

Working paper

2019-16

Statistics and Econometrics
ISSN 2387-0303

Comparing Forecasts of Extremely Large Conditional Covariance Matrices

Guilherme V. Moura, André A.P. Santos, Esther Ruiz

Serie disponible en



<http://hdl.handle.net/10016/12>

Creative Commons Reconocimiento-
NoComercial- SinObraDerivada 3.0 España
([CC BY-NC-ND 3.0 ES](http://creativecommons.org/licenses/by-nc-nd/3.0/es/))

Comparing Forecasts of Extremely Large Conditional Covariance Matrices

Guilherme V. Moura^a, André A. P. Santos^b, Esther Ruiz^{c,*}

*^aDepartment of Economics
Universidade Federal de Santa Catarina*

*^bBig Data Institute
Universidad Carlos III de Madrid
and
Department of Economics
Universidade Federal de Santa Catarina*

*^cDepartment of Statistics
Universidad Carlos III de Madrid*

Abstract

Modelling and forecasting high dimensional covariance matrices is a key challenge in data-rich environments involving even thousands of time series since most of the available models suffer from the curse of dimensionality. In this paper, we challenge some popular multivariate GARCH (MGARCH) and Stochastic Volatility (MSV) models by fitting them to forecast the conditional covariance matrices of financial portfolios with dimension up to 1000 assets observed daily over a 30-year time span. The time evolution of the conditional variances and covariances estimated by the different models is compared and evaluated in the context of a portfolio selection exercise. We conclude that, in a realistic context in which transaction costs are taken into account, modelling the covariance matrices as latent Wishart processes delivers more stable optimal portfolio compositions and, consequently, higher Sharpe ratios.

Keywords: Covariance forecasting, GARCH, Minimum-variance portfolio, Portfolio turnover, Risk-adjusted return, Stochastic volatility.

JEL classification: C53; G17.

*Corresponding author. e-mail: ortega@est-econ.uc3m.es

1. Introduction

Conditional covariance matrices are a fundamental ingredient in many financial problems that require modelling and forecasting the covariation among financial returns. However, the number of off-diagonal elements of the covariance matrices increases exponentially with the number of assets and, consequently, modeling and forecasting the covariation among financial returns becomes challenging when the number of assets involved is large, say in the order of hundreds or even thousands. Most of the available econometric specifications for dynamic covariance matrices suffer from the curse of dimensionality, having difficulties in the estimation of the model parameters. Moreover, it is important to keep in mind that conditional covariance matrices should be defined on the manifold of symmetric positive-definite matrices, therefore raising further problems when dealing with financial systems with a large number of variables.

Most models for conditional covariance matrices can be classified within one of two main families: multivariate GARCH (MGARCH) and multivariate Stochastic Volatility (MSV). The most popular family of models for conditional covariance matrices is the MGARCH family. Since the seminal contribution of Bollerslev, Engle and Wooldridge (1988), a variety of MGARCH specifications have been proposed in the literature; see Bauwens, Laurent and Rombouts (2006) and Silvennoinen and Teräsvirta (2009) for excellent surveys and Engle (2009) and Francq and Zakoian (2019) for complete texts about the topic. Recently, Almeida, Hotta and Ruiz (2018) and Boudt, Galanos, Payseur and Zivot (2019) survey feasible MGARCH models for large-scale applications and discuss the trade-off between feasibility and flexibility. However, the systems considered in these surveys only contain 10 assets. In any case, Almeida, Hotta and Ruiz (2018) conclude that the most flexible MGARCH parameterization is the popular Dynamic Conditional Correlation (DCC) model originally proposed by Engle (2002). Note that the positivity of conditional covariance matrices can be easily guaranteed in the context of the DCC model. Very recently, Engle, Ledoit and Wolf (2019) robustify the DCC model with correlation targeting by estimating the unconditional correlation matrix using the nonlinear shrinkage (NLS) approach of Ledoit and Wolf (2012) and estimating the dynamic parameters using the composite likelihood method of Pakel, Shephard, Sheppard and Engle (2019). Engle, Ledoit and Wolf (2019) implement their robustified

DCC model in a large-scale portfolio selection problem with up to 1000 assets without taking into account transaction costs.

Alternatively, prompted by the flexibility and success of univariate stochastic volatility (Carnero, Peña and Ruiz, 2004), generalizations of the univariate state-space model for variances to a multivariate setting have received a great deal of attention since the original proposal of Harvey, Ruiz and Shephard (1994); see, for example, Asai, McAleer and Yu (2006), Yu and Meyer (2006), Chib, Omori and Asai (2009) and Kastner (2019). As with any multivariate specification of conditional covariances, it is important to guarantee positive definiteness of MSV models. One attractive specification of MSV models, originally proposed by Uhlig (1994, 1997), which guarantees positive definiteness of conditional covariance matrices, is based on treating unobserved dynamic precision matrices as Wishart processes. However, some of the Wishart MSV (WMSV) specifications suffer from the curse of dimensionality and can only be implemented to systems with relatively small cross-sectional dimensions. For example, Philipov and Glickman (2006a) consider a system with 12 series while Philipov and Glickman (2006b) deal with 88 series. Asai and McAleer (2009) and Gouriéroux, Jasiak and Sufana (2009) have empirical applications with systems of 3 variables and, finally, Golosnoy, Gribisch and Liesenfeld (2012) analyze 5 series. The original WMSV specification proposed by Uhlig (1994, 1997) is interesting when dealing with very large systems of returns because the dynamic dependence of the covariance matrices is controlled by just one single parameter that can be estimated by Maximum Likelihood (ML); see Kim (2014) and Moura and Noriller (2019).

