Beyond GMV: Raising the bar for evaluating covariance matrix estimators*

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, 2024

Abstract

When validating variance-covariance (VCV) estimators based on the ex-post volatility of the global minimum variance (GMV) portfolio, we confirm the academic backing for considering shrinkage and covariance dynamics in modeling the VCV for equity portfolio construction. Yet, the underlying test portfolios are often impractical due to their high leverage, concentration, turnover, and transaction costs. Resorting to more practical GMV and risk-parity portfolios, we reveal a considerably reduced opportunity set for alternative VCV estimators. Although the implicit shrinkage in asset weight constraints makes further shrinkage redundant, accounting for the dynamics in the time-series dependence of asset returns remains statistically relevant.

JEL classification: C13, C55, C58, G11

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

*Corresponding author: Clint Howard, *email-address:* c.howard@robeco.com, *Postal address:* Weena 850, 3014DA Rotterdam, South Holland, The Netherlands
We are grateful to David Blitz, Khuram Chaudhry, Jeffrey Durieux, Amit Goyal, Danny Huang, Rob Huisman, Rasmus Lönn, Bart van der Grient, Sam van Meer, Stefan Mittnik, Michael Rockinger, Philipp Schiele, Jesper van Winden, and participants at the CEQURA Conference on Advances in Financial and Insurance Risk Management, 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 risk-based portfolio optimization 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 a myriad of alternative VCV estimators to address these limitations. These estimators are typically evaluated empirically using the ex-post volatility and Sharpe ratio of Markowitz's (1952) global minimum variance (GMV) portfolio.[1] With the exception of 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. 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 unconstrained GMV portfolio and ex-post volatility as the standard evaluation criteria for the practical relevance of alternative VCV estimators. Instead, we propose the use of a more realistic weight-constrained long-only GMV portfolio with transaction cost penalties and to consult its after-cost performance measures.

Using constituents of the S&P 500 index from January 1990 to December 2021, we first evaluate whether the choice of optimal VCV matrix estimator differs across various risk-based portfolios, where optimality is determined by the ex-post portfolio volatility. We then explore the pratical relevance of the risk-based portfolios by investigating characteristics beyond ex-post volatility. We evaluate the risk-adjusted returns, asset weights, turnover, transaction costs, factor exposures, and portfolio similarity, characteristics that have frequently been ignored in the portfolio management literature (Lesmond, Schill and Zhou, 2004; Frazzini, Israel and Moskowitz, 2012; Novy-Marx and Velikov, 2016).

To assess the practical relevance of VCV estimators, we explore GMV and risk-parity

[1] 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).

portfolios. Risk parity portfolios are effectively looking to also minimize portfolio variance but subject to a diversification constraint, hence, they naturally extend the set of test portfolios. As for GMV, consider the traditional unconstrained GMV portfolio, the long-only GMV portfolio, and we introduce a long-only GMV portfolio with a maximum-weight constraint and a transaction cost penalty to ensure a less concentrated portfolio after transaction costs. As for risk parity, we evaluate two risk-parity allocations that aim to maximize risk diversification: the equal risk contribution (ERC) portfolio of Maillard, Roncalli and Teïletche (2010) and the hierarchical risk-parity (HRP) portfolio of López de Prado (2016). ERC aims to ensure equal risk contribution from each asset in the portfolio, while HRP assumes a hierarchical structure between assets. HRP exploits this hierarchical structure by using hierarchical clustering to filter out the most important asset links and uses the found clusters to allocate to assets with an inverse variance scheme.

The literature on VCV estimators has advanced various streams of VCV improvements. To capture a wide range of both traditional and state-of-the-art VCV estimators, we focus on three key design choices: (i) shrinkage, (ii) time-dynamics, and (iii) factor structure. Combining these choices results in estimators with varying degrees of complexity. First, we use shrinkage methods to reduce the instability and prevent singularity of the VCV matrix estimators. We investigate the linear (LS) and non-linear (NLS) shrinkage estimators of Ledoit and Wolf (2004b, 2022a). We find that for the unconstrained GMV portfolio, linear shrinkage is effective in reducing ex-post portfolio volatility relative to the sample VCV estimator. In line with prior literature (Ledoit and Wolf, 2012, 2017), we document that non-linear shrinkage significantly outperforms linear shrinkage. However, the shrinkage benefits from linear and non-linear shrinkage diminish significantly when moving from unconstrained GMV portfolios to long-only risk-based portfolios, resonating with the shrinkage implicit in long-only constraints (Jagannathan and Ma, 2003).

The second design choice is to move from static to dynamic estimators to account for dynamic time-series dependence in asset returns. We focus on the DCC model of Engle (2002) combined with non-linear shrinkage methods (denoted DCC-NLS), following Engle et al. (2019). We benchmark this complex estimator to a simple dynamic VCV estimator driven by an exponentially weighted moving average scheme (i.e., the RiskMetrics approach

(RiskMetrics, 1996)). We consider both the standard RiskMetrics estimator (RM) as well as RiskMetrics augmented with non-linear shrinkage (RM-NLS). We find that these two dynamic estimators significantly outperform their static counterparts (the NLS estimator) in terms of minimizing ex-post volatility in all test portfolios, except for the unconstrained GMV portfolios. In fact, the performance of the RM estimator is on par with that of RM-NLS for all but the unconstrained GMV portfolio, indicating that further shrinkage is redundant. The more intricate DCC-NLS estimator only significantly outperforms the simpler RM-NLS estimator for the two risk-parity test portfolios, but not for the tested GMV portfolios. These results suggest using a simple approach, such as RiskMetrics, may achieve comparable performance to more sophisticated models such as DCC.

