A Review of Mean Variance Optimization Methods

Haoyu Lei*

June 1, 2020

Abstract

I review mean-variance optimization (MVO) methods by replicating Goto and Xu (2015) and extending the results to 201912. I have three findings: (1) The Graphical Lasso method consistently delivers low out-of-sample (OOS) risk while exploiting hedging relations among assets, across all samples. (2) The Jagannathan and Ma (2003) nonnegativity constraint method is not as disappointing in reducing OOS risk as previously documented. In 100 randomized samples, it reduces similar amount of risks as Glasso in tests ending in 201012 and 201912, and produces the highest Sharpe ratios in both tests using these samples. (3) The recent bull market had an impact on method performance from 201012-201912 in the time series, as revealed by reductions in OOS risk, and increases in Sharpe ratios. However, cross-sectionally, the relative performance of these MVO methods remain stable.

*The University of Hong Kong.

Contents

I. Introduction **3**

II. Data **4**

III. Methodology **5**

A. Mean Variance Optimization Methods . . . 5

A.1 Sample Estimator . . . 5

A.2 Goto and Xu (2015) Glasso Estimator . . . 5

A.3 Jagannathan and Ma (2003) Nonnegativity Constraint . . . 5

A.4 Ledoit and Wolf (2004) Shrinkage Estimator . . . 6

A.5 DeMiguel, Garlappi, and Uppal (2009) $1/N$ Estimator . . . 6

A.6 Industry Factor Model Estimator . . . 6

B. Backtest Procedures . . . 6

IV. Results and Discussions **7**

A. Out-of-Sample Portfolio Risk . . . 7

B. Out-of-Sample Sharpe Ratio . . . 7

C. Distribution of Portfolio Weights . . . 7

D. Portfolio Turnover . . . 8

E. Transaction Cost-Adjusted Certainty-Equivalent Returns . . . 8

F. 100 Randomized Samples: Distribution of Realized Volatility, Sharpe Ratio and Weights, and Risk Reduction versus Sparsity . . . 9

V. Conclusion **10**

VI. Appendix **11**

A. Ledoit and Wolf (2004) Shrinkage Estimator . . . 11

A.1 Formula for Shrinkage Target . . . 11

A.2 Formula for Shrinkage Intensity . . . 11

List of Tables

1 Data Description from Goto and Xu (2015), and for Tests Ending in 201012 14
2 Data Description for Tests Ending in 201912 14
3 Out-of-Sample Periods, Regularization Parameters, and Sparsity for Estimated $\hat{\Psi}_\rho$ from Goto and Xu (2015) 15
4 Out-of-Sample Periods, Regularization Parameters, and Sparsity for Estimated $\hat{\Psi}_\rho$ in Tests Ending in 201012 15
5 Out-of-Sample Periods, Regularization Parameters, and Sparsity for Estimated $\hat{\Psi}_\rho$ in Tests Ending in 201912 16
6 Out-of-Sample Portfolio Risk (monthly) from Goto and Xu (2015) 17
7 Out-of-Sample Portfolio Risk (monthly) in Tests Ending in 201012 17
8 Out-of-Sample Portfolio Risk (monthly) in Tests Ending in 201912 18
9 Out-of-Sample Sharpe Ratio (monthly) from Goto and Xu (2015) 19
10 Out-of-Sample Sharpe Ratio (monthly) in Tests Ending in 201012 19
11 Out-of-Sample Sharpe Ratio (monthly) in Tests Ending in 201912 20
12 Distribution of Portfolio Weights from Goto and Xu (2015) 21
13 Distribution of Portfolio Weights in Tests Ending in 201012 22
14 Distribution of Portfolio Weights in Tests Ending in 201912 23
15 Portfolio Turnover (monthly) from Goto and Xu (2015) 24
16 Portfolio Turnover (monthly) for Tests Ending in 201012 24
17 Portfolio Turnover (monthly) for Tests Ending in 201912 25
18 T_COST-ADJUSTED_CERs (annual %) from Goto and Xu (2015) . . . 26
19 T_COST-ADJUSTED_CERs (annual %) in Tests Ending in 201012 . . . 26
20 T_COST-ADJUSTED_CERs (annual %) in Tests Ending in 201912 . . . 27
21 Distribution of Realized Volatility, Sharpe Ratio and Weights, and Risk Reduction versus Sparsity for Individuals from Goto and Xu (2015) . . . 28
22 Distribution of Realized Volatility, Sharpe Ratio and Weights, and Risk Reduction versus Sparsity for Individuals in Tests Ending in 201012 . . . 29
23 Distribution of Realized Volatility, Sharpe Ratio and Weights, and Risk Reduction versus Sparsity for Individuals in Tests Ending in 201912 . . . 30

I. Introduction

The portfolio theory (Markowitz, 1952) construct mean-variance efficient portfolios given investors' expected asset returns and covariances. However, estimation of these moments is usually difficult. The covariance estimates exhibit significant sampling errors and specification errors, resulting in poor out-of-sample (OOS) performance. Therefore, considerable efforts were made to improve covariance matrix estimation.

In the recent decade, the U.S. equity market has experienced considerable changes, for example, lower market volatility, higher returns, and thus higher reward-to-risk ratio. It is therefore of great academic and practical interests to ask whether these changes affected the performance of previously documented mean-variance optimization (MVO) methods, and more importantly, why. To the best of my knowledge, this is a piece of work that remains to be done.

With recent data, this paper reexamines mean-variance optimization methods shown to achieve some degrees of success towards reducing covariance matrix estimation errors, thus OOS portfolio risk. They include Jagannathan and Ma's (2003) nonnegativity-constraint method (JM), Ledoit and Wolf's (2004) shrinkage estimation method (LW), DeMiguel, Garlappi, and Uppal's (2009) naïve diversification rule (EW), industry factor model (IND), and more recently, Goto and Xu's (2015) Glasso estimation method (Glasso), which directly reduces inverse covariance matrix estimation errors. The sample covariance matrix (Sample) is also included as a baseline.

I conduct the reexamination by comparing OOS performance of global minimum variance (GMV) portfolios. I use GMV portfolios instead of tangency portfolios to control for errors arising from estimating expected returns. Specifically, I discuss two comparisons along each performance measure. First, controlling for backtest period, I summarize the results from Goto and Xu (2015) and those from my replication, to examine the discrepancies possibly attributable to differences in datasets and replication procedures. Second, controlling for datasets and replication procedures, I vary backtest periods, compare results till 201012 with those till 201912, and inspect whether the previously documented results still hold after 2010.

I summarize the main findings from this empirical exercise:

1. Glasso consistently shows superior ability to reduce OOS risk while exploiting hedging relations among assets, across all samples. OOS Sharpe ratios and certainty-equivalent returns of Glasso portfolios are decently high among its peers, but not as impressive as their reductions in OOS variances.

2. JM is not so worse in reducing OOS risk as previously documented. In 100 randomized samples, its risk reduction ability is on par with Glasso in tests ending in 201012, and second-best to Glasso in tests ending in 201912. It also delivers the highest Sharpe ratios in both tests with 100 samples.

3. Since Dec. 2010, there has been a 10-year boom in the U.S. market. The changes are embodied in my findings, as revealed by reductions in OOS risk, and increases in Sharpe ratios. However, the cross-sectional differences remain stable, i.e., the relative strengths among these MVO methods did not change significantly.

The rest of the paper is organized as follows. Section II presents data sources. Section III introduces (inverse) covariance matrix estimation methods, backtest procedures, and metrics for comparison. Section IV discusses results from the two comparisons. Section V concludes. The Appendix lists method to vectorize some scalar formulas.

II. Data

Tables 1 and 2 list the data sets I use to test mean-variance optimization methods. They are identical except for the last column, Time Period, which indicates the testing period for each data set. The data sets can be divided into two groups: the first group of data sets, data set 1-3, include large portfolios as assets. The second group, data set 4, consists of 100 data sets with individual stocks as test assets.

