $\sigma_{i,t}$ is the conditional standard deviation of the $i-$th component of $\mathbf{r}_t$ at time $t-1$. $(\epsilon_t)_{t\in\mathbb{Z}}$ is an independent, normally distributed zero-mean process in $\mathbb{R}^d$ with components having a unit conditional standard deviation by construction: i.e. $\epsilon_t \sim_{\text{i.i.d.}} \mathcal{N}(0, R_t)$. To simplify the notation, conditional means of $\mathbf{r}_t$ have been set to zero in (2.1).[1] The conditional covariance matrix of ϵ_t at time $t-1$ is denoted by R_t. Therefore, we obtain the following standard factorization of the conditional covariance matrix of $\mathbf{r}_t$:

$$H_t = Cov_{t-1}(\mathbf{r}_t) = D_t R_t D_t. \tag{2.2}$$

This way of writing the conditional covariance matrix H_t of the returns vector at time t suggests a two-step specification and estimation strategy. In such a procedure one first specifies D_t and then focuses on models for the conditional correlation matrix R_t. Given that the focus of our empirical study is on correlations, the model assumed for the individual volatilities is the same across the different model settings for correlations introduced below. In particular, we fit the tree-structured GARCH model proposed by Audrino and Bühlmann (2001) and used in Audrino and Trojani (2011) to each individual return's conditional variance $\sigma_{i,t}^2$, $i = 1, \ldots, d$.[2] The next subsections introduce the different models for correlations R_t to be considered in our study. This list of models is surely not complete, but contains the most relevant, widely used, and best performing approaches introduced in the literature.

All the models we investigate here can be estimated using simple variants of the two-step procedure described in Engle (2002). This procedure applies a quasi-likelihood estimation in which one has to estimate first the parameters involved in the univariate conditional covariance processes and then in a second step the parameters underlying the correlation structure using

[1] Since this study focuses on correlation prediction, this is not a crucial assumption.

[2] As a robustness check, we estimated the individual returns' conditional variance dynamics in D_t using the component GARCH-MIDAS approach proposed by Engle et al. (2009). Final results for correlation prediction were qualitatively the same as the ones described in Section 3. All details about this first step estimation can be obtained by the authors upon request.

the estimated standardized residuals $\hat{\epsilon}_t = \hat{D}_t^{-1}\mathbf{r}_t$ and correlation targeting. The asymptotic properties of the two-step estimator have already been widely discussed in the literature. A detailed discussion is beyond the scope of this study. We refer the interested reader to the specific papers introducing the different models we consider.

2.1.1 Bollerslev's constant conditional correlations (CCC) model

The first model we consider is the constant conditional correlation (CCC) model introduced by Bollerslev (1990). The model specification for the correlation matrix R_t in (2.2) can be easily obtained by setting

$$R_t \equiv \overline{Q} \tag{2.3}$$

constant over time. Thus, in this model both the long-run and the short-run components affecting correlations dynamics are assumed to be constant. One may argue that considering a model with constant correlations is a contradiction when nowadays there is huge empirical evidence that correlations are time-varying. Why then assume constant correlation dynamics? For at least two reasons: first, because this model is still a very popular benchmark used in practice given its (computational) simplicity, in particular when dealing with high-dimensional volatility matrices. Second, because incorporating the idea of mixed data sampling (MIDAS) for the long-run component of correlations in a simple CCC setting yields slowly time-varying correlations that are in fact very accurate for some known data sets, and may also be for others. The inclusion of the CCC model allows us, as in the previous empirical studies, to compare the performance of the other more flexible alternatives introduced below against a simple, misspecified benchmark.

2.1.2 Engle's dynamic conditional correlations (DCC) model

The dynamic conditional correlations (DCC) model proposed by Engle (2002) was the first step in the direction of allowing correlation dynamics to be time-varying and therefore more closely matched to reality. The DCC model is still the benchmark model used both in academic studies and in the private sector because of its parsimony in terms of parameters and, therefore, computational feasibility and immediateness even in large dimensions.[3]

[3]We also estimated the corrected version of the CCC model proposed by Aielli (2008) and (2009). In line with the results already mentioned by the author, we did not find any particular statistical and economic difference between the forecasting power of the standard DCC and the corrected DCC models. For that reason results are

The DCC model for the correlation matrix R_t in (2.2) can be specified as follows:

$$R_t = diag[Q_t]^{-1/2} Q_t \; diag[Q_t]^{-1/2} \tag{2.4}$$

with

$$Q_t = (1 - \phi - \lambda)\overline{Q} + \phi \epsilon_{t-1}\epsilon'_{t-1} + \lambda Q_{t-1} \;, \tag{2.5}$$

where the parameters $\phi \geq 0$ and $\lambda \geq 0$ are imposed to satisfy $\phi + \lambda < 1$ for stationarity reasons, and $\overline{Q}$ is the unconditional covariance matrix of the residuals ϵ_t. Fixing $\phi = 0$ and $\lambda = 0$ we get the CCC model as a special case.

The elements $q_{i,j,t}, \; i,j = 1,\ldots,d$, that build the matrix Q_t in (2.5) are interpreted as the short-run correlations between assets i and j. They are specified to follow some sort of autoregressive dynamics. Therefore, in the classical DCC model, only the short-term movements of the correlation dynamics are explicitly taken into account and modeled. In contrast, the long-run correlations between assets i and j are assumed to be constant over time and are associated with the elements $\overline{q}_{i,j}, i,j = 1,\ldots,d$, that build the matrix $\overline{Q}$. In particular, rewriting equation (2.5) element by element as

$$q_{i,j,t} - \overline{q}_{i,j} = \phi(\epsilon_{i,t-1}\epsilon_{j,t-1} - \overline{q}_{i,j}) + \lambda(q_{i,j,t-1} - \overline{q}_{i,j}), \; i,j = 1,\ldots,d, \tag{2.6}$$

suggests the idea of some short-term movements of correlations around a constant long-run relationship. The assumption of constant long-run correlations will be relaxed in the following two approaches.