The main contribution of this paper is to compare the robustified DCC model of Engle, Ledoit and Wolf (2019) and the WMSV model of Uhlig (1994, 1997) in the context of the same portfolio selection exercise considered by the former authors. The portfolios are constructed in the context of the entire universe of NYSE, NASDAQ and AMEX stock returns observed daily from 1970 to 2016. We consider investment universes of $N \in \{100, 500, 1000\}$ assets and obtain optimal minimum variance portfolios re-balanced on a monthly basis. Conditional covariance matrices are also computed using the popular RiskMetrics' approach (see J.P.Morgan/Reuters (1996), Mina and Xiao (2001), Zumbach (2007b,a) and Alexander (2008)) as well as the standard sample unconditional covariance

estimator based on a rolling window scheme; see Adams, Füss and Glück (2017). We show that the correlations estimated by the WMSV model are smaller, smoother and have less dispersion than those estimated by any of the other three specifications of the conditional covariance matrices. We also conduct a detailed out-of-sample evaluation of the resulting minimum variance optimal portfolios considering not only their portfolio risk but also the portfolio risk-adjusted return measured by the Sharpe ratio. Similar as in DeMiguel, Garlappi and Uppal (2009) and Kirby and Ostdiek (2012), the exercise is carried out not only by comparing the variances of the portfolios but also their turnover ratios as well as the impact of the presence of transaction costs when evaluating the Sharpe ratio. We conclude that, in concordance with the results in Engle, Ledoit and Wolf (2019), optimal portfolios based on conditional covariances obtained with the robustified DCC model outperform all competitors in terms of standard deviation of portfolio returns. However, these portfolios have a larger turnover and, consequently, when transaction costs are taken into account, they have a lower Sharpe ratio in comparison to those obtained with the WMSV model. As a consequence of the smoother correlations, the portfolios selected using the WMSV specification of the conditional covariance matrices have smaller turnover and, consequently, large Sharpe ratios.

Other horse races among forecasts of conditional covariance matrices for financial returns have been carried out in the literature; see Laurent, Rombouts and Violante (2012, 2013) and Almeida, Hotta and Ruiz (2018). However, Laurent, Rombouts and Violante (2012, 2013) consider $N = 10$ and $N = 3$, respectively, and compare the MGARCH models in terms of predictive ability. Also, Almeida, Hotta and Ruiz (2018) only consider systems with up to $N = 10$ assets and compare the forecast performance of MGARCH specifications. Therefore, our main contribution with respect to these comparisons is to compare the models in the context of truly large portfolios based not only on forecasts but also on the portfolio performance using models from both MGARCH and MSV families.

The rest of the paper is organized as follows. Section 2 describes the alternative specifications considered to forecast conditional covariances. In Sections 3 and 4, we fit the models for the conditional covariances considered and compare portfolio performance, respectively. Finally, Section

5 concludes.

2. Covariance matrix specifications

Consider that the $N \times 1$ vector of returns observed at time t , $t = 1, \dots, T$ is given by

$$r_t = H_t^{1/2} \varepsilon_t \quad (1)$$

where ε_t is an $N \times 1$ Gaussian white noise vector with covariance I_N , the $N \times N$ identity matrix, and H_t is the $N \times N$ positive definite conditional covariance matrix of r_t at time t . Next, we briefly describe the alternative specifications to forecast covariance matrices of large systems of returns considered in this paper.

2.1. DCC model

We consider the DCC specification proposed by Engle, Ledoit and Wolf (2019) that merges the original DCC model of Engle (2002) with the shrinkage principle, which is largely applied to portfolio optimization problems in order to obtain covariance matrices less prone to estimation error, specially in high dimensional problems; see, for instance, Ledoit and Wolf (2004a, 2017a). In the DCC model, H_t is decomposed as the product of D_t and Ψ_t as follows

$$H_t = D_t \Psi_t D_t, \quad (2)$$

where D_t is an $N \times N$ diagonal matrix with its i -th diagonal element, $h_{i,t}$, being the conditional standard deviation of the i -th asset. We assume that each $h_{i,t}^2$ follows a univariate GARCH(1,1) process although a variety of univariate conditional variance specifications could be used for this purpose. Finally, Ψ_t is the conditional correlation matrix of the devolitized residuals, $s_t = (r_{1,t}/h_{1,t}, \dots, r_{N,t}/h_{N,t})'$, which is governed by the following correlation targeting dynamics

$$\Psi_t = (1 - \alpha - \beta)C + \alpha s_{t-1} s_{t-1}' + \beta \Psi_{t-1}, \quad (3)$$

where α and β are scalar parameters that guide the dynamics of all correlations and C is the unconditional covariance matrix of the vector of devolatilized residuals, s_t .

Estimation of the DCC model is carried out in three steps. In the first step, QML estimates of the parameters of the univariate GARCH(1,1) models for each asset are obtained. The estimated volatilities are used to devolatilize the return series.

In the second step, the unconditional covariance matrix, C , is estimated. Engle (2002) proposes estimating C by the sample covariance matrix of the devolatilized residuals, s_t . It is known, however, that the standard sample covariance estimator is prone to estimation error. To circumvent this problem, Engle, Ledoit and Wolf (2019) propose estimating C by using the nonlinear shrinkage (NLS) approach of Ledoit and Wolf (2012), denoted by \hat{C} . Although Engle, Ledoit and Wolf (2019) estimate \hat{C} using the QuEST function described in Ledoit and Wolf (2017b), we obtain \hat{C} using the analytical nonlinear shrinkage approach of Ledoit and Wolf (2019) which is much faster and has similar accuracy. Note that, in spite of the fact that devolatilized returns are used as inputs and, regardless of the estimator of C implemented, the diagonal elements of the estimated C matrix tend to slightly deviate from one. Therefore, every column and every row of the estimated C matrix has to be divided by the square root of the corresponding diagonal entry, so as to produce a proper correlation matrix. From now on, the DCC model, in which C is estimated by \hat{C} , will be denoted as DCC-NLS model.

Finally, in the third step, once the unconditional covariance matrix, C , is estimated, the parameters α and β of the correlation-targeting dynamics in (3) are estimated by the composite likelihood (CL) method of Pakel, Shephard, Sheppard and Engle (2019). We compute the log-likelihood by summing up the log-likelihood of all contiguous pairs of assets. Therefore, only $N - 1$ bivariate log-likelihoods should be computed.¹

We discuss in Section ?? of the Supplementary Material two variants for the estimation of the DCC model, namely i) the original DCC proposal of Engle (2002) in which C is estimated by the sample covariance matrix of devolatilized residuals, denoted by DCC-Sample, and ii) the estimator

¹In order to estimate the DCC-NLS model, we use and adapt some of the Matlab codes of the MFE Toolbox developed by Professor Kevin Sheppard from Oxford University and available in his web page.

of C considered in Engle, Ledoit and Wolf (2019) in which C is estimated by the linear shrinkage (LS) approach of Ledoit and Wolf (2004b), denoted as DCC-LS.