$\boxed{\text{Tables 1 and 2 about here.}}$

The first group of data sets are well-known in the literature. Data set 1 is 100 portfolios formed on size and book-to-market ratio. Data set 2 contains 48 Fama-French industry portfolios. Data set 3 is a combination of 1 and 2. Missing values are replaced with CRSP value-weighted market returns. They can be found on Professor French's website. I include these data sets for three reasons. First, they are standard data sets commonly used in empirical tests. Second, they serve as benchmark data sets to evaluate the discrepancies between my results and the results established in Goto and Xu (2015), thus surrecting the credibility of my findings. Data sets based on randomly sampled individual stocks are not sufficient to meet this purpose. Third, a mean-variance optimization method may have different performance on well-diversified portfolios and individual stocks with high idosyncratic volatility, so for the sake of completeness, it is necessary to consider both types of assets.

The second group is 100 data sets of 100 randomly sampled individual stocks. The sampling procedures follow Jagannathan and Ma (2003) and Goto and Xu (2015). At the end of Dec. 1982, I choose 100 stocks at random from NYSE and AMEX without replacement, with prices greater than $5 and price records available in the past 10 years. In test periods, I replace missing values with CRSP value-weighted market returns. Finally, I repeat the procedures above 100 times to generate 100 samples. It should be noted that a considerable part of individual stocks were delisted in subsequent periods, and replacing these missing returns with CRSP value-weighted market returns is equivalent to reinvesting in the market index.

Although replacement with market returns is a standard practice, there are some critical issues. First, the samples may suffer from significant multicollinearity, as towards the end of testing periods, a considerable portion of test asset return series are replaced with a single market return series. There is a nuanced distinction between such multicollinearity and the commonly-recognised multicollinearity arising naturally from individual stocks' data generating processes: the former is simply an undesirable by-product when we clean the data.

Second, such a treatment may impede a test from revealing the true ability of an optimization method to reduce the OOS risk of *individual-stock-based* portfolios. Consider a hypothetical sample where all individual stocks had been delisted before 201011. Then in 201012, an efficient portfolio will be formed on 100 perfectly correlated, well-diversified market portfolios, which does not satisfy the original purpose of our tests. For reference, on average, 74.43% of stocks were missing or delisted across individual samples in 201012; the figure increased to 80.40% in 201912.

To conclude, when assessing the OOS performance of different mean variance optimization methods, replacing missing values with market returns with a lot of delistings in a sample is likely to overweigh methods that better handle the *perfect* multicollinearity problem, and risk veiling methods addressing the true multicollinearity among individual stocks. A workaround may be to resample a new series when a delisting happens, and in practice, this will be equivalent to reinvesting in a new individual stock, instead of market index. However, this is another story.

III. Methodology

A. Mean Variance Optimization Methods

The GMV portfolio weights are given by

$$w_{GMV} = \frac{\Sigma^{-1}\iota}{\iota^{\intercal}\Sigma^{-1}\iota}, \tag{1}$$

where Σ^{-1} is the inverse covariance matrix, or the mean-variance optimizer, and ι is a vector of 1s.

A.1 Sample Estimator

The sample estimator, $\hat{S}^{-1}$, is the inverse of sample covariance matrix.

A.2 Goto and Xu (2015) Glasso Estimator

Goto and Xu (2015) note that the parameters in the mean-variance optimizer Ψ reveal hedge relationships among assets in the portfolio, i.e., each row i in Ψ denotes the coefficients obtained by regressing returns from asset i on the rest of the assets ("hedge regressions"), and then scaled by inverse residual variance. The multicollinearity problem in hedge regressions can be tackled by imposing l_1 panelties on the magnitude of parameters, which is Lasso regression. Furthermore, graphical lasso (Glasso) helps restrict the resulting mean-variance optimizer to be positive definite while keeping the benefits of Lasso.

A.3 Jagannathan and Ma (2003) Nonnegativity Constraint

The vector of optimal GMV portfolio weights, w, is derived by solving the optimization problem

$$\min_{w} w'Sw \tag{2}$$

$$\text{s.t.} \sum_{i} w_i = 1 \tag{3}$$

$$w_i \geq 0, i = 1, 2, ..., N, \tag{4}$$

where S is the sample covariance matrix.

A.4 Ledoit and Wolf (2004) Shrinkage Estimator

Ledoit and Wolf (2004) propose to shrink the sample covariance matrix towards a constant correlation "target" matrix, resulting in a shrinkage covariance matrix estimator. The Appendix outlines the steps to derive the estimator and vectorize some formulas given in scalar form.

A.5 DeMiguel, Garlappi, and Uppal (2009) $1/N$ Estimator

DeMiguel, Garlappi, and Uppal (2009) employ the naïve diversification rule to generate GMV portfolio weights. Each asset receives a weight of $1/N$, and the weight vector is given by

$$w = \frac{1}{\iota^\intercal \iota}\iota, \tag{5}$$

where ι is a vector of 1s with length N being the total number of assets. Equivalently, the estimator is

$$\hat{\Sigma}_{EW}^{-1} = I, \tag{6}$$

where I is an $N \times N$ identity matrix.

A.6 Industry Factor Model Estimator

Besides testing the methods above on portfolios, I test the industry factor model estimator on 100 samples of 100 randomly chosen individual stocks. Each stock is mapped to one of the Fama-French 30 industry portfolios based on their SIC codes. At the beginning of each month, I invert the covariance matrix of past 120-month industry portfolio returns, $\hat{\Sigma}_{IND}$, to get the industry factor model estimator. Notice that because of its nature, this method does not apply to backtests using portfolios as test assets.

B. Backtest Procedures

A backtest involves two parts: a training period, and a testing period. Only the Glasso method requires training period to search for a regularization parameter, ρ. Tables 4 and 5 summarize the OOS periods, regularization parameter, and OOS sparsity for estimated $\hat{\Psi}_\rho$. Table 4 is dedicated for testing periods ending in 201012, and Table 5 is for periods ending in 201912.

Tables 4 and 5 about here.

I take a "sliding window" approach during the testing period. For each of the methods listed above, at the beginning of each month, I form a sample of past 120-month asset returns, and estimate a mean variance optimizer. At month-end, I calculate GMV portfolio statistics using the estimated optimizer. After that, at the beginning of next month, I exclude the most remote observation and include the most recent observation in the previous sample, to form a new sample. This sample is again used to estimate an optimizer and calculate end-of-month statistics. I repeat the procedures until the end of testing period and obtain statistics (or series of statistics) for subsequent comparison and analysis.

IV. Results and Discussions

In this section, I will discuss the OOS performance of MVO methods on various metrics.

A. Out-of-Sample Portfolio Risk

Tables 6, 7, and 8 summarize the OOS portfolio risk performance. Overall, Glasso achieves the best OOS risk reduction across all data sets during both 196307-201012 and 196307-201912, while Sample and EW are consistently ranked among the worst performers.

Tables 6, 7, and 8 about here.

JM performs better than originally reported in test period 196307-201012 in all data sets, especially in SZBM100. I investigate two possible reasons: differences in linear programming packages or backtest procedures. First, I tried two different R packages, osqp and quadprog, and they report similar OOS variances. Second, the differences in variances among other methods because of backtest procedures are not as significant, comparing Tables 6 with 7. Therefore, it is relatively safe to rule out these two reasons.

B. Out-of-Sample Sharpe Ratio

Tables 9, 10, and 11 list the OOS Sharpe ratios. Among portfolio-based data sets, Glasso and LW attain the highest Sharpe ratios in data sets 1 and 3 in tests ending in 201012; and in data sets 1, 2, and 3 in tests ending in 201912. In individual data sets, Glasso and LW exhibit slightly weaker performance than EW and JM before 201012. Nevertheless, the differences become less significant up and until 201912.