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. Factor models aim to improve the stability of VCV estimators by shrinking the component of the VCV matrix that is not driven by the risk factors. This reduces the dilution of the signals by structure-free shrinkage estimators (López de Prado, 2019) at the expense of introducing additional biases through the assumed factor structure (Ledoit and Wolf, 2003). Our analysis indicates that minimum-variance portfolios do not directly benefit from a factor structure. The factor structure shows more promise when moving beyond GMV. The exact factor model estimator achieves the third-lowest volatility for ERC and HRP. Moreover, alternative portfolio management applications, such as factor-based portfolios, are likely to benefit from a factor structure to align the VCV estimator with the very factors that are driving active portfolio risk,

Figure 1 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 14.3% and reduces down to 10.9% when applying the RiskMetrics approach together with non-linear 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 13.4% and 13.9%, 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 non-linear 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 1 about here>

We contribute to the literature that explores the intersection of VCV matrix estimators and risk-based portfolio construction by considering a large asset universe setting. For small asset universes, Ardia et al. (2017) and Jain and Jain (2019) apply three traditional VCV matrix estimators on a suite of risk-based portfolios (e.g., GMV and ERC).[2] However, their VCV matrix estimators become numerically unstable or even singular when the number of assets exceeds the sample size. This yields sub-optimal solutions with poor ex-post performance due to the magnification of estimation errors by the optimization algorithm (Michaud, 1989; López de Prado, 2016; Ledoit and Wolf, 2004*b*). López de Prado (2016; 2020) refers to this phenomenon as Markowitz's curse: the diversification benefits are overpowered by estimation errors in correlations and variances.[3] Importantly, Kan and Zhou (2007) show that the estimation errors in the variance become much more severe than errors in the mean when the number of assets increases for a constant sample size. This illustrates the need for well-conditioned VCV estimators for large asset universes. We expand on this literature by applying a wide range of state-of-the-art VCV estimators on several risk-based portfolio construction methods in a relevant investment universe, where we consider performance

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

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

measures beyond ex-post volatility.

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. Compared to the other GMV portfolios, this portfolio gives up some volatility reduction in return for higher net Sharpe ratios and more realistic portfolio characteristics.

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 set-up of the empirical analysis including the data descriptions and performance metrics, and the results. Section 5 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.

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, and Σ_t is the VCV matrix of dimension $N_t \times N_t$ at date t. This problem, that we refer to as unconstrained GMV (labeled GMV UNC), has the following analytical solution:

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

where ι is a vector of ones with length N_t. 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, thereby reducing net portfolio returns. Second, long-only constraints bring implicit shrinkage of the VCV matrix estimator (Jagannathan and Ma, 2003), thus help to mitigate the adverse portfolio effects of estimation error.[4]

2.1.2. Transaction cost penalty and maximum-weight constraints

Long-only GMV portfolios can 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 transaction cost penalty parameter. 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.[5] We estimate the stock-specific transaction costs $c_{t,i}$ using

[4] Zhao, Ledoit and Jiang (2023) compare direct shrinkage of the VCV matrix to imposing gross-exposure constraints. They find that non-linear 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 non-linear shrinkage methods have N_t degrees of freedom.

[5] Ledoit and Wolf (2022*a*) 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

the model of Briere, Lehalle, Nefedova and Raboun (2020) that requires open-high-low-close price data.

The maximum-weight constraints of 1% ensure that at least 100 positions are held and thus force the portfolio to be less concentrated/more diversified. Asset weight constraints also impose implicit shrinkage on the VCV. The specification we present here extends the formulation used in Ledoit and Wolf (2022*a*) by incorporating long-only and maximum-weight constraints alongside a transaction cost penalty. These constraints produce 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.[6]

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:

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

base their choice on the Sharpe Ratio. Because we consider minimum-variance portfolios, specifically, we base the choice of λ on a trade-off between the ex-post volatility and the average transaction costs rather than the Sharpe Ratio.

[6]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.

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

Importantly, this problem can be expressed as a minimum-variance optimization problem subject to a diversification constraint, see Maillard et al. (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 $w_{i,t} = \frac{\tilde{w}_{i,t}}{\sum_{i=1}^{N_t} \tilde{w}_{i,t}}$ and c_t is a constant. 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, it holds that:

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

thus the ERC portfolio can be interpreted as a middle-ground portfolio between GMV and EW portfolios.

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 on 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.[7]

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 and is a natural candidate as DeMiguel, Garlappi and Uppal (2009) show that this portfolio can be difficult to beat out-of-sample. The VW portfolio assigns asset weights proportional to their market cap and this market portfolio is a classic reference point. Finally, the IV portfolio minimizes portfolio variance without accounting for risk diversification benefits by assigning weights inverse to the stock's historical volatility, that is 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. Drawing inspiration from the literature, we consider estimators that increase complexity along three dimensions: (i) shrinkage, (ii) time-dynamics, and (iii) factor structure. This section introduces salient VCV matrix estimators that we use in our subsequent horse race of different risk-based portfolio allocations.

3.1. Sample estimators in large asset universes

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

[7] We refer the reader to López de Prado (2016) for further details.

$$S = \widehat{\mathbb{E}}\left[(r_t - \bar{r})(r_t - \bar{r})'\right] = \frac{1}{T-1}\sum_{t=1}^{T}(r_t - \bar{r})(r_t - \bar{r})', \tag{7}$$

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*).[8] This is again a manifestation of Markowitz's curse, where error maximization 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*b*) 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 non-linear.[9] The shrinkage estimators that we consider are solely driven by the asset return data.[10]

[8]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.

