Direct Portfolio Weight Estimator: Mitigating Specification Risk with Realized Utility

Ekaterina Kazak[†] Yifan Li[‡] Ingmar Nolte[§]

University of Manchester University of Manchester Lancaster University

Sandra Nolte[¶]

Lancaster University

Abstract

Estimation noise is a well-known issue in empirical portfolio modelling. However, existing models suffer from large forecasting errors which dominate the theoretical gain. In this paper, we propose a direct weight estimator (DWE), which accounts for forecasting risk and avoids the over-parametrization problem by forecasting a one-dimensional portfolio measure directly. We define a forecasting error based on realized measures and optimize for a weight vector which results in a more precise forecast and at the same time is not far from the optimal portfolio solution. The DWE is shown to outperform commonly used approaches in both simulation and empirical studies.

JEL classification: G11, C58, C22, C53

[*]Financial support by the MEiN Program of the University of Konstanz is gratefully acknowledged. The authors would like to thank the participants of the Konstanz-Lancaster Econometrics workshop German Statistical Week and CFE conference in London for their helpful comments. All remaining errors are ours.

[†]Department of Economics, School of Social Sciences, Arthur Lewis Building, Oxford Road, Manchester M13 9PL, email: ekaterina.kazak@manchester.ac.uk (corresponding author).

[‡]Alliance Manchester Business School, Booth Street West, Manchester M15 6PB, UK. Phone: +44 (0) 161 306 6402, email: yifan.li@manchester.ac.uk.

[§]Lancaster University Management School, Bailrigg, Lancaster, LA1 4YX, UK. Phone: +44 (0)1524 592644, email: i.nolte@lancaster.ac.uk.

[¶]Lancaster University Management School, Bailrigg, Lancaster, LA1 4YX, UK. Phone: +44 (0)1524 592644, email: s.nolte@lancaster.ac.uk.

1 Introduction

It has been long recognized in the literature that modelling underlying volatility of a return process is one of the key elements of the investment decision. The fundamental idea is based on the assumption that the investment problem can be formulated in terms of the first two moments of the return process. When it comes to a portfolio allocation problem, not only the latent volatility process of individual stock returns needs to be estimated and forecasted, but also the covariance of the portfolio constituents.

The financial econometrics literature on multivariate volatility estimation and forecasting is very rich (Bauwens et al., 2006). A commonly used class of Dynamic Conditional Correlation (DCC) models by Tse and Tsui (2002) and Engle (2002) has been often criticised for a curse of dimensionality. These models use a so-called DRD decomposition, which splits the covariance matrix into a diagonal matrix of individual stock variances D and a conditional correlation matrix R, whereby D and R are estimated and modelled separately. Most of the DCC specifications require a highly parametric two-step likelihood estimation subject to non-trivial stationarity and positive definiteness constraints, which become intractable when the number of assets in the portfolio is large.

As the full parametric specification of the DCC is often infeasible, one of the most common ways to approach the problem of over-parametrization is to simplify the parametric specification of correlations and/or variances at the cost of model flexibility, e.g. a scalar DCC model is not able to capture volatility spillover effects. Another direction of volatility forecasting looks into the shrinkage approaches. Ledoit and Wolf (2004; 2014) wrote a series of papers on different shrinkage approaches aiming at stabilizing the high dimensional covariance estimation problem; among recent contributions Engle et al. (2019) use eigenvalue shrinkage to stabilize the inverse of estimated covariance matrix inverse for further use in the global minimum variance portfolio. Amongst other improvements in the modelling of multivariate volatility are the contributions of Hautsch et al. (2015); Jin and Maheu (2013); Golosnoy et al. (2012); Hafner and Linton (2010), however the curse of dimensionality prevents any of the multivariate models to significantly dominate benchmark DCC models in a general set up.

Another direction in the portfolio optimization literature tackles the problem of over-parametrization by regularizing the portfolio weight vector directly. An example among many other contributions is Jagannathan and Ma (2003), who propose to impose a norm-constraint on the portfolio optimization problem to account for no short sale constraints. Similarly Brodie et al. (2009), Li (2015), Goto and Xu (2015) use the ℓ_1-penalization (lasso) and Yen (2015) the ℓ_2-penalization (ridge) to constrain the portfolio weights.

The major drawback of both shrinkage approaches is the choice of tuning parameters, which is often non-trivial and derived either under unrealistic assumptions on the data generating process or relies on the estimates of the first and second moments of the return process. As an alternative of the return-based estimation Bollerslev et al. (2018b) recently proposed a HAR-DRD model, which became feasible due to the availability of the high-frequency data. Intra-day returns allow for *ex post* nonparametric and consistent estimation of the daily realized covariance matrices. With a simple HAR model of Corsi (2009) one can reliably forecast univariate realized variances and correlations, which are combined together in a DRD decomposition. The HAR-DRD model is extremely easy to estimate, however, it is subject to the same critique as the DCC models, namely the number of parameters to estimate grows with the portfolio's dimension.

Therefore, in this paper, we propose a novel method of forecasting the portfolio weight vector. In the first step, we utilize realized covariance matrices to construct the ex post optimal univariate portfolio performance measure of interest. In the second step, this series is forecasted with a simple univariate HAR model. We then use a constraint optimization to obtain the forecast of the weight vector which is optimal with respect to the chosen portfolio performance measure. Our proposed direct weight estimation approach is easy to implement, relies on the univariate forecast irrespectively of the portfolio dimension and is flexible enough to adapt to a variety of portfolio performance criteria.

The remainder of the paper is organized as follows. Section 2 introduces the novel direct weight estimation (DWE) approach. Section 3 illustrates the properties of DWE in a simulation study. Section 4 examines its performance in an empirical study and discusses the restricted DWE. Section 5 concludes.

2 Model

We are interested in the following dynamic portfolio choice problem. Starting with a market with N distinct assets, we observe the N-dimensional asset prices P_t on a daily basis, where t denotes the index of the trading days. The corresponding return of holding the assets on day t is defined as $r_t = P_t/P_{t-1} - 1$. We shall make the following assumption on the return process:

Assumption 1

On a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$, define an N-dimensional vector-valued daily return process $\{r_t\}_{t=1,2,\dots}$. For each interval $[t-1, t]$, r_t is generated as follows:

$$r_t = \mu_t + \int_{t-1}^{t} \Theta_s dW_s, \qquad (1)$$

where μ_t is a $\mathcal{F}_{t-1}$-predictable bounded random variable, Θ_s is an $\mathcal{F}_s$-adapted N-by-N right-continuous spot covolatility process, and W_s is an N-dimensional Wiener process. Denote $\Sigma_t = \int_{t-1}^{t} \Theta_s \Theta_s' ds$ as the quadratic covariation of r_t on $[t-1, t]$, we assume that for all $\omega \in \mathbb{R}^N$ and t, it holds almost surely that $0 < \omega' \Sigma_t \omega < \infty$ and that Σ_t is weakly stationary and ergodic.

In essence, the above assumption states that r_t is the endpoint of a continuous semi-martingale with some drift μ_t and a positive-definite and stationary quadratic covariation matrix Σ_t. The continuous assumption is only for notational convenience and should not be viewed as a restriction here, as the quadratic covariation is also well-defined in the presence of square-integrable finite-activity jumps (see, for example, Chapter 1 of Aït-Sahalia and Jacod (2014)), albeit with more cumbersome expressions. The assumption implies that $\mathrm{E}_{t-1}[r_t] = \mu_t$ and $\mathrm{V}_{t-1}[r_t] = \mathrm{E}_{t-1}[\Sigma_t]$ by construction and the Ito isometry, where $\mathrm{E}_{t-1}[\cdot] = \mathrm{E}[\cdot|\mathcal{F}_{t-1}]$ is the $\mathcal{F}_{t-1}$-conditional expectation operator. For notational convenience we will denote $\Omega_t = \mathrm{E}_{t-1}[\Sigma_t]$.

At time t, a representative investor would like to invest all her capital into the N assets based on the information set $\mathcal{F}_t$. She then waits till time $t+1$ and rebalances the position according to $\mathcal{F}_{t+1}$, and the procedure iterates indefinitely. Let us use the N-by-1 vector ω_t to denote the weights assigned to each asset at time t based on $\mathcal{F}_{t-1}$, which satisfies $\omega_t \in \mathcal{S}$, where $\mathcal{S} = \{\omega \in \mathbb{R}^N : \omega'\iota = 1\}$ and ι is an N-by-1 vector of ones. The classical von Neumann-Morgenstern theorem states that at time t, the investor should maximize the conditional expectation of her utility of holding the portfolio $\omega_t' r_t$ to solve for the portfolio weights:

$$\omega_t^\star = \underset{\omega_t \in \mathcal{S}}{\arg\max}\ \mathrm{E}_{t-1}[\mathcal{U}(\omega_t' r_t)], \quad s.t.\ \omega_t'\iota = 1. \tag{2}$$

where $\mathcal{U}(\cdot)$ is the utility function of the investor. To simplify the analysis, we shall assume that the investor has a simple mean-variance conditional utility function:

$$\mathrm{E}_{t-1}[U_t(\omega_t;\gamma)] = \omega_t'\mu_t - \frac{\gamma}{2}\omega_t'\Omega_t\omega_t, \tag{3}$$

where γ is the Arrow-Pratt risk-aversion coefficient. This can be considered as a second-order Taylor expansion around $\mathrm{E}_{t-1}[\omega_t' r_t]$ that approximates a general utility function (Bollerslev et al., 2018a).

In this paper, we mainly discuss the case with $\gamma = \infty$, which suggests that the investor is infinitely risk averse and cares only about the risk of the portfolio. Although the general $\gamma > 0$ case can be dealt with in a similar manner using our method, setting $\gamma = \infty$ allows us to focus on optimizing the portfolio variance without the knowledge of μ_t, which is known to be notoriously difficult to model precisely (Jagannathan and Ma, 2003). In this case, we effectively have $U_t(\omega_t) = -\omega_t'\Omega_t\omega_t$, and the argument $\gamma = \infty$ is suppressed for notational convenience. The optimal *ex ante* weight vector that solves $\max_{\omega_t \in \mathcal{S}} U_t(\omega_t)$ is the well-known Global Minimum Variance Portfolio (GMVP) weights:

$$\omega_t^\star = \frac{\Omega_t^{-1}\iota}{\iota'\Omega_t^{-1}\iota}, \tag{4}$$

which requires the knowledge of Ω_t. However, modelling Ω_t based only on the daily return process proves to be very challenging due to the curse of dimensionality.