Tables 9, 10, and 11 about here.

Notice that in the time series, there have been noticeable increases in Sharpe ratios after 201012 comparing the performance in two test periods. The key reason for this enhancement is more likely driven by a bull market in the past 10 years, rather than the outstanding ability of an MVO method to increase OOS return per unit of risk.

Looking at cross-sections, the differences in Sharpe ratios between different MVO methods remain relatively stable in data sets 1-3, and less stable in data set 4. Glasso consistently outperforms the rest methods in SZBM100 and IND48, and is marginally worse than LW in the two data sets combined. In individual-stock-based samples, Glasso is consistently better than Sample, LW, IND and worse than JM. The evidence is less clear on whether Glasso or EW is better at improving OOS Sharpe ratios.

C. Distribution of Portfolio Weights

Tables 12, 13 and 14 show the distribution of portfolio weights. For each method applied to a data set, I pool all portfolio weights, and report their statistics: minimum, maximum, 1%/5%/95%/99% percentile weights. I also report Herfindahl index, which measures the concentration of portfolio weights in a pool: the larger the index is, the less evenly-distributed are the weights.

Tables 12, 13 and 14 about here.

In every data set, Glasso and JM generate more evenly-distributed weights, and less extreme weights, during all test periods. However, Glasso produces more reasonable negative weights, indicating its ability to exploit hedging relations among assets, whereas JM assigns only positive weights to quite few assets, refusing to invest in the rest. In contrast, LW is inferior in terms of minimum and maximum weights, and Herfindahl index; Sample is even worse.

D. Portfolio Turnover

Following DeMiguel, Garlappi, Nogales, and Uppal (2009), I calculate portfolio turnover as

$$\text{Turnover} = \frac{1}{T-\tau-1}\sum_{t=\tau}^{T-1}\sum_{j=1}^{N}(|w_{j,t+1}^{i} - w_{j,t^{+}}^{i}|), \tag{7}$$

where T is the total number of testing periods, $\tau = 120$ is the sliding window size, $w_{j,t}^{i}$ is the weight of asset i at the end of period t, and $w_{j,t^{+}}^{i}$ is the weight of asset i right before rebalancing at the end of period $t+1$. The turnover can be interpreted as the average fraction of wealth being transacted on rebalancing dates. Tables 15, 16, 17 report portfolio turnovers.

Tables 15, 16, 17 about here.

The EW strategy remains to have the lowest turnover, because it doesn't take too many trades to maintain a level of $1/N$ for each asset. JM follows, with the second lowest turnover. After that, Glasso and LW portfolios have higher turnovers because they take advantage of short positions. Finally, the Sample strategy delivers the highest turnover, since it usually takes extreme positions as a result of high estimation errors.

E. Transaction Cost-Adjusted Certainty-Equivalent Returns

Tables 18, 19, 20 report transaction cost-adjusted certainty-equivalent returns (T_COST-ADJUSTED_CERs). We can interpret it as the excess return after transaction costs an investor expect from a portfolio, assuming no arbitrage and he has a quadratic utility. The higher this measure of a portfolio, the more desirable risk-return profile the portfolio has.

The formula of T_COST-ADJUSTED_CER is given by

$$\text{T\_COST-ADJUSTED\_CERs} = \hat{\mu} - \frac{\gamma}{2}\hat{\sigma}^{2} - \text{T\_COST}, \tag{8}$$

where $\hat{\mu}$ is the annualized mean excess return, $\hat{\sigma}^2$ is the annualized excess return variance, and

$$\text{T\_COST} = \text{Turnover} \times 12 \times 50\text{bps}, \tag{9}$$

which is the annualized turnover times 50 basis points for each trade.

Tables 18, 19, 20 about here.

Glasso achieves the best economic gains in two (1, 3) out of the first three data sets, in the original results and during all periods, while JM dominates in the second data set. Although JM is less desirable in the originally reported results, it is on par with Glasso and EW strategies in 100 randomized portfolios. The Sample method consistently realizes the worst performance, as expected.

F. 100 Randomized Samples: Distribution of Realized Volatility, Sharpe Ratio and Weights, and Risk Reduction versus Sparsity

I take a closer look at the performance of MVO methods on data set 4, 100 randomized samples. Specifically, I form five panels, which are in Tables 21, 22, 23.

Tables 21, 22, 23 about here.

Panel A, B, and C show the summary statistics of OOS return variances, Herfindahl index of portfolio weights, and Sharpe ratios of 100 randomized portfolios for each MVO method. Specifically, a panel reports the mean, standard deviation, minimum and maximum values. It also reports the number of 100 portfolios on which GMV-$\hat{\Psi}_\rho$ has a measure greater than and less than, respectively, an alternative portfolio. Panel D and E examine the relationship between $\hat{\Psi}_\rho$'s OOS risk reduction ability and its sparsity.

In Panel A, GMV-$\hat{\Psi}_\rho$ portfolios have a relatively stable distribution, with low first and second moments. In the test ending in 201012, Glasso's and JM's performances are similar, as revealed in the 100 comparisons: JM has lower variances in 52 comparisons, and Glasso wins in the rest 48 times. However, Glasso dominates in the test ending in 201912. Therefore, there is stronger evidence that Glasso reduces OOS variance better than the rest of the MVO methods.

GMV-$\hat{\Psi}_\rho$ is also the most impressive in terms of Herfindahl index. In Panel B, GMV-$\hat{\Psi}_\rho$ portfolios assign the second-most evenly distributed weights to their assets, following GMV-$1/N$ portfolios, which by definition are guaranteed to have the lowest Herfindahl. This result means that the Glasso method is better than others at exploiting the hedge relationships and allocate reasonable proportion of wealth to different assets.

The evidence of OOS Sharpe ratios are mixed in Panel C. In three tables, GMV-$\hat{\Psi}_\rho$ portfolios have higher Sharpe ratios in most of the cases than GMV-LW, GMV-IND, and GMV-Sample portfolios. The rankings among GMV-$\hat{\Psi}_\rho$, GMV-$1/N$, and GMV-JM are mixed, however. First, Glasso is worse than EW in 100 comparisons in the original results and tests ending in 201012, but the situation is reversed in tests ending in 201912. This discrepancy is less likely to result from replication errors, but more of a result from changes in the data generating process. Therefore, it indicates that Glasso is more sophisticated in exploiting the covariance structure among individual stocks than EW after 201012.

Second, Glasso outperforms JM in the original report, but underperforms in replications ending in both 201012 and 201912. Combining the evidence in A., differences in replication details have a higher chance to explain the discrepancies, rather than changes in data.

Finally, I take average of monthly out-of-sample $\sigma_{\hat{S}^{-1}} - \sigma_{\hat{\Psi}_\rho}$ (or risk reduction), and monthly sparsity for each of the 100 samples, to form a panel of 100 observations. The observations can be viewed as relatively independent because I drew each of the samples at random. I then regress $\sigma_{\hat{S}^{-1}} - \sigma_{\hat{\Psi}_\rho}$ on sparsity. I report the results in Panel E: the relationship turns out to be not as significant as originally reported. In both tests ending in 201012 and 201912, a 1% increase in sparsity has effectively no effect on OOS risk reduction. It is perhaps because there was an "average" process, and the results may be more significant if I used the original observations instead.

I further group observations by sparsity level, and observe the average risk reductions in the bottom 30%, middle 40%, and top 30% levels. They are reported in Panel E. Contrary to the impression from observing Panel D, there are economically risk reductions when a sample move from bottom 30% to middle 40%. The reduction is 0.118%and 0.099% in tests ending in 201012 and 201912, respectively. But further moving the sample to top 30% will deterioriate risk reduction, although the magnitude of deterioriation is not significant.

