Optimal and naive diversification in an emerging market: Evidence from China's A-shares market

Abstract: This paper empirically investigates the out-of-sample performance of the 1/N naive rule and the Markowitz mean-variance strategies in the largest emerging market (i.e., China's A-shares market) and provides three new findings. First, we show that some mean-variance optimization strategies can outperform the 1/N rule in China's A-shares market, while minimum-variance strategies cannot. Using Certainty Equivalent Return (CER) instead of Sharpe ratios do not change our results qualitatively. Second, we find an obvious advantage of mean-variance optimization when N is large. Third, when transaction costs are taken into account, the profitability of the unconstrained mean-variance optimizations almost vanishes, while the profitability of the mean-variance optimizations with the short-sale constraint remains. Our results are robust to using a shorter estimation window of about 60 months. These results provide support for the use of optimal diversification strategies in emerging markets.

Keywords: portfolio choice; mean-variance optimization; 1/N naive diversification; China
JEL classification: C44, D81, G11, G12

1. Introduction

The relative performance of the mean-variance strategies and 1/N rule depends on the tradeoff between bias and variance. On the one hand, since Markowitz's (1952) seminal paper, mean-variance optimization has become a cornerstone of modern portfolio theory (Brandt, 2009). On the other hand, the simple naive 1/N diversification rule, which is arguably from the ancient Babylonian Talmud 1500 years ago and states that an investor should always equally divide her wealth across N assets under consideration, has been surprisingly widely used by individual investors (Benartzi and Thaler, 2001; Huberman and Jiang, 2006). The former is theory-based and usually asymptotically unbiased, but requires estimating model parameters from the data and thus has sizeable variance in small samples: the well-known parameter uncertainty or estimation error problem. The latter does not rely on any theory or data, is biased but has zero variance.

Although there is a growing body of research on the relative performance of the mean-variance strategies and 1/N rule[1], there are still many open questions. For example, many observers have documented evidence of the lukewarm performance of mean-variance strategies from Developed Markets (DMs)[2], but there is little empirical evidence as to whether mean-variance strategies outperform the 1/N rule in Emerging Markets (EMs), probably because the assets in EMs do not have a sufficiently long data history to form mean-variance strategies. This is an important question given the substantial mispricing but short history in EMs. A couple of papers conjecture that the magnitude of mispricing may promote the performance of the mean-variance

---

[1] For recent examples, pls refer to Wang et al. (2015), Ackermann et al. (2017), Bessler et al. (2017), Yan and Zhang (2017), and Platanakis et al. (2018).

[2] For instance, DeMiguel et al. (2009, p1915) find that "*none is consistently better than the 1/N rule in terms of Sharpe ratio…*", by evaluating 14 variants of mean-variance optimization across both simulated and real datasets from DMs.

strategies (Tu and Zhou, 2011; Yan and Zhang, 2017), while short history for assets prevent a precise estimation of the moments and hence deteriorate the performance of the mean-variance strategies (see, c.f., DeMiguel et al., 2009).

Furthermore, there is a debate about the role of N in the literature. For instance, Huberman and Jiang (2006) note that "*Participants tend to allocate their contributions evenly across the funds they use, with the tendency weakening with the number of funds used*", which is in stark contrast with DeMiguel et al. (2009, p1920) who find "*the results show that the naive 1/N strategy is more likely to outperform the strategies from the optimizing models when: (i) N is large*", while Yan and Zhang (2017) propose that "*As the number of assets N increases, there is a tradeoff between precisely estimating the covariance matrix and exploiting mispriced assets*".

Finally, although it has been argued that mean-variance strategies with the short-sale constraint perform better than the unconstrained mean-variance strategies and the minimum-variance strategies[3] perform better than the mean-variance strategies in DMs, it remains unclear whether these strategies perform better than the mean-variance strategies and the 1/N rule in EMs.

To answer these questions, we focus on the equity market from the largest EM in the world (i.e., China[4]), while the extant literature mostly relies on simulated data and/or data from DMs. China's A-shares market has the largest capitalization among EMs, and relatively long enough

[3] Minimum-variance strategies can be seen as a special case of the mean-variance strategies. For instance, DeMiguel et al. (2009) note "*Also, although this strategy does not fall into the general structure of mean-variance expected utility, its weights can be thought of as a limiting case of Equation (3), if a mean-variance investor either ignores expected returns or, equivalently, restricts expected returns so that they are identical across all assets*".

