Beyond GMV: The practical evaluation of covariance matrices*

M. Sipke Dom[1], Clint Howard[2,3], Maarten Jansen[3] and Harald Lohre[3,4]

[1]*Erasmus University Rotterdam*
[2]*University of Technology Sydney*
[3]*Robeco Quantitative Investments*
[4]*Lancaster University Management School*

January, 2025

Abstract

We empirically evaluate variance-covariance (VCV) estimators for equity portfolio construction and challenge traditional validation methods based on unconstrained global minimum-variance (GMV) portfolios. While traditional methods support using shrinkage and modeling covariance dynamics, the resulting portfolios are often impractical due to high leverage, concentration, and costs. By examining more realistic constrained GMV and risk parity portfolios, we find a significantly reduced opportunity for alternative VCV estimators to outperform the sample estimator. Asset weight constraints provide implicit shrinkage, making further explicit shrinkage largely unnecessary. However, accounting for time-series dynamics in asset returns remains statistically relevant for volatility reduction. We propose using a weight-constrained, long-only GMV portfolio with transaction cost penalties as a more suitable test portfolio for the practical evaluation of VCV estimators. Our findings emphasize the importance of considering both statistical robustness and practical implementation when selecting VCV estimators for portfolio management.

JEL classification: C13, C55, C58, G11

Keywords: minimum-variance portfolio, risk-based portfolios, shrinkage, factor models, portfolio optimization

*Corresponding author: Maarten Jansen, *email-address:* m.jansen@robeco.com, *Postal address:* Weena 850, 3014DA Rotterdam, South Holland, The Netherlands
We are grateful to David Blitz, Khuram Chaudhry, Gianluca De Nard, Jeffrey Durieux, Amit Goyal, Bart van der Grient, Danny Huang, Rob Huisman, Rasmus Lönn, , Matthias Hanauer, Sam van Meer, Stefan Mittnik, Michael Rockinger, Philipp Schiele, Jesper van Winden, Pim van Vliet, and participants at the CEQURA Conference on Advances in Financial and Insurance Risk Management, the INQUIRE Practitioner Seminar, the J.P. Morgan's Macro Quant and Derivatives Conference in London, and the Research Seminars at Erasmus University Rotterdam and Robeco for helpful discussions and comments. Note that this paper expresses the authors' views which do not necessarily coincide with those of Robeco.

1. Introduction

The key ingredient to equity portfolio construction is the variance-covariance (VCV) matrix of asset returns. The natural candidate to use is the sample covariance matrix; however, this estimator is prone to error and not suitable when the number of assets under consideration is large (Clarke, De Silva and Thorley, 2006; Lee, 2011; Ardia, Bolliger, Boudt and Gagnon-Fleury, 2017; Jurczenko and Teiletche, 2018). A rich literature proposes many alternative VCV estimators to address these limitations. These estimators are typically evaluated empirically using the ex post volatility of Markowitz's (1952) global minimum-variance (GMV) portfolio.[1] Except for a long-only constraint, studies proposing new VCV estimators rarely impose additional constraints on the GMV test portfolio, making the resulting portfolios unrealistic to implement in practice due to high leverage, concentration, turnover, and transaction costs.[2] We explore recent enhancements in VCV matrix estimation in equity universes from a practitioner's perspective across a range of risk-based portfolios, including constrained GMV and risk parity portfolios. We challenge the use of the ex post volatility of the unconstrained GMV portfolio as the primary empirical evaluation criteria for the performance of VCV estimators. Instead, we propose the use of a more realistic weight-constrained long-only GMV portfolio with transaction cost penalties and to consult additional performance measures to assess the practical relevance.

Using the 500 largest U.S. stocks from January 1990 to December 2021, we empirically evaluate a selection of VCV matrix estimators across various risk-based portfolio construction methods. The key performance metric for VCV matrix estimators is the ex post volatility of the test portfolios. The ex post evaluation metric should be aligned with the ex ante objective, which for the risk-based portfolios is minimizing portfolio volatility (Lee, 2011; Engle et al., 2019). In addition, we investigate portfolio characteristics beyond ex post volatility to verify whether these portfolios are realistic to hold in practice. We evaluate their risk-adjusted returns, asset weight concentration, portfolio turnover, transaction costs, and factor exposures,

[1] This is predicated on the out-of-sample properties of these portfolios which has been established in prior research (Haugen and Baker, 1991; Jagannathan and Ma, 2003).

[2] Some recent examples are Engle, Ledoit and Wolf (2019), Trucíos, Zevallos, Hotta and Santos (2019), Conlon, Cotter and Kynigakis (2021), De Nard, Ledoit and Wolf (2021), and De Nard, Engle, Ledoit and Wolf (2022).

characteristics that are commonly ignored in the portfolio management literature (Lesmond, Schill and Zhou, 2004; Frazzini, Israel and Moskowitz, 2012; Novy-Marx and Velikov, 2016). An evaluation metric we do not consider is the consistency of the VCV estimators. While this metric is widely used in the literature to evaluate VCVs (see, e.g., Fan, Liao and Mincheva (2013); Ledoit and Wolf (2004*b*, 2022*a*); Caner, Medeiros and Vasconcelos (2023)), it needs to be determined using simulations and may not significantly affect performance in practice (Engle et al., 2019; Ledoit and Wolf, 2024).

Our selection of VCV estimators consists of both traditional and state-of-the-art VCV estimators focusing on three key design choices: (i) shrinkage, (ii) time-dynamics, and (iii) factor structure. Combining these elements results in estimators with varying degrees of complexity. We first investigate the linear (LS) and nonlinear (NLS) shrinkage estimators of Ledoit and Wolf (2004b, 2024). Next, we account for dynamic time-series dependence in asset returns using the dynamic conditional correlation (DCC) model of Engle (2002) combined with nonlinear shrinkage methods (denoted DCC-NLS), following Engle et al. (2019) and the RiskMetrics (RM) approach (RiskMetrics, 1996) using an exponentially weighted moving average scheme. Finally, we impose two different factor structures to model the VCV matrix of a large number of assets assuming a small number of driving risk factors.

We construct a set of GMV and risk parity test portfolios to evaluate the VCVs. Specifically, we consider the traditional unconstrained GMV portfolio, a long-only GMV portfolio, and we introduce a long-only GMV portfolio with a maximum-weight constraint and a transaction cost penalty. Building on GMV portfolios, risk parity portfolios effectively aim to minimize portfolio variance subject to a diversification constraint and thus serve as a natural alternative to GMV portfolios. We evaluate two risk parity allocations: the equal risk contribution (ERC) portfolio of Maillard, Roncalli and Teïletche (2010) and the hierarchical risk parity (HRP) portfolio of López de Prado (2016).

Our main findings can be summarized as follows. For the traditional unconstrained GMV portfolios, we find a superior performance for sophisticated estimators like nonlinear shrinkage and DCC-NLS. The nonlinear shrinkage estimator obtains a significantly lower volatility than the simpler sample and linear shrinkage estimators, while DCC-NLS achieves the best performance overall. Although these findings are in line with the academic literature, we show that the unconstrained GMV portfolio is not a realistic test portfolio due its extremely high leverage, concentration, and turnover. Specifically, there is a considerable variation in the gross exposure and turnover, with a gap of nearly 600% gross exposure and 300% turnover across the estimators.

As the test portfolios become more realistic, there is a clear reduction in the range of volatilities between the best and worst estimators from 6.37% (unconstrained GMV) to 0.56% (weight-constrained GMV portfolio with a transaction cost penalty). This is an important result, as it shows that the discriminatory power of VCV estimators declines in more realistic test portfolios. This result can be partly explained by the implicit shrinkage embedded in portfolio constraints, which renders further shrinkage unnecessary (Jagannathan and Ma, 2003). Accounting for time-series dynamics remains statistically significant, as dynamic estimators are consistently the best-performing. Moreover, the sample estimator achieves similar results to the shrinkage estimators, which suggests that this estimator benefits most from implicit shrinkage and can achieve superior net risk-adjusted returns as a result. Finally, while the long-only GMV portfolio generally yields higher net Sharpe ratios than the unconstrained portfolio, these portfolios still give rise to large transaction costs and remain highly concentrated in the number of assets. This shows that the long-only constraint is not sufficient. We suggest imposing additional weight constraints and a transaction cost penalty to the GMV portfolio or considering risk parity portfolios for the practical evaluation of VCV estimators.

The intersection of VCV matrix estimators and risk-based portfolio construction has been explored in various contexts, with studies like Nakagawa, Imamura and Yoshida (2018) applying VCV estimators to multiple risk-based portfolios, and Ardia et al. (2017) and Jain and Jain (2019) examining traditional VCV estimators on risk-based portfolios for small asset universes.[3] However, as the number of assets increases, traditional VCV estimators face challenges such as numerical instability and "Markowitz's curse", where diversification benefits are overpowered by estimation errors (Michaud, 1989; López de Prado, 2016; Ledoit and Wolf, 2004*b*).[4] These challenges highlight the need for well-conditioned VCV estimators

[3]Ardia et al. (2017) consider six asset universes with 7–30 assets and Jain and Jain (2019) consider five different asset universes with ten assets each.

[4]This curse is derived from a high condition number, i.e., the absolute value of the ratio between the maximum

in large asset universes, as emphasized by Kan and Zhou (2007), who demonstrate the increased severity of variance estimation errors compared to mean errors as the number of assets grows. Our study applies a wide range of state-of-the-art VCV estimators on multiple risk-based portfolio construction methods in a large investment universe. We extend beyond ex post volatility to consider broader performance measures and jointly evaluate complex VCV estimators across numerous risk-based portfolio construction techniques. Importantly, we demonstrate how the selection of optimal VCV estimator can differ significantly based on the chosen evaluation portfolio.

An extensive body of literature proposes modifications to VCV estimators for large-scale portfolio optimization, including shrinkage methods (Ledoit and Wolf, 2003, 2017, 2020, 2022*a*), factor models (Fan et al., 2013), robust optimization (DeMiguel, Martin-Utrera and Nogales, 2013), and dynamic models (Engle, Ledoit and Wolf, 2019; De Nard, Ledoit and Wolf, 2021; De Nard, Engle, Ledoit and Wolf, 2022). Most of these studies evaluate their innovations based on the ex post volatility of unconstrained GMV portfolios. Recent literature has also explored alternative portfolio construction techniques, such as weight constraints (Jagannathan and Ma, 2003; DeMiguel, Garlappi, Nogales and Uppal, 2009), gross leverage constraints (Fan, Zhang and Yu, 2012; Zhao, Ledoit and Jiang, 2023), long-only constraints (Ledoit and Wolf, 2017) and transaction cost penalties (Ledoit and Wolf, 2017), with Boyd, Johansson, Kahn, Schiele and Schmelzer (2024) providing a comprehensive summary of various adjustments to the standard Markowitz portfolio optimization paradigm.