2.1.3 Tree-structured dynamic conditional correlations (TreeDCC) model

Different extensions of the DCC model have been proposed in the literature: we decided to consider in this study only the forerunner of this family (i.e. Engle's DCC model) and the most recent generalization of it proposed by Audrino and Trojani (2011), namely the tree-structured DCC model, for the following reasons (not ordered by importance):

- it is a complete data driven approach that includes as nested models both the standard CCC and DCC approaches;

- its empirically proven superior forecasting ability over several different competitors;

reported only for the classical DCC model.

- it allows for, in connection with MIDAS, a natural specification and estimation of the different effects under investigation: short- and long-run components that may be subjected to structural breaks of a different nature;

- and for the sake of comprehensibility and brevity in discussion of the results, allowing it to dispense with many alternatives.

The tree-structured DCC model specification assumes that the dynamics of R_t are subjected to structural breaks defined according to several multivariate thresholds and follow locally Engle-type DCC processes. In order to keep the model tractable in large dimensions as well, it is assumed that thresholds in the R_t dynamics depend on ϵ_{t-1} only via the average

$$\overline{\rho}_{t-1} = \frac{1}{d(d-1)} \sum_{u \neq v} \epsilon_{t-1,u} \epsilon_{t-1,v}$$

of the cross products of the component of ϵ_{t-1}.

To define the parametric threshold function R_t in the model, let $\mathcal{P} = \{\mathcal{R}_1, .., \mathcal{R}_w\}$ be a partition of the state space $\mathbb{G}$ of all relevant predictor variables $\mathbf{X}_{t-1}$.[4] In particular, the dynamics of R_t are specified as follows:

$$R_t = \sum_{i=1}^{w} c_i R_{it} I_{[\mathbf{X}_{t-1} \in \mathcal{R}_i]} + \left(1 - \sum_{i=1}^{w} c_i I_{[\mathbf{X}_{t-1} \in \mathcal{R}_i]}\right) Id_n \tag{2.7}$$

where $c_1, \ldots, c_n \in [0,1]$, Id_n is the $d-$dimensional identity matrix, and the parametric processes for R_{it}, $1 = 1, \ldots, n$, are given by:

$$R_{it} = diag[Q_{it}]^{-1/2} Q_{it} \; diag[Q_{it}]^{-1/2} \tag{2.8}$$

with

$$Q_{it} = (1 - \phi_i - \lambda_i)\overline{Q} + \phi_i \epsilon_{t-1} \epsilon'_{t-1} + \lambda_i Q_{it-1} \; , \tag{2.9}$$

[4] The state space is defined to be as broad as possible including all predictor variables that are thought to be relevant for correlation prediction, such as, for example, the multivariate past return vector $\mathbf{r}_{t-1}$. The class of admissible partitions $\mathcal{P}$ for the correlation function is composed of rectangular partition cells that are delimited by a set of multivariate thresholds. In order to construct such rectangular partition cells, a binary tree in which every terminal node represents a cell $\widetilde{\mathcal{R}}_i$ is used. Estimation of the threshold function in the correlation dynamics is achieved by a high dimensional model selection procedure that determines the optimal number and the structure of the relevant thresholds in the underlying partition. This model selection scheme is not computationally feasible if applied directly to the multivariate time series $(\epsilon_t \epsilon'_t)_{t \in \mathbb{Z}}$. A natural way to reduce estimation complexity is to notice that the partition $\mathcal{P}$ is identical to the one implied by a corresponding tree-structured univariate model for the time series $(\overline{\rho}_t)_{t \in \mathbb{Z}}$. For all details we refer to Audrino and Trojani (2011).

parameters $\phi_i, \lambda_i \geq 0$ such that $\phi_i + \lambda_i < 1$ for all $i = 1, .., w$, and $\overline{Q}$ is defined as in the classical DCC model. When the partition is trivial, the correlation dynamics are globally defined, no relevant structural break is found, and one obtains the standard DCC model by setting $c_1 = \ldots = c_w = 1$.

Moreover, by setting $\phi_i = \lambda_i = 0$ for $i = 1, .., w$, one can rewrite R_t as:

$$R_t = \sum_{i=1}^{w} c_i \overline{Q} I_{[\mathbf{X}_{t-1} \in \mathcal{R}_i]} + \left(1 - \sum_{i=1}^{w} c_i I_{[\mathbf{X}_{t-1} \in \mathcal{R}_i]}\right) Id_n. \tag{2.10}$$

where $\overline{Q}$ is a fixed $d-$dimensional correlation matrix. In this case, we obtain a piecewise constant correlation matrix defined by a multivariate threshold function over the partition $\mathcal{P}$. Once again in the case that the partition is trivial and correlations are constant overall, we get the standard CCC model. Finally, when $\phi_i > 0$ or $\lambda_i > 0$ for $i = 1, .., w$ and $\mathcal{P}$ is not a trivial partition, by setting $c_1 = \ldots = c_w = 1$ we obtain a tree-structured DCC–model where correlations are subjected to regime-shifts and are locally following Engle's DCC-dynamics over the distinct partitioning cells $\mathcal{R}_i$.

In the tree-structured DCC model, the short-run component of correlations is modeled using local DCC-type of processes. The assumption of a constant long-run component implicitly stated by the standard DCC model is relaxed in favor of a more realistic idea of dynamic long-run movements in correlations that are locally constant. Both components may be subjected to structural breaks defined as some multivariate thresholds of the relevant predictor space.

A disadvantage of that model is that two effects, namely a time-varying long-run component and structural breaks, are modeled in such a way that they are strongly interconnected and can not be separated. In fact, the long-run movements of correlations are time-varying only as a consequence of the presence of structural breaks. This problem can be solved by using mixed data sampling, as is presented in the next section.