[4] Currently, China is one of the largest economies in the world and the largest EM. In 2014, the IMF ranks China as the largest economy by purchasing power parity, and the second largest by nominal GDP. China has the largest population, and is the engine of the world economy over the past three decades, with an average GDP growth rate above 8%. China is not only the largest exporter of goods, but also the second largest importer of goods.

history to form mean-variance strategies. We contribute to the prior literature by providing three new findings. First, we show that some mean-variance optimization strategies (i.e., "mv-c" and "dp-c") can outperform the 1/N rule in China's A-shares market, while minimum-variance strategies cannot (in consistency with Wang et al., 2015; Yan and Zhang, 2017). This is not surprising, since only minimum risk is of concern and the information about expected returns is sacrificed, although the minimum-variance type of strategies require estimating only the variance-covariance matrix but not the expected returns, and is less vulnerable to estimation risk than other sample-based mean-variance efficient portfolios (e.g., Green and Hollifield, 2002; Jagannathan and Ma, 2003; DeMiguel et al., 2009). Using Certainty Equivalent Return (CER) instead of Sharpe ratios do not change our results qualitatively.

Second, we find an obvious advantage of some mean-variance optimizations (i.e., "mv-c" and "dp-c") when N is large, which is different from the case of non-mispricing in DeMiguel et al. (2009) but in consistency with Huberman and Jiang (2006) and Yan and Zhang (2017). According to Yan and Zhang (2017), as N grows, there is a tradeoff between the estimation of the covariance matrix and the potential benefits of exploiting mispriced assets. While the former makes it difficult for Markowitz strategies to outperform the 1/N rule, the latter offers Markowitz strategies advantages over the $1/N$ rule in the presence of mispricing. Given a sufficiently large mispricing, the increase of N always leads to Markowitz strategies a big advantage over the $1/N$ rule.

Finally, when transaction costs are taken into account, the profitability of the unconstrained mean-variance optimizations almost vanishes, while the profitability of the mean-variance optimizations with the short-sale constraint remains. Our results are robust to using a shorter estimation window of about 60 months.

Our paper speaks to three strands of the literature. It relates to studies that investigate the underperformance of the Markowitz strategies relative to the 1/N rule, from the perspective of estimation errors (Brown, 1979; Jobson and Korkie, 1980; Michaud, 1989; Jorion, 1992). As estimation errors can lead to unstable portfolio weights away from optimal over time, it is not uncommon to argue that the Markowitz strategies can be outperformed by the 1/N rule. For instance, the contribution of DeMiguel et al. (2009) is "*to show that because the effect of estimation error on the weights is so large, even the models designed explicitly to reduce the effect of estimation error achieve only modest success*". We show that the underperformance of Markowitz strategies brought by the estimation errors might be offset by the benefits of exploring the considerable mispricing in EMs, which is less discussed in this area.

A small but growing literature has attempted to explain the dominant role of the 1/N rule, including Tu and Zhou (2011), Fletcher (2011), Kirby and Ostdiek (2012), Pflug et al. (2012), and Yan and Zhang (2017). Fletcher (2011) examines whether optimal diversification strategies outperform the 1/N strategy in U.K. stock returns. We add to Fletcher (2011) as the market efficiency (mispricing) may vary from country to country, and urge for further search of international evidence, especially from less efficient markets such as China. We rely on real data from China's A-shares market, while Tu and Zhou (2011), Yan and Zhang (2017) focus on mispricing with simulated data.

Finally, this study also connects to the strand of literature regarding active versus passive management in the markets with different extents of efficiency. We directly test the conjecture that (due to larger mispricing in EMs) mean-variance strategies should be more successful in EMs than in DMs and provide supportive empirical evidence, especially for two mean-variance optimizations with the short-sale constraint (i.e., "mv-c" and "dp-c"). This finding is in line with the literature such as Harvey (1995), Morck et al. (2000), Van der Hart et al. (2003), Griffin et al. (2010), and Dyck et al. (2013).

The rest of the paper unfolds as follows. In section 2, we describe our datasets. Section 3 describes the portfolio rules under our consideration. Section 4 presents empirical results while Section 5 concludes.

2. Datasets

In this section, we list the five datasets of monthly returns in China's A-shares market which we use to study the out-of-sample performance of the relative performance of the mean-variance strategies and 1/N rule. The five empirical datasets are listed in Table 1. Although China's stock trading started from 1990, the period of 1990 to 1995 is only the priming stage, the amount and scale of stocks are small, and their ups and downs are too noisy at this stage. We focus on the sample of excess returns of China's A-shares market from January 1996 to December 2016 in our empirical analysis (252 months). We use risk-free returns to compute excess returns on portfolios and obtain the risk-free returns from the CSMAR database, which uses treasury bonds as basis.