V. Conclusion

I reexamined the performance of MVO methods on two groups of data sets: the first ones are based on large portfolios, and the second ones are based on individual stocks. I did so by replicating Goto and Xu (2015), and make comparisons both in the cross section and the time series. In the cross section, Glasso is sophisticated at reducing OOS risk while exploiting hedging relations among assets. JM is better at risk reduction than previously documented. In 100 randomized samples, it reduces similar amount of risk in tests ending in 201012, and slightly worse than Glasso in tests ending in 201912. It also produces the highest Sharpe ratios in both tests. The time series reveals the market boom in the recent decade, reflecting an overall decrease in OOS risk and increase in Sharpe ratios. But these changes did not affect the relative strengths of MVO methods, because the MVO-method differences in OOS risk and Sharpe ratios remain relatively stable.

VI. Appendix

A. Ledoit and Wolf (2004) Shrinkage Estimator

The formula for shrinkage target and shrinkage intensity given in the original paper are presented in terms of elements of a matrix. To implement the formula more efficiently in R (i.e., eliminate as many for loops as possible), I adjusted the original formula and added a few shorthands. Details are available in the R code.

A.1 Formula for Shrinkage Target

Some notations:

- $Y \equiv \{y_{it}\}, 1 \le i \le N, 1 \le t \le T$ is an $N \times T$ panel of stock/portfolio returns
- $\tilde{Y} \equiv \{\tilde{y}_{it}\}$ is an $N \times T$ panel of *demeaned* stock/portfolio returns
- $R \equiv \{r_{ij}\}, 1 \le i, j \le N$ is the sample correlation matrix
- $S \equiv \{s_{ij}\}, 1 \le i, j \le N$ is the sample covariance matrix

The average sample correlation is given by

$$\bar{r} = \frac{2}{(N-1)N} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} r_{ij} = \frac{1}{(N-1)N} \left\{ \left(\sum_{i=1}^{N} \sum_{j=1}^{N} R_{ij} \right) - N \right\} \tag{10}$$

Given $f_{ii} = s_{ii}$, $f_{ij} = \bar{r}\sqrt{s_{ii}s_{jj}}$, and $r_{ij} = s_{ij}/\sqrt{s_{ii}s_{jj}}$, we have

$$F_{ii} = S_{ii} \qquad \text{and} \qquad F_{ij} = \bar{r}\frac{S_{ij}}{R_{ij}} \tag{11}$$

A.2 Formula for Shrinkage Intensity

Given $\hat{\kappa} = (\hat{\pi} - \hat{\rho})/\hat{\gamma}$, we need to estimate each of the elements on the RHS.

$\hat{\pi}$ Start with $\hat{\pi}_{ij}$,

$$\begin{aligned} \hat{\pi}_{ij} &= \frac{1}{T} \sum_{t=1}^{T} \{(y_{it} - \bar{y}_{i\cdot})(y_{jt} - \bar{y}_{j\cdot}) - s_{ij}\}^2 \\ &= \frac{1}{T} \sum_{t=1}^{T} \{\tilde{Y}_{it}\tilde{Y}_{jt} - S_{ij}\}^2 \\ &= \frac{1}{T} \sum_{t=1}^{T} \{\tilde{Y}_{it}^2\tilde{Y}_{jt}^2 - 2S_{ij}\tilde{Y}_{it}\tilde{Y}_{jt} + S_{ij}^2\} \\ &= \frac{1}{T} \sum_{t=1}^{T} \{\tilde{Y}_{it}^2\tilde{Y}_{jt}^2\} - \frac{2S_{ij}}{T} \sum_{t=1}^{T} \{\tilde{Y}_{it}\tilde{Y}_{jt}\} + S_{ij}^2 \end{aligned} \tag{12}$$

Denote $Z = \tilde{Y}^2$, $A = \tilde{Y}\tilde{Y}^\intercal$, and $B = ZZ^\intercal$, we have

$$\hat{\pi}_{ij} = \frac{1}{T}B_{ij} - \frac{2S_{ij}}{T}A_{ij} + S_{ij}^2 \tag{13}$$

Therefore,

$$\hat{\pi} = \frac{1}{T}\sum_{i=1}^{N}\sum_{j=1}^{N} B_{ij} - \frac{2}{T}\sum_{i=1}^{N}\sum_{j=1}^{N} S_{ij}A_{ij} + \sum_{i=1}^{N}\sum_{j=1}^{N} S_{ij}^2 \tag{14}$$

Notice the last term is the frobenius norm of matrix S.

$\hat{\rho}$ Here are the original expressions:

$$\hat{\rho} = \sum_{i=1}^{N} \hat{\pi}_{ii} + \sum_{i=1}^{N}\sum_{j=1,j\neq i}^{N} \frac{\bar{r}}{2}\left(\sqrt{\frac{s_{jj}}{s_{ii}}}\hat{\vartheta}_{ii,ij} + \sqrt{\frac{s_{ii}}{s_{jj}}}\hat{\vartheta}_{jj,ij}\right) \tag{15}$$

$$\hat{\vartheta}_{ii,ij} = \frac{1}{T}\sum_{t=1}^{T}\{(y_{it}-\bar{y}_{i\cdot})^2 - s_{ii}\}\{(y_{it}-\bar{y}_{i\cdot})(y_{jt}-\bar{y}_{j\cdot}) - s_{ij}\} \tag{16}$$

$$\hat{\vartheta}_{jj,ij} = \frac{1}{T}\sum_{t=1}^{T}\{(y_{jt}-\bar{y}_{j\cdot})^2 - s_{jj}\}\{(y_{it}-\bar{y}_{i\cdot})(y_{jt}-\bar{y}_{j\cdot}) - s_{ij}\} \tag{17}$$

A formula for $\hat{\pi}_{ij}$ is given in the last section, so the first term of $\hat{\rho}$ is

$$\sum_{i=1}^{N}\hat{\pi}_{ii} = \frac{1}{T}\sum_{i=1}^{N} B_{ii} - \frac{2}{T}\sum_{i=1}^{N} S_{ii}A_{ii} + \sum_{i=1}^{N} S_{ii}^2 \tag{18}$$

Now look at the second term of $\hat{\rho}$.

$$\begin{aligned}\hat{\vartheta}_{ii,ij} &= \frac{1}{T}\sum_{t=1}^{T}\{\tilde{Y}_{it}^2 - S_{ii}\}\{\tilde{Y}_{it}\tilde{Y}_{jt} - S_{ij}\} \\ &= \frac{1}{T}\sum_{t=1}^{T}\{\tilde{Y}_{it}^3\tilde{Y}_{jt}\} - \frac{S_{ij}}{T}\sum_{t=1}^{T}\tilde{Y}_{it}^2 - \frac{S_{ii}}{T}\sum_{t=1}^{T}\{\tilde{Y}_{it}\tilde{Y}_{jt}\} + S_{ii}S_{ij}\end{aligned} \tag{19}$$

Let $W = \tilde{Y}^3$, and $C = W\tilde{Y}^\intercal$, we can further simplify the expression:

$$\hat{\vartheta}_{ii,ij} = \frac{1}{T}C_{ij} - \frac{1}{T}S_{ij}A_{ii} - \frac{1}{T}S_{ii}A_{ij} + S_{ii}S_{ij} \tag{20}$$