The recent developments in econometric methods for high-frequency data bring a new solution to this problem. With high-frequency data, we can construct a sequence of daily realized covariances $\{\hat{\Sigma}_t\}_{t=1,2,\dots}$ that estimates the quadratic covariation matrices $\{\Sigma_t\}_{t=1,2,\dots}$. Under the assumption that these estimators are conditionally unbiased such that $\mathrm{E}_{t-1}[\hat{\Sigma}_t] = \mathrm{E}_{t-1}[\Sigma_t] = \Omega_t$, Ω_t can be interpreted as the linear predictor of $\hat{\Sigma}_t$. The availability of $\{\hat{\Sigma}_t\}_{t=1,2,\dots}$ allows us to consider a slightly easier problem:

GMVP allocation with high-frequency data: Given observations of daily realized covariances $\{\Sigma_t\}_{t=1,2,\dots}$ and let us denote $\mathcal{F}_t^\Sigma = \sigma(\{\Sigma_s : 0 \leq s \leq t\})$ as the filtration generated by the daily realized covariances up to time t. The GMVP allocation problem is to find an $\mathcal{F}_t^\Sigma$-adapted weight vector $\omega_t \in \mathcal{S}$ which maximizes $E_{t-1}[RU_t(\omega_t)]$, where

$RU_t(\omega_t)$ is the *realized utility*[1] (RU) defined as:

$$RU_t(\omega_t) = -\omega_t'\hat{\Sigma}_t\omega_t. \tag{5}$$

Here one can also enlarge the filtration $\mathcal{F}_t^{\Sigma}$ to include more information in determining ω_t. Under the assumption that $\hat{\Sigma}_t$ is conditionally unbiased, one sees that $E_{t-1}[RU_t(\omega_t)] = U_t(\omega_t)$, which reconciles with the previous task in Eq. (2).

For a particular choice of ω_t, its performance can be evaluated by the unconditional RU:

$$\mathrm{E}[RU_t(\omega_t)] = \mathrm{E}[-\omega_t'\hat{\Sigma}_t\omega_t] = \mathrm{E}[U_t(\omega_t)] \leq \mathrm{E}[U_t(\omega_t^\star)], \tag{6}$$

where the second equality holds by the assumption of conditional unbiasedness of $\hat{\Sigma}_t$, and the last equality is obtained iff $\omega_t = \omega_t^\star$. Importantly, since RUs are observed as long as $\hat{\Sigma}_t$ is available, the unconditional RU can be estimated consistently with a sufficiently long out-of-sample period by the law of large numbers, similar to a standard forecasting exercise. This allows us to obtain a more precise portfolio performance measure than the portfolio variance based on daily returns, i.e. $\mathrm{V}[\omega_t' r_t]$.

The most popular solution to the GMVP problem with high-frequency data in the literature is the so-called 'plug-in' method. The method consists of two steps. First, construct an $\mathcal{F}_{t-1}^{\Sigma}$-adapted predictor $\hat{\Omega}_t$ of $\hat{\Sigma}_t$, usually using some least squares-based models. Second, plug $\hat{\Omega}_t$ into Eq. (4) to obtain the estimated GMVP weights:

$$\hat{\omega}_t^\star = \frac{\hat{\Omega}_t^{-1}\iota}{\iota'\hat{\Omega}_t^{-1}\iota}. \tag{7}$$

[1]Note that for the general case with $\gamma > 0$, the realized utility can be defined as $RU_t(\omega_t;\gamma) = \omega_t' r_t - \frac{\gamma}{2}\omega_t'\hat{\Sigma}_t\omega_t$.

The rationale is straightforward: given the conditionally unbiasedness of $\hat{\Sigma}_t$, Ω_t is the mean-variance best linear predictor of $\hat{\Sigma}_t$, and $\hat{\Omega}_t$ is the sample counterpart of Ω_t. Clearly, as $\hat{\Omega}_t$ approaches Ω_t, $\hat{\omega}_t^\star$ converges to $\omega_t^\star$, the optimal GMVP weight vector.

However, there are two major problems with the plug-in approach. First, to produce the estimate $\hat{\Omega}_t$, one needs to simultaneously model $N(N+1)/2$ unique time series while ensuring the positive-definiteness of $\hat{\Omega}_t$ in order to compute its inverse. The existing literature typically employs some simplification or regularization in constructing the forecasting model, which inevitably introduces a substantial amount of misspecification risk for modelling Ω_t. Second, $\hat{\omega}_t^\star$ is by design an *indirect* estimator of $\omega_t^\star$, as the prediction model typically estimates $\hat{\Omega}_t$ by minimizing its distance to $\hat{\Sigma}_t$ in a mean-squared sense, which does not directly translate into a higher expected utility due to the highly nonlinear functional form of $\hat{\omega}_t^\star$ in terms of $\hat{\Omega}_t$.

The problems of the plug-in approach motivates us to propose a *direct weight estimator* (DWE) of the portfolio weights. We utilize the fact that for any $\omega \in \mathcal{S}$, $RU_t(\omega)$ is observed given $\hat{\Sigma}_t$. Therefore, for a particular ω, we can compute an $\mathcal{F}_{t-1}^{\Sigma}$-adapted forecast of $RU_t(\omega)$, denoted as $\widehat{RU}_t(\omega)$. The DWE approach attempts to find $\hat{\omega}_t$ that solves the following maximization problem:

$$\hat{\omega}_t^{DWE} = \underset{\omega \in \mathcal{S}}{\arg\max}\ \widehat{RU}_t(\omega). \tag{8}$$

Intuitively, $\hat{\omega}_t^{DWE}$ is the weight vector that produces the largest predicted RU. In practice, the above problem can be solved easily by a gradient-based numerical optimization algorithm.

The DWE approach is more appealing than the plug-in method in several aspects: (1) instead of modelling the full realized covariance matrices, for each ω we only model a univariate time series $\{RU_s(\omega)\}_{s=1:t-1}$ without the need to predict the realized covariance matrix, which is a much simpler forecasting problem. One can thus fully exploit existing univariate prediction models to reduce misspecification. (2) as opposed to the plug-in approach whose forecasting target is $\hat{\Sigma}_t$ which does not directly translate into a higher utility, the DWE approach by construction directly maximizes the predicted RU, an estimator of the expected utility. (3) the number of free parameters to estimate is only $N-1$ for the DWE approach, which grows linearly with N and is free from the curse of dimensionality. (4) the DWE approach can be easily extended to account for constraints or modifications to the target function, e.g. a short-selling constraint or a transaction cost-adjusted certainty equivalent, as one can simply add parameter constraints and modify the target function of the numerical optimization algorithm. On the contrary, such an estimate may not always be possible for plug-in-type estimators as a closed-form solution of the weight vector as a function of Ω_t may not exist.

2.1 Implementation Details

We now explain how the optimization problem of Eq. (8) is solved in detail. Suppose we construct the daily realized covariance measures $\{\hat{\Sigma}_s\}_{s=1:t-1}$ from day 1 to day $t-1$, and we would like to estimate the GMVP weights $\omega_t^{\star}$. For any vector $\omega \in \mathcal{S}$, we can form the time series of realized utility, $\{RU_s(\omega)\}_{s=1:t-1}$. Note that given $\hat{\Sigma}_t$, the realized utility is bounded above by $RU_s(\omega) \leq RU_s(\omega_t^*) = \frac{1}{\iota'\hat{\Sigma}_t^{-1}\iota}$, where $\omega_t^* = \frac{\hat{\Sigma}^{-1}\iota}{\iota'\hat{\Sigma}^{-1}\iota}$ is the *ex post* GMVP weight vector that maximizes the realized utility. The upper bound allows us to construct loss functions to measure the optimality of some ω relative to the *ex post* optimal choice ω_t^*. Among the many possible choices, we propose to use the log-distance loss function, defined as:

$$ld_t(\omega) = \ln \frac{RU_t(\omega_t^*)}{RU_t(\omega)} = \ln[(\omega'\hat{\Sigma}_t\omega)(\iota\hat{\Sigma}_t^{-1}\iota)] \geq 0, \forall \omega \in \mathcal{S}. \tag{9}$$

It is a loss function in the sense that $ld_t(\omega) = 0$ iff $\omega = \omega_t^*$. We caution that $\hat{\omega}_t^*$ is different from the *ex ante* optimal choice $\hat{\omega}_t^\star$, as $\hat{\omega}_t^*$ requires the knowledge at time t which is not available to the investor at time $t-1$. However, since both $RU_t(\omega_t^\star)$ and $RU_t(\omega_t^*)$ are not functions of ω_t, using $RU_t(\omega_t^*)$ instead of the unobserved $RU_t(\omega_t^\star)$ as the benchmark of the loss function does not alter the maximization problem.

For any fixed ω, we can construct the sequence of loss $\{ld_s(\omega)\}_{s=1:t-1}$ and predict the loss at time t using a standard HAR model:

$$ld_t(\omega) = c + \beta^{(1)} ld_{t-1}^{(1)}(\omega) + \beta^{(5)} ld_{t-1}^{(5)}(\omega) + \beta^{(22)} ld_{t-1}^{(22)}(\omega) + u_t, \tag{10}$$

where $ld_{t-1}^{(k)}(\omega) = \frac{1}{k}\sum_{s=1}^{k} ld_{t-s}(\omega)$. Fitting the above model with OLS, we obtain the ω-dependent $\mathcal{F}_{t-1}^{\Sigma}$-adapted parameter estimates $\hat{c}(\omega)$, $\hat{\beta}^{(1)}(\omega)$, $\hat{\beta}^{(5)}(\omega)$, and $\hat{\beta}^{(22)}(\omega)$. Based on the HAR model, the predicted one-day ahead loss at time t is then given by:

$$\hat{\mathrm{E}}[ld_t(\omega)|\mathcal{F}_{t-1}^{\Sigma}] = \hat{c}(\omega) + \hat{\beta}^{(1)}(\omega) ld_{t-1}^{(1)}(\omega) + \hat{\beta}^{(5)}(\omega) ld_{t-1}^{(5)}(\omega) + \hat{\beta}^{(22)} ld_{t-1}^{(22)}(\omega). \tag{11}$$

Setting $\widehat{RU}_t(\omega) \propto \exp(-\mathrm{E}[ld_t(\omega)|\mathcal{F}_{t-1}^{\Sigma}])$,[2] the DWE approach solves the following problem numerically:

$$\hat{\omega}_t^{DWE} = \arg\max_{\omega \in \mathcal{S}} \widehat{RU}_t(\omega) \Leftrightarrow \arg\min_{\omega \in \mathcal{S}} \hat{\mathrm{E}}[ld_t(\omega)|\mathcal{F}_{t-1}^{\Sigma}], \tag{12}$$

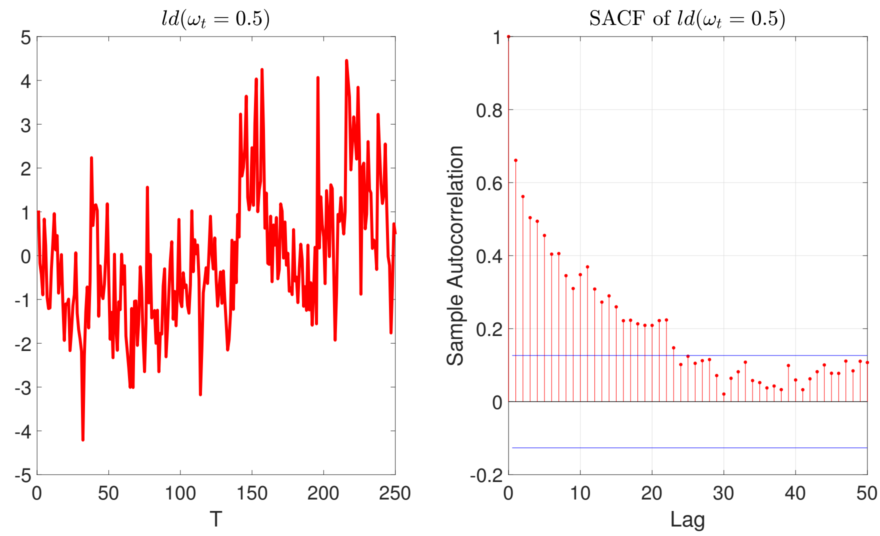
and the rightmost problem can be solved easily by standard gradient-based constraint optimization algorithms, since $\hat{\mathrm{E}}[ld_t(\omega)|\mathcal{F}_{t-1}^{\Sigma}]$ is a continuous function of ω by construction.