2.1 SMB, HML and MKT portfolios

The "MKT/SMB/HML" is also called "Fama-French benchmark factors". We obtain these three variables for China's A-shares market via the Resset database:

1) MKT: The market premium of China's A-shares market.
2) HML: The average return of the portfolio that consists of two value portfolios minus two growth portfolios (high minus low).
3) SMB: The average return of the portfolio that includes three small portfolios minus three big portfolios (small minus big).

2.2 Industry portfolios

The "Industry" dataset includes monthly excess returns on 15 industry portfolios in China's A-shares market. The 15 industries in this dataset are picked from the 19 classes which are classified according to the Industry Classification Standard produced by China Securities Regulatory Commission (CSRC). These 15 industries we consider are Agriculture, Forestry, Animal Husbandry and Fishery; Mining; Manufacturing; Electric Power, Heat, Gas and Water Production and Supply; Construction; Wholesale and Retail Trade; Transport, Storage and Postal Service; Accommodation and Catering; Information Transmission, Software and Information Technology Services; Financial Industry; Real Estate; Leasing and Commercial Service; Water Conservancy, Environment and Public Facility Management; Culture, Sports and Entertainment; Diversified Industry. We obtain this dataset from the Resset Database.

2.3 Size and book-to-market-sorted portfolios

The "P25" consists of the portfolios which are the intersections of 5 portfolios formed on size (market equity, ME) and 5 portfolios formed on the ratio of book equity divided by market equity (BE/ME). We obtain this dataset of monthly returns on 25 portfolios from the Resset Database.

Following Wang (2005), we exclude the five portfolios which contain the largest firms and form a dataset ("P20") that consists of 20 portfolios because the 25 portfolios are almost linear correlated with MKT, SMB and HML. We use this dataset to construct the "FF-1 factor" dataset, in which we add the MKT into the "P20" and let the investor take CAPM into account.

3. Methodology

This section describes the portfolio strategies and the measures of portfolio strategies in this paper. The portfolio strategies cover the 1/N rule and the major Markowitz strategies utilized in DeMiguel et al. (2009) and Tu and Zhou (2011). One of the most important characteristics of China's stock market is that it prohibits short-selling until very recently. Thus, short-sale constrained portfolios may be particularly suitable for China's stock market. Following DeMiguel et al. (2009), we report the results for the short-sale constrained portfolios, but neither for the multi-prior robust portfolio rule such as the one in Garlappi et al. (2007), nor for the portfolios shrinking the elements of the variance-covariance matrix such as Best and Grauer (1991, 1992), Chan, Karceski, and Lakonishok (1999) and Ledoit and Wolf (2004a, 2004b), as Jagannathan and Ma (2003) show that these portfolios are equivalent to imposing a short-sale constraint on the minimum-variance portfolio. Regarding Tu and Zhou (2011), we only consider the optimal combination of the 1/N rule and the sample tangency strategy and the optimal combination of the 1/N rule and Kan and Zhou (2007), as (i) they are the only two analytically tractable; (ii) they are the only two considered in latter literature such as Fletcher (2011) or Moorman (2014). We further exclude the VT and RRT strategies proposed by Kirby and Ostdiek (2012), as they require neither optimization nor covariance matrix inversion, two most notable characteristics of the Markowitz strategies. Different from DeMiguel et al. (2009), we consider an additional portfolio for the purely statistical approach relying on Bayesian diffuse-priors (Barry, 1974, Bawa et al., 1979) as well as its short-sale constrained variant. As a result, we have 13 representative portfolio rules in total.

3.1 Portfolio rules

In this subsection, we briefly discuss the portfolio rules considered in this study.

The Naive Portfolio ("$Naive, \mathbf{1}/N,$ or ew")

The *naive* portfolio is an even allocation of wealth across N risky assets

$$x_e = \mathbf{1}_N/N \tag{1}$$

It is equivalent to holding an equally weighted index consisting of the N risky assets. It involves no estimation risk since it does not require parameter estimation and model optimization. In theory, it normally deviates from the mean-variance optimal portfolio and thus has suboptimal performance. In practice, its performance depends on the trade-off between its sub-optimality and immunity to estimation risk.

