Journal of Banking and Finance 118 (2020) 105882

Contents lists available at ScienceDirect

Journal of Banking and Finance

journal homepage: www.elsevier.com/locate/jbf

Comparing high-dimensional conditional covariance matrices: Implications for portfolio selection

Guilherme V. Moura[a], André A.P. Santos[b], Esther Ruiz[c]

[a] Department of Economics Universidade Federal de Santa Catarina, Brazil
[b] Big Data Institute Universidad Carlos III de Madrid and Department of Economics Universidade Federal de Santa Catarina, Brazil
[c] Department of Statistics Universidad Carlos III de Madrid, Spain

ARTICLE INFO

Article history:
Received 28 November 2019
Accepted 4 June 2020
Available online 10 June 2020

JEL classification:
C53
G17

Keywords:
GARCH
Minimum-variance portfolio
Mean-variance portfolio
Risk-adjusted returns
Stochastic volatility
Turnover-constrained portfolios

ABSTRACT

Portfolio selection based on high-dimensional covariance matrices is a key challenge in data-rich environments with the curse of dimensionality severely affecting most of the available covariance models. We challenge several multivariate Dynamic Conditional Correlation (DCC)-type and Stochastic Volatility (SV)-type models to obtain minimum-variance and mean-variance portfolios with up to 1000 assets. We conclude that, in a realistic context in which transaction costs are taken into account, although DCC-type models lead to portfolios with lower variance, modeling the covariance matrices as latent Wishart processes with a shrinkage towards the diagonal covariance matrix delivers more stable optimal portfolios with lower turnover and higher information ratios. Our results reconcile previous findings in the portfolio selection literature as those claiming for equicorrelations, a smooth dynamic evolution of correlations or correlations close to zero.

1. Introduction

The covariation among financial returns is a fundamental ingredient in many procedures to obtain optimal portfolio weights. However, the number of off-diagonal elements of the conditional covariance matrices increases exponentially with the number of assets and, consequently, modelling the covariation among financial returns becomes challenging when the number of assets involved is large, say in the order of hundreds or even thousands. In many cases, the cross-sectional dimension is similar to the temporal dimension and, consequently, simple estimators of the covariance matrices are poorly conditioned with some small eigenvalues. Alternatively, more sofisticated 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 large financial systems.

Some strategies for portfolio selection overcome the curse of dimensionality by avoiding the computation of covariances. The first of these strategies is the naive equally-weighted (EW) portfolio, which does not require any optimization strategy and/or estimation of parameters. According to DeMiguel et al. (2009), when dealing with portfolios with $N \in (3, 50)$ monthly assets and a estimation window of $T = 120$ months, the EW portfolio is difficult to beat by alternative mean-variance portfolios, according to the Sharpe ratio, the certainty-equivalent return or the turnover; see also Hjalmarsson and Manchev (2012) for the outperformance of EW portfolios. Second, Kirby and Ostdiek (2012) propose a volatility timing (VT) strategy which implies using a diagonal covariance specification with the portfolio weights of each asset being proportional to the inverse of the variances. They show that, in a mean-variance portfolio, assuming zero correlations and taking into account only changes in volatilities generate portfolios with low turnover that outperform naive diversification even in the presence of high transaction costs.

Alternatively, a large number of portfolio selection strategies take into account the covariation between assets and, consequently, require estimation of large conditional covariance matrices. Engle et al. (2019) show that, in a large-scale portfolio selection problem with up to 1000 assets, a robustified Dynamic Conditional Correlation (DCC) model generate portfolios with lower variances with respect to those obtained with a number of competitors. The robustified DCC model is based on the DCC specifica-

E-mail address: ortega@est-econ.uc3m.es (E. Ruiz).

https://doi.org/10.1016/j.jbankfin.2020.105882

tion of the dynamic evolution of the conditional covariance matrices with correlation targeting. The estimation of the unconditional correlation matrix is carried out using the non-linear shrinkage (NLS) approach of Ledoit and Wolf (2012) and the estimation of the dynamic parameters uses the composite likelihood method of Pakel et al. (2020). Note that the positivity of conditional covariance matrices can be easily guaranteed in the context of the DCC model. Very recently, De Nard et al. (2020) injects factor structure into the estimation of time-varying, large-dimensional covariance matrices of stock returns by modelling the idiosyncratic noises via the robustified DCC model of Engle et al. (2019). They show that the inclusion of a factor structure yields more efficient portfolios with smaller turnover.

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 et al. (1994). One attractive specification of multivariate Stochastic Volatility models, originally proposed by Uhlig (1994, 1997), is based on treating unobserved dynamic precision matrices as Wishart processes. The Wishart stochastic volatility (WSV) specification guarantees positive definiteness of conditional covariance matrices. Furthermore, this specification 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 easily estimated by Maximum Likelihood (ML); see Kim (2014) and Moura and Noriller (2019).[1] The WSV specification models the evolution of the covariance matrices as a dynamic exponential weighted moving average (EWMA) as in the popular estimator implemented by RiskMetrics; see J.P.Morgan/Reuters (1996), Mina and Xiao (2001), Zumbach (2007a,b) and Alexander (2008). The main difference between RiskMetrics and WSV covariance matrices is that the discounting parameter of the former is fixed while, in the latter, it is estimated and depends on the portfolio dimension. Given that the estimated discounting parameter of WSV is closer to one than that of the RiskMetrics, the initial covariance matrix plays a role when computing the conditional covariances using WSV. If this initial covariance is diagonal and N is large, then the estimated conditional correlations are shrunk towards zero. In this case, the estimated covariance matrices are close to be diagonal as proposed by Kirby and Ostdiek (2012).

The main contribution of this paper is to compare empirically the EW and VT portfolios with minimum-variance and mean-variance portfolios based on conditional covariance matrices obtained with the robustified DCC model of Engle et al. (2019), the factor model of De Nard et al. (2020) and the WSV model of Uhlig (1994, 1997). Additionaly, conditional covariance matrices are also computed using the popular RiskMetrics' approach and the standard sample unconditional covariance estimator based on a rolling window scheme. We evaluate the performance of the different conditional covariance models in delivering minimum-variance and mean-variance portfolios. In order to take into account in an explicit way the impact of transaction costs during the portfolio formation process, we follow DeMiguel et al. (2020) and also implement and evaluate the performance of alternative estimators of the covariances when selecting turnover-constrained versions of the minimum-variance and mean-variance portfolios. As in Engle et al. (2019) and De Nard et al. (2020), 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 portfolios re-balanced on a monthly basis. Similar as in DeMiguel et al. (2009) and Kirby and Ostdiek (2012), the out-of-sample evaluation and comparison of the portfolios is carried out not only by comparing their standard deviations but also their Sharpe ratios. The impact of the presence of transaction costs is considered by evaluating information ratios (IRs). We show that, for large dimensions, the correlations estimated by the WSV model initialized with a diagonal covariance matrix, are smaller, smoother and have less dispersion than those estimated by any of the other specifications of the conditional covariance matrices considered. As a result, the portfolios selected using the WSV specification with a diagonal initial conditional covariance matrix have smaller turnover and, consequently, larger IRs on an after-fee basis. We also show that, in concordance with the results in Engle et al. (2019), optimal portfolios based on conditional covariances obtained with the robustified DCC model outperform all competitors in terms of standard deviations of portfolio returns when $N = 100$ or 500. However, optimal portfolios based on conditional covariances obtained with the approximate factor DCC model have smaller standard deviations when $N = 1000$; see De Nard et al. (2020). However, these portfolios have a larger turnover and, consequently, even when mild transaction costs are taken into account, they have a lower IRs in comparison to those obtained with the WSV model. These results are potentially relevant for portfolio managers to chose the most adequate strategy for portfolio selection depending on their objective and portfolio dimension.

The rest of the paper is organized as follows. Section 2 describes the alternative specifications considered to forecast conditional covariances while Section 3 describes the portfolio policies considered as well as the measures for portfolio performance. The main contributions of this paper appear in Sections 4 and 5 in which we estimate the conditional correlations of large systems of returns using the alternative models considered and compare the performance of different portfolios constructed using the estimated correlations, respectively. Finally, Section 6 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

$$r_t = H_t^{1/2} \varepsilon_t \tag{1}$$

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. In this section, we briefly describe the alternative specifications to forecast covariance matrices of large systems of returns considered in this paper.

2.1. Constant conditional covariance matrices

If covariance matrices were constant over time, $H_t = H$ is the unconditional covariance matrix of asset returns, which can be estimated by the sample covariance matrix of returns. Alternative estimators of the unconditional covariance matrix are considered in the Supplementary Material.