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

[10]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).

3.2.1. Linear shrinkage

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

$$\widehat{\Sigma} = (1-\kappa)S + \kappa q I_N, \tag{8}$$

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.[11] 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. Non-linear 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*c*). In contrast, non-linear shrinkage methods do not require assumptions on the characteristics of the true VCV matrix. Although non-linear shrinkage methods are more complex, they have significantly better out-of-sample performance (Ledoit and Wolf, 2022*c*). We use the Quadratic Inverse Shrinkage (QIS) estimator of Ledoit and Wolf (2022*b*), referred to as NLS throughout this paper. This estimator shrinks the inverse eigenvalues of the VCV matrix.[12]

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.

[11] We also consider linear shrinkage toward a constant correlation matrix (Ledoit and Wolf, 2003). Results are generally inferior compared to the LS estimator.

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

Similar to Engle et al. (2019), we apply non-linear 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*b*) NLS estimator and do not consider other shrinkage methods since Engle et al. (2019) find non-linear shrinkage to be most effective.

3.3.1. Dynamic conditional correlation model

We estimate the dynamic conditional correlation model with non-linear shrinkage (DCC-NLS) of Engle et al. (2019). The model extends Engle's (2019) DCC model, which models the time-varying conditional volatilities and correlations using a generalized autoregressive conditional heteroskedasticity (GARCH)-like process. Non-linear 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. 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. Although the NLS estimator of Ledoit and Wolf (2022*b*) is designed for sample VCV matrices in an iid sample, using NLS allows for the RM-NLS estimator to serve as a middle ground between the RM and DCC-NLS estimators.

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:

$$r_t = \alpha + Bf_t + u_t, \quad \forall t \in \{1, \dots, T\}, \tag{9}$$

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:

$$\Sigma_t = B'\Sigma_{f,t}B + \Sigma_{u,t}, \tag{10}$$

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 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 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 (2022c) 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

[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).

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. Horse racing VCV estimators in risk-based portfolios

4.1. Empirical design

4.1.1. Data

Our asset universe comprises constituents of the S&P 500 index 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 Refinitiv Datastream. 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 and VIX data from the FRED.[14]

VCV matrices are estimated using a moving window approach. While we have investigated windows of one, three, and five years, we focus on presenting results based on the three-year window, thus enabling a streamlined exposition. 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 part of the S&P 500 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.

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 2 presents the distribution of estimated transaction costs. The range of transaction costs is fairly wide in the early 1990s, with maximum stock-specific transaction costs exceeding 200bps. The median transaction cost started out at 80bps but quickly dropped to 50bps

[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.

around 2000 and below 20bps in 2005. Specifically, the median transaction cost is 4.3bps since 2005. These numbers are in line with other studies, e.g., Corwin and Schultz (2012), Abdi and Ranaldo (2017), and Ardia, Guidotti and Kroencke (2022).

<Insert Figure 2 about here>

4.1.2. 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. With regards to portfolio concentration, we calculate three measures. The average number of positions with an absolute weight greater than 0.001% (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 comparison point when comparing portfolio concentration within and across different portfolio construction methods.

4.1.3. Covariance estimator and portfolio overview

Table 1 provides an overview of the VCV estimators and risk-based portfolios used in the subsequent analyses.

<Insert Table 1 about here>

4.2. Evaluating VCV matrix estimators by ex-post portfolio volatility

Table 2 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 three-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 2 about here>

4.2.1. 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, i.e., the S&P 500, has an annualized volatility of 18.6%, while the equal-weighted portfolio's volatility is one percentage point higher (19.7%); the inverse variance portfolio has a volatility of 16.9%. Against this backdrop, the unconstrained GMV obtains an even lower volatility (14.3%).

Importantly, moving from the sample estimator to the shrinkage estimators clearly reduces the ex-post volatilities of the unconstrained GMV portfolio further, giving volatilies between 11.0% (NLS) and 11.9% (LS). Given that the unconstrained GMV portfolios can short assets, it is not surprising to see considerably lower ex-post volatilities than that of the long-only benchmark portfolios. Nevertheless, enforcing long-only constraints (while increasing the overall level of portfolio volatility) still brings about substantial volatility reduction relative to the benchmark portfolios. In addition, the range of volatilities is considerably shrunk relative to the unconstrained base case. However, imposing long-only constraints diminishes this added value of shrinkage estimators and the differences in ex-post volatility between the two shrinkage estimators. Under long-only constraints, the sample and shrinkage estimators both reach an ex-post volatility of 12.5%. The implicit shrinkage due to imposing the long-only constraints can explain this loss in volatility reduction for the shrinkage estimators (Jagannathan and Ma, 2003). Although the shrinkage estimators seem less useful for improving the stability of the estimator in a long-only setting, they may still help prevent singularity of the VCV matrix when the asset universe is very large.

In terms of linear versus non-linear shrinkage, the NLS estimator only clearly outperforms the LS estimator for the unconstrained GMV portfolio. For the long-only portfolios, the performance of NLS is roughly similar to that of LS (even though the GMV CON indicates a 5% significant improvement). Hence, we find no conclusive evidence of which static structure-free shrinkage estimator is best in terms of minimizing ex-post volatility for all tested risk-based portfolios.

4.2.2. 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 and the DCC-NLS estimators are among the worst performers in terms of minimizing ex-post volatility for the unconstrained GMV portfolio. Hence, the choice of shrinkage estimator is more important than the choice of dynamic estimator in the unconstrained GMV portfolio. The combination of RM with the NLS estimator yields the lowest ex-post volatility but the outperformance of RM-NLS over the static NLS estimator is not significant on a 5% level.