Our proposed test portfolio is similar to that of Zhao et al. (2023) and Ledoit and Wolf (2024). Zhao et al. (2023) impose gross exposure constraints of varying degrees and find that a DCC-NLS style estimator is the preferred estimator for their constrained GMV test portfolios. When Zhao et al. (2023) reduce the gross exposure of their test portfolios, they find that the difference in volatility reduction between the sample estimator and preferred DCC-NLS estimator is smaller. Ledoit and Wolf (2024) incorporate gross exposure constraints and a transaction cost penalty into the objective function of their test portfolios, however they do not test a long-only strategy. They find that the more advanced DCC-NLS estimators indeed

and minimum eigenvalues of the correlation/VCV matrix. If this ratio is high, a small change of an element in the VCV matrix leads to a completely different inverse.

outperform the basic NLS estimators for unconstrained GMV portfolios and a 130% long, 30% short GMV portfolio. We build upon both studies by constructing test portfolios which incorporate all these dimensions (gross exposure constraints, transaction cost penalties), as well as applying explicit asset-level weight constraints and constructing alternative risk parity portfolios. By designing a test portfolio that is both simple and closely resembles real-world constraints, our research provides a comprehensive understanding of VCV estimator performance in practical portfolio management contexts.

Our findings emphasize the importance of prudent test portfolio selection when evaluating VCV matrix estimators. We find discrepancies between the optimal VCV matrix estimator among different risk-based portfolios. As such, VCV estimators that have only been tested on unconstrained GMV portfolios, which is the standard practice in the academic literature, may not be the best choice for portfolios with meaningful investment constraints. Indeed, unconstrained and long-only GMV portfolios come with high levels of portfolio concentration and turnover, they show poor risk-adjusted returns, rendering them unsuitable candidates for evaluating VCV estimators. Our maximum-weight-constrained long-only GMV portfolio with transaction cost penalty mitigates these adverse properties and is thus a more suitable test portfolio for evaluating new VCV matrix estimators on. For practitioners, our findings emphasize the necessity to benchmark innovations in VCV estimators against portfolios resembling those that will be used in practice. For academics, we illustrate the importance of presenting the performance of VCV estimators across a range of risk-based portfolios when establishing the practical relevance of innovations in these estimators.

The remainder of this paper is structured as follows. Section 2 and Section 3 describe the risk-based portfolios and VCV matrix estimators used in the analysis, respectively. Section 4 presents the the empirical design including the data descriptions and performance metrics, and Section 5 portrays the results from horse-racing VCV estimators. Section 6 concludes.

2. Designing test portfolios for evaluating variance-covariance estimators

Classical Markowitz (1952) mean-variance portfolio optimization that trades off expected risk and return often suffers from estimation error, producing concentrated portfolios that may disappoint ex post from a risk-adjusted performance perspective. Avoiding the forecasting of expected returns and, in turn, focusing on risk-based portfolio allocation has thus become a popular area of research. In the realm of mean-variance portfolio optimization, risk-based portfolio allocation boils down to investigating minimum-variance portfolios, in which the VCV matrix is the key determinant of the resulting portfolio. Estimating VCVs is also prone to estimation error, especially in large asset universes where estimation error of the VCV becomes larger than that of expected returns (Kan and Zhou, 2007). Academic researchers have embraced the GMV portfolio as a natural candidate to judge the success of any effort to improve the accuracy of VCV estimation.

Although the GMV is a salient use-case in the study of risk-based portfolio allocations, there are other contenders that could benefit from more precise risk measurement and management. Therefore, we carefully lay out the different notions of risk-based allocation. We start with introducing the classical GMV problem along with a set of general constraints. We then move beyond GMV and present alternative allocation schemes designed to maximize risk diversification. Such risk parity strategies can be considered GMV portfolios that are subject to diversification constraints. While these portfolios strive to equalize single stock risk contributions to the overall portfolio, we also look into HRP portfolios that leverage the hierarchical structure inherent in the VCV. Lastly, we use simple allocation strategies like $1/N$ or inverse volatility for benchmarking the more involved strategies. For the portfolios that do not restrict short selling, we use their analytical formulae, and we resort to convex optimization for determining the long-only portfolios. Throughout this section N_t denotes the number of assets in the asset universe on date t. Table 1 provides an overview of the risk-based portfolios and VCV estimators used in the subsequent analyses.

<Insert Table 1 about here>

2.1. Global minimum-variance (GMV) portfolio

The GMV portfolio minimizes the ex ante portfolio variance:

$$\min_{w_t} w_t' \Sigma_t w_t \quad \text{s.t.} \quad \iota' w_t = 1, \tag{1}$$

where w_t is the vector of asset weights at date t, Σ_t is the VCV matrix of dimension $N_t \times N_t$ at date t, and ι is a vector of ones with length N_t. This optimization problem, that we refer to as unconstrained GMV (labeled GMV UNC), has the following analytical solution:

$$w_t^* = \frac{\Sigma_t^{-1}\iota}{\iota'\Sigma_t^{-1}\iota}. \tag{2}$$

The GMV portfolio corresponds to a single point on Markowitz's (1952) efficient frontier and requires only the VCV as input. The unconstrained GMV portfolio is simplistic by nature (e.g., due to the absence of asset weight restrictions). Still, various studies show that standard GMV portfolios often yield superior out-of-sample performance compared to other mean-variance portfolios, even when performance is measured not only in terms of minimal ex post risk, but also in terms of risk-adjusted returns (Jagannathan and Ma, 2003; Haugen and Baker, 1991).

2.1.1 Long-only constraints

Imposing long-only constraints (i.e., $w_t \geq 0$) to the GMV portfolio (labeled GMV LO) can be beneficial for two reasons. First, large leveraged portfolios are generally riskier to investors and often require higher portfolio turnover to implement, thereby reducing net portfolio returns. Second, long-only constraints bring implicit shrinkage of the VCV matrix estimator (Jagannathan and Ma, 2003), and thus helps to mitigate the adverse portfolio effects of estimation error. Zhao et al. (2023) compare direct shrinkage of the VCV matrix to imposing gross exposure constraints. They find that nonlinear shrinkage of the VCV matrix remains beneficial even if moderate gross exposure constraints are imposed as long as some short positions are allowed. This is because the constraints only adhere to one degree of freedom (i.e., the magnitude of the gross exposure constraints), whereas nonlinear shrinkage methods have N_t degrees of freedom.

2.1.2 Transaction cost penalty and maximum-weight constraints

Long-only GMV portfolios can still be overly concentrated and do not explicitly control for transaction costs. To construct more practically relevant portfolios, we investigate the long-only GMV portfolio with a transaction cost penalty in conjunction with maximum-weight constraints of one percent:

$$\min_{w_t} w_t' \Sigma_t w_t + \lambda \cdot \left(\sum_{i=1}^{N_t} c_{t,i} |w_{i,t} - w^*_{i,t-1}| + \tau_t^{\text{fix}} \right) \tag{3}$$
$$\text{s.t.} \quad \iota' w_t = 1, \quad w_t \geq 0, \quad w_{i,t} \leq 0.01, \quad \forall i = 1, \ldots, N_t,$$

where $c_{t,i}$ are the estimated transaction costs of asset i, $w^*_{i,t-1}$is the weight of asset i one day prior to the rebalancing date, τ_t^{fix} are the fixed transaction costs due to assets leaving the asset universe on date t, and λ is the transaction cost penalty parameter. At the initial date, $t = 0$, $w^*_{i,t-1}$ is set to zero. After evaluating a grid of transaction cost penalties $\lambda = 10^{-j}$ for $j \in \{2, 3, 4, 5, 6\}$, we use $\lambda = 10^{-3}$ for each of the VCV estimators in the main analysis. Ledoit and Wolf (2024) use different values for λ for their static and dynamic estimators. Since they investigate mean-variance portfolios with transaction cost penalties and an expected return constraint, they base their choice on the Sharpe ratio. Because we consider GMV portfolios, specifically, we instead base the choice of λ on a trade-off between the ex post volatility and the average transaction costs . We estimate the stock-specific transaction costs $c_{t,i}$ using the model of Briere, Lehalle, Nefedova and Raboun (2020) that requires open-high-low-close price data.

The maximum-weight constraints of one percent ensure that at least 100 positions are held and thus force the portfolio to be overdiversified, whilst forcing the asset weights to sum to one. Asset weight constraints also impose implicit shrinkage on the VCV. The specification we present here extends the formulation used in Ledoit and Wolf (2024) by incorporating long-only and maximum-weight constraints alongside a transaction cost penalty. These constraints resonate with an investment objective more closely aligned with a large institutional investor who is predominantly long-only and sensitive to transaction costs. We label this GMV portfolio variant GMV CON.[5]

[5]We separately run portfolios with maximum-weight constraints and transaction cost penalties. We find similar results to the combined case, and thus do not report them.

2.2. Risk parity portfolios

One feature that makes VCV matrix estimators attractive in the construction of risk-based portfolios is the option to take diversification into account via the pairwise information contained in the asset covariances. GMV portfolios implicitly aim to maximize risk diversification by minimizing the ex ante portfolio variance. We investigate two alternative risk-based portfolios that explicitly aim for optimally risk-diversified portfolios.

2.2.1 Equal Risk Contribution (ERC) portfolio

ERC portfolios aim for an allocation in which every asset contributes equally to the total portfolio risk, meaning that the risk contribution ($\mathrm{RC}_{i,t}$) to the portfolio by any asset i on date t is equal to $1/N_t$. We numerically optimize the ERC portfolio by minimizing the variance of the risk contributions:

$$\min_{w_t} \sum_{i=1}^{N_t} \left(\underbrace{\frac{w_{i,t}[\Sigma_t w_t]_i}{w_t' \Sigma_t w_t}}_{\mathrm{RC}_{i,t}} - \frac{1}{N_t} \right)^2 \tag{4}$$

$$\text{s.t.} \quad \iota' w_t = 1, w_{i,t} \geq 0, \quad \forall i \in \{1, \ldots, N_t\}.$$

Importantly, this problem can be expressed as a minimum-variance optimization problem subject to a diversification constraint, see Maillard, Roncalli and Teïletche (2010). Specifically,

$$\min_{\tilde{w}_t} \tilde{w}_t{}' \Sigma \tilde{w}_t \tag{5}$$

$$\text{s.t.} \quad \sum_{i=1}^{N_t} \log(\tilde{w}_{i,t}) \geq c_t, \quad \tilde{w}_{i,t} \geq 0, \quad \forall i \in \{1, \ldots, N_t\},$$

where c_t is a constant and $\tilde{w}_{i,t}$ are the unscaled optimal weights that are then scaled to sum up to one: $w_{i,t} = \frac{\tilde{w}_{i,t}}{\sum_{i=1}^{N_t} \tilde{w}_{i,t}}$. Hence, it is natural to consider risk parity as test portfolios for evaluating VCV matrix estimators where success then is gauged in terms of the resulting impact on portfolio volatility and diversification. The ex ante volatility of ERC portfolios can be directly related to GMV and equally-weighted (EW) portfolios, specifically, Maillard et al. (2010) show that:

$$\sigma_{GMV,t} \leq \sigma_{ERC,t} \leq \sigma_{EW,t}. \tag{6}$$

where $\sigma_{i,t} = w'_{i,t}\Sigma_{i,t}w_{i,t}$ with $i \in GMV, ERC, EW$ are the ex ante portfolio volatilities. This ranking holds after plugging in an estimated VCV matrix and portfolio weights, computed with in-sample asset returns. Given this in-sample relation, the ERC portfolio can be interpreted as a middle-ground portfolio between GMV and EW portfolios. The ranking does not necessarily hold for the ex post portfolio volatilities, which still have to be verified with out of sample asset returns.