2.2. Riskmetrics

One of the most popular specifications of the conditional covariance matrix, H_t, in equation (1), prominent in the industry and among market participants, is based on the RiskMetrics 1994 (hereafter RM-1994) methodology; see J.P.Morgan/Reuters (1996), Mina and Xiao (2001) and Alexander (2008). According to RM-1994, one-step-ahead conditional covariance matrices are obtained

[1] The Wishart specification has been recently used by Hautsch and Voigt (2019) to model large realized covariance matrices.

as an exponential weighted moving average (EWMA) of quadratic forms of past returns as follows

$$H_t = (1-\lambda)\sum_{i=1}^{t-1}\lambda^{i-1}r_{t-i}r'_{t-i}. \tag{2}$$

where $\lambda = 0.94$ for daily data. Note that, given that $\lambda = 0.94$, the weight placed in older observations is decreasing exponentially.[2] Zumbach (2007a,b) extend RM-1994 to the RM-2006 approach by proposing a (pseudo) long-memory model for the covariance matrices in which the weights of past quadratic forms of returns decay hyperbolically rather than exponentially. According to RM-2006, the conditional covariance matrices are obtained as a weighted sum of EWMAs rather than a single EWMA, as follows

$$H_t = \sum_{i=1}^{14}\omega_i\Sigma_{it-1}, \tag{3}$$

where the weights are given by $\omega_i = \frac{1}{c}\left[1-\frac{\ln(\tau_i)}{7.35}\right]$ with c being a normalization constant that ensures that $\sum_{i=1}^{14}\omega_i = 1$ and $\tau_i = 4\times\sqrt{2}^{i-1}$. The EWMAs in equation (3) are given by

$$\Sigma_{it} = \lambda_i\Sigma_{it-1} + (1-\lambda_i)r_t r'_t, \tag{4}$$

where $\lambda_i = \exp\frac{-1}{\tau_i}$.[3]

It is important to note that, if there are more observations than assets, the RiskMetrics covariance matrices are at least positive semi-definite. However, as the number of assets grows, the ratio of the largest to the smallest eigenvalues of the covariance matrices deteriorates.

2.3. DCC Model

We consider the DCC specification proposed by Engle et al. (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 follows

$$H_t = D_t\Psi_t D_t, \tag{5}$$

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 $h^2_{i,t}$ 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 returns, which is governed by the following correlation targeting dynamics

$$\Psi_t = (1-\alpha-\beta)C + \alpha s_{t-1}s'_{t-1} + \beta\Psi_{t-1}, \tag{6}$$

where $s_t = (r_{1,t}/h_{1,t},\ldots,r_{N,t}/h_{N,t})'$ and α and β are scalar parameters guiding the dynamics of all correlations and C is the unconditional covariance matrix of 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 devolatize 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 devolatized returns, s_t. It is known, however, that the standard sample covariance estimator is prone to estimation error. To circumvent this problem, Engle et al. (2019) propose estimating C by using the non-linear shrinkage (NLS) approach of Ledoit and Wolf (2012), denoted by $\hat{C}$. Although Engle et al. (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 (2020) which is much faster and has similar accuracy. Note that, in spite of the fact that devolatized 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.[4]

Finally, in the third step, once the unconditional covariance matrix, C, is estimated, the parameters α and β of the correlation-targeting dynamics in (6) are estimated by the composite likelihood method of Pakel et al. (2020). The log-likelihood is computed by summing up the log-likelihood of all contiguous pairs of assets. Therefore, only $N-1$ bivariate log-likelihoods should be computed.[5]

2.4. Wishart stochastic covariances

Given that, in equation (1), the conditional means of returns are assumed to zero, we follow Windle and Carvalho (2014) and adopt only the WSV part of the model proposed by Uhlig (1994) and Uhlig (1997). The WSV model specifies a multiplicative law of motion for the stochastic precision matrix, H_t^{-1}, which is driven by a singular multivariate Beta distribution shock as follows

$$H_t^{-1} = \frac{d+1}{d}\mathcal{U}(H_{t-1}^{-1})'\Theta_t\mathcal{U}(H_{t-1}^{-1}), \tag{7}$$

$$H_1^{-1} \sim \mathcal{W}_N(d,[dS_0]^{-1}), \tag{8}$$

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. Finally, $\mathcal{W}_N$ denotes the N-dimensional Wishart distribution and $S_0^{-1} = E[H_1^{-1}]$.

Uhlig (1997) shows that, in the context of the model in (7), the nonlinear filtering of the latent precision matrices can be computed analytically. More specifically, Uhlig (1994) extends the study of Wishart and multivariate beta distributions to the singular case, which allows Uhlig (1997) to exploit the conjugacy between the multivariate Normal, the Wishart, and the singular multivariate beta distributions to show that one-step-ahead prediction densities and filtered densities have analytical expressions.[6] In particular, the predictive density of the precision matrix is given by

$$p(H_{t+1}^{-1}|r_t) \sim \mathcal{W}_N(d,[dS_{t+1}]^{-1}), \tag{9}$$

where S_{t+1} evolves acording to

$$S_{t+1} = \frac{d}{d+1}S_t + \frac{1}{d+1}r_t r'_t. \tag{10}$$

[2] The 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 multivariate stochastic volatility model of Harvey et al. (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 et al. (1994), the specification is for log-variances while in the RiskMetrics methodology variances are modelled directly.

[3] We use the Matlab routine riskmetrics2006 available in the MFE Toolbox provided by Kevin Sheppard; see Sheppard (2012) for details.

[4] Alternative estimators of the matrix C are considered in the Supplementary Material.

[5] In 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.

[6] See the Supplementary material in the online appendix for details.

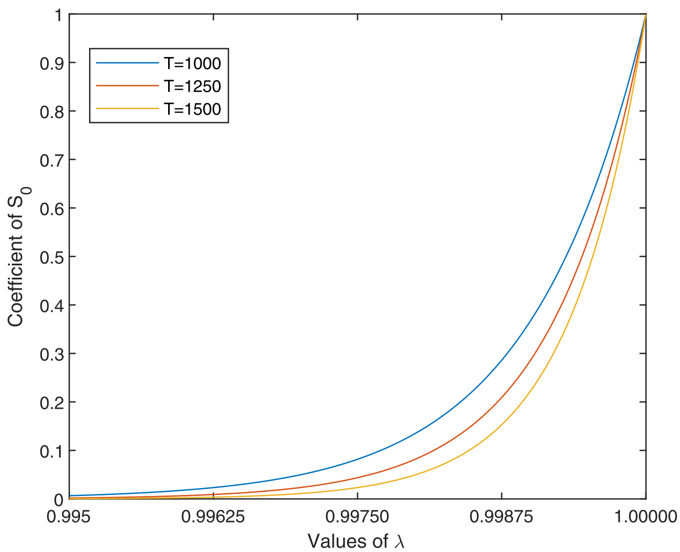

Fig. 1. Coefficient of the initial covariance matrix S_0. The Figure plots the weight placed in the initial covariance matrix, S_0, as a function of the decay parameter, λ, for sample sizes of $T \in \{1000, 1250, 1500\}$ in the computation of the one-step-ahead conditional covariance matrix, H_{T+1}. .

It is worth noting that the dynamics of the precision matrix is governed by a unique parameter, d, that can be estimated by Maximum Likelihood (ML); see Kim (2014). We refer the reader to Section of the Supplementary Material that brings additional details regarding filtering and ML estimation of the WSV model considered in this paper.

Substituting backwards in equation (10), it is possible to obtain the following expression for S_t

$$S_t = \lambda^t S_0 + (1-\lambda)\sum_{i=1}^{t-1}\lambda^{i-1} r_{t-i} r'_{t-i}, \tag{11}$$

where $\lambda = \frac{d}{d+1} < 1$. In expression (11), the evolution of S_t is the same as that implied by the RM-1994 covariance matrices in (2). The main difference between (2) and (11) is that in the former, the discount parameter is fixed at $\lambda = 0.94$ while in the latter, λ is estimated. Given that $\lambda = \frac{d}{d+1}$ and $d > N+1$, if N is large, $\lambda \approx 1$. Consequently, (11) implies that the estimated conditional covariance matrices are shrunk towards the initial covariance matrix, S_0. The shrinkage of H_{T+1} towards S_0 is very sensitive to the parameter λ. Fig. 1, which plots λ^T for different values of T, shows that if $\lambda < 0.99$ and $T > 1000$, the one-step-ahead forecasts of the conditional variances do not depend on the initial covariance matrices.[7] However, if $\lambda > 0.998$, the forecasts are shrunk towards the initial covariance matrix. S_0 plays a larger role in the one-step-ahead forecast of the conditional covariance matrix when λ is very close to one. Fig. 1 also shows that small increments in λ have a substantial impact on the influence of S_0 on the one-step-ahead forecast of the conditional covariance matrix.

Given that, for values of $\lambda \approx 1$, the one-step-ahead forecasts of the conditional covariance matrices are shrunk towards S_0, it is crucial to chose an adequate initial covariance matrix. We consider two alternative initial matrices S_0. First, S_0 is given by an equicorrelation covariance matrix in which, as in Engle and Kelly (2012), the correlation between any two returns is equal to the average sample correlation between all returns in the portfolio. Second, following Uhlig (1997) and Kim (2014), we set S_0 to a diagonal matrix whose diagonal elements are given by the in-sample variance of each return. This second specification, denoted as Shrunk WSV (SWSV), implies that the one-step-ahead forecasts of the correlation are shrunk towards zero. Note that, even though, in practice, the covariances among asset returns are often different from zero, the error incurred in estimating those quantities can lead to noisy estimates specially when the cross-section dimension is large; see Kirby and Ostdiek (2012), Stivers and Sun (2016), and Santos (2019) for a discussion.