For the long-only GMV portfolios, however, both RM and DCC-NLS are among the best three estimators. Given that the benefits of shrinkage diminish once long-only constraints are imposed, the additional flexibility of the dynamic estimators seem a meaningful resort to further improve ex-post volatility. The good performance of the HRP portfolio with the DCC-NLS estimator is a natural extension of the findings of Jain and Jain (2019), who find the DCC estimator optimal in small asset universes.

4.2.3. 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.

4.2.4. Portfolio risk comparison

Having covered all VCV matrix design choices, we next discuss the general risk properties of the risk-based portfolios, as measured by the ex-post volatility. The choice of VCV matrix estimator is most impactful for the unconstrained GMV portfolio. Once long-only constraints are introduced, the opportunity set for VCV estimators is reduced, as the implicit shrinkage takes effect. This is apparent in GMV LO, but gets more prominent in GMV CON, only allowing for a narrow range of ex-post volatilities from 13.4% (RM-NLS) to 13.9% (EFM).

In unreported results, we compare the estimation windows and observe that the ex-post volatility decreases when the estimation window decreases for the long-only GMV, ERC, and HRP portfolios.[15] This indicates that the adverse effects of the higher concentration ratios are smaller after applying shrinkage than the benefits of estimating the VCV matrix excluding older asset returns. Conversely, the ex-post volatility of the unconstrained GMV portfolio increases or remains relatively constant when decreasing the estimation period for all VCV estimators except EFM.

4.2.5. Subperiod analysis

Panels C and D of Table 2 show how the ex-post volatility of the risk-based portfolios differs in high- and low-volatility periods. When the VIX index is above (below) its five-year moving average we consider the period a high-volatility (low-volatility) period and we thus

[15] These results are available upon request.

divide the sample period into 52% high-volatility periods and 48% low-volatility periods, respectively. Not surprisingly, the level of ex-post volatilities is the highest during high-volatility periods. For instance, average market volatility then stands at 25.0% whilst low volatility periods see a figure of 10.6%. 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.

4.3. 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 portfolio's 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.

4.3.1. Global minimum variance portfolios

Table 3 presents the detailed performance statistics for the benchmark portfolios, the four 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 1,062%. Notably, modeling a more structured VCV helps reducing such gross-exposure (with the EFM variant having the lowest gross-exposures, 236%). Still, these portfolios display unduly high portfolio turnover, ranging from 24.0% (EFM) to 556.3% (AFM-DCC-NLS) in terms of one-way monthly turnover. As a consequence, corresponding transaction costs eat into the GMV UNC performance, considerably reducing net returns and Sharpe ratios.

<Insert Table 3 about here>

Finally, the top ten names of GMV UNC make up a large share of the portfolio weights and the effective portfolio weights range between 3.0 (Sample) and 47.0 (EFM), highlighting the poor diversification properties of unconstrained GMV portfolios.[16] 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.

By design, long-only GMV variants come with a reduced gross-exposure of 100%. Still, turnover statistics are elevated and range from 13.0% (EFM) to 95.1% (DCC-NLS)—suggesting that the consideration of covariance dynamics can become costly. Indeed, such DCC-modeling results in the lowest net Sharpe ratio for the DCC-NLS model (0.46). Modeling dynamics via RiskMetrics is less turnover-intensive, rendering the RM-NLS variant the long-only GMV portfolio with the highest net Sharpe ratio (0.63). Notwithstanding, long-only GMV portfolios are generally too concentrated to be considered viable alternatives in practice; the effective portfolio weights are consistently below 32.0. This finding is in line with Clarke et al. (2011), who rationalize that the long-only GMV portfolio tends to only select assets with low market exposures.

To compare the long-only GMV portfolio for different VCV estimators, Table 4 shows the average number of assets in the top ten largest portfolio positions of both two portfolios and their mutual tracking errors. Given the explicit shrinkage effects of long-only constraints, the static structure-free shrinkage estimators display the lowest mutual tracking errors with the sample estimator (0.49% for LS, 1.49 for NLS). Of their top ten portfolio positions, 9.65 overlap with the sample estimator for LS and 8.56 for NLS, on average. The estimators with the largest volatility differentials have the highest tracking errors and largest number of non-overlapping positions. This holds for the best estimators in terms of minimizing ex-post volatility DCC-NLS and the worst estimator EFM (a mutual tracking error of 8.48% and only 3.10 mutual positions).

<Insert Table 4 about here>

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

[16] POS, MAXW, and WEFF are all computed using absolute portfolio weights.

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.[17] Unsurprisingly, 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, resulting in consistently higher net Sharpe ratio compared to the GMV LO portfolios. On average, the transaction costs drop by more than 50% compared to the long-only portfolios. 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. Table A.1 shows the effect of different transaction cost penalties on ex-post volatility and transaction costs. 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.

4.3.2. Beyond GMV portfolios

Table 3 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. Except for the DCC-NLS VCV, ERC shows single-digit turnover numbers. 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 (488), the HRP portfolio is more concentrated than the ERC portfolio as its top ten names consume between 7.0–11.7% of the total portfolio (relative to 5.3–7.1% for ERC). Moreover, the number of effective portfolio positions is consistently the highest for the ERC portfolio ranging from 366.0 (EFM) to 411.0 (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 1–3bps for ERC which is on par with naive EW strategies but around 4 times higher than the transaction cost of the VW market portfolio.

