UNIVERSITÀ DEGLI STUDI FIRENZE

DISIA
DIPARTIMENTO DI STATISTICA, INFORMATICA, APPLICAZIONI "GIUSEPPE PARENTI"

A Dynamic Conditional Approach to Portfolio Weights Forecasting

Fabrizio Cipollini, Giampiero M. Gallo,
Alessandro Palandri

DISIA WORKING PAPER
2020/06

© Copyright is held by the author(s).

A Dynamic Conditional Approach to Portfolio Weights Forecasting

Fabrizio Cipollini[1], Giampiero M. Gallo[2,*] and Alessandro Palandri[1]

[1]Dipartimento di Statistica, Informatica, Applicazioni (DiSIA) *G. Parenti*, Università degli Studi di Firenze

[2]Italian Court of Audits, and New York University Florence

This Version: May 12, 2020

Abstract

We build the time series of optimal realized portfolio weights from high-frequency data and we suggest a novel Dynamic Conditional Weights (DCW) model for their dynamics. DCW is benchmarked against popular *model-based* and *model-free* specifications in terms of weights forecasts and portfolio allocations. Next to portfolio variance, certainty equivalent and turnover, we introduce the break-even transaction costs as an additional measure that identifies the range of transaction costs for which one allocation is preferred to another. By comparing minimum-variance portfolios built on the components of the Dow Jones 30 Index, the proposed DCW overall attains the best allocations with respect to the measures considered, for any degree of risk-aversion, transaction costs and exposure.

Keywords: Portfolio Allocation, Realized Volatility, Realized Correlations, Dynamic Conditional Modeling, Portfolio Weights Modeling.

JEL classification: C32, C53, G11, G17.

*Address correspondence to: Giampiero M. Gallo, Corte dei conti, Sezione regionale di controllo per la Lombardia, Via Marina 5, 20121 Milan, Italy, Email: `giampiero.gallo@nyu.edu`. The views expressed in the article are those of the authors and do not involve the responsibility of the Corte dei conti.

1 Introduction

A key aspect of active portfolio management is forecasting the weights that optimize portfolio holdings with respect to some representative measure of the investor's preferences. Since Markowitz (1952), these forecasts of the optimal portfolio weights are generally derived from the forecasts of conditional moments of asset returns. The availability of realized measures from high-frequency data allows for *model-based* and *model-free* forecasting of conditional variance-covariance (var-cov) matrices and, by successive manipulation thereof, of optimal portfolio weights: see, e.g., Aït-Sahalia *et al.* (2010), Christensen *et al.* (2010), Barndorff-Nielsen *et al.* (2011), Zhang (2011) and Bibinger *et al.* (2014), among others.

The *model-based* approaches are inspired by the logic behind Multivariate-GARCH (MGARCH) models[1] with the substantial difference that information is extracted from realized measures rather than low-frequency estimates of the second moments, such as the outer product of the vector of returns (or their residuals after some filtration). Examples of these approaches are the fractionally integrated processes of Chiriac and Voev (2011), the vector autoregressions of Callot *et al.* (2017) and the specifications based on the Wishart distribution of Gourieroux *et al.* (2009), Golosnoy *et al.* (2012), Noureldin *et al.* (2012) and Jin and Maheu (2010), among others. Within this framework it is not uncommon to separately model conditional variances[2] and correlation matrices to achieve a good balance between parameter parsimony and richness in the description of the second order dynamics.

On the other hand, *model-free* approaches, also referred to as *nonparametric*, impose driftless random-walk dynamics to the conditional second moments, and thus eliminate the parameter estimation problem altogether. However, for large cross-sectional dimensions, the lag-1 realized var-cov matrices may result in extreme portfolio weights, poor portfolio performance out-of-sample (OOS) and even positive-semidefiniteness (psd). To mitigate this problem, various shrinkage approaches are available: the most direct is to impose constraints on the portfolio weights[3] as in Jagannathan and Ma (2003), El Karoui (2010), Fan *et al.* (2012) and Gandy and

[1]For a review of MGARCH models see Bauwens *et al.* (2006).

[2]Amongst the various approaches to volatility modeling that make use of realized measures are the Heterogeneous Autoregressive model (HAR) of Corsi (2009) and Corsi *et al.* (2012), the Multiplicative Error Model (MEM) of Engle (2002b) and Brownlees *et al.* (2012) and the HEAVY of Shephard and Sheppard (2010), as a particular case of the vector-MEM of Cipollini *et al.* (2013). For a survey see Andersen *et al.* (2006) and Park and Linton (2012), among others.

[3]While the target is usually defined by portfolio weights with no short-sale (positivity) constraints, other alternatives are also possible, i.e. equal weights.

Veraart (2013). Shrinkage of the realized var-cov matrix has been proposed by Fan *et al.* (2008), Fan *et al.* (2011), Ledoit and Wolf (2012), Tao *et al.* (2011), Tao *et al.* (2013), Fan *et al.* (2016) and Aït-Sahalia and Xiu (2017), to name a few. Ideas behind these approaches may also be traced back to the MGARCH literature and consist of imposing a factor structure to the returns and a sparse error var-cov matrix with blocks defined by some characteristics of the assets such as sector, industry, etc.

In this paper we introduce the Dynamic Conditional Weights (DCW), an approach which emerges when expressing the autoregressive representation of the portfolio-variance optimization problem in terms of a time-independent weighting matrix. The result is a specification in which the forecast of the conditional portfolio weights derives from a linear function of past conditional weights and past realized (hence observable) weights. When associated with suitable estimation procedures, the main advantage of DCW with respect to standard *model-based* approaches is the circumvention of the curse of dimensionality problem. With respect to the *model-free* approaches, DCW does not require the imposition of particular var-cov matrix structures nor discretionary choices about the level of shrinkage.

Focusing on the minimum-variance allocation, empirical results show that DCW outperforms *model-based* and *model-free* approaches in terms of out-of-sample portfolio variance, certainty equivalence and turnover (De Miguel *et al.*, 2009). Since transaction costs may significantly alter the outlook in the performance of the approaches, we introduce the *Break-Even Transaction Costs* as a more comprehensive measure of forecasting performance confirming the goodness of the DCW allocation in terms of minimal portfolio variance and turnover. Furthermore, since the *model-based* and *model-free* literatures have proceeded on somewhat parallel tracks,[4] a contribution of this paper is a comparison across approaches, assessing the quality of the respective forecasts and portfolio allocations.

The paper is organized as follows. Section 2 introduces the optimal portfolio allocation problem. The *model-based* and *model-free* approaches are discussed in Section 3. Section 4 introduces the direct modeling of the portfolio weights. Measures of performance, data and results are presented in Section 5. Section 6 concludes.

[4]For example, the most recent contributions to the *model-free* literature, such as Fan *et al.* (2016) and Aït-Sahalia and Xiu (2017), do not benchmark their approaches to any of the *model-based*.

2 Minimum Variance Portfolio

Following Aït-Sahalia and Xiu (2017), Fan *et al.* (2016), Behr *et al.* (2013) and Fan *et al.* (2012), Fan *et al.* (2008), among others,[5] we focus on minimizing portfolio variance, which allows for a clean evaluation of the contribution of modeling and forecasting second moments to the optimal allocation. Furthermore, the minimum-variance portfolio has often been found to perform equally well as, if not better than, the mean-variance portfolio, even when measured in terms of Sharpe ratios: see De Miguel *et al.* (2009) and De Miguel *et al.* (2014).

Letting Ω_t be the time $(t-1)$–conditional variance-covariance matrix of the $(M \times 1)$ vector of returns r_t in excess of the risk-free rate, the optimal *relative* weights that minimize portfolio variance are given by:

$$\omega_t = \frac{\Omega_t^{-1}\iota}{\iota'\Omega_t^{-1}\iota} \tag{1}$$

where ι is the $(M \times 1)$ unit vector. Such weights are optimal for investors maximizing the following quadratic utility:

$$V_t = -\frac{\gamma}{2}\omega_t'\Omega_t\omega_t \quad \text{s.t.} \quad \iota'\omega_t = 1 \tag{2}$$

where γ is the investor's risk-aversion. Although inconsequential in the utility specification of equation (2), the level of risk-aversion γ becomes relevant in the presence of transaction costs: see Section 5.5. While the *model-based* and the *model-free* literature have focused on generating forecasts of Ω_t to plug in equation (1), the proposed DCW will model and forecast ω_t directly.

Throughout the paper we consider the portfolio allocation problem of a *day trader* type of investor who closes positions at the end of each trading day. By so doing, we can ignore pre–market or after–hours exchanges which follow different price formation dynamics and for which high-frequency observations are not available. Moreover, it allows us to neatly bypass all potential problems arising from short positions that stretch over long periods of time.

[5]Other examples are Bednarek and Patel (2018), Maillet *et al.* (2015), Candelon *et al.* (2012), Scherer (2011), Clarke *et al.* (2011), etc.

3 Projecting Covariances

3.1 Model-Based Approaches

Let S_t be a realized measure of the var-cov matrix of M assets at time t and $\Omega_t \equiv \mathbb{E}_{t-1}[S_t]$ be its conditional expectation in $(t-1)$. *Model-based* approaches provide dynamic structures for Ω_t in terms of lags of Ω_t and S_t. In general, *model-based* approaches inspired by the most popular MGARCH models generate positive definite (pd) predictions Ω_t under the weakest condition that the realizations S_{t-1} are psd of rank one[6]. However, there is an inherent trade-off to the modeling of pd matrices: parsimony of model parameters *versus* richness in the description of the second order dynamics. In fact, the number of parameters to be jointly estimated is generally a power function of the cross-sectional dimension M. For example, in the Dynamic Conditional Correlations of Engle (2002a), with *targeting*[7] the order is M^2, M^1 and M^0 for the full, diagonal and scalar matrices of parameters, respectively. Nevertheless, the dimensionality problem may be circumvented altogether by the element-by-element modeling of the conditional variance-covariance matrices in the order prescribed by the Sequential Conditional Correlations (SCC) decomposition of Palandri (2009).

Representative specifications of the *model-based* class to be considered in the empirical analysis that follows are Volatility Timing (VT) and Dynamic Conditional Correlations (DCC) based on realized measures. While both approaches model the conditional variances of the returns, only DCC also models the conditional correlation matrix while VT sets it equal to the identity matrix. We model the M conditional variances using the benchmark HAR specification of Corsi (2009). Let $s^2_{i,t}$ be a realized measure of the variance of asset i at time t and $\sigma^2_{i,t} \equiv \mathbb{E}_{t-1}[s^2_{i,t}]$ its conditional expectation at $(t-1)$, then:

$$\sigma^2_{i,t} = \alpha_{i,0} + \alpha_{i,1} \cdot s^2_{i,t-1} + \alpha_{i,2} \cdot \frac{1}{5}\sum_{j=1}^{5} s^2_{i,t-j} + \alpha_{i,3} \cdot \frac{1}{22}\sum_{j=1}^{22} s^2_{i,t-j}$$

which links the conditional variance $\sigma^2_{i,t}$ to past realizations over daily, weekly and monthly time intervals.