2.5. Approximate factor models

Very recently, De Nard et al. (2020) propose estimating conditional covariance matrices of large dimensions by blending an approximate factor model (AFM) with time-varying conditional heteroscedastic idiosyncratic noises and show empirically that this combination yields portfolios with minimum-variance among those considered by them. Assuming that there is just one unique common factor, De Nard et al. (2020) propose the following AFM-DCC-NLS model to estimate conditional covariance matrices

$$r_t = A + Bf_t + u_t, \tag{12}$$

where $A = (\alpha_1, \ldots, \alpha_N)$ and $B = (\beta_1, \ldots, \beta_N)$ are $N \times 1$ vectors of constants and factor loadings, respectively, f_t is the market factor assumed to have zero mean and variance σ_f^2 and u_t is an $N \times 1$ conditionally Gaussian white noise vector with covariance matrix Σ_{ut}[8] The conditional covariance matrix of r_t is given by:

$$H_t = \sigma_f^2 BB' + \Sigma_{ut}. \tag{13}$$

Estimation of the AFM-DCC-NLS model is performed in two steps. In the first step, after estimating by OLS the parameters in A and B in (12), the residuals, $\hat{u}_t$, are obtained as usual. In the second step, $\hat{u}_t$ are used to estimate Σ_{ut}, the time-varying conditional covariance matrix of the residuals, using the DCC-NLS described above.

3. Large-scale portfolios

Define the return of a portfolio at time t as $R_t = w'_t r_t$ with $w_t = (w_{1,t}, \ldots, w_{N,t})'$ being the portfolio weights at time t obtained at time $t-1$. In this section, we describe the portfolio selection policies considered in this paper and the criteria for evaluating their performance.

3.1. Equally-weighted and volatility timing portfolios

We start by considering two very simple portfolio strategies. First, we consider equal-weigted (EW) portfolios with the weight of each asset being equal to $w_{it} = \frac{1}{N}, \forall i, t$.

The second portfolio considered, proposed by Kirby and Ostdiek (2012), is the volatility timing (VT) (or inverse variance) portfolio. In this portfolio, the weight of each asset is proportional to the inverse of its variance, as follows

$$w_{it} = \frac{(1/\sigma_i^2)}{\sum_{i=1}^N (1/\sigma_i^2)}, \forall t, \tag{14}$$

where σ_i^2 is the variance of the i-th asset. The VT portfolio is equivalent to the solution of the minimum-variance portfolio described below obtained when all off-diagonal covariance elements of H_t are equal to zero. Two interesting aspects of the VT portfolio are that i) it does not require any covariance matrix inversion,

[7] In the empirical application following later in this paper, $T = 1250$.

[8] We consider the AFM model with only one factor (the market factor), since the results in De Nard et al. (2020) suggest that additional factors do not improve the performance of portfolios. The market factor is defined as the return of the market portfolio in excess of the risk-free rate and is obtained from Kenneth French's data library.

and ii) it does not generate negative weights. In practice σ_i^2 is estimated by the sample variance of the i-th asset.

3.2. Minimum-variance portfolio

The third portfolio selection 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 et al. (2006, 2011) are extensive practitioner-oriented studies on the composition and performance of minimum-variance portfolios; see also Engle and Kelly (2012) who evaluate whether equicorrelation is better than different correlations using minimum-variance portfolios and Kastner (2019) who compares alternative covariance matrices when $N = 300$ in terms of minimum-variance portfolios. The minimum-variance portfolio problem is defined as follows

$$\begin{aligned}&\min_{w_t} w_t' H_t w_t\\ &\text{subject to } w_t'\iota = 1,\end{aligned} \tag{15}$$

where ι is an appropriately sized vector of ones. The solution to (15) is given by

$$w_t = \frac{\iota' {H_t}^{-1}}{\iota' {H_t}^{-1}\iota}. \tag{16}$$

In practice, feasible portfolio weights, $\hat{w}_t$, are obtained by replacing in equation (16), the unknown covariance matrix, H_t by an estimate, $\hat{H}_t$, which is obtained at time $t-1$ using each of the specifications described above.

3.3. Mean-variance with a momentum signal portfolio

We also consider the mean-variance with a momentum signal portfolio proposed by Engle et al. (2019), which is based on an investor who wishes to minimize portfolio risk subjected to a target portfolio return. This portfolio optimization problem is defined as follows

$$\begin{aligned}&\min_{w_t} w_t' H_t w_t\\ &\text{subject to} \begin{cases} w_t' m = b \\ w_t'\iota = 1,\end{cases}\end{aligned} \tag{17}$$

where m is the signal variable and b is the target return. The solution to (17) is given by

$$w_t = {H_t}^{-1}\frac{m(Cb - D) + \iota(E - Db)}{EC - D^2} \tag{18}$$

where $C = \iota' {H_t}^{-1}\iota$, $D = m{H_t}^{-1}\iota$ and $E = m{H_t}^{-1}m$. In practice, a large number of variables can be used to construct the signal. We follow Engle et al. (2019) and use the well-known momentum factor of Jegadeesh and Titman (1993). The momentum signal of a given stock is computed as the geometric average of the previous 252 returns on that stock but excluding the most recent 21 returns, that is, the geometric average over the previous year but excluding the previous month. Collecting the individual momentums of all the N stocks contained in the portfolio universe yields the return-predictive signal m. The target return b is computed as the arithmetic average of the momentums of the stocks belonging to the top-quintile stocks according to momentum. Finally, in practice, the unknown covariance matrix H_t should be replaced by an estimated $\hat{H}_t$.

3.4. Turnover-constrained portfolios

Finally, we consider an alternative formulation of the minimum-variance and mean-variance portfolio policies in which transaction costs are taken into account explicitly during the portfolio formation process as proposed by DeMiguel et al. (2020). The idea behind the turnover-constrained portfolios is to solve the investor's portfolio problem while simultaneously taking into account the impact of transaction costs by adding a penalization term given by the portfolio turnover on the portfolio's objective function.

The turnover-constrained minimum-variance and mean-variance problems are defined as in (15) and (17), respectively, with the objective function modified as follows:

$$\min_{w_t} w_t' H_t w_t + \kappa ||w_t - w_t^*||_1, \tag{19}$$

where $||a||_1 = \sum_{i=1}^N |a|$ is the 1-norm of the N-dimensional vector a and w_t^* is the portfolio obtained at time $t-1$ and after taking into account the changes in asset prices between periods $t-1$ and t. The constant κ controls for the degree to which the portfolio turnover is penalized when solving (19). We set $\kappa = 1 \times 10^{-3}$.

One important aspect of the turnover-constrained portfolios is that they require numerical solutions. For that purpose, we use the CVX software (CVX Research, 2012).

3.5. Evaluation of portfolio performance

The evaluation of the portfolios' performance is based on the one-step-ahead portfolio returns, R_t. We first compute their overtime mean and standard deviation, $\bar{R}^P$ and σ^P, respectively. A portfolio is prefered when its variance is as small as possible.

Following, Gasbarro et al. (2007), DeMiguel and Nogales (2009), Zakamouline and Koekebakker (2009), Behr et al. (2013) and Hautsch and Voigt (2019), among many others, we also evaluate the portfolios by the risk-adjusted returns measured by the information ratio (IR), which is defined as follows[9]

$$IR = \frac{\bar{R}^P}{\sigma^P}. \tag{20}$$

A superior covariance forecasting model should provide portfolios with low variance and/or large IRs.

Additionally, many authors point out the importance of taking into account the impact of transaction costs on the performance of optimal portfolios; see, for example, Han (2006) and Hautsch and Voigt (2019). Consequently, we also evaluate the performance of the portfolios by computing the mean, standard deviation and IR of the returns net of transaction costs. Following Della Corte et al. (2008), Kirby and Ostdiek (2012), and Thornton and Valente (2012), we compute the portfolio return net of transaction costs as follows

$$R_t^P = (1 - c \cdot turnover_t)(1 + R_t) - 1, \tag{21}$$

where $turnover_t = \sum_{j=1}^N \left(\left|w_{j,t} - w_{j,t}^*\right|\right)$ is the portfolio turnover at time t, defined as the fraction of wealth traded between periods $t-1$ and t with w_t^* being the allocation vector at period $t-1$ after taking into account the changes in asset prices between periods $t-1$ and t. Finally, c is the fee that must be paid for each transaction that is measured in terms of basis points (b.p.). Note that, in practice, transaction costs depend on (possible time-varying) institutional rules and the liquidity supply in the market. It is unavoidable that transaction costs are underestimated or overestimated in individual assets. 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

[9] Note that in some of these works, they use the Sharpe ratio instead of the IR; see, for example, Goodwin (1998) and Israelsen (2005) for the definition of the Sharpe ratio and IR.