2.2. Wishart Multivariate Stochastic Volatility

Consider that returns are given by equation (1). In the WMSV model proposed by Uhlig (1994), the precision matrix is given by

$$H_t^{-1} = \frac{d+1}{d} \mathcal{U}(H_{t-1}^{-1})' \Theta_t \mathcal{U}(H_{t-1}^{-1}), \quad (4)$$

where $\mathcal{U}(H_t^{-1})$ is the upper triangular matrix obtained from the Cholesky decomposition of H_t^{-1} and Θ_t are random iid draws from an N -dimensional singular multivariate beta distribution, $\mathcal{B}_N(\frac{d}{2}, \frac{1}{2})$, as defined by Uhlig (1994), with $d > N - 1$ being a scalar parameter defining its degrees of freedom. The initial condition for the covariance matrix is given by

$$H_1^{-1} \sim \mathcal{W}_N(d+1, [(d+1)S_0]^{-1}), \quad (5)$$

where \mathcal{W}_N denotes the N -dimensional Wishart distribution and $S_0 = E[H_0^{-1}]$. We follow Uhlig (1997) and Kim (2014) and set S_0 to a diagonal matrix whose elements are given by the inverse in-sample variance of each series.² Uhlig (1997) shows that, according to equation (4), the dynamics of the inverse covariance matrices are very similar to those of random walks.

The dynamics of the precision matrix are governed by a unique parameter, d , that can be estimated by Maximum Likelihood (ML); see Kim (2014) and Moura and Noriller (2019). We refer the reader to Section ?? of the Supplementary Material that brings additional details regarding filtering and ML estimation of the WMSV model considered in this paper.

2.3. RiskMetrics

One of the most popular specifications of the conditional covariance matrix, H_t , in equation (1) is based on the RiskMetrics 1994 (hereafter RM-1994) methodology; see J.P.Morgan/Reuters

²We have also considered alternative initial conditional covariance matrices with correlations different from zero. The results were always worse in terms of portfolio performance.

(1996), Mina and Xiao (2001) and Alexander (2008). In short, according to RM-1994, one-step-ahead conditional covariance matrices are obtained as an exponential weighted moving average (EWMA) of quadratic forms of past returns with the weighting designed to give more weight to more recent observations as follows

$$H_t = (1 - \lambda) \sum_{i=1}^{t-1} \lambda^i r_{t-i} r'_{t-i}. \quad (6)$$

where $\lambda = 0.94$ for daily data. Note that, given that $\lambda = 0.94$, the weight placed in older observations is decreasing. This could be an important limitation of RM-1994 when trying to invert large covariance matrices making the “effective window” considered in the EWMA specification in (6) to be small depending of the cross-section dimension of r_t . In spite of this limitation, EWMA is prominent in the industry and among market participants.³

The RM-1994 approach is extended to the RM-2006 approach by Zumbach (2007b,a) in order to incorporate additional stylized facts such as long memory of the volatility process. We follow Engle, Ledoit and Wolf (2019) and specify H_t according to the RM-2006 approach as one of the benchmark covariance specifications.⁴

2.4. Unconditional Covariance matrix

Finally, we also specify H_t as the unconditional covariance matrix of asset returns that is estimated by the sample covariance of returns.

As a robustness check, the unconditional covariance matrix is also estimated by the linear shrinkage (LS) method of Ledoit and Wolf (2004b) and by the analytical nonlinear shrinkage (NLS) method of Ledoit and Wolf (2019).

³The EWMA filter is a particular case of the filter obtained if the Kalman filter were implemented when the parametric model for conditional covariance matrices is the MSV of Harvey, Ruiz and Shephard (1994) with all variances and covariances restricted to have the same variances of the transition noise and such that the smoothing parameter is 0.94. The only difference is that in the model proposed by Harvey, Ruiz and Shephard (1994), the specification is for log-variances while in the RiskMetrics methodology variances are modeled directly.

⁴We use the Matlab routine `riskmetrics2006` available in the MFE Toolbox provided by Kevin Sheppard.

3. Fitting and forecasting large covariance matrices

In this section, we fit the conditional covariance models described in Section 2 to a large cross-section dimension system with up to 1000 assets traded in the US stock market. The data set consists of prices of all NYSE, AMEX and NASDAQ stocks observed daily from 01/01/1970 to 12/31/2016.

The models are recursively estimated using a rolling window scheme based on investment universes with $N \in \{100, 500, 1000\}$ assets starting using data observed from 01/01/1970 to 12/11/1974 with $T = 1250$ observations. The investment universes are obtained as follows. We find the set of stocks that have a complete return history over the most recent $T = 1250$ days as well as a complete return “future” over the next 21 days. We then look for possible pairs of highly correlated stocks, that is, pairs of stocks with returns with a sample correlation exceeding 0.95 over the past 1250 days. With such pairs, if they should exist, we remove the stock with the lower volume on investment date h . Of the remaining set of stocks, we then pick the largest N stocks (as measured by their market capitalization on investment date h) as our investment universe. In this way, the investment universe changes slowly from one investment date to the next. In line with Brandt, Santa-Clara and Valkanov (2009), we do not include the risk-free asset in the investment opportunity set as including this asset boils down to a change in the scale of the stock portfolio weights and is not interesting per se. The parameters are re-estimated every month (we adopt the common convention that 21 consecutive days constitute one month). Therefore, for each model and investment universe, N , we perform a total of 505 rolling window estimations. Using these estimates, at each day from 12/12/1974 to 12/31/2016, we obtain the corresponding one-step-ahead predictions of the covariance matrices, with a total of 10,605 predictions.

Figure 1 plots the time series evolution of the out-of-sample 25th, 50th and 75th percentiles of the one-step-ahead pairwise correlations when $N = 500$ for each of the four specifications of the conditional covariance matrices described in Section 2.⁵ To compute these quantities, at each moment of time t , we vectorize the lower triangular part of the one-step-ahead covariance matrix

⁵The results obtained with $N = 100$ and $N = 1000$ are qualitatively similar.

obtained with each of the models and compute the percentiles of the (vectorized) correlations. The first conclusion from Figure 1 is that the median level of the correlations estimated by the WMSV model is clearly lower than those obtained when the correlations are estimated by any of the other three alternative specifications considered. Note that the median level of the WMSV correlations is around 0.1 while the level is around 0.3 for any of the other specifications.