[17] The 10% threshold is marginally exceeded for some estimators due to rounding.

The emerging net returns are on par or higher than those of the market portfolio, yet, the corresponding net Sharpe ratios are very much comparable to that of the GMV LO portfolios and smaller than the GMV CON portfolios, owing to the higher risk level of the risk-parity strategies.

Examining the effect of the choice of VCV estimator for ERC and HRP portfolios, we observe a spread of only 0.4% (ERC) and 0.8% (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 AFM-DCC-NLS, 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.

4.3.3. 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 portfolio's 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 RM-NLS estimator estimated over the three-year estimation window.[18] 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 market betas explain the different return levels of the GMV portfolios with

[18]Other VCV estimators and estimation windows give similar results that are available upon request.

the average return across VCV estimators increasing from only 8.7% (GMV UNC) to 10.0% (GMV LO) and 11.4% (GMV CON). 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 0.2–0.3. 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 a mild Profitability exposure for the long-only GMV portfolio, and significant exposures in GMV CON for the Profitability, Investment and Momentum factors. Taken together such factor exposures explain two thirds of the variation in long-only GMV portfolio returns and 80% 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. 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, both risk-parity portfolios come with positive Size and negative Momentum exposure. Also, the ERC portfolio yields significant Value exposures. With ERC being a middle-ground portfolio in between equal-weighting and minimum-variance, it is natural for it to inherit some of the EW portfolio's implicit rebalancing characteristic that prompts selling winners and buying losers. Overall, more than 90% 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. 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.

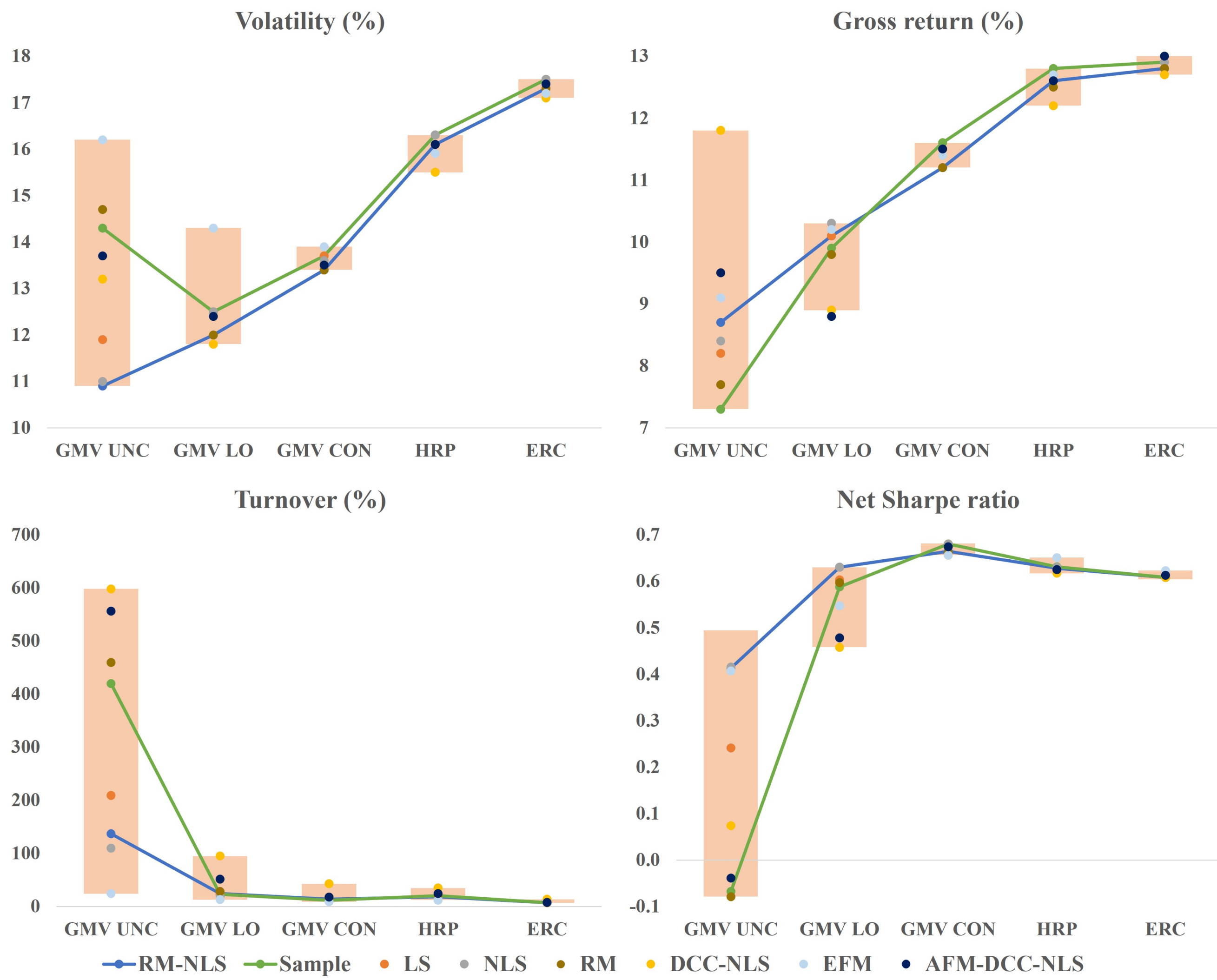
Against this backdrop, we focus on more realistic test portfolios, including more diversified GMV portfolios as well as two risk-parity propositions. Constraining portfolio weights brings more realistic GMVs that have lower turnover and costs, though portfolio concentration is still high. Importantly, while the implicit shrinkage imposed by the asset weight constraints renders the opportunity set for linear and non-linear 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.