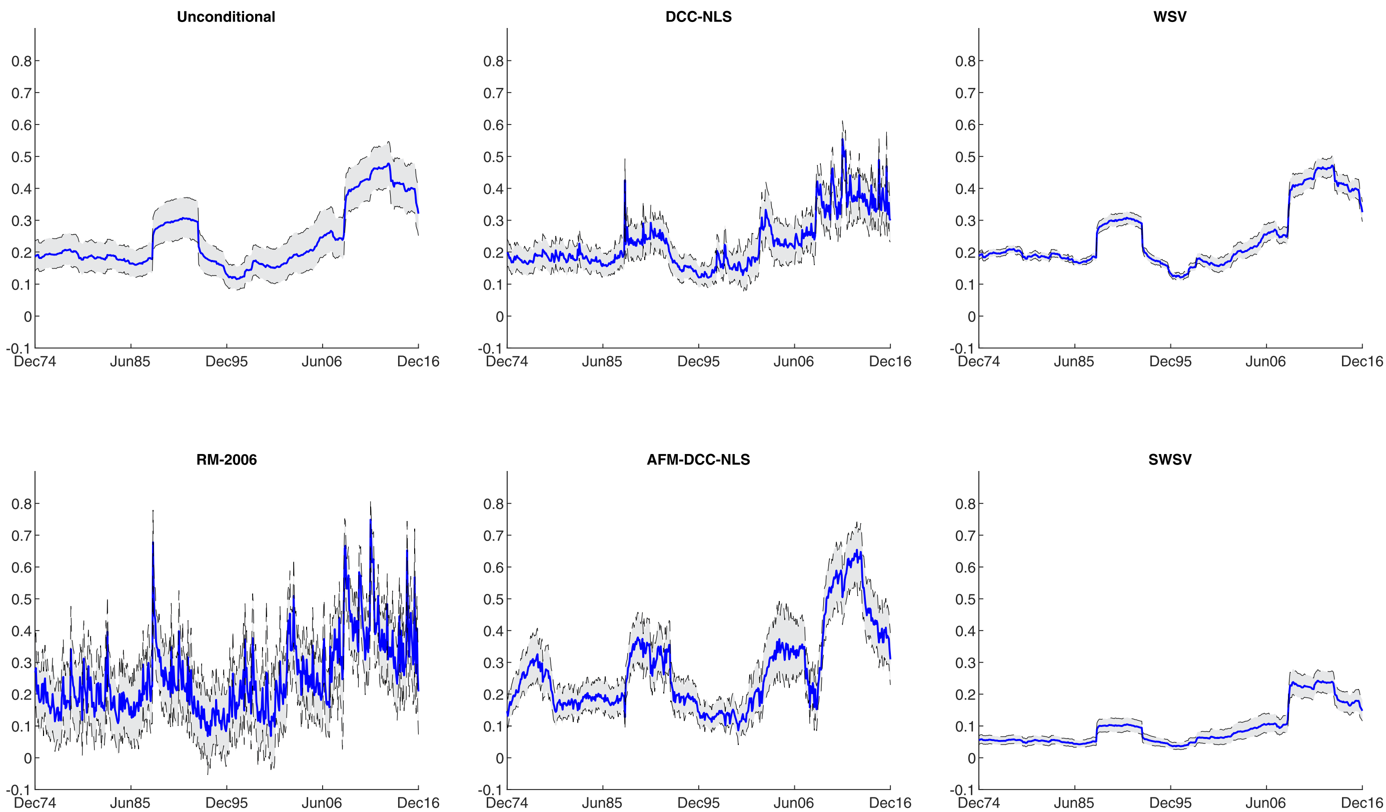

Fig. 2. Estimated pairwise correlations. The Figure plots the evolution of the out-of-sample one-step-ahead median pairwise correlations (solid blue line) along with the 25th and 75th percentiles (dashed lines) when $N = 500$ obtained with different specifications of the conditional covariance matrices.

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." Hautsch and Voigt (2019) mention that scenarios with $c < 100$ can be associated with rather small transaction costs; see also Kirby and Ostdiek (2012) who consider $c = 50$ b.p. To be conservative, in order to take into account the impact of proportional transaction costs, we consider the cases in which $c \in \{0, 5, 10\}$ b.p.

Finally, for each portfolio universe, N, and level of transaction costs, c, we test for the statistical significance of the difference between the standard deviation (IR) of each of the portfolios considered and the corresponding standard deviation of the DCC-NLS portfolio, by using the two-sided p-value of the prewhitened HAC_{PW} test described by Ledoit and Wolf (2011) and Ledoit and Wolf (2008) for the variance and the IR, respectively.

4. Empirical estimation of large conditional covariance matrices

We fit the conditional covariance models described in Section 2 to a large system with up to 1000 assets traded in the US stock market. The data set consists of returns (including dividends) of all NYSE, AMEX and NASDAQ stocks observed daily from 01/01/1970 to 12/31/2016. It is important to note that, although the empirical results may depend on the particular assets analyzed as well as on their observation frequency, given the very large number of individual returns observed daily considered, we can expect the results to be rather general and of interest to portfolio managers; see also Engle et al. (2019) and De Nard et al. (2020) who analyze the same data set observed over different spans of time.

The covariance models (unconditional, RM-2006, DCC-NLS, AFM-DCC-NLS, WSV and SWSV) are recursively estimated every month (we adopt the common convention that 21 consecutive days constitute a month), at investment dates $h = 1, \ldots, 505$, 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. Following Engle et al. (2019) and De Nard et al. (2020), 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 et al. (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. 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.

Fig. 2 plots the time series evolution of the median and 25th and 75th percentiles of the one-step-ahead pairwise estimated correlations when $N = 500$ for each of the six specifications of the conditional covariance matrices considered.[10] The first conclusion from Fig. 2 is that the median level of the correlations estimated

[10] The results obtained with $N = 100$ and $N = 1000$ are qualitatively similar and are available in the Supplementary Material.

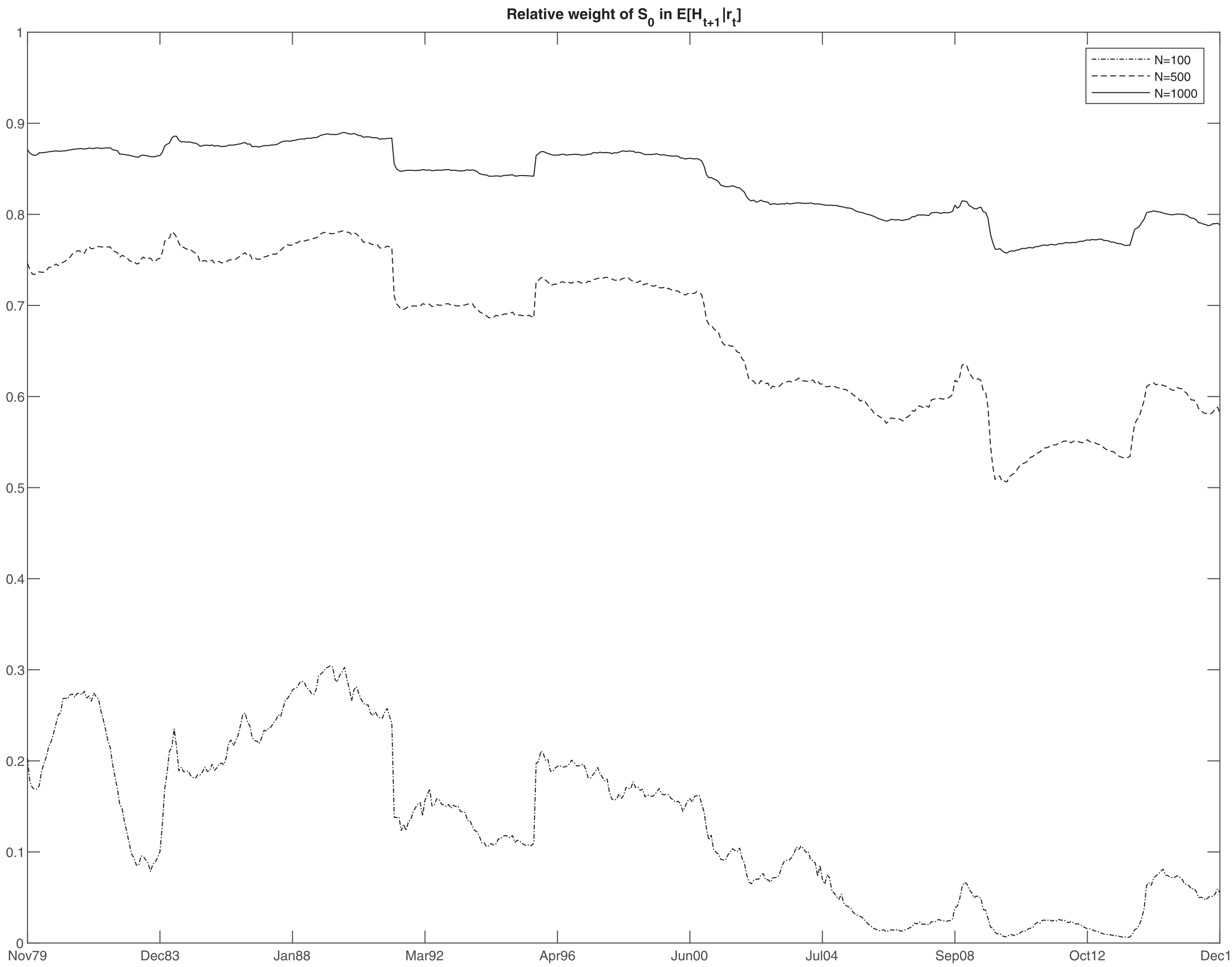

Fig. 3. Relative weight of S_0 The Figure plots the evolution of the relative weight of S_0 in the one-step-ahead forecast of the SWSV model.

by the SWSV model is clearly lower than those obtained when the correlations are estimated by any of the other alternative specifications considered. Note that the median level of the SWSV correlations is around 0.1 while the level is around 0.3 for any of the other specifications. This result is expected if we take into account that the initial covariance matrix of the SWSV model is a diagonal matrix and look at the evolution of the weights of S_0 plotted in Fig. 3. These weights, although decreasing over time, move around 0.7. Therefore, as mentioned in Section 2, the SWSV correlations are strongly shrunk towards zero and the weight of the most recent information on cross-products of returns is relatively low. As a consequence, the SWSV correlations are not only lower than for any other estimator of the covariances but also rather smooth and with very low cross-sectional dispersion.

