☆

When "Time Varying" Volatility meets "Transaction Cost" in Portfolio Selection

W. Qiao[a], D. Bu[b], A. Gibberd[c], Y. Liao[d,], T. Wen[e], E. Li[f]

[a]*School of Economics and Management, Ningbo University of Technology, China*
[b]*Department of Applied Finance, Macquarie University, Australia*
[c]*Department of Mathematics & Statistics, Lancaster University, UK*
[d]*Department of Applied Finance, Macquarie University, Australia*
[e]*College of Economics and Managment, Southwest University, China*
[f]*Sydney Boys High School, Australia*

June 13, 2023

Abstract

We propose a new strategy for mean-variance portfolio selection that tackles transaction costs and change detection in covariance matrix simultaneously. The new strategy solely rebalances the portfolio when change points are detected in the covariance matrix, striking an optimal trade-off between rebalancing the portfolio to capturing the recent information in return data and avoiding excessive trading. Our empirical results suggest favorable out-of-sample performance of the new strategy in terms of portfolio variance, portfolio turnovers and portfolio sharpe ratio with transaction cost. We also show that these gains come from the improved accuracy for covariance matrix prediction and the ability for tracking significant changes in covariance matrix.

*JEL classification:*C13, C51, C61, G11

Keywords: Covariance Matrix; Structural Change; Transaction Cost; Sparse Time Variation; Large Scale Portfolio

---

☆We thank seminar and conference participants at Lancaster University. Qiao acknowledges research support from Annual General Project of the National Social Science Fund (NO21BJY055), National Statistical Science Research Project (NO: 2022LY069), and Ningbo Small Business Growth Research Base Project (NO: 2020XQY007). Bu and Liao acknowledge research support from the Australian Research Council under project No. LP200301118.

Email addresses: qiaowen1982@163.com (W. Qiao), di.bu@mq.edu.au (D. Bu), a.gibberd@lancaster.ac.uk (A. Gibberd), yin.liao@mq.edu.au. (Y. Liao), tingwen1225@163.com (T. Wen), ethanli.syd@gmail.com (E. Li)

I. Introduction

We motivate our study from a mean-variance investor perspective, who adjusts their allocation based on changes in the estimated conditional covariance matrix of returns. The purpose of these portfolio strategies is to overcome the challenge that changing market conditions present to traditional static asset allocation, and hence requires investors to actively adjust portfolio weights based on sample information of the volatility dynamics. However, the presence of transaction costs makes otherwise optimal portfolio rebalancing costly, outweighing the benefit of tracking changes in volatility. This presented investors with a dilemma: how should we construct the portfolio in the presence of both transaction costs and structural changes in volatility? To address this issue, we propose a new portfolio strategy that can detect significant changes (or structural breaks) in the asset return covariance matrix even in a large dimension, and thereby rebalances portfolios solely when a change point has been detected in the covariance matrix. We examine the empirical performance of the new strategy through comparing it with other alternatives, and document that our new strategy simultaneously lowers portfolio risk exposure and turnovers, largely improving portfolio out-of-sample sharpe ratio with transaction costs. These results suggest that investors can benefit from change detection in covariance matrix when facing up both transaction costs and structural breaks in volatilities. To link with existing literatures, we offer two alternative interpretations to the new strategy from Bayesian portfolio choice and portfolio choice under turnover penalization perspectives.

To mitigate the impact of transaction cost in portfolio allocation, there is a large number of existing studies in the literature. For example, the strategy discussed in Gârlranu and Pedersen (2013) and Gârlranu and Pedersen (2016) suggest that investors should trade only partially toward the desired position, e.g. trading only 15% toward the zero-cost optimal targets each day. Another available strategy (Kirby and Ostdiek 2012) allows the sensitivity of portfolio weights to volatility changes to be adjusted via a tuning parameter, and thus, the portfolio turnover can be controlled to a desired low level. These strategies successfully control over portfolio turnovers through reducing trading frequency, but a common drawback here is the use of a constant trading reduction rate over time which ignores the time-varying trade-off between portfolio risk reduction and the increase of transaction costs. Intuitively, when the asset return volatility evolves smoothly, such as what we have seen during the market calm periods, using a large reduction rate for trading can be appropriate as it can remarkably reduce the transaction cost but without increasing portfolio risk exposure too much. Conversely, in more turbulent periods, the same level of trading reduction rate may not be appropriate as the portfolio weights need to be adjusted faster to adapt with the significantly changed market conditions. Therefore, to better strike the time-varying trade-off between portfolio risk and transaction cost, we need a mechanism to locally detect significant variations in asset return covariance matrix, and then suggest investors when significant updates on portfolio weights are needed.

So far, the canonical approach to assessing time variations in asset return volatility is to use a rolling window estimation (see DeMiguel et al. 2009b,a, Kirby and Ostdiek 2012, Kourtis et al. 2012, Goto and Xu 2015, for example). This analysis involves recursively estimating sample covariance matrix of the asset returns by re-weighting new observations according to the rolling window. Subsequently, analysis can be performed directly on the covariance matrix estimate to infer the dependence structure of asset returns when new observations arise. While rolling windows are a valuable tool for investigating dynamic changes, there are two main issues associated with its use. Firstly, the choice of window length can be a difficult parameter to tune. It is advisable to set the window length to be large enough to allow for a robust estimation but without making it too large, which can result in overlooking short-term fluctuations. Secondly, the rolling window faces the potential issue of variability between temporally adjacent estimates. This arises as a direct consequence of the fact that each covariance matrix across the rolling windows is estimated independently without any mechanism present to encourage temporal homogeneity. This additional variability can jeopardise the accuracy of the estimation as well as hugely increase turnovers in the context of portfolio allocation.

To address these issues, we propose a *regularized rolling window (RRW)* approach that regularizes the standard rolling window estimation using a penalty term that assists exploiting the temporal similarity between consecutive window estimates, resulting in a piecewise constant estimate. Specifically, our approach estimates the covariance matrix using traditional quasi-maximum likelihood but with an additional constraint to shrink the difference between contemporaneous and lagged estimates. With the constraint, the overall size of element-wise differences between the consecutive covariance matrix estimates are penalized, producing two-fold estimation effects: (i) the time variation of any element of the covariance matrix (e.g. the $(i, j)th$ entry) is set to be zero if the sample variation is below than a constraint whose strength relates to a turning parameter defined in the RRW optimisation problem; (ii) the time variation of any element of the covariance matrix is decreased toward zero by the magnitude of the threshold when the sample variation is above the threshold. As such, the penalty term leads the covariance matrix estimate to achieve both shrinkage and sparsity in time variations, eliminating unnecessary evolutions embedded in the standard rolling window analysis. The motivation of our RRW approach is similar to recent advances in machine learning and statistics on sparsity pursuit in time-varying regression parameters: the "fused lasso" estimator of Tibshirani et al. (2005) that operates in a linear (least-squares) regression setting and acts to shrink the insignificant changes in consecutive parameter estimates towards zero. We extend the idea to a multivariate setting, and employ the matrix version of "fused lasso", that is, "graphical fused lasso" algorithm of Gibberd and Nelson (2017) to solve the estimation problem.

Next, we apply our new covariance matrix estimate to mean-variance portfolio decision making to develop a new portfolio allocation strategy. Under our approach, the portfolios are re-balanced monthly based solely on the significant changes detected in the covariance matrix. Otherwise, the portfolio remains as what it was in the last period. We control the sensitivity of change detection in covariance matrix via a tuning parameter, and the tuning parameter can be hence treated as a measure of change detection aggressiveness and allows us to keep the turnover of the proposed strategies to a level competitive with other existed strategies, e.g. $1/N$ naive diversification. To link with existing literatures, we offer two alternative interpretations for the new strategy from both Bayesian portfolio choice and portfolio choice with turnover penalization perspectives. A recent study of Hautsch and Voigt (2019) established a link between turnover penalization and covariance shrinkage in portfolio allocation. We show that a zero-cost mean-variance portfolio formed using our new covariance matrix estimator is equivalent to a sample mean-variance portfolio achieved through using a turnover penalization, where (i) the turnover is measured by a weighted quadratic transaction cost, and (ii) the weights on transaction costs are determined by a term measuring the difference between the standard and our regularized rolling window based covariance matrix estimates. A large difference between the standard and our regularized rolling sample estimates implies that the real covariance structure is relatively stable so that the regularization removes hugely the unnecessary evolutions generated by the standard rolling window analysis. Regularizing the weighted transaction costs, therefore, ensures that the strength of penalization on portfolio turnover is time-varying, depending on the magnitude of time variation in covariance matrix. To the best of our knowledge, the time varying turnover penalization is novel in the literature, as previous studies (e.g. DeMiguel et al. 2009b, DeMiguel and Olivares-Nadal 2018, Engle et al. 2012) usually adopted a constant tuning parameter to control the level of penalization on portfolio turnovers. Turning to the Bayesian interpretation, we know Bayesian investors often employ useful prior information about quantities of interest. We show that the covariance matrix estimate from our RRW approach can be interpreted as being the *maximum a-posteriori (MAP)* Bayesian estimate associated with Gaussian likelihood for asset returns and a Laplace prior on the time change of the inverse covariance matrix. The resulting portfolios are hence a Bayesian portfolio formed by the investors who have a prior belief on the changes of covariance matrix, and where they construct a portfolio that maximizes the posterior distribution of the change in the covariance matrix.

Lastly, we investigate the economic value of our new portfolio allocation strategy using a range of real data sets. We evaluate the out-of-sample empirical gains associated with investing in mean-variance portfolios using our new covariance matrix estimate, where the portfolio expected returns are measured using the sample estimate which is a constant. Therefore, the portfolio weights mainly focus on our estimation on the structure of covariance matrix. We find that the new strategy outperforms a set of commonly used mean-variance portfolio alternatives, including the standard rolling sample strategy where the covariance matrix is measured by the rolling sample estimate, the $1/N$ naive diversification strategy, the portfolio optimization with covariance matrix forecasts from dynamic models, e.g., Exponential weighted moving average (EWMA) model, and several recently developed new covariance matrix estimator and portfolio optimization techniques, i.e., the linear shrinkage estimator by Ledoit and Wolf (2003), the nonlinear shrinkage estimator by Ledoit and Wolf (2017), and the mean-variance portfolio optimization technique suggested by Ao et al. (2018), in terms of both out-of-sample portfolio risk and turnover control. More importantly, the new strategy earns significantly larger sharpe ratio with considering transaction cost. We further provide additional insights into the gains generated by our new strategy, and attribute the gains to the improved covariance matrix estimation accuracy as well as the better timing ability in significant changes of the covariance matrix.

We make a methodological contribution and an empirical contribution. From a methodological perspective, our RRW approach relates to a burgeoning literature on estimating covariance matrix using the shrinkage technique developed in statistics and machine learning fields.

The idea behind is to shrink an unbiased estimator towards a lower variance (or more stable) target so that the shrunk estimator can strike a balance between mis-specification biases and estimation risk. These existing shrinkage estimators, so far, mainly focus on addressing the static single-period estimation problem, e.g. the "sparse" estimator of Goto and Xu (2015) that shrinks the off-diagonal elements of the (inverse) covariance matrix towards zero and thus reduces the cross-sectional dimension (or the number) of assets in the portfolio selection. The estimators of Ledoit and Wolf (2003) and Chan et al. (1999) shrink the sample covariance matrix towards a more parsimonious target matrix, such as a constant correlation matrix or a covariance matrix with industry factor structure. Our RRW estimator extends the shrinkage idea to a dynamic setting, casting attention on the time change of the covariance matrix. The improved covariance matrix estimator is hence particular conducive to dynamic portfolio selection. In addition, our approach also links with the literature on estimating the dynamic covariance matrix. The aforementioned EWMA model is the most parsimonious dynamic model for covariance matrix, and the recent two studies (Engle et al. 2019, Kastner 2019) have rich this literature on large dynamic covariance matrix estimation. Compared with the dynamic models where the covariance matrix is continuously changing over time, our approach imposes sparsity on the time variation to reduce random variation in the estimates, and this property is particularly favorable for reducing transaction costs in the portfolio allocation. From an empirical perspective, our study links with the work of Fleming et al. (2001), Fleming et al. (2003) and Moreira and Muir (2017) who study portfolio allocation in the context of a short-term mean-variance investor. We go well beyond the results in these papers by focusing on discrete significant changes in covariance matrix, rather than continuous small changes. From portfolio allocation perspective, large change detection in covariance matrix has two advantages over the use of continuous dynamic models. First, the detected change points in covariance matrix enables investors to assess when significant updates to portfolio weights are required, allowing them to avoid excessive trading and the associated large transaction costs. Second, change detection can better reflect structural breaks in covariance matrix, improving investors extreme-market-timing ability.