2.2.2 Hierarchical Risk Parity (HRP) portfolio

The notion of a hierarchical structure in financial markets is becoming increasingly popular in modern portfolio theory. Mantegna (1999) established an economically meaningful taxonomy for stocks in the S&P 500 universe using hierarchical clustering on the correlation matrix of its assets. Tumminello, Aste, Di Matteo and Mantegna (2005) show that the instability of VCV matrix characteristics can be reduced through hierarchical clustering. Standard VCV estimators do not take into account that certain assets are close substitutes of one another (López de Prado, 2016). Not distinguishing between assets in the asset universe can produce ill-conditioned VCV estimators that are prone to Markowitz's curse. López de Prado (2016) assumes a hierarchical structure of the financial assets, which alleviates this problem by reducing the number of links between each of the assets to one, resulting in a minimum spanning tree with $N-1$ edges. López de Prado (2016) proposes to build a hierarchical risk parity (HRP) procedure that can be summarized in three steps: (i) tree clustering, (ii) quasi-diagonalization, and (iii) recursive bisection. We refer the reader to López de Prado (2016) for further details on the estimation of the HRP portfolio.

2.3. Benchmark portfolios

We use three simple portfolio allocation schemes as benchmark portfolios: (i) equally weighted (EW), (ii) value-weighted (VW), and (iii) inverse-variance (IV). The EW strategy assigns equal weights to all assets in the portfolio:

$$w_{i,t}^{EW} = \frac{1}{N_t}. \tag{7}$$

The EW strategy is a natural candidate as DeMiguel, Garlappi and Uppal (2009) show that this portfolio can be difficult to beat out of sample based on Sharpe ratio, certainty-equivalent return, or turnover. The VW portfolio assigns asset weights proportional to their market cap and this market portfolio is a classic reference point:

$$w_{i,t}^{VW} = \frac{mcap_{i,t}}{\Sigma_{i=i}^{N_t} mcap_{i,t}}, \tag{8}$$

where $mcap_{i,t}$ is the market capitalization of stock i at time t. Finally, the IV portfolio minimizes portfolio variance without accounting for risk diversification benefits by assigning weights inverse to the stock's historical volatility:

$$w_{i,t}^{IV} = \frac{1/\sigma_{i,t}}{1/\Sigma_{i=1}^{N_t} \sigma_{i,t}}, \tag{9}$$

where $\sigma_{i,t}$ is the stock i's historical volatility estimated over the same period as the VCVs. This can be seen as a special case of a fully-invested GMV portfolio where all off-diagonal elements of the VCV are set to zero.

3. Estimating large VCV matrices

The key ingredient to determining risk-based portfolio allocations is the VCV matrix. However, accurate estimation of the VCV matrix is challenging, particularly in large asset universes. Therefore, it is crucial to produce well-conditioned VCV estimators. The literature has put forward different methods to estimate the VCV with high precision but which come with varying degrees of complexity. In this paper, we compare the performance of the standard sample estimator to a set of salient shrinkage-based estimators that are introduced in this section. Drawing inspiration from the literature, we increase the complexity of the shrinkage-based estimators along three dimensions: (i) shrinkage type, (ii) time-dynamics, and (iii) factor structure. Our selection of VCV estimators encompasses most commonly used ones in the literature.[6]

3.1. Sample estimators in large asset universes

The most common VCV matrix estimator is the unbiased sample estimator, $S \in \mathbb{R}^{N\times N}$:

$$S = \widehat{\mathbb{E}}\left[(r_t - \bar{r})(r_t - \bar{r})'\right] = \frac{1}{T-1}\sum_{t=1}^{T}(r_t - \bar{r})(r_t - \bar{r})', \tag{10}$$

where $r_t \in \mathbb{R}^N$ are the asset returns at time t and $\bar{r} = \frac{1}{T}\sum_{t=1}^{T} r_t$.

Using the sample estimator for portfolio allocation in large asset universes is problematic. Large asset universes are synonymous with high concentration ratios (that is, the number of assets over the number of observations), making portfolios that require inverting the VCV matrix infeasible due to the (near) singularity of the sample estimator. However, even when the VCV matrix is non-singular, unbiased sample estimators of the VCV matrix are well-known for producing unstable portfolios with poor out-of-sample performance in large asset universes (Jobson and Korkie, 1980; Brandt, 2010; Ledoit and Wolf, 2004*b*).[7] This is again a manifestation of Markowitz's curse, where error maximization largely drives the resultant portfolios.

3.2. Shrinkage estimators

To reduce estimation errors in the sample mean, Stein (1956) and James and Stein (1961) established the concept of shrinkage. Ledoit and Wolf (2003; 2004*a*; 2004*b*; 2012; 2015; 2017; 2020; 2022*a*) apply shrinkage to VCV matrix estimation with the aim of achieving more stable risk-based portfolio performance when the concentration ratio is high. Simple shrinkage involves combining or averaging two "extreme" estimators to create a better performing and more stable combined estimator. There are two main approaches to shrinkage: linear and

[6] Two notables classes of estimators that we do not cover are the POET estimator by Fan et al. (2013) and precision matrix estimators. The former is effectively an approximate factor model (see Section 3.4) with two key differences: (i) latent factors instead of observable factors, and (ii) thresholding instead of shrinkage of the residual covariance matrix. Precision matrix estimators, such as Caner et al. (2023); Friedman et al. (2008); Kourtis et al. (2012), directly estimate the inverse variance-covariance matrix and forego the need to invert the VCV matrix. We do not consider these estimators because not all our portfolio allocation problems require the inversion of the VCV matrix.

[7] The concentration ratio could simply be reduced by increasing T. However, this solution is generally not applicable to financial time series, such as stock returns, given limited data availability and non-stationarity.

nonlinear.[8] The shrinkage estimators that we consider are solely driven by the asset return data.[9]

3.2.1 Linear shrinkage

The simplest version of linear shrinkage is shrinking the $N \times N$ sample VCV matrix toward a scalar multiple of the identity matrix (Ledoit and Wolf, 2004*b*):

$$\widehat{\Sigma} = (1 - \kappa)S + \kappa q I_N, \tag{11}$$

where $\kappa \in [0, 1]$ is the shrinkage intensity, $q \in \mathbb{R}$ is a scalar, and $I_N \in \mathbb{R}^{N \times N}$ is the identity matrix. We set q equal to the average of the univariate variances of the sample VCV matrix estimator, enforcing qI_N as the shrinkage target. Throughout the remainder of this paper, we refer to the linear shrinkage estimator as LS.[10] The shrinkage intensity κ is calculated based on Ledoit and Wolf (2004*b*) such that it is a consistent estimator of the asymptotically optimal intensity.

3.2.2 Nonlinear shrinkage

A drawback of linear shrinkage is that the shrinkage target has to be determined a priori based on assumed characteristics of the unknown true VCV matrix (Ledoit and Wolf, 2022*b*). In contrast, nonlinear shrinkage methods do not require assumptions on the characteristics of the true VCV matrix. Although nonlinear shrinkage methods are more complex, they show significantly better out-of-sample performance (Ledoit and Wolf, 2022*b*). We use the Quadratic Inverse Shrinkage (QIS) estimator of Ledoit and Wolf (2022*a*), referred to as NLS throughout this paper. This estimator shrinks the inverse eigenvalues of the VCV matrix.[11]

[8] We refer to Ledoit and Wolf (2022*b*) for a comprehensive overview of various shrinkage estimators developed over the past 15 years.

[9] Because not all portfolios in our analysis require VCV matrix inversions, we do not consider direct shrinkage estimators of the inverse VCV matrix. For examples of shrinkage estimators of the inverse VCV matrix, see the works of Friedman, Hastie and Tibshirani (2008), Kourtis, Dotsis and Markellos (2012), and DeMiguel, Martin-Utrera and Nogales (2013).

[10] We also consider linear shrinkage toward a constant-correlation matrix (Ledoit and Wolf, 2004*a*). Results are generally inferior compared to the LS estimator.

[11] We find similar results when implementing the analytical NLS estimator of Ledoit and Wolf (2020).

3.3. Dynamic estimators

The estimators introduced so far are static and ignore time-variation of the VCV matrix. Static estimators introduce biases towards older asset returns and ignore the time-variability and clustering of volatility. In this vein, dynamic estimators allow for time-varying conditional covariance matrices by assigning different weights to older and more recent asset returns. Similar to Engle et al. (2019), we apply nonlinear shrinkage to the dynamic estimators to prevent in-sample overfitting and ensure a non-singular matrix. For consistency, we use the Ledoit and Wolf (2022*a*) NLS estimator and do not consider other shrinkage methods since Engle et al. (2019) find nonlinear shrinkage to be most effective.

3.3.1 Dynamic conditional correlation model

We estimate the dynamic conditional correlation model with nonlinear shrinkage (DCC-NLS) of Engle et al. (2019). The model extends Engle's (2002) DCC model, which models the time-varying conditional volatilities and correlations using a generalized autoregressive conditional heteroskedasticity (GARCH)-like process. Nonlinear shrinkage (in our case: the NLS estimator) is applied to the correlation targeting matrix to prevent negative eigenvalues of the matrix. Furthermore, we use the averaged forecast approach of De Nard et al. (2021) to convert the DCC estimator (which is a prediction for the next day) to a prediction for the next month.

3.3.2 RiskMetrics

Accounting for dynamics based on the described DCC-NLS estimator introduces significant complexity. Therefore, we also consider a simple dynamic estimator that is popular among practitioners, the RiskMetrics (1996) estimator (RM). This estimator weights the observations by a decay parameter $\xi^{-(T-t)}$. Here, T is the current date and t is the date of some earlier observation. We set $\xi = 0.99734$, which roughly corresponds to a half-life of one year for daily return data. We refer to this estimator as RM.

Because the RM estimator, by itself, does not account for poor conditioning of the VCV matrix, shrinkage may be beneficial in a large asset universe setting. Therefore, we apply the

NLS estimator to the estimated RM VCV matrix and we label this the RM-NLS estimator. Estimation follows a two-stage process. First, we estimate the sample covariance matrix using RiskMetrics, thus introducing exponential decay in the estimation of the volatilities and pairwise correlations. Second, we apply NLS to the estimated RM covariance matrix. Although the NLS estimator of Ledoit and Wolf (2022*a*) is designed for sample VCV matrices in an i.i.d. sample, using NLS allows for the RM-NLS estimator to serve as a middle ground between the RM and DCC-NLS estimators.[12]

3.4. Factor models

Factor models are derived from asset pricing theory and focus on specifying a functional form of stock returns. Factor models reduce dimensionality in asset pricing by attempting to explain the cross-sectional information of a large number of asset returns (N) based on a parsimonious set of factors (K). Linear factor models can be represented as:

$$r_t = \alpha + Bf_t + u_t, \quad \forall t \in \{1, \ldots, T\}, \tag{12}$$

where $\alpha \in \mathbb{R}^{N_t}$ is often assumed $\mathbf{0}$, $f_t \in \mathbb{R}^K$ are the factor returns, $B \in \mathbb{R}^{N_t \times K}$ is the loadings matrix, $u_t \sim \mathcal{N}_{N_t}(0, \Sigma_{u,t})$ are the idiosyncratic errors, and T is the sample size. Exploiting this linear factor structure, the VCV matrix of asset returns, Σ_t, can be written as:

$$\Sigma_t = B'\Sigma_{f,t}B + \Sigma_{u,t}, \tag{13}$$

where $\Sigma_{f,t}$ is the $K \times K$ factor returns VCV matrix and $\Sigma_{u,t}$ is the $N \times N$ residual VCV matrix on date t.