Figure 1 also shows that both the DCC-NLS and RM-2006 models yield correlations that are highly time-varying and noisy. This finding corroborates those obtained both analytically and empirically in Adams, Füss and Glück (2017) who suggest that some popular conditional correlation models such as the DCC can generate spurious fluctuations in correlations. In contrast, the median correlations implied by the WMSV model evolve in a much more smooth way. They only jump to a higher value between 1991 and 1995 and between 2004 and 2008; see Adams, Füss and Glück (2017) who argue that financial correlations are mostly constant over time with financial shocks leading to breaks that shift the level of correlations. The median correlations obtained when estimating the unconditional covariance matrix are also rather smooth.

Finally, the 25th and 75th percentiles of the pairwise correlations plotted in Figure 1 give a sense of the dispersion of these pairwise correlations at each moment of time. We can observe that the dispersion of the pairwise estimated correlations is much smaller when the WMSV specification is implemented supporting the Dynamic Equicorrelation (DECO) model proposed by Engle and Kelly (2012). Note that, as pointed out by Engle and Kelly (2012), the assumption of equicorrelation makes it possible to estimate arbitrarily large covariance matrices with ease. They also show that DECO models can improve portfolio selection compared to unrestricted dynamic correlation structure as in the DCC model.

According to the results in this section, the WMSV model generates one-step-ahead correlations that are close to the equicorrelation assumption and that simultaneously evolve smoothly, only jumping at particular moments of time. However, given the extremely large dimension of the returns, it is difficult to carry out a forecasting evaluation based on these one-step-ahead pairwise correlations. Consequently, in the next section, we evaluate the performance of the DCC, WMSV, RM-2006 and unconditional one-step-ahead pairwise correlations in an economically meaningful

way by using them to construct minimum variance portfolios. We will see that the differences in the estimated pairwise conditional correlations have important implications for the performance of the optimal portfolios.

4. Large scale portfolio selection

In this section, we perform a large scale portfolio selection exercise to compare the specifications of the conditional covariance matrices described in Section 2 based on a long out-of-sample evaluation period. We discuss the portfolio policy considered in the paper and the methodology for evaluating the portfolio performance. A superior covariance forecasting model should provide portfolios with lower variance and larger Sharpe ratios. Finally, we discuss the empirical results.

Our approach to portfolio construction is largely inspired by Engle, Ledoit and Wolf (2019) and it is reproduced here for the sake of completeness. Portfolio weights are updated on a monthly basis in order to avoid excessive turnover levels associated to daily re-balancing. During a month, we assume that there are no transactions. Denote investment dates by $h = 1, \dots, 505$. As explained in Section 3, at any investment date h , each of the specifications for the conditional covariance H_t considered are estimated using the most recent $T = 1250$ daily returns and the corresponding one-step-ahead forecasts are obtained considering investment universes with $N \in \{100, 500, 1000\}$ assets.

The portfolio policy considered is based on an investor who adopts the minimum variance criterion in order to decide her portfolio allocations. A very large body of literature in portfolio optimization considers this particular policy. For instance, Clarke, De Silva and Thorley (2006, 2011) are extensive practitioner-oriented studies on the performance and on the composition of minimum variance portfolios. Engle and Kelly (2012) also evaluate whether equicorrelation is better than different correlations using minimum variance portfolios. Finally, Kastner (2019) compares alternative covariance matrices when $N = 300$ in terms of minimum variance portfolios.

The minimum variance portfolio problem is defined as follows

$$\begin{aligned} \min_{w_t \in \mathbb{R}^N} \quad & w_t' H_t w_t \\ \text{subject to} \quad & w_t' \iota = 1, \end{aligned} \tag{7}$$

where w_t is the vector of portfolio weights for time t chosen at time $t - 1$, H_t is the conditional covariance matrix of asset returns at time t , and ι is an appropriately sized vector of ones. The solution to (7) is given by

$$w_t = \frac{\iota' H_t^{-1}}{\iota' H_t^{-1} \iota}. \tag{8}$$

In practice, however, feasible portfolio weights, \hat{w}_t are obtained by replacing the unknown covariance matrix, H_t , in equation (8) by an estimate, \hat{H}_t , which, in our case, is obtained at time $t - 1$ using each of the specifications described above.

As a robustness check, we discuss in Section ?? of the Supplementary Material the implementation of alternative portfolio policies. In particular, as in Engle, Ledoit and Wolf (2019), we consider the mean-variance policy with a momentum signal. We also consider the equally-weighted and value-weighted policies, and two alternative versions of the volatility timing policy proposed in Kirby and Ostdiek (2012). The results of the robustness checks are reported in Section ?? of the Supplementary Material.

To take into account the impact of transaction costs on the performance of optimal portfolios (Han, 2006), we follow Della Corte, Sarno and Thornton (2008), Kirby and Ostdiek (2012), and Thornton and Valente (2012) and compute the portfolio return net of transaction costs as follows

$$R_t^P = (1 - c \cdot turnover_t) (1 + \hat{w}_t' r_{t+1}) - 1, \tag{9}$$

where $turnover_t = \sum_{j=1}^N (|\hat{w}_{j,t+1} - \hat{w}_{j,t}|)$ is the portfolio turnover at time t , defined as the fraction of wealth traded between periods t and $t + 1$, and c is the fee that must be paid for each transaction that is measured in terms of basis points (b.p.). French (2008) estimates the trading cost in 2006, based on stocks traded on NYSE, AMEX, and NASDAQ, including “total commissions, bid-ask

spreads, and other costs investors pay for trading services”, and finds that this cost has dropped significantly over time going “from 146 basis points in 1980 to a tiny 11 basis points in 2006.” In order to take into account the impact of proportional transaction costs, we consider the cases in which $c = 0$ and 10 b.p.. We believe that setting $c = 10$ b.p. is an accurate description of the proportional cost paid in order to trade stocks of large firms, and that professional trading firms should be able to trade at even lower transaction costs; see Kirby and Ostdiek (2012) and Ledoit and Wolf (2017a) for a discussion. We also compute the risk-adjusted portfolio return net of transaction costs measured by the SR, which is defined as follows

$$SR = \frac{\bar{R}^P}{\sigma^P}, \quad (10)$$

where \bar{R}^P and σ^P are the average and the standard deviation of portfolio returns net of transaction costs, respectively.⁶

Table 1 reports the average turnover and the average and the standard deviation of the minimum variance portfolio return net of transaction costs, \bar{R}^P and σ^P , respectively, as well as the average SR, computed over the out-of-sample period under the two scenarios of proportional transaction costs considered, namely, $c = 0$ and $c = 10$ b.p. and under the different covariance matrix specifications considered. Panel A brings results for portfolios with $N = 100$ assets, whereas Panels B and C report results for portfolios with $N = 500$ and $N = 1000$ assets, respectively. Table 1 also reports the Sharpe ratios (SR) and turnovers obtained in monthly terms using only out-of-sample observations.