The rest of the paper is organized as follows. Section 2 provides a small simulation study to address the motivation of our new RRW method. Section 3 introduces the RRW method for change detection in conditional covariance matrix and provides discussion on several empirical implementation issues. Section 4 develops our new volatility timing strategy using the RRW covariance matrix estimator and offers two alternative interpretations on the new strategy. Section 5 describes the data and presents the empirical analysis results. Section 6 examines the robustness of our results. Section 7 concludes.

II. Econometric Methodology

A. The Regularised Rolling Window Approach

We start with the standard rolling window approach, where the covariance matrix at each time t is estimated by minimizing a loss function (or negative log-likelihood) with return observations from the time window $[t-1-h, t-1]$, where h is the window length. Specifically, we consider

$$\hat{\Sigma}_t^{-1} := \arg\min_{\Sigma_t^{-1}} \left[l(\Sigma_t^{-1}) \right] \ .$$

with the loss

$$l(\Sigma_t^{-1}) := -\log\det(\Sigma_t^{-1}) + \mathrm{trace}(\hat{S}_t \Sigma_t^{-1}) \ . \tag{1}$$

where $\hat{S}_t$ and Σ_t^{-1} denote the sample and our estimate of the covariance matrix, respectively. Thus, the time-varying covariance matrix through each time point are estimated recursively by including new observations according to the rolling window. Clearly, with the rolling window analysis, the covariance matrix is estimated independently across estimation windows without any mechanism present to encourage temporal similarity. A potential issue arises that the estimates for two adjacent time points might be largely different due to estimation errors which contradicts with the reality especially when the market is relatively stable. Additionally, the extra variability caused by the independent estimation across rolling windows can potentially mask significant changes in covariance matrix.

To address these empirical features, we propose to minimize a penalized loss function which contains the negative log-likelihood as shown in Equation 1 and an additional penalty term that regularizes the difference between the contemporaneous and the lagged estimates produced by previous estimation window:

$$l(\Sigma_t^{-1}) = -\log\det(\Sigma_t^{-1}) + \mathrm{trace}(\hat{S}_t \Sigma_t^{-1}) + \lambda \|\Sigma_t^{-1} - \hat{\Sigma}_{t-1}^{-1}\|_1 \ , \tag{2}$$

where $\hat{\Sigma}_t^{-1}$ is the lagged estimate from the previous estimation window, and the ℓ_1 norm, defined

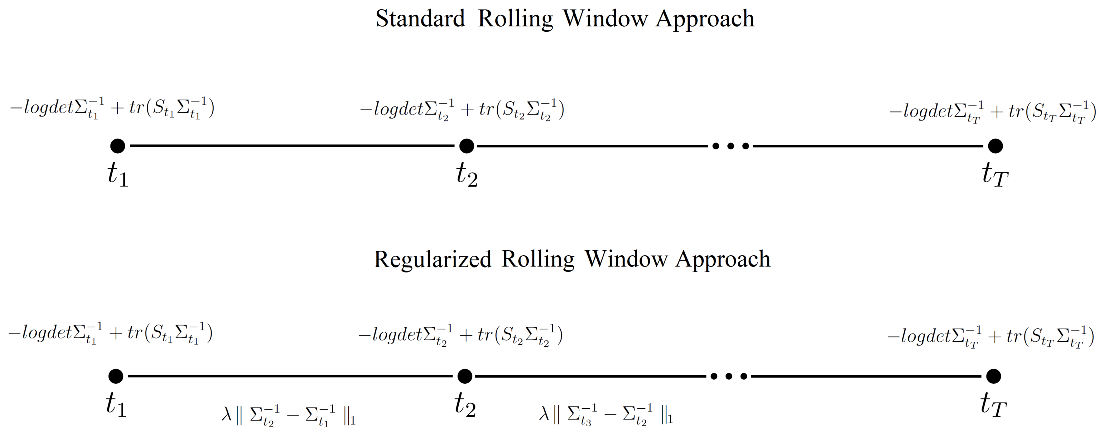

Figure 1: **A graphical comparison between the standard and regularized "rolling window" approach.** The top panel shows the estimation procedure of standard rolling window approach and the bottom panel shows the estimation procedure of the regularized rolling window approach.

as

$$\|\Sigma_t^{-1} - \Sigma_{t-1}^{-1}\|_1 := \sum_{i,j} |\Sigma_{ij,t}^{-1} - \Sigma_{ij,t-1}^{-1}| \ ,$$

measuring the difference between the current and lagged inverse covariance matrix estimates (that is, the sum of the absolute values of edgewise differences between the two estimates.). The λ is a tuning parameter that controls the degree of regularization on the difference, becoming a soft threshold.

Clearly, the new approach nests the standard rolling window estimation as a special case when the regularization parameter (λ) is equal to zero. Figure 1 gives a graphical interpretation about the relation between our RRW approach and the standard rolling window analysis. The regularization term, $\lambda\|\Sigma_t^{-1} - \hat{\Sigma}_{t-1}^{-1}\|_1$, restricts the difference between the current inverse covariance matrix Σ_t^{-1} and the estimate of previous period $\hat{\Sigma}_{t-1}^{-1}$. For off-diagonal entries, if the consecutive difference in the inverse-covariance is below a pre-determined threshold, e.g. related to the size of λ, we set the difference to be zero, when the difference is above the threshold we shrink the difference towards zero. As such, the regularization achieves both sparsity and shrinkage in time variations of the covariance matrix estimation, encouraging a temporally stable (or piecewise constant) estimator which has capacity to reflect big changes in the matrix at element level, but is less sensitive to small ones.

B. Empirical Implementation

We next discuss several empirical issues related to the implementation of the RRW approach. Firstly, we consider the problem of solving Equation 2. Due to the structure of the penalty terms, this loss function is convex, however, because of the ℓ_1 norm it is not continuously differentiable. In this context, traditional optimization algorithms used for maximum likelihood estimation, or generalized method of moments, etc., cannot be adopted without further modifications. Instead, we employ the popular alternating directions method of multipliers (ADMM) algorithm (Boyd et al. 2010) to solve the optimisation problem. The ADMM method is a form of augmented Lagrangian algorithm that is particularly well suited to addressing the highly structured nature of problems such as the one proposed here, for instance Danaher et al. (2013), Gibberd and Nelson (2017) also use this approach for fused estimation of inverse covariance matrices. We provide more detail on the estimation procedure using the ADMM algorithm in Appendix A.

The second major challenge when implementing the RRW approach relates to the selection of the regularisation parameter λ. In this paper, we employ a heuristic parameter-tuning technique inspired by the Akaike information criterion (AIC). We define the AIC for each window $t = 1, \dots T$ as

$$\mathrm{AIC}_t(\lambda) = -\log\det(\hat{\Sigma}_t^{-1}) + \mathrm{trace}(\hat{S}_t\hat{\Sigma}_t^{-1}) + K_t, \tag{3}$$

where K_t is an estimate of the "degrees of freedom". In practice, we use the first $T = 120$ months as a training period (with window width $M = 60$) and use it to search for the value of λ that minimizes the average, $(T-M)^{-1}\sum_{t=M}^{T}\mathrm{AIC}_t(\lambda)$. Following this training period, we adhere to this choice throughout the out-of-sample testing period[1].

III. The Portfolio Problem

In this section, we use the RRW based covariance matrix estimates to develop a new volatility timing strategy. To better understand the motivation of the new strategy, we offer two alternative interpretations from the portfolio choice with turnover penalization and the Bayesian

[1]Goto and Xu (2015) stated that dynamic selection of the regulariser may improve the performance, but they did not adopt this approach due to the intensive computation burden associated with it. We refer to their statement here to justify our choice.

portfolio choice perspectives. We conclude the section by introducing several commonly used mean-variance strategies and performance metrics which we will use to evaluate RRW portfolios.

A. Global Minimum Variance Portfolio

To develop our mean-variance portfolio strategies, we consider a risk-averse investor who allocates wealth across N risky assets plus a riskless asset, e.g. cash. The investor uses conditional mean-variance analysis to make his allocation decisions and re-balances his portfolio monthly. Let R_{t+1}, $\mu = E[R_{t+1}]$, and $\Sigma_t = E_t[(R_{t+1} - \mu)(R_{t+1} - \mu)']$ denote an $N \times 1$ vector of risky asset returns, the expected value of R_{t+1}, and the conditional covariance matrix of R_{t+1}. For each date t, to minimize conditional volatility subject to a given expected return, the investor solves the following quadratic program:

$$
\begin{aligned}
w_{t+1} &:= \min_{w}[w^\top \Sigma_t w] \\
s.t. \quad \mu_p &= w^\top \mu + (1 - w^\top 1) r_{\text{free}},
\end{aligned}
\tag{4}
$$

where w is an $N \times 1$ vector of portfolio weights on the risky assets, r_{free} is the return on the riskless asset, and μ_p is the target expected return. The solution to this optimization problem is given by:

$$
w_{t+1} = \frac{\mu_p \Sigma_t^{-1} \mu}{\mu^\top \Sigma_t^{-1} \mu}\,, \tag{5}
$$

where w delivers the risky asset weights, and the weight on the riskless asset is $1 - w^\top 1$.

The trading strategy implicit in Equation 5 identifies the dynamically re-balanced portfolio that has a global minimum variance (GMV). We follow Ledoit and Wolf (2017) to estimate the GMV portfolio, because it is a clean problem in terms of evaluating the quality of a covariance matrix estimator, since it abstracts from having to estimate the vector of expected returns at the same time.

B. A Transaction Cost Interpretation of RRW

In Hautsch and Voigt (2019), a link between turnover penalization and covariance shrinkage in portfolio allocation is investigated. Specifically, they show that the optimization problem with quadratic transaction costs can be interpreted as a classical mean-variance problem without transaction costs, however, where the covariance matrix is regularized towards the identity matrix. Following this line of interpretation, we show that how optimal MV portfolio achieved using our RRW based covariance matrix estimate without transaction costs links with MV portfolio optimization problem with a time-varying quadratic transaction costs, where the time dependence is determined by the difference between our regularized and the sample covariance matrix.