The choices of the loss function and the forecasting model require some discussion. First, taking logarithms ensures the positivity of the predicted portfolio variance. Also, we find that a linear prediction model is more appropriate for the log-difference loss than,

[2]Note that this is a naive forecast of $RU_t(\omega)/RU_t(\omega_t^*)$ which is not mean-squared optimal due to the log-transformation. Bårdsen and Lütkepohl (2011) show that the naive forecast is preferable than the mean-squared optimal one in the presence of specification and estimation uncertainty, which rationalizes our choice here.

e.g., a linear difference loss or a ratio-based loss, generating much higher forecasting power.[3] The HAR model is not a requirement of our approach, and in principle any univariate time series prediction model can be used here. The HAR model is chosen due to its computational simplicity (as the predicted loss can be expressed in closed form) and ability to capture long-range dependence commonly observed in empirical volatility time series.

Figure 1: Time series and autocorrelation plots of distance measure for $\omega = 0.5$



The left panel represents the time series of the log distances to realized utility from Eq. (9). The right panel plots the SACF of realized distances from Eq. (9). The considered portfolio is formed of $N = 2$ assets using the data introduced in Section 4. The distances are evaluated at the equally weighted portfolio $(\omega_{1,t} = 0.5, \omega_{2,t} = 0.5)$.

Figure 1 depicts the time series and the sample autocorrelation function (SACF) of the log-distance from Eq. (9), which we chose as a loss function to measure the quality of the portfolio weight vector. The underlying data is introduced in Section 4. The decay in the sample ACF motivates the choice of the HAR model and the time series plot itself motivates the choice of the log-transformation.

The DWE approach provides a very flexible framework to account for various features in the portfolio allocation exercise. For example, one can consider a short-selling constraint

[3]For the GMVP problem, one can alternatively minimize the predicted log portfolio variance, which is numerically identical to the log-difference approach.

for the portfolio weights ω, in the spirit of Fan et al. (2012). Let $b \in [0, 1]$ denote the maximum allowed percentage of short selling for an individual stock, we can restrict the parameter space $\mathcal{S}$ to $\mathcal{S}^b = \{\omega \in \mathcal{S} : \min(\omega) \geq -b\}$, which simply imposes an additional lower bound to ω and is trivial to implement in practice. Multi-step portfolio allocation can also be considered by increasing the forecasting horizon for the HAR model in Eq. (11). Moreover, a penalized DWE approach can be designed to mitigate the impact of transaction costs by adding a regularization term to Eq. (12). We elaborate the design of the regularization and the effectiveness of transaction cost reduction using this feature in Section 4.2. Lastly, the DWE estimator can be easily adapted to other portfolio allocation settings, e.g. the general $\gamma > 0$ case with or without a risk-free asset. One simply needs to construct an appropriate loss function for the corresponding realized utility given a fixed ω, and the DWE estimator can be computed as the weight vector that minimizes the predicted loss.

3 Simulation study

In order to illustrate the advantages of the DWE approach over the plug-in type of weight estimators, we simulate the following HAR-DRD return process r_t with the covariance matrix Σ_t and dynamic conditional correlations:

$$\Sigma_t = D_t R_t D_t, \quad D_t = \begin{pmatrix} \sigma_{11,t} & . & 0 \\ 0 & \ddots & 0 \\ 0 & . & \sigma_{NN,t} \end{pmatrix}, \quad R_t = \begin{pmatrix} 1 & \rho_{12,t} & \cdots & \rho_{1N,t} \\ \rho_{12,t} & 1 & \cdots & \rho_{2N,t} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{1N,t} & \cdot\cdot & \rho_{(N-1)N,t} & 1 \end{pmatrix}, \tag{13}$$

$$\ln \sigma^2_{jj,t} = \beta_{j,0} + \beta_{j,1} \ln \sigma^2_{jj,t-1} + \beta_{j,2} \frac{1}{5} \sum_{i=1}^{5} \ln \sigma^2_{jj,t-i} + \beta_{j,3} \frac{1}{22} \sum_{i=1}^{22} \ln \sigma^2_{jj,t-i} + \varepsilon_{j,t}, \tag{14}$$

$$vech(R_t) = \alpha_0 + \alpha_1 vech(R_{t-1}) + \alpha_2 \frac{1}{5} \sum_{i=1}^{5} vech(R_{t-i}) + \alpha_3 \frac{1}{22} \sum_{i=1}^{22} vech(R_{t-i}) + \varepsilon_t, \tag{15}$$

$$r_t \sim \mathcal{N}(\mu, \Sigma_t), \quad \varepsilon_{j,t} \sim t(\nu_j), \quad \varepsilon_t \sim t(\nu), \tag{16}$$

where $j = 1, ..., N$, the $vech(\cdot)$ operator denotes the vector-form of the lower triangular part of a matrix, and $t(\nu)$ denotes a Student-t distributed random variable with ν degrees of freedom.

We simulate 100 realizations of the HAR model with a maximum N of 93 using Eq. (13) to Eq. (16), where the parameters of the data generating process are calibrated on the S&P100 data described in Section 4 below. For each simulated path, we simulate a series of 1763 observations in total, where the number of observations is chosen to match the empirical analysis of this paper. We construct one-step-ahead forecasts of GMVP portfolio weights based on our DWE approach and several competing estimators use a rolling window approach with daily rebalancing, where the in-sample estimation window length is $T = 1000$ with an out-of-sample evaluation horizon of $H = 63$. The following competing portfolio weight estimators are considered:

1. the plug-in approach with a correctly specified HAR-DRD model, estimated on the history of observed $\Sigma_{1:t}$, where a one-step ahead forecast of the covariance matrix $\hat{\Omega}_t = \hat{D}_{t+1|t}\hat{R}_{t+1|t}\hat{D}_{t+1|t}$ is calculated based on the estimated correctly specified HAR-DRD Eq. (14) - Eq. (15);

2. the plug-in approach with a misspecified HAR-DRD model where the HAR and scalar HAR in Eq. (14) - Eq. (15) are replaced with an AR(1) model;

3. the plug-in approach with a "shrinkage to market" covariance matrix estimated on returns by Ledoit and Wolf (2004), where the last available estimate $\hat{\Omega}_{t-1}$ is used for the weight forecast.

We consider several evaluation metrics to assess the quality of the GMVP portfolio weight vector estimates. First, we compute the expected utility difference between the *ex post* optimal portfolio against the portfolio formed by the estimated weight vector, which is defined as:

$$\mathrm{E}[d(\hat{\omega}_t)] = \hat{\mathrm{E}}[RU_t(\omega_t^\star)] - \hat{\mathrm{E}}[RU_t(\hat{\omega}_t)] = \hat{\mathrm{E}}\left[\frac{1}{H}\sum_{t=1}^{H}\left(\hat{\omega}_t'\Sigma_t\hat{\omega}_t - \frac{1}{\iota'\Sigma_t^{-1}\iota}\right)\right], \tag{17}$$

where $\hat{\mathrm{E}}[\cdot]$ stands for the Monte Carlo simulated expectation. Intuitively, a model with smaller $\mathrm{E}[d(\hat{\omega}_t)]$ on average generates GMVP portfolio with variance closer to the *ex post* optimal one, hence it generates a higher expected utility on average and should be preferred. Second, we compute the root mean squared error (RMSE) of the estimated portfolio variance against the *ex post* optimal portfolio variance:

$$RMSE(\hat{\omega}_t) = \hat{\mathrm{E}}\left[\sqrt{\frac{1}{H}\sum_{t=1}^{H}\left(\hat{\omega}_t'\Sigma_t\hat{\omega}_t - \frac{1}{\iota'\Sigma_t^{-1}\iota}\right)^2}\right]. \tag{18}$$

The RMSE accounts for both the accuracy and the stability of the estimated GMVP portfolio variance. Although the variance of the estimated GMVP portfolio variance does not enter into the investor's utility function, from a forecasting perspective, it is frequently used as a loss function to compare the performance of different predictive models. Third, we evaluate the expected total transaction cost (TTC) of the estimated weight vectors, which measures the practicality of the portfolio allocation methods. This measure is defined as:

$$\mathrm{E}[TTC(\hat{\omega})] = \hat{\mathrm{E}}\Big[c \cdot \sum_{t=2}^{H} \sum_{j=1}^{N} |\hat{\omega}_{j,t} - \hat{\omega}_{j,t^+}|'\iota\Big], \tag{19}$$

where TTC denotes transaction costs at period t, $\hat{\omega}_{j,t^+}$ is the actual (after a price change at t) portfolio weight for asset j before rebalancing at t and c the cost per transaction (50 basis points, DeMiguel et al. (2009)). We present the simulation results in Table 1.

To allow for better interpretation of the results, in Table 1 we use the correct HAR-DRD model as the benchmark and present the evaluation metrics of the other three methods relative to the benchmark. Specifically, any evaluation metric less than one indicates that the method outperforms the benchmark. Concluding from the performance of the benchmark, we see that the correctly specified HAR generates both smaller and more stable GMVP variances as N increases, evidenced by the declining $\mathrm{E}[d(\hat{\omega}_t)]$ and $RMSE(\hat{\omega}_t)$ as a function of N. Unsurprisingly, increasing N also leads to higher transaction costs, as more assets need to be traded to rebalance the portfolio.