In DCC, the conditional var-cov matrix is decomposed into standard-deviation D_t and correlation R_t matrices: $\Omega_t = D_t R_t D_t$. The elements of D_t are populated with

[6] Predictions may fail to be pd for extremely large psd realizations for which $\Omega_t \propto S_{t-1}$.

[7] Variance *targeting*, proposed by Engle and Mezrich (1996), is the setting of the model's unconditional variance to its sample counterpart. In the multivariate case, *targeting* is particularly convenient as it eliminates $M(M+1)/2$ parameters from variance specifications and $M(M-1)/2$ parameters from correlation specifications.

the square-root of the HAR variances, while the elements of R_t are modeled jointly using the Dynamic Conditional Correlation (1,1) specification[8] with *targeting*:

$$R_t = \left(\overline{P} - A\overline{P}A' - B\overline{P}B'\right) + AP_{t-1}A' + BR_{t-1}B', \tag{3}$$

where $\overline{P}$ is the sample average of the realized correlation matrices P_t and A and B are either full, diagonal or scalar matrices of parameters. With a perspective on portfolios constructed over a vast number of assets, it should be noted that, with parameters of order M^0, scalar DCC with *targeting* is the only scalable specification of the three. We estimate model parameters both by least-squares and Gaussian quasi-maximum-likelihood[9]. Finding that the former is orders of magnitude faster than the latter and delivers superior OOS results, we only report and discuss the findings pertaining to the least-squares estimation.

3.2 Model-Free Approaches

Underlying the *model-free* approaches is the assumption that the realized var-cov matrices follow a driftless random-walk process from which $\Omega_t = S_{t-1}$. Although this assumption eliminates estimation and scalability problems, the literature on volatility (GARCH, MGARCH and Realized Variance models) has invariably shown how stationary specifications consistently outperform the random-walk, both in-sample (IS) and OOS. Furthermore, in contrast to the *model-based* specifications, problems do arise using the random-walk when S_{t-1} is psd. With respect to this issue, Aït-Sahalia and Xiu (2017) adopt the following three approaches.

The first consists of aggregating daily realized measures into k-period (for an arbitrary k, e.g. bimonthly) var-cov matrices, which delivers pd k-period measures unsuitable at the daily level. Furthermore, the exclusion of the overnight movements, not captured by the realized measures, while irrelevant from the perspective of a day trader, may accumulate undesirable effects over k-periods.

The second is to express the psd S_{t-1} as the sum of two rank-deficient matrices: the first arising from a factor structure[10] of the data, as in Fan *et al.* (2008), Fan *et al.*

[8] Although (1,1) specifications are well suited in the empirical applications, as highlighted by Hansen and Lunde (2005), further lags may be considered.

[9] DCC estimated by quasi-maximum-likelihood is the HEAVY of Noureldin *et al.* (2012) without the overnight components.

[10] Factor decompositions have been studied extensively in the MGARCH literature and generated the Factor-GARCH family of models: see Diebold and Nerlove (1989), Engle *et al.* (1990), Alexander and Chibumba (1997), Sentana (1998), among others. Although, due to the *curse of dimensionality*, Factor-GARCH models have not been particularly successful in dealing with both flexibility and

(2011), Fan *et al.* (2016) and Aït-Sahalia and Xiu (2017), and the second being the residual var-cov matrix. Calibrating the shrinkage of the latter toward a diagonal or block-diagonal structure[11] allows to achieve pd of the recombined matrix.

The third consists of controlling for portfolio exposure EC (where $EC = 1$ and $EC = \infty$ are the no short-selling, respectively, the unconstrained portfolios) by adding the constraint $\sum_{i=1}^{M} |w_{i,t}| \leq EC$ to the optimization problem in equation (2). Proposed, among others, by Jagannathan and Ma (2003), exposure constraints help reduce the effects of estimation and forecast errors, portfolio turnover and associated transaction costs. However, exposure constraints alone are not enough to turn the optimization problem in (2) from ill- to well-behaved when var-cov matrices are rank-deficient.

For what matters here, as a benchmark specification for the *model-free* class we adopt the raw realized S_{t-1} with exposure constraints. In so doing, we rely on some findings - presented independently by Fan *et al.* (2016, Figure 5), and Aït-Sahalia and Xiu (2017, Figure 6) - showing that factor structures and shrinkages do not bring about significant improvements in the OOS results. Following on their outcome that the optimal exposure is $EC = 2$, we investigate exposure constraints between 1 and 2, i.e. $EC = \{1.00, 1.25, 1.50, 1.75, 2.00\}$, keeping the case of no constraints on the weights ($EC = \infty$) as a reference. In fact, as EC increases, transaction costs become larger and larger; thus $EC > 2$ is suboptimal as the resulting portfolios exhibit larger OOS variances and higher transaction costs. By the same token, values of $EC < 2$ should not be discarded *a priori* as a larger OOS variance may be offset, in some measure, by lower transaction costs.

4 Dynamic Conditional Weights Modeling

Dynamic Conditional Weights (DCW) is an approach directed at the daily time series of realized optimal portfolio weights

$$\nu_t \equiv (\iota' S_t^{-1} \iota)^{-1} S_t^{-1} \iota. \tag{4}$$

The weights ν_t are observable in t (from the observability of the realized S_t) and minimize the portfolio realized variance $\nu_t' S_t \nu_t$. The time series profile of ν_t for a few tickers may be graphically appraised in Figure 1 (Apple, Boeing, Johnson and

feasibility, the new idea of the *model-free* literature is to decompose the realizations into factors and residual components to shrink.

[11]Block-diagonal structures based on characteristics such as sector, industry, etc. had already been investigated in the MGARCH literature: for example, see Billio *et al.* (2006) and Billio and Caporin (2009).

Johnson, and Merck): they display different ranges (same scale is used across) around a changing level, venture into negative territory and, as other financial time series, are characterized by persistence and some short-lived variability.

In order to define the dynamic structure of the DCW we move from an autoregressive representation of the portfolio-variance minimization problem[12] and we express it in terms of a time-independent weighting matrix. To see the details, let us recall that ω_t is the vector of weights minimizing the portfolio conditional variance $d(\Omega_t, \omega_t) \equiv \omega_t' \Omega_t \omega_t$, where $\Omega_t = \mathbb{E}_{t-1}[S_t]$ is the time $(t-1)$–conditional expectation of S_t. Furthermore, let $\{d(S_{t-i}, \omega_t) \equiv \omega_t' S_{t-i} \omega_t\}_{i=0}^{\infty}$ be the sequence of stationary realized portfolio variances, given the portfolio weights ω_t. Thus, the autocorrelation structure of $d(S_t, \omega_t)$ may be satisfactorily represented as an AR(r), so that an expression for $d(\Omega_t, \omega_t)$ can be:

$$d(\Omega_t, \omega_t) = c_t + \sum_{i=1}^{r} \theta_{t,i} d(S_{t-i}, \omega_t), \tag{5}$$

where c_t and $\theta_{t,i}$ (shorthand for $c(\omega_t)$ and $\theta_i(\omega_t)$ respectively) are such that $c_t \geq 0$ and $\theta_{t,i} \geq 0$ $\forall i$ to satisfy necessary and sufficient conditions for the positivity of $d(\Omega_t, \omega_t)$.

Adding and subtracting ν_{t-i} to ω_t, the generic element $d(S_{t-i}, \omega_t)$ may be rewritten as:

$$\begin{aligned} d(S_{t-i}, \omega_t) &= [\nu_{t-i} + (\omega_t - \nu_{t-i})]' S_{t-i} [\nu_{t-i} + (\omega_t - \nu_{t-i})] \\ &= \left(\iota' S_{t-i}^{-1} \iota\right)^{-1} + 2\left(\iota' S_{t-i}^{-1} \iota\right)^{-1} \iota' (\omega_t - \nu_{t-i}) + (\omega_t - \nu_{t-i})' S_{t-i} (\omega_t - \nu_{t-i}) \\ &= \left(\iota' S_{t-i}^{-1} \iota\right)^{-1} + (\omega_t - \nu_{t-i})' S_{t-i} (\omega_t - \nu_{t-i}) \end{aligned} \tag{6}$$

with $\iota' (\omega_t - \nu_{t-i}) = 0$ due to the portfolio weights adding to unity by construction. Similarly, in view of the fact that:

$$1 = \left(\iota' S_0^{-1} \iota\right)^{-1} + (\omega_t - \nu_0)' S_0 (\omega_t - \nu_0),$$

for some symmetric and pd matrix S_0 and $\nu_0 = \left(\iota' S_0^{-1} \iota\right)^{-1} S_0^{-1} \iota$, equation (5) may be rewritten as:

$$\begin{aligned} d(\Omega_t, \omega_t) &= d_0 + (\omega_t - \nu_0)' c_t S_0 (\omega_t - \nu_0) + \sum_{i=1}^{r} (\omega_t - \nu_{t-i})' \theta_{t,i} S_{t-i} (\omega_t - \nu_{t-i}) \\ &= d_0 + m(\omega_t)' W_t \, m(\omega_t) \end{aligned} \tag{7}$$

[12] In the Appendix A.1 we derive identical DCW dynamics from the maximization of a quadratic utility dependent on portfolio returns r_t. The corresponding realized portfolio weights, which are a function of both returns and var-cov matrix, allow for a seamless merger with the literature focusing on estimation-error reduction in the vector of average returns: see, e.g., the Bayes-Stein shrinkage portfolio of Jorion (1985) and Jorion (1986), the Bayesian portfolio based on belief in an asset-pricing model of Pastor (2000) and Pastor and Stambaugh (2000), the portfolio implied by asset-pricing models with unobservable factors of MacKinlay and Pastor (2000), and the three-fund portfolio of Kan and Zhou (2007).

where d_0 is the sum of all the terms that do not depend on ω_t, the vector

$$m(\omega_t) \equiv ((\omega_t - \nu_0)', (\omega_t - \nu_{t-1})', \ldots, (\omega_t - \nu_{t-r})')',$$

and the matrix W_t is block-diagonal with blocks $(c_t S_0, \theta_{t,1} S_{t-1}, \ldots, \theta_{t,r} S_{t-r})$.

Minimizing $m(\omega_t)' W_t\, m(\omega_t)$ wrt ω_t in such a context loops back to the usual solution $\omega_t = (\iota' \Omega_t^{-1} \iota)^{-1} \Omega_t^{-1} \iota$, whose implementation requires the separate specification of the dynamics of Ω_t.