Let $M = \{m_{ij} : m_{ij} \equiv \sqrt{s_{jj}/s_{ii}}\}$ be a matrix of "scaling factors", then $M^\intercal = \{m_{ij} : m_{ij} = \sqrt{s_{ii}/s_{jj}}\}$. Apply the double sums and the corresponding scaling factor matrix:

$$\begin{aligned}\sum_{i=1}^{N}\sum_{\substack{j=1\\j\neq i}}^{N} M_{ij}\hat{\vartheta}_{ii,ij} &= \sum_{i=1}^{N}\sum_{\substack{j=1\\j\neq i}}^{N} M_{ij}\left\{\frac{1}{T}C_{ij} - \frac{1}{T}S_{ij}A_{ii} - \frac{1}{T}S_{ii}A_{ij} + S_{ii}S_{ij}\right\} \\ &= \frac{1}{T}\sum_{i=1}^{N}\sum_{\substack{j=1\\j\neq i}}^{N} M_{ij}\left(C_{ij} - A_{ii}S_{ij}\right) + \sum_{i=1}^{N}\sum_{\substack{j=1\\j\neq i}}^{N} M_{ij}S_{ii}\left(S_{ij} - \frac{1}{T}A_{ij}\right) \\ &= \frac{1}{T}\sum_{i=1}^{N}\sum_{\substack{j=1\\j\neq i}}^{N} M_{ij}\left(C_{ij} - A_{ii}S_{ij}\right) + \sum_{i=1}^{N}\sum_{\substack{j=1\\j\neq i}}^{N} \frac{S_{ij}}{R_{ij}}\left(S_{ij} - \frac{1}{T}A_{ij}\right)\end{aligned} \tag{21}$$

Similarly, we can derive:

$$\sum_{i=1}^{N}\sum_{\substack{j=1\\j\neq i}}^{N} M_{ji}\hat{\vartheta}_{jj,ij} = \frac{1}{T}\sum_{i=1}^{N}\sum_{\substack{j=1\\j\neq i}}^{N} M_{ji}\left(C_{ji} - A_{jj}S_{ij}\right) + \sum_{i=1}^{N}\sum_{\substack{j=1\\j\neq i}}^{N} \frac{S_{ij}}{R_{ij}}\left(S_{ij} - \frac{1}{T}A_{ij}\right) \tag{22}$$

Add up the two terms above, scale by $\bar{r}/2$, and simplify, we get the simplified expression for the second part of $\hat{\rho}$:

$$\begin{aligned}\frac{\bar{r}}{2}\sum_{i=1}^{N}\sum_{\substack{j=1\\j\neq i}}^{N}\left\{M_{ij}\hat{\vartheta}_{ii,ij} + M_{ji}\hat{\vartheta}_{jj,ij}\right\} &= \frac{\bar{r}}{2T}\sum_{i=1}^{N}\sum_{j=1}^{N}\left\{2M_{ij}C_{ij} - A_{ii}M_{ij}S_{ij} - A_{jj}M_{ij}^{\intercal}S_{ij}\right\}\\ &\quad - \frac{\bar{r}}{T}\sum_{i=1}^{N}\left\{C_{ii} - A_{ii}S_{ii}\right\} + \bar{r}\sum_{i=1}^{N}\sum_{\substack{j=1\\j\neq i}}^{N}\frac{S_{ij}}{R_{ij}}\left(S_{ij} - \frac{1}{T}A_{ij}\right)\end{aligned} \tag{23}$$

$\hat{\gamma}$ The formula is straightforward:

$$\hat{\gamma} = \sum_{i=1}^{N}\sum_{j=1}^{N}(F_{ij} - S_{ij})^2 = \|F - S\| \tag{24}$$

$\hat{\kappa}$

$$\hat{\kappa} = \frac{\hat{\pi} - \hat{\rho}}{\hat{\gamma}} \tag{25}$$

$\hat{\delta}^*$ Finally,

$$\hat{\delta}^* = \max\left\{0, \min\left\{\frac{\hat{\kappa}}{T}, 1\right\}\right\} \tag{26}$$

Data Set	Descriptor 1	Description 2	Market 3	N 4	N/T 5	Time Period 6
1	SZBM100	100 (10×10) portfolios formed on size and BM	U.S.	100	0.833	July 1963-Dec. 2010
2	IND48	48 industry portfolios	U.S.	48	0.400	July 1963-Dec. 2010
3	SZBM100 + IND48	Combination of SZBM100 and IND48	U.S.	148	1.233	July 1963-Dec. 2010
4	Individuals	100 stocks from NYSE and AMEX	U.S.	100	0.833	July 1973-Dec. 2010

Table 1: Data Description from Goto and Xu (2015), and for Tests Ending in 201012

Data Set	Descriptor 1	Description 2	Market 3	N 4	N/T 5	Time Period 6
1	SZBM100	100 (10×10) portfolios formed on size and BM	U.S.	100	0.833	July 1963-Dec. 2019
2	IND48	48 industry portfolios	U.S.	48	0.400	July 1963-Dec. 2019
3	SZBM100 + IND48	Combination of SZBM100 and IND48	U.S.	148	1.233	July 1963-Dec. 2019
4	Individuals	100 stocks from NYSE and AMEX	U.S.	100	0.833	July 1973-Dec. 2019

Table 2: Data Description for Tests Ending in 201912

		Out-of-Sample Analysis Period			Proposed Optimizer: $\hat{\Psi}_\rho$	
		Training Period	Testing Period	T_f	ρ	Sparsity
Data Set	Descriptor	1	2	3	4	5
1	SZBM100	July 1973-June 1983	July 1983-Dec. 2010	330	1.7	45.0%
2	IND48	July 1973-June 1983	July 1983-Dec. 2010	330	1.3	32.2%
3	SZBM100 + IND48	July 1973-June 1983	July 1983-Dec. 2010	330	1.9	44.0%
4	Individuals	Jan. 1983-Dec. 1992	Jan. 1993-Dec. 2010	216	5.9	32.4%

Table 3: Out-of-Sample Periods, Regularization Parameters, and Sparsity for Estimated $\hat{\Psi}_\rho$ from Goto and Xu (2015)

		Out-of-Sample Analysis Period			Proposed Optimizer: $\hat{\Psi}_\rho$	
		Training Period	Testing Period	T_f	ρ	Sparsity
Data Set	Descriptor	1	2	3	4	5
1	SZBM100	July 1973-June 1983	July 1983-Dec. 2010	330	1.7	75.6%
2	IND48	July 1973-June 1983	July 1983-Dec. 2010	330	1.3	54.0%
3	SZBM100 + IND48	July 1973-June 1983	July 1983-Dec. 2010	330	1.9	79.7%
4	Individuals	Jan. 1983-Dec. 1992	Jan. 1993-Dec. 2010	216	6.3	53.3%

Table 4: Out-of-Sample Periods, Regularization Parameters, and Sparsity for Estimated $\hat{\Psi}_\rho$ in Tests Ending in 201012

		Out-of-Sample Analysis Period			Proposed Optimizer: $\hat{\Psi}_\rho$	
		Training Period	Testing Period	T_f	ρ	Sparsity
Data Set	Descriptor	1	2	3	4	5
1	SZBM100	July 1973-June 1983	July 1983-Dec. 2019	438	1.7	75.7%
2	IND48	July 1973-June 1983	July 1983-Dec. 2019	438	1.3	54.6%
3	SZBM100 + IND48	July 1973-June 1983	July 1983-Dec. 2019	438	1.9	79.8%
4	Individuals	Jan. 1983-Dec. 1992	Jan. 1993-Dec. 2019	324	6.3	42.6%

Table 5: Out-of-Sample Periods, Regularization Parameters, and Sparsity for Estimated $\hat{\Psi}_\rho$ in Tests Ending in 201912