Fig. 2 also shows that the RM-2006, DCC-NLS and AFM-DCC-NLS models yield correlations that are highly time-varying and noisy; see Adams et al. (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 unconcditional and SWSV and WSV models evolve in a much more smooth way. Adams et al. (2017) argue that financial correlations are mostly constant over time with financial shocks leading to breaks that shift the level of correlations. The relatiosnhips between two different assets seem to not change drastically in a short period of time. Furthermore, it is well established in the related literature that, in time of crises, the degree of co-movement between asset returns changes, partly as a result of generally increased uncertainty. In Fig. 2, we can observe that the correlations estimated by the unconditional and WSV models and, in particular, by the SWSV model jump to a higher level in October 1987 after the Black Monday when all markets fall together, generating highly correlated negative returns. These correlations keep being larger during 5 years, until October 1992. The second jump corresponds to October 2008 that coincides with the start of the Global Financial crisis. In 2016 the correlations are still relatively high without recovering the levels previous to the crisis. Note that the median correlations estimated using the unconditional and WSV models are very similar as a result of the WSV correlations being shrunk towards the unconditional correlations.

Finally, the 25th and 75th percentiles of the pairwise correlations plotted in Fig. 2 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 WSV and SWSV specifications are 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. The dispersion of the pairwise estimated correlations is larger for the correlations estimated using the unconditional and the RM-2006 covariance matrices.

According to the results in this section, the SWSV model generates one-step-ahead correlations that are close to the equicorrelation assumption and also close to zero with a smooth time variation, only jumping at particular moments of time. The WMS correlations are also close to equicorrelation and smooth although larger than those of SWSV. The correlations estimated by the unconditional variance are also smooth but the dispersion is much larger. RM-2006 pairwise correlations not only have a large cross-sectional dispersion but also are very variable through time. Finally, the dispersion and temporal variability of DCC-based correlation is in between those of the unconditional and RM-2006 specifications.

In the next section, we evaluate the performance of these estimated pairwise correlations in an economically meaningful way by using them to construct minimum-variance and mean-variance portfolios. We will show that the differences in the estimated pairwise conditional correlations have important implications for the performance of the optimal portfolios.

5. Empirical portfolio selection and evaluation

In this section, we perform a large-scale portfolio selection exercise. Our approach to portfolio construction is largely inspired by Engle et al. (2019) and De Nard et al. (2020) with portfolio weights updated on a monthly basis in order to avoid excessive turnover levels associated to daily re-balancing.[11] For each investment universe with $N \in \{100, 500, 1000\}$ assets, EW and VT portfolios are constructed. Also, for each N, one-step-ahead forecasts of the conditional covariance matrices are obtained by each of the covariance models considered in Section 4. Then, minimum-variance and mean-variance with a momentum signal portfolios are constructed both with and without turnover constrains.

5.1. Minimum-variance portfolios

Table 1 reports the average turnover and the annualized average, standard deviation and IR of returns net of transaction costs of the EW, VT and minimum-variance portfolios. The returns have been computed over the out-of-sample period under the three scenarios of proportional transaction costs, namely, $c = 0, 5$ and 10 b.p. 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. Consider first the results for the average turnover. Obviously, regardless of N, the average turnover is very similar for EW and VT portfolios, around 0.10, and clearly smaller than for the minimum-variance portfolios constructed with any of the estimators of the covariance matrices considered. 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 SWSV specification consistently outperforms all competitors as it delivers minimum-variance portfolios with a much lower turnover. On the other hand, the maximum average turnover is always achieved in portfolios based on the RM-2006 covariances. With respect to portfolios based on unconditional covariances, we can observe that they have larger turnovers as N increases. The turnover of portfolios based on AFM-DCC-NLS and DCC-NLS is in between those of the unconditional and RM-2006 and those of the SWSV and WSV portfolios. Furtermore, note that the differences among average turnovers are large. For instance, the results in Panel C show that the turnover of SWSV portfolios with $N = 1000$ assets is 0.26, whereas the same figure for the AFM-DCC-NLS specification is 1.70. The RM-2006 specification achieved the worst results in terms of turnover (7.30). It seems that the turnover is smaller as the temporal variability and cross-sectional dispersion of the estimated conditional cross-correlations are smaller.

We move now to evaluate and compare the minimum-variance portfolios in terms of their average standard deviations. Table 1 reports the results of testing whether the difference between the standard deviation of each portfolio and that of the portfolio constructed using the DCC-NLS covariance matrix is significant. For relatively small porfolios, when $N = 100$, the standard deviation is minimum when the portfolio is constructed using the covariance matrices estimated using DCC-NLS regardless of whether there are transaction costs while, in concordance with results in De Nard et al. (2020), if the portfolio is large ($N = 1000$) the standard deviation is minimized by the AFM-DCC-NLS portfolios. Note that the differences between the standard deviations of the DCC-NLS and AFM-DCC-NLS portfolios are not significantly different when the universe investment is relatively small while they are significantly different when the universe is large. For medium-size portfolios, the standard deviation is minimized by DCC-NLS when there are not transaction costs and by AFM-DCC-NLS when transaction costs are taken into account although the differences between the standard deviations of these portfolios are not significantly different. We can also observe that, regardless of the size of the investment universe, N and the presence of transaction costs, the EW and VT portfolios as well as those based on the unconditional, RM-2006 and Wishart-based covariance matrices always have standard deviations that are significantly different and larger than that of the corresponding DCC-NLS portfolio. Summarizing, if one wants to choose a minimum-variance portfolio with minimum standard deviation, the conditional covariances should be estimated using the AFM-DCC-NLS model if $N = 1000$ or the DCC-NLS when $N = 100$.

As mentioned before, the portfolios are also evaluated in terms of their IRs. The results reported in Table 1 show that the IR for the portfolios constructed using the SWSV estimates of the covariances consistently display the highest IRs, regardless of N and the level of transaction costs. There is just one exception, in cases where transaction costs are absent and $N = 1000$, the IR of the AFM-DCC-NLS is the largest. Note that, in this case, the difference between the IR of the AFM-DCC-NLS and that of the SWSV portfolios is very small. It is also important to note that the differences between the IR of the SWSV and that of the DCC-NLS portfolios are significantly different from zero when transaction costs are taken into account. As expected, if $c = 0$, these differences are not significative.

In summary, in the presence of transaction costs, the portfolio IR is clearly maximized when the covariances are estimated using the SWSV model. This is the result of the SWSV portfolio having simultaneously smaller turnover and not very large standard deviations. Given that DCC-NLS based portfolios have large turnovers, their information ratios are smaller than those of portfolios based on SWSV 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 Fig. 2 that the correlation implied by the SWSV model evolve in a smoother way in comparison to those obtained with DCC-NLS covariance models. Furthermore, the cross-sectional dispersion of the pair-wise correlations is also much smaller. This helps under-

[11] De Nard et al. (2020) also propose reducing the turnover by using "averaged forecasting". When the frequency of the observed returns is daily but the portfolio is held for a month, the covariance matrices are averaged over one-step-ahead forecasts over the 21 periods. However, this procedure is not trivial to implement and we do not pursue it further.

Table 1
Performance of minimum-variance portfolios. The Table reports performance statistics for equally weighted (EW), variance targeting (VT) and minimum-variance portfolios with $N \in \{100, 500, 1000\}$ assets obtained with several covariance models. Information ratios (IR) are computed using returns net of transaction costs of 0, 5, and 10 b.p. Mean returns, standard deviation, and IR are reported in annual terms whereas turnovers are in monthly terms. All figures 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. One, two, and three asterisks denote that the standard deviation (IR) is statistically different with respect the standard deviation (IR) of the benchmark DCC-NLS model at the 10%, 5%, and 1% levels, respectively. The smallest (largest) standard deviation (IR) are highlighted in **bold**.