To avoid modeling Ω_t directly, our suggested solution is to apply the minimization problem to the expression $m(\omega_t)' W m(\omega_t)$ with a time-independent W, so that the associated first order conditions may be written as $Q\, m(\omega_t) = 0$ with $Q = (Q_0, Q_1, Q_2, \ldots, Q_r)$ where Q_i are $(M \times M)$ matrices that do not depend on ω_t. Solving for ω_t gives:

$$\omega_t = \widetilde{a} + \sum_{i=1}^{r} \widetilde{A}_i \nu_{t-i}$$

where parameters are in the $(M \times 1)$ vector $\widetilde{a}$ and in the $(M \times M)$ matrices $\widetilde{A}_i$.

By analogy to other financial time series models, one can replace this AR(r) representation (presumably needing a large r) with a more parsimonious representation for the vector of expected portfolio weights ω_t as an ARMA(p, q)[13]:

$$\omega_t = \kappa + \sum_{i=1}^{p} A_i^* \nu_{t-i} + \sum_{j=1}^{q} B_j^* \omega_{t-j} \tag{8}$$

which gives the DCW functional form of the optimal ω_t, given the time-independent W.

Finally, taking expectations of both sides of equation (8) and letting $\omega \equiv \mathbb{E}[\nu_t]$ it follows that:

$$\omega_t = \left[I - \sum_{i=1}^{p} A_i^* - \sum_{j=1}^{q} B_j^* \right] \omega + \sum_{i=1}^{p} A_i^* \nu_{t-i} + \sum_{j=1}^{q} B_j^* \omega_{t-j} \tag{9}$$

The specification using full matrix coefficients guarantees $\iota' \omega_t = 1$ for all t only if, beside $\iota' \omega = 1$ and $\iota' \omega_0 = 1$, we impose the restrictions $A_i^{*\prime} \iota = a_i \iota$ and $B_j^{*\prime} \iota = b_j \iota$ for all i and j (i.e., in each coefficient matrix all columns must add to the same value). This condition is satisfied if the coefficient matrices are scalar but not in the diagonal case, for which the normalization $\omega_t / (\iota' \omega_t)$ is used. Notice that, regardless of the structure

[13]While, in general, a (p, q) parameterization does not necessarily coincide with the (∞) parameterization, just like in the case of a GARCH(p, q) *vs* an ARCH(∞), the former will provide more stable estimates and accurate forecasts than the latter, all the more when truncated.

of the coefficient matrices, DCW requires the modeling of only M dynamic components (the portfolio weights) in contrast to standard *model-based* approaches which require modeling of M conditional variances and $M(M-1)/2$ conditional correlations.

The parameters in (9) may be estimated using various approaches. Two of them are worth discussing briefly: in the first, estimates are obtained by minimizing the sample portfolio variance itself. This approach has the appealing feature of selecting the model parameters that IS minimize the same function used to evaluate the OOS performance. Its main drawback is that the objective function has to be optimized with respect to all parameters jointly, making the estimation particularly cumbersome and difficult to apply to large cross-sectional dimensions M. The second approach is least-squares estimation, performed by the IS minimization of the square distance between predictions ω_t and realizations ν_t. When associated to a diagonal parameterization of the matrices A_i^* and B_j^*, it further allows for the equation-by-equation ARMA estimation of the model parameters. Therefore, in view of its applicability to a vast number of assets M, in what follows we focus on the diagonal DCW specification estimated by least-squares. Furthermore, we will concentrate on the DCW(1,1) parameterization, with a twofold motivation: to provide a fair comparison to the *model-based* DCC(1,1) and to work with a baseline specification that is easy to scale to a large number of assets. It follows that the empirical performance of DCW, presented in Section 5, is likely to improve if a case-by-case best IS specification is derived from any combination of standard selection techniques such as Information Criteria, pruning statistically insignificant parameters, and Box-Jenkins–type procedures based on the properties of the residual ACFs and PACFs.

5 Empirical Application

The data used for portfolio selection pertain to $M = 28$ of the 30 constituents of the Dow Jones 30 Index. The sample has 11 years of high-frequency daily observations from 01/03/2005 to 12/31/2015 for a total of 2768 days. Two series, with tickers TRV and V, are not included in the study because they are not available for the full sample period[14]. Tickers of the 28 included stocks are: AAPL, AXP , BA, CAT, CSCO, CVX, DD, DIS, GE, GS, HD, IBM, INTC, JNJ, JPM, KO, MCD, MMM, MRK, MSFT, NKE, PFE, PG, UNH, UTX, VZ, WMT, XOM. The raw tick-by-tick TAQ data are

[14]TRV data are available only from 02/26/2007 while V data are missing from 08/04/2006 to 02/26/2007.

cleaned using the procedure of Brownlees and Gallo (2006) from which realized kernel covariances are computed following the approach of Barndorff-Nielsen *et al.* (2011). Details on this procedure may be found in Appendix A.2. The sample is split into six 5-year IS periods: 2005-2009 to 2010-2014, with about 1260 observations each. Each model specification is estimated IS and ensuing OOS forecasts are computed for the following year (about 252 observations).

In Section 5.1, we discuss the portfolio weights forecast performance for the diagonal DCW(1,1) of equation (9), the scalar DCC(1,1) of equation (3) and the plain RW of Section 3.2. DCW and DCC with full matrices of coefficients are not considered due to their limited applicability to large cross-sectional dimensions. On the other hand, scalar DCW and diagonal DCC are estimated but not reported as their OOS performance is inferior to that of diagonal and scalar, respectively. Instead, the choice of plain RW as benchmark of the *model-free* approaches is motivated by the findings of Fan *et al.* (2016), Figure 5, and Aït-Sahalia and Xiu (2017), Figure 6, which show that none of the proposed alternatives consistently outperforms the plain $\Omega_t = S_{t-1}$ in attaining the minimum OOS portfolio variance.

In Sections 5.2-5.6 we comment on the resulting portfolio performances in terms of various standard measures as in Bollerslev *et al.* (2018) and the novel *break-even transaction costs*.

5.1 Portfolio Weights

Table 1 reports descriptive statistics of the equation-by-equation ARMA estimates of the parameters of the diagonal DCW(1,1) of equation (9). Persistence, estimated by $A^* + B^*$, is in-line with that of realized variances, with B^* substantially larger than A^*. Specifically, over the 168 estimates, the maximum A^* is 0.35 while the minimum B^* is approximately 0.5. Descriptive statistics for the IS R^2 may be found in Table 2. Over the six IS periods, portfolio weights exhibit different degrees of predictability with R^2 ranging from 4% to 50%. Overall predictability of realized portfolio weights is attested by the average R^2 which ranges between 20% and 28%, depending on the IS period. Given that each realized weight is made of 756 covariances and 28 variances, reported R^2 are found to be in line with those reported by the *model-based* literature for the modeling of realized variances (higher) and covariances (lower).

We measure OOS performance in terms of R^2 as in Welch and Goyal (2008):

$$R^2 = 1 - \frac{MSE_A}{MSE_N}$$

where the MSE_A is the OOS mean-squared forecasting-error of the model whose weights forecasts are being evaluated and MSE_N is the *reference* measure. Notice that, in comparing the performance of competing specifications, the rankings of the OOS R^2 are unaffected by the choice of MSE_N. Here, we calculate MSE_N with respect to the *ex-post* OOS mean of the portfolio weights. Summary statistics of DCW performance are presented in Table 3, with forecasts explaining 10% of the OOS variability in the realized portfolio weights, on average. While for some stocks, DCW forecasts explain more than 20%, for others they are as low as -7%. OOS R^2 associated to portfolio weights forecasts from DCW, DCC and RW are summarized by the kernel density estimates in Figure 4. It clearly emerges that the OOS weights forecasts of DCW are superior to those of DCC, which exhibit R^2 centered at zero and a long left tail. The approximately symmetric distribution of RW OOS R^2, centered around -40%, provides further evidence of the relatively poor performance of random walk dynamics.

To see how the DCW with $EC = \infty$ forecasts behave in practice, we organize one–step ahead results for individual stocks by taking their absolute value and rescaling them to sum up to one. The outcome is then aggregated by sector and ordered according to the average importance over the period considered. The graphical representation of the cumulative relative importance of sectors (value) is influenced by the corresponding cardinality (i.e. value = average $\times$ # of tickers); each sector position is readable as the difference from the lower line (the top line being 1). Over the entire period 2010–2015 (Figure 2), the relative importance of Services is fairly stable around 0.23; the next sector is Consumer Goods whose importance oscillates around 0.20, although it shows a higher variability and a temporary diminished importance during 2013; Healthcare is next 0.15 and it shows a diminishing importance with a drastic reduction of its values right after the beginning of 2013. Technology has an average importance of 0.16 with a fairly stable value over the whole period; Basic Materials has a relative importance of 0.08; Industrial Goods has an overall value 0.12: its relative importance seems to increase after the beginning of 2013 for about one year, and then, again, during the first half of 2015. Finally, Financials has a relative importance of 0.06. Breaking the results by year (Figure 3), we get a more detailed view of the evolution of this relative importance: first and foremost the confirmation that Services and Consumer Goods alternate in the top position (4, respectively 2 times). Financial is always in the weakest position (with a substantial gain in 2015); Health Care is fairly prominent in the first four years (reaching the second ranking in 2013), but it rapidly deteriorates in 2014 and even more so in 2015. Technology jumps to the third position in 2014 and 2015.

5.2 Portfolio Variance (PV)

One measure of OOS performance is the average portfolio variance[15] that emerges from choosing model κ:

$$\mathsf{PV}_\kappa = \frac{1}{T}\sum_{t=1}^{T} \widehat{\omega}'_{\kappa,t} S_t \widehat{\omega}_{\kappa,t}$$

where $\widehat{\omega}_{\kappa,t}$ is the time t forecast of the optimal portfolio weights from model κ, S_t is the time t realized variance-covariance matrix and $t = 1, \ldots, T$ is the OOS period.

From Table 4[16], VT produces smaller portfolio variances than those of the Naive equally weighted portfolio (by between 7.10% and 17.86%, with an average of 14.34% over the six-year period). Overall, the *model-free* RW exhibits PV improvements between 24.61% ($EC \leq \infty$) and 29.04% ($EC \leq 1.50$). For any value of EC, the portfolio variances of the *model-based* DCC are lower than those of RW by between 1.53% ($EC \leq 1.25$) and 9.08% ($EC \leq \infty$). DCW portfolio variance without exposure constraints ($EC \leq \infty$) is the smallest. It is smaller than that of RW for any value of EC, by between 0.80% ($EC \leq \infty$) and 13.30% ($EC \leq 1.25$). It is smaller than that of DCC for all EC above 1.50 (by between 1.16% when $EC \leq 1.50$ and 4.65% when $EC \leq \infty$), but larger for all EC below 1.25 (between 0.74% when $EC = 1.25$ and 1.53% when $EC = 1.00$). Should the investor be able to select the EC parameter *ad hoc*, as is the case for some parameters of the *model-free* approaches, RW portfolio variance would be reduced by 3.44% by DCC and 7.89% by DCW.