2.1.4 Mixed data sampling (MIDAS) for the long-run component of correlations

The piecewise constant assumption for the long-run component of the correlations introduced in the tree-structured DCC model may be relaxed even further to a smoother functional form. This task can be accomplished with the use of mixed data sampling (MIDAS), recently introduced in the forecasting volatility literature by Ghysels et al. (2006) and Forsberg and Ghysels (2007), among others. MIDAS can be applied to the constant matrix $\overline{Q}$ in the three settings introduced above, which leads to the following:

- In the CCC setting, to correlations where the long-run component is time-varying whereas the short-term movements are assumed to be constant. Thus, the CCC-MIDAS model is a model with dynamic correlations.

- In the DCC setting, to a relaxation of the constant assumption for the long-run component of correlations. That is, that in the DCC-MIDAS model both short- and long-run components are time-varying.

- In the treeDCC setting, to a generalization of the functional form for the long-run component of correlations that is now changing more smoothly. In the treeDCC-MIDAS model all three effects (time-varying short- and long-run components, and structural breaks) are explicitly and separately taken into account.

In the MIDAS approach for correlations, the constant assumption for the out of the diagonal elements $\overline{q}_{i,j},\ i,j = 1,\dots,d$, building the matrix $\overline{Q}$ in (2.3), (2.5), and (2.9) is relaxed in favor of a judiciously chosen weighted average of historical correlations, that is:

$$\overline{q}_{i,j,t} = \sum_{l=1}^{K_{ij}} \varphi_l(w_{ij}) c_{i,j,t-l} \tag{2.11}$$

where the historical correlations $c_{i,j,t}$ are defined as

$$c_{i,j,t} = \frac{\sum_{k=t-N_{ij}}^{t} \epsilon_{i,k}\epsilon_{j,k}}{\sqrt{\sum_{k=t-N_{ij}}^{t} \epsilon_{i,k}^2}\sqrt{\sum_{k=t-N_{ij}}^{t} \epsilon_{j,k}^2}}$$

and the weighting scheme involves so called Beta weights as

$$\varphi_l(w_{ij}) = \frac{(1-\frac{1}{K_{ij}})^{w_{ij}-1}}{\sum_{k=1}^{K_{ij}}(1-\frac{k}{K_{ij}})^{w_{ij}-1}}.$$

The main idea is that the long-run component of correlations can be filtered from some empirical proxies given by weighted averages of cross-products of residuals. K_{ij} and N_{ij} are tuning parameters of the MIDAS filters that should be optimized during the estimation and may be naturally interpreted as number of lags of realized correlations considered by the filter and number of daily non-overlapping returns needed to compute each realized correlation, respectively.[5] For simplicity reasons, we assume that these parameters are the same across assets, i.e. $N_{ij} = N$ and $K_{ij} = K$ for all $i, j = 1, \dots, d$. w_{ij} are the parameters defining the decay patterns of the

[5] We refer to Colacito et al. (2010) for a complete description of the optimization strategy of the tuning parameters.

weights. Similarly, we consider these parameters to be the same for correlations between the same type of assets. In particular, in the industry portfolio example we will consider MIDAS filters of order 1 (same decay across all assets) and MIDAS filters of order 2 (two different decay parameters allowed: the first one for the correlations between different industry portfolios, and the second one for correlations between an industry portfolio and the long term bond). In the US stock market example we will focus only on MIDAS filters of order 1, given that all assets belong to the same stock market and therefore it is reasonable to assume that the long-run components of correlations across stocks show similar patterns.

The interpretation of the CCC-, DCC-, and treeDCC-MIDAS models is similar to what we derived in (2.6) for the DCC approach: two components for correlations are extracted, one pertaining to short-term movements that are estimated using CCC or DCC models, the other pertaining to a secular, slowly varying component that may well be linked to macroeconomic fundamentals and is estimated via MIDAS. The short-term dynamics of correlations are supposed to move around the long-run component. Moreover, MIDAS in connection with treeDCC allows such components to switch across different regimes caused by structural breaks.

2.2 The data sets

In our empirical study we test the forecasting accuracy of the different competitors on two updated data sets that have already been used in the literature to show the goodness of fit of the models introduced above. In particular we will focus our analysis on correlations between Fama-French industry portfolios and the long-term US bond, and correlations among different US stock assets traded in the S&P 500. In both cases, the forecasting power of the different components affecting correlation dynamics, namely short- and long-run components and structural breaks, is applied in a very difficult situation that may be interpreted as stress testing. In fact, the out-of-sample period considered is the most recent years and therefore includes the late-2000s financial crisis. More details on the data are provided in the next two subsections.

2.2.1 Industry portfolios and the 10 year bond

The data set we analyze is exactly the same already investigated in Colacito et al. (2010). Our in-sample period starts on July 15, 1971, and ends on June 30, 2006, for a total of 8703 daily observations. This period is the same as the one analyzed in Colacito et al. (2010). The out-of-sample period goes from July 2006 to December 2009, for a total of 875 days.

We investigate the forecasting impact of explicitly modeling the different components affecting correlations on two three-dimensional problems, in both cases considering two industry portfolios and the 10 year bond. Data for the 10 industry portfolios and all details on how the different industry portfolios are constructed can be freely downloaded from the Kenneth French web-page. Exactly as in Colacito et al. (2010), we focus on correlations among the Energy and Hi-Tech portfolios vs. the 10 year bond in the first example and among the Manufacturing and Shops portfolios vs. the 10 year bond in the second example. Given that three assets and two different markets (i.e. stock and bond markets) are involved, we estimate models using one and two MIDAS filters. The use of two MIDAS filters allows us to take into account possible different decays in the weights when looking at correlations between the two industry portfolios and correlations between an industry portfolio and the 10 year bond.