Proposition 1. *RRW Equivalence to Transaction Cost Penalisation*

In the case of our RRW approach, the resulting covariance matrix is derived from the following optimization problem:

$$\begin{aligned} \hat{\Sigma}_{t,RRW} \quad &:= \quad \arg\max_{\Sigma_t} \left[\log\det(\Sigma_t) - \text{tr}(\Sigma_{t-1}^{-1}, \Sigma_t - \hat{S}_t)\right] \ . \\ s.t. \quad & \|\Sigma_t - \hat{S}_t\|_\infty \leq \lambda \end{aligned} \tag{6}$$

where $\|X\|_\infty := \max_{ij} |X_{ij}|$ *is the dual norm of* $\|X\|_1$*. Thus, the RRW covariance matrix estimator* $\hat{\Sigma}_{t,RRW}$ *can be expressed associated with the sample covariance estimator plus some difference, i.e.* $\hat{S}_t$ *as* $\hat{\Sigma}_{t,RRW} = \hat{S}_t + \lambda\Delta_t$*, where* $\Delta_t := (\hat{\Sigma}_{t,RRW} - \hat{S}_t)/\lambda$*. The proof of the above is given in the Appendix B and follows from basic duality properties of the RRW optimization problem. Plugging the RRW estimator into the standard mean-variance portfolio optimization, the portfolio allocation* w^*_{t+1} *can then be stated as*

$$\begin{aligned} w^*_{t+1} \quad &= \quad \arg\min_{w} \left[w^\top \hat{\Sigma}_{t,RRW} w - w^\top \mu\right] \\ &= \quad \arg\max_{w} \left[w^\top \mu^* - w^\top \hat{S}_t w - \lambda(w - w_t^+)^\top \Delta_t (w - w_t^+)\right] \ , \end{aligned} \tag{7}$$

where

$$\mu_t^\star = \mu - 2\Delta_t w_t^+ \ ,$$

and the weights are normalized such that $w^\top 1 = 1$*.*

Intuitively, our RRW estimator regularizes the sample estimator $\hat{S}_t$ through an additional matrix Δ_t, with λ serving as shrinkage parameter. Note that the regularization effect of RRW estimator exhibits some similarity to the implications of the shrinkage approach proposed by Ledoit and Wolf (2003), but the RRW approach replaces the identity matrix in the Ledoit and Wolf estimator with Δ_t which is time-varying. The resulting zero-cost mean-variance portfolio using our RRW estimator is thus equivalent to a portfolio formed with a time-varying penalty on transaction cost. Figure 2 provides an example of time evolution of $\lambda{\Delta_t}^2$ and highlights recession periods according to the NBER business cycle classification in grey. Clearly, in contrast with the constant identity matrix, the figure shows the Δ_t changes over time and peaks at the recession periods.

To further illustrate the effect of time-varying transaction cost penalty on portfolio optimization, Figure. 3 plots the surface of mean-variance optimization function over time with both our time-varying and traditional constant quadratic transaction cost penalty. The optimal solution for first asset allocation (that is, w_1) is also provided in the lower panel of the figure. The NBER defined recession periods are also highlighted in grey. We observe that the function surface under the time-varying transaction cost penalty (top left panel) exhibits much more stable than the one using constant penalty (top right panel) in most of time without losing capacity to reflect abrupt changes of market conditions during recession periods, such as the observed spikes in the function surface during 2008 crisis period. These findings are further corroborated by the time series plot of w_1 in the lower panel of the figure. Taken together, the time-varying transaction cost penalty resulting from our RRW covariance matrix estimator helps to impose a market-condition-dependent regularization on the transaction cost, assisting investors to achieve better marketing timing and better balance between portfolio risk exposure and turnovers.

C. Bayesian Portfolio Interpretation

Kyung et al. (2010) and Wang (2012) respectively give Bayesian interpretations for regularised regression via the lasso (Tibshirani 1996) and the graphical lasso (Friedman et al. 2008). Since these estimators are closely aligned with the RRW optimisation problem we can follow their line of reasoning to give a Bayesian interpretation for our RRW covariance matrix estimator as well as the resulting portfolios.

Assumption 1. *We start with the assumption that future stock returns are independently normally distributed according the previously estimated covariance, i.e.*

$$r_{t+1} \sim \mathcal{N}(0, \Sigma_t) \, , \tag{8}$$

[2] To generate the examples in Figures. 2 and 3, we construct a two-asset portfolio using two randomly selected assets from the Fama-French 48 Industry portfolio dataset. We consider both the max and min elements of the matrix Δ_t to illustrate its time evolution in Figure 2

and $r_{t+1} \perp r_t$*, i.e. there is no auto-correlation structure in returns.*

Assumption 2. *In order to understand the RRW estimator, we now make a further assumption, and put ourselves in the shoes of an investor who has a prior belief that the inverse covariance may change in a sparse manner over time, i.e. changes will not be at every time-step, but occur rarely. Specifically, we will assume that the temporal variation of the inverse-covariance follows a Laplace (double exponential) distribution:*

$$p(\Theta_{t+1} - \hat{\Theta}_t|\rho) = Z^{-1} \prod_{i<j} \Big\{ f_{\mathrm{DE}}(\Theta_{t;ij} - \hat{\Theta}_{t-1;ij}|\rho) \Big\} \times \prod_{i=1}^{N} \Big\{ f_{\mathrm{Exp}}(\Theta_{t+1;ii} - \hat{\Theta}_{t;ii}|\rho/2) \Big\} \mathbf{1}_{\Theta_t \succeq 0} \,,$$

where $f_{\mathrm{DE}}(x|\rho) = (\rho/2)\exp(-\rho|x|)$ *has the form of the double exponential density,* $f_{\mathrm{Exp}}(x|\rho) = \rho\exp(-\rho x)\mathbf{1}_{x>0}$ *has the form of the exponential density, and* Z *is a normalising constant. The notation* $\mathbf{1}_{\Theta \succeq 0}$ *is used to denote the indicator function, in this case for the space of positive definite matrices for* Θ_t*.*

Proposition 2. *Given that assumptions 1 and 2 hold, and we further assume* $\Sigma_t = \hat{\Sigma}_t$ *in Equation 8, then one can interpret the RRW estimator (minimiser of Eq.2) as being the maximum-a-posteriori (MAP) estimate for the inverse covariance at time* $t+1$*. Specifically, assume that investor believes the temporal variation of inverse covariance matrix has a prior distribution as above, there exists a threshold parameter* ρ *such that our RRW estimator is the mode of the posterior distribution of* Θ_t*.*

Now it is clear that under our framework, choosing the portfolio that maximizes the posterior distribution of the change of (inverse) covariance matrix guarantees that the investor is choosing the portfolio with the highest probability of being the MV portfolio given the investors prior distribution on the temporal change of the (inverse) covariance matrix and the observed asset-return data. In other words, in our setting, the investor chooses the portfolio that maximizes the posterior probability (i.e. the posterior mode) of the change of (inverse) covariance matrix. This interpretation is a bit different from the traditional Bayesian portfolio choice literature in which the investor either chooses the portfolio that maximizes expected utility with respect to the posterior distribution of stock returns (for instance, Jorion 1986) or the portfolio that maximizes the posterior distribution of portfolio weights directly (see DeMiguel et al. 2009a, Tu and Zhou 2010). In our framework the investor has a prior belief on the change of (inverse) covariance matrix rather than on the asset-return distribution or on the portfolio weights. Consequently, while the Bayesian investor in the traditional setting chooses the portfolio that maximizes expected utility with respect to the posterior distribution of asset returns, or chooses the portfolio that maximizes portfolio weights with respect to the posterior distribution of portfolio weights, in our setting the investor chooses the portfolio that maximizes the posterior distribution of the (inverse) covariance matrix changes.

D. Performance Evaluation Metrics

To measure the economic value of our new approach, we compare its performance with several competing GMV strategies using a series of performance evaluation metrics. We firstly compare the GMV strategies constructed using different covariance matrix estimator. There are mainly two broad avenues for estimating a covariance matrix: structure-based and structure-free estimation. We start with other structure-free estimation, including the standard rolling window covariance estimator, and thus, the corresponding GMV portfolio refers to as $\text{GMV}_{\text{sample}}$ (Chan et al. 1999, DeMiguel et al. 2009b,a, Kirby and Ostdiek 2012, Kourtis et al. 2012, Goto and Xu 2015). To further reduce the estimation error of the sample covariance matrix, we use the shrunk version of the sample covariance estimator (Ledoit and Wolf 2003) that shrinks the sample estimate towards an identity matrix, where the corresponding GMV portfolio is denoted as GMV_{Lin}. A more advanced nonlinear covariance shrinkage estimator (Ledoit and Wolf 2017) was also considered for comparison, where the constructed GMV portfolio denoted by $\text{GMV}_{\text{NonLin}}$. In addition, we include a robust covariance estimator, the Minimum Covariance Determinant (Rousseeuw 1984), and the constructed GMV portfolio refers to as $\text{GMV}_{\text{robust}}$. Counterparts for the RRW approach can be generated by using the shrunk sample estimator for $\hat{S}_t$ in Equation 1. The resulting GMV portfolio is defined as GMV_{RRW}. Secondly, we consider GMV portfolio strategies using structure-based covariance estimation. We start with the GMV portfolios using covariance forecasts from dynamic models, e.g. the *Exponentially Weighted Moving Average (EWMA)* model (Zakamulin 2015), denoted by GMV_{EWMA}. This method of estimating the covariance matrix is popularized by the RiskMetrics group, and Zakamulin (2015) also find that the simple EWMA covariance matrix forecast performs comparably with the multivariate GARCH forecast. In this approach, the exponentially weighted covariance matrix is estimated using the following recursive form:

$$\hat{S}_t^{\text{EWMA}} = (1 - \lambda_{\text{EWMA}})\varepsilon_{t-1}\varepsilon_{t-1}^{\top} + \lambda_{\text{EWMA}}\hat{S}_{t-1}^{\text{EWMA}} \quad , \tag{9}$$

where $0 < \lambda_{\text{EWMA}} < 1$ is the decay constant, and ε_{t-1} is the return residual. We follow the recommendations of the RiskMetrics group and estimate the EWMA covariance-matrix using $\lambda_{EWMA} = 0.97$. This comparison is of particular useful for examining whether investors benefit from sparse rather than continuous time variation assumption on covariance matrix[3]. Thirdly, we consider the naïve $1/N$ strategy, denoted by $\text{GMV}_{\text{equal}}$. The naïve $1/N$ strategy demonstrates favorable out-of-sample performance and has been found very hard to beat in practice, especially in the presence of high transaction costs. We use it as a benchmark in order to examine the ability of our new approach in terms of controlling transaction costs. In addition, we consider an alternative mean-variance portfolio optimization techniques suggested by Ao et al. (2018) that relies on a novel unconstrained regression representation of the mean-variance optimization problem combined with high-dimensional sparse-regression methods. The resulting GMV portfolio is denoted as $\text{GMV}_{\text{Sparreg}}$.

Next, we evaluate portfolio out-of-sample performance from several perspective. First, we test out-of-sample performance in terms of risk exposure and turnovers. The portfolio risk exposure is measured by the variance of out-of-sample portfolio returns, and the portfolio turnovers are measured by

$$\text{Turnover} = \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{N} (|\hat{w}_{i,t+1} - \hat{w}_{i,t^+}|), \tag{10}$$

where $\hat{w}_{i,t+1}$ and $\hat{w}_{i,t}$ are the desired portfolio weights in asset i at time t and $t+1$, after rebalancing, and $\hat{w}_{i,t^+}$ is the portfolio weight before rebalancing at $t+1$. The turnover quantity defined can be interpreted as the average percentage of wealth traded in each period. Lastly, we assess whether our new strategy has economic gain. We lastly calculate the Sharpe ratio after deducting transaction costs. We first take transaction costs into account and compute the returns net of transaction costs. Following Robert et al. (2012), we compute the portfolio return net of transaction costs in each period as follows:

$$r_{net,t} = (1 - \sum_{i} c_{t,i} |w_{i,t+1} - w_{i,t^+}|)(1 + r_t) - 1, \tag{11}$$

[3]We also use a covariance estimator proposed by Foster and Nelson (1996) that was used in the paper by Fleming et al. (2001). This estimator is a weighted rolling average of the squares and cross products of past re-turn innovations that nests most ARCH, GARCH, and stochastic volatility models as special cases. We follow Fleming et al. (2001) to determine the weights through minimizing the asymptotic mean squared error (MSE) of the estimator. We do not report the corresponding GMV portfolio here, as the portfolio performs closely to the GMV_{EWMA} and the results are available upon request.

where $c_{t,i}$ is a cost level that measures the transaction cost per dollar traded for trading asset i, and r_t is the portfolio return without transaction cost in period t. For the cost level $c_{t,i}$, we follow Robert et al. (2012) to set $c_{t,i}$ to decrease linearly from 0.6% to 0.1% during the period 1967 to 1990, and set it to be a constant 0.1% from 1991 to 2017. The Sharpe ratio after deducting transaction cost is thus defined as

$$S\hat{R}_t = (\hat{r}_{net,t} - r_{\text{free}})/\hat{\sigma}_{net,t} \quad , \tag{12}$$

where $\sigma_{\hat{net},t}$ is the standard deviation of portfolio return net of transaction cost over the out-of-sample testing period.