The Tangency Portfolio ("mv")

Suppose that an investor can invest in N risky assets and one risk-free asset. Let R_t denote the N-vector of excess returns at t and assume that they are independent and identically distributed with mean equal to the N-dimensional vector μ and covariance Σ. Let x denote the vector of fractions of wealth allocated to risky assets. The remainder $(1 - \mathbf{1}'_N x)$ is invested in the risk-free asset where $\mathbf{1}$ is the $N \times 1$ vector of ones. The relative weights in the investor's portfolio with only risky assets are

$$w = \frac{x}{|\mathbf{1}'_N x|} \tag{2}$$

where we take the absolute value of the sum of portfolio weights to make sure that the relative weights w have the same signs as x in exceptional cases, in which, the sum of portfolio weights $\mathbf{1}'_N x$ is negative.

The investor's preference is represented by the general form of quadratic utility function:

$$U(x) = x'\mu - \frac{\gamma}{2} x'\Sigma x \tag{3}$$

where γ is the investor's risk aversion coefficient.

The optimal portfolio weights that maximize the above objective function are

$$x^* = \frac{\Sigma^{-1}\mu}{\gamma} \tag{4}$$

x^* is known as respectively the tangency portfolio and has the highest Sharpe ratio. In practice, we have to estimate μ and Σ. The simplest way to do that is using their sample analogs

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} R_t\,,\ \ \hat{\Sigma} = \frac{1}{T-N-2}\sum_{t=1}^{T}(R_t - \hat{\mu})(R_t - \hat{\mu})' \tag{5}$$

where $R_t, t = 1, \dots, T$ are historical excess returns.

Alternatively, we can estimate Σ as follows

$$\tilde{\Sigma} = \frac{1}{T}\sum_{t=1}^{T}(R_t - \hat{\mu})(R_t - \hat{\mu})' \ \text{ or } \ \bar{\Sigma} = \frac{1}{T-1}\sum_{t=1}^{T}(R_t - \hat{\mu})(R_t - \hat{\mu})' \tag{6}$$

They are known as the maximum likelihood estimator and the unbiased estimator of Σ. They differ from $\hat{\Sigma}$ by a scale factor and are asymptotically equivalent to $\hat{\Sigma}$.

Theoretically, the tangency portfolio is obtained by plugging the true mean and covariance matrix of excess returns into (4), and it should have the largest Sharpe ratio. In practice, we can estimate the tangency portfolio by plugging sample estimates $\hat{\mu}$ and $\hat{\Sigma}$ in (6) into (4). The resulting portfolio is the sample tangency portfolio ("*mv*"):

$$\hat{x}^* = \frac{\hat{\Sigma}^{-1}\hat{\mu}}{\gamma} \tag{7}$$

The portfolio obtained by substituting sample estimates of μ and Σ into (3) is known as the plug-in estimator. It has poor out-of-sample performance and extremely unstable weights over time. Many extensions of the plug-in rule are essentially different ways of refining the plug-in estimators.

The Minimum-Variance Portfolio ("$\boldsymbol{min}$")

Theoretically, the minimum-variance portfolio is

$$x^{min} = \frac{\Sigma^{-1}\mathbf{1}_N}{\mathbf{1}_N'\Sigma^{-1}\mathbf{1}_N} \tag{8}$$

It is the solution to the following minimization problem

$$Minimize\, x'\,\Sigma x,\ s.t.\ \mathbf{1}'_N x = 1 \tag{9}$$

Plugging the estimate of covariance $\hat{\Sigma}$ into x^{MVP}, we obtain its sample counterpart ("*min*")

$$\hat{x}^{min} = \frac{\hat{\Sigma}^{-1}\mathbf{1}_N}{\mathbf{1}'_N\hat{\Sigma}^{-1}\mathbf{1}_N} \tag{10}$$

This portfolio can be considered as a special case of mean-variance portfolios, as shown in DeMiguel et al. (2009). This portfolio differs from other Markowitz strategies in that it requires estimating only the covariance matrix of asset returns and is less vulnerable to estimation risk than other sample mean-variance efficient portfolios; however, it may well have a lower Sharpe ratio since it has the minimum risk.

Bayesian Diffuse-Prior Portfolio ("dp")

This approach assumes a diffuse prior about the unknown mean and the variance-covariance matrix, which results in an optimal portfolio that allocates less weight to risky assets. Barry (1974), Klein and Bawa (1976), and Brown (1979) show that if the prior is chosen to be diffuse, and the conditional likelihood is normal, then the predictive distribution is a student-t distribution. Hence, while still using the historical mean to estimate expected returns, this approach inflates the covariance matrix by a factor of (1+1/M).