2.2.2 US equity returns

We consider the same 9-dimensional data set already studied in Audrino and Trojani (2011), that is the multivariate time series of (annualized) daily log-returns for the following US stocks: Alcoa, Citigroup, Hasbro, Harley Davidson, Intel, Microsoft, Nike, Pfizer, and Exxon. Data are for the sample period between January 2, 2001 and November 15, 2010, amounting to 2483 trading days. The source of the data is *Tick Data*, a division of Nexa Technologies, Inc. (see the webpage `http://www.tickdata.com`). Using these tick-by-tick data, we construct so-called realized covariances and correlations as more accurate proxies of the unavailable and unobservable true covariances and correlations using the method presented in Corsi and Audrino (2007). This will allow us to extend the number of performance measures we can use to evaluate the different models, as shown below in Section 2.3.

We split the sample into two subperiods. The first one consists of 1507 trading days, from January 2, 2001 to December 29, 2006. Data from this subperiod are used for in-sample estimation. The second subperiod consists of the remaining 976 observations, up to November 15, 2010, and is used for out-of-sample performance evaluation.

For this application we test the forecasting power of the different models using only MIDAS filters of order 1, given that all assets belong to a common market.

2.3 The evaluation criteria

The criteria we use to discriminate between the forecasting power of the models (and therefore of the components that are supposed to affect correlation dynamics) are of two different types: statistical and economic measures.

2.3.1 Statistical performance

To quantify and compare the out-of-sample fit of the different models, we compute several goodness-of-fit statistics for conditional covariances. Since the individual volatility processes are kept identical for all models we consider in our horse race, this comparison allows us to investigate the additional explanatory power of the components affecting the correlation dynamics. We consider the following goodness-of-fit measures:

- The multivariate negative log-likelihood statistic (NL),
- The multivariate version of the classical mean absolute error statistic (MAE),
- The multivariate version of the classical mean squared error statistic (MSE).

The last two performance measures require the specification of sensible values for the unknown true conditional covariance matrix. A powerful way of computing good proxies for this matrix is by means of the so-called realized covariance approach. For this one needs high-frequency data, as for example in our second application in which we use tick-by-tick return data to compute the realized covariance between returns using the methodology proposed in Corsi and Audrino (2007). Unfortunately such high-frequency data are not available for our first data set. Using accurate proxies for the unobservable conditional covariance matrix allows us to get reliable rankings of the models with respect to the performance measures considered, avoiding possible misleading results implied by an unfortunate choice of the loss function; see Laurent et al. (2009) for more details.

To evaluate whether differences in performance among the models considered are statistically significant, we perform a series of tests recently proposed in the literature. In particular, we test whether the additional flexibility allowed by the use of specific dynamics for the short-run and long-run components, as well as the explicit inclusion of structural breaks, yields to significant improvements in the forecasting accuracy of the correlation predictions using the following methods:

- We perform the superior predictive ability (SPA) test introduced by Hansen (2005). This allows us to verify whether each of the models considered is significantly outperformed by one (or more) of the alternatives.

- We construct model confidence sets (MCS) at different confidence levels as proposed by Hansen et al. (2003) and (2011). The MCS approach has been introduced with the goal of characterizing the best subset of models with respect to a pre-specified performance measure out of a set of competing ones.

2.3.2 Economic significance

We investigate the economic significance of including the different components affecting correlations in the models using the portfolio selection application first introduced in Engle and Colacito (2006) and extensively used thereafter in the academic literature; see, among others, Bandi et al. (2008), Colacito et al. (2010) or Audrino and Trojani (2011).

The idea is that an investor solves the classical mean-variance portfolio allocation problem to get the optimal weights of the different constituents of the portfolio. That is, the investor minimizes the expected variance of the portfolio under the constraint that the weights add to some scalar value. Clearly, to solve such an optimization problem, the investor should know the one-period ahead conditional covariance matrix of the assets composing the portfolio under investigation. Given that this matrix is unknown, conditional covariances one-period ahead are predicted using the different approaches illustrated in Section 2.1.

Engle and Colacito (2006) show that in order to quantify the gains from superior covariance prediction one can look at the ratio

$$\frac{\sigma_t - \sigma_t^*}{\sigma_t^*} \geq 0,$$

where σ_t and σ_t^* are the portfolio volatility when the investor chooses the weights according to the covariance predictions given by a pre-specified model and using the true unknown covariance matrix, respectively. This ratio can be interpreted as the percentage reduction in portfolio investment that could have been achieved by knowing the true covariance matrix.

Moreover, when two alternative models (and therefore two alternative series of covariance predictions) are at one's disposal, one can test differences in the following way. For each period, the mean-variance problem is solved and portfolio weights are constructed based on each covariance matrix prediction. Denote the portfolio returns attained according to each one-step ahead covariance matrix prediction as π_t^j, where j identifies the model used for predicting the covariance matrix. One can perform a test on the series of the differences of the squared returns on the two portfolios $u_t = (\pi_t^{j_1})^2 - (\pi_t^{j_2})^2$ using generalizations of the standard Diebold and Mariano (1995) test. The null hypothesis of equal portfolio variance is simply a test that the mean of the differences u_t is zero.

3 Empirical results

Similarly to what has been done in Section 2.2, we split the discussion of the results in two subsections corresponding to the two data sets under investigation. Given that the focus of this study is on out-of-sample forecasting ability, the in-sample estimation results are not presented and discussed. Nevertheless, all in-sample estimation results (in terms of values and significance of the estimated parameters) are qualitatively the same as those discussed in the previous studies published in the literature, namely Colacito et al. (2010) for the DCC-MIDAS and Audrino and Trojani (2011) for the treeDCC approach. Results of the in-sample estimation can be obtained by the authors upon request.

3.1 Industry portfolios and the 10 year bond

We first focus on the three-dimensional example of energy portfolio returns, hi-tech portfolio returns, and 10 year bond differences. A representative picture of the type of correlation estimates we get is illustrated in Figure 1. As described in the introduction to this paper, it is quite difficult to identify significant differences among the correlation estimates obtained using the models. The immediate consequence is that it is almost impossible to disentangle the type of correlation dynamics that can be attributed to the short- and long-run components. In estimating the treeDCC approach, we found a significant regime-shift in correlations when past hi-tech portfolio returns are low (i.e. below the 25%-quantile of the realized returns in the sample) or high (i.e. above the same threshold). Correlation dynamics are found to be locally of a DCC-type. Nevertheless, correlations obtained using this tree-structured two-regimes DCC model are qualitatively indistinguishable when superimposed on those obtained using the classical DCC model in a picture similar to Figure 1.