Figure 1: 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. 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.



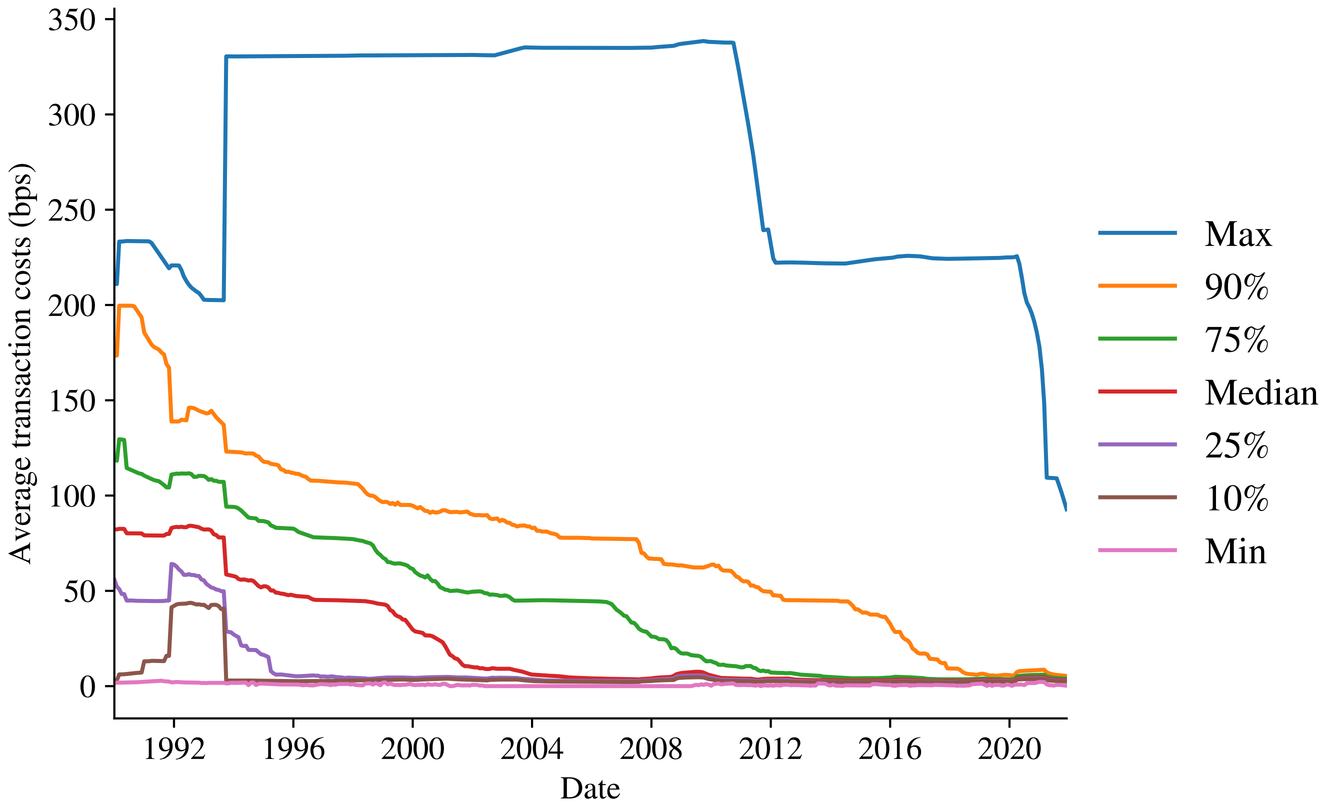


Figure 2: Quantile plots of the estimated stock-specific transaction costs
This plot shows the cross-sectional minimum, median, maximum, and the top/bottom 10% and 25% quantiles of the stock-specific transaction costs in basis points. The sample period is from January 1, 1990, to December 31, 2021.

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

Panel A: VCV matrix estimators

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

Panel B: Risk-based portfolios

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 transaction cost penalty
ERC	Equal risk contribution portfolio
HRP	Hierarchical risk-parity portfolio

Table 2: Ex-post volatility of the risk-based portfolios

This table presents the ex-post volatility of risk-based portfolios 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) periods. High- and low-volatility periods 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) Sample versus LS (2) LS versus NLS, (3) Sample versus RM, (4) NLS versus RM-NLS, (5) RM-NLS versus DCC-NLS, (6) Sample versus EFM, and (7) DCC-NLS versus AFM-DCC-NLS.