IV. Empirical Analysis

A. Data

We employ two data sets: one from Ken French's Web site for portfolio investing analysis, including the returns on 25 value-weighted portfolios of stocks sorted by size and book-to-market (that is, 25FF), the 48 industry value-weighted portfolios (that is, 48Ind), 100 portfolios formed on size and book-to-market (that is, 100SBM), 100 Portfolios formed on size and operating profitability (that is, 100SOP), and 100 Portfolios formed on size and investment (that is, 100SI). For close-to-close returns, we use data from 1967 to 2017 downloaded from Ken French's Web site.

The second data set consists of individuals stocks from the Center for Research in Security Prices (CRSP). We follow Ledoit and Wolf (2017) to consider the following portfolio sizes: $N = 30, 50, 100, 250, 500$. This range covers the majority of the important stock indexes, from the Dow Jones Industrial Average to the S&P 500. We first determine the 500 largest stocks, as measured by their market value on the investment date t that have a complete return history over the most recent $T = 250$ days as well as a complete return "history" over the next 21 days. Out of these 500 stocks, we then select N at random: these N randomly selected stocks constitute the investment universe for the upcoming 21 days. The corresponding portfolios are labelled as $IndN$.

Table 1 summarizes the sample information for each data set, and also provides information regarding the question whether the sparse assumption on time variation of covariance matrix is supported by the real data. Looking at the column 8 and 9 of the table, where we report the optimal value of λ we choose for each data set (as selected via AIC), and the average percentage of non-changed off-diagonal elements between consecutive covariance matrix estimates throughout the whole sample periods. The latter measures the degree of sparsity in time variation of the covariance matrix. We find that the degree of sparsity ranges from 35.78% to 23.55% across these data sets, meaning that a significant fraction of elements in covariance matrix did not significantly change between consecutive periods. Hence, the sparse time variation assumption appears reasonable in practice.

B. Covariance forecasting performance

In this section, we evaluate different covariance matrix forecasting methods by directly evaluating their forecast accuracy. In particular, we compare the performances of different methods by performing out-of-sample forecasts of the monthly covariance matrix. That is, the covariance matrix for month t is forecast based on information available at the end of month $t-1$. More specifically, our forecasts are based on the rolling window estimation scheme by using a look back period of 60 months. We follow Zakamulin (2015) to evaluate the covariance matrix forecast accuracy using the mean squared forecast error (MSFE) as defined below:

$$\begin{aligned} SFE &= \sum_{i=1}^{n}\sum_{j=1}^{i}(\sigma_{ij,t}-\hat{\sigma}_{ij,t})^2, \\ MSFE &= 1/t\sum_{t=1}^{M} SFE_t \end{aligned} \tag{13}$$

where $\sigma_{ij,m}$ and $\hat{\sigma}_{ij,t}$ denote the monthly realized and a forecast of the covariance matrix for month t. The SFE denotes the squared forecasting error and the MSFE denotes the mean squared forecasting errors. M is the number of months in the out-of-sample evaluation period.

Table 2 reports the MSFE produced by different covariance-matrix forecasting methods for each data set separately. The results demonstrate that i). the rolling window based sample covariance method and the rolling window based sample covariance with shrinkage method provide almost identical forecast accuracy. This finding is consistent with Disatnik and Benninga (2007) and Zakamulin (2015) that the linear shrinkage does not produce a significantly better forecast of the covariance matrix; ii). the nonlinear shrinkage method largely reduced the forecasting errors, followed by the robust covariance estimator; and iii). the dynamic model $EWMA$ performs the best in the covariance matrix forecast, and our RRW method performs similarly to the $EWMA$ in most cases.

C. Portfolio performance

We now turn to evaluating the out-of-sample portfolio performance. We start with portfolio investing by focusing on the five Fama-French portfolio data sets: 25FF, 48Ind, 100SBM, 100SOP, and 100SI. A move towards analysis on the individual stock level is given in the next section. In each month t, we construct the GMV portfolios using stock returns from past $M = 60$ months (5 years)[4]. We hold such portfolios for 1 month and calculate the portfolio returns for out-of-sample month $t + 1$. We continue this process by adding the return for the next period in the data set and dropping the earliest return from the estimation window.

Our primary interest is in the ability of the proposed GMV strategy in reducing the out-of-sample portfolio risk. We first construct the time series of out-of-sample returns for our GMV_{RRW} and other seven competing GMV portfolios. We then compare out-of-sample return variance to see whether GMV_{RRW} achieves out-of-sample risk reduction. Last, we test the significance of any difference between GMV_{RRW} and other alternatives using the stationary bootstrap of Politis and Romano (1994). Panel A of Table 3 reports the monthly out-of-sample risk for each GMV portfolio strategy, and Panel B gives the difference test results. From the table, we observe: i). compared to the standard sample GMV portfolio, $\text{GMV}_{\text{sample}}$, our portfolio ($\text{GMV}_{\text{RRW}}$) significantly reduces the portfolio out-of-sample risk, suggesting that the regularization increases the ability of the rolling sample estimates in change detection of the covariance matrix, and thereby improves the portfolio performance by better controlling portfolio risk exposure. For example, the portfolio risk decreases from 27.629($\%^2$) to 13.707($\%^2$) for 25FF, and from 26.5335($\%^2$) to 13.157($\%^2$) for 48IND; ii). compared with the portfolios with other structure-free covariance estimators, such as linear and nonlinear covariance shrinkage estimators (i.e., GMV_{Lin} and $\text{GMV}_{\text{Nonlin}}$), the robust covariance estimators, (i.e., $\text{GMV}_{\text{robust}}$), the RRW covariance estimator still reduces the portfolio risk in all the cases, further supporting the benefit of exploiting the significant change in covariance matrix for portfolio construction. iii). compared with the naïve $1/N$ strategy ($\text{GMV}_{\text{equal}}$), our portfolio ($\text{GMV}_{\text{RRW}}$) still achieves lower risk, implying that tracking the variance changes is favorable to portfolio out-of-sample risk reduction; iv). compared with the GMV portfolio formed using covariance matrix forecasts

[4]The choice of the rolling estimation window size, $M = 60$, follows the standard practice in the literature. To save space, we report the results only for $M = 60$. We have also conducted an analysis using a longer estimation window of $M = 120$ and found the results are generally robust. The results are available upon request.

from the EWMA model, GMV_{EWMA}, our portfolio once again offers smaller risk, e.g. 13.707 v.s. 14.422 in 25FF, and 13.157 v.s. 13.8565 in 48IND. This supports the assumption that sparse rather than continuous changes in the covariance matrix does not weaken, but rather strengthens the portfolio out-of-sample risk reduction; v). compared with $\text{GMV}_{\text{Sparreg}}$, the ($\text{GMV}_{\text{RRW}}$) achieved smaller risk when the portfolio dimension is relatively small and performs similarly for large portfolios (i.e., the portfolio dimension is greater than 100).

We next turn to investigating the ability of our RRW strategy in controlling portfolio turnover. We calculate the monthly portfolio turnovers as stated in Equation 10 for the eight portfolios and report these results in Panel B of Table 3. It is not surprising we observe that the equally weighted $1/N$ portfolio provides the lowest turnover for all the data sets. The naïve diversification requires only a very small amount of trades to maintain the equal weights. On the contrary, the sample portfolio $\text{GMV}_{\text{sample}}$ and the other GMV portfolio strategy, i.e., GMV_{EWMA}, $\text{GMV}_{\text{robust}}$, GMV_{Lin}, always suffer large turnovers, because it requires active trading to adapt with the changing covariance matrix in order to achieve the best risk diversification. Our portfolio (GMV_{RRW}) significantly reduces the portfolio turnovers, even compared with the portfolios with more advanced covariance estimator and optimization technique ($\text{GMV}_{\text{Nonlin}}$ and $\text{GMV}_{\text{Sparreg}}$). The favorable performance in portfolio turnover control verifies that the RRW approach offers a more stable estimate for the covariance matrix that significantly reduces the portfolio turnovers and thus the associated transaction costs.

Finally, we assess the sharpe ratio after deducting transaction costs. Following standard practice of Robert et al. (2012), we set the transaction cost per dollar traded to decrease linear from 0.6% to 0.1% during the period 1967 to 1990, and set it to be a constant 0.1% from 1991 to 2017. Panel C of Table 3 reports the results. We observe that the portfolios GMV_{RRW} still outperform all the alternatives by retaining the highest Sharpe ratios across all the data sets. The Sharpe ratio (after deducting the transaction costs) of the portfolio GMV_{RRW} ranges from 0.212 to 0.289, followed by the portfolio $\text{GMV}_{\text{Sparreg}}$.

We also conduct a portfolio analysis operating based on individual assets. We randomly select 500 samples from CRSP stocks and randomly select N stocks from them to form portfolios, where $N = 30, 50, 100, 250, 500$. Table 3 shows that the portfolio (GMV_{RRW} achieves lower out-of-sample portfolio risk than other alternatives in all the cases, which offers compelling evidence for the ability of GMV_{RRW} to achieve significant reduction in out-of-sample portfolio risk. We turn to portfolio turnovers and observe that GMV_{RRW} achieves the lowest turnovers in all the cases. In terms of the Sharpe ratio after deducting transaction costs, we find that the GMV_{RRW} provides the highest Sharpe ratio after transaction cost in all the data sets. These results once again suggest that the RRW estimator provides larger economic gains for a panel of samples that have a relatively stable covariance matrix structure and calls for a higher level of temporal stability regularization.

Since some fund managers face a no-short-sales constraint, we now impose a lower bound of zero on all portfolio weights. Table 4 reports the results. The ranking of the portfolios across the three performance evaluation metrics are pretty much similar, except that the $\text{GMV}_{\text{sample}}$ performs much better. These findings are consistent with Jagannathan and Ma (2003) who demonstrate theoretically that imposing a no-short-sales constraint corresponds to an implicit shrinkage of the sample covariance matrix in the context of estimating the global minimum-variance portfolio.

D. Decomposing the Performance Gain: Estimation Accuracy and the Ability to Time Significant Changes

While the above findings support the existence of additional economic gains from using our RRW estimator in GMV strategies, in this section, we attempt to answer the question: where are the gains generated from? We explain the advantage of our RRW estimator from two perspectives: *estimation accuracy* and the ability o *timing significant changes.*

Firstly, we attribute the better portfolio performance to the improved covariance matrix forecasts. We examine forecasting accuracy of our RRW approach based covariance matrix estimate, compared with the shrunk rolling sample estimate and the forecasts from EWMA model using the following log predictive likelihood:

$$l_t(\Sigma^{-1}) = \ln(\det(\hat{\Sigma}_{t-1}^{-1})) - \tilde{r}_t^\top \hat{\Sigma}_{t-1}^{-1} \tilde{r}_t, \tag{14}$$

where T is the total number of out-of-sample testing periods. $\widetilde{r}_t$ denotes the demeaned return vector at time t, and Σ_{t-1} is the covariance matrix estimate for time t but made at time $t-1$. We average $l_t(\Sigma^{-1})$ across the whole out-of-sample testing period, that is, $L(\Sigma^{-1}) = (1/T)\sum_{t=1}^{T} l_t(\Sigma^{-1})$. We calculate out-of-sample log predictive likelihood for each covariance matrix estimator, and then test the significance of the difference between ones from our estimator and the other alternatives. Table 2 reports the testing results. Column 2 of the table shows that our RRW estimate has a significantly higher predictive likelihood than does the shrunk rolling sample estimates in all the data sets, proving that imposing temporal similarity regularization in rolling window approach reduces the covariance matrix predictive errors. Column 3 shows that the RRW estimate outperforms the EWMA based estimate, suggesting that allowing for piecewise constancy is conducive to increase the predictive accuracy for out-of-sample covariance matrix. Overall, these results confirm that the RRW approach improves forecasting accuracy of the covariance matrix, leading to the better out-of-sample portfolio performance.