What then are the consequences for correlation prediction accuracy? Results for the out-of-sample negative log-likelihood, as well as AIC and BIC criteria based on the in-sample log-likelihood are summarized in Panel A of Table 1. AIC and BIC are reported because they allow for a direct comparison with the results found in Colacito et al. (2010). p-values of the SPA tests for the null hypothesis of each model not being significantly outperformed by any of the alternatives are reported in parentheses.

TABLE 1 ABOUT HERE.

Results of the AIC and BIC criteria are the same as those presented in Colacito et al. (2010): they support the inclusion of a MIDAS specification for the long-run component of correlations. We also found evidence that only one MIDAS filter is needed and supported by the data (except for the CCC-setting).

Turning to the analysis of the predictive accuracy of the models in a true out-of-sample forecasting exercise, we see that differences among the models are very small. Clear statistically significant gains measured by the SPA tests are found only when including an explicit MIDAS long-run component for correlations in the CCC setting. Nevertheless, this result can be simply interpreted as a rejection of the assumption of constant correlations in favor of more realistic time-varying correlation dynamics. Given the estimates shown in Figure 1, we may not be sure that such time-varying dynamics can be attributed to the long-run component of correlations. The simple DCC model yields similar time-varying dynamics for correlations but focuses only on the short-run component. MIDAS does not seem to improve the accuracy of the correlation predictions when added to DCC-type short-term dynamics, both in the absence and presence of structural breaks. By contrast, some marginal improvements can be observed when introducing regime-shifts: the two-regimes treeDCC model in fact yields the smallest out-of-sample negative log-likelihood.

To end the statistical evaluation of the predictions, we constructed model confidence sets (MCS) at different significance levels based on the out-of-sample negative log-likelihood. At all relevant levels, the only model that was eliminated from the set of possible alternatives building the MCS was the simple CCC model, once again implying that (i) the assumption of constant correlations is clearly violated, and (ii) no other significant differences can be found among the other models. A simple CCC-MIDAS approach performs equally well as the other more flexible alternatives for this data set.

The results on the statistical performance are confirmed in testing the economic significance of the differences. No significant difference is found when the simple CCC model is left out from the set of possible alternatives. The null hypothesis of equal portfolio variance tested using Diebold and Mariano-type of tests is never rejected. Moreover, efficiency gains are never larger than a few basis points.

The same analysis is performed on the three-dimensional example of manufacturing portfolio returns, shops portfolio returns, and 10 year bond differences. Results of our forecasting exercise are summarized in Table 1, Panel B. Once again, results of the AIC and BIC criteria are in line with those presented in Colacito et al. (2010), except for the fact that we do not find any additional value in considering a second MIDAS filter in the different approaches.[6] All results and conclusions that may be derived from the out-of-sample experiment are similar to those described and discussed above for the first three-dimensional example on the energy portfolio, the hi-tech portfolio and the 10 year bond. For the sake of brevity, we do not repeat the whole discussion here. It is important to notice, however, the fact that in this case modeling simultaneously long-term correlation movements and regime-shifts based on threshold structured structural breaks yield to marginal improvements in the forecasting power of the correlation models.

3.2 US equity returns

As a second forecasting exercise we consider nine of the most frequently traded US stocks belonging to the S&P 500. In this case differences among the correlation estimates obtained using the models are more evident as highlighted in Figures 2 and 3, which may be taken as representative pictures of the behavior of correlation estimates among the stocks. The different components affecting correlation dynamics seem to play a more relevant role in this case and can be naturally disentangled and interpreted as long- and short-run movements, respectively.

Estimating the treeDCC model we identified a reach threshold structure similar to those showed in Audrino and Trojani (2011), although not exactly the same given that the in-sample time period under investigation in this study is longer. The final treeDCC model has four regimes depending on the realizations of the past lagged returns of the Microsoft, Exxon, and Citigroup stocks. The short-term component of correlation dynamics is found to be locally constant in these four regimes. The long-term movements of correlations are slowly varying with drastic regime shifts when a break occurs.

[6] This difference may be due to the different approach we use for estimating individual volatilities in the first step of the two-step procedure.

Results for the out-of-sample negative log-likelihood, as well as the out-of-sample aggregated mean absolute error (MAE) and aggregated mean squared error (MSE) are summarized in Table 2, Panel A. For the MAE and MSE statistics we consider as an accurate proxy of the true unknown covariance matrix the one constructed from high-frequency data. Similar to Table 1, p-values of the SPA tests are reported in parentheses.

TABLE 2 ABOUT HERE.

Results of the statistical performances clearly show that models that do not take the short-term DCC structure of correlations into account are clearly outperformed. Allowing for a more flexible specification of the long-term component, as it is done via MIDAS or regime-shifts, seems not to be crucial for reaching a superior predictive performance. More, such increased flexibility of the models yields in most cases to overfitting and a deterioration of the final accuracy of the correlation predictions. These results are consistent across the different performance measures considered.

These results are partially confirmed when constructing MCS at different significance levels, as is shown in Table 2, Panel B. In fact, the optimal MCS with respect to the out-of-sample negative log-likelihood include only the DCC-based models. Predictive differences among the models are more difficult to see when MCS are constructed using the out-of-sample aggregated mean absolute and squared errors: In these cases only the very simple CCC model and the CCC-MIDAS(1) models seem to be eliminated when constructing the MCS. A possible reason for this may be a low power of the procedure in constructing MCS due to the approximation of the true unknown covariance matrix.