Panel A: Benchmark portfolios					
	VW	EW	IV		
Full sample	18.6	19.7	16.9		
High volatility regime	25.0	26.4	22.7		
Low volatility regime	10.6	11.3	9.6		
Panel B: Full sample					
	GMV UNC	GMV LO	GMV CON	ERC	HRP
Sample	14.3	12.5	13.6	17.5	16.3
LS	**11.9****	12.5	13.7	17.5	16.3
NLS	**11.0****	12.5	13.6*	17.5**	16.3
RM	14.7	**12.0****	**13.4****	17.3**	**16.1****
RM-NLS	**10.9**	**12.0****	**13.4****	**17.3****	16.1**
DCC-NLS	13.2	**11.8**	**13.5**	**17.1****	**15.5****
EFM	16.2	14.3	13.9	**17.2****	**15.9****
AFM-DCC-NLS	13.7	12.4	13.5	17.4	16.1
Panel C: High volatility regime					
	GMV UNC	GMV LO	GMV CON	ERC	HRP
Sample	17.6	16.4	18.1	23.4	21.8
LS	**14.9****	16.4	18.1	23.4	21.8
NLS	**14.2****	16.4	18.0**	23.4**	21.9
RM	18.1	**15.7****	**17.8****	23.2**	**21.5****
RM-NLS	**14.1**	**15.7****	**17.8****	**23.2****	21.6**
DCC-NLS	17.1	**15.6**	**18.0**	**22.8****	**20.7****
EFM	20.5	18.3	18.3	**23.0****	**21.3****
AFM-DCC-NLS	17.1**	16.2	18.0	23.4	21.5
Panel D: Low volatility regime					
	GMV UNC	GMV LO	GMV CON	ERC	HRP
Sample	10.8	7.9	8.2	10.0	9.3
LS	**8.6****	7.9**	8.2	10.0	9.4
NLS	**7.3****	7.9	8.2	10.0**	9.4
RM	11.1	**7.5****	**8.0****	9.9**	9.3**
RM-NLS	**7.2***	**7.5****	**8.0****	**9.9****	9.3**
DCC-NLS	8.8	**7.1***	**8.0**	**9.9**	**8.9****
EFM	11.3	9.8	8.7	**9.8****	**9.1****
AFM-DCC-NLS	10.1	7.8	8.1	10.0	**9.2**

Table 3: 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 S&P 500 universe using a three 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 >0.001% (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	TC	TO	POS	MAXW	WEFF	GEXP
VW	-	18.6	12.2	0.54	0.54	0.38	1.4	488	21.7	115	100
EW	-	19.7	13.4	0.57	0.56	1.59	6.8	488	2.1	488	100
IV	-	16.9	12.9	0.64	0.63	1.31	6.1	488	6.0	365	100
GMV UNC	Sample	14.3	7.3	0.36	-0.07	74.15	419.5	487	33.5	3	1062
	LS	11.9	8.2	0.51	0.24	38.40	208.8	487	27.6	7	704
	NLS	11.0	8.4	0.57	0.42	19.87	109.7	487	20.8	17	454
	RM	14.7	7.7	0.38	-0.08	80.70	459.5	487	29.6	3	1110
	RM-NLS	10.9	8.7	0.61	0.41	24.90	137.0	487	20.9	15	489
	DCC-NLS	13.2	11.8	0.73	0.07	104.09	598.0	487	53.7	4	764
	EFM	16.2	9.1	0.43	0.41	5.05	24.0	487	24.1	47	236
	AFM-DCC-NLS	13.7	9.5	0.54	-0.04	95.15	556.3	487	48.1	3	962
GMV LO	Sample	12.5	9.9	0.62	0.59	4.88	23.0	43	65.4	19	100
	LS	12.5	10.1	0.63	0.60	4.60	21.9	48	60.5	23	100
	NLS	12.5	10.3	0.66	0.63	3.88	19.0	60	49.4	32	100
	RM	12.0	9.8	0.64	0.60	5.93	28.4	41	66.9	18	100
	RM-NLS	12.0	10.1	0.67	0.63	5.01	24.5	56	51.9	30	100
	DCC-NLS	11.8	8.9	0.57	0.46	16.20	95.1	33	75.1	13	100
	EFM	14.3	10.2	0.56	0.55	2.65	13.0	45	56.3	26	100
	AFM-DCC-NLS	12.5	10.3	0.66	0.63	3.86	19.0	58	50.2	32	100
GMV CON	Sample	13.7	11.6	0.69	0.68	1.49	11.4	114	10.1	105	100
	LS	13.7	11.5	0.69	0.68	1.44	11.3	115	10.1	105	100
	NLS	13.6	11.5	0.69	0.68	1.36	10.9	118	10.0	106	100
	RM	13.4	11.2	0.67	0.66	1.79	14.0	114	10.1	105	100
	RM-NLS	13.4	11.2	0.68	0.66	1.66	13.5	117	10.0	106	100
	DCC-NLS	13.5	11.4	0.69	0.66	5.07	42.8	113	10.1	104	100
	EFM	13.9	11.4	0.66	0.66	1.22	8.9	114	10.1	105	100
	AFM-DCC-NLS	13.6	11.5	0.69	0.68	1.37	10.9	117	10.0	106	100
ERC	Sample	17.5	12.9	0.62	0.61	1.50	6.8	488	5.3	410	100
	LS	17.5	12.9	0.62	0.61	1.50	6.8	488	5.3	411	100
	NLS	17.5	12.9	0.62	0.61	1.50	6.8	488	5.4	407	100
	RM	17.3	12.8	0.62	0.61	1.57	7.1	488	5.4	407	100
	RM-NLS	17.3	12.8	0.62	0.61	1.57	7.1	488	5.5	404	100
	DCC-NLS	17.1	12.7	0.62	0.61	2.72	13.5	488	5.7	398	100
	EFM	17.2	13.0	0.63	0.62	1.56	7.6	488	7.1	366	100
	AFM-DCC-NLS	17.5	12.8	0.61	0.60	1.62	7.5	488	5.4	408	100
HRP	Sample	16.3	12.8	0.65	0.63	4.20	20.0	488	7.9	320	100
	LS	16.3	12.7	0.65	0.63	4.02	19.3	488	7.6	327	100
	NLS	16.3	12.7	0.65	0.63	3.35	16.4	488	7.0	339	100
	RM	16.1	12.5	0.65	0.63	4.43	21.1	488	8.1	317	100
	RM-NLS	16.1	12.6	0.65	0.63	3.73	18.1	488	7.1	336	100
	DCC-NLS	15.5	12.2	0.65	0.62	6.41	34.7	488	11.7	267	100
	EFM	15.9	12.7	0.66	0.65	2.38	11.7	488	10.0	277	100
	AFM-DCC-NLS	16.3	12.7	0.65	0.63	3.31	16.2	488	7.0	339	100

Table 4: Mutual tracking error and mutual positions of the long-only GMV portfolio This table presents the average number of assets that are in the top ten largest positions of both portfolios, in the upper right matrix. The mutual tracking error is presented in the lower left matrix. Mutual tracking error and positions are calculated in a pairwise manner for each VCV estimator for the long-only GMV portfolio. All values are computed for the S&P 500 universe using a three year estimation period over the full out-of-sample period from January 1, 1995, to December 31, 2021.