Data Set	Descriptor	*Panel A.* *Return Variance (%^{2})*						*Panel B.* $\sigma_{ALT} - \sigma$ *for* $ALT = \hat{\Psi}_\rho$				
		$\sigma^2_{\hat{\Psi}_\rho}$	$\sigma^2_{\hat{S}^{-1}}$	σ^2_{EW}	σ^2_{JM}	σ^2_{LW}	σ^2_{IND}	$\sigma_{\hat{S}^{-1}}$	σ_{EW}	σ_{JM}	σ_{LW}	σ_{IND}
1	SZBM100	13.30	65.30	25.43	58.57	21.98	NA	4.43***	1.40***	4.01*	1.04***	NA
2	IND48	12.45	17.59	22.88	16.22	13.15	NA	0.66***	1.25***	0.50*	0.10	NA
3	SZBM100 + IND48	10.70	$T < N$	24.12	16.43	12.74	NA	$T < N$	1.64***	0.78***	0.30	NA
4	Individuals	10.10	19.44	23.79	13.40	11.36	11.01	1.21***	1.72***	0.45***	0.19***	0.14***

Table 6: Out-of-Sample Portfolio Risk (monthly) from Goto and Xu (2015)

Data Set	Descriptor	*Panel A.* *Return Variance (%^{2})*						*Panel B.* $\sigma_{ALT} - \sigma$ *for* $ALT = \hat{\Psi}_\rho$				
		$\sigma^2_{\hat{\Psi}_\rho}$	$\sigma^2_{\hat{S}^{-1}}$	σ^2_{EW}	σ^2_{JM}	σ^2_{LW}	σ^2_{IND}	$\sigma_{\hat{S}^{-1}}$	σ_{EW}	σ_{JM}	σ_{LW}	σ_{IND}
1	SZBM100	12.97	55.93	25.65	18.07	17.95	NA	3.88***	1.46***	0.65***	0.63***	NA
2	IND48	12.37	17.58	22.73	13.43	13.26	NA	0.67***	1.25***	0.15	0.12***	NA
3	SZBM100 + IND48	10.60	$T < N$	24.21	12.82	11.75	NA	$T < N$	1.66***	0.33***	0.17*	NA
4	Individuals	10.52	16.26	19.56	10.52	11.56	25.27	0.79	1.18	0.00	0.16	1.78

Table 7: Out-of-Sample Portfolio Risk (monthly) in Tests Ending in 201012

		Panel A. Return Variance ($\%^2$)						*Panel B.* $\sigma_{ALT} - \sigma$ *for* $ALT = \hat{\Psi}_\rho$				
Data Set	Descriptor	$\sigma^2_{\hat{\Psi}_\rho}$	$\sigma^2_{\hat{S}^{-1}}$	σ^2_{EW}	σ^2_{JM}	σ^2_{LW}	σ^2_{IND}	$\sigma_{\hat{S}^{-1}}$	σ_{EW}	σ_{JM}	σ_{LW}	σ_{IND}
1	SZBM100	12.01	52.72	23.98	16.11	15.70	NA	3.79***	1.43***	0.55**	0.50***	NA
2	IND48	11.11	16.19	20.68	11.95	12.11	NA	0.69***	1.21***	0.12	0.15***	NA
3	SZBM100 + IND48	9.55	$T < N$	22.44	11.47	10.71	NA	$T < N$	1.65***	0.30***	0.18**	NA
4	Individuals	9.38	13.70	17.10	9.64	10.48	21.16	0.64	1.07	0.04	0.17	1.54

Table 8: Out-of-Sample Portfolio Risk (monthly) in Tests Ending in 201912

Data Set	Descriptor	*Panel A. Monthly SRs* $\hat{\Psi}_\rho$	$\hat{S}^{-1}$	$1/N$	JM	LW	IND	*Panel B. Differences in SRs between GMV-$\hat{\Psi}_\rho$ and the Other Five Portfolios* $\hat{S}^{-1}$	$1/N$	JM	LW	IND
1	SZBM100	0.260	0.112	0.113	0.065	0.215	NA	0.148**	0.147***	0.195***	0.045	NA
2	IND48	0.126	0.070	0.129	0.151	0.119	NA	0.056**	-0.003	-0.025	0.007	NA
3	SZBM100 + IND48	0.267	$T < N$	0.119	0.123	0.315	NA	$T < N$	0.148***	0.144***	-0.048	NA
4	Individuals	0.147	0.100	0.181	0.101	0.105	0.138	0.047***	-0.034***	0.046***	0.042***	0.010**

Table 9: Out-of-Sample Sharpe Ratio (monthly) from Goto and Xu (2015)

Data Set	Descriptor	*Panel A. Monthly SRs* $\hat{\Psi}_\rho$	$\hat{S}^{-1}$	$1/N$	JM	LW	IND	*Panel B. Differences in SRs between GMV-$\hat{\Psi}_\rho$ and the Other Five Portfolios* $\hat{S}^{-1}$	$1/N$	JM	LW	IND
1	SZBM100	0.281	0.158	0.128	0.154	0.240	NA	0.180***	0.184**	0.144*	0.058	NA
2	IND48	0.125	0.068	0.123	0.122	0.110	NA	0.074***	0.030	0.006	0.018	NA
3	SZBM100 + IND48	0.271	$T < N$	0.128	0.158	0.280	NA	$T < N$	0.183	0.122	-0.003	NA
4	Individuals	0.118	0.084	0.132	0.143	0.103	0.089	0.034	-0.014	-0.025	0.014	0.028

Table 10: Out-of-Sample Sharpe Ratio (monthly) in Tests Ending in 201012

		Panel A. Monthly SRs						*Panel B. Differences in SRs between GMV-$\hat{\Psi}_\rho$ and the Other Five Portfolios*				
Data Set	Descriptor	$\hat{\Psi}_\rho$	$\hat{S}^{-1}$	$1/N$	JM	LW	IND	$\hat{S}^{-1}$	$1/N$	JM	LW	IND
1	SZBM100	0.317	0.222	0.146	0.194	0.279	NA	0.143**	0.198***	0.135	0.049	NA
2	IND48	0.183	0.120	0.147	0.169	0.174	NA	0.079	0.061	0.017	0.012	NA
3	SZBM100 + IND48	0.323	$T < N$	0.147	0.200	0.344	NA	$T < N$	0.210	0.132	-0.016	NA
4	Individuals	0.190	0.147	0.169	0.203	0.170	0.129	0.043	0.022	-0.012	0.020	0.061

Table 11: Out-of-Sample Sharpe Ratio (monthly) in Tests Ending in 201912

Data Set	Descriptor	Portfolio	Minimum	1%	5%	95%	99%	Maximum	Herfindahl ($\%^2$)
1	SZBM100	GMV-$\hat{\Psi}_\rho$	-0.200	-0.129	-0.098	0.131	0.185	0.258	4,782
		GMV-$\hat{S}^{-1}$	-1.693	-0.820	-0.533	0.550	0.860	1.591	VLN
		GMV-JM	0.000	0.000	0.000	0.000	0.467	1.000	8,515
		GMV-LW	-0.457	-0.267	-0.155	0.192	0.341	0.676	12,354
2	IND48	GMV-$\hat{\Psi}_\rho$	-0.235	-0.151	-0.105	0.161	0.375	0.510	4,105
		GMV-$\hat{S}^{-1}$	-0.804	-0.383	-0.202	0.246	0.531	0.872	VLN
		GMV-JM	0.000	0.000	0.000	0.000	0.948	1.000	8,029
		GMV-LW	-0.200	-0.144	-0.103	0.175	0.488	0.655	5,182
3	SZBM100 + IND48	GMV-$\hat{\Psi}_\rho$	-0.158	-0.096	-0.064	0.085	0.136	0.252	3,169
		GMV-$\hat{S}^{-1}$				$T < N$			
		GMV-JM	0.000	0.000	0.000	0.000	0.167	1.000	7,891
		GMV-LW	-0.364	-0.172	-0.101	0.139	0.243	0.493	8,626
4	Individuals	GMV-$\hat{\Psi}_\rho$	-0.075	-0.039	-0.024	0.060	0.144	0.315	1,250
		GMV-$\hat{S}^{-1}$	-28.364	-0.448	-0.141	0.184	0.517	27.720	VLN
		GMV-JM	0.000	0.000	0.000	0.001	0.396	0.965	5,922
		GMV-LW	-0.166	-0.077	-0.047	0.080	0.203	0.469	2,561