Next, we demonstrate the advantage of our RRW approach in timing significant changes of the covariance matrix. The top panel of Figure 4 plots the temporal variation of our RRW (the solid line) and the shrunk sample (the dotted line) covariance matrix estimators, measured by the trace of the estimated covariance structures ($\sum_{ii} \hat{\Sigma}_{ii,t}$). We highlight in grey the recession periods according to the NBER business cycle classification. Clearly, the RRW estimator is more stable than the sample counterpart during the calm period, such as the period between 2002-2007, but without loosing capacity to reflect large changes during the recession period, such as the GFC period from 2008-2010.

To further illustrate this point, we split the whole sample into "good" and "bad" economic periods according to the NBER business cycle classification. Then, we compare out-of-sample portfolio turnover and risk exposure of our RRW strategy with other alternatives. The results are reported in Table 5. If our strategy has better capacity in timing significant changes, we expect that it allows more aggressive updates in portfolio weights and thus larger increase in portfolio turnovers compared with the $1/N$ strategy during "bad" periods. On the contrary, during "good" periods, it should have more conservative response to the change of covariance matrix, leading much less portfolio turnovers compared with the shrunk rolling sample strategy. Looking at the portfolio turnovers reported in the table, we observe that during "good" periods, the portfolio turnover of GMV_{RRW} is quite close to that of the $1/N$ portfolio, but the both are much less than that of $\text{GMV}_{\text{sample}}$. For example, for the data of $25FF$, the portfolio turnover of GMV_{RRW} is only 0.019 that is close to 0.017 of the $1/N$ portfolio, but much less than 0.047 of the $\text{GMV}_{\text{sample}}$. On the other hand, during the bad periods, the portfolio turnover of GMV_{RRW} hugely increases to 1.033, but the turnover of the $1/N$ portfolio only slightly increases to 0.018. Comparatively, the sample portfolio $\text{GMV}_{\text{sample}}$ has a constantly higher portfolio turnovers, which is also increasing during the bad period. The bottom panel of Figure 4 plots the portfolio turnovers using our RRW (the solid line) and the sample (the dotted line) strategies against the business cycles. Our RRW strategy always offers less portfolio turnovers, but reasonably increases turnovers when the market is in distress. We also test the significance of different improvement from the second best estimator to our our RRW estimator (denoted as $\mathrm{MV}_{\mathrm{RRW}}$) between the good and bad periods. We find that the improvement achieved by the RRW estimator appears more pronounced during the bad periods.

To conclude the section we examine the portfolio risk and whether the better ability of $\mathrm{GMV}_{\mathrm{RRW}}$ to time changes results in a better control on portfolio risk exposure. We notice that $\mathrm{GMV}_{\mathrm{RRW}}$ always achieves the lowest risk, and the $1/N$ strategy outperforms the sample strategy in "good" periods by achieving lower risk, and vice versa in "bad" periods. In summary, we observe that our RRW covariance estimator has a great ability to highlight significant changes, helping the resulting portfolio to strike a better balance between portfolio turnovers and risk exposure.

V. Robustness Checks

A. The Estimation Window Length

To some extent, one may argue that the RRW estimator uses longer samples (from the previous estimation window) than other competing methods used in empirical evaluations of the out-of-sample portfolio performance. This makes it difficult to evaluate the performance gain from the RRW approach over other methods. Does the performance gain come from the particular regularization, or does it comes from the use of a longer sample? We examine a simple way to address this question, that is, we adopt the standard rolling window approach in other competing methods with a fixed longer estimation window, e.g. $M = 120$ months. Table 6 reports the results and two patterns are observed: i) all the competing methods perform better with longer estimation window. This is not surprising as the longer sample of observations provides more historical information which helps achieving more robust estimation; ii) the RRW estimator still outperforms all the other competitors, confirming that through exploiting similarity between two consecutive estimation windows using the temporal similarity regularization, the temporally stable estimator largely reduces "spurious" time variations caused by estimation errors based on the standard rolling window approach. The resulting portfolio, therefore, exhibits more stable out-of-sample performance.

B. Weekly Return Data

We use monthly stock returns in the benchmark analysis; here, we evaluate the performance of the different portfolios regarding weekly return data for the five data sets to see whether the results are robust to the return data frequency. We report the portfolio performances in Table 7. We find that our results are generally robust to the use of weekly data. For instance, we find that even with weekly data, our regularized portfolios with $\gamma = 5$ generally outperform the alternatives. When we compare the performance of the portfolios for monthly and weekly return data, we find that the portfolios perform slightly better with monthly than with weekly data. We believe the reason is that the benefit of more frequently adjusting the hedge trades is offset by the higher transaction costs.

VI. Conclusion

In this paper, we propose a regularized rolling window approach to estimate the time-varying covariance matrix, and construct novel minimum-variance portfolios. Through imposing a temporal variation constraint on the standard rolling window based sample estimates this new method is both simple and interpretable, whilst also yielding superior out-of-sample forecasts for the covariance matrix and being capable of detecting significant changes in covariance matrix. We demonstrate that in the presence of both structural changes and transaction cost, the resulting minimum-variance portfolio achieves simultaneously low risk exposure and turnover, earning significant economic gains compared to a set of commonly used alternatives. These results support our initial motivation of this study: in the presence of transaction cost, investors can benefit from significant change detection in volatility.

References

Ao, M., Yingying, L., Zheng, X., 09 2018. Approaching mean-variance efficiency for large portfolios. The Review of Financial Studies 32 (7), 2890–2919.

Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., 2010. Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends. Mach. Learn. 3, 1–122.

Chan, L., Karceski, J., Lakonishok, J., 1999. On portfolio optimization: Forecasting covariances and choosing the risk model. Review of Financial Studies 12, 937–974.

Danaher, P., Wang, P., Witten, D. M., 2013. The joint graphical lasso for inverse covariance estimation across multiple classes. Journal of the Royal Statistical Society: Series B (Statistical Methodology).

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

DeMiguel, V., Garlappi, L., Uppal, R., 2009b. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? Review of Financial Studies 22, 1915–1953.

DeMiguel, V., Olivares-Nadal, A. V., 2018. Technical note-a robust perspective on transaction costs in portfolio optimization. Operations Research, forthcoming.

Disatnik, D. J., Benninga, S., 2007. Shrinking the covariance matrix. The Journal of Portfolio Management 33 (4), 55–63.

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

Engle, R. F., Robert, F., Jekrey, R., 2012. Measuring and modeling execution cost and risk. Journal of Portfolio Management 38 (2), 14–28.

Fleming, J., Kirby, C., Ostdiek, B., 2001. The economic value of volatility timing. Journal of Finance 56 (1), 329–352.

Fleming, J., Kirby, C., Ostdiek, B., 2003. The economic value of volatility timing using realized volatility. Journal of Financial Economics 67 (3), 473 – 509.

Foster, D., Nelson, D. B., 1996. Continuous record asymptotics for rolling sample variance estimators. Econometrica 64 (1), 139–74.

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

Gârlranu, N., Pedersen, L., 2013. Dynamic trading with predictable returns and transaction costs. The Journal of Finance LXVIII (6), 2309–2340.

Gârlranu, N., Pedersen, L., 2016. Dynamic portfolio choice with frictions. Journal of Economic Theory 165, 487–516.

Gibberd, A. J., Nelson, J. D. B., 2017. Regularized Estimation of Piecewise Constant Gaussian Graphical Models: The Group-Fused Graphical Lasso. Journal of Computational and Graphical Statistics 26 (3), 623–634.

Goto, S., Xu, Y., 2015. Improving mean variance optimization through sparse hedging restrictions. Journal of Financial and Quantitative Analysis 50, 1415–1441.

Hautsch, N., Voigt, S., 2019. Large-scale portfolio allocation under transaction costs and model uncertainty. Journal of Econometrics.

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

Jorion, P., 1986. Bayes-stein estimation for portfolio analysis. Journal of Financial and Quantative Analysis 21, 279–292.

Kastner, G., 2019. Sparse bayesian time-varying covariance estimation in many dimensions. Journal of Econometrics 210 (1), 98–115.

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

Kourtis, A., Dotsis, G., Markellos, R., 2012. Parameter uncertainty in portfolio selection: Shrinking the inverse covariance matrix. The Journal of Banking and Finance 36, 2522–2613.

Kyung, M., J. Gill, M. G., Casalla, G., 2010. Penalized regression, standard error, and bayesian lasso. Bayesian Analysis 5, 369–412.

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

Ledoit, O., Wolf, M., 06 2017. Nonlinear shrinkage of the covariance matrix for portfolio selection: Markowitz meets goldilocks. The Review of Financial Studies 30 (12), 4349–4388.

Moreira, A., Muir, T., 2017. Volatility-managed portfolios. The Journal of Finance 72 (4), 1611–1644.

Politis, D., Romano, J., 1994. The stationary bootstrap. Journal of the American Statistical Association, 1303–1313.

Robert, E., Robert, F., Jeffrey, R., 2012. Measuring and modeling execution cost and risk. The Journal of Portfolio Management 38 (2), 14–28.

Rousseeuw, P. J., 1984. Least median of squares regression. Journal of the American Statistical Association 79 (388), 871–880.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological) 58 (1), 267–288.

Tibshirani, R., Saunders, M., Rosset, S., Zhu, J., Knight, K., 2005. Sparsity and smoothness via the fused lasso. Journal of Royal Statistics Society B 67 (1), 91–108.

Tu, J., Zhou, G., 2010. Incorporating economic objectives into bayesian priors: Portfolio choice under parameter uncertainty. Journal of Financial and Quantitative Analysis 45 (4), 959–986.

Wang, H., dec 2012. Bayesian Graphical Lasso Models and Efficient Posterior Computation. Bayesian Analysis 7 (4), 867–886.

Zakamulin, V., 2015. A test of covariance-matrix forecasting methods. The Journal of Portfolio Management 41 (3), 97–108.

Table 1: **Data Description.** This table lists the data sets used in our empirical analysis. Column 2 provides the abbreviation used to refer to the testing portfolios. Column 3 gives more detailed descriptions about the data sets. Column 4 reports the number of stocks in each data set, and Column 5 reports the length of sample period. Column 6 and 7 present the training and testing period in out-of-sample analysis. Column 8 and 9 give the optimal value we used for the regularization parameter (λ) in RRW approach and the average percentage of non-changed off-diagonal elements between consecutive covariance matrix estimates throughout the whole sample period.