		Number of Assets							
		Sample	LS	NLS	RM	RM-NLS	DCC-NLS	EFM	AFM-DCC-NLS
Mutual Tracking Error (%)	Sample		9.65	8.56	7.51	7.29	4.31	5.11	7.20
	LS	0.49		8.79	7.53	7.43	4.32	5.23	7.21
	NLS	1.49	1.11		7.26	7.61	4.22	5.65	6.89
	RM	1.94	1.99	2.35		8.78	4.67	4.80	6.60
	RM-NLS	2.13	1.94	1.71	1.36		4.58	5.20	6.47
	DCC-NLS	5.17	5.11	5.19	4.91	4.92		3.10	5.40
	EFM	7.06	6.98	6.60	7.29	6.89	8.48		4.82
	AFM-DCC-NLS	2.75	2.73	2.99	2.98	3.10	4.60	7.30	

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 using a three 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 (%)	0.00	0.04	0.02	0.05	0.02	0.01	0.01	0.01
	(0.09)	(0.72)	(0.28)	(0.29)	(0.19)	(0.07)	(0.16)	(0.08)
MKT	0.99	1.01	0.86**	0.34**	0.51**	0.64**	0.90**	0.84**
	(-1.21)	(0.46)	(-6.23)	(-11.17)	(-12.79)	(-10.40)	(-4.25)	(-7.02)
SMB	-0.15**	0.10**	0.05	-0.02	0.04	0.05	0.09**	0.08**
	(-14.18)	(3.62)	(1.92)	(-0.32)	(0.88)	(1.33)	(3.42)	(2.83)
HML	-0.01	0.16**	0.11*	-0.09	-0.04	-0.04	0.11**	0.07
	(-1.41)	(4.62)	(2.52)	(-0.83)	(-0.58)	(-0.71)	(2.94)	(1.62)
RMW	0.05**	0.12**	0.21**	0.13	0.15*	0.22**	0.21**	0.25**
	(3.83)	(3.35)	(5.77)	(1.40)	(2.32)	(3.85)	(5.38)	(6.44)
CMA	0.06**	0.10*	0.13**	0.10	0.15	0.18**	0.13**	0.16**
	(3.82)	(2.11)	(2.89)	(0.90)	(1.83)	(2.66)	(2.98)	(3.37)
MOM	-0.02**	-0.17**	-0.09**	0.05	-0.01	-0.05*	-0.10**	-0.08**
	(-2.98)	(-8.57)	(-4.78)	(0.76)	(-0.26)	(-1.97)	(-5.63)	(-4.91)
LOWVOL	0.01	0.07**	0.17**	0.30**	0.28**	0.24**	0.11**	0.15**
	(1.02)	(3.62)	(7.63)	(4.88)	(7.52)	(7.20)	(5.30)	(6.96)
R^2	0.99	0.96	0.94	0.37	0.67	0.80	0.94	0.93

Beyond GMV: Raising the bar for evaluating covariance matrix estimators

-Online Appendix-

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*

Table A.1: Ex-post volatility and transaction cost of the GMV CON portfolio with different transaction cost penalty weights.

The results presented are for the S&P 500 universe using a three year estimation period over the full out-of-sample period from January 1, 1995, to December 31, 2021. λ denotes the weight of the transaction cost penalty in the optimization problem. We present annualized portfolio volatility (%) (Vol.) and annualized portfolio transaction costs (bps) (TC).

	Vol.					TC				
λ	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
Sample	13.8	13.7	13.6	13.6	13.6	0.52	1.49	2.65	3.25	3.35
LS	13.8	13.7	13.6	13.6	13.6	0.49	1.44	2.58	3.20	3.30
NLS	13.8	13.6	13.6	13.6	13.6	0.46	1.36	2.44	3.02	3.12
RM	13.6	13.4	13.4	13.4	13.4	0.59	1.79	3.31	3.99	4.09
RM-NLS	13.6	13.4	13.4	13.4	13.4	0.55	1.66	3.16	3.81	3.91
DCC-NLS	13.6	13.5	13.5	13.5	13.5	1.47	5.07	9.09	10.09	10.21
EFM	14.0	13.9	13.9	13.9	13.9	0.44	1.22	2.02	2.40	2.46
AFM-DCC-NLS	13.7	13.5	13.5	13.5	13.5	0.68	2.19	4.08	4.80	4.91

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. (2022), Efficient estimation of bid-ask spreads from open, high, low, and close prices, *Working paper* .

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.

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**. Article 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. 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.

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*), Markowitz portfolios under transaction costs, *Working Paper Series/Department of Economics* (420).

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

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

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.

López de Prado, M. (2019), Estimation of theory-implied correlation matrices, *Working paper* .

López de Prado, M. (2020), *Machine Learning for Asset Managers*, Cambridge University Press, chapter 7, pp. 92–104.

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.

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.