The first important result we observe in Table 1 is that, regardless of the dimension of the portfolio, N , and of whether transactions costs are considered or not, the minimum variance portfolios obtained with the correlations estimated by RM-2006 have more risk (measured in terms

⁶As suggested in Ledoit and Wolf (2008), we could test the statistical significance of the differences between the portfolio variances and Sharpe ratios of the portfolios derived from the alternative conditional covariance matrix specifications using tests based on the stationary bootstrap procedure of Politis and Romano (1994), which allows to formally compare optimal portfolios obtained with alternative conditional covariance specifications in terms of their sample characteristics. However, given that we are obtaining $L = 10605$ out-of-sample one-step-ahead portfolio returns, even tiny differences between the variances and Sharpe ratios will be significant. Consequently, we do not report the results of these tests.

of standard deviations) and smaller Sharpe ratios than when the portfolios are obtained by any of the other three specifications of the conditional covariances considered.

Second, we observe that the differences between portfolios based on different covariance matrices are larger as the number of assets in the portfolio, N , is larger. The standard deviations and Sharpe ratios of the portfolios constructed using the unconditional, DCC-NLS and WMSV specifications are rather similar when $N = 100$. However, if $N = 500$ or 1000 , the standard deviations of the unconditional portfolios are much larger than those of the portfolios computed using either the DCC-NLS or the WMSV covariance matrices. The conclusions are similar with respect to Sharpe Ratios. If $N = 100$, the Sharpe Ratios of the portfolios constructed using the unconditional, DCC-NLS and WMSV covariance matrices are very similar. However, if $N = 500$ or 1000 , the Sharpe Ratios of the portfolios constructed using the unconditional covariances are smaller than those of the DCC-NLS and WMSV portfolios.

Third, if $N = 500$ or 1000 , taking into account transaction costs is important to compare the performance of DCC-NLS and WMSV portfolios. If there are no transaction costs, i.e. $c = 0$, and $N = 100$, the performance of DCC-NLS and WMSV portfolios, both in terms of risk and risk-adjusted returns, is similar. If $N = 500$ or 1000 , the portfolios based on DCC-NLS covariances have slightly smaller standard deviations than those based on WMSV covariances while the former have slightly larger Sharpe ratios. On the other hand, if transaction costs are considered, i.e. $c = 10$, the conclusions are the same with respect to standard deviations. However, the picture is very different when looking at the Sharpe ratios. Given that DCC-NLS based portfolios have large turnovers, their Sharpe ratios are smaller than those of portfolios based on WMSV covariances. According to the modern portfolio theory, portfolio re-balancing occurs in response to changes in the correlations among asset returns. In other words, when the correlation among assets change, so does the optimal portfolio composition. In this sense, higher levels of portfolio turnovers can be a consequence of frequent and/or abrupt changes in the correlations implied by an underlying covariance model. We observe in Figure 1 that the correlation implied by the WMSV model evolve in a smoother way in comparison to those obtained with DCC-NLS and RM-2006 covariance models. This helps understanding why the WMSV model leads to optimal portfolios that demand

less re-balancing, which attenuates the impact of trading costs and leads to higher after-fee risk-adjusted returns measured by the Sharpe ratio. The large turnovers of DCC-NLS based portfolios are related to the variability of the correlations, forcing \hat{w}_t to vary through time implying larger transaction costs.

In order to provide a visual inspection of this particular result, we plot in Figure 2 the out-of-sample monthly turnovers of the minimum variance portfolios with $N = 1000$ assets obtained with the WMSV and the DCC-NLS specifications. We observe that the turnovers associated to the WMSV specification are consistently much lower than those obtained with the DCC-NLS covariance matrix throughout the whole out-of-sample period.

One remarkable aspect of the results reported in Table 1 is that the turnover of the minimum variance portfolios are substantially different among alternative covariance specifications. We find that the WMSV specification consistently outperforms all competitors as it delivers minimum variance portfolios with a much lower turnover. For instance, the results in Panel C shows that the turnover of portfolios with $N = 1000$ assets is 0.26, whereas the same figure for the DCC-NLS specification is 2.15. The RM-2006 specification achieved the worst results in terms of turnover (7.30).

Taken together, the results reported in Table 1 point to the WMSV specification outperforms competing specifications specially in terms of risk-adjusted performance net of transaction costs. The turnovers of the WMSV portfolios are clearly smaller and, consequently, the Sharpe ratios are larger than those of DCC-NLS portfolios. For instance, the results in Panel C indicate that when $N = 1000$, the SR obtained with the WMSV model in the presence of transactions costs is 0.56 and this figure is substantially smaller in comparison to all other specifications considered.

Summarizing, in concordance with the results reported by Engle, Ledoit and Wolf (2019), we can conclude that portfolios based on the DCC-NLS specification of the conditional covariance matrices achieve the lowest standard deviation of returns in all cases considered. However, we observe that the risk-adjusted returns, measured by the Sharpe ratio, reveal that the WMSV specification outperformed all competitors and that the differences in performance are more pronounced as we move to portfolios with higher dimensions and take into account the presence of transaction costs.

The documented outperformance net of transaction costs returns of the portfolios obtained with the WMSV model vis-a-vis those obtained with the alternative covariance specifications is intimately related to the lower level of turnover achieved with the WMSV specification, which helps avoiding an excessive deterioration of portfolio performance due to the presence of transaction costs.