Data Set	Abbreviation	Description	N	T	Training Period	Testing Period	λ	Percentage of unchanged matrix off-diagonal elements
1	25FF	25 size and BM portfolios	25	553	Jan. 1967-Dec. 1972	Jan.1973-Dec.2017	0.5	23.55%
2	48IND	48 industry portfolios	48	553	Jan. 1967-Dec. 1972	Jan.1973-Dec.2017	0.5	30.23%
3	100SBM	100 size and book-to-market portfolios	100	553	Jan. 1967-Dec. 1972	Jan.1973-Dec.2017	0.5	31.75%
4	100SOP	100 size and operating profitability portfolios	100	553	Jan. 1967-Dec. 1972	Jan.1973-Dec.2017	0.5	32.33%
5	100SI	100 size and investment portfolios	100	553	Jan. 1967-Dec. 1972	Jan.1973-Dec.2017	0.5	31.86%
6	Ind30	30 stocks from CRSP	30	553	Jan. 1992-Dec. 1997	Jan.1998-Dec.2017	0.4	26.74%
7	Ind50	50 stocks from CRSP	50	553	Jan. 1992-Dec. 1997	Jan.1998-Dec.2017	0.4	25.28%
8	Ind100	100 stocks from CRSP	100	553	Jan. 1992-Dec. 1997	Jan.1998-Dec.2017	0.5	30.27%
9	Ind25	250 stocks from CRSP	250	553	Jan. 1992-Dec. 1997	Jan.1998-Dec.2017	0.6	35.78%
10	Ind500	500 stocks from CRSP	500	553	Jan. 1992-Dec. 1997	Jan.1998-Dec.2017	0.5	32.29%

Table 2: **Out-of-sample covariance matrix prediction.** This table reports the MSFE produced by different covariance-matrix forecasting methods. COV_{sample} denotes the rolling window based sample estimation of the covariance matrix; COV_{EWMA} denotes the EWMA covariance; COV_{Lin} denotes the linear shrinkage covariance estimator by Ledoit and Wolf (2003); COV_{Nonlin} denotes the nonlinear shrinkage covariance estimator by Ledoit and Wolf (2017); COV_{robust} denotes the robust covariance estimator, the Minimum Covariance Determinant; and COV_{RRW} denotes our RRW covariance estimator. In all methods, the length of the rolling estimation window is 120 months.

Data Set	COV_{sample}	COV_{Lin}	COV_{Nonlin}	COV_{EWMA}	COV_{robust}	COV_{RRW}
25FF	0.0085	0.0083	0.0061	0.0049	0.0078	0.0059
48IND	0.0092	0.0091	0.0065	0.0058	0.0086	0.0063
100SBM	0.0105	0.0099	0.0093	0.0083	0.0098	0.0088
100SOP	0.0101	0.0099	0.0095	0.0087	0.0097	0.0089
100SI	0.0106	0.0103	0.0099	0.0091	0.0095	0.0092
Ind30	0.0073	0.0069	0.0066	0.0055	0.0069	0.0056
Ind50	0.0081	0.0079	0.0070	0.0059	0.0077	0.0062
Ind100	0.0098	0.0097	0.0083	0.0073	0.0089	0.0076
Ind250	0.0107	0.0105	0.0098	0.0094	0.0098	0.0095
Ind500	0.0112	0.0110	0.0105	0.0099	0.0106	0.0101

Table 3: **Out-of-Sample Portfolio Performance (without short-sales restrictions).** For each data set, this table reports the out-of-sample monthly risk (return variances) (Panel A), overall turnover (Panel B), and monthly sharpe ratio (Panel C) for the following eight portfolios: portfolio using our RRW estimator (denoted as GMV_{RRW}); portfolio using rolling window based shrunk sample estimate of the covariance matrix (denoted as GMV_{sample}); equally weighted portfolio (denoted as GMV_{equal}); portfolio using EWMA based estimates of covariance matrix (denoted as MV_{EWMA}); portfolio using the linear shrinkage covariance estimator by Ledoit and Wolf (2003) (denoted as GMV_{Lin}); portfolio using the nonlinear shrinkage covariance estimator by Ledoit and Wolf (2017) (denoted as GMV_{Nonlin}); portfolio using the robust covariance estimator, the Minimum Covariance Determinant (denoted as GMV_{robust}); portfolio using the mean-variance portfolio optimization technique (sparse regression) suggested by Ao, Yingying, and Zheng (2019) (denoted as $GMV_{Sparreg}$). We test the significance of any difference between GMV_{RRW} and other alternatives using the stationary bootstrap of Politis and Romano (1994). We only report the significance between GMV_{RRW} and the closest alternative due to the restricted space. *, **, and *** indicate significant differences at the 10%, 5% and 1% levels, respectively.

Data Set	GMV_{sample}	GMV_{equal}	GMV_{EWMA}	GMV_{Lin}	GMV_{Nonlin}	GMV_{robust}	$GMV_{Sparreg}$	GMV_{RRW}
Panel A: Portfolio Risk (monthly return variance (%^{2}))								
25FF	27.6290	25.5890	14.4220	16.7870	13.7765	14.2210	13.7870	13.7070***
48IND	26.5335	24.5754	13.8565	16.1266	13.2125	13.6635	13.2469	13.1570***
100SBM	30.7495	28.4814	16.0654	18.6949	15.4325	15.8420	15.3594	15.2400**
100SOP	31.5997	29.2671	16.4984	19.2026	15.7219	16.2686	15.7723	15.6730*
100SI	32.1107	29.7457	16.7994	19.5412	15.9980	16.5664	16.0632	15.8910***
Ind30	26.2023	24.2682	13.6804	15.9227	13.2125	13.4898	13.0783	12.9960**
Ind50	28.6714	26.5576	14.9865	17.4371	14.2987	14.7782	14.3285	14.2030**
Ind100	33.8437	31.3454	17.6700	20.5663	17.2128	17.4239	16.8924	16.7860
Ind250	37.2119	34.4657	19.4330	22.6167	18.6960	19.1624	18.5781	18.4520*
Ind500	39.4282	36.5185	20.5904	23.9637	19.7674	20.3037	19.6847	19.5510*
Panel B: Portfolio Turnover								
25FF	0.291	0.017	0.320	0.057	0.049	0.062	0.028	0.020***
48IND	0.346	0.020	0.461	0.063	0.053	0.078	0.034	0.025***
100SBM	0.538	0.031	0.592	0.106	0.087	0.115	0.041	0.037***
100SOP	0.655	0.038	0.721	0.128	0.086	0.139	0.058	0.045***
100SI	0.467	0.027	0.512	0.091	0.077	0.099	0.039	0.032***
Ind30	0.335	0.020	0.368	0.066	0.047	0.055	0.031	0.023***
Ind50	0.597	0.035	0.656	0.117	0.072	0.098	0.055	0.041***
Ind100	0.844	0.049	0.928	0.165	0.097	0.139	0.064	0.058***
Ind250	0.931	0.055	1.024	0.183	0.104	0.154	0.077	0.064***
Ind500	1.033	0.060	1.136	0.202	0.162	0.171	0.085	0.071***
Panel C: Portfolio Sharpe Ratio with transaction cost (monthly)								
25FF	0.197	0.255	0.223	0.209	0.276	0.267	0.281	0.289***
48IND	0.166	0.228	0.189	0.176	0.228	0.225	0.235	0.243***
100SBM	0.150	0.209	0.170	0.159	0.0210	0.204	0.215	0.220**
100SOP	0.161	0.211	0.182	0.171	0.225	0.218	0.229	0.236***
100SI	0.164	0.168	0.186	0.174	0.219	0.222	0.234	0.239
Ind30	0.213	0.305	0.241	0.226	0.297	0.288	0.311	0.312
Ind50	0.189	0.245	0.215	0.201	0.261	0.257	0.268	0.277**
Ind100	0.195	0.266	0.221	0.207	0.269	0.264	0.275	0.286***
Ind250	0.167	0.239	0.189	0.177	0.231	0.226	0.239	0.244*
Ind500	0.145	0.198	0.164	0.154	0.201	0.196	0.206	0.212**

Table 4: **Out-of-Sample Portfolio Performance (with short-sales restrictions).** For each data set, this table reports the out-of-sample monthly risk (return variances) (Panel A), overall turnover (Panel B), and monthly sharpe ratio (Panel C) for the following eight portfolios: portfolio using our RRW estimator (denoted as MV_{RRW}); portfolio using rolling window based shrunk sample estimate of the covariance matrix (denoted as MV_{sample}); equally weighted portfolio (denoted as MV_{equal}); portfolio using EWMA based estimates of covariance matrix (denoted as MV_{EWMA}); portfolio using the linear shrinkage covariance estimator by Ledoit and Wolf (2003) (denoted as MV_{Lin}); portfolio using the nonlinear shrinkage covariance estimator by Ledoit and Wolf (2017) (denoted as MV_{Nonlin}); portfolio using the robust covariance estimator, the Minimum Covariance Determinant (denoted as MV_{robust}); portfolio using the mean-variance portfolio optimization technique (sparse regression) suggested by Ao, Yingying, and Zheng (2019) (denoted as $MV_{Sparreg}$). We test the significance of any difference between GMV_{RRW} and other alternatives using the stationary bootstrap of Politis and Romano (1994). We only report the significance between GMV_{RRW} and the closest alternative due to the restricted space. *, **, and *** indicate significant differences at the 10%, 5% and 1% levels, respectively.

Data Set	MV_{sample}	MV_{equal}	MV_{EWMA}	MV_{Lin}	MV_{Nonlin}	MV_{robust}	$MV_{Sparreg}$	MV_{RRW}
				Panel A: Portfolio Risk (monthly return variance ($\%^2$))				
25FF	14.2198	25.5890	14.4222	16.7871	14.1908	14.2208	13.7871	13.7072***
48IND	13.6554	24.5754	13.8565	16.1266	13.5321	13.6635	13.2469	13.1570**
100SBM	15.7986	28.4815	16.0653	18.6957	15.6565	15.8432	15.3601	15.2421**
100SOP	16.2541	29.2673	16.4986	19.2014	15.9892	16.2697	15.7734	15.6731**
100SI	16.5543	29.7457	16.7994	19.5412	16.2341	16.5664	16.0632	15.8910***
Ind30	13.4767	24.2692	13.6804	15.9227	13.3226	13.4898	13.0783	12.9960**
Ind50	14.7656	26.5589	14.9875	17.4373	14.5673	14.7794	14.3299	14.2030*
Ind100	17.4121	31.3461	17.6700	20.5663	17.1287	17.4239	16.8924	16.7876***
Ind250	19.1545	34.4670	19.4330	22.6167	18.8789	19.1632	18.5781	18.4520**
Ind500	20.2998	36.5165	20.5904	23.9637	19.8786	20.3067	19.6854	19.5514**
				Panel B: Portfolio Turnover				
25FF	0.069	0.017	0.334	0.068	0.056	0.077	0.034	0.029**
48IND	0.387	0.020	0.478	0.075	0.067	0.096	0.058	0.042**
100SBM	0.543	0.031	0.604	0.114	0.096	0.134	0.063	0.051***
100SOP	0.678	0.038	0.754	0.153	0.121	0.156	0.074	0.062**
100SI	0.473	0.027	0.533	0.104	0.098	0.113	0.066	0.058*
Ind30	0.358	0.020	0.387	0.078	0.069	0.087	0.031	0.026*
Ind50	0.613	0.035	0.672	0.134	0.092	0.103	0.075	0.062**
Ind100	0.885	0.049	0.945	0.184	0.123	0.154	0.086	0.071***
Ind250	0.989	0.055	1.067	0.206	0.134	0.167	0.092	0.079***
Ind500	1.115	0.060	1.154	0.215	0.167	0.184	0.099	0.084**
				Panel C: Portfolio Sharpe Ratio with transaction cost (monthly)				
25FF	0.156	0.255	0.215	0.202	0.268	0.254	0.273	0.286**
48IND	0.134	0.228	0.176	0.164	0.220	0.213	0.228	0.241***
100SBM	0.142	0.209	0.163	0.146	0.212	0.211	0.214	0.219**
100SOP	0.152	0.201	0.176	0.168	0.209	0.206	0.216	0.233***
100SI	0.160	0.168	0.179	0.170	0.220	0.213	0.225	0.236**
Ind30	0.209	0.305	0.235	0.215	0.289	0.276	0.306	0.310
Ind50	0.175	0.245	0.207	0.199	0.221	0.243	0.245	0.257**
Ind100	0.186	0.266	0.213	0.193	0.249	0.252	0.261	0.274**
Ind250	0.174	0.239	0.177	0.164	0.210	0.218	0.223	0.239**
Ind500	0.139	0.198	0.152	0.141	0.185	0.184	0.198	0.205***

Table 5: **Portfolio performance during "good" and "bad" periods.** This table reports the out-of-sample overall turnover and monthly sharpe ratio for the following eight portfolios: portfolio using our RRW estimator (denoted as MV_{RRW}); portfolio using rolling window based shrunk sample estimate of the covariance matrix (denoted as MV_{sample}); equally weighted portfolio (denoted as MV_{equal}); portfolio using EWMA based estimates of covariance matrix (denoted as MV_{EWMA}); portfolio using the linear shrinkage covariance estimator by Ledoit and Wolf (2003) (denoted as MV_{Lin}); portfolio using the nonlinear shrinkage covariance estimator by Ledoit and Wolf (2017) (denoted as MV_{Nonlin}); portfolio using the robust covariance estimator, the Minimum Covariance Determinant (denoted as MV_{robust}); portfolio using the mean-variance portfolio optimization technique (sparse regression) suggested by Ao, Yingying, and Zheng (2019) (denoted as $MV_{Sparreg}$). The "good" (Panel A) and "bad" (Panel B) periods are based on NBER business cycle classifications. We test the significance of different improvement from the second best estimator to our our RRW estimator (denoted as MV_{RRW}) between the good and bad periods. *, **, and *** indicate significant differences at the 10%, 5% and 1% levels, respectively.