Finally, we verify whether the statistical differences in the accuracy of the correlation predictions across the models translate to relevant economic gains from the viewpoint of an investor who is willing to use such predictions. In testing the series of squared portfolio returns pairwise differences, the null hypothesis of equal portfolio variance can never be rejected. Efficiency gains are once again on the order of few basis points (never larger than 15 bp), yielding to the conclusion that, economically speaking, the very simple CCC model also remains very competitive in our stress testing experiment.

4 Conclusions

In this empirical study, we tested the forecasting performance of several correlation models introduced in the recent literature in a very hard exercise: in fact the out-of-sample period in our applications corresponded to the late-2000s financial crisis. In particular, we focused on model specifications for a flexible estimation of only one or simultaneously both long-run and short-run components affecting correlation dynamics. In a subsequent extension, we increased flexibility of the correlation models even more by allowing the components to be subjected to structural breaks of a different nature. We estimated the long-run component via mixed data sampling (MIDAS), the short-term movement according to the dynamic conditional correlation (DCC) model introduced by Engle (2002), and we determined the type and nature of the structural breaks using the recent tree-structured DCC model introduced by Audrino and Trojani (2011).

The main results of our study can be summarized as follows. First, we found that for some data sets it is very difficult to disentangle the part of the correlation dynamics related to the long-run component from the one that may be estimated using a simple DCC model for short-term correlation behavior. This superimposition of the two components has as a consequence a similar forecasting performance across the models. By contrast, testing the models on US equity returns, we found that separately modeling the long- and short-run components of correlations yields superior results in term of statistical accuracy of the predictions. In this case the two effects can be naturally separated and interpreted. The most important part that models must correctly estimate seems to be related to the short-term movements of correlations. In fact we found that the simple DCC model was never outperformed by its competitors. This result may be due to the out-of-sample period we chose, corresponding to the late-2000s financial crisis, for which having an accurate perception of the short-term correlation changes is probably more important than correctly identifying the long-trend showed by correlations.

Second, allowing for structural breaks marginally increased the forecasting power of the models, although it never yielded to statistically significant improvements in the accuracy of the correlation predictions. Third, when we tested whether such (statistically significant) differences in correlations also might have been economically profitable for private investors we did not find any significant economic gains on the top of a few basis points. Once again this suggests that it is very difficult to improve the accuracy of correlations predictions obtained by very simple models in periods of market turbulence.

Our results concern only the late-2000s financial crisis. In the future we plan to investigate in greater depth the differences between the short- and long-run components of correlations and the performance of the different approaches in other stress-testing exercises. To this end, we plan to consider other data sets and periods of market turmoil, induced by different causes than those related to the late-2000s financial crisis. This may lead to a different appraisal of the predictive power of the main components affecting correlation dynamics.

References

Aielli, G.P. (2008). Consistent estimation of large scale dynamic conditional correlation. Working paper 48, University of Messina.

Aielli, G.P. (2009). Dynamic conditional correlations: on properties and estimation. Working paper, University of Florence.

Audrino, F. and Bühlmann, P. (2001). Tree-Structured GARCH Models. *Journal of the Royal Statistical Society, Series B*, **63**, No. 4, 727-744.

Audrino, F. and Trojani, F. (2011). A general multivariate threshold GARCH model with dynamic conditional correlations. *Journal of Business and Economic Statistics* **29**, Issue 1, 138-149.

Bandi, F., Russell, J.R. and Zhu, Y. (2008). Using high-frequency data in dynamic portfolio choice. *Econometric Reviews* **27**, 163-198.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *The Review of Economics and Statistics* **72**, 498-505.

Colacito, R., Engle, R.F., and Ghysels, E. (2010). A component model for dynamic correlations. *Journal of Econometrics*, forthcoming.

Corsi, F. and Audrino, F. (2007). Realized correlation tick-by-tick. Discussion Paper, Department of Economics, University of St. Gallen.

Diebold, F.X. and Mariano, R.S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics* **13**, 253-263.

Engle, R.F. (2002). Dynamic conditional correlation - a simple class of multivariate GARCH models. *Journal of Business and Economic Statistics* **20**, 339-350.

Engle, R. and Colacito, R. (2006). Testing and valuing dynamic correlations for asset allocation. *Journal of Business and Economic Statistics* **24**, 238-253.

Engle, R., Ghysels, E., and Sohn, B. (2009). On the economic sources of stock market volatility. NYU Working Paper No. FIN-08-043.

Forsberg, L. and Ghysels, E. (2007). Why do absolute returns predict volatility so well? *Journal of Financial Econometrics* **5**, Issue 1, 31-67.

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

Hansen, P. R. (2005). A test for superior predictive ability. *Journal of Business & Economic Statistics* **23**, 365–380.

Hansen, P.R., Lunde, A. and Nason, J.M. (2003). Choosing the best volatility models: The model confidence set approach. *Oxford Bulletin of Economics and Statistics* **65**, 839-861.

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

Laurent, S., Rombouts, J.V.K. and Violante, F. (2009). Consistent ranking of multivariate volatility models. CORE Discussion Paper.

Ledoit, O., Santa-Clara, P. and Wolf, M (2003). Flexible multivariate GARCH modeling with an application to international stock markets. *The Review of Economics and Statistics*, **85**, Issue 3, 735-747.

Pelletier, D. (2006). Regime switching for dynamic correlations. *Journal of Econometrics* **131**, 445-473.