For $\Sigma_{u,t}$, we consider both an exact factor model (EFM) and an approximate factor model (AFM) design. The difference between EFMs and AFMs is fundamentally a bias-variance trade-off. EFMs assume that the factors fully explain cross-sectional asset risk, i.e., $\Sigma_{u,t}$ is a diagonal matrix of static sample variances. AFMs assume a less stringent structure with a

[12] This approach shares some similarities with the DCC-NLS model. Both estimators estimate the conditional standard deviations using a GARCH-like process and both estimators shrink the correlations using the NLS shrinkage estimator. The key difference is that DCC-NLS applies shrinkage to estimate the unconditional correlation matrix, whereas we apply it to the conditional correlation matrix estimated using RM.

sparse residual VCV matrix, which we model using the DCC-NLS estimator. This estimator originates from De Nard et al. (2021) and is denoted by AFM-DCC-NLS. De Nard et al. (2021) find a one-factor structure to be optimal for their AFM-DCC-NLS. Therefore, our application also uses a one-factor model with the market factor to construct $\Sigma_{f,t}$ for both our EFM and AFM.[13] Ledoit and Wolf (2022*b*) and De Nard et al. (2021) find no improvements in the performance of their factor-model-based estimator when they allow for time-variation of the factor VCV matrix. Because the main objective of this paper is to compare common VCV estimator choices from the literature rather than finding the overall 'best' estimator, we also do not model a time-varying factor VCV matrix or factor loadings.

4. Empirical design

4.1. Data

Our asset universe comprises of U.S. stocks with share codes 10 or 11 that are listed on the NYSE, AMEX, and NASDAQ stock exchanges from January 1, 1990, through December 31, 2021. The data consists of daily stock-level returns, market capitalization, and open-high-low-close prices, sourced from the Center of Research in Security Prices. We take daily and monthly Treasury bill rates, market, and factor returns from the Kenneth French Data Library. Treasury bill rates are used as the risk-free rate to calculate excess returns. Lastly, we obtain returns of the Low-Volatility factor from Robeco Datasets and VIX data from the FRED.[14]

VCV matrices are estimated using a rolling window approach with a five-year estimation period. VCV matrix estimates are provided to each portfolio construction method, where we rebalance portfolios on the last trading day of each month. Portfolios are then held for one month and subsequently rebalanced. On the first trading day of the month, the asset universe includes all assets that are among the $N \in \{500, 1,000\}$ largest in terms of market

[13] We tested various specifications including Principal Components Analysis and the Fama-French three-factor model (Fama and French, 1993), and found that our results are robust to the choice of factor model. We focus on the one-market-factor model for brevity and to keep our methodology consistent with the findings of De Nard et al. (2021).

[14] See the websites https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html, https://www.robeco.com/en-int/insights/2022/04/data-sets-volatility-sorted-portfolios, and https://fred.stlouisfed.org/series/VIXCLS.

cap on that day, and have data available on every day during the in-sample period (that is, during the estimation rolling window). If asset return data is missing in the out-of-sample period, we set it equal to zero. Because the results for the 500 and 1,000 stocks universes lead to the same conclusions, we show the results for the 500 stocks universe in this paper and refer to the Online Supplementary Material for the results of the 1,000 stocks universe.

To gauge the practicality of the resulting risk-based portfolio allocation we compute portfolio performance net of transaction costs. The asset-specific transaction costs are estimated using the model of Briere et al. (2020) that requires open-high-low-close price data. Figure 1 presents the distribution of estimated transaction costs. The average median transaction cost throughout the sample is 4.1bps. Clear spikes in transaction costs can be seen after 2008 and 2020 coinciding with the GFC and COVID-19 crisis, respectively. These numbers are in line with other studies, e.g., Corwin and Schultz (2012), Abdi and Ranaldo (2017), Ardia, Guidotti and Kroencke (2024), and Ledoit and Wolf (2024).

<Insert Figure 1 about here>

4.2. Performance metrics

We evaluate each portfolio in the upcoming horse race using several ex post performance metrics. In line with Lee (2011), the main performance criterion is the annualized ex post portfolio volatility, computed as the standard deviation of the out-of-sample portfolio returns. As ERC portfolios can be reformulated as minimum-variance optimizers, ex post volatility also closely aligns with the ex ante objective of all risk-based portfolios. We use the pairwise variance test of Ledoit and Wolf (2011) to determine whether differences in ex post volatility between two estimators are statistically significant.

To keep the amount of testing tractable, we do not test every possible pairwise combination of estimators. Instead, as the complexity of the estimator increases we typically use a simpler model as the benchmark to test against. Specifically, the LS estimator is tested against the Sample estimator; the NLS estimator against the LS estimator; the RM-NLS estimator against the NLS estimator; the DCC-NLS estimator against the RM-NLS estimator; the EFM estimator against the Sample estimator; and the AFM-RM-NLS estimator against the

RM-NLS estimator. The ex post p-values are adjusted using the Holm (1979) correction to control for the family-wise error rate, accounting for the seven comparisons we make.

In addition to ex post volatility, we calculate (risk-adjusted) returns (both gross and net of transaction costs), portfolio concentration, and turnover, which are all of practical interest to investors.[15] With regards to portfolio concentration, we calculate three measures. The average number of positions with an absolute weight greater than 0.1% (POS), the average monthly sum of the largest absolute 10 of positions (MAXW), and the average monthly effective portfolio weights (i.e., the inverse of the sum of the squared portfolio weights, denoted WEFF). We note that for highly leveraged long-short portfolios, the measurement of WEFF can become distorted. However, WEFF still provides a meaningful reference point when comparing portfolio concentration within and across different portfolio construction methods. We refer to Table 2 for an overview of the performance metrics.

<Insert Table 2 about here>

5. Horse racing VCV estimators in risk-based portfolios

5.1. Evaluating VCV matrix estimators by ex post portfolio volatility

Table 3 shows the ex post volatility of the benchmark portfolios, risk-based portfolios, and selected VCV matrix estimators detailed in Table 1. Specifically, for each risk-based portfolio and VCV estimator combination, we use a five-year VCV estimator as input to the risk-based portfolio method. We then calculate the outperformance of each pair over the designated benchmark model.

<Insert Table 3 about here>

5.1.1 How does the choice of shrinkage method affect performance?

To set the stage, we first report the ex post volatilities of the three benchmark strategies. The market portfolio (value-weighted) has an annualized volatility of 18.92%, while the equal-weighted portfolio's volatility is one half percentage point higher (19.38%); the inverse

[15] We also calculate the 1% and 5% Value-at-Risk (VAR) and Expected Shortfall. These results are not reported as they result in conclusions largely consistent with ex post volatility.

variance portfolio has a volatility of 16.84%. Against this backdrop, the unconstrained GMV clearly outperforms the benchmark portfolios with a volatility of 10.91% for the sample estimator. This is not surprising, as the unconstrained GMV portfolios have the ability to short assets, thereby mitigating exposure to systematic risk factors such as the market factor. When long-only constraints are imposed, there is an increase in the overall portfolio volatility. Nevertheless, the sample estimator continues to outperform the benchmark portfolios even after imposing long-only constraints across the GMV and HRP test portfolios. The only exception is the ERC portfolio for which the inverse variance portfolio achieves a lower volatility.

Moving from the sample estimator to the shrinkage estimators clearly reduces the ex post volatilities of the unconstrained GMV portfolio further, yielding volatilities between 10.08% (NLS) and 10.49% (LS). In terms of linear versus nonlinear shrinkage, the NLS estimator thus clearly outperforms the LS estimator for the unconstrained GMV portfolio.

The range of volatilities decreases, and the value of shrinkage estimators diminishes when considering other test portfolios. These portfolios impose portfolio constraints, which already provide implicit shrinkage (Jagannathan and Ma, 2003). In all cases, the shrinkage estimators yield similar ex post volatilities as the sample estimator. Therefore, there is no conclusive evidence of which static structure-free shrinkage estimator is best for minimizing ex post volatility beyond the unconstrained base case. However, shrinkage estimators may still be useful in long-only portfolios to prevent singularity of the VCV matrix when the asset universe is very large.

5.1.2 Do dynamic estimators outperform static estimators?

Moving from (structure-free) static estimators to (structure-free) dynamic estimators, the discrepancy between the unconstrained GMV portfolios and the remaining long-only portfolios becomes most evident. The simple RM estimator is among the worst performers in terms of minimizing ex post volatility for the unconstrained GMV portfolio. Combining the RM estimator with NLS shrinkage results in a volatility lower than that of the static shrinkage estimators. The more sophisticated DCC-NLS estimator achieves the lowest volatility (9.33%) out of the VCV estimator candidates, which is also significantly lower than for the RM-NLS estimator. Hence, both shrinkage and the choice of dynamic estimator is important for the unconstrained GMV portfolio.

For the long-only GMV portfolios, each dynamic estimator outperforms the static estimators, with RM and RM-NLS performing equally well. Since the benefits of shrinkage decrease when long-only constraints are imposed, the additional flexibility of dynamic estimators seems to be a meaningful way to further reduce ex post volatility. This holds true for all other test portfolios, as the structure-free dynamic estimators result in significantly lower volatilities than their static counterparts. The DCC-NLS achieves the lowest ex post volatility among all estimators across all but the GMV CON test portfolios.

5.1.3 Do minimum-variance problems benefit from a factor structure?

Lastly, we investigate whether accounting for a factor structure is important in VCV estimation. It turns out that—setting aside the two risk parity portfolios—EFMs are among the worst estimators in terms of minimizing ex post volatility. This suggests that the additional bias outweighs the benefits of reduced estimation errors; arguing in favor of AFMs rather than EFMs to limit this structural bias (if one does not wish to shrink the components of the VCV matrix estimator driven by the risk factors). However, AFM-DCC-NLS does not improve upon the structure-free DCC-NLS estimator in any of the risk-based portfolios.[16]

5.1.4 Subperiod analysis

Panels C and D of Table 3 show how the ex post volatility of the risk-based portfolios differs in high- and low-volatility regimes. When the VIX index is above (below) its five-year moving average we consider the period a high-volatility (low-volatility) regime and

[16]We note that this finding contrasts that of De Nard et al. (2021) who find that AFM-DCC-NLS further reduces the volatility for GMV UNC over the DCC-NLS estimator. We replicate their results in the 500 stocks universe over their sample period of 1978-2017 and find very similar values in terms of volatility (8.1%), turnover (129%) and gross exposure (345%) for the GMV UNC portfolio under the AFM-DCC-NLS estimator. However, our DCC-NLS estimator finds a slightly lower volatility than that of De Nard et al. (2021) (8.1% compared to 8.3%), which changes the ranking of our estimators compared to theirs for GMV UNC. These differences can be attributed to the shrinkage estimator used (we use QIS Ledoit and Wolf (2022*a*); they use QuEST Ledoit and Wolf (2017)) and the sample selection processes, (we use the end-of-month calendar dates to determine our out-of-sample periods; De Nard et al. (2021) assume fixed estimation and out-of-sample periods). Our extension to 1978-2017 data suggests that the factor structure benefits emerge from the 1978-1990 period.

we thus divide the sample period into 46% high-volatility regimes and 54% low-volatility regimes, respectively. Not surprisingly, the level of ex post volatilities is the highest during high-volatility regimes. For instance, average market volatility then stands at 25.32% whilst low volatility regimes see a figure of 10.92%. By and large, the evaluation of estimators by volatility regime is consistent with the presented full sample evidence. For instance, the RM, RM-NLS and DCC-NLS estimators consistently outperform the other estimators in the long-only GMV portfolios and are among the best performers in the risk parity portfolios.