5.3 Certainty Equivalent Return (CEQ)

Another common measure of OOS performance is the certainty equivalent return. It is defined as the certain return that an investor is willing to accept to switch from model κ_1 to κ_2:

$$\mathsf{CEQ}_{\kappa_1 \to \kappa_2} = \gamma \cdot \frac{1}{2}\left(\mathsf{PV}_{\kappa_1} - \mathsf{PV}_{\kappa_2}\right) \tag{10}$$

[15] A common measure of the OOS portfolio performance is the Sharpe ratio which highlights the reward-to-risk. However, given that in this study we concentrate exclusively on the contribution of the conditional second moments to optimal portfolio formation, we deem it more appropriate to use a measure of OOS performance that captures second moment effects only.

[16] From here on, in our comments we focus in particular on the overall results reported in the column labeled ALL, unless otherwise stated.

Reporting $\mathsf{CEQ}_{\kappa_1 \to \kappa_2}$ for $\gamma = 1$ allows for the immediate calculation of the certainty equivalent return for any value of risk-aversion[17] simply by rescaling the reported value by γ. While the rankings it generates within this framework are no different from those of PV, CEQ may still be helpful in quantifying the differences in portfolio variances by translating them into returns.

As shown in Table 5, VT exhibits OOS certainty equivalences, with respect to Naive, that range between 0.95 and 6.84 average daily basis points and 3.63 basis points over the whole OOS period. Switching from VT to RW the certainty equivalence ranges from 5.33 ($EC \leq \infty$) to 6.29 ($EC \leq 1.50$) basis points. The switch from RW to DCC also exhibits positive CEQ for any exposure constraint and between 0.24 ($EC \leq 1.25$) and 1.48 ($EC \leq \infty$) basis points. The switch from DCC to DCW exhibits positive CEQ for EC at and above 1.50, between 0.17 ($EC \leq 1.50$) and 0.69 ($EC \leq \infty$) basis points, but -0.11 and -0.24 basis points for ($EC \leq 1.25$) and ($EC = 1.00$), respectively. Again, should the investor be able to select the EC parameter *ad hoc*, switching from RW to DCC and DCW would correspond to 1.48 and 2.17 daily basis points, respectively.

5.4 Turnover (TO)

In this study, where the focus is on daily trading with no overnight holdings, we have zero portfolio weights prior to rebalancing. Hence, average turnover is given by

$$\mathsf{TO}_\kappa = \frac{1}{T} \sum_{t=1}^{T} \sum_{j=1}^{M} |\widehat{\omega}_{\kappa,j,t}|$$

and captures the average portfolio exposure $\overline{EC}_\kappa$ of forecasting model κ: in the optimal portfolio allocation literature it is commonly reported, as it is of relevant interest for investors. This measure, which does not include assets' returns, is derived in Appendix A.3 where we also show its precision up to two orders of magnitude.

In Table 6, Naive and VT have turnovers of 1.00, by construction. Of the other strategies, for any exposure constraint EC, RW exhibits the highest turnover[18] TO, followed by DCC and, last, DCW. In the presence of transaction costs, the implications for the investor are that RW gives rise to the most expensive allocations, followed by those of DCC and DCW.

[17] For example, De Miguel *el al.* (2009) consider risk-aversion coefficients of $\gamma = \{1, 2, 3, 4, 5, 10\}$.

[18] Turnover is a reflection of the dispersion of the resulting portfolio weights and, in the presence of estimation and forecasting errors, it may be taken as an indirect measure of their magnitude.

5.5 Break-Even Transaction Costs (BETC)

To provide a comprehensive view of portfolio performance which includes both PV and the associated transaction costs from TO, we introduce break-even transaction costs as a novel measure of portfolio performance. Specifically, BETC identifies the transaction costs for which two portfolio allocations are indifferent and consecutively the range over which one allocation is preferred to the other. As shown in Appendix A.4, with markup transaction costs τ, average transaction costs $\overline{TC}_\kappa$ for model κ may be approximated up to two orders of magnitude by:

$$\overline{TC}_\kappa \approx 2\tau \cdot \mathsf{TO}_\kappa$$

which, combined with equation (10), allows to derive the *net* certainty equivalent return:

$$NCEQ_{\kappa_1 \to \kappa_2} = \gamma \cdot \frac{1}{2} \left(\mathsf{PV}_{\kappa_1} - \mathsf{PV}_{\kappa_2} \right) + 2\tau \left(\mathsf{TO}_{\kappa_1} - \mathsf{TO}_{\kappa_2} \right) \tag{11}$$

The *break-even transaction cost* (BETC) is defined as the value of $\tau/\gamma > 0$ that sets equation (11) to zero:

$$\mathsf{BETC}_{\kappa_1 \to \kappa_2} = -\frac{1}{4} \cdot \frac{\mathsf{PV}_{\kappa_1} - \mathsf{PV}_{\kappa_2}}{\mathsf{TO}_{\kappa_1} - \mathsf{TO}_{\kappa_2}},$$

and hence it combines PV and TO to identify transaction costs per units of risk-aversion (τ/γ) for which one approach is preferred to another. BETC are reported in Table 7: entries may be simply multiplied by γ to obtain transaction costs corresponding to risk-aversions different from unity. VT is preferred to Naive for any level of the transaction costs τ: smaller PV and equal TO. RW is preferred to VT for any τ only with no short-selling constraints $EC = 1.00$. In the other cases, RW is preferred to VT for greater risk-aversion γ and non-negligible transaction costs. Both DCC and DCW are preferred to RW for any γ and τ. This is due to the fact that their (estimated) shrinkage produces smaller PV and lower TO, both indicative of portfolio weights of higher quality. DCW is always preferred (any τ/γ) to DCC for EC at and above 1.50. It is interesting to note that the year 2010 contains some influential data connected to the flash crash of May 6 which impact on the results: while we opted for not arbitrarily correcting for those specific values, the overall results on 2011–2015 confirm that DCW is to be preferred[19] to RW and DCC for any EC and τ/γ.

[19] Excluding 2010, for the triplet RW, DCC and DCW we have $\mathsf{PV} = \{0.293206, 0.289647, 0.287181\}$, for $EC \leq 1.25$, and $\mathsf{PV} = \{0.307414, 0.301346, 0.300040\}$, for $EC = 1.00$. Similarly, $\mathsf{TO} = \{1.23, 1.23, 1.09\}$, for $EC \leq 1.25$, and $\mathsf{TO} = 1$ for all three when $EC = 1.00$. Hence, there is no τ/γ for which RW is preferred to DCC or DCC is preferred to DCW.

5.6 Utility Levels

In Figure 5 we report the utility levels[20] associated with the various approaches. Specifically, for a given strategy, we begin by plotting the utilities associated with a given strategy as a function of τ/γ for the various exposure constraints EC. We then construct the envelope of each strategy as a function of τ/γ. The envelopes are piecewise linear curves (the utilities are linear in τ/γ) which identify the maximal utility attainable by each strategy for the *ex post* optimal exposure constraint EC. Finally, we report the envelope differences of RW, DCC and DCW with respect to VT. For low τ/γ, high EC allocations are preferred as they produce portfolios with smaller variances. On the other hand, when τ/γ is high, low EC allocations are preferred as the increase in portfolio variance is more than compensated by the decrease in the associated transaction costs. From the first graph of Figure 5, DCC and DCW outperform RW for any τ/γ, while DCC is preferred to DCW for τ/γ greater than 3 basis points. Once again, excluding the year 2010 produces slightly different results as shown in the bottom panel of Figure 5: DCW is preferred to both DCC and RW, for any τ/γ. Furthermore, as τ/γ increases and $EC = 1.00$ becomes optimal for all approaches, the allocation gains of DCW over DCC tend to vanish while their difference with respect to RW remains sizeable.

6 Conclusions

In this paper we motivate the use of Dynamic Conditional Weights by deriving its dynamic structure by expressing the autoregressive representation of the portfolio-variance minimization problem in terms of a time-independent weighting matrix. We evaluate portfolio weights forecasts from the proposed approach against those of representative *model-based* and *model-free* specifications which forecast conditional var-cov matrices to calculate the optimal weights. Specifically, the scalar DCC with HAR variance dynamics as a manageable representative model for the *model-based* class and the daily RW as the benchmark specification for the *model-free* class. We find the DCW portfolio allocations to have lower variance PV and turnover TO than DCC and RW, for any value of the exposure constraints EC. The proposed BETC criterion, which allows the joint evaluation of strategy performance and implementation costs,

[20] In the presence of transaction costs, the utility function of equation (2) becomes $V_t = -2\tau \mathsf{TO}_\kappa - 0.5\gamma \mathsf{PV}_\kappa$. Reported utility levels are those associated to the *rank-preserving* transformation $\widetilde{V}_t = -2(\tau/\gamma)\mathsf{TO}_\kappa - 0.5\mathsf{PV}_\kappa$.

highlights that, for any realistic level of transaction costs, investors would switch from RW to DCC and, with the exception of the no short-selling case $EC = 1.00$, would switch from DCC to DCW.

While measures of portfolio performance are of primary interest in a portfolio management framework, our analysis suggests not to overlook the forecasts of the portfolio weights. In fact, considering that portfolio measures not only capture how close the weights forecasts are to the *realizations*, but also their diversification effects, a given strategy may perform relatively well because it provides one of the infinitely many good diversifications despite poorly forecasting the optimal weights. This may be the case for the RW at the heart of the *model-free* approaches: considering the lack of evidence supporting random walk dynamics for variances and covariances and the relatively poor performance of the associated weights forecasts, performance of the RW portfolio allocations may mostly reflect diversification.

The Dynamic Conditional Weights approach is readily extendible (Appendix A.1) to the general case of a quadratic utility maximization to incorporate the advances in the reduction of estimation-error in the vector of average returns as in De Miguel *et al.* (2014), Bouaddi and Taamouti (2013), Behr *et al.* (2013), Behr *et al.* (2012), Kirby and Ostdiek (2012), Tu and Zhou (2011), De Miguel *et al.* (2010) and Brandt *et al.* (2009), among others. Another noteworthy extension of the proposed approach is the one that ensues when the weighting matrix is allowed to exhibit *time-dependence*. The resulting DCW dynamics will display a higher degree of flexibility by allowing for time-varying parameters, a route suggested by Bollerslev *et al.* (2016) to alleviate model misspecification in the context of var-cov modeling. We leave these and other refinements to future research.

References

Aït-Sahalia, Y., Fan, J. and Xiu, D. (2010) High-frequency covariance estimates with noisy and asynchronous financial data, *Journal of the American Statistical Association*, **105**, 1504–1517.

Aït-Sahalia, Y. and Xiu, D. (2017) Using principal component analysis to estimate a high dimensional factor model with high-frequency data, *Journal of Econometrics*, **201**, 384–399.