Regarding the performance of the misspecified HAR method, we see that misspecifying the original HAR structure by AR(1) leads to a 15-20% increase in $\mathrm{E}[d(\hat{\omega}_t)]$ and $RMSE(\hat{\omega}_t)$ and a more than 50% increase in the transaction cost for almost every N. This result demonstrates the problem of a potentially misspecified forecasting model for the conditional variance-covariance matrix—even a mild misspecification can deteriorate the performance of the resulting GMVP portfolio and may greatly inflate the transaction cost required.

Table 1: Simulation results of GMVP portfolio weight estimators.

N	3	13	23	33	43	53	63	73	83	93
Panel 1: $\mathrm{E}[d(\hat{\omega}_t)] \times 10^6$										
Correct HAR	3.964	1.656	0.902	0.650	0.557	0.461	0.430	0.380	0.350	0.320
Misspecified HAR	1.202	1.150	1.163	1.158	1.153	1.155	1.173	1.172	1.169	1.149
Ledoit Wolf (2004)	5.091	3.258	3.411	4.315	5.322	5.276	5.566	5.352	5.213	5.135
DWE	1.034	1.125	1.127	1.158	1.139	1.147	1.150	1.158	1.166	1.164
Panel 2: $RMSE(\hat{\omega}_t) \times 10^6$										
Correct HAR	7.670	2.009	1.017	1.059	0.594	0.489	0.459	0.401	0.373	0.335
Misspecified HAR	1.295	1.145	1.163	1.143	1.157	1.159	1.199	1.194	1.191	1.164
Ledoit Wolf (2004)	3.991	3.018	3.200	3.476	5.273	5.197	5.481	5.281	5.075	5.014
DWE	0.987	1.111	1.138	1.159	1.147	1.153	1.151	1.159	1.163	1.161
Panel 3: $\mathrm{E}[TTC(\hat{\omega})]$										
Correct HAR	0.043	0.063	0.065	0.065	0.069	0.069	0.071	0.072	0.074	0.077
Misspecified HAR	1.424	1.523	1.529	1.533	1.513	1.526	1.534	1.547	1.549	1.530
Ledoit Wolf (2004)	0.054	0.043	0.039	0.040	0.039	0.039	0.038	0.038	0.038	0.036
DWE	1.059	1.009	0.973	0.958	0.932	0.928	0.912	0.907	0.932	0.940

Numbers in the table correspond to the three simulated evaluation metrics as defined in Eq. (17) to Eq. (19) for the four GMVP portfolio weight estimators averaged over 100 simulation draws with N ranging from 3 to 93. For each estimator and N, the one-step ahead forecast of the portfolio weights is computed over an evaluation horizon of $H = 63$ observations and an in-sample estimation window length $T = 1000$. The original values of the evaluation metrics are presented only for the correct HAR model in blue, which serves as the benchmark for comparison. The evaluation metrics of the three other methods are relative to the benchmark, i.e. they are divided by the corresponding values of the benchmark model.

Comparing the misspecified HAR method to our DWE method, we find that our method almost uniformly dominates the misspecified HAR model for all three evaluation metrics across different N. In terms of $\mathrm{E}[d(\hat{\omega}_t)]$ and $RMSE(\hat{\omega}_t)$, the improvement over the misspecified HAR method is perhaps numerically small and we cannot beat the correct HAR model. However, we stress that for $N \geq 23$, the DWE approach can significantly reduce the transaction cost relative to the misspecified HAR model, and the transaction costs are even smaller than those of the correct model specification. As the forecasting model for the realized covariances is likely to be misspecified in empirical analysis, the simulation provides strong evidence to support the advantage of the DWE method in practice, since it provides both more accurate and more stable estimated GMVP portfolio variances and incurs less transaction costs compared to the misspecified model across different choices of N. This finding is also confirmed in our empirical analysis in Section 4 below.

As to the Ledoit and Wolf (2004) approach, it is clear that this method fails to capture the dynamics of the conditional covariances of the assets, leading to a substantially inflated $\mathrm{E}[d(\hat{\omega}_t)]$ and $RMSE(\hat{\omega}_t)$ relative to the correct model. Being a shrinkage-type estimator, the Ledoit and Wolf (2004) approach enjoys much lower transaction costs than the other three dynamic rebalancing methods as expected. However, we note that one can explicitly include transaction costs as a regularization term in the objective function of the DWE approach to further reduce the transaction cost of the DWE approach, which we discuss in Section 4.2. In detail, we show that a restricted version of the DWE estimator can reduce the transaction cost to a level that is comparable with a shrinkage-based approaches without a substantially inflated portfolio variance.

4 Empirical Evidence

We now evaluate the DWE method and the plug-in competitors based on actual data. We use the stocks contained in the S&P100 together with the SPY index and use daily data from January 2014 until December 2020. We construct realized covariance matrices following the flat-top realized kernel by Varneskov (2016) for the 93 stocks, including SPY, which result in a time series of 1763 observations. Table 8 in the Appendix reports descriptive statistics of all the stocks used in this paper. Daily returns exhibit the expected properties of left skewness and over-kurtosis.

For the commonly used approaches for estimating GMVP weights we first estimate a one-step ahead covariance matrix and then use it in a plug-in formula to calculate the weights $\hat{\omega}_t$. We consider:

1. a sample covariance matrix estimator computed based on T in-sample returns, where the last available estimate is used for the weight forecast;

2. the last available realized covariance matrix of the sample $\hat{\Sigma}_t$, where the last available RC is used for the weight forecast;

3. HAR-DRD model with HAR variance estimates and a sample estimator for the correlations;

4. HAR-DRD model with HAR variance estimates and a scalar HAR for the correlations (Bollerslev et al., 2018b).

We adopt the standard assumption for the rolling window evaluation that the one-step ahead forecast $\hat{\omega}_t$ is used to compute the out-of-sample return for the next period: $\hat{r}^p_{t+1} = \hat{\omega}'_{t+1|t} r_{t+1}$. The estimation window is shifted one period ahead H times resulting in the $H \times 1$ vector of the out-of-sample portfolio returns $\{\hat{r}^p_{t+1}, \ldots, \hat{r}^p_{t+H}\}$.

4.1 DWE

4.1.1 Realized portfolio variance

The DWE approach is explicitly designed to minimize the portfolio variance *realized* on the next day:

$$\widehat{rv}^p_{t+1} = \hat{\omega}'_{t+1|t} \hat{\Sigma}_{t+1} \hat{\omega}_{t+1|t}, \tag{20}$$

where $\hat{\omega}_{t+1|t} = \hat{E}[\omega_{t+1}|\mathcal{F}_t]$ denotes a one-step ahead forecast of portfolio weights and $\hat{\Sigma}_{t+1}$ denotes the realized covariance matrix at day $t + 1$. Therefore, we firstly compare the estimators in terms of the mean realized portfolio variance over the out-of-sample period, which as H grows estimates the unconditional realized utility from Eq. (6):

$$\bar{\hat{rv}}^p = \frac{1}{H}\sum_{h=1}^{H}\hat{rv}^p_{t+h}, \tag{21}$$

where H denotes the number of out-of-sample periods considered.

To compare the performance of different weight estimators across the portfolio dimension we randomly select 100 unique subsets of size N from the pool of 93 assets. We then report the average realized portfolio variance across 100 random portfolios. This guarantees that for a given estimator the portfolio performance comparison is independent of the initial asset selection.

Table 2: Realized portfolio variance for $T = 1000$.

N	Sample cov	RC_{t-1}	HAR-DRD CCC	HAR-DRD DCC	DWE
3	0.000156	0.000149	0.000153	0.000150	**0.000148**
13	0.000105	0.000099	0.000108	0.000099	**0.000092**
23	0.000098	0.000097	0.000107	0.000091	**0.000082**
33	0.000095	0.000105	0.000114	0.000086	**0.000077**
43	0.000094	0.000122	0.000129	0.000081	**0.000074**
53	0.000094	0.000158	0.000162	0.000078	**0.000073**
63	0.000093	0.000228	0.000224	0.000075	**0.000074**
73	0.000093	0.000326	0.000349	0.000074	**0.000072**
83	0.000094	0.000382	0.000511	0.000073	**0.000070**
93	0.000094	0.000358	0.000440	0.000073	**0.000069**

Numbers in the table correspond to the average realized GMVP portfolio variance computed on out-of-sample portfolio returns of portfolios of size N across 100 unique subsets of 93 assets. The last row reports the results over the whole available pool of assets. For each portfolio the realized variance is computed over an evaluation horizon of $H = 763$ observations and an in-sample estimation window length $T = 1000$. Numbers in bold correspond to the smallest realized portfolio variance for a given portfolio size N.

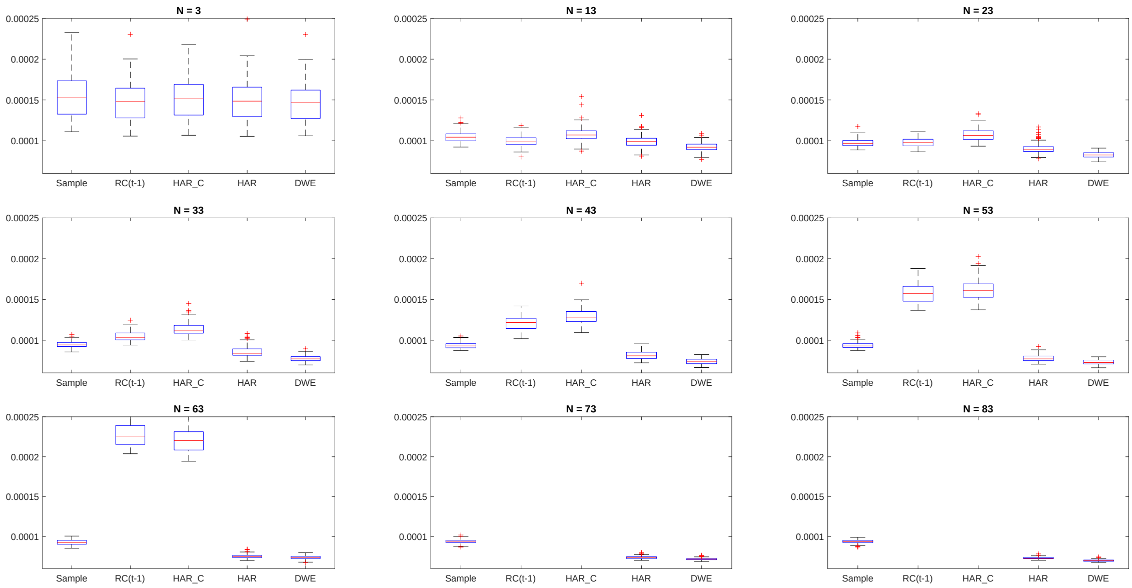
Table 2 reports the average realized portfolio variance $\bar{\hat{rv}}^p$ across 100 randomly formed portfolios of size N for an in-sample estimation window of approximately 4 years and the out-of-sample evaluation horizon H of 3 years. The last row corresponds to the mean realized portfolio variance evaluated on the whole pool of 93 assets. For a given portfolio size N there is a clear hierarchy between HAR-DRD models and the proposed DWE approach: introducing dynamics into the conditional correlation matrix certainly facilitates the decrease in the out-of-sample portfolio variance as the HAR-DCC model dominates the HAR-CCC with a constant conditional correlation matrix. Moreover, the proposed DWE approach outperforms both HAR models. In fact, across all the considered portfolio sizes N the proposed DWE approach results in the lowest mean realized daily portfolio variance. Moreover, the realized portfolio variance decreases with the dimension N, indicating the benefits of diversification. For the HAR-DRD with CCC model and the plug-in strategy with the previous realized correlation RC_{t-1} the increase in the portfolio size for $N > 43$ results in the increase of the portfolio variance, indicating that diversification gains are suppressed by the increase in modelling errors.