Looking at the robustness checks reported in the Supplementary Material accompanying this paper, related to the implementation of alternative covariance specifications and additional portfolio policies, we show that the results are reassuring. We observe that even though the portfolios obtained with the volatility timing policy delivered lower turnover in comparison to those obtained with the WMSV model, the former performed worse in terms of risk and risk-adjusted returns. Furthermore, the results point to the outperformance of the optimal portfolios obtained with WMSV in terms of risk-adjusted returns net of transaction costs when additional covariance specifications and portfolio policies are taken into account. We also observe that the DCC-NLS specification delivers less risky mean-variance portfolios in all instances, therefore corroborating the previous results for the minimum variance portfolios reported in Table 1 as well as the results reported in Engle, Ledoit and Wolf (2019). It is worth noting, however, that the turnover of the mean-variance portfolios are, in the vast majority of the cases, higher with respect to those obtained with the minimum variance portfolios. This is due to the fact that the mean-variance problem is known to be very sensitive to estimation of the mean returns (e.g. Jagannathan and Ma, 2003). Very often, the estimation error in the mean returns degrades the overall portfolio performance and introduces an undesirable level of portfolio turnover. In fact, existing evidence suggests that the performance of optimal portfolios that do not rely on estimated mean returns is better; see DeMiguel, Garlappi and Uppal (2009). As expected, the results reported in the Supplementary Material reveal that the risk-adjusted performance of mean-variance portfolios is, in fact, substantially affected by the presence of transaction costs. When transaction costs are taken into account, the SR of the mean-variance portfolios obtained with the DCC-NLS model decreases from 0.54 to 0.42. The best performance in terms of SR is obtained with the WMSV model (0.50). Additionally, the WMSV model is the only specification able to generate portfolios with higher Sharpe ratio with respect to the equally-weighted and value-weighted portfolios both in the absence and in the presence of

transaction costs. Finally, when comparing among alternative unconditional estimators, we observe that the LS outperforms the unconditional, and the NLS outperforms both. This result suggests that the nonlinear shrinkage developed in Ledoit and Wolf (2012, 2019) is in fact an improvement with respect to the linear shrinkage as well as with respect to the traditional sample covariance estimator. A similar finding is observed when comparing among alternative DCC specifications. We observe that the DCC-LS outperforms the DCC-Sample, and the DCC-NLS outperforms both.

5. Concluding remarks

Modeling and forecasting high dimensional conditional covariance matrices in a data-rich environment is challenging. Most models for dynamic covariance matrices suffer from the curse of dimensionality, which creates difficulties in the estimation process when considering applications involving hundreds or thousands of time series. We compare the one-step-ahead correlations obtained from the DCC WMSV, RM-2006 and unconditional covariance models in an empirical application based on daily returns of NYSE, NASDAQ and AMEX stocks, with up to 1000 assets. We show that the pairwise correlations obtained using the WMSV model are more stable over time and have less cross-sectional dispersion than those obtained by any of the other three specifications considered. We evaluate the performance of the correlations using them to select minimum variance portfolios. The WMSV correlations deliver more stable optimal portfolios weights, implying a lower turnover in comparison to the alternative conditional covariance specifications considered. We find that the risk-adjusted performance of the WMSV model is consistently superior to that of alternative specifications, mainly when considering trading costs.

6. Acknowledgements

Guilherme V. Moura is supported by the Brazilian Government through grants number 424942-2016-0 (CNPQ) and 302865-2016-0 (CNPQ). André A.P. Santos is supported by the Brazilian Government through grants number 303688-2016-5 (CNPQ) and 420038-2018-3 (CNPQ). Esther Ruiz is supported by the Spanish Government through grant number ECO2015-70331-C2-2-R (MINECO/FEDER).

Tables and Figures

Table 1: Performance of minimum variance portfolios

The Table reports performance statistics for minimum variance portfolios with $N \in \{100, 500, 1000\}$ assets obtained with a set of covariance models. Sharpe ratios are computed using returns net of transaction costs of 0 and 10 b.p. All figures are reported in monthly terms and are based on out-of-sample observations. The out-of-sample period goes from 12/12/1974 to 12/31/2016 (10,605 daily observations) resulting in a total of 505 months. Portfolio weights are updated on a monthly basis.

	No transaction costs				Transaction costs = 10 b.p.		
	Mean ret. (%)	Std. dev. (%)	Sharpe ratio	Turnover	Mean ret. (%)	Std. dev. (%)	Sharpe ratio
<i>Panel A: N=100 assets</i>							
Unconditional	1.09	3.43	0.32	0.63	1.02	3.43	0.30
RM-2006	0.82	3.70	0.22	2.80	0.54	3.72	0.14
DCC-NLS	0.99	3.27	0.30	2.23	0.77	3.28	0.23
WMSV	0.99	3.29	0.30	0.75	0.91	3.29	0.28
<i>Panel B: N=500 assets</i>							
Unconditional	1.04	2.69	0.39	1.78	0.86	2.70	0.32
RM-2006	0.98	3.13	0.31	3.71	0.61	3.15	0.19
DCC-NLS	1.04	2.25	0.46	2.41	0.80	2.27	0.35
WMSV	1.18	2.46	0.48	0.38	1.15	2.46	0.47
<i>Panel C: N=1000 assets</i>							
Unconditional	0.98	3.29	0.30	5.97	0.39	3.35	0.12
RM-2006	0.94	3.68	0.25	7.30	0.21	3.75	0.06
DCC-NLS	1.06	1.91	0.56	2.15	0.84	1.92	0.44
WMSV	1.28	2.26	0.57	0.26	1.26	2.26	0.56