		No transaction costs			Transaction costs = 5 b.p.			Transaction costs = 10 b.p.		
	Turnover	Mean (%)	Std. dev. (%)	IR	Mean (%)	Std. dev. (%)	IR	Mean (%)	Std. dev. (%)	IR
Panel A: N=100 assets										
EW	0.11	18.66	17.27***	1.08	18.59	17.27***	1.08	18.52	17.26***	1.07*
VT	0.11	17.48	15.43***	1.13	17.42	15.43***	1.13	17.35	15.43***	1.12**
Unconditional	0.63	13.05	11.88**	1.10	12.67	11.88**	1.07	12.30	11.88**	1.04**
RM-2006	2.80	9.80	12.82***	0.76***	8.13	12.84***	0.63***	6.45	12.87***	0.50***
DCC-NLS	2.08	12.12	**11.33**	1.07	10.88	**11.34**	0.96	9.63	**11.36**	0.85
AFM-DCC-NLS	1.30	12.67	11.51	1.10	11.90	11.52	1.03	11.12	11.52	0.96
WSV	0.66	11.79	11.47	1.03	11.39	11.47	0.99	10.99	11.47	0.96
SWSV	0.25	14.63	11.77**	**1.24****	14.48	11.77**	**1.23*****	14.33	11.77**	**1.22*****
Panel B: N=500 assets										
EW	0.10	22.48	16.60***	1.35	22.42	16.60***	1.35	22.36	16.60***	1.35
VT	0.09	20.01	14.30***	1.40	19.96	14.30***	1.40	19.91	14.30***	1.39
Unconditional	1.78	12.46	9.32***	1.34*	11.39	9.33***	1.22	10.33	9.34***	1.11
RM-2006	3.71	11.71	10.84***	1.08***	9.49	10.86***	0.87***	7.26	10.92***	0.67***
DCC-NLS	2.31	12.28	**7.80**	1.57	10.89	**7.81**	1.39	9.51	7.85	1.21
AFM-DCC-NLS	1.73	12.13	7.81	1.55	11.10	**7.81**	1.42	10.06	**7.83**	1.28
WSV	0.54	8.03	9.09***	0.88***	7.70	9.09***	0.85***	7.38	9.09***	0.81***
SWSV	0.38	14.22	8.53***	**1.67**	13.99	8.53***	**1.64****	13.77	8.53***	**1.61*****
Panel C: N=1000 assets										
EW	0.10	25.64	16.63***	1.54**	25.58	16.63***	1.54***	25.52	16.63***	1.53***
VT	0.08	22.04	14.01***	1.57**	21.99	14.01***	1.57	21.94	14.01***	1.57
Unconditional	5.97	11.78	11.41***	1.03***	8.20	11.46***	0.72***	4.63	11.60***	0.40***
RM-2006	7.30	11.25	12.75***	0.88***	6.88	12.80***	0.54***	2.51	12.98***	0.19***
DCC-NLS	2.10	12.48	6.61	1.89	11.22	6.63	1.69	9.97	6.66	1.50
AFM-DCC-NLS	1.70	12.70	**6.38*****	**1.99**	11.68	**6.38*****	1.83	10.66	**6.41*****	1.66
WSV	0.43	6.77	8.12***	0.83***	6.51	8.12***	0.80***	6.25	8.12***	0.77***
SWSV	0.26	15.39	7.83***	1.97	15.23	7.83***	**1.95*****	15.08	7.83***	**1.93*****

standing why the SWSV 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 IR. The large turnovers of DCC-NLS based portfolios are related to the variability of the correlations, forcing w_t to vary through time implying larger transaction costs. In order to provide a visual inspection of this particular result, we plot in Fig. 4 the out-of-sample monthly turnovers of the minimum-variance portfolios with $N = 1000$ assets obtained with the SWSV and the DCC-NLS specifications. We observe that the turnovers associated to the SWSV specification are consistently much lower than those obtained with the DCC-NLS covariance matrix throughout the whole out-of-sample period.

Taken together, the results reported in Table 1 show that the SWSV specification outperforms competing specifications specially in terms of risk-adjusted performance net of transaction costs. The turnovers of the SWSV portfolios are clearly smaller while simultaneously the standard deviations are not very large, and, consequently, the IRs are significantly larger than those of minimum-variance portfolios constructed with alternative models for conditional covariance matrices. For instance, the results in Panel C indicate that when $N = 1000$, the IR obtained with the SWSV model in the presence of transactions costs of 10 b.p. is 1.93 and this figure is substantially higher in comparison to all other specifications considered.

Summarizing, in concordance with the results reported by De Nard et al. (2020), we can conclude that portfolios based on the AFM-DCC-NLS specification of the conditional covariance matrices achieve the lowest standard deviation of returns when N is large. When $N = 100$, the minimum standard deviation is obtained by the minimum-variance portfolio obtained using the DCC-NLS covariances. However, we observe that the risk-adjusted re-

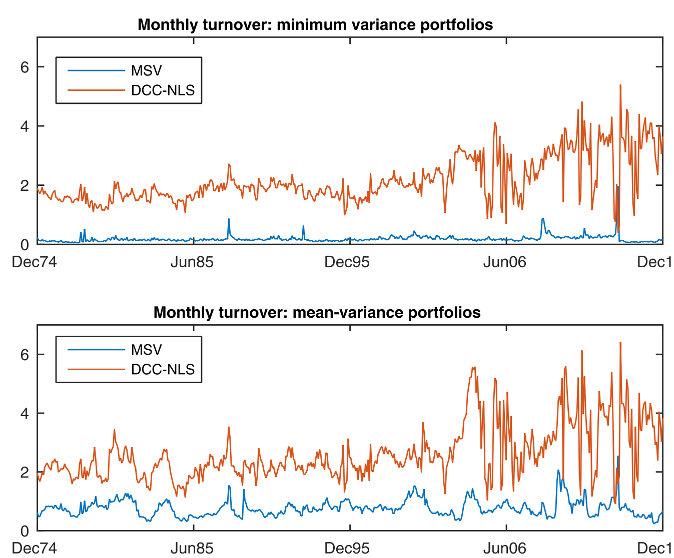

Fig. 4. Monthly portfolio turnover. The Figure plots out-of-sample monthly turnover of the minimum-variance (top panel) and mean-variance (bottom panel) portfolios with $N = 1000$ assets obtained with the WMS^2V (blue lines) and the DCC-NLS (red lines) estimated covariance matrices. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

turns, measured by the IR, reveal that the SWSV 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

Table 2
Performance of turnover-constrained minimum-variance portfolios. The Table reports performance statistics for turnover-constrained minimum-variance portfolios with $N \in \{100, 500, 1000\}$ assets obtained with several covariance models. Information ratios (IR) are computed using returns net of transaction costs of 0, 5, and 10 b.p. Mean returns, standard deviation, and IR are reported in annual terms whereas turnovers are in monthly terms. All figures 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. One, two, and three asterisks denote that the standard deviation (IR) is statistically different with respect the standard deviation (IR) of the benchmark DCC-NLS model at the 10%, 5%, and 1% levels, respectively. The smallest (largest) standard deviation (IR) are highlighted in **bold**.

		No transaction costs			Transaction costs = 5 b.p.			Transaction costs = 10 b.p.		
	Turnover	Mean (%)	Std. dev. (%)	IR	Mean (%)	Std. dev. (%)	IR	Mean (%)	Std. dev. (%)	IR
Panel A: N=100 assets										
Unconditional	0.58	13.05	11.88***	1.10	12.70	11.88***	1.07	12.35	11.88***	1.04**
RM-2006	2.63	9.78	12.79***	0.76***	8.20	12.80***	0.64***	6.63	12.83***	0.52***
DCC-NLS	2.02	12.11	**11.32**	1.07	10.90	**11.33**	0.96	9.68	**11.35**	0.85
AFM-DCC-NLS	1.24	12.69	11.51	1.10	11.95	11.51	1.04	11.20	11.52	0.97
WSV	0.62	11.79	11.47	1.03	11.42	11.46	1.00	11.05	11.47	0.96
SWSV	0.23	14.63	11.77***	**1.24****	14.49	11.77***	**1.23*****	14.35	11.77**	**1.22*****
Panel B: N=500 assets										
Unconditional	1.39	12.45	9.30***	1.34*	11.61	9.30***	1.25	10.78	9.31***	1.16
RM-2006	2.03	11.81	10.38***	1.14***	10.60	10.39***	1.02**	9.38	10.41***	0.90**
DCC-NLS	2.12	12.29	**7.77**	1.58	11.02	**7.78**	1.42	9.74	**7.81**	1.25
AFM-DCC-NLS	1.55	12.19	7.80	1.56	11.27	7.81	1.44	10.34	7.82	1.32
WSV	0.45	8.01	9.10***	0.88***	7.74	9.10***	0.85***	7.47	9.10***	0.82***
SWSV	0.31	14.18	8.55***	**1.66**	13.99	8.55***	**1.64****	13.81	8.55***	**1.62*****
Panel C: N=1000 assets										
Unconditional	2.72	11.42	10.46***	1.09***	9.79	10.47***	0.94***	8.17	10.51***	0.78***
RM-2006	1.32	11.44	9.66***	1.19***	10.65	9.66***	1.10***	9.86	9.66***	1.02***
DCC-NLS	1.81	12.46	6.56	1.90	11.38	6.57	1.73**	10.30	6.60	1.56
AFM-DCC-NLS	1.42	12.77	**6.36****	**2.01**	11.92	**6.37****	1.87	11.07	**6.38****	1.73*
WSV	0.33	6.74	8.15***	0.83***	6.54	8.15***	0.80***	6.35	8.15***	0.78***
SWSV	0.19	15.33	7.87***	1.95	15.22	7.87***	**1.93****	15.11	7.87***	**1.92*****

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

As mentioned above, we also construct minimum-variance portfolios with turnover restrictions. The corresponding results are reported in Table 2. We can observe that, with the level of penalization $\kappa = 1 \times 10^{-3}$, the turnover-constrained minimum-variance portfolio display, as expected, lower turnovers. The reduction in turnover is larger the larger N. If $N = 1000$, we observe, on average across covariance models, a reduction of 56% in the portfolio turnover in comparison to the turnover-unconstrained portfolios. It is remarkable the large reduction in the turnovers of the portfolios constructed using RM-2006, which is 7.3 in the unconstrained portfolio and 1.32 in the constrained one. Similarly, the standard deviations of the constrained portfolios are smaller than those of the corresponding unsconstrained ones. The reduction in the standard deviation is usually mild for all portfolios. Finally, when looking at IRs, they are slightly larger. The only exception is RM-2006, for which, when $N = 1000$, we observe a large increase with its IR moving from 0.88 to 1.19 when the portfolio is restricted. However, note that these changes in the measures of portfolio performance do not change the main conclusions obtained when the portfolio turnovers were unconstrained.