Table 12: Distribution of Portfolio Weights from Goto and Xu (2015)

Data Set	Descriptor	Portfolio	Minimum	1%	5%	95%	99%	Maximum	Herfindahl ($\%^2$)
1	SZBM100	GMV-$\hat{\Psi}_\rho$	-0.187	-0.116	-0.087	0.113	0.164	0.216	3,821
		GMV-$\hat{S}^{-1}$	-1.680	-0.824	-0.546	0.577	0.853	1.545	115,519
		GMV-JM	0.000	0.000	0.000	0.080	0.240	0.462	1,892
		GMV-LW	-0.199	-0.127	-0.093	0.142	0.203	0.335	5,394
2	IND48	GMV-$\hat{\Psi}_\rho$	-0.192	-0.134	-0.098	0.148	0.349	0.459	3,561
		GMV-$\hat{S}^{-1}$	-0.726	-0.370	-0.204	0.249	0.528	0.841	10,768
		GMV-JM	0.000	0.000	0.000	0.109	0.518	0.709	3,036
		GMV-LW	-0.238	-0.149	-0.104	0.178	0.487	0.655	5,230
3	SZBM100 + IND48	GMV-$\hat{\Psi}_\rho$	-0.146	-0.088	-0.058	0.072	0.118	0.232	2,600
		GMV-$\hat{S}^{-1}$				$T < N$			
		GMV-JM	0.000	0.000	0.000	0.020	0.164	0.678	2,857
		GMV-LW	-0.169	-0.110	-0.078	0.112	0.177	0.404	5,103
4	Individuals	GMV-$\hat{\Psi}_\rho$	-0.146	-0.041	-0.023	0.053	0.095	0.342	733
		GMV-$\hat{S}^{-1}$	-83.327	-0.398	-0.119	0.165	0.501	76.240	265,575
		GMV-JM	0.000	0.000	0.000	0.060	0.175	0.668	1,593
		GMV-LW	-0.512	-0.094	-0.053	0.085	0.190	0.705	2,682
		GMV-IND	-0.564	-0.113	-0.054	0.081	0.143	0.639	2,222

Table 13: Distribution of Portfolio Weights in Tests Ending in 201012

Data Set	Descriptor	Portfolio	Minimum	1%	5%	95%	99%	Maximum	Herfindahl ($\%^2$)
1	SZBM100	GMV-$\hat{\Psi}_\rho$	-0.187	-0.117	-0.086	0.115	0.165	0.216	3,807
		GMV-$\hat{S}^{-1}$	-1.701	-0.820	-0.524	0.557	0.838	1.521	109,414
		GMV-JM	0.000	0.000	0.000	0.067	0.261	0.676	2,263
		GMV-LW	-0.199	-0.125	-0.092	0.144	0.212	0.337	5,413
2	IND48	GMV-$\hat{\Psi}_\rho$	-0.192	-0.129	-0.092	0.153	0.297	0.459	3,302
		GMV-$\hat{S}^{-1}$	-0.722	-0.349	-0.197	0.258	0.446	0.841	10,451
		GMV-JM	0.000	0.000	0.000	0.119	0.407	0.709	2,850
		GMV-LW	-0.240	-0.145	-0.099	0.186	0.370	0.654	4,911
3	SZBM100 + IND48	GMV-$\hat{\Psi}_\rho$	-0.146	-0.083	-0.056	0.071	0.112	0.232	2,413
		GMV-$\hat{S}^{-1}$				$T < N$			
		GMV-JM	0.000	0.000	0.000	0.017	0.184	0.682	2,716
		GMV-LW	-0.173	-0.111	-0.079	0.115	0.176	0.402	5,269
4	Individuals	GMV-$\hat{\Psi}_\rho$	-0.157	-0.048	-0.025	0.053	0.103	0.342	750
		GMV-$\hat{S}^{-1}$	-83.327	-0.297	-0.100	0.147	0.443	76.240	188,070
		GMV-JM	0.000	0.000	0.000	0.059	0.200	0.741	1,832
		GMV-LW	-0.988	-0.106	-0.055	0.085	0.215	1.421	3,055
		GMV-IND	-0.564	-0.108	-0.053	0.082	0.154	0.639	2,297

Table 14: Distribution of Portfolio Weights in Tests Ending in 201912

Data Set	Descriptor	$\hat{\Psi}_\rho$	$\hat{S}^{-1}$	$1/N$	JM	LW	IND
1	SZBM100	0.534	7.970	0.025	0.282	1.220	NA
2	IND48	0.298	0.778	0.033	0.062	0.327	NA
3	SZBM100 + IND48	0.527	$T < N$	0.028	0.066	1.140	NA
4	Individuals	0.160	5.220	0.033	0.109	0.319	0.281

Table 15: Portfolio Turnover (monthly) from Goto and Xu (2015)

Data Set	Descriptor	$\hat{\Psi}_\rho$	$\hat{S}^{-1}$	$1/N$	JM	LW	IND
1	SZBM100	0.454	7.293	0.024	0.120	0.582	NA
2	IND48	0.269	0.775	0.033	0.077	0.332	NA
3	SZBM100 + IND48	0.458	$T < N$	0.027	0.087	0.759	NA
4	Individuals	0.136	4.901	0.031	0.102	0.384	0.368

Table 16: Portfolio Turnover (monthly) for Tests Ending in 201012

Data Set	Descriptor	$\hat{\Psi}_\rho$	$\hat{S}^{-1}$	$1/N$	JM	LW	IND
1	SZBM100	0.434	7.282	0.023	0.110	0.566	NA
2	IND48	0.256	0.775	0.032	0.077	0.328	NA
3	SZBM100 + IND48	0.429	$T < N$	0.026	0.085	0.787	NA
4	Individuals	0.123	3.649	0.025	0.092	0.375	0.358

Table 17: Portfolio Turnover (monthly) for Tests Ending in 201912

Data Set	Descriptor	$\hat{\Psi}_\rho$	$\hat{S}^{-1}$	$1/N$	JM	LW	IND
1	SZBM100	4.17	-56.55	-0.94	-13.32	-1.83	NA
2	IND48	-0.19	-6.42	0.34	2.08	-0.73	NA
3	SZBM100 + IND48	4.11	$T < N$	-0.38	0.64	2.82	NA
4	Individuals	1.54	-31.99	3.27	-0.24	-0.75	0.43

Table 18: T_COST-ADJUSTED_CERs (annual %) from Goto and Xu (2015)

Data Set	Descriptor	$\hat{\Psi}_\rho$	$\hat{S}^{-1}$	$1/N$	JM	LW	IND
1	SZBM100	5.51	-46.36	-0.07	1.69	3.31	NA
2	IND48	-0.05	-6.44	0.03	0.89	-1.14	NA
3	SZBM100 + IND48	4.63	$T < N$	0.11	2.43	3.41	NA
4	Individuals	0.55	-30.37	0.92	1.69	-1.68	-4.48

Table 19: T_COST-ADJUSTED_CERs (annual %) in Tests Ending in 201012

Data Set	Descriptor	$\hat{\Psi}_\rho$	$\hat{S}^{-1}$	$1/N$	JM	LW	IND
1	SZBM100	6.97	-40.21	1.21	3.86	5.16	NA
2	IND48	2.45	-3.71	1.59	2.97	1.69	NA
3	SZBM100 + IND48	6.52	$T < N$	1.63	4.67	5.56	NA
4	Individuals	3.39	-19.57	3.08	4.01	1.12	-1.50