Date set	MV_{sample}	MV_{equal}	MV_{EWMA}	MV_{Lin}	MV_{Nonlin}	MV_{robust}	$MV_{Sparreg}$	MV_{RRW}
				Panel A: Good period				
Portfolio turnover								
25FF	0.256	0.010	0.306	0.048	0.032	0.051	0.023	0.013
48IND	0.321	0.013	0.445	0.052	0.043	0.060	0.021	0.017
100SBM	0.505	0.022	0.563	0.097	0.086	0.102	0.025	0.020
100SOP	0.629	0.028	0.707	0.103	0.053	0.115	0.037	0.029
100SI	0.432	0.020	0.501	0.078	0.067	0.084	0.032	0.022
Ind30	0.307	0.015	0.342	0.054	0.033	0.043	0.027	0.017
Ind50	0.566	0.029	0.633	0.102	0.052	0.077	0.038	0.032
Ind100	0.822	0.033	0.918	0.148	0.076	0.113	0.052	0.034
Ind250	0.916	0.043	1.005	0.166	0.098	0.142	0.054	0.045
Ind500	1.007	0.049	1.105	0.199	0.097	0.160	0.071	0.052
Portfolio sharpe ratio with transaction cost								
25FF	0.207	0.264	0.236	0.254	0.289	0.284	0.291	0.306
48IND	0.178	0.239	0.199	0.216	0.241	0.237	0.243	0.256
100SBM	0.165	0.209	0.191	0.228	0.225	0.221	0.227	0.234
100SOP	0.187	0.234	0.194	0.221	0.230	0.225	0.238	0.247
100SI	0.192	0.218	0.192	0.207	0.241	0.234	0.246	0.252
Ind30	0.229	0.322	0.247	0.234	0.313	0.296	0.327	0.331
Ind50	0.199	0.267	0.253	0.218	0.287	0.263	0.279	0.285
Ind100	0.209	0.279	0.264	0.219	0.280	0.271	0.284	0.293
Ind250	0.167	0.251	0.221	0.189	0.239	0.238	0.244	0.259
Ind500	0.163	0.213	0.198	0.167	0.209	0.206	0.217	0.228
				Panel B: Bad period				
Portfolio turnover								
25FF	0.312	0.030	0.332	0.068	0.065	0.076	0.053	0.044
48IND	0.363	0.041	0.476	0.077	0.081	0.086	0.072	0.063**
100SBM	0.552	0.052	0.603	0.118	0.097	0.128	0.066	0.063***
100SOP	0.671	0.051	0.734	0.139	0.095	0.145	0.074	0.067***
100SI	0.486	0.037	0.527	0.107	0.085	0.106	0.053	0.048***
Ind30	0.349	0.032	0.379	0.078	0.061	0.064	0.051	0.044
Ind50	0.618	0.048	0.665	0.126	0.089	0.108	0.067	0.058**
Ind100	0.865	0.054	0.939	0.177	0.099	0.151	0.078	0.063***
Ind250	0.954	0.063	1.041	0.198	0.123	0.167	0.089	0.072***
Ind500	1.052	0.078	1.153	0.224	0.131	0.189	0.094	0.088***
Portfolio sharpe ratio with transaction cost								
25FF	0.184	0.243	0.223	0.198	0.229	0.232	0.233	0.245
48IND	0.152	0.219	0.174	0.159	0.202	0.207	0.207	0.218**
100SBM	0.137	0.203	0.163	0.154	0.203	0.201	0.206	0.217***
100SOP	0.148	0.202	0.171	0.163	0.199	0.204	0.194	0.215***
100SI	0.144	0.174	0.165	0.162	0.205	0.186	0.225	0.213***
Ind30	0.209	0.298	0.223	0.218	0.298	0.276	0.305	0.304
Ind50	0.169	0.231	0.206	0.198	0.250	0.248	0.251	0.261**
Ind100	0.181	0.255	0.213	0.199	0.259	0.252	0.268	0.276***
Ind250	0.148	0.221	0.175	0.164	0.225	0.227	0.226	0.232***
Ind500	0.132	0.184	0.157	0.148	0.198	0.192	0.201	0.205***

Table 6: **Robustness check using longer estimation window for competing methods.** For each data set, this table reports the out-of-sample overall turnover and monthly sharpe ratio for the following eight portfolios: portfolio using our RRW estimator (denoted as GMV_{RRW}); portfolio using rolling window based shrunk sample estimate of the covariance matrix (denoted as $\text{GMV}_{\text{sample}}$); equally weighted portfolio (denoted as $\text{GMV}_{\text{equal}}$); portfolio using EWMA based estimates of covariance matrix (denoted as GMV_{EWMA}); portfolio using the linear shrinkage covariance estimator by Ledoit and Wolf (2003) (denoted as GMV_{Lin}); portfolio using the nonlinear shrinkage covariance estimator by Ledoit and Wolf (2017) (denoted as $\text{GMV}_{\text{Nonlin}}$); portfolio using the robust covariance estimator, the Minimum Covariance Determinant (denoted as $\text{GMV}_{\text{robust}}$); portfolio using the mean-variance portfolio optimization technique (sparse regression) suggested by Ao, Yingying, and Zheng (2019) (denoted as $\text{GMV}_{\text{Sparreg}}$). The RRW estimator is formed with rolling window of $M = 60$, and other covariance matrix estimators are formed with rolling window of $M = 120$. The transaction cost of each is calculated as 50 basis points times monthly turnover times 12 (to annualize). We test the significance of any difference between GMV_{RRW} and other alternatives using the stationary bootstrap of Politis and Romano (1994). We only report the significance between GMV_{RRW} and the closest alternative due to the restricted space. *, **, and *** indicate significant differences at the 10%, 5% and 1% levels, respectively.

Date set	$\text{GMV}_{\text{sample}}$	$\text{GMV}_{\text{equal}}$	GMV_{EWMA}	GMV_{Lin}	$\text{GMV}_{\text{Nonlin}}$	$\text{GMV}_{\text{robust}}$	$\text{GMV}_{\text{Sparreg}}$	GMV_{RRW}
Panel A: Portfolio Risk (monthly return Variance ($\%^2$))								
25FF	28.1230	25.4329	14.4114	16.7763	14.123	14.2202	13.7854	13.7056**
48IND	26.2431	24.4389	13.8438	16.1234	13.436	13.6621	13.2321	13.1564***
100SBM	29.5673	28.4764	16.0443	18.5348	15.567	15.8414	15.3434	15.2387**
100SOP	30.5976	29.2564	16.4765	19.2012	16.987	16.2673	15.7653	15.6654***
100SI	31.1124	29.7322	16.7761	19.5332	16.342	16.5658	16.0621	15.8876***
Ind30	25.1213	24.2543	13.6632	15.9212	13.231	13.4873	13.0776	12.9877**
Ind50	27.5543	26.5438	14.9765	17.4364	14.565	14.7764	14.3256	14.2012*
Ind100	32.7675	31.3321	17.6432	20.5543	17.098	17.4134	16.8876	16.7765***
Ind250	36.6543	34.3217	19.4221	22.6154	19.078	19.1543	18.5654	18.4435**
Ind500	38.8987	36.5097	20.5765	23.9621	19.998	20.3012	19.6783	19.5467**
Panel B: Portfolio Turnover								
25FF	0.294	0.021	0.332	0.059	0.065	0.078	0.035	0.021***
48IND	0.352	0.019	0.470	0.061	0.076	0.084	0.043	0.027***
100SBM	0.541	0.034	0.597	0.112	0.089	0.125	0.043	0.039*
100SOP	0.656	0.043	0.732	0.132	0.112	0.143	0.061	0.051**
100SI	0.468	0.032	0.519	0.097	0.098	0.107	0.043	0.035*
Ind30	0.321	0.019	0.376	0.069	0.059	0.064	0.033	0.022***
Ind50	0.589	0.037	0.664	0.121	0.099	0.103	0.058	0.048**
Ind100	0.852	0.051	0.925	0.173	0.131	0.153	0.067	0.061*
Ind250	0.940	0.049	1.029	0.189	0.146	0.165	0.079	0.069***
Ind500	1.024	0.062	1.143	0.213	0.157	0.178	0.087	0.070**
Panel C: Portfolio Sharpe Ratio with transaction cost (monthly)								
25FF	0.199	0.261	0.228	0.212	0.285	0.271	0.294	0.294
48IND	0.165	0.229	0.194	0.181	0.231	0.229	0.247	0.253**
100SBM	0.154	0.212	0.175	0.164	0.216	0.201	0.223	0.232**
100SOP	0.157	0.215	0.186	0.175	0.226	0.217	0.231	0.231
100SI	0.168	0.171	0.190	0.179	0.225	0.228	0.228	0.243**
Ind30	0.219	0.312	0.245	0.231	0.301	0.293	0.305	0.323***
Ind50	0.193	0.251	0.218	0.211	0.254	0.248	0.274	0.289**
Ind100	0.199	0.272	0.230	0.210	0.267	0.253	0.286	0.271***
Ind250	0.161	0.244	0.195	0.184	0.225	0.219	0.241	0.253***
Ind500	0.143	0.206	0.171	0.163	0.210	0.208	0.212	0.221***

Table 7: **Robustness test using weekly returns.** For each data set, this table reports the out-of-sample overall turnover and monthly sharpe ratio for the following eight portfolios: portfolio using our RRW estimator (denoted as MV_{RRW}); portfolio using rolling window based shrunk sample estimate of the covariance matrix (denoted as MV_{sample}); equally weighted portfolio (denoted as MV_{equal}); portfolio using EWMA based estimates of covariance matrix (denoted as MV_{EWMA}); portfolio using the linear shrinkage covariance estimator by Ledoit and Wolf (2003) (denoted as MV_{Lin}); portfolio using the nonlinear shrinkage covariance estimator by Ledoit and Wolf (2017) (denoted as MV_{Nonlin}); portfolio using the robust covariance estimator, the Minimum Covariance Determinant (denoted as MV_{robust}); portfolio using the mean-variance portfolio optimization technique (sparse regression) suggested by Ao, Yingying, and Zheng (2019) (denoted as $MV_{Sparreg}$). The transaction cost of each is calculated as 50 basis points times weekly turnover times 52 (to annualize). We test the significance of any difference between GMV_{RRW} and other alternatives using the stationary bootstrap of Politis and Romano (1994). We only report the significance between GMV_{RRW} and the closest alternative due to the restricted space. *, **, and *** indicate significant differences at the 10%, 5% and 1% levels, respectively.