Jorion's (1986) Bayes-Stein Shrinkage Estimators ("$Jorion, or\ bs$")

Jorion's (1986) Bayes-Stein estimators of mean and covariance are

$$\hat{\mu}^{bs} = \left(1-\hat{\delta}\right)\hat{\mu} + \hat{\delta}\hat{\mu}^{min}\mathbf{1}_N \tag{11}$$

$$\hat{\Sigma}^{bs} = \left(1+\frac{1}{T+\hat{\tau}}\right)\hat{\Sigma} + \frac{\hat{\tau}}{T(T+1+\hat{\tau})}\frac{\mathbf{1}_N\mathbf{1}'_N}{\mathbf{1}'_N\hat{\Sigma}^{-1}\mathbf{1}_N} \tag{12}$$

where

$$\hat{\tau} = \frac{T\hat{\delta}}{1-\hat{\delta}} \tag{13}$$

$$\hat{\delta} = \frac{N+2}{N+2+T(\hat{\mu}-\hat{\mu}^{min})'\hat{\Sigma}^{-1}(\hat{\mu}-\hat{\mu}^{min}\mathbf{1}_N)} \tag{14}$$

We see that the estimator $\hat{\mu}^{bs}$ shrinks the sample average return $\hat{\mu}$ to $\hat{\mu}^{min}$, the return on the sample global minimum-variance portfolio. The resulting portfolio ("*Jorion, or bs*") is obtained by plugging $\hat{\mu}^{bs}$ and $\hat{\Sigma}^{bs}$ into (10). Jorion (1986) reports that this portfolio outperforms the global minimum-variance portfolio, the Bayesian diffuse-prior portfolio and the plug-in estimators in terms of expected utility loss.

Kan and Zhou(2007)'s Three-Fund Rule ("Kan-Zhou, or kz3")

If estimation errors in two risky portfolios are not perfectly correlated, the estimation risk can be reduced by combining them. Kan and Zhou (2007) use the sample global minimum-variance to diversify the estimation risk of the sample tangency portfolio. Their optimal portfolio can be expressed as

$$x^{Kan-Zhou} = \frac{(T-N-1)(T-N-4)}{\gamma T(T-2)}\left[\eta\hat{\Sigma}^{-1}\hat{\mu} + (1-\eta)\mu^{min}\hat{\Sigma}^{-1}\mathbf{1}_N\right] \tag{15}$$

where

$$\eta = \frac{\psi^2}{\psi^2+\frac{N}{T}} \tag{16}$$

$$\psi^2 = (\mu-\mu^{min})'\Sigma^{-1}(\mu-\mu^{min}) \tag{17}$$

μ^{min} is the expected excess return on the global minimum-variance portfolio.

This rule implies that investors just need to allocate their wealth into three funds: the sample global minimum-variance portfolio, the sample tangency portfolio and the risk-free asset. For this reason, Kan and Zhou (2007) name their portfolio the three-fund rule. According to this rule, the higher N/T is, the larger the fraction of wealth invested in the global minimum-variance portfolio will be. The reason is as the number of assets increases relative to the sample length, the tangency portfolio will become more difficult to estimate. Consequently, the optimal portfolio will have a greater reliance on the global minimum-variance portfolio.

In theory, this portfolio is superior to either the sample tangency portfolio or the sample global minimum-variance portfolio. However, μ and Σ are unknown and thus $x^{Kan-Zhou}$ in (15) have to be estimated using sample data. Following Kan and Zhou (2007), we calculate $\hat{x}^{Kan-Zhou}$ ("*Kan-Zhou* or kz3") using the following estimates of η and ψ

$$\hat{\eta} = \frac{\hat{\psi}^2}{\hat{\psi}^2 + \frac{N}{T}} \tag{18}$$

$$\hat{\psi}^2 = \frac{(T-N-1)\bar{\psi}^2 - (N-1)}{T} + \frac{2(\bar{\psi}^2)^{\frac{N-1}{2}}(1+\bar{\psi}^2)^{-\frac{T-2}{2}}}{TB_{\bar{\psi}^2/(1+\bar{\psi}^2)}((N-1)/2,(T-N+1)/2)} \tag{19}$$

$$\bar{\psi}^2 = (\hat{\mu} - \hat{\mu}^{min})'\hat{\Sigma}^{-1}(\hat{\mu} - \hat{\mu}^{min}) \tag{20}$$

where $B_x(a,b) = \int_0^x y^{a-1}(1-y)^{b-1}dy$ is the incomplete $Beta$ function.