5.2. Mean-variance portfolios with momentum signal

The performance results for the mean-variance portfolios with momentum signal are reported in Table 3. We can observe that both the turnovers and the standard deviations of the mean-variance portfolios are larger for all specifications of the conditional covariance matrices considered. The larger turnovers of the mean-variance portfolios could be due to the fact that the mean-variance problem is known to be very sensitive to estimation of the mean returns; see 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 et al. (2009). As expected, the results reported in Table 3 reveal that the risk-adjusted performance of mean-variance portfolios is, in fact, substantially affected by the presence of transaction costs. However, the main conclusions for the minimum-variance portfolios are corroborated. The DCC-NLS and AFM-DCC-NLS specifications deliver less risky mean-variance portfolios when N is large, therefore corroborating the previous results for the minimum-variance portfolios reported in Table 1 as well as the results reported in De Nard et al. (2020). However, the best performance in terms of IR when transaction costs are considered is obtained with the SWSV model (1.73 when $N = 1000$ and transaction cost is 10 b.p.).

The results of the turnover-constrained mean-variance portfolios are reported in Table 4. We can observe that, with the level of penalization $\kappa = 1 \times 10^{-3}$, the turnover-constrained mean-variance portfolios display, in general, the same numbers as those reported in Table 3 for the corresponding unconstrained portfolios.

5.3. Robustness checks

In this subsection, we carry out several robustness checks to validate the main conclusions. First, we consider alternative estimators of the covariance matrices. In particular, on top of estimating the unconditional covariance matrices using the sample covariance, we also estimate it by the linear shrinkage (Unconditional-LS) method of Ledoit and Wolf (2004b) and by the analytical non-linear shrinkage (Unconditional-NLS) method of Ledoit and Wolf (2020). We also consider 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 devolatized residuals, denoted by DCC-Sample, and ii) the estimator of C considered in Engle et al. (2019) in which C is estimated by the linear shrinkage (LS) approach of Ledoit and Wolf (2004b), denoted as DCC-LS. Finally, we also consider the fac-

Table 3
Performance of mean-variance portfolios with momentum signal. The Table reports performance statistics for mean-variance portfolios with momentum signal with $N \in \{100, 500, 1000\}$ assets obtained with a set of covariance models. Information ratios (IR) are computed using returns net of transaction costs of 0, 5, and 10 b.p. Mean returns, standard deviation, and IR are reported in annual terms whereas turnovers are in monthly terms. All figures 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. One, two, and three asterisks denote that the standard deviation (IR) is statistically different with respect the standard deviation (IR) of the benchmark DCC-NLS model at the 10%, 5%, and 1% levels, respectively. The smallest (largest) standard deviation (IR) are highlighted in **bold**.

		No transaction costs			Transaction costs = 5 b.p.			Transaction costs = 10 b.p.		
	Turnover	Mean (%)	Std. dev. (%)	IR	Mean (%)	Std. dev. (%)	IR	Mean (%)	Std. dev. (%)	IR
Panel A: N=100 assets										
Unconditional	1.21	14.73	14.18***	1.04	14.01	14.18***	0.99	13.29	14.18**	0.94
RM-2006	3.76	14.11	15.76***	0.90*	11.86	15.77***	0.75**	9.60	15.82***	0.61**
DCC-NLS	2.49	14.90	**13.72**	1.09	13.41	**13.73**	0.98	11.92	**13.75**	0.87
AFM-DCC-NLS	1.78	15.12	13.94	1.09	14.06	13.94	1.01	12.99	13.95	0.93
WSV	1.32	14.04	13.77	1.02	13.25	13.78	0.96	12.46	13.78	0.90
SWSV	0.83	16.61	14.51***	**1.14**	16.12	14.51***	**1.11***	15.62	14.51***	**1.08*****
Panel B: N=500 assets										
Unconditional	2.34	13.84	10.61***	1.30**	12.44	10.62***	1.17**	11.04	10.64***	1.04*
RM-2006	5.14	15.69	12.97***	1.21**	12.60	12.99***	0.97***	9.52	13.07***	0.73***
DCC-NLS	2.66	14.32	**8.96**	1.60	12.72	**8.98**	1.42	11.13	**9.02**	1.23
AFM-DCC-NLS	2.19	13.93	9.03	1.54	12.62	9.04	1.40	11.30	9.06	1.25
WSV	1.13	10.52	10.30***	1.02***	9.85	10.30***	0.96***	9.17	10.31***	0.89***
SWSV	0.95	16.39	10.17***	**1.61**	15.82	10.17***	**1.56**	15.25	10.17***	**1.50*****
Panel C: N=1000 assets										
Unconditional	7.06	10.45	13.12***	0.80***	6.22	13.19***	0.47***	1.99	13.37***	0.15***
RM-2006	16.26	-6.69	27.86***	-0.24***	-16.43	27.97***	-0.59***	-26.17	28.40***	-0.92***
DCC-NLS	2.48	14.12	**7.55**	1.87	12.64	7.57	1.67	11.15	7.61	1.47
AFM-DCC-NLS	2.18	14.34	7.43	**1.93**	13.04	**7.45**	1.75	11.73	**7.48**	1.57
WSV	1.03	8.91	9.25***	0.96***	8.29	9.25***	0.90***	7.68	9.26***	0.83***
SWSV	0.84	17.36	9.46***	1.83	16.85	9.46***	**1.78**	16.35	9.46***	**1.73*****

Table 4
Performance of turnover-constrained mean-variance portfolios. The Table reports performance statistics for turnover-constrained mean-variance portfolios with momentum signal with $N \in \{100, 500, 1000\}$ assets obtained with a set of covariance models. Information ratios (IR) are computed using returns net of transaction costs of 0, 5, and 10 b.p. Mean returns, standard deviation, and IR are reported in annual terms whereas turnovers are in monthly terms. All figures 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. One, two, and three asterisks denote that the standard deviation (IR) is statistically different with respect the standard deviation (IR) of the benchmark DCC-NLS model at the 10%, 5%, and 1% levels, respectively. The smallest (largest) standard deviation (IR) are highlighted in **bold**.

		No transaction costs			Transaction costs = 5 b.p.			Transaction costs = 10 b.p.		
	Turnover	Mean (%)	Std. dev. (%)	IR	Mean (%)	Std. dev. (%)	IR	Mean (%)	Std. dev. (%)	IR
Panel A: N=100 assets										
Unconditional	1.16	14.76	14.18***	1.04	14.06	14.18***	0.99	13.36	14.18**	0.94
RM-2006	3.59	14.11	15.72***	0.90	11.96	15.74***	0.76**	9.81	15.78***	0.62**
DCC-NLS	2.43	14.89	**13.71**	1.09	13.44	**13.72**	0.98	11.98	**13.74**	0.87
AFM-DCC-NLS	1.73	15.14	13.94*	1.09	14.11	13.94	1.01	13.07	13.95	0.94
WSV	1.27	14.04	13.78	1.02	13.28	13.78	0.96	12.52	13.78	0.91
SWSV	0.81	16.62	14.52***	**1.15**	16.14	14.52***	**1.11***	15.65	14.52***	**1.08*****
Panel B: N=500 assets										
Unconditional	1.96	13.92	10.60***	1.31**	12.74	10.60***	1.20**	11.57	10.62***	1.09
RM-2006	3.03	15.98	12.44***	1.28**	14.16	12.45***	1.14*	12.35	12.48***	0.99*
DCC-NLS	2.48	14.37	**8.94**	**1.61**	12.89	**8.96**	1.44	11.40	**9.00**	1.27
AFM-DCC-NLS	2.01	14.03	9.03	1.55	12.83	9.04	1.42	11.63	9.06	1.28
WSV	1.03	10.57	10.32***	1.02***	9.96	10.32***	0.96***	9.34	10.32***	0.90***
SWSV	0.87	16.44	10.18***	**1.61**	15.92	10.18***	**1.56**	15.40	10.19***	**1.51*****
Panel C: N=1000 assets										
Unconditional	3.62	10.47	12.13***	0.86***	8.30	12.16***	0.68***	6.13	12.21***	0.50***
RM-2006	2.66	6.35	14.68***	0.43***	4.76	14.69***	0.32***	3.16	14.71***	0.22***
DCC-NLS	2.20	14.19	7.52	1.89	12.88	7.53	1.71	11.56	7.57	1.53
AFM-DCC-NLS	1.90	14.47	**7.43**	**1.95**	13.33	**7.44**	**1.79**	12.19	**7.46**	1.63
WSV	0.89	8.99	9.28***	0.97***	8.46	9.28***	0.91***	7.93	9.28***	0.85***
SWSV	0.72	17.44	9.49***	1.84	17.01	9.49***	**1.79**	16.58	9.49***	**1.75****

tor specification of De Nard et al. (2020) with Σ_{ut} modelled by the Whishart model described in Section 2.4. We denote this specification as AFM-WSV when the initial covariance matrix, S_0, is a equicorrelation matrix and AFM-SWSV when S_0 is diagonal. The results for these additional specifications of the covariance matrices are reported in Section C of the Supplementary Material.

In Section C of the Supplementary Material, we also discuss the implementation of alternative portfolio policies. In particular, we consider the value-weighted policies, and an alternative version of the volatility timing policy proposed in Kirby and Ostdiek (2012).

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. When comparing among alternative unconditional estimators, we observe that the unconditional-LS outperforms the unconditional, and the unconditional-NLS outperforms both. This result suggests that the non-linear shrinkage developed in Ledoit and Wolf (2012, 2020) 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.