Alexander, C. and Chibumba, A. (1997) Multivariate orthogonal factor GARCH, *Mimeo, University of Sussex.*

Andersen, T. G., Bollerslev, T., Christoffersen, P. F. and Diebold, F. X. (2006) Volatility and correlation forecasting, in *Handbook of Economic Forecasting* (Eds.) G. Elliott, C. W. J. Granger and A. Timmermann, North Holland.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. (2011) Multivariate realised kernels: Consistent positive semi-definite estimators of the covariation of equity prices with noise and non-synchronous trading, *Journal of Econometrics*, **162**, 149–169.

Bauwens, L., Laurent, S. and Rombouts, J. V. K. (2006) Multivariate GARCH models: a survey, *Journal of Applied Econometrics*, **21**, 79–109.

Bednarek, Z. and Patel, P. (2018) Understanding the outperformance of the minimum variance portfolio, *Finance Research Letters*, **24**, 175 – 178.

Behr, P., Guettler, A. and Miebs, F. (2013) On portfolio optimization: Imposing the right contraints, *Journal of Banking and Finance*, **37**, 1232–1242.

Behr, P., Guettler, A. and Truebenbach, F. (2012) Using industry momentum to improve portfolio performance, *Journal of Banking and Finance*, **36**, 1414–1423.

Bibinger, M., Hautsch, N., Malec, P. and Reiß, M. (2014) Estimating the quadratic covariation matrix from noisy observations: Local method of moments and efficiency, *The Annals of Statistics*, **42**, 1312–1346.

Billio, M. and Caporin, M. (2009) A generalized dynamic conditional correlation model for portfolio risk evaluation, *Mathematics and Computers in Simulation*, **79**, 2566–2578.

Billio, M., Caporin, M. and Gobbo, M. (2006) Flexible dynamic conditional correlation multivariate GARCH models for asset allocation, *Applied Financial Economics Letters*, **2**, 123–130.

Bollerslev, T., Patton, A. J. and Quaedvlieg, R. (2016) Exploiting the errors: A simple approach for improved volatility forecasting, *Journal of Econometrics*, **192**, 1–18.

Bollerslev, T., Patton, A. J. and Quaedvlieg, R. (2018) Modeling and forecasting (un)reliable realized covariances for more reliable financial decisions, *Journal of Econometrics*, **207**, 71–91.

Bouaddi, M. and Taamouti, A. (2013) Portfolio selection in a data-rich environment, *Journal of Economic Dynamics and Control*, **37**, 2943–2962.

Brandt, M., Santa-Clara, P. and Valkanov, R. (2009) Parametric portfolio policies: Exploiting characteristics in the cross-section of equity returns, *Review of Financial Studies*, **22**, 3411–3447.

Brownlees, C. T., Cipollini, F. and Gallo, G. M. (2012) Multiplicative Error Models, in *Volatility Models and Their Applications* (Eds.) L. Bauwens, C. Hafner and S. Laurent, Wiley, pp. 223–247.

Brownlees, C. T. and Gallo, G. M. (2006) Financial econometric analysis at ultra–high frequency: Data handling concerns, *Computational Statistics and Data Analysis*, **51**, 2232–2245.

Callot, L., Kock, A. and Medeiros, M. (2017) Modeling and forecasting large realized covariance matrices and portfolio choice, *Journal of Applied Econometrics*, **32**, 140–158.

Candelon, B., Hurlin, C. and Tokpavi, S. (2012) Sampling error and double shrinkage estimation of minimum variance portfolios, *Journal of Empirical Finance*, **19**, 511–527.

Chiriac, R. and Voev, V. (2011) Modelling and forecasting multivariate realized volatility, *Journal of Applied Econometrics*, **26**, 922–947.

Christensen, K., Kinnerbrock, S. and Podolskij, M. (2010) Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data, *Journal of Econometrics*, **159**, 116–133.

Cipollini, F., Engle, R. F. and Gallo, G. M. (2013) Semiparametric vector MEM, *Journal of Applied Econometrics*, **28**, 1067–1086.

Clarke, R., de Silva, H. and Thorley, S. (2011) Minimum-variance portfolio composition, *The Journal of Portfolio Management*, **Winter**, 1 – 16.

Corsi, F. (2009) A simple approximate long-memory model of realized volatility, *Journal of Financial Econometrics*, **7**, 174–196.

Corsi, F., Audrino, F. and Renò, R. (2012) HAR modeling for realized volatility forecasting, in *Handbook of Volatility Models and their Applications* (Eds.) L. Bauwens, C. Hafner and S. Laurent, John Wiley & Sons, pp. 371–382.

De Miguel, V., Garlappi, L. and Uppal, R. (2009) Optimal versus naïve diversification: How inefficient is the 1/N strategy?, *Review of Financial Studies*, **22**, 1915–1953.

De Miguel, V., Martin-Utrera, A. and Nogales, F. (2014) Parameter uncertainty in multiperiod portfolio optimization with transaction costs, *Journal of Financial and Quantitative Analysis*, **50**, 1443–1471.

De Miguel, Y., V. Plyakha, Uppal, R. and Vikov, G. (2010) Improving portfolio selection using option-implied volatility and skewness, *Journal of Financial and Quantitative Analysis*, **48**, 1813–1845.

Diebold, F. and Nerlove, M. (1989) The dynamics of exchange rate volatility: A multivariate latent factor ARCH model, *Journal of Applied Econometrics*, **1**, 1–21.

El Karoui, N. (2010) High-dimensional effects in the markowitz problem and other quadratic programs with linear constraints: Risk underestimation, *The Annals of Statistics*, **38**, 3487–3566.

Engle, R., Ng, V. and M., R. (1990) Asset pricing with a factor–ARCH covariance structure: empirical estimates for treasury bills, *Journal of Econometrics*, **45**, 213–238.

Engle, R. F. (2002a) Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models, *Journal of Business & Economic Statistics*, **20**, 339–350.

Engle, R. F. (2002b) New frontiers for ARCH models, *Journal of Applied Econometrics*, **17**, 425–446.

Engle, R. F. and Mezrich, J. (1996) GARCH for groups, *Risk*, **9**, 36–40.

Fan, J., Fan, Y. and Lv, J. (2008) High dimensional covariance matrix estimation using a factor model, *Journal of Econometrics*, **147**, 186–197.

Fan, J., Furger, A. and Xiu, A. (2016) Incorporating global industrial classification standard into portfolio allocation: A simple factor-based large covariance matrix estimator with high frequency data, *Journal of Business and Economic Statistics*, **34**, 489–503.

Fan, J., Liao, Y. and Mincheva, M. (2011) High-dimensional covariance matrix estimation in approximate factor models, *Annals of Statistics*, **39**, 3320–3356.

Fan, J., Zhang, J. and Yu, K. (2012) Vast portfolio selection with gross-exposure constraints, *Journal of the American Statistical Association*, **107**, 592–606.

Gandy, A. and Veraart, L. (2013) The effect of estimation in high-dimensional portfolios, *Mathematical Finance*, **23**, 531–559.

Golosnoy, V., Gribisch, B. and Liesenfeld, R. (2012) The conditional autoregressive Wishart model for multivariate stock market volatility, *Journal of Econometrics*, **167**, 211–223.

Gourieroux, C., Jasiak, J. and Sufana, R. (2009) The Wishart autoregressive process of multivariate stochastic volatility, *Journal of Econometrics*, **150**, 167–181.

Hansen, P. R. and Lunde, A. (2005) A forecast comparison of volatility models: Does anything beat a GARCH(1,1)?, *Journal of Applied Econometrics*, **20**, 873–889.

Jagannathan, R. and Ma, T. (2003) Risk reduction in large portfolios: Why imposing wrong constraints helps, *Journal of Finance*, **58**, 1651–1684.

Jin, X. and Maheu, J. M. (2010) Modeling realized covariances and returns, *Journal of Financial Econometrics*, **11**, 335–369.

Jorion, P. (1985) International portfolio diversification with estimation risk, *Journal of Business*, **58**, 259–292.

Jorion, P. (1986) Bayes-Stein estimation for portfolio analysis, *Journal of Financial and Quantitative Analysis*, **21**, 279–292.

Kan, R. and Zhou, G. (2007) Optimal portfolio choice with parameter uncertainty, *Journal of Financial and Quantitative Analysis*, **42**, 621–656.

Kirby, C. and Ostdiek, B. (2012) It's all in the timing: Simple active portfolio strategies that outperform naïve diversification, *Journal of Financial and Quantitative Analysis*, **47**, 437–467.

Ledoit, O. and Wolf, M. (2012) Nonlinear shrinkage estimation of large-dimensional covariance matrices, *Annals of Statistics*, **40**, 1024–1060.

MacKinlay, A. and Pastor, L. (2000) Asset pricing implications for expected returns and portfolio selection, *The Review of Financial Studies*, **13**, 883–916.

Maillet, B., Tokpavi, S. and Vaucher, B. (2015) Global minimum variance portfolio optimisation under some model risk: A robust regression-based approach, *European Journal of Operational Research*, **244**, 289 – 299.

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

Noureldin, D., Shephard, N. and Sheppard, K. (2012) Multivariate high-frequency-based volatility (HEAVY) models, *Journal of Applied Econometrics*, **27**, 907–933.

Palandri, A. (2009) Sequential conditional correlations: Inference and evaluation, *Journal of Econometrics*, **153**, 122–132.

Park, S. and Linton, O. (2012) Realized volatility: Theory and applications, in *Handbook of Volatility Models and their Applications* (Eds.) L. Bauwens, C. Hafner and S. Laurent, John Wiley & Sons, pp. 319–345.

Pastor, L. (2000) Portfolio selection and asset pricing models, *The Journal of Finance*, **55**, 179—223.

Pastor, L. and Stambaugh, R. F. (2000) Comparing asset pricing models: An investment perspective, *The Journal of Financial Economics*, **56**, 335–381.

Scherer, B. (2011) A note on the returns from minimum variance investing, *Journal of Empirical Finance*, **18**, 652–660.

Sentana, E. (1998) The relation between conditional heteroskedastic factor models and factor GARCH models, *The Econometrics Journal*, **1**, 1–9.

Shephard, N. and Sheppard, K. (2010) Realising the future: forecasting with high frequency based volatility (HEAVY) models, *Journal of Applied Econometrics*, **25**, 197–231.

Tao, M., Wang, Y., Yao, Q. and Zou, J. (2011) Large volatility matrix inference via combining low-frequency and high-frequency approaches, *Journal of the American Statistical Association*, **106**, 1025–1040.

Tao, M., Wang, Y. and Zou, H. (2013) Optimal sparse volatility matrix estimation for high-dimensional itô processes with measurement errors, *The Annals of Statistics*, **141**, 1816–1864.

Tu, J. and Zhou, G. (2011) Markowitz meets talmud: A combination of sophisticated and naïve diversification strategies, *Journal of Financial Economics*, **99**, 204—215.