The proposed DWE indirect weight estimation approach outperforms the competitors not only in mean, but also across the 100 randomly drawn portfolios. Figure 2 depicts the boxplots of average realized portfolio variances reflecting the distribution across randomly formed portfolios. Each panel in the figure corresponds to a portfolio size N and each boxplot corresponds to a weight estimation strategy and reflects the distribution of the out-of-sample portfolio variance across 100 randomly selected portfolios. The improvements of the DWE become more obvious with the increased portfolio dimension N, e.g. for $N \geq 43$ the largest average realized portfolio variance of DWE is lower than the smallest one of all the other approaches but the HAR-DRD with a DCC specification, which is outperformed in mean.

4.1.2 Variance of out-of-sample returns

We now consider a commonly used portfolio performance measure, namely the out-of-sample portfolio variance, which is computed in a rolling window of $T = 1000$ in-sample observations. In the absence of realized covariances, different portfolio strategies are then

Figure 2: Boxplots of the realized portfolio variance for $T = 1000$.



Boxplots of the out-of-sample portfolio variances computed on net portfolio returns over a 100 randomly drawn portfolios of size N. For each randomly drawn portfolio the realized portfolio variance is computed over an evaluation horizon of $H = 763$ observations and an in-sample estimation window length $T = 1000$. X-axes denote different ways of computing portfolio weights: Sample denotes the plug-in weight estimator as in Eq. (7) with a sample covariance matrix; RC_{t-1} for plug-in approach with the last available realized correlation matrix; HAR_C and HAR are the plug-in weights with HAR-DRD approaches by Bollerslev et al. (2018b) with constant and dynamic correlations respectively; DWE for the proposed approach. Note, the limits of the y-axes are fixed for comparison, thus for $N \geq 73$ the boxplots for RC_{t-1} and HAR_C are no longer visible.

evaluated based on the out-of-sample variance $\hat{\sigma}^2_{os}$ of portfolio returns given by:

$$\hat{\sigma}^2_{os} = \frac{1}{H-1} \sum_{h=1}^{H} \left(\hat{r}^p_{t+h} - \hat{\mu}_{os} \right)^2, \tag{22}$$

$$\text{where:} \qquad \hat{\mu}_{os} = \frac{1}{H} \sum_{h=1}^{H} \hat{r}^p_{t+h} = \frac{1}{H} \sum_{h=1}^{H} \hat{\omega}'_{t+h|t} r_{t+h}.$$

Table 3 reports the average out-of-sample variances of portfolios of different sizes N for an estimation window of approximately 4 years and the out-of-sample evaluation horizon H of 3 years. The last row corresponds to the out-of-sample portfolio return variance evaluated on the whole pool of 93 assets. The proposed DWE approach results in the smallest out-of-sample portfolio variance for all portfolio sizes N.

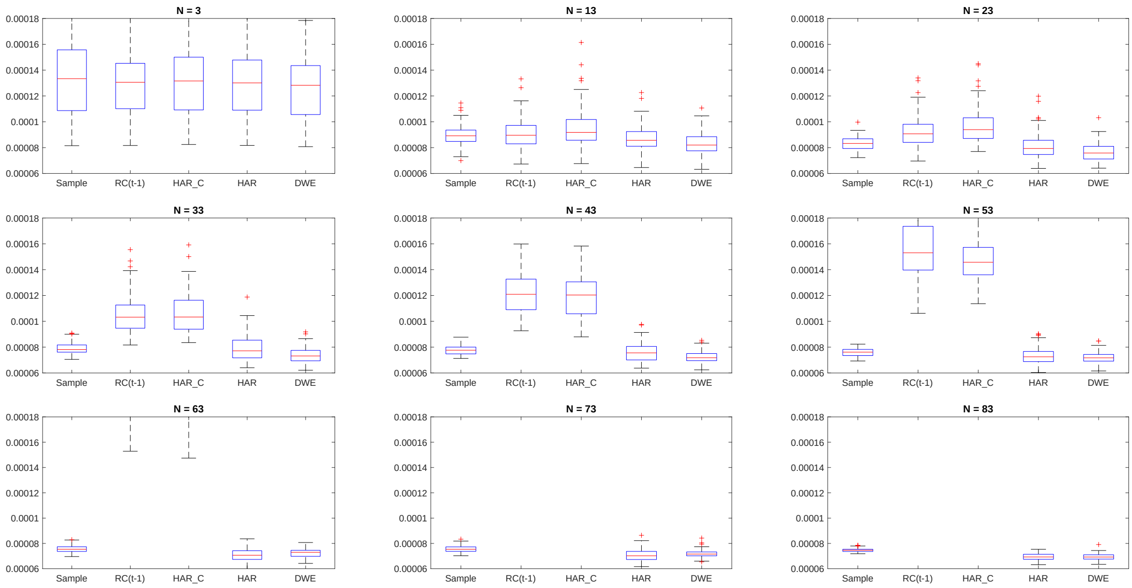
Figure 3 provides more insights into the differences in the performance of the considered weight estimators. The visual comparison of the boxplots across the portfolio dimension

Table 3: Variance of out-of-sample portfolio returns for $T = 1000$.

N	Sample cov	RC_{t-1}	HAR-DRD CCC	HAR-DRD DCC	DWE
3	0.000136	0.000131	0.000132	0.000130	**0.000129**
13	0.000090	0.000091	0.000095	0.000088	**0.000083**
23	0.000083	0.000093	0.000097	0.000082	**0.000076**
33	0.000079	0.000105	0.000106	0.000079	**0.000074**
43	0.000078	0.000123	0.000122	0.000076	**0.000072**
53	0.000076	0.000157	0.000150	0.000073	**0.000072**
63	0.000075	0.000215	0.000205	0.000072	**0.000071**
73	0.000075	0.000316	0.000308	0.000072	**0.000070**
83	0.000075	0.000369	0.000400	0.000070	**0.000069**
93	0.000075	0.000355	0.000426	0.000069	**0.000067**

Numbers in the table correspond to the average out-of-sample GMVP portfolio variance computed on out-of-sample portfolio returns of portfolios of size N across 100 unique subsets of 93 assets. The last row reports the results over the whole available pool of assets. For each portfolio the out-of-sample return is computed over an evaluation horizon of $H = 763$ observations and an in-sample estimation window length $T = 1000$. Numbers in bold correspond to the smallest out-of-sample portfolio variance for a given portfolio size N.

Figure 3: Boxplots of the variance of the out-of-sample portfolio returns for $T = 1000$.



Boxplots of the portfolio variances computed on out-of-sample portfolio returns over a 100 randomly drawn portfolios of size N. For each randomly drawn portfolio the out-of-sample portfolio variance is computed over an evaluation horizon of $H = 763$ observations and an in-sample estimation window length $T = 1000$. X-axes denote different ways of computing portfolio weights: Sample denotes the plug-in weight estimator as in Eq. (7) with a sample covariance matrix; RC_{t-1} for plug-in approach with the last available realized correlation matrix; HAR_C and HAR are the plug-in weights with HAR-DRD approaches by Bollerslev et al. (2018b) with constant and dynamic correlations respectively; DWE for the proposed approach. Note, the limits of the y-axes are fixed for comparison, thus for $N \geq 73$ the boxplots for RC_{t-1} and HAR_C are no longer visible.

N reveals that the difference in the performance of the estimators becomes apparent for larger portfolios $N \geq 33$. Similarly to the distribution of the realized portfolio variance, in terms of the variance of the out-of-sample portfolio returns the closest competitor for the DWE approach is the HAR-DRD model with dynamic structure on the conditional correlation matrix. However, even for the larger portfolio dimensions $N \geq 63$, when the medians of DWE and HAR are close to each other, the distribution of portfolio variance of the DWE has a lower interquartile range compared to HAR, which indicates that the performance of the DWE is more robust to the assets selection compared to HAR.

4.1.3 Transaction costs

We now compare different weight estimation strategies in terms of another empirically relevant metric, namely the transaction costs incurred by dynamically rebalancing the portfolio, which is also used in our simulation. For each randomly drawn portfolio, we compute the transaction cost (TC) as:

$$TC_t = c \cdot \sum_{j=1}^{N} |\hat{\omega}_{j,t} - \hat{\omega}_{j,t^+}|, \tag{23}$$

where TC_t denotes transaction costs at period t, $\hat{\omega}_{j,t^+}$ the actual (after a price change at t) portfolio weight before rebalancing at t and c the cost per transaction (50 basis points, DeMiguel et al. (2009)). For every randomly drawn portfolio we compute the total transaction costs over the out-of-sample period as a sum $\sum_{h=2}^{H} TC_h$:

Tables 4 reports the average TTC across 100 randomly formed portfolios of size N for an in-sample estimation windows of $T = 1000$. The larger the asset space, the more expensive is the daily rebalancing of the portfolio. Similarly to the comparisons of the out-of-sample portfolio variance, there is a clear hierarchy between the models. Introducing dynamics into the conditional correlation matrix for the plug-in approach leads to a reduction in turnover costs. The indirect DWE weight estimation dominates both of the HAR-DRD model specifications and produces lower transaction costs. However, a

Table 4: Transaction costs for $T = 1000$.

N	Sample cov	RC_{t-1}	HAR-DRD CCC	HAR-DRD DCC	DWE
3	0.027	1.153	0.771	0.517	0.451
13	0.063	3.777	3.481	1.513	1.276
23	0.104	6.618	6.385	2.334	2.029
33	0.150	10.021	9.834	3.168	2.759
43	0.193	14.229	14.039	3.861	3.343
53	0.241	19.819	19.534	4.576	3.806
63	0.289	26.786	26.269	5.232	4.030
73	0.335	33.901	33.150	5.799	4.258
83	0.386	38.486	38.592	6.410	5.225
93	0.437	48.595	47.626	6.978	4.661

Numbers in the table correspond to the average total transaction costs computed according to Eq. (23) for portfolios of size N across 100 unique subsets of 93 assets. The last row reports the results over the whole available pool of assets. For each portfolio the TTC is computed over an evaluation horizon of $H = 763$ observations and an in-sample estimation window length $T = 1000$.

simple sample covariance matrix estimator appears to result in more stable weights over time reflected by low turnover costs. Even though the plug-in weight estimation with a static forecast of sample covariance matrix results into a higher realized variance and out-of-sample return variance, it might be beneficial for the investor to control the amount of rebalancing.