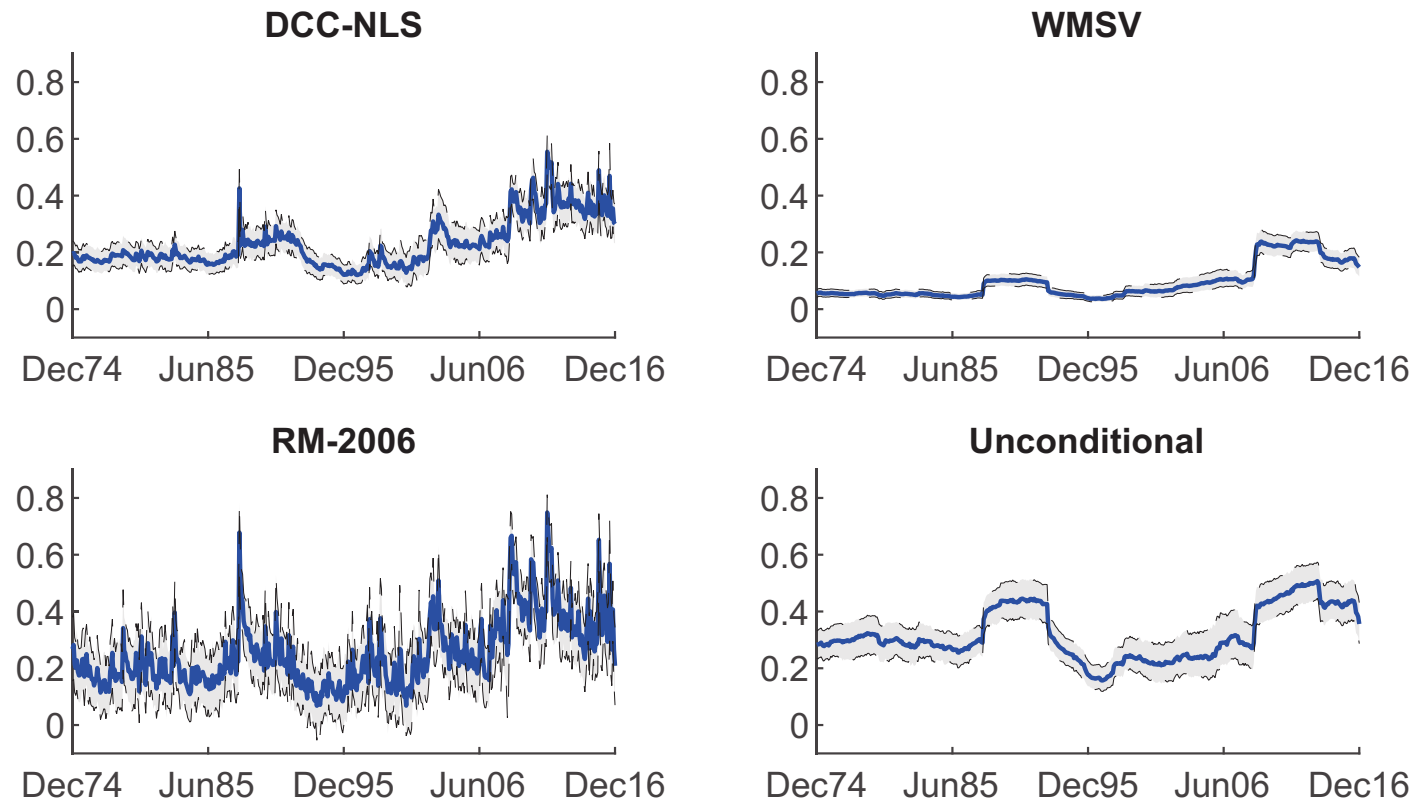


Figure 1: Median pairwise correlations

The Figure plots the evolution of the out-of-sample one-step-ahead median pairwise correlations (solid blue line) along with the 25th and 75th percentiles (dashed lines) when $N = 500$ obtained with different specifications of the conditional covariance matrices.

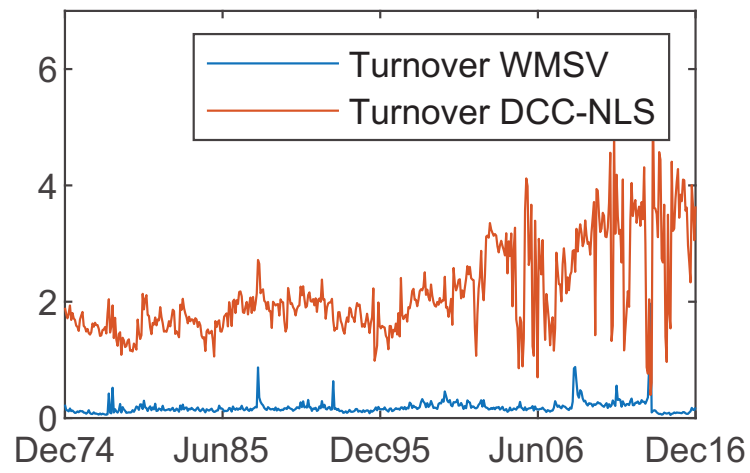


Figure 2: Monthly portfolio turnover

The Figure plots out-of-sample monthly turnover of the minimum variance with $N = 1000$ assets obtained with the WMSV and the DCC-NLS specifications.