Date set	MV_{sample}	MV_{equal}	MV_{EWMA}	MV_{Lin}	MV_{Nonlin}	MV_{robust}	$MV_{Sparreg}$	MV_{RRW}
Panel A: Portfolio Risk (monthly return Variance ($\%^2$))								
25FF	28.2345	25.4432	14.4214	16.7787	14.187	14.2214	13.7867	13.7043***
48IND	26.2543	24.4397	13.8564	16.1256	13.546	13.6654	13.2332	13.1555***
100SBM	29.5786	28.4851	16.0567	18.5353	15.678	15.8423	15.3456	15.2371**
100SOP	30.5981	29.2654	16.4897	19.2022	16.121	16.2689	15.7653	15.6662***
100SI	31.1235	29.7434	16.7878	19.5345	16.342	16.5665	16.0632	15.8889***
Ind30	25.1245	24.2665	13.6786	15.9232	13.231	13.4868	13.0787	12.9881**
Ind50	27.5567	26.5546	14.9897	17.4378	14.564	14.7753	14.3264	14.2023***
Ind100	32.7675	31.3431	17.6554	20.5554	17.223	17.4123	16.8889	16.7779***
Ind250	36.6654	34.3342	19.4234	22.6165	19.025	19.1554	18.5644	18.4441***
Ind500	38.9074	36.5123	20.5776	23.9785	20.128	20.3023	19.6772	19.5475***
Panel B: Portfolio Turnover								
25FF	0.298	0.028	0.345	0.063	0.075	0.089	0.046	0.024***
48IND	0.345	0.021	0.488	0.072	0.087	0.092	0.051	0.032***
100SBM	0.543	0.029	0.601	0.106	0.138	0.144	0.054	0.041**
100SOP	0.667	0.047	0.756	0.145	0.143	0.156	0.072	0.055***
100SI	0.472	0.054	0.523	0.104	0.103	0.114	0.053	0.032***
Ind30	0.322	0.032	0.386	0.077	0.065	0.076	0.031	0.020***
Ind50	0.594	0.055	0.681	0.134	0.106	0.112	0.055	0.049**
Ind100	0.848	0.065	0.944	0.186	0.144	0.165	0.072	0.071
Ind250	0.946	0.051	1.032	0.192	0.168	0.173	0.083	0.065***
Ind500	1.038	0.069	1.156	0.233	0.177	0.189	0.089	0.072**
Panel C: Portfolio Sharpe Ratio with transaction cost (monthly)								
25FF	0.194	0.263	0.229	0.224	0.287	0.272	0.299	0.301
48IND	0.163	0.234	0.195	0.198	0.244	0.231	0.251	0.267***
100SBM	0.156	0.221	0.176	0.153	0.226	0.214	0.232	0.245***
100SOP	0.158	0.210	0.183	0.186	0.229	0.213	0.236	0.242**
100SI	0.172	0.178	0.191	0.183	0.231	0.232	0.234	0.241**
Ind30	0.223	0.320	0.248	0.244	0.305	0.298	0.316	0.324**
Ind50	0.198	0.262	0.219	0.226	0.269	0.251	0.279	0.290***
Ind100	0.204	0.288	0.233	0.234	0.276	0.259	0.292	0.292
Ind250	0.167	0.251	0.198	0.196	0.231	0.223	0.245	0.255***
Ind500	0.146	0.213	0.175	0.165	0.215	0.216	0.218	0.224***

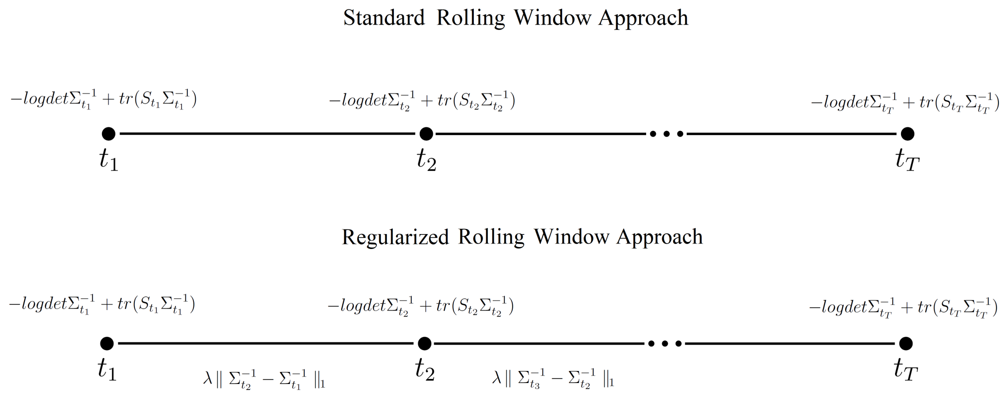

Figure 1: **A graphical comparison between the standard and regularized "rolling window" approach.** The top panel shows the estimation procedure of standard rolling window approach and the bottom panel shows the estimation procedure of the regularized rolling window approach.

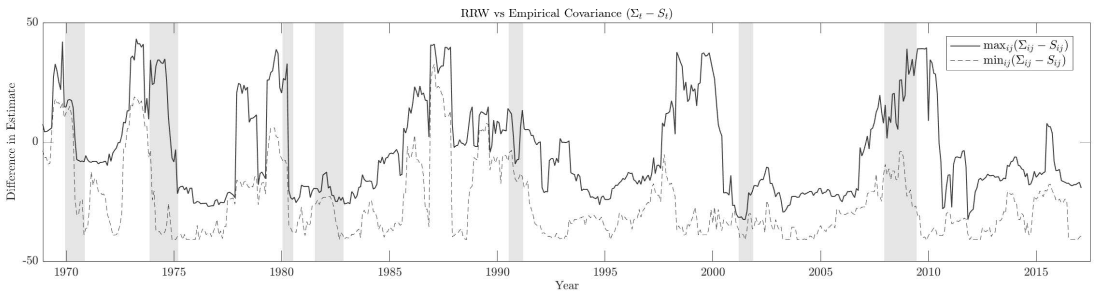

Figure 2: Plot of how $\max_{ij}(\Sigma_{RRW,t} - S_t)$ and $\min_{ij}(\Sigma_{RRW,t} - S_t)$ vary as a function of time. According to Eq. 6 the RRW estimator should always maintain $\max_{ij}(|[\hat{\Sigma}_{RRW,t} - \hat{S}_t]_{ij}|) \leq \lambda$. Note how the difference tends to increase before or around recession period, these are periods where jumps in the portfolio position (and estimated covariance) are likely to occur, see Figs. 3 for comparison.

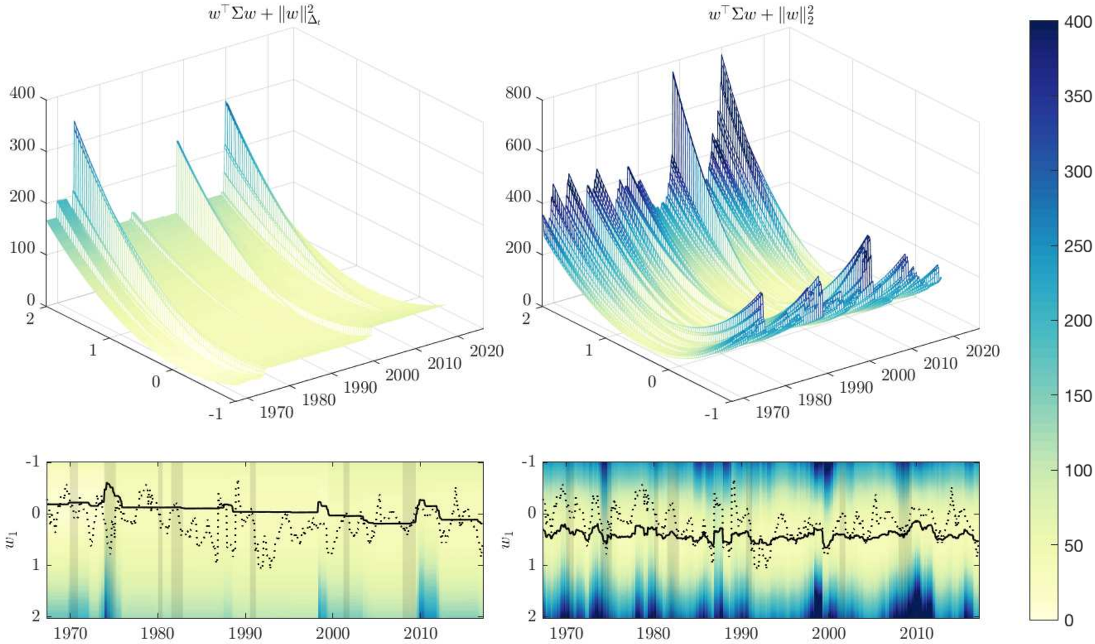

Figure 3: Surface plots of the portfolio optimization objective function over business cycles. The optimum portfolio lies at the minimiser of these objective functions and the corresponding portfolio allocation solution for first asset (that is, w_1) is indicated in the lower panels by the solid line v.s w_1 produced by minimizing $w'\Sigma w$ without transaction cost (the dashed line). We set estimation window length $M = 12$ and $\lambda = 40$ to allow comparison with Figs. 2. The grey overlaid bands (in the lower panes) denote recession periods. Note: we here use the notation $\|w\|^2_{\Delta_t} := (w - w_t^+)^\top \Delta_t (w - w_t^+)$.

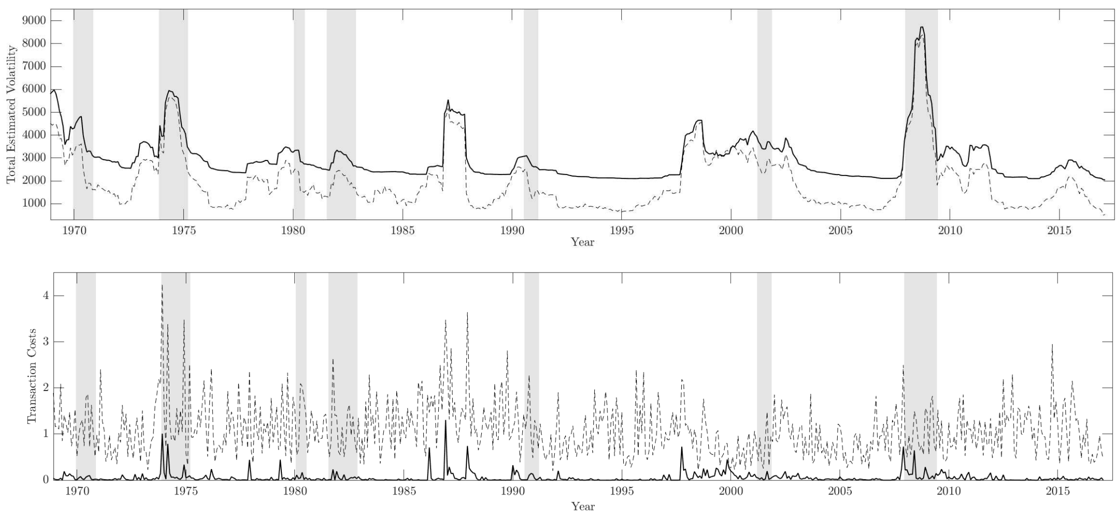

Figure 4: Top: Estimated Total Market Volatility ($\sum_{i=1}^{N} \hat{\Sigma}_{ii}$) for both the standard rolling window (dashed), and regularised rolling window (solid). Bottom: Transaction costs as measured via portfolio weights $\|w_t - w_{t-1}\|_1$. Note: In this case, we set $M = 12$ and $\lambda = 40$ to highlight the changes in variance which can occur in periods of recession.