5.2. Risk-based portfolios in practice

Having evaluated risk-based portfolios from an ex post volatility perspective, we wonder about their practical relevance. To this end, we investigate the risk-based portfolios' performance statistics, specifically looking into the performance drag imposed by portfolio turnover. We also scrutinize the distributions of portfolio weights to gauge overall portfolio concentration and diversification.

5.2.1 Unconstrained global minimum-variance portfolios

Table 4 presents the detailed performance statistics for the benchmark portfolios, the three GMV portfolios, and the two risk parity alternatives. As expected, the unconstrained GMV portfolio achieves the lowest ex post volatility. Although GMV UNC allows for considerable reductions in portfolio volatility, one has to be mindful that these reductions rely on highly levered long and short portfolio positions. For instance, the unconstrained GMV based on the sample VCV comes with an average gross exposure of 723%. Notably, modeling a more structured VCV helps reducing such gross exposure (with the EFM variant having the lowest gross exposures, 232%). Still, these portfolios display unduly high portfolio turnover, ranging from 25.7% (EFM) to 302.0% (RM) in terms of one-way monthly turnover.

<Insert Table 4 about here>

Despite relatively high transaction costs and turnover, the GMV UNC portfolio has one candidate VCV estimator that stands out in terms of risk-adjusted returns : the DCC-NLS estimator. The transaction costs reduce the Sharpe ratio of the GMV UNC portfolio with

DCC-NLS estimator from 0.958 to 0.850, which still dwarfs the Sharpe ratios of all other portfolios considered in the analysis. This is in line with the finding of Ledoit and Wolf (2024), who show the DCC-NLS estimator outperforming the NLS estimator and EW portfolio in terms of maximizing the net Sharpe ratio.

Nevertheless, poor diversification properties highlight the impractical nature of the GMV UNC portfolio regardless of the choice of VCV estimator. For instance, the top ten names of GMV UNC make up a large share of the portfolio weights and the effective portfolio weights range between 5.0 (RM) and 48.0 (EFM).[17] The sobering verdict is that GMV UNC is of little practical relevance and thus not particularly informative for the choice of VCV estimators in actual portfolio management settings. Therefore, we investigate more realistic GMV variants, one with long-only constraint and one with tighter maximum stock weights as well as a transaction cost penalty.

5.2.2 Long-only GMV portfolios

By design, long-only GMV variants come with a reduced gross exposure of 100%. Still, turnover statistics are elevated and range from 11.8% (EFM) to 90.7% (DCC-NLS)—suggesting that the consideration of covariance dynamics can become costly. From a turnover perspective, AFM-DCC-NLS improves upon the DCC-NLS estimator by reducing the turnover nearly by half at the expense of a higher average volatility. Modeling dynamics via RiskMetrics or RM-NLS is even less turnover-intensive with an average turnover of 28.2 and 25.7%, respectively. Notwithstanding, long-only GMV portfolios are generally too concentrated to be considered viable alternatives in practice; the effective portfolio weights are consistently below 30.0. This finding is in line with Clarke, De Silva and Thorley (2011), who rationalize that the long-only GMV portfolio tends to only select assets with low market exposures.

Against this backdrop, we next enforce more diversified GMV portfolios by complementing the long-only constraint using upper weights constraints ($w_t \leq 1\%$) as well as a transaction cost penalty. As a result, portfolio concentration is considerably reduced, seeing the top ten names making up almost exactly 10% for all considered VCV estimators.[18] Unsurprisingly,

[17] POS, MAXW, and WEFF are all computed using absolute portfolio weights.
[18] The 10% threshold is marginally exceeded for all estimators due to rounding.

the GMV portfolio optimization sees most portfolio names testing the upper bound of 1%, resulting in portfolios that have hardly more than 100 names in total. In addition to improved portfolio diversification, one enjoys reduced turnover and transaction costs, but these result in higher net Sharpe ratio compared to the GMV LO portfolios for the static estimators only. The transaction costs drop for each of the estimators except for the EFM estimator. Since we apply the same penalty $\lambda = 10^{-3}$ to all VCV estimators, the dynamic estimators still have a higher level of turnover and transaction costs. [19]While applying a higher penalty to the dynamic estimators would result in equal levels of transaction costs, this may come at the cost of higher volatility. Naturally, the implicit shrinkage brings about a reduction in the opportunity set for any given VCV estimator.

5.2.3 Beyond GMV portfolios

Table 4 also presents the performance statistics and portfolio characteristics of the two risk parity alternatives. While the HRP portfolio has similar turnover figures as the long-only GMV portfolio, the classic ERC portfolios display the lowest turnover statistics across all tested strategies. Naturally, dynamic modeling of the VCV calls for higher turnover with risk parity strategies, albeit at a lower level. Although ERC and HRP have, on average, the same amount of non-negligible positions (500), the HRP portfolio is more concentrated than the ERC portfolio as its top ten names consume between 7.2–14.1% of the total portfolio (relative to 5.3–6.4% for ERC). Moreover, the number of effective portfolio positions is consistently the highest for the ERC portfolio ranging from 382.0 (EFM) to 415.0 (LS) making the ERC portfolio the least concentrated among the risk-based portfolios. This finding is in line with the notion that this portfolio serves as a middle ground between the EW and long-only GMV portfolios. As a result, the estimated transaction costs are 0.38–0.61bps for ERC which is on par with naive EW strategies but around 5 times higher than the transaction cost of the VW market portfolio. The emerging net returns are slightly lower than those of the market portfolio and the corresponding net Sharpe ratios are smaller than the long-only GMV portfolios, owing to the higher risk level of the risk parity strategies.

[19]Table A.1 in the Online Supplementary Material shows the effect of different transaction cost penalties on ex post volatility and transaction costs.

Examining the effect of the choice of VCV estimator for ERC and HRP portfolios, we observe a spread of only 0.5% (ERC) and 1.8% (HRP) in ex post volatility across the VCV estimators. This indicates a significantly reduced opportunity set for the VCV estimators. We find that shrinkage does generally not result in significant volatility reduction relative to the sample estimator. However, with the exception of RM for the HRP portfolio, the dynamic estimators all significantly outperform their static counterparts. ERC and HRP are also the only portfolios where EFM improves upon the sample estimator. Finally, the number of effective positions is (second) lowest for DCC-NLS, this means that this estimator diverges most from the equally weighted benchmark. Paired with the fact that this estimator also yields the lowest ex post volatility, this result shows that optimal diversification, powered by a well-conditioned VCV estimator, may improve upon naive $1/N$ diversification.

5.2.4 Risk-based portfolio selection and factor investing

Regardless of the chosen VCV estimator, we have demonstrated that long-only minimum-variance investing would have generated higher risk-adjusted returns than a naive market portfolio (VW). To rationalize this observation, we run a style factor regression of the risk-based portfolios' returns to investigate their salient systematic factor exposures. Table 5 reports regression results based on a multi-factor regression featuring an intercept (α) and seven off-the-shelve factors: Market (MKT), Size (SMB), Value (HML), Profitability (RMW), Investment (CMA), and Momentum (MOM) factors from the Kenneth French library as well as the Low-Volatility (LOWVOL) factor. Here, we focus on risk-based portfolio returns based on the DCC-NLS estimator estimated over the five-year estimation window. Indeed, all variations of the GMV portfolio display very low Market betas that are significantly lower than 1.0, ranging from 0.3 for the unconstrained GMV to around 0.6 for the constrained variants. The low Market betas are in line with Clarke et al. (2011) and Scherer (2011), who find that GMV portfolios only select assets with low Market exposures.

<Insert Table 5 about here>

Moreover, all GMV portfolios load positively on the Low-Volatility factor with highly significant betas around 0.10–0.24. Such exposures are expected given the risk-based portfolio objective that renders the GMV portfolios implicitly exploiting the Low-Volatility anomaly. Outside Market and Low-Volatility factor exposures, we only observe significant exposures in GMV CON for the Profitability and Investment factors. Taken together such factor exposures explain only 47% of the variation in long-only GMV portfolio returns and 75% of the GMV CON portfolio returns, which in both cases leave no significant alpha.

Notably, the two risk parity variants show some similar factor exposures as GMV CON. We also document strong loadings on the Low-Volatility, Profitability, and Investment factors and a Market beta that is significantly lower than 1.0. Unlike the long-only GMV portfolio, the ERC portfolio comes with positive Size exposure. Overall, more than 85% of the variation in risk parity portfolio return can be attributed to such common factor exposures, again leaving an insignificant alpha over the sample period.

5.3. Results summary

Figure 2 summarizes our key results across the seven VCV estimators and five risk-based portfolios that we explore. First, we confirm the academic evidence that the unconstrained GMV portfolio benefits from more complexity in VCV modeling. The according ex post portfolio volatility based on the sample VCV is 10.91% and reduces down to 9.33% when applying DCC together with nonlinear shrinkage. Second, the implicit VCV shrinkage that arises from incorporating practical considerations into the portfolio optimization limits the ex post volatility reduction of more involved VCV estimation choices, such as accounting for time-series dynamics and imposing factor model structures. For instance, the constrained GMV setting shrinks the ex post volatility range to lie in between 12.99% and 13.55%, and the shrinkage of this opportunity set is even more pronounced for the other portfolio statistics. Third, unconstrained GMV portfolios are very concentrated and have unduly high turnover which eats up any gross return benefit. Indeed, after accounting for turnover and transaction costs, the Sharpe ratio improvement of more complex VCV estimators over the sample estimator is reduced, particularly for the more constrained test portfolios. Specifically, a VCV estimator that combines a simple dynamics model, such as RiskMetrics with nonlinear shrinkage, performs in line with more complex VCV estimators on realistic test portfolios. Ultimately, these findings highlight practitioners' need for an alternative test portfolio to evaluate VCV estimators.

<Insert Figure 2 about here>

6. Conclusion

The estimation of VCV matrices lies at the heart of risk-based portfolio optimization. The associated estimation risk unduly exacerbates error maximization in portfolio construction, particularly when dealing with large asset universes. A plethora of advanced VCV estimators have been suggested in the literature, and we study the practical value-add of key contenders. The tested estimators differ in salient model features regarding their approach to shrinkage, dynamics, and factor structure.

The key contribution of the present paper is to question the common practice in the literature of validating VCV estimators based on the ex post performance of an (unconstrained) GMV portfolio. We confirm the latter use-case benefits from more complexity in VCV modeling, resulting in a reduction of ex post volatilities that tends to boost the GMV portfolio's Sharpe ratios. Yet, we argue that there are more portfolio characteristics to consider for demonstrating the practical value-add of new VCV estimators. Specifically, unconstrained GMV portfolios come with very high turnover and thus transaction costs that (more than) erode any observed gross benefits. Moreover, the ensuing portfolios concentrate portfolio weights on very few names, defeating the core principle of portfolio diversification.