References

- Adams, Z., Füss, R., Glück, T., 2017. Are correlations constant? Empirical and theoretical results on popular correlation models in finance. *Journal of Banking & Finance* 84, 9–24.
- Alexander, C., 2008. *Market Risk Analysis: Practical Financial Econometrics*. volume 2. John Wiley & Sons.
- Almeida, D., Hotta, L., Ruiz, E., 2018. MGARCH models: Trade-off between feasibility and flexibility. *International Journal of Forecasting* 34, 45–63.
- Asai, M., McAleer, M., 2009. The structure of dynamic correlations in multivariate stochastic volatility models. *Journal of Econometrics* 150, 182–192.
- Asai, M., McAleer, M., Yu, J., 2006. Multivariate stochastic volatility: a review. *Econometric Reviews* 25, 145–175.
- Bauwens, L., Laurent, S., Rombouts, J., 2006. Multivariate GARCH models: a survey. *Journal of Applied Econometrics* 21, 79–109.
- Bollerslev, T., Engle, R.F., Wooldridge, J.M., 1988. A capital asset pricing model with time-varying covariances. *Journal of Political Economy* 96, 116–31.
- Boudt, K., Galanos, A., Payseur, S., Zivot, E., 2019. Multivariate GARCH models for large-scale applications: A survey, in: *Handbook of Statistics*. Elsevier, pp. 193–242.
- Brandt, M.W., Santa-Clara, P., Valkanov, R., 2009. Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns. *The Review of Financial Studies* 22, 3411–3447.
- Carnero, M.A., Peña, D., Ruiz, E., 2004. Persistence and kurtosis in GARCH and stochastic volatility models. *Journal of Financial Econometrics* 2, 319–342.
- Chib, S., Omori, Y., Asai, M., 2009. *Multivariate stochastic volatility*. Springer Verlag.
- Clarke, R.G., De Silva, H., Thorley, S., 2006. Minimum-variance portfolios in the us equity market. *The Journal of Portfolio Management* 33, 10–24.
- Clarke, R.G., De Silva, H., Thorley, S., 2011. Minimum-variance portfolio composition. *Journal of Portfolio Management* 37, 31.
- Della Corte, P., Sarno, L., Thornton, D.L., 2008. The expectation hypothesis of the term structure of very short-term rates: Statistical tests and economic value. *Journal of Financial Economics* 89, 158–174.
- DeMiguel, V., Garlappi, L., Uppal, R., 2009. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies* 22, 1915–1953.
- Engle, R., 2002. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics* 20, 339–350.
- Engle, R., 2009. *Anticipating correlations: a new paradigm for risk management*. Princeton University Press.
- Engle, R., Kelly, B., 2012. Dynamic equicorrelation. *Journal of Business & Economic Statistics* 30(2), 212–228.
- Engle, R.F., Ledoit, O., Wolf, M., 2019. Large dynamic covariance matrices. *Journal of Business & Economic Statistics* 37, 363–375.
- Franco, C., Zakoian, J.M., 2019. *GARCH models: structure, statistical inference and financial applications*. John Wiley & Sons.
- French, K.R., 2008. Presidential address: The cost of active investing. *The Journal of Finance* 63, 1537–1573.
- Golosnoy, V., Gribisch, B., Liesenfeld, R., 2012. The conditional autoregressive Wishart model for multivariate stock market volatility. *Journal of Econometrics* 167, 211–223.
- Gouriéroux, C., Jasiak, J., Sufana, R., 2009. The Wishart autoregressive process of multivariate stochastic volatility. *Journal of Econometrics* 150, 167–181.
- Han, Y., 2006. Asset allocation with a high dimensional latent factor stochastic volatility model. *Review of Financial Studies* 19, 237–271.
- Harvey, A., Ruiz, E., Shephard, N., 1994. Multivariate stochastic variance models. *The Review of Economic Studies* 61, 247–264.
- Jagannathan, R., Ma, T., 2003. Risk reduction in large portfolios: Why imposing the wrong constraints helps. *The Journal of Finance* 58, 1651–1684.
- J.P.Morgan/Reuters, 1996. *Riskmetrics technical document* .
- Kastner, G., 2019. Sparse Bayesian time-varying covariance estimation in many dimensions. *Journal of Econometrics* 210, 98–115.
- Kim, D., 2014. Maximum likelihood estimation for vector autoregressions with multivariate stochastic volatility. *Economics Letters* 123, 282–286.
- Kirby, C., Ostdiek, B., 2012. It’s all in the timing: simple active portfolio strategies that outperform naive diversification. *Journal of Financial and Quantitative Analysis* 47, 437–467.

- Laurent, S., Rombouts, J.V., Violante, F., 2012. On the forecasting accuracy of multivariate GARCH models. *Journal of Applied Econometrics* 27, 934–955.
- Laurent, S., Rombouts, J.V., Violante, F., 2013. On loss functions and ranking forecasting performances of multivariate volatility models. *Journal of Econometrics* 173, 1–10.
- Ledoit, O., Wolf, M., 2004a. Honey, I shrunk the sample covariance matrix. *Journal of Portfolio Management* 30, 110–119.
- Ledoit, O., Wolf, M., 2004b. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis* 88, 365–411.
- Ledoit, O., Wolf, M., 2008. Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance* 15, 850–859.
- Ledoit, O., Wolf, M., 2012. Nonlinear shrinkage estimation of large-dimensional covariance matrices. *The Annals of Statistics* 40, 1024–1060.
- Ledoit, O., Wolf, M., 2017a. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. *The Review of Financial Studies* 30, 4349–4388.
- Ledoit, O., Wolf, M., 2017b. Numerical implementation of the QuEST function. *Computational Statistics & Data Analysis* 115, 199–223.
- Ledoit, O., Wolf, M., 2019. Analytical nonlinear shrinkage of large-dimensional covariance matrices. Forthcoming, *Annals of Statistics* .
- Mina, J., Xiao, J.Y., 2001. Return to RiskMetrics: the evolution of a standard. *RiskMetrics Group* 1, 1–11.
- Moura, G.V., Noriller, M.R., 2019. Maximum likelihood estimation of a TVP-VAR. *Economics Letters* 174, 78–83.
- Pakel, C., Shephard, N., Sheppard, K., Engle, R., 2019. Fitting vast dimensional time-varying covariance models. Forthcoming, *Journal of Business & Economic Statistics* .
- Philipov, A., Glickman, M.E., 2006a. Factor multivariate stochastic volatility via Wishart processes. *Econometric Reviews* 25, 311–334.
- Philipov, A., Glickman, M.E., 2006b. Multivariate stochastic volatility via Wishart processes. *Journal of Business & Economic Statistics* 24, 313–328.
- Politis, D., Romano, J., 1994. The stationary bootstrap. *Journal of the American Statistical Association* 89, 1303–1313.
- Silvennoinen, A., Teräsvirta, T., 2009. Multivariate GARCH models, in: *Handbook of Financial Time Series*. Springer, pp. 201–229.
- Thornton, D.L., Valente, G., 2012. Out-of-sample predictions of bond excess returns and forward rates: An asset allocation perspective. *Review of Financial Studies* 25, 3141–3168.
- Uhlig, H., 1994. On singular Wishart and singular multivariate Beta distributions. *The Annals of Statistics* 22, 395–405.
- Uhlig, H., 1997. Bayesian vector autoregressions with stochastic volatility. *Econometrica* 65, 59–74.
- Yu, J., Meyer, R., 2006. Multivariate stochastic volatility models: Bayesian estimation and model comparison. *Econometric Reviews* 25, 361–384.
- Zumbach, G., 2007a. A gentle introduction to the RM2006 methodology. *RiskMetrics Group* .
- Zumbach, G., 2007b. The RiskMetrics 2006 methodology. Technical report, *RiskMetrics Group* .