Since $\hat{x}^{Kan-Zhou}$ is subject to estimation risk, there is no guarantee that it outperforms the sample tangency portfolio and the sample global minimum-variance portfolio.

Combination of $1/N$ with the Sample Tangency Portfolio ("$CML, or\ ew-mv$,")

DeMiguel et al. (2009) find that the seemingly naive equally weighted portfolio x_e in (3) shows very impressive performance relative to the sophisticated ones under their consideration. In view of this, Tu and Zhou (2011) propose using the $1/N$ portfolio x_e to hedge the estimation risk in the sample tangency portfolio. The resulting rule ("ew $-$ mv") is

$$\hat{x}^{CML} = \frac{\hat{\pi}_2}{\hat{\pi}_1 + \hat{\pi}_2} x_e + \frac{\hat{\pi}_1}{\hat{\pi}_1 + \hat{\pi}_2} \hat{x}^* \tag{21}$$

where

$$\hat{\pi}_1 = x_e^T \hat{\Sigma} x_e - \frac{2}{\gamma} x_e' \hat{\mu} + \frac{1}{\gamma^2} \hat{\theta}^2 \tag{22}$$

$$\hat{\pi}_2 = \frac{1}{\gamma^2}(c_1 - 1)\hat{\theta}^2 + \frac{c_1}{\gamma^2}\frac{N}{T} \quad (23)$$

$$c_1 = \frac{(T-2)(T-N-2)}{(T-N-1)(T-N-4)} \quad (24)$$

$$\hat{\theta}^2 = \frac{(T-N-2)\bar{\theta}^2 - N}{T} + \frac{2(\bar{\theta}^2)^{\frac{N}{2}}(1+\bar{\theta}^2)^{-\frac{T-2}{2}}}{MB_{\bar{\theta}^2/(1+\bar{\theta}^2)}(N/2,(T-N)/2)} \quad (25)$$

$$\bar{\theta}^2 = \hat{\mu}'\hat{\Sigma}^{-1}\hat{\mu} \quad (26)$$

$x_e = \mathbf{1}_N/N$ is the weight vector of the $1/N$ portfolio, $\hat{x}^*$ refers to the sample tangency portfolio and $B_x(a,b) = \int_0^x y^{a-1}(1-y)^{b-1}dy$ is the incomplete $Beta$ function.

Combination of $1/N$ with the Three-Fund Portfolio ("CKZ, $or\ ew - kz3$")

Tu and Zhou (2011) also mix the $1/N$ rule x_e with the three-fund rule $x^{Kan-Zhou}$ in (15). Using sample analogs of μ and Σ, this mixture ("ew $-$ **kz3**") can be written as

$$\hat{x}^{CKZ} = \left(1 - \frac{\hat{\pi}_1 - \hat{\pi}_{13}}{\hat{\pi}_1 - 2\hat{\pi}_{13} + \hat{\pi}_3}\right)x_e + \frac{\hat{\pi}_1 - \hat{\pi}_{13}}{\hat{\pi}_1 - 2\hat{\pi}_{13} + \hat{\pi}_3}\hat{x}^{Kan-Zhou} \quad (27)$$

where

$$\hat{\pi}_3 = \frac{\hat{\theta}^2}{\gamma^2} - \frac{1}{\gamma^2 c_1}(\hat{\theta}^2 - \frac{N\hat{\eta}}{T}) \quad (28)$$

$$\hat{\pi}_{13} = \frac{\hat{\theta}^2}{\gamma^2} - \frac{1}{\gamma}w_e'\hat{\mu} + \frac{1}{\gamma c_1}(\hat{\eta}w_e'\hat{\mu} + (1-\hat{\eta})\hat{\mu}^{min}w_e'\mathbf{1}_N - \frac{1}{\gamma}[\hat{\eta}\hat{\mu}'\hat{\Sigma}^{-1}\hat{\mu} + (1-\hat{\eta})\hat{\mu}^{min}\hat{\mu}'\hat{\Sigma}^{-1}1_N]) \quad (29)$$

where $\hat{x}^{Kan-Zhou}$ is the three-fund rule estimated using (18), (19) and (20), and $\hat{\eta}$, $\hat{\pi}_1$, c_1 and $\hat{\theta}^2$ are given respectively by (18), (22), (24) and (25). Since shrinkage factors in (51) are estimated using sample analogs, this rule may fail to beat the $1/N$ rule or the three-fund rule due to estimation errors.