Against this backdrop, we focus on more realistic test portfolios, including more diversified GMV portfolios as well as two risk parity propositions. Constraining portfolio weights brings more realistic GMVs that have lower turnover and costs, though portfolio concentration is still high. Importantly, while the implicit shrinkage imposed by the asset weight constraints renders the opportunity set for linear and nonlinear shrinkage methods marginal, dynamic covariance modeling is still rewarded. Interestingly, VCV modeling via the RiskMetrics approach is found to be on par with more intricate DCC-NLS modeling when considering constrained GMV portfolios.

Lastly, we investigate two salient risk parity strategies, ERC and HRP. Both improve upon the tested GMV variants in terms of portfolio diversification and turnover. By design, these strategies operate at a higher absolute level of volatility, and we find them offering even less opportunity for the various VCV estimators to impact the associated ex post volatilities. By and large, the relative ranking of VCVs for the risk parity use-cases mirrors that of the constrained GMV use-cases, and dynamic VCV modeling is still deemed relevant (with DCC-NLS outperforming RiskMetrics-NLS in a statistically significant manner). In terms of economic significance, however, the more sophisticated estimators hardly improve the ex post volatility of these more realistic portfolios, especially when compared to the unconstrained GMV case.

Our findings emphasize the importance of risk-based portfolio selection when evaluating VCV estimators. As of this writing, the ex post volatility of the traditional GMV portfolio is the key validation criterion in the academic literature. We recommend against directly implementing estimators that were empirically found to be optimal solely based on this criterion alone. Instead, estimators should be evaluated based on the objective of the risk-based investor whilst enforcing meaningful investment constraints. At the minimum, we propose using a long-only GMV portfolio with maximum-weight constraints and a transaction cost penalty as the starting point for evaluating VCV matrix estimators in large asset universes. Such realistic test portfolios suggest that the overall room for improvement from a given VCV estimator is limited, but one might though make a difference that sometimes is deemed statistically significant.

7. Disclosure of interest

Clint Howard, Maarten Jansen, and Harald Lohre are employees of Robeco Institutional Asset Management. Robeco is a global investment management firm that may or may not use similar techniques as described in this paper. The views expressed here are those of the authors and not necessarily those of Robeco. We have disclosed these interests fully to Taylor & Francis.

8. Declaration of funding

No funding was received.

References

Abdi, F. and Ranaldo, A. (2017), A simple estimation of bid-ask spreads from daily close, high, and low prices, *Review of Financial Studies* **30**(12), 4437–4480.

Ardia, D., Bolliger, G., Boudt, K. and Gagnon-Fleury, J.-P. (2017), The impact of covariance misspecification in risk-based portfolios, *Annals of Operations Research* **254**(1), 1–16.

Ardia, D., Guidotti, E. and Kroencke, T. A. (2024), Efficient estimation of bid-ask spreads from open, high, low, and close prices, *Journal of Financial Economics* **161**, 103916.

Boyd, S., Johansson, K., Kahn, R., Schiele, P. and Schmelzer, T. (2024), Markowitz portfolio construction at seventy, *Journal of Portfolio Management* **50**(8), 117–160.

Brandt, M. W. (2010), Portfolio choice problems, *in* 'Handbook of Financial Econometrics: Tools and Techniques', Elsevier, chapter 5, pp. 269–336.

Briere, M., Lehalle, C.-A., Nefedova, T. and Raboun, A. (2020), Modeling transaction costs when trades may be crowded: A Bayesian network using partially observable orders imbalance, *Machine Learning for Asset Management: New Developments and Financial Applications* pp. 387–430.

Caner, M., Medeiros, M. and Vasconcelos, G. F. (2023), Sharpe ratio analysis in high dimensions: Residual-based nodewise regression in factor models, *Journal of Econometrics* **235**(2), 393–417.

Clarke, R., De Silva, H. and Thorley, S. (2011), Minimum-variance portfolio composition, *Journal of Portfolio Management* **37**(2), 31–45.

Clarke, R. G., De Silva, H. and Thorley, S. (2006), Minimum-variance portfolios in the U.S. equity market, *Journal of Portfolio Management* **33**(1), 10–24.

Conlon, T., Cotter, J. and Kynigakis, I. (2021), Machine learning and factor-based portfolio optimization, *Michael J. Brennan Irish Finance Working Paper Series Research Paper No. 21–6* .

Corwin, S. A. and Schultz, P. (2012), A simple way to estimate bid-ask spreads from daily high and low prices, *Journal of Finance* **67**(2), 719–760.

De Nard, G., Engle, R. F., Ledoit, O. and Wolf, M. (2022), Large dynamic covariance matrices: Enhancements based on intraday data, *Journal of Banking & Finance* **138**, 106426.

De Nard, G., Ledoit, O. and Wolf, M. (2021), Factor models for portfolio selection in large dimensions: The good, the better and the ugly, *Journal of Financial Econometrics* **19**(2), 236–257.

DeMiguel, V., Garlappi, L., Nogales, F. J. and Uppal, R. (2009), A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms, *Management Science* **55**(5), 798–812.

DeMiguel, V., Garlappi, L. and Uppal, R. (2009), Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?, *Review of Financial Studies* **22**(5), 1915–1953.

DeMiguel, V., Martin-Utrera, A. and Nogales, F. J. (2013), Size matters: Optimal calibration of shrinkage estimators for portfolio selection, *Journal of Banking & Finance* **37**(8), 3018–3034.

Engle, R. (2002), Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models, *Journal of Business & Economic Statistics* **20**(3), 339–350.

Engle, R. F., Ledoit, O. and Wolf, M. (2019), Large dynamic covariance matrices, *Journal of Business & Economic Statistics* **37**(2), 363–375.

Fama, E. F. and French, K. R. (1993), Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* **33**(1), 3–56.

Fan, J., Liao, Y. and Mincheva, M. (2013), Large covariance estimation by thresholding principal orthogonal complements, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **75**(4), 603–680.

Fan, J., Zhang, J. and Yu, K. (2012), Vast portfolio selection with gross-exposure constraints, *Journal of the American Statistical Association* **107**(498), 592–606.

Frazzini, A., Israel, R. and Moskowitz, T. J. (2012), Trading costs of asset pricing anomalies, *Fama-Miller Working Paper, Chicago Booth Research Paper* (14-05).

Friedman, J., Hastie, T. and Tibshirani, R. (2008), Sparse inverse covariance estimation with the graphical lasso, *Biostatistics* **9**(3), 432–441.

Haugen, R. A. and Baker, N. L. (1991), The efficient market inefficiency of capitalization–weighted stock portfolios, *Journal of Portfolio Management* **17**(3), 35–40.

Holm, S. (1979), A simple sequentially rejective multiple test procedure, *Scandinavian Journal of Statistics* pp. 65–70.

Jagannathan, R. and Ma, T. (2003), Risk reduction in large portfolios: Why imposing the wrong constraints helps, *Journal of Finance* **58**(4), 1651–1683.

Jain, P. and Jain, S. (2019), Can machine learning-based portfolios outperform traditional risk-based portfolios? The need to account for covariance misspecification, *Risks* **7**(3), 74.

James, W. and Stein, C. (1961), Estimation with quadratic loss, *in* 'Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability: Contributions to the Theory of Statistics', Vol. 1, University of California Press, pp. 361–380.

Jobson, J. D. and Korkie, B. (1980), Estimation for Markowitz efficient portfolios, *Journal of the American Statistical Association* **75**(371), 544–554.

Jurczenko, E. and Teiletche, J. (2018), Active risk-based investing, *Journal of Portfolio Management* **44**(3), 56–65.

Kan, R. and Zhou, G. (2007), Optimal portfolio choice with parameter uncertainty, *Journal of Financial and Quantitative Analysis* **42**(3), 621–656.

Kourtis, A., Dotsis, G. and Markellos, R. N. (2012), Parameter uncertainty in portfolio selection: Shrinking the inverse covariance matrix, *Journal of Banking & Finance* **36**(9), 2522–2531.

Ledoit, O. and Wolf, M. (2003), Improved estimation of the covariance matrix of stock returns with an application to portfolio selection, *Journal of Empirical Finance* **10**(5), 603–621.

Ledoit, O. and Wolf, M. (2004*a*), Honey, I shrunk the sample covariance matrix, *Journal of Portfolio Management* **30**(4), 110–119.

Ledoit, O. and Wolf, M. (2004*b*), A well-conditioned estimator for large-dimensional covariance matrices, *Journal of Multivariate Analysis* **88**(2), 365–411.

Ledoit, O. and Wolf, M. (2011), Robust performances hypothesis testing with the variance, *Wilmott* **2011**(55), 86–89.

Ledoit, O. and Wolf, M. (2012), Nonlinear shrinkage estimation of large-dimensional covariance matrices, *Annals of Statistics* **40**(2), 1024–1060.

Ledoit, O. and Wolf, M. (2015), Spectrum estimation: A unified framework for covariance matrix estimation and PCA in large dimensions, *Journal of Multivariate Analysis* **139**, 360–384.

Ledoit, O. and Wolf, M. (2017), Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets Goldilocks, *Review of Financial Studies* **30**(12), 4349–4388.

Ledoit, O. and Wolf, M. (2020), Analytical nonlinear shrinkage of large-dimensional covariance matrices, *Annals of Statistics* **48**(5), 3043–3065.

Ledoit, O. and Wolf, M. (2022*a*), Quadratic shrinkage for large covariance matrices, *Bernoulli* **28**(3), 1519–1547.

Ledoit, O. and Wolf, M. (2022*b*), The power of (non-)linear shrinking: A review and guide to covariance matrix estimation, *Journal of Financial Econometrics* **20**(1), 187–218.

Ledoit, O. and Wolf, M. (2024), Markowitz portfolios under transaction costs, *Working Paper Series/Department of Economics* (420).

Lee, W. (2011), Risk-based asset allocation: A new answer to an old question?, *Journal of Portfolio Management* **37**(4), 11–28.

Lesmond, D. A., Schill, M. J. and Zhou, C. (2004), The illusory nature of momentum profits, *Journal of Financial Economics* **71**(2), 349–380.

López de Prado, M. (2016), Building diversified portfolios that outperform out-of-sample, *Journal of Portfolio Management* **42**(4), 59–69.

Maillard, S., Roncalli, T. and Teïletche, J. (2010), The properties of equally weighted risk contribution portfolios, *Journal of Portfolio Management* **36**(4), 60–70.

Mantegna, R. N. (1999), Hierarchical structure in financial markets, *European Physical Journal B-Condensed Matter and Complex Systems* **11**(1), 193–197.

Markowitz, H. (1952), Portfolio selection, *Journal of Finance* **7**(1), 77–91.

Michaud, R. O. (1989), The Markowitz optimization enigma: Is 'optimized' optimal?, *Financial Analysts Journal* **45**(1), 31–42.

Nakagawa, K., Imamura, M. and Yoshida, K. (2018), Risk-based portfolios with large dynamic covariance matrices, *International Journal of Financial Studies* **6**(2), 52.

Novy-Marx, R. and Velikov, M. (2016), A taxonomy of anomalies and their trading costs, *Review of Financial Studies* **29**(1), 104–147.

RiskMetrics (1996), RiskMetrics – Technical document. Available at https://www.msci.com/documents/10199/ (last accessed: September 10, 2022).

Scherer, B. (2011), A note on the returns from minimum variance investing, *Journal of Empirical Finance* **18**(4), 652–660.