Welch, I. and Goyal, A. (2008) A comprehensive look at the empirical performance of equity premium prediction, *The Review of Financial Studies*, **21**, 1455–1508.

Zhang, L. (2011) Estimating covariation: Epps effect and microstructure noise, *Journal of Econometrics*, **160**, 33–47.

A Appendix

A.1 Dynamic Conditional Weights for Quadratic Utility

Consider the problem of an investor forecasting portfolio weights ω_t in $(t-1)$ to maximize the quadratic utility:

$$V(r_t, S_t, \omega_t) = \omega_t' r_t - \frac{\gamma}{2}\omega_t' S_t \omega_t$$

where r_t and S_t are the realized vector of returns and var-cov matrix, respectively. Let $\{V(r_{t-i}, S_{t-i}, \omega_t)\}_{i=0}^{\infty}$ be the time series of realized utilities $V(r_{t-i}, S_{t-i}, \omega_t) = \omega_t' r_{t-i} - 0.5\gamma\omega_t S_{t-i}\omega_t$, reconstructed at time t, given the portfolio weights ω_t. Following the same steps of Section 4, assume the autocorrelation structure of $V(r_t, S_t, \omega_t)$ is captured, with the desired degree of precision, by an AR(r) and take $\mathbb{E}_{t-1}$ of both sides of the equality:

$$V\left(\mathbb{E}_{t-1}(r_t), \Omega_t, \omega_t\right) = c_t + \sum_{i=1}^{r} \theta_{t,i} V\left(r_{t-i}, S_{t-i}, \omega_t\right) \tag{12}$$

where $c_t \geq 0$ and $\theta_{t,i} \geq 0$, $\forall i$ to guarantee positivity of the conditional utility $V\left(\mathbb{E}_{t-1}(r_t), \Omega_t, \omega_t\right)$. Let $\delta_{t,i} = (\omega_t - \nu_{t-i})$ where $\nu_{t-i} = \gamma^{-1} S_{t-i}^{-1} r_{t-i}$ is the vector of realized optimal portfolio weights that maximize $V(r_{t-i}, S_{t-i}, \nu_{t-i})$, then the generic element $V\left(r_{t-i}, S_{t-i}, \omega_t\right)$ may be rewritten as:

$$\begin{aligned} V\left(r_{t-i}, S_{t-i}, \omega_t\right) &= \omega_t' r_{t-i} - \frac{\gamma}{2}\nu_{t-i}' S_{t-i}\nu_{t-i} - \gamma\delta_{t,i}' S_{t-i}\nu_{t-i} - \frac{\gamma}{2}\delta_{t-i}' S_{t-i}\delta_{t-i} \\ &= \omega_t' r_{t-i} - \frac{\gamma}{2}\nu_{t-i}' S_{t-i}\nu_{t-i} - \delta_{t,i}' r_{t-i} - \frac{\gamma}{2}\delta_{t-i}' S_{t-i}\delta_{t-i} \\ &= \nu_{t-i}' r_{t-i} - \frac{\gamma}{2}\nu_{t-i}' S_{t-i}\nu_{t-i} - \frac{\gamma}{2}\left(\omega_t - \nu_{t-i}\right)' S_{t-i}\left(\omega_t - \nu_{t-i}\right) \end{aligned}$$

Similarly, in view of the fact that:

$$1 = \frac{1}{2\gamma} r_0' S_0^{-1} r_0 - \frac{\gamma}{2}\left(\omega_t - \nu_0\right)' S_0\left(\omega_t - \nu_0\right)$$

for some vector r_0 and some symmetric and pd matrix S_0 with $\nu_0 = \gamma^{-1} S_0^{-1} r_0$, equation (12) may be rewritten as:

$$\begin{aligned} V\left(\mathbb{E}_{t-1}(r_t), \Omega_t, \omega_t\right) &= d_0 - \frac{\gamma}{2} c_t\left(\omega_t - \nu_0\right)' S_0\left(\omega_t - \nu_0\right) \\ &- \frac{\gamma}{2}\sum_{i=1}^{r} \theta_{t,i}\left(\omega_t - \nu_{t-i}\right)' S_{t-i}\left(\omega_t - \nu_{t-i}\right) \end{aligned}$$

where d_0 is the sum of all the terms that do not depend on ω_t. Following *mutatis mutandis* the steps of Section 4 leads to the same dynamic specification of ω_t as in equation (9) with the only difference being in the definition of the realized weights ν_{t-i}.

A.2 Data Handling

For each trading day t, let $\{r_j\}_{j=1}^{J}$ be the collection of the $(M \times 1)$ return-vectors resulting from price-vectors synchronized according to Barndorff-Nielsen *et al.* (2011). The daily realized kernel variance-covariance matrix is then computed as:

$$S = \sum_{h=-l}^{l} k\left(\frac{h}{H}\right) \Gamma_h$$

where $l = \min(H, J-1)$ and H is:

$$H = \frac{1}{M} \sum_{i=1}^{M} 3.51 \cdot J^{3/5} \left(\frac{(2J)^{-1} \sum_{j=1}^{J} r_{i,j}^2}{\sum_{j=1}^{\widetilde{J}} \widetilde{r}_{i,j}^2} \right)^{2/5}$$

with $r_{i,j}$ are the i-th elements of the vectors r_j. Similarly, $\widetilde{r}_{i,j}$ are the i-th elements of the vectors $\widetilde{r}_j$ where $\{\widetilde{r}_j\}_{j=1}^{\widetilde{J}}$ is the collection of return-vectors in the j-th bin of equally spaced 15 minute intervals. Γ_h and the Parzen kernel $k(x)$ are given by:

$$\Gamma_h = \begin{cases} \displaystyle\sum_{j=h+1}^{J} r_j r'_{j-h} & \text{if } h \geq 0 \\ \displaystyle\sum_{j=-h+1}^{J} r_{j+h} r'_j & \text{if } h < 0 \end{cases} \quad ; \quad k(x) = \begin{cases} 1 - 6x^2 + 6x^3 & \text{if } x \in [0, 1/2] \\ 2(1-x)^3 & \text{if } x \in (1/2, 1] \\ 0 & \text{otherwise} \end{cases}$$

A.3 Turnover Approximation

For our *day trader* type of investor, who opens (closes) all positions at the beginning (end) of the trading day, the average turnover of forecasting model κ over T trading days is given by:

$$2\mathsf{TO}_\kappa = \frac{1}{T} \sum_{t=1}^{T} \sum_{j=1}^{M} |\widehat{\omega}_{\kappa,j,t}| + |\widehat{\omega}^c_{\kappa,j,t}|$$

where $\widehat{\omega}^c_{\kappa,j,t}$ is the value of the portfolio weight $\widehat{\omega}_{\kappa,j,t}$ at the end of the trading day. Value of the weights at *close* is related to value at *open* by $\widehat{\omega}^c_{\kappa,j,t} = (1 + r^{oc}_{j,t}) \cdot \widehat{\omega}_{\kappa,j,t}$, where $r^{oc}_{j,t}$ is the *open-close* return. It follows that turnover may be rewritten as:

$$2\mathsf{TO}_\kappa = \frac{1}{T} \sum_{t=1}^{T} \sum_{j=1}^{M} (2 + r^{oc}_{j,t}) \, |\widehat{\omega}_{\kappa,j,t}|$$

Let $\overline{\overline{r}}$ be the daily weighted average return over the entire time series and across all assets:

$$\overline{\overline{r}} \equiv \frac{\frac{1}{T \cdot M} \sum_{t=1}^{T} \sum_{j=1}^{M} |\widehat{\omega}_{\kappa,j,t}| \cdot r^{oc}_{j,t}}{\frac{1}{T \cdot M} \sum_{t=1}^{T} \sum_{j=1}^{M} |\widehat{\omega}_{\kappa,j,t}|}$$

Then:

$$\frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{M} r_{j,t}^{oc}|\widehat{\omega}_{\kappa,j,t}| = \overline{\overline{r}} \cdot \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{M}|\widehat{\omega}_{\kappa,j,t}|$$

from which it follows that turnover associated with model κ may be rewritten as:

$$\begin{aligned} 2\mathsf{TO}_{\kappa} &= (2+\overline{\overline{r}}) \cdot \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{M}|\widehat{\omega}_{\kappa,j,t}| \\ &\approx 2 \cdot \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{M}|\widehat{\omega}_{\kappa,j,t}| \end{aligned}$$

given that $\overline{\overline{r}}$ is usually a very small number. The main advantage of this approximation is not to rely on portfolio returns: since none of the competing models is optimized with respect to returns, including them in the measures of performance would only amount to adding noise to the analysis. In addition, to have an idea of the goodness of our approximation, consider the case in which the weighted average return is p over a year's time, which corresponds to the daily average $\overline{\overline{r}} \approx 0.004 \cdot p$. It follows that the percentage approximation error ξ of our turnover measure is $\xi = -p(500+p)^{-1}$. Thus, even in the presence of a 100% annual return, the percentage error of our measure of turnover is less than 0.2%.

A.4 Transaction Costs Approximation

The approximation of average transaction cost associated with model κ follows directly from that of turnover:

$$\begin{aligned} \overline{TC}_{\kappa} &= \tau \cdot 2\mathsf{TO}_{\kappa} \\ &\approx 2\tau \cdot \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{M}|\widehat{\omega}_{\kappa,j,t}| \end{aligned}$$

Table 1:

Summary statistics of diagonal DCW parameter estimates over the five IS periods. The, equation by equation, sum of parameters $A^* + B^*$ captures persistence.

	A^*	B^*	$A^* + B^*$		A^*	B^*	$A^* + B^*$
Mean	0.16585	0.77946	0.94531	5% perc.	0.10437	0.60701	0.85147
Minimum	0.07146	0.49990	0.77485	95% perc.	0.24781	0.85961	0.98286
Maximum	0.35466	0.89259	0.98924	Std. Dev.	0.04287	0.07248	0.04051

Table 2:

Summary statistics of IS R^2 for the diagonal DCW specification.

	2005-09	2006-10	2007-11	2008-12	2009-13	2010-14
Mean	0.26927	0.28780	0.28042	0.23246	0.21310	0.20822
Minimum	0.08605	0.09961	0.07387	0.07048	0.04422	0.06710
Maximum	0.46056	0.49641	0.51059	0.48423	0.43555	0.41336
5% perc.	0.09369	0.11291	0.11389	0.07357	0.05967	0.07272
95% perc.	0.44997	0.48114	0.48190	0.44085	0.41778	0.39195
Std. Dev.	0.10319	0.11272	0.09711	0.09085	0.09794	0.09212

Table 3: Summary statistics of OOS R^2 for the diagonal DCW specification. The OOS total sums of squares are calculated from the OOS averages.