Short-sale constrained portfolios

In this paper, we consider four strategies that constrain short selling, i.e., the constrained mean-variance, constrained minimum-variance, constrained Bayesian diffuse-prior, and Bayes-Stein shrinkage portfolios. They are constructed by adding a non-negativity constraint on the weights of the assets in the optimizing process.

To constrain the portfolios, we impose the constraint: $x_i \geq 0, i = 1, \ldots, N$ in the optimization. When it comes to the classic mean-variance portfolio, we can yield the Lagrangian:

$$\mathcal{L} = x_t^T \mu_t - \frac{\gamma}{2} x_t^T \Sigma_t x_t + x_t^T \lambda_t, \tag{30}$$

where λ_t is the $\mathrm{N} \times 1$ vector of Lagrange multipliers for the short-sale constraints. Then, after we rearrange the mentioned equation, we can find that the difference between constrained weights and unconstrained weights is the adjusted mean vector which is shown as follows: $\tilde{\mu}_t = \mu_t + \lambda_t$. As the constraint for asset i binds, $\lambda_{t,i} > 0$, the expected excess returns is thus increased by the multiplier from $\mu_{t,i}$ to $\tilde{\mu}_{t,i} = \mu_{t,i} + \lambda_{t,i}$. This change of mean is equivalent to the "shrinkage" approach.

When it comes to the minimum-variance portfolio, Jagannathan and Ma (2003) studied the impact of imposing short-sale constraints on this portfolio and indicated that we can replace the covariance matrix using:

$$\hat{\Sigma}_{\mathrm{min-c}} = \hat{\Sigma}_{\mathrm{t}} - \lambda_t 1_N^T - 1_N \lambda_t^T, \tag{31}$$

where λ_t plays the same role as it does in mean-variance portfolios. This time the sample covariance is reduced by the multiplier.

Apart from the constrained mean-variance portfolios and constrained minimum-variance portfolios, we also consider the approach to constrain Bayesian diffuse-prior and Bayes-Stein portfolios, and we find that is similar to that of *mv* portfolio and *min* portfolio.

Table 2 provides an overview of all the portfolio rules described above. Among all rules, the $1/N$ rule is the only one free from parameter estimation. However, it may deviate substantially from the mean-variance optimal portfolio. All the other rules require estimating input parameters and thus are vulnerable to estimation risk.

3.2 Sharpe ratio and Certainty Equivalent Return (CER)

In this subsection we present the methods we use in this paper to evaluate the performance of the trading strategies above, namely the out-of-sample Sharpe ratio, and Certainty Equivalent Return (CER). We consider both the pre-cost and post-cost cases.

Sharpe ratio

Sharpe ratio is a measure for the excess return which is adjusted by the risk:

$$\widehat{\mathrm{SR}}_{\mathrm{i}} = \frac{\hat{\mu}_i}{\widehat{\sigma}_{\mathrm{i}}}. \tag{32}$$

For model i, we use the sample standard deviation $\widehat{\sigma}_{\mathrm{i}}$ to divide the sample mean of out-of-sample excess returns $\hat{\mu}_i$, to obtain the out-of-sample Sharpe ratio.

Certainty Equivalent Return (CER)

The second measure we use to evaluate the performance of the portfolios is the Certainty Equivalent Return (CER). This is defined as the fixed return that an investor would like to accept rather than investing in a specific portfolio of risky assets. Thus, CER represents the risk tolerance of an investor:

$$\widehat{\mathrm{CER}}_{\mathrm{i}} = \hat{\mu}_i - \frac{\gamma}{2}\hat{\sigma}_i^2. \tag{33}$$

where $\hat{\mu}_i$ is the sample mean of out-of-sample excess returns of model i and $\hat{\sigma}_i^2$ is the sample covariance of these returns. In addition, γ is the risk aversion, which is the same as we set before.

The definition of CER comes from the equation that computes the expected utility of an investor in Markowitz portfolios and we assume that all investors tend to maximize their utilities in all strategies. The value of CER represents the rate of return an investor is willing to accept instead of investing in the risky asset. If the risk is huge and leads to a negative CER, in practice investor will not invest in portfolios that have negative CER.