Stein, C. (1956), Variate normal distribution, *in* 'Proceedings of the 3rd Berkeley Symposium on Mathematical Statistics and Probability: Contributions to the Theory of Statistics', Vol. 1, University of California Press, p. 197.

Trucíos, C., Zevallos, M., Hotta, L. K. and Santos, A. A. (2019), Covariance prediction in large portfolio allocation, *Econometrics* **7**(2), 19.

Tumminello, M., Aste, T., Di Matteo, T. and Mantegna, R. N. (2005), A tool for filtering information in complex systems, *Proceedings of the National Academy of Sciences* **102**(30), 10421–10426.

Zhao, Z., Ledoit, O. and Jiang, H. (2023), Risk reduction and efficiency increase in large portfolios: Gross-exposure constraints and shrinkage of the covariance matrix, *Journal of Financial Econometrics* **21**(1), 73–105.

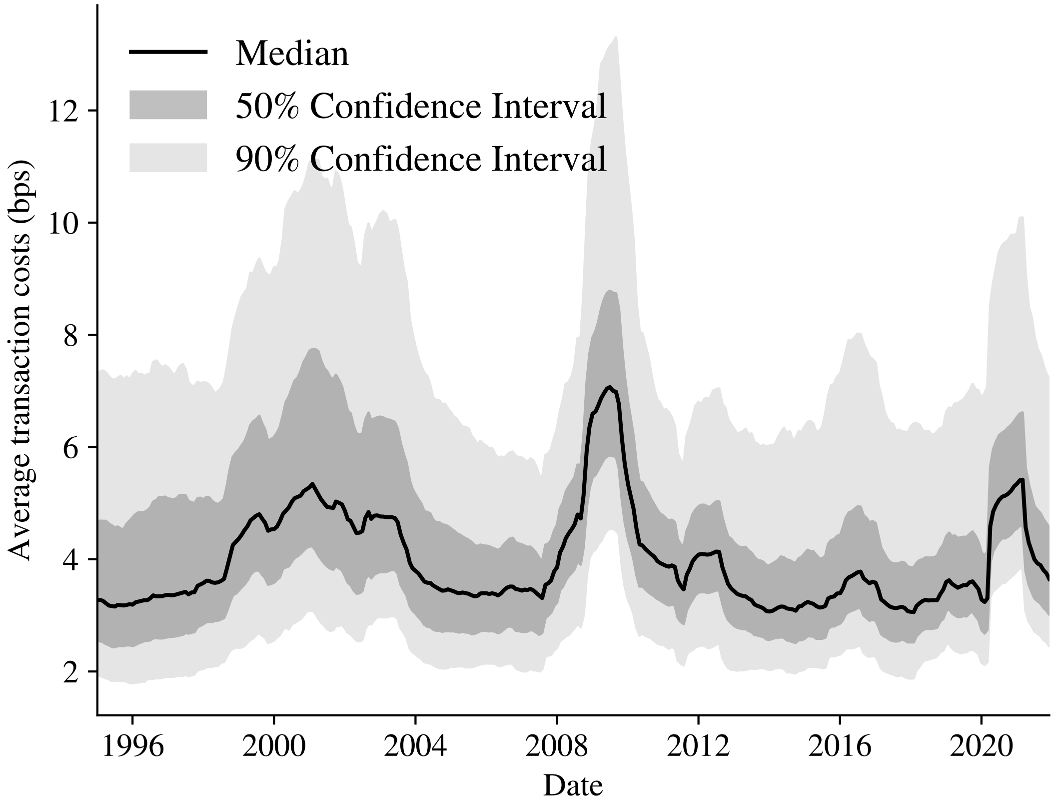

Figure 1: Quantile plots of the estimated stock-specific transaction costs
This plot shows the cross-sectional median and the top/bottom 5% and 25% quantiles of the stock-specific transaction costs in basis points. The sample period is from January 1, 1995, to December 31, 2021.

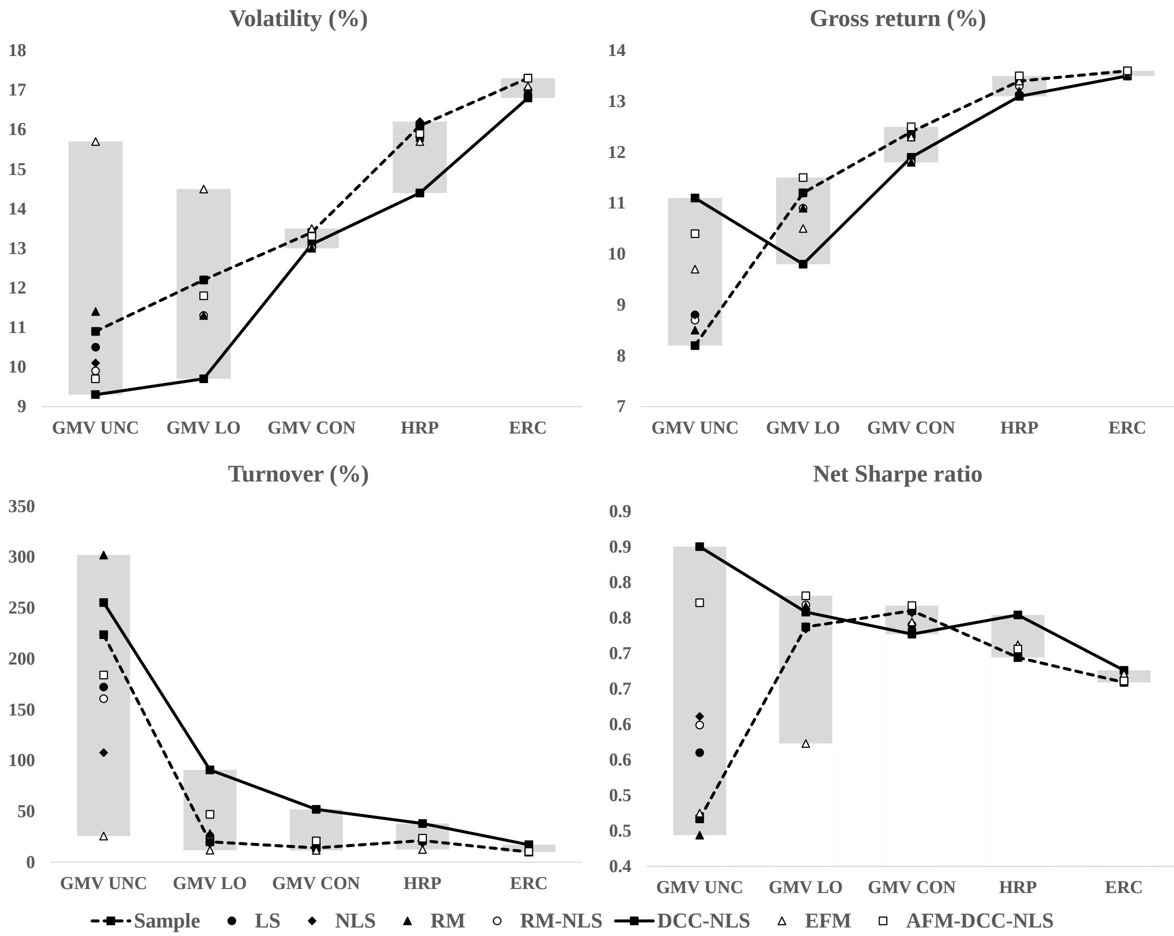

Figure 2: Key portfolio properties across tested VCV estimators and portfolios
This figure presents the key portfolio properties obtained from implementing alternative VCV estimators for different test portfolios with 500 US stocks. Details on the used acronyms for estimators and test portfolios can be taken from Table 1. The upper left panel shows ex post volatility, the upper right shows annualized gross returns, the lower left panel shows average monthly portfolio one-way turnover, and the lower right panel shows net Sharpe ratios. Within each panel the range of observed portfolio statistics is highlighted by a colored bar. The portfolio metrics are calculated over the full out-of-sample period from January 1, 1995, to December 31, 2021.

Table 1: Description of the risk-based portfolios and VCV matrix estimators
This table presents a descriptive glossary of the risk-based portfolio construction methods (Panel A) and VCV matrix estimators methods (Panel B) that we use. The GMV UNC portfolio does not use long-only constraints. All other portfolios use a long-only constraint. The far right Benchmark column in Panel B displays the choice of benchmark estimator that each VCV matrix estimator is evaluated against. The estimators below the horizontal line in Panel B make use of a factor structure.

Panel A: VCV matrix estimators

Portfolio	Description
GMV UNC	Unconstrained global minimum-variance portfolio
GMV LO	Global minimum-variance portfolio with long-only constraints
GMV CON	Global minimum-variance portfolio with long-only constraints, maximum-weight constraints and a transaction cost penalty
ERC	Equal risk contribution portfolio
HRP	Hierarchical risk parity portfolio

Panel B: Risk-based portfolios

Estimator	Description	Benchmark
Sample	Sample estimator	-
LS	Linear shrinkage toward a scalar multiple of the identity matrix	Sample
NLS	Quadratic shrinkage of the inverse eigenvalues	LS
RM	RiskMetrics estimator	Sample
RM-NLS	RiskMetrics estimator with NLS	NLS
DCC-NLS	Dynamic conditional correlation model with NLS	RM-NLS
EFM	Exact factor model with diagonal residual VCV matrix	Sample
AFM-DCC-NLS	Approximate factor model with the DCC-NLS estimator used for the correlation targeting matrix	DCC-NLS

Table 2: Description of performance metrics

This table presents a descriptive glossary of the presented portfolio performance metrics. For the calculation of these metrics, we use the daily portfolio returns (R_t), the daily risk-free rate ($R_{f,t}$), the net returns after taking transaction costs into account ($R_{net,t}$), the asset weights at the beginning of the month after rebalancing ($w_{i,t}$), the asset weights at the end of the month before rebalancing ($w^*_{i,t}$), and the monthly transaction costs (TC_t) computed using the approach of Briere et al. (2020).

Acronym	Metric	Definition
Vol.	Volatility	$Vol = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left(R_t - \frac{1}{T}\sum_{t=1}^{T} R_t\right)^2}$
Ret.	Expected returns	$Ret = \frac{1}{T}\sum_{t=1}^{T} R_t$
SR	Sharpe ratio	$SR = \frac{1}{T}\sum_{t=1}^{T}\left(R_t - R_{f,t}\right)/Vol$
NSR	Net Sharpe ratio	$NSR = \frac{1}{T}\sum_{t=1}^{T}(R_{net,t} - R_{f,t})/\text{Vol}$
TO	Monthly turnover	$TO = \frac{1}{T}\sum_{t=1}^{T}\lvert w_{t+1} - w_t^*\rvert$
TC	Monthly transaction costs	$TC = \frac{1}{T}\sum_{t=1}^{T} TC_t$
POS	Average monthly number of non-negligible positions	$POS = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_t} 1_{\lvert w_{i,t}\rvert \geq 10bps}$
MAXW	Average monthly top 10 stock positions	$MAXW = \frac{1}{T}\sum_{t=1}^{T}\sum_{j=1}^{10} w_{j,t}^{top10}$
WEFF	Effective portfolio weights	$WEFF = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{\sum_{i=1}^{N_t} w_{i,t}^2}$
GEXP	Gross exposure	$GEXP = \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{N_t}\lvert w_{i,t}\rvert$

Table 3: Ex post volatility of the risk-based portfolios
This table presents the ex post volatility of risk-based portfolios of the universe with 500 US stocks driven by various VCV matrix estimators over the full out-of-sample period (Panel B) from January 1, 1995, to December 31, 2021, and during high- (Panel C) and low-volatility (Panel D) regimes. High- and low-volatility regimes correspond to dates on which the VIX index is above or below its five-year moving average, respectively. Panel A presents the realized volatility of the three benchmark portfolios. The top three performers of each risk-based portfolio are in bold face. * and ** indicate a 5% and 1% statistically significant reduction in ex post volatility after a Holm correction for multiple testing. Differences in ex post volatility are tested as follows: (1) LS versus Sample (2) NLS versus LS, (3) RM versus Sample, (4) RM-NLS versus NLS, (5) DCC-NLS versus RM-NLS, (6) EFM versus Sample, and (7) AFM-DCC-NLS versus DCC-NLS.