	2010	2011	2012	2013	2014	2015
Mean	0.12704	0.12773	0.08409	0.05893	0.10469	0.12473
Minimum	-0.06260	-0.02693	-0.05123	-0.04390	-0.06979	-0.03695
Maximum	0.33471	0.40899	0.37336	0.21401	0.25642	0.36444
Std. Dev.	0.09520	0.10447	0.09785	0.06753	0.08564	0.09834
5% perc.	-0.03601	-0.01474	-0.04148	-0.04130	-0.05593	-0.02137
95% perc.	0.32166	0.36616	0.31808	0.20497	0.24894	0.33456

Table 4:

Average OOS daily variances PV of the portfolio strategies. For Naive and VT, exposure EC is 1.00 by definition. For RW, DCC and DCW results are presented without exposure constraints ($EC \leq \infty$), with $EC \leq \{2.00, 1.75, 1.50, 1.25\}$ and no short-selling ($EC = 1.00$).

Model	2010	2011	2012	2013	2014	2015	All
Naive	0.765768	0.940376	0.337162	0.247769	0.265182	0.477073	0.505796
VT	0.629024	0.776975	0.277726	0.228755	0.242804	0.443216	0.433283
$EC \leq \infty$							
RW	0.410682	0.440253	0.220276	0.249329	0.223546	0.415370	0.326658
DCC	0.363172	0.424437	0.199700	0.211095	0.206934	0.376261	0.297010
DCW	0.368602	0.404762	0.180981	0.200283	0.190229	0.353792	0.283197
$EC \leq 2.00$							
RW	0.398498	0.429203	0.215218	0.241141	0.216777	0.386768	0.314685
DCC	0.363755	0.424242	0.199700	0.211045	0.206517	0.375678	0.296899
DCW	0.373493	0.418564	0.181196	0.199287	0.190273	0.359317	0.287114
$EC \leq 1.75$							
RW	0.389666	0.431561	0.209392	0.235116	0.211803	0.381235	0.309880
DCC	0.365464	0.429841	0.199629	0.210688	0.205754	0.374272	0.297687
DCW	0.379125	0.428851	0.181590	0.198752	0.190562	0.362648	0.290351
$EC \leq 1.50$							
RW	0.387697	0.441894	0.202447	0.227201	0.206624	0.378370	0.307464
DCC	0.370794	0.442001	0.198367	0.209836	0.204315	0.372569	0.299733
DCW	0.390087	0.445715	0.183172	0.198466	0.191394	0.368115	0.296260
$EC \leq 1.25$							
RW	0.396078	0.466123	0.198625	0.218991	0.202397	0.379484	0.310385
DCC	0.385483	0.467761	0.196376	0.207064	0.202168	0.374460	0.305651
DCW	0.410382	0.478605	0.187242	0.199859	0.193726	0.376030	0.307905
$EC = 1.00$							
RW	0.416423	0.509316	0.206378	0.223728	0.205606	0.391573	0.325618
DCC	0.399477	0.510957	0.201358	0.208639	0.201557	0.383752	0.317734
DCW	0.435044	0.516830	0.193881	0.205725	0.196801	0.386463	0.322585

Table 5:

Average OOS daily certainty equivalent CEQ, expressed in *basis points*, relative to the change of strategy indicated by $\rightarrow$. CEQ are calculated for a risk-aversion coefficient of $\gamma = 1$ and may be computed for different values of γ by simple multiplication. For Naive and VT, exposure EC is 1.00 by definition. For RW, DCC and DCW results are presented without exposure constraints ($EC \leq \infty$), with $EC \leq \{2.00, 1.75, 1.50, 1.25\}$ and no short-selling ($EC = 1.00$).

Model	2010	2011	2012	2013	2014	2015	All
Naive → VT	6.84	8.17	2.97	0.95	1.12	1.69	3.63
$EC \leq \infty$							
VT → RW	10.92	16.84	2.87	-1.03	0.96	1.39	5.33
RW → DCC	2.38	0.79	1.03	1.91	0.83	1.96	1.48
DCC → DCW	-0.27	0.98	0.94	0.54	0.84	1.12	0.69
$EC \leq 2.00$							
VT → RW	11.53	17.39	3.13	-0.62	1.30	2.82	5.93
RW → DCC	1.74	0.25	0.78	1.50	0.51	0.55	0.89
DCC → DCW	-0.49	0.28	0.93	0.59	0.81	0.82	0.49
$EC \leq 1.75$							
VT → RW	11.97	17.27	3.42	-0.32	1.55	3.10	6.17
RW → DCC	1.21	0.09	0.49	1.22	0.30	0.35	0.61
DCC → DCW	-0.68	0.05	0.90	0.60	0.76	0.58	0.37
$EC \leq 1.50$							
VT → RW	12.07	16.75	3.76	0.08	1.81	3.24	6.29
RW → DCC	0.85	-0.01	0.20	0.87	0.12	0.29	0.39
DCC → DCW	-0.96	-0.19	0.76	0.57	0.65	0.22	0.17
$EC \leq 1.25$							
VT → RW	11.65	15.54	3.96	0.49	2.02	3.19	6.14
RW → DCC	0.53	-0.08	0.11	0.60	0.01	0.25	0.24
DCC → DCW	-1.24	-0.54	0.46	0.36	0.42	-0.08	-0.11
$EC = 1.00$							
VT → RW	10.63	13.38	3.57	0.25	1.86	2.58	5.38
RW → DCC	0.85	-0.08	0.25	0.75	0.20	0.39	0.39
DCC → DCW	-1.78	-0.29	0.37	0.15	0.24	-0.14	-0.24

Table 6:

Average daily OOS daily turnover TO of the portfolio strategies. For Naive and VT, exposure EC is 1.00 by definition. For RW, DCC and DCW results are presented without exposure constraints ($EC \leq \infty$), with $EC \leq \{2.00, 1.75, 1.50, 1.25\}$ and no short-selling ($EC = 1.00$).

Model	2010	2011	2012	2013	2014	2015	All
Naive	1.00	1.00	1.00	1.00	1.00	1.00	1.00
VT	1.00	1.00	1.00	1.00	1.00	1.00	1.00
$EC \leq \infty$							
RW	2.06	2.09	1.84	1.83	1.80	2.17	1.97
DCC	1.68	1.74	1.49	1.43	1.47	1.69	1.59
DCW	1.49	1.55	1.34	1.27	1.21	1.42	1.38
$EC \leq 2.00$							
RW	1.87	1.84	1.77	1.73	1.70	1.83	1.79
DCC	1.67	1.69	1.49	1.43	1.46	1.66	1.57
DCW	1.42	1.45	1.32	1.23	1.18	1.29	1.32
$EC \leq 1.75$							
RW	1.71	1.69	1.67	1.62	1.60	1.67	1.66
DCC	1.61	1.61	1.49	1.42	1.44	1.60	1.53
DCW	1.36	1.39	1.28	1.20	1.15	1.23	1.27
$EC \leq 1.50$							
RW	1.49	1.49	1.49	1.46	1.43	1.46	1.47
DCC	1.47	1.46	1.43	1.38	1.38	1.45	1.43
DCW	1.26	1.29	1.22	1.15	1.11	1.14	1.19
$EC \leq 1.25$							
RW	1.25	1.25	1.25	1.22	1.22	1.22	1.23
DCC	1.25	1.25	1.25	1.22	1.20	1.22	1.23
DCW	1.13	1.15	1.12	1.07	1.05	1.05	1.09
$EC = 1.00$							
RW	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DCC	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DCW	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 7:

Average OOS daily break-even transaction costs BETC, expressed in *basis points*, relative to the change of strategy indicated by $\rightarrow$. BETC are calculated for a risk-aversion coefficient of $\gamma = 1$ and may be computed for different values of γ by simple multiplication. $<$ and $>$ define the range of transaction costs for which the strategy on the right of $\rightarrow$ is preferred to that on the left. The entry A (N) indicates that the strategy to the right of $\rightarrow$ is preferred for Any (No) value of τ/γ. For Naive and VT, exposure EC is 1.00 by definition. For RW, DCC and DCW results are presented without exposure constraints ($EC \leq \infty$), with $EC \leq \{2.00, 1.75, 1.50, 1.25\}$ and no short-selling ($EC = 1.00$).

Model	2010	2011	2012	2013	2014	2015	All
Naive → VT	A	A	A	A	A	A	A
$EC \leq \infty$							
VT → RW	<5.15	<7.72	<1.71	N	<0.60	<0.60	<2.75
RW → DCC	A	A	A	A	A	A	A
DCC → DCW	<0.71	A	A	A	A	A	A
$EC \leq 2.00$							
VT → RW	<6.62	<10.35	<2.03	N	<0.93	<1.70	<3.75
RW → DCC	A	A	A	A	A	A	A
DCC → DCW	A	A	A	A	A	A	A
$EC \leq 1.75$							
VT → RW	<8.43	<12.53	<2.55	N	<1.29	<2.31	<4.67
RW → DCC	A	A	A	A	A	A	A
DCC → DCW	>1.37	A	A	A	A	A	A
$EC \leq 1.50$							
VT → RW	<12.31	<17.10	<3.84	N	<2.10	<3.52	<6.69
RW → DCC	A	A	A	A	A	A	A
DCC → DCW	>2.30	>0.55	A	A	A	A	A
$EC \leq 1.25$							
VT → RW	<23.29	<31.09	<7.91	<1.11	<4.59	<7.24	<13.36
RW → DCC	A	N	A	A	A	A	A
DCC → DCW	>5.19	>2.71	A	A	A	>0.23	>0.40
$EC = 1.00$							
VT → RW	A	A	A	A	A	A	A
RW → DCC	A	N	A	A	A	A	A
DCC → DCW	N	N	A	A	A	N	N

Figure 1: Realized portfolio weights over the entire 2005-2015 period for Apple (top-left), Boeing (top-right), Johnson & Johnson (bottom-left) and Merck (bottom-right).

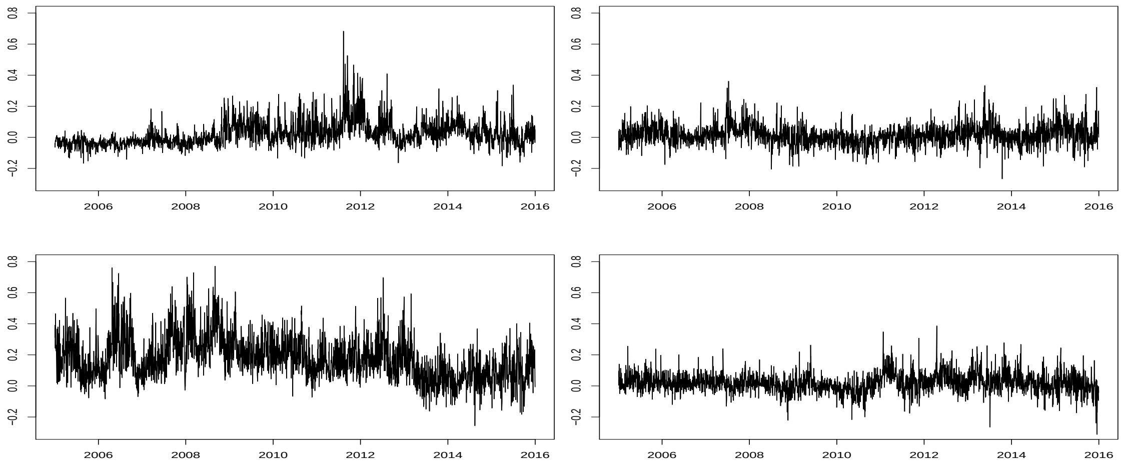

Figure 2: Graphical representation of cumulative relative importance of sectors over the entire period 2010–2015. One–step ahead weight forecasts for individual stocks in the DJ30 from DCW are taken in absolute value and then rescaled to sum up to one. Sector values are obtained by aggregation and then ordered (bottom to top) according to the average relative importance; single sector positions are readable as a difference from the lower line (top line =1).