CCC-MIDAS vs. DCC estimates: Industry portfolios returns

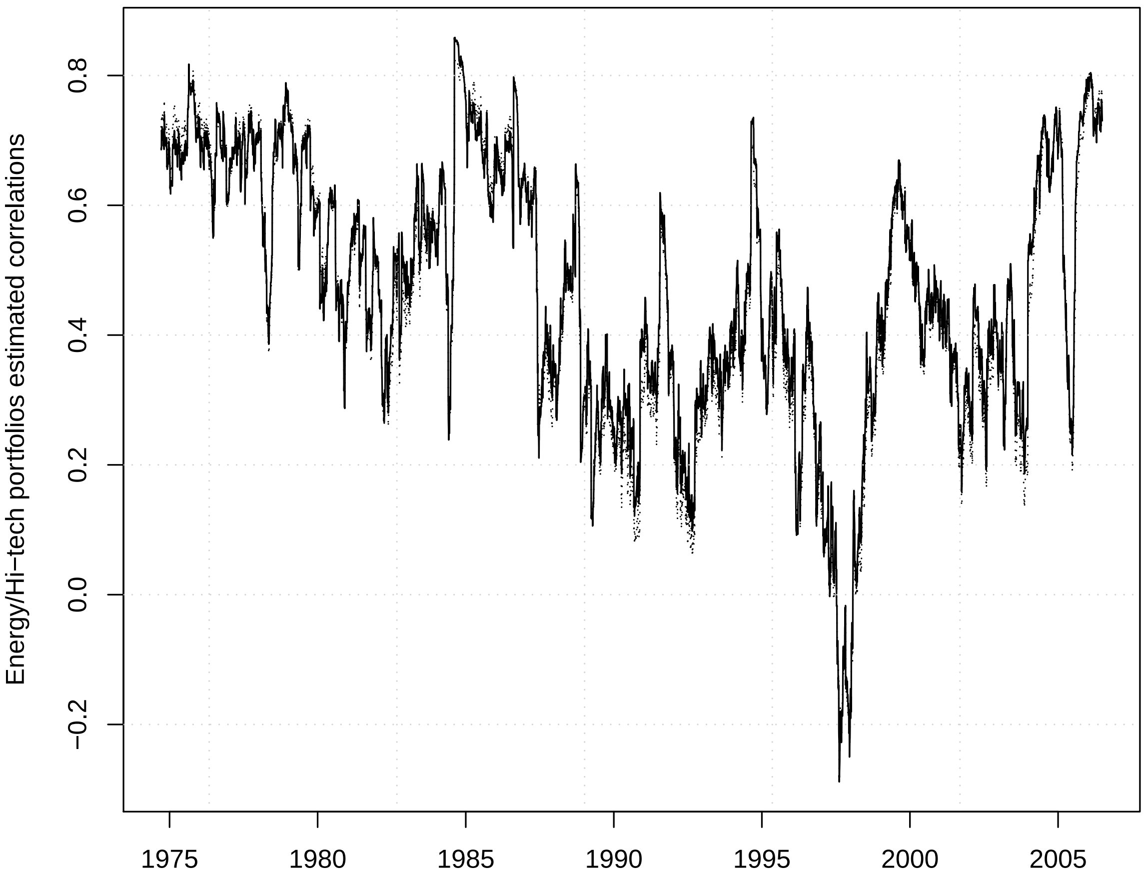

Figure 1: Estimated dynamic correlations between the energy and hi-tech Fama-French industry portfolios for the period between July 15 1971 and June 30 2006 for a total of 8703 daily observations. Correlations are estimated using a component model in a constant conditional correlations (CCC) setting and MIDAS for the long-run correlation component (solid line), and Engle's dynamic conditional correlation (DCC) model (dotted line).

CCC-MIDAS vs. DCC estimates: US stock returns

Figure 2: Estimated dynamic correlations between Nike and Microsoft returns for the period starting in January 2001 and ending in December 2006 for a total of 1507 daily observations. Correlations are estimated using a component model in a constant conditional correlations (CCC) setting and MIDAS for the long-run correlation component (solid line), and Engle's dynamic conditional correlation (DCC) model (dotted line).