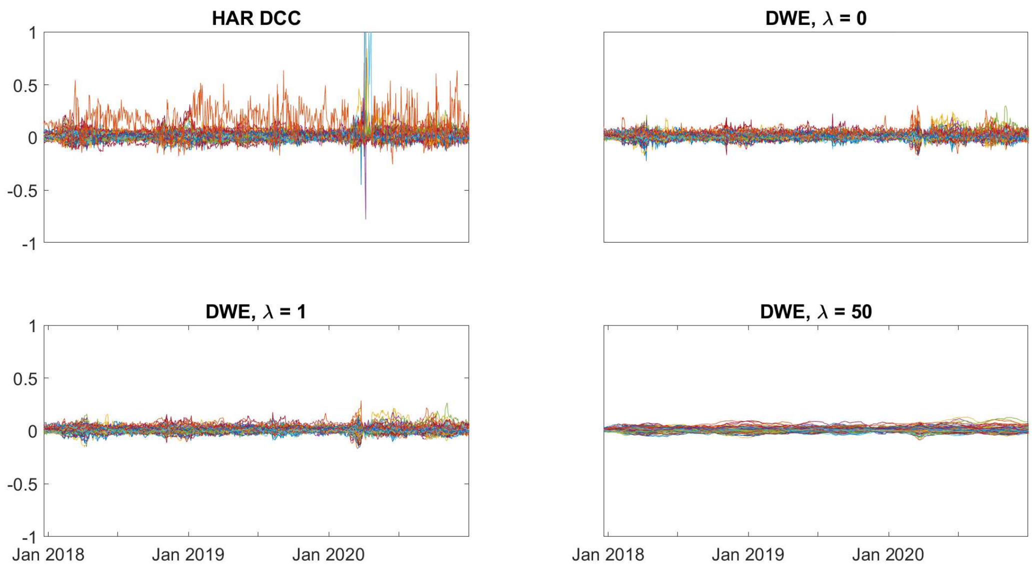
4.2 Restricted DWE

To showcase the flexibility of the proposed direct weight estimation approach we introduce a restricted DWE estimator, which imposes an ℓ_2-norm on the difference between the current and the previous portfolio weight:

$$\hat{\omega}_{t,k} = \arg\min_{\hat{\omega}_{\cdot,k}} \hat{\mathrm{E}}_{t-1}[ld(\hat{\omega}_{t,k})] + \lambda \sum_{i=1}^{N} (\hat{\omega}_{t,k} - \hat{\omega}_{t-1,k})^2, \tag{24}$$

where λ is a tuning parameter which controls the stability of the weights over time: the larger λ, the more stable the weight forecasts across time, which results in lower transaction costs. $\hat{\mathrm{E}}_{t-1}[ld(\hat{\omega}_{t,k})]$ for this example denotes a one-step ahead HAR-forecast of the log-distance: $\hat{\alpha}_0 + \hat{\alpha}_1 ld(\hat{\omega}_{t-1,k}) + \hat{\alpha}_2 \frac{1}{5}\sum_{j=1}^{5} ld(\hat{\omega}_{t-j,k}) + \hat{\alpha}_3 \frac{1}{21}\sum_{j=1}^{21} ld(\hat{\omega}_{t-j,k})$. We utilise the same dataset and as above for a given portfolio size N we form 100 unique portfolios out of the whole pool of the assets. The reported results are averages over the 100 randomly drawn portfolios.

Figure 4: Time series plots of HAR-DCC weight estimates and DWE with $\lambda = 0, 1, 50$.



Different panels on the plot correspond to the time series of estimated weights over the 3 years of the out-of-sample period for $N = 93$. The upper left panel corresponds to the unrestricted HAR DCC weight estimator and other panels correspond to the DWE with different values of the the tuning parameter λ.

Figure 4 plots the time series of the estimated weights for unrestricted estimators (the upper panels) and restricted DWE estimators (lower panels). As reported in Table 4 the estimated weights of unrestricted DWE estimator with $\lambda = 0$ are more stable over time compared to HAR DCC weights which corresponds to the lower amount of transaction costs. With the increase in the tuning parameter λ, the estimated weights become more stable over time resulting in the reduction of turnover costs. The trade-off the investor faces is between the out-of-sample portfolio risk and the turnover costs of daily rebalancing, which for the restricted DWE estimator is controlled by the tuning parameter λ.

We now compare our restricted estimator to the standard methods, which aim at the intertemporal weight stabilization:

1. Plug-in estimator with a shrinkage to market covariance matrix estimated on T in-sample returns by Ledoit and Wolf (2004), where the last available estimate $\hat{\Sigma}_t$ is used for the weight forecast;

2. Plug-in estimator with non-linear eigenvalue shrinkage of the covariance matrix estimated on T in-sample returns by Ledoit and Wolf (2020), where the last available estimate $\hat{\Sigma}_t$ is used for the weight forecast;

3. Equally weighted portfolio, which is rebalanced daily.

Table 5 reports the mean realized portfolio variance and Table 6 reports the mean out-of-sample return variance for the restricted estimators.

Table 5: Realized portfolio variance for $T = 1000$.

N	Ledoit and Wolf 2004	Ledoit and Wolf 2020	DWE $\lambda = 0.1$	DWE $\lambda = 0.5$	DWE $\lambda = 1$	DWE $\lambda = 10$	DWE $\lambda = 50$	Equally weighted
3	0.000156	0.000156	**0.000149**	**0.000149**	0.000150	0.000153	0.000155	0.000166
13	0.000103	0.000104	**0.000093**	**0.000093**	0.000094	0.000097	0.000099	0.000116
23	0.000095	0.000097	**0.000082**	**0.000082**	**0.000082**	0.000086	0.000089	0.000109
33	0.000091	0.000094	0.000077	**0.000076**	**0.000076**	0.000080	0.000084	0.000107
43	0.000089	0.000092	0.000074	**0.000073**	**0.000073**	0.000078	0.000082	0.000105
53	0.000088	0.000092	0.000073	**0.000072**	**0.000072**	0.000078	0.000080	0.000104
63	0.000087	0.000090	**0.000074**	**0.000074**	0.000075	0.000077	0.000078	0.000104
73	0.000086	0.000091	**0.000072**	**0.000072**	**0.000072**	0.000073	0.000075	0.000103
83	0.000085	0.000089	**0.000070**	**0.000070**	**0.000070**	0.000071	0.000072	0.000103
93	0.000085	0.000089	**0.000069**	**0.000069**	**0.000069**	0.000070	0.000071	0.000103

Numbers in the table correspond to the average realized GMVP portfolio variance computed on out-of-sample portfolio returns of portfolios of size N across 100 unique subsets of 93 assets. The last row reports the results over the whole available pool of assets. For each portfolio the realized variance is computed over an evaluation horizon of $H = 763$ observations and an in-sample estimation window length $T = 1000$. Numbers in bold correspond to the smallest realized portfolio variance for a given portfolio size N.

We consider several values of the tuning parameter λ to determine how much would the realized portfolio variance and the variance of the out-of-sample portfolio returns change with the increase in the tuning parameter. For both tables the numbers in bold indicate the smallest portfolio risk for a given size N across the models. For the realized portfolio variance the extremely large penalties λ of the DWE estimator increase the portfolio risk, whereas for the out-of-sample portfolio variance the increase in λ seems to be even rewarding for larger portfolio sizes. Notably, the restricted DWE estimator outperforms both shrinkage approaches and the equally weighted portfolio independently of the tuning parameter choice.

Table 6: Variance of out-of-sample portfolio returns for $T = 1000$.

N	Ledoit and Wolf 2004	Ledoit and Wolf 2020	DWE $\lambda = 0.1$	DWE $\lambda = 0.5$	DWE $\lambda = 1$	DWE $\lambda = 10$	DWE $\lambda = 50$	Equally weighted
3	0.000136	0.000136	**0.000130**	0.000130	0.000131	0.000133	0.000136	0.000147
13	0.000089	0.000090	**0.000084**	**0.000084**	**0.000084**	0.000085	0.000087	0.000104
23	0.000083	0.000083	0.000077	**0.000076**	**0.000076**	0.000077	0.000078	0.000098
33	0.000078	0.000079	0.000074	0.000074	0.000073	**0.000072**	0.000074	0.000096
43	0.000077	0.000077	0.000072	**0.000071**	**0.000071**	**0.000071**	0.000072	0.000094
53	0.000075	0.000075	0.000072	0.000071	**0.000070**	**0.000070**	0.000071	0.000093
63	0.000074	0.000075	0.000072	0.000072	0.000072	**0.000070**	**0.000070**	0.000093
73	0.000073	0.000074	0.000071	0.000071	0.000070	**0.000068**	**0.000068**	0.000093
83	0.000072	0.000073	0.000069	0.000068	0.000068	**0.000064**	0.000066	0.000092
93	0.000072	0.000073	0.000067	0.000066	0.000065	**0.000064**	0.000066	0.000092

Numbers in the table correspond to the average out-of-sample GMVP portfolio variance computed on out-of-sample portfolio returns of portfolios of size N across 100 unique subsets of 93 assets. The last row reports the results over the whole available pool of assets. For each portfolio the out-of-sample return is computed over an evaluation horizon of $H = 763$ observations and an in-sample estimation window length $T = 1000$. Numbers in bold correspond to the smallest out-of-sample portfolio variance for a given portfolio size N.

Table 7 reports the average total transaction costs for the restricted estimator. The penalisation of ℓ_2-norm on the intertemporal difference of the portfolio weights reduces the amount of transaction costs to the level of the shrinkage estimators without paying a huge price in terms of the out-of-sample portfolio risk. The exact choice of λ depends however on the individual preferences and the utility function of the investor.

Tables in Appendix A.2 and A.3 provide robustness checks of the same analysis as above but for a shorter sample size $T = 250$. We find that the proposed direct weight estimation approach performs similarly well for smaller sample sizes. The restricted version still outperforms its shrinkage competitors in terms of the out-of-sample investment risk, keeping the transaction cost level low at the same time.

Table 7: Transaction costs for $T = 1000$.