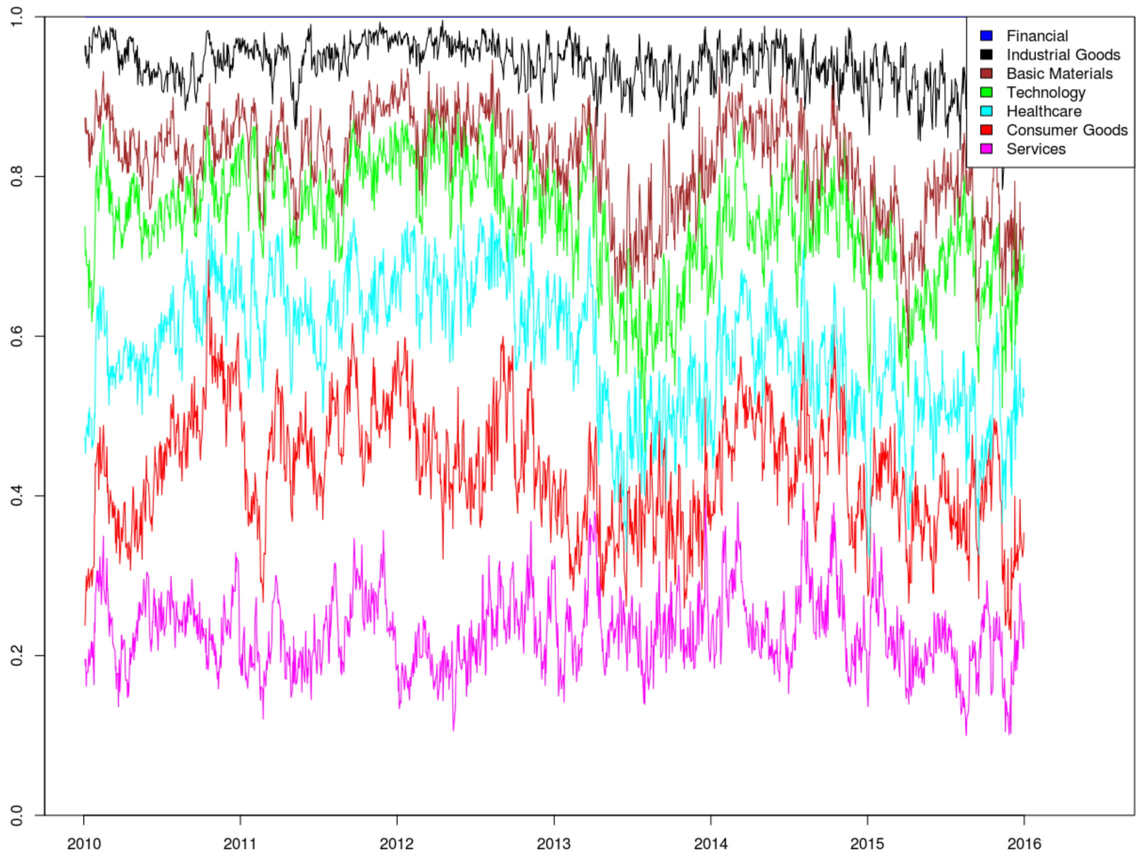

Figure 3: Graphical representation of cumulative relative importance of sectors by year 2010 to 2015. One–step ahead weight forecasts for individual stocks in the DJ30 from DCW are taken in absolute value and then rescaled to sum up to one. Sector values are obtained by aggregation and then ordered (bottom to top) according to the average relative importance; single sector positions are readable as a difference from the lower line (top line =1).

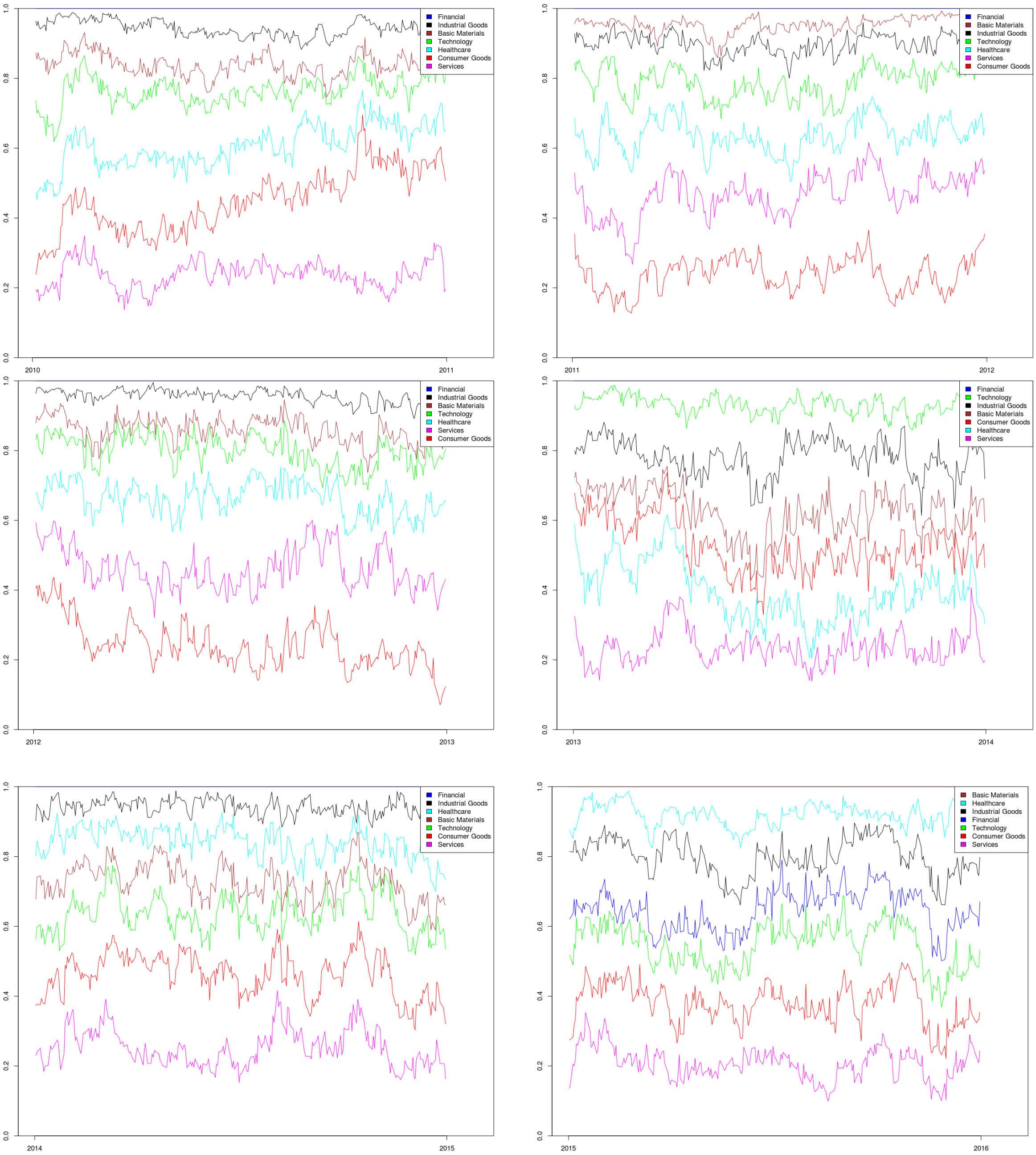

Figure 4: Density representation of OOS R^2 of forecasted portfolio weights across assets and time periods.

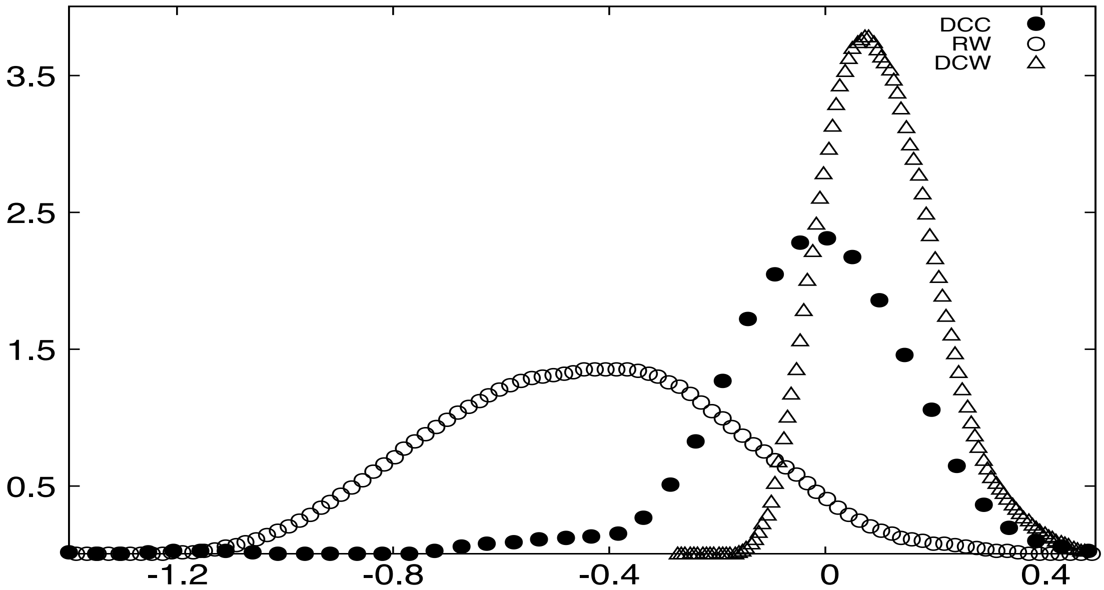

Figure 5: Graphical representation of utility envelopes of RW, DCC and DCW with respect to VT. On the horizontal axis are the transaction costs per units of risk-aversion τ/γ and on the vertical axis are the utilities measured with respect to that of VT. The first graph represents the entire period 2010–2015, while the second graph excludes the year 2010.