CCC-MIDAS vs. DCC-MIDAS estimates: US stock returns

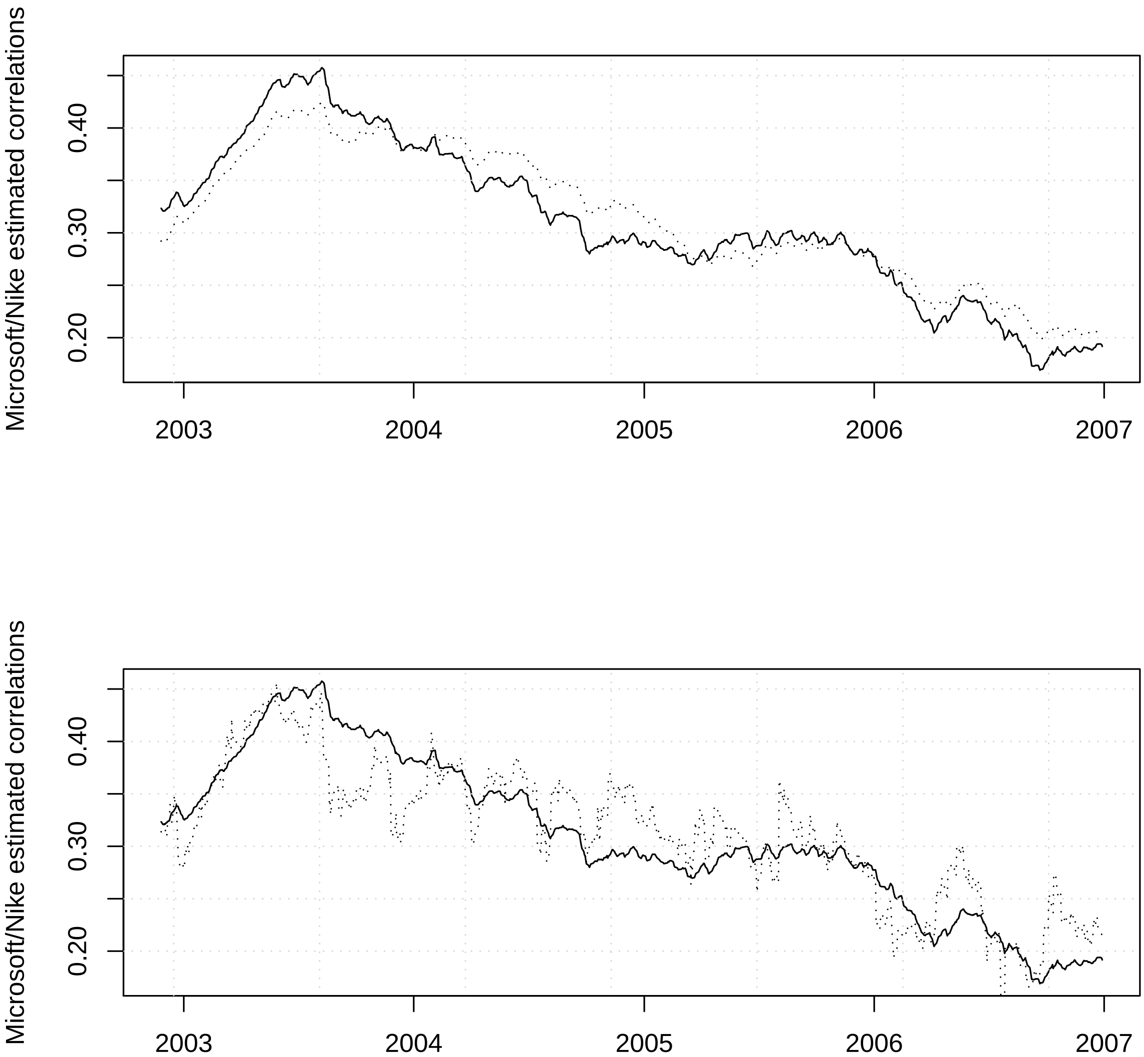

Figure 3: Estimated dynamic correlations between Nike and Microsoft returns for the period starting in January 2001 and ending in December 2006 for a total of 1507 daily observations. Correlations are estimated using a component model in a constant conditional correlations (CCC, solid line) and a dynamic conditional correlation (DCC, dotted line) setting and MIDAS for the long-run correlation component. The upper plot shows the estimated long-run component, whereas final estimated correlations are presented in the bottom panel. Note that in the CCC-MIDAS approach they are the same given that no short-run dynamics are estimated.

Panel A. Energy, Hi-Tech, and 10 Year Bond

Model	AIC	BIC	OS-LogLik
CCC	25547.74	25735.37	2035.116 (0)
CCC-MIDAS(1)	24522.86	24717.44	1839.736 (0.0379)
CCC-MIDAS(2)	24507.98	24709.51	1830.525 (0.6065)
DCC	24406.92	24608.45	1833.819 (0.127)
DCC-MIDAS(1)	24378.54	**24589.96**	1838.683 (0.2003)
DCC-MIDAS(2)	24380.38	24595.81	1838.568 (0.2287)
treeDCC	24387.36	24602.79	**1829.494** (0.8346)
treeDCC-MIDAS(1)	**24374.80**	24597.18	1836.787 (0.3601)
treeDCC-MIDAS(2)	24376.58	24605.91	1836.605 (0.4016)

Panel B. Manufacturing, Shops, and 10 Year Bond

Model	AIC	BIC	OS-LogLik
CCC	19183.15	19370.78	1547.373 (0)
CCC-MIDAS(1)	18729.99	18924.58	1344.597 (0.1064)
CCC-MIDAS(2)	18699.65	18901.19	1342.955 (0.2757)
DCC	17932.01	18133.54	1341.167 (0.7503)
DCC-MIDAS(1)	17887.61	18096.09	1337.191 (0.3445)
DCC-MIDAS(2)	17889.61	18105.04	1337.322 (0.3124)
treeDCC	17836.17	18051.60	1339.775 (0.9514)
treeDCC-MIDAS(1)	17774.33	**17996.71**	1336.432 (0.6984)
treeDCC-MIDAS(2)	**17774.09**	18003.42	**1336.188** (0.8243)

Table 1: Statistical performance results for predicting correlations among industry portfolios and the 10 year bond. In particular we consider the energy portfolio, the hi-tech portfolio, and the 10 year bond (Panel A), and the manufacturing portfolio, the shops portfolio, and the 10 year bond (Panel B). The in-sample time period starts on July 15, 1971, and ends on June 30, 2006 (8703 days) whereas the out-of-sample period goes from July 2006 to December 2009 (875 days). AIC, BIC, and OS-LogLik denote the Akaike information criterion, the Bayesian-Schwartz information criterion, and the out-of-sample negative log-likelihood, respectively. The p-values of the SPA tests are reported in parentheses.

Panel A. US Equity Returns: Statistical Performance

Model	OS-LogLik	OS-MAE	OS-MSE
CCC	40778.76 (0)	618.1319 (0)	1820.347 (0.0154)
CCC-MIDAS(1)	40740.28 (0)	628.7089 (0)	1820.381 (0.0009)
DCC	**40684.18** (0.5069)	**612.2878** (0.1506)	1793.672 (0.9337)
DCC-MIDAS(1)	40692.97 (0.3078)	618.6040 (0.027)	**1792.806** (0.195)
treeDCC	40745.53 (0)	606.4264 (0.5057)	1806.591 (0.4327)
treeDCC-MIDAS(1)	40739.92 (0.0001)	613.1020 (0.0648)	1808.674 (0.0472)

Panel B. US Equity Returns: Model Confidence Sets

Performance	MCS
OS-LogLik	DCC and DCC-MIDAS(1) (at all confidence levels)
OS-MAE	all models except CCC (5%, 10%) and CCC-MIDAS(1) (10%)
OS-MSE	all models except CCC-MIDAS(1) (10%)

Table 2: Statistical performance results for predicting correlations for a nine-dimensional US equity stock example. In Panel A, results are reported for the out-of-sample negative log-likelihood (OS-LogLik), the out-of-sample aggregated mean absolute errors (OS-MAE), and the out-of-sample aggregated mean squared errors (OS-MSE). In parentheses the p-values of the SPA tests are shown. The obtained model confidence sets (MCS) at different significance levels are described in Panel B. The in-sample time period goes from January 2001 until the end of 2006 (1507 days), whereas the out-of-sample period considers the time period between January 2007 and the November 15, 2010 (976 days).