Table 20: T_COST-ADJUSTED_CERs (annual %) in Tests Ending in 201912

	$\hat{\Psi}_\rho$	$\hat{S}^{-1}$	$1/N$	JM	LW	IND
Panel A. Out-of-Sample Return Variance (monthly; %^{2})						
Mean	10.10	19.44	23.79	13.40	11.37	11.01
Standard deviation	2.23	6.09	1.19	4.69	2.67	2.56
Minimum	6.05	9.53	21.21	6.52	6.48	6.52
Maximum	14.92	58.09	26.91	31.51	17.17	15.87
No. of GMV-$\hat{\Psi}_\rho$ > ALT		0	0	5	5	5
No. of GMV-$\hat{\Psi}_\rho$ < ALT		100	100	95	95	95
Panel B. Herfindahl Index of Optimized Portfolio Weights (%^{2})						
Mean	1,250	VLN	100	5,922	2,561	1,909
Standard deviation	329	VLN	0	1,773	614	395
Minimum	807	VLN	100	3,247	1,625	1,356
Maximum	2,022	VLN	100	8,854	3,875	2,866
No. of GMV-$\hat{\Psi}_\rho$ > ALT		0	100	0	0	0
No. of GMV-$\hat{\Psi}_\rho$ < ALT		100	0	100	100	100
Panel C. Out-of-Sample Sharpe Ratio (monthly)						
Mean	0.147	0.100	0.181	0.101	0.115	0.138
Standard deviation	0.031	0.057	0.006	0.045	0.043	0.035
Minimum	0.065	-0.046	0.167	-0.012	0.019	0.036
Maximum	0.237	0.232	0.198	0.197	0.229	0.229
No. of GMV-$\hat{\Psi}_\rho$ > ALT		81	12	85	91	67
No. of GMV-$\hat{\Psi}_\rho$ < ALT		19	88	15	9	33

Panel D. Risk Reduction (monthly % $\sigma_{\hat{S}^{-1}} - \sigma_{\hat{\Psi}_\rho}$) Regressed on Sparsity

Intercept	t-Statistic	Sparsity	t-Statistic	R^2
-2.198	(-2.13)	0.112	(3.30)	10%

Panel E. Risk Reduction (monthly % $\sigma_{\hat{S}^{-1}} - \sigma_{\hat{\Psi}_\rho}$) across Sparsity Groups

Sparsity Group	Bottom 30%	Middle 40%	Top 30%
Risk Reduction	1.078	1.105	1.468

Table 21: Distribution of Realized Volatility, Sharpe Ratio and Weights, and Risk Reduction versus Sparsity for Individuals from Goto and Xu (2015)

	$\hat{\Psi}_\rho$	$\hat{S}^{-1}$	$1/N$	JM	LW	IND
Panel A. Out-of-Sample Return Variance (monthly; %^{2})						
Mean	10.52	16.26	19.56	10.52	11.56	25.27
Standard deviation	1.76	3.61	1.08	2.72	2.88	5.55
Minimum	6.64	9.16	16.70	5.52	6.27	16.96
Maximum	13.51	25.15	22.33	17.30	18.66	52.54
No. of GMV-$\hat{\Psi}_\rho$ > ALT		0	0	52	23	0
No. of GMV-$\hat{\Psi}_\rho$ < ALT		100	100	48	77	100
Panel B. Herfindahl Index of Optimized Portfolio Weights (%^{2})						
Mean	733	265,575	100	1,593	2,682	2,222
Standard deviation	142	198,564	0	702	520	313
Minimum	423	32,175	100	878	1,785	1,619
Maximum	1,093	1,033,762	100	3,249	4,014	3,276
No. of GMV-$\hat{\Psi}_\rho$ > ALT		0	100	0	0	0
No. of GMV-$\hat{\Psi}_\rho$ < ALT		100	0	100	100	100
Panel C. Out-of-Sample Sharpe Ratio (monthly)						
Mean	0.118	0.084	0.132	0.143	0.103	0.089
Standard deviation	0.030	0.059	0.009	0.031	0.047	0.032
Minimum	0.042	-0.033	0.105	0.073	0.004	0.009
Maximum	0.195	0.209	0.153	0.228	0.211	0.179
No. of GMV-$\hat{\Psi}_\rho$ > ALT		76	34	14	68	72
No. of GMV-$\hat{\Psi}_\rho$ < ALT		24	66	86	32	28

Panel D. Risk Reduction (monthly % $\sigma_{\hat{S}^{-1}} - \sigma_{\hat{\Psi}_\rho}$) Regressed on Sparsity

Intercept	t-Statistic	Sparsity	t-Statistic	R^2
0.379	(1.024)	0.007	(1.072)	1.16%

Panel E. Risk Reduction (monthly % $\sigma_{\hat{S}^{-1}} - \sigma_{\hat{\Psi}_\rho}$) across Sparsity Groups

Sparsity Group	Bottom 30%	Middle 40%	Top 30%
Risk Reduction	0.696	0.814	0.803

Table 22: Distribution of Realized Volatility, Sharpe Ratio and Weights, and Risk Reduction versus Sparsity for Individuals in Tests Ending in 201012

	$\hat{\Psi}_\rho$	$\hat{S}^{-1}$	$1/N$	JM	LW	IND
Panel A. Out-of-Sample Return Variance (monthly; %^{2})						
Mean	9.38	13.70	17.10	9.64	10.48	21.16
Standard deviation	1.37	2.77	0.87	2.31	2.33	4.45
Minimum	6.25	7.81	14.67	5.33	5.97	14.70
Maximum	12.00	19.78	19.40	14.36	15.25	41.31
No. of GMV-$\hat{\Psi}_\rho$ > ALT		0	0	38	20	0
No. of GMV-$\hat{\Psi}_\rho$ < ALT		100	100	62	80	100
Panel B. Herfindahl Index of Optimized Portfolio Weights (%^{2})						
Mean	750	188,070	100	1,832	3,055	2,297
Standard deviation	144	136,223	0	664	595	337
Minimum	418	31,573	100	1,038	2,070	1,666
Maximum	1,108	693,907	100	3,389	4,398	3,483
No. of GMV-$\hat{\Psi}_\rho$ > ALT		0	100	0	0	0
No. of GMV-$\hat{\Psi}_\rho$ < ALT		100	0	100	100	100
Panel C. Out-of-Sample Sharpe Ratio (monthly)						
Mean	0.190	0.147	0.169	0.203	0.170	0.129
Standard deviation	0.025	0.047	0.008	0.026	0.039	0.029
Minimum	0.135	0.039	0.146	0.144	0.091	0.015
Maximum	0.253	0.259	0.191	0.267	0.275	0.196
No. of GMV-$\hat{\Psi}_\rho$ > ALT		85	82	24	73	99
No. of GMV-$\hat{\Psi}_\rho$ < ALT		15	18	76	27	1

Panel D. Risk Reduction (monthly % $\sigma_{\hat{S}^{-1}} - \sigma_{\hat{\Psi}_\rho}$) Regressed on Sparsity

Intercept	t-Statistic	Sparsity	t-Statistic	R^2
0.498	(2.052)	0.003	(0.534)	0.29%

Panel E. Risk Reduction (monthly % $\sigma_{\hat{S}^{-1}} - \sigma_{\hat{\Psi}_\rho}$) across Sparsity Groups

Sparsity Group	Bottom 30%	Middle 40%	Top 30%
Risk Reduction	0.562	0.661	0.645

Table 23: Distribution of Realized Volatility, Sharpe Ratio and Weights, and Risk Reduction versus Sparsity for Individuals in Tests Ending in 201912

References

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

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

Goto, S. & Xu, Y. (2015). Improving mean variance optimization through sparse hedging restrictions. *Journal of Financial and Quantitative Analysis*, *50*(6), 1415–1441.

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

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

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