Introducing a common factor and modelling the covariance of the idiosyncratic errors using WSV models does not improve the results with respect to implementing these models directly to returns.

Finally, we observe that even though the portfolios obtained with the volatility timing policy delivered lower turnover in comparison to those obtained with the SWSV 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 SWSV in terms of risk-adjusted returns net of transaction costs when additional covariance specifications and portfolio policies are taken into account. Additionally, the SWSV model is the only specification able to generate portfolios with higher IR with respect to the equally-weighted and value-weighted portfolios both in the absence and in the presence of transaction costs.

6. Concluding remarks

Modelling and forecasting large 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 unconditional, RM-2006, DCC, AFM-DCC, WSV and SWSV 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 SWSV model are close to zero, more stable over time and have less cross-sectional dispersion than those obtained by any of the other specifications considered. We evaluate the performance of the correlations using them to select minimum-variance and mean-variance portfolios, as well as turnover-constrained minimum-variance and mean-variance portfolios. The SWSV 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 SWSV model is consistently superior to that of alternative specifications, specially when considering trading costs. Furthermore, one further attractive of SWSV is that it is computationally very simple even in the presence of very large portfolios.

In concordance with results in Engle et al. (2019), if the portfolio manager wants to chose the portfolio with minimum variance, then she should choose the portfolio in which the covariance matrices have been estimated using the DCC model or, if the number of assets is very large, the factor modification proposed by De Nard et al. (2020). However, if she prefers to choose the portfolio with maximum IR, then the portfolio selected using the Wishart covariance matrices is superior. It is up to her to decide which is her criteria to choose the portfolio.

In this paper, we also reconcile previous finding on portfolio optimization as those concluding that portfolios based on zero correlations can be optimal (DeMiguel, Garlappi and Uppal, 2009; Hjalmarsson and Manchev, 2012; Kirby and Ostdiek, 2012), or discussing the advantages of equicorrelation (Engle and Kelly, 2012). We also find similar results as those in Adams et al. (2017) who finds a smooth temporal evolution of pairwise correlations. These results are potentially relevant for portfolio managers to chose the most adequate portfolio selection strategy depending on their objective portfolio dimensions.

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).

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.jbankfin.2020.105882.

References

Adams, Z., Füss, R., Glück, T., 2017. Are correlations constant? Empirical and theoretical results on popular correlation models in finance. J. Bank. Finance 84, 9–24.
Alexander, C., 2008. Market Risk Analysis: Practical Financial Econometrics, 2. John Wiley & Sons.
Behr, P., Guetter, A., Miebs, F., 2013. On portfolio optimization: imposing the right constraints. J. Bank. Finance 37 (4), 1232–1242.
Brandt, M.W., Santa-Clara, P., Valkanov, R., 2009. Parametric portfolio policies: exploiting characteristics in the cross-section of equity returns. Rev. Financ. Stud. 22 (9), 3411–3447.
Carnero, M.A., Peña, D., Ruiz, E., 2004. Persistence and kurtosis in GARCH and stochastic volatility models. J. Financ. Econom. 2 (2), 319–342.
Clarke, R.G., De Silva, H., Thorley, S., 2006. Minimum-variance portfolios in the us equity market. J. Portf. Manag. 33 (1), 10–24.
Clarke, R.G., De Silva, H., Thorley, S., 2011. Minimum-variance portfolio composition. J. Portf. Manag. 37 (2), 31.
CVX Research, I., 2012. CVX: Matlab Software for Disciplined Convex Programming, Version 2.0.
De Nard, G., Ledoit, O., Wolf, M., 2020. Factor models for portfolio selection in large dimensions: The good, the better and the ugly. Journal of Financial Econometrics. Forthcoming
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. J. Financ. Econ. 89 (1), 158–174.
DeMiguel, V., Garlappi, L., Uppal, R., 2009. Optimal versus naive diversification: how inefficient is the 1/N portfolio strategy? Rev. Financ. Stud. 22 (5), 1915–1953.
DeMiguel, V., Martin-Utrera, A., Nogales, F.J., Uppal, R., 2020. A portfolio perspective on the multitude of firm characteristics. Review of Financial Studies. Forthcoming
DeMiguel, V., Nogales, J., 2009. Portfolio selection with robust estimation. Oper. Res. 57 (3), 560–577.
Engle, R., 2002. Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. J. Bus. Econ. Stat. 20 (3), 339–350.
Engle, R., Kelly, B., 2012. Dynamic equicorrelation. J. Bus. Econ. Stat. 30 (2), 212–228.
Engle, R.F., Ledoit, O., Wolf, M., 2019. Large dynamic covariance matrices. J. Bus. Econ. Stat. 37 (2), 363–375.
French, K.R., 2008. Presidential address: the cost of active investing. J. Finance 63 (4), 1537–1573.
Gasbarro, D., Wong, W., Zumwalt, J., 2007. Stochastic dominance of ishares. Eur. J. Finance 13 (1), 89–101.
Goodwin, T., 1998. The information ratio. Financ. Anal. J. 54 (4), 34–43.
Han, Y., 2006. Asset allocation with a high dimensional latent factor stochastic volatility model. Rev. Financ. Stud. 19 (1), 237–271.
Harvey, A., Ruiz, E., Shephard, N., 1994. Multivariate stochastic variance models. Rev. Econ. Stud. 61 (2), 247–264.
Hautsch, N., Voigt, S., 2019. Large-scale portfolio allocation under transaction costs and model uncertainty. J. Econom. 212 (1), 221–240.
Hjalmarsson, E., Manchev, P., 2012. Characteristic-based mean-variance portfolio choice. J. Bank. Finance 36 (5), 1392–1401.
Israelsen, C., 2005. A refinemen of the sharpe ratio and information ratio. J. Asset Manag. 5 (6), 423–427.
Jagannathan, R., Ma, T., 2003. Risk reduction in large portfolios: why imposing the wrong constraints helps. J. Finance 58 (4), 1651–1684.
Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. J Finance 48 (1), 65–91.
Morgan/Reuters, J.P., 1996. Riskmetrics Technical Document.
Kastner, G., 2019. Sparse bayesian time-varying covariance estimation in many dimensions. J. Econom. 210 (1), 98–115.
Kim, D., 2014. Maximum likelihood estimation for vector autoregressions with multivariate stochastic volatility. Econ. Lett. 123 (3), 282–286.

Kirby, C., Ostdiek, B., 2012. It'S all in the timing: simple active portfolio strategies that outperform naive diversification. J. Financ. Quant. Anal. 47 (2), 437–467.
Ledoit, O., Wolf, M., 2004. Honey, i shrunk the sample covariance matrix. J. Portf. Manag. 30 (4), 110–119.
Ledoit, O., Wolf, M., 2004. A well-conditioned estimator for large-dimensional covariance matrices. J. Multivar. Anal. 88 (2), 365–411.
Ledoit, O., Wolf, M., 2008. Robust performance hypothesis testing with the sharpe ratio. J. Empir. Finance 15 (5), 850–859.
Ledoit, O., Wolf, M., 2011. Robust performance hypothesis testing with the variance. Wilmott Mag. 55, 86–89.
Ledoit, O., Wolf, M., 2012. Nonlinear shrinkage estimation of large-dimensional covariance matrices. Ann. Stat. 40 (2), 1024–1060.
Ledoit, O., Wolf, M., 2017. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets Goldilocks. Rev. Financ. Stud. 30 (12), 4349–4388.
Ledoit, O., Wolf, M., 2017. Numerical implementation of the QuEST function. Comput. Stat. Data Anal. 115, 199–223.
Ledoit, O., Wolf, M., 2020. Analytical nonlinear shrinkage of large-dimensional covariance matrices. Annals of Statistics (264). Forthcoming
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. Econ Lett 174, 78–83.
Pakel, C., Shephard, N., Sheppard, K., Engle, R., 2020. Fitting vast dimensional time–varying covariance models. J. Bus. Econ. Stat. doi: 1080/07350015.2020.1713795.
Santos, A.A., 2019. Disentangling the role of variance and covariance information in portfolio selection problems. Quant. Finance 19 (1), 57–76.
Sheppard, K., 2012. Forecasting High Dimensional Covariance Matrices. John Wiley & Sons.
Stivers, C., Sun, L., 2016. Mitigating estimation risk in asset allocation: diagonal models versus 1/n diversification. Financ. Rev. 51 (3), 403–433.
Thornton, D.L., Valente, G., 2012. Out-of-sample predictions of bond excess returns and forward rates: an asset allocation perspective. Rev. Financ. Stud. 25 (10), 3141–3168.
Uhlig, H., 1994. On singular wishart and singular multivariate beta distributions. Ann. Stat. 22 (1), 395–405.
Uhlig, H., 1997. Bayesian vector autoregressions with stochastic volatility. Econometrica 65 (1), 59–74.
Windle, J., Carvalho, C., 2014. A tractable state-space model for symmetric positive-definite matrices. Bayesian Anal. 9 (4), 759–792.
Zakamouline, V., Koekebakker, S., 2009. Portfolio performance evaluation with generalized sharpe ratios: beyond the mean and variance. J. Bank. Finance 33 (7), 1241–1254.
Zumbach, G., 2007. A Gentle Introduction to the RM2006 Methodology. RiskMetrics Group.
Zumbach, G., 2007. The RiskMetrics 2006 Methodology. RiskMetrics Group. Technical report