Panel A: Benchmark portfolios

	VW	EW	IV
Full sample	18.92	19.38	16.84
High volatility regime	25.32	25.75	22.51
Low volatility regime	10.92	11.54	9.76

Panel B: Full sample

	GMV UNC	GMV LO	GMV CON	ERC	HRP
Sample	10.91	12.24	13.36	17.31	16.11
LS	10.49**	12.23	13.36	17.31	16.14
NLS	10.08**	12.22	13.36	17.29**	16.16
RM	11.36	**11.31****	**13.00****	**17.04****	**15.80****
RM-NLS	**9.91**	**11.33****	**12.99****	**17.02****	15.83**
DCC-NLS	**9.33***	**9.74****	**13.10**	**16.77****	**14.40****
EFM	15.70	14.50	13.55	17.07**	**15.71****
AFM-DCC-NLS	**9.69**	11.78	13.34	17.30	15.89

Panel C: High volatility regime

	GMV UNC	GMV LO	GMV CON	ERC	HRP
Sample	13.45	15.95	17.60	23.06	21.47
LS	13.05**	15.95	17.60	23.07	21.53
NLS	12.79*	15.93	17.58	23.03**	21.54
RM	14.03	**14.71****	**17.19****	**22.69****	**21.04****
RM-NLS	**12.58**	**14.75****	**17.16****	**22.66****	21.08**
DCC-NLS	**12.09**	**12.52****	**17.34**	**22.28****	**19.11****
EFM	19.88	18.59	17.72	22.75**	**20.94****
AFM-DCC-NLS	**12.37**	15.21	17.59	23.07	21.21

Panel D: Low volatility regime

	GMV UNC	GMV LO	GMV CON	ERC	HRP
Sample	8.19	7.86	8.23	10.18	9.45
LS	7.72**	7.82**	8.23	10.18	9.45
NLS	7.03**	7.82	8.25	10.17**	9.48
RM	8.49	**7.30****	**7.89****	10.06**	9.32**
RM-NLS	**6.90**	**7.29****	**7.89****	**10.04****	9.33**
DCC-NLS	**6.12****	**6.54****	**7.91**	**9.99****	**8.63****
EFM	11.04	9.81	8.58	**10.01****	**9.22****
AFM-DCC-NLS	**6.65**	7.80	8.19	10.13	**9.26**

Table 4: Performance overview of the risk-based portfolios
This table presents the performance statistics for the risk-based portfolios. The first three rows present the performance statistics for the three benchmark portfolios. The results presented are for the universe with 500 US stocks using a five year estimation period over the full out-of-sample period from January 1, 1995, to December 31, 2021. We present annualized portfolio volatility (%) (Vol.), annualized portfolio return (%) (Ret.), annualized portfolio Sharpe ratio (SR), annualized portfolio Sharpe ratio net of transaction costs (NSR), annualized portfolio transaction costs (bps) (TC), average monthly portfolio one-way turnover (%) (TO), average monthly number of positions with weight >0.1% (POS), average monthly sum of the largest 10% of positions (%) (MAXW), the average monthly effective portfolio weights (WEFF), and average monthly gross portfolio exposure (%) (GEXP).

Portfolio	Estimator	Vol.	Ret.	SR	NSR	TO	TC	POS	MAXW	WEFF	GEXP
VW	-	18.9	14.0	0.627	0.626	2.5	0.09	219	21.5	119	100
EW	-	19.4	14.2	0.623	0.620	10.3	0.42	500	2.0	500	100
IV	-	16.8	13.7	0.688	0.685	8.6	0.30	398	6.0	367	100
GMV UNC	Sample	10.9	8.2	0.554	0.467	223.7	7.50	470	34.4	6	723
	LS	10.5	8.7	0.630	0.560	172.2	5.84	466	30.8	9	602
	NLS	10.1	8.8	0.659	0.611	107.7	3.75	457	26.1	17	447
	RM	11.4	8.5	0.557	0.444	302.0	10.27	473	34.6	5	812
	RM-NLS	9.9	8.8	0.670	0.599	160.7	5.67	463	29.1	13	518
	DCC-NLS	9.3	11.1	0.958	0.850	255.3	8.00	422	54.9	12	349
	EFM	15.7	9.7	0.484	0.475	25.7	0.91	412	24.3	48	232
	AFM-DCC-NLS	9.7	10.4	0.848	0.771	184.0	5.96	440	41.6	13	420
GMV LO	Sample	12.2	11.2	0.743	0.737	20.1	0.55	40	66.0	18	100
	LS	12.2	11.2	0.743	0.737	19.6	0.55	44	62.0	21	100
	NLS	12.2	11.2	0.739	0.734	18.3	0.52	54	51.6	29	100
	RM	11.3	10.9	0.775	0.766	28.2	0.75	38	67.3	17	100
	RM-NLS	11.3	10.9	0.776	0.768	25.7	0.71	48	56.0	25	100
	DCC-NLS	9.7	9.8	0.788	0.758	90.7	2.35	29	78.0	11	100
	EFM	14.5	10.5	0.576	0.573	11.8	0.31	40	56.4	24	100
	AFM-DCC-NLS	11.8	11.5	0.794	0.781	47.1	1.22	36	70.0	16	100
GMV CON	Sample	13.4	12.4	0.764	0.760	14.0	0.44	110	10.1	104	100
	LS	13.4	12.3	0.764	0.760	13.9	0.44	111	10.1	105	100
	NLS	13.4	12.3	0.761	0.757	13.5	0.43	113	10.1	106	100
	RM	13.0	11.8	0.745	0.739	18.2	0.57	110	10.1	104	100
	RM-NLS	13.0	11.8	0.745	0.740	17.7	0.55	112	10.1	105	100
	DCC-NLS	13.1	11.9	0.742	0.727	51.9	1.60	110	10.1	104	100
	EFM	13.5	12.3	0.747	0.744	11.7	0.36	109	10.1	104	100
	AFM-DCC-NLS	13.3	12.5	0.774	0.767	21.0	0.67	111	10.1	105	100
ERC	Sample	17.3	13.6	0.662	0.659	10.1	0.38	481	5.3	414	100
	LS	17.3	13.6	0.662	0.659	10.1	0.38	482	5.3	415	100
	NLS	17.3	13.6	0.662	0.659	10.1	0.38	480	5.5	410	100
	RM	17.0	13.5	0.668	0.665	10.4	0.39	481	5.6	407	100
	RM-NLS	17.0	13.5	0.668	0.665	10.4	0.39	480	5.7	403	100
	DCC-NLS	16.8	13.5	0.680	0.676	17.2	0.61	475	6.0	393	100
	EFM	17.1	13.6	0.674	0.671	10.0	0.37	464	6.4	382	100
	AFM-DCC-NLS	17.3	13.6	0.664	0.661	10.5	0.39	478	5.8	397	100
HRP	Sample	16.1	13.4	0.700	0.694	21.3	0.70	372	7.9	323	100
	LS	16.1	13.5	0.702	0.696	20.9	0.69	373	7.7	326	100
	NLS	16.2	13.5	0.704	0.698	18.8	0.62	377	7.2	338	100
	RM	15.8	13.2	0.700	0.694	22.9	0.74	370	8.2	317	100
	RM-NLS	15.8	13.3	0.704	0.699	20.6	0.68	373	7.6	330	100
	DCC-NLS	14.4	13.1	0.765	0.754	38.1	1.17	345	14.1	238	100
	EFM	15.7	13.4	0.715	0.712	12.5	0.42	353	10.0	272	100
	AFM-DCC-NLS	15.9	13.5	0.713	0.706	23.6	0.77	366	8.6	305	100

Table 5: Risk-based portfolio multi-factor exposures

This table presents the results from regressing the time series of monthly excess portfolio returns derived from the portfolio/estimator combinations against five asset pricing factors: Market minus risk-free rate (MKT), Size (SMB), Value (HML), Momentum (WML), and Low-Volatility (LOWVOL). The portfolios are computed for the universe with 500 US stocks using a five year estimation window over the full-out-of-sample period from January 1, 1995, to December 31, 2021. The GMV, ERC and HRP portfolios are estimated using the RM-NLS VCV estimator (similar results are obtained for the other estimators). t-values are shown in parentheses, and computed using HC3 standard errors. * and ** indicate statistical significance at the 5% and 1% level, respectively. For the MKT factor, the t-values and significance level are shown with respect to 1.0.

Portfolio	VW	EW	IV	GMV UNC	GMV LO	GMV CON	ERC	HRP
α (%)	0.09	0.04	0.02	0.22	0.11	0.01	0.02	0.07
	(1.20)	(0.47)	(0.27)	(1.73)	(0.84)	(0.12)	(0.20)	(0.65)
MKT	0.96	0.97	0.83**	0.34**	0.38**	0.6**	0.86**	0.74**
	(-1.68)	(-0.73)	(-5.16)	(-12.78)	(-12.86)	(-9.67)	(-3.98)	(-6.59)
SMB	-0.12**	0.10**	0.04	-0.01	0.01	0.03	0.08**	0.05
	(-4.98)	(3.35)	(1.35)	(-0.18)	(0.21)	(1.01)	(2.82)	(1.48)
HML	-0.04	0.07	0.07	-0.06	-0.06	-0.01	0.06	-0.01
	(-1.51)	(1.63)	(1.43)	(-0.83)	(-0.92)	(-0.22)	(1.17)	(-0.12)
RMW	0.07*	0.12**	0.20**	0.16	0.13	0.19**	0.18**	0.19**
	(1.97)	(2.79)	(4.53)	(1.95)	(1.61)	(2.80)	(3.79)	(3.58)
CMA	0.04	0.07	0.12*	0.07	0.14	0.15*	0.12*	0.14*
	(0.89)	(1.20)	(2.16)	(0.83)	(1.56)	(1.98)	(2.10)	(1.97)
MOM	-0.01	-0.06**	-0.04*	0.04	0.02	0.01	-0.02	-0.02
	(-1.18)	(-3.20)	(-2.12)	(1.28)	(0.65)	(0.43)	(-1.11)	(-0.88)
LOWVOL	0.01	0.04	0.16**	0.21**	0.18**	0.24**	0.10**	0.15**
	(0.62)	(1.21)	(5.35)	(3.80)	(3.36)	(6.47)	(3.37)	(4.37)
R^2	0.94	0.91	0.89	0.43	0.47	0.75	0.89	0.85