N	Ledoit and Wolf 2004	Ledoit and Wolf 2020	DWE $\lambda = 0.1$	DWE $\lambda = 0.5$	DWE $\lambda = 1$	DWE $\lambda = 10$	DWE $\lambda = 50$	Equally weighted
3	0.027	0.027	0.386	0.276	0.212	0.060	0.030	0.029
13	0.056	0.061	1.159	0.906	0.740	0.235	0.079	0.033
23	0.088	0.099	1.851	1.452	1.192	0.391	0.129	0.034
33	0.121	0.140	2.492	1.939	1.588	0.523	0.174	0.034
43	0.153	0.177	3.045	2.353	1.920	0.634	0.213	0.034
53	0.186	0.216	3.430	2.648	2.162	0.728	0.249	0.034
63	0.219	0.253	3.669	2.888	2.387	0.835	0.287	0.034
73	0.251	0.294	3.977	3.244	2.746	1.008	0.345	0.034
83	0.282	0.327	4.948	4.137	3.544	1.306	0.443	0.034
93	0.314	0.364	4.402	3.717	3.202	1.221	0.427	0.034

Numbers in the table correspond to the average total transaction costs computed according to Eq. (23) for portfolios of size N across 100 unique subsets of 93 assets. The last row reports the results over the whole available pool of assets. For each portfolio the TTC is computed over an evaluation horizon of $H = 763$ observations and an in-sample estimation window length $T = 1000$.

5 Conclusions

This paper introduces a novel direct portfolio weight estimation approach (DWE), which is particularly beneficial for higher dimensional portfolios. The main contribution of the proposed weight estimation is the reduction of the multidimensional forecasting into a one-dimensional problem using realized utility.

Realized covariance matrices are used to construct a time series of previously optimal portfolio variances and the optimal weight vector forecast is recovered directly from the series of realized portfolio variances through a constrained optimization. This guarantees that the DWE maximises the predicted realized utility, an estimator of expected utility. The main advantage of the proposed direct weight estimation approach is the mitigation of the curse of the dimensionality problem. For the method to function we only require a univariate time series forecasting model, and thus the number of parameters to be estimated does not depend on the number of assets in the portfolio. In the empirical application we have illustrated that a simple HAR model is a good enough forecasting tool for the proposed estimator to outperform the main competitors which are based on both realized covariance forecasting and shrinkage of the covariance matrices.

The proposed method is extremely flexible and can be adapted to any portfolio performance measure of interest which can be constructed ex-post using realized returns. We have demonstrated that the DWE can be easily extended to control for the amount of portfolio rebalancing, which leads to a reduction of transaction costs. And as a potential improvement of the method one could investigate whether the univariate forecasting model used in the proposed method can be improved upon with machine learning techniques.

References

Aït-Sahalia, Y. and J. Jacod (2014): *High-frequency financial econometrics*, Princeton University Press.

Bårdsen, G. and H. Lütkepohl (2011): "Forecasting levels of log variables in vector autoregressions," *International Journal of Forecasting*, 27, 1108–1115.

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

Bollerslev, T., B. Hood, J. Huss, and L. H. Pedersen (2018a): "Risk Everywhere: Modeling and Managing Volatility," *The Review of Financial Studies*, 31, 2729–2773.

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

Brodie, J., I. Daubechies, C. D. Mol, D. Giannone, and I. Loris (2009): "Sparse and stable Markowitz portfolios," *Proceedings of the National Academy of Sciences*, 106, 12267–12272, PMID: 19617537.

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

DeMiguel, V., L. Garlappi, and R. Uppal (2009): "Optimal Versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy?" *Review of Financial Studies*, 22, 1915–1953.

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

Engle, R. F., O. Ledoit, and M. Wolf (2019): "Large dynamic covariance matrices," *Journal of Business & Economic Statistics*, 37, 363–375.

Fan, J., J. Zhang, and K. Yu (2012): "Vast Portfolio Selection With Gross-Exposure Constraints," *Journal of the American Statistical Association*, 107, 592–606.

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

Goto, S. and Y. Xu (2015): "Improving Mean Variance Optimization through Sparse Hedging Restrictions," *Journal of Financial and Quantitative Analysis*, 50, 1415–1441.

Hafner, C. M. and O. Linton (2010): "Efficient estimation of a multivariate multiplicative volatility model," *Journal of Econometrics*, 159, 55–73.

Hautsch, N., L. M. Kyj, and P. Malec (2015): "Do high-frequency data improve high-dimensional portfolio allocations?" *Journal of Applied Econometrics*, 30, 263–290.

Jagannathan, R. and T. Ma (2003): "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps," *The Journal of Finance*, 58, 1651–1684.

Jin, X. and J. M. Maheu (2013): "Modeling realized covariances and returns," *Journal of Financial Econometrics*, 11, 335–369.

Ledoit, O. and M. Wolf (2004): "Honey, I Shrunk the Sample Covariance Matrix," *The Journal of Portfolio Management*, 30, 110–119.

——— (2014): "Nonlinear Shrinkage of the Covariance Matrix for Portfolio Selection: Markowitz Meets Goldilocks," Working paper, Department of Economics, University of Zurich.

——— (2020): "Analytical nonlinear shrinkage of large-dimensional covariance matrices," *Annals of Statistics*, 48, 3043–3065.

Li, J. (2015): "Sparse and Stable Portfolio Selection with Parameter Uncertainty," *Journal of Business & Economic Statistics*, 33, 381–392.

Tse, Y. K. and A. K. C. Tsui (2002): "A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations," *Journal of Business & Economic Statistics*, 20, 351–362.

Varneskov, R. T. (2016): "Flat-top realized kernel estimation of quadratic covariation with nonsynchronous and noisy asset prices," *Journal of Business & Economic Statistics*, 34, 1–22.

Yen, Y.-M. (2015): "Sparse Weighted-Norm Minimum Variance Portfolios," *Review of Finance*, 20, 1259–1287.

A Appendix

A.1 Descriptive statistic

Table 8: Descriptive statistic

	Returns			Realized Covariance	
Ticker	Mean	Skewness	Kurtosis	Average variance	Average correlation
AAPL	0.0005	-0.1497	6.4232	0.0002	0.2702
ABBV	0.0000	-0.4049	7.0031	0.0003	0.1944
ABT	0.0003	-0.4526	8.9318	0.0002	0.2904
ACN	0.0006	0.4866	12.9030	0.0002	0.3233
ADBE	0.0005	-0.3117	6.5015	0.0002	0.2304
AIG	-0.0002	-0.5365	14.6920	0.0003	0.2718
ALL	0.0002	0.0173	11.0979	0.0002	0.3294
AMGN	0.0000	0.1813	6.6015	0.0002	0.2114
AMT	0.0002	0.2202	11.4673	0.0002	0.2007
AMZN	-0.0002	-0.3121	5.8850	0.0002	0.2433
AXP	-0.0002	0.4598	15.1984	0.0002	0.3295
BA	-0.0005	-0.9317	14.6866	0.0003	0.2414
BAC	-0.0002	-0.0912	7.4706	0.0002	0.2859
BK	0.0000	-0.2275	6.3893	0.0002	0.2954
BLK	-0.0002	-0.2301	12.3504	0.0002	0.2738
BMY	-0.0002	-0.4521	6.3678	0.0002	0.2166
BRK_B	-0.0004	-0.2212	7.8560	0.0001	0.4127
C	-0.0003	-0.1221	8.5162	0.0003	0.2720
CAT	0.0002	-0.2360	7.1445	0.0002	0.2476
CHTR	0.0005	-0.0459	5.9367	0.0003	0.1569
CL	0.0003	0.2390	11.8459	0.0001	0.2715
CMCS_A	0.0002	0.1483	8.0642	0.0002	0.2461
COF	-0.0002	-0.3628	10.3964	0.0003	0.2749
COP	-0.0004	0.3319	8.3627	0.0004	0.1729
COST	0.0003	-0.0652	8.0229	0.0001	0.2763
CRM	0.0001	-0.0603	6.9354	0.0003	0.1965
CSCO	0.0003	0.0397	10.2466	0.0002	0.3086
CVS	-0.0003	-0.0705	6.5906	0.0002	0.2458
CVX	-0.0004	-1.0226	23.3743	0.0002	0.2430
DHR	0.0001	-0.6288	11.8489	0.0001	0.3196
DIS	-0.0003	-0.3321	11.0017	0.0002	0.2985
DUK	0.0002	-0.1082	13.7510	0.0001	0.1564
EMR	-0.0001	-0.0596	14.1665	0.0002	0.2931
EXC	0.0003	0.3519	14.1422	0.0002	0.1590
F	-0.0012	-0.0253	7.2240	0.0003	0.2412
FB	0.0003	-0.1495	5.4549	0.0003	0.2130
FDX	-0.0001	0.1376	9.8415	0.0002	0.2506
GD	0.0000	-0.0323	7.3266	0.0002	0.2782
GE	-0.0009	-0.0106	8.2354	0.0003	0.2557
GILD	-0.0008	0.0066	6.3041	0.0003	0.1896
GM	-0.0009	-0.0418	6.8607	0.0003	0.2176
GOOG	0.0001	-0.5207	5.6488	0.0002	0.2812
GS	0.0000	0.0049	7.7132	0.0002	0.2786
HD	0.0003	-0.3000	10.0399	0.0002	0.2997
HON	0.0000	0.0485	9.9954	0.0002	0.3560
IBM	-0.0001	-0.1310	7.4394	0.0001	0.3340
INTC	0.0007	0.5830	9.8934	0.0002	0.2546
JNJ	0.0000	-0.9633	11.6151	0.0001	0.3004
JPM	0.0000	0.0840	6.7562	0.0002	0.3257

KMI	-0.0009	-0.5889	12.7430	0.0003	0.1861
KO	0.0001	-0.7802	13.3800	0.0001	0.2868
LLY	0.0004	0.2805	8.6657	0.0002	0.2204
LMT	-0.0001	0.1999	14.5327	0.0002	0.2625
LOW	0.0002	-1.2937	18.4849	0.0002	0.2557
MA	0.0000	-0.5586	11.0735	0.0002	0.2944
MCD	0.0003	0.7493	15.5476	0.0001	0.2714
MDLZ	0.0001	0.1956	6.3841	0.0002	0.2613
MDT	-0.0002	-0.3879	6.5070	0.0002	0.2857
MET	-0.0002	-0.3230	9.9861	0.0002	0.2721
MMM	-0.0001	-0.6260	12.7065	0.0001	0.3403
MO	0.0001	-0.8173	13.3655	0.0002	0.2160
MRK	-0.0003	-0.1954	7.7219	0.0002	0.2658
MS	-0.0001	0.1162	7.0550	0.0003	0.2649
MSFT	0.0004	-0.2448	6.6762	0.0002	0.2858
NEE	0.0005	0.3198	11.8549	0.0002	0.1541
NFLX	0.0003	0.1275	5.4112	0.0004	0.1444
NKE	0.0001	-0.0303	7.1928	0.0002	0.2527
NVDA	0.0006	-0.2058	6.5883	0.0004	0.1682
ORCL	0.0004	0.3310	9.7702	0.0002	0.3085
PEP	0.0003	-0.7301	26.4735	0.0001	0.2796
PFE	-0.0004	-0.4543	7.7439	0.0002	0.2667
PG	0.0003	-0.0283	12.3701	0.0001	0.2641
PM	0.0002	-0.2862	9.5891	0.0002	0.2205
QCOM	0.0002	1.1946	23.3548	0.0002	0.2357
SBUX	0.0002	-0.1219	7.5749	0.0002	0.2678
SLB	-0.0007	0.0977	8.1619	0.0004	0.1684
SO	0.0005	0.7008	23.5597	0.0001	0.1550
SPG	-0.0009	-2.0751	36.4763	0.0003	0.1552
T	-0.0003	-0.2502	7.3867	0.0001	0.2741
TGT	0.0000	0.0721	7.3790	0.0002	0.2114
TMO	0.0002	-0.4640	7.0247	0.0002	0.2597
TSLA	0.0003	0.3980	6.0261	0.0007	0.1170
TXN	0.0005	0.0199	6.9090	0.0002	0.2678
UNH	0.0002	-0.0329	8.4897	0.0002	0.2192
UNP	0.0003	0.0722	6.3295	0.0002	0.2471
UPS	0.0001	0.5287	10.3149	0.0002	0.3016
USB	0.0000	0.0795	12.3455	0.0002	0.3282
V	0.0000	-0.1663	7.5423	0.0002	0.3068
VZ	0.0000	0.2728	7.2173	0.0001	0.2516
WFC	-0.0003	-0.3321	10.1079	0.0002	0.3039
WMT	0.0003	-0.0624	13.9357	0.0001	0.2544
XOM	-0.0005	-0.1676	7.7057	0.0002	0.2665
SPY	0.0001	-0.5152	9.9201	0.0001	0.7513