Transaction costs

In practice, the expenses always exist when transactions happen, no matter the investor purchases or sells the asset. The size of the transaction costs depends on the size of the commission fee and the bid/ask spread. If a portfolio has great change in its weights of assets every time it is rebalanced, the impact can be obvious on the excess return of the whole portfolio. As the return is reduced by trading volume, the cost of trading will inevitably influence the investor's decision making. In order to minimize transaction costs, investors may accept little deviations from the "optimal" weights they calculate to reduce the trading volume. We define turnover as the average amount of trading across N assets:

$$\text{Turnover} = \frac{1}{T-M}\sum_{t=1}^{T-M}\sum_{j=1}^{N}\left(\left|\hat{w}_{k,j,t+1} - \hat{w}_{k,j,t+}\right|\right), \tag{34}$$

where we use $\hat{w}_{k,j,t}$ to denote the weight of asset j at time t in model k; $\hat{w}_{k,j,t+}$ is the weight after generating the return and before rebalancing between t and $\mathrm{t}+1$; and $\hat{w}_{k,j,t+1}$ is the weight after rebalancing at the beginning of time $\mathrm{t}+1$. Turnover is computed based on the absolute difference between weights because the directions of the transaction could be different when rebalancing. In the case of naive portfolio, at the beginning of every period we rebalance the weight of every asset to 1/N, but $\hat{w}_{k,j,t+}$ may be different because prices of assets may change differently between t and

$t + 1$. Turnover can be seen as the average percentage of assets traded in each period. We consider the performance after transaction fee is paid. We set the rate of transaction cost equal to 48.7 basis points, which is the same as the rules produced by China Securities Regulatory Commission (CSRC).

Let $R_{k,p}$ be the vector of returns from model k which contains N assets before rebalancing, so $R_{k,p,t+1} = \sum_{j=1}^{N} R_{j,t+1}\widehat{w}_{k,j,t}$. After the portfolio is rebalanced at time $t + 1$, the transaction appears and the transaction fee is paid. The proportion of the absolute change of weight is $\left|\widehat{w}_{k,j,t+1} - \widehat{w}_{k,j,t+}\right|$. Assuming that s is the proportional transaction fee, the cost is $s \times \sum_{j=1}^{N}\left|\widehat{w}_{k,j,t+1} - \widehat{w}_{k,j,t+}\right|$. The return after the transaction cost is:

$$R^{*}_{k,p,t+1} = R_{k,p,t+1} - s \times \sum_{j=1}^{N}\left|\widehat{w}_{k,j,t+1} - \widehat{w}_{k,j,t+}\right|. \tag{35}$$

4. Empirical results

In this section we evaluate the performance of different portfolios in China's A-shares market by Sharpe ratio, and CER, respectively. We focus on the case in which the risk-aversion parameter $\gamma = 3$ and the estimation window of 120 months ($M = 120$) but also consider the alternatives such as $M = 60$, $\gamma = 1$.

4.1 Sharpe ratios

We present the Sharpe ratios of our baseline models in Tables 3 and 4, where we follow the literature in this area and use the estimation window of 120 months ($M = 120$) and 60 months ($M = 60$), respectively. Panel A and B report the pre-cost and post-cost cases, respectively. Several interesting observations can be made. First, in terms of the pre-cost out-of-sample Sharpe ratio, the naive strategy ("ew") outperforms the sample mean-variance portfolio ("mv") for first two datasets ("MKT/SMB/HML" and "Industry") but not the latter three datasets based on size and book-to-market-sorted portfolios, which is in contrast with DeMiguel et al. (2009). We speculate that it is due to the different number of investable assets (i.e., N) in each dataset. Specifically, there are only 3 (15) investable assets in the first (second) dataset, but at least 20 assets in the latter three assets.

Second, the above finding carries to several other mean-variance strategies such as the "kz3", "ew-mv", "ew-kz3", "dp", "bs", "mv-c" and "dp-c", while the minimum-variance type of strategies (i.e., "min", "ew-min" and "min-c") cannot outperform the 1/N rule (in consistency with Wang et al., 2015; Yan and Zhang, 2017). This is not surprising, since only minimum risk is of concern and the information about expected returns is sacrificed, although the minimum-variance type of strategies requires estimating only the variance-covariance matrix but not the expected returns.

Third, in the first two datasets (i.e., "MKT/SMB/HML" and "Industry"), we find that the three combinations of mean-variance portfolios (i.e., "kz3", "ew-mv", "ew-kz3") can improve the performance of mean-variance portfolio, which is in consistency with Kan and Zhou (2007), and Tu and Zhou (2011). However, we disagree with Kan and Zhou (2007), and Tu and Zhou (2011) as we find the extent of improvement is not large enough to outperform the naive portfolio in China’s A-shares market. In the latter three datasets based on size and book-to-market-sorted portfolios, however, the three combinations of mean-variance portfolios (i.e., "kz3", "ew-mv", "ew-kz3") cannot improve the mean-variance portfolio.

Fourth, constraints improve the out-of-sample