The table reports the sample moments of the returns used in Section 4 and time series average realized variances and correlations with the other assets.

A.2 Unrestricted estimators, T = 250

Table 9: Realized portfolio variance for $T = 250$.

N	Sample cov	RC_{t-1}	HAR-DRD CCC	HAR-DRD DCC	DWE
3	0.0000662	0.0000648	0.0000656	0.0000643	**0.0000641**
13	0.0000392	0.0000382	0.0000408	0.0000360	**0.0000357**
23	0.0000366	0.0000369	0.0000403	0.0000321	**0.0000316**
33	0.0000370	0.0000391	0.0000427	0.0000300	**0.0000298**
43	0.0000382	0.0000443	0.0000484	0.0000288	**0.0000285**
53	0.0000406	0.0000541	0.0000579	0.0000282	**0.0000281**
63	0.0000422	0.0000676	0.0000705	0.0000276	**0.0000275**
73	0.0000471	0.0000769	0.0000798	0.0000273	**0.0000272**
83	0.0000504	0.0000775	0.0000828	**0.0000269**	0.0000274
93	0.0000556	0.0001350	0.0001309	**0.0000266**	0.0000269

Numbers in the table correspond to the average realized GMVP portfolio variance computed on out-of-sample portfolio returns of portfolios of size N across 100 unique subsets of 93 assets. The last row reports the results over the whole available pool of assets. For each portfolio the realized variance is computed over an evaluation horizon of $H = 763$ observations and an in-sample estimation window length $T = 250$. Numbers in bold correspond to the smallest realized portfolio variance for a given portfolio size N.

Table 10: Variance of out-of-sample portfolio returns for $T = 250$.

N	Sample cov	RC_{t-1}	HAR-DRD CCC	HAR-DRD DCC	DWE
3	**0.000058**	**0.000058**	**0.000058**	**0.000058**	**0.000058**
13	0.000036	0.000036	0.000038	**0.000035**	0.000036
23	0.000034	0.000034	0.000037	**0.000033**	**0.000033**
33	0.000033	0.000036	0.000039	**0.000031**	**0.000031**
43	0.000033	0.000040	0.000043	**0.000031**	**0.000031**
53	0.000034	0.000046	0.000048	**0.000030**	**0.000030**
63	0.000034	0.000054	0.000055	0.000030	**0.000029**
73	0.000035	0.000061	0.000060	0.000030	**0.000029**
83	0.000036	0.000064	0.000062	**0.000029**	0.000030
93	0.000037	0.000119	0.000114	0.000029	**0.000027**

Numbers in the table correspond to the average out-of-sample GMVP portfolio variance computed on out-of-sample portfolio returns of portfolios of size N across 100 unique subsets of 93 assets. The last row reports the results over the whole available pool of assets. For each portfolio the out-of-sample return is computed over an evaluation horizon of $H = 763$ observations and an in-sample estimation window length $T = 250$. Numbers in bold correspond to the smallest out-of-sample portfolio variance for a given portfolio size N.

Table 11: Transaction costs for $T = 250$.

N	Sample cov	RC_{t-1}	HAR-DRD CCC	HAR-DRD DCC	DWE
3	0.040	1.204	0.723	0.490	0.471
13	0.155	3.617	3.218	1.335	1.446
23	0.284	6.140	5.784	2.030	2.227
33	0.434	9.006	8.662	2.714	2.841
43	0.592	12.310	11.925	3.274	3.311
53	0.771	16.251	15.756	3.842	3.649
63	0.966	20.565	19.907	4.337	3.815
73	1.190	24.479	23.690	4.869	4.304
83	1.427	27.266	26.424	5.295	5.533
93	1.691	41.147	40.055	5.731	4.941

Numbers in the table correspond to the average total transaction costs computed according to Eq. (23) for portfolios of size N across 100 unique subsets of 93 assets. The last row reports the results over the whole available pool of assets. For each portfolio the TTC is computed over an evaluation horizon of $H = 763$ observations and an in-sample estimation window length $T = 250$.

A.3 Restricted estimators, T = 250

Table 12: Realized portfolio variance for $T = 250$.

N	Ledoit and Wolf 2004	Ledoit and Wolf 2020	DWE $\lambda = 0.1$	DWE $\lambda = 0.5$	DWE $\lambda = 1$	DWE $\lambda = 10$	DWE $\lambda = 50$	Equally weighted
3	0.000066	0.000066	**0.000065**	**0.000065**	**0.000065**	**0.000065**	**0.000065**	0.000071
13	0.000038	0.000038	0.000036	0.000036	0.000036	0.000036	**0.000035**	0.000043
23	0.000035	0.000035	**0.000031**	**0.000031**	**0.000031**	0.000032	**0.000031**	0.000040
33	0.000033	0.000033	**0.000029**	**0.000029**	**0.000029**	0.000031	**0.000029**	0.000038
43	0.000032	0.000033	**0.000028**	**0.000028**	**0.000028**	0.000034	0.000029	0.000037
53	0.000031	0.000033	**0.000028**	**0.000028**	**0.000028**	0.000034	0.000030	0.000037
63	0.000031	0.000032	**0.000027**	**0.000027**	**0.000027**	0.000036	0.000029	0.000036
73	0.000030	0.000033	**0.000027**	**0.000027**	**0.000027**	0.000034	**0.000027**	0.000036
83	0.000030	0.000032	0.000027	0.000027	0.000027	0.000033	**0.000026**	0.000036
93	0.000030	0.000032	0.000027	**0.000026**	**0.000026**	0.000031	**0.000026**	0.000036

Numbers in the table correspond to the average realized GMVP portfolio variance computed on out-of-sample portfolio returns of portfolios of size N across 100 unique subsets of 93 assets. The last row reports the results over the whole available pool of assets. For each portfolio the realized variance is computed over an evaluation horizon of $H = 763$ observations and an in-sample estimation window length $T = 250$. Numbers in bold correspond to the smallest realized portfolio variance for a given portfolio size N.

Table 13: Variance of out-of-sample portfolio returns for $T = 250$.

N	Ledoit and Wolf 2004	Ledoit and Wolf 2020	DWE $\lambda = 0.1$	DWE $\lambda = 0.5$	DWE $\lambda = 1$	DWE $\lambda = 10$	DWE $\lambda = 50$	Equally weighted
3	**0.000058**	**0.000058**	0.000060	0.000060	0.000060	0.000060	0.000060	0.000067
13	0.000036	0.000036	0.000036	0.000036	0.000036	0.000036	**0.000035**	0.000044
23	0.000033	0.000033	**0.000032**	**0.000032**	**0.000032**	0.000034	**0.000032**	0.000041
33	0.000032	0.000032	**0.000031**	**0.000031**	**0.000031**	0.000034	**0.000031**	0.000040
43	0.000031	0.000031	**0.000030**	**0.000030**	0.000031	0.000037	0.000032	0.000039
53	0.000031	**0.000030**	**0.000030**	**0.000030**	**0.000030**	0.000038	0.000033	0.000039
63	0.000031	0.000030	**0.000029**	**0.000029**	**0.000029**	0.000040	0.000032	0.000038
73	0.000030	0.000029	0.000029	**0.000028**	**0.000028**	0.000038	0.000029	0.000038
83	0.000030	0.000029	0.000030	0.000029	**0.000028**	0.000037	**0.000028**	0.000038
93	0.000030	0.000029	0.000027	**0.000026**	**0.000026**	0.000035	0.000028	0.000038

Numbers in the table correspond to the average out-of-sample GMVP portfolio variance computed on out-of-sample portfolio returns of portfolios of size N across 100 unique subsets of 93 assets. The last row reports the results over the whole available pool of assets. For each portfolio the out-of-sample return is computed over an evaluation horizon of $H = 763$ observations and an in-sample estimation window length $T = 250$. Numbers in bold correspond to the smallest out-of-sample portfolio variance for a given portfolio size N.

Table 14: Transaction costs for $T = 250$.

N	Ledoit and Wolf 2004	Ledoit and Wolf 2020	DWE $\lambda = 0.1$	DWE $\lambda = 0.5$	DWE $\lambda = 1$	DWE $\lambda = 10$	DWE $\lambda = 50$	Equally weighted
3	0.038	0.039	0.418	0.299	0.228	0.060	0.025	0.022
13	0.113	0.138	1.341	1.092	0.915	0.302	0.091	0.025
23	0.184	0.231	2.041	1.697	1.441	0.546	0.166	0.025
33	0.252	0.325	2.647	2.210	1.890	0.792	0.243	0.025
43	0.318	0.405	3.099	2.621	2.270	1.049	0.322	0.025
53	0.383	0.488	3.386	2.886	2.510	1.264	0.391	0.025
63	0.451	0.561	3.600	3.104	2.727	1.457	0.454	0.025
73	0.512	0.641	4.119	3.568	3.144	1.649	0.528	0.025
83	0.574	0.707	5.353	4.603	4.005	1.917	0.625	0.025
93	0.634	0.776	4.709	4.064	3.533	1.804	0.616	0.025

Numbers in the table correspond to the average total transaction costs computed according to Eq. (23) for portfolios of size N across 100 unique subsets of 93 assets. The last row reports the results over the whole available pool of assets. For each portfolio the TTC is computed over an evaluation horizon of $H = 763$ observations and an in-sample estimation window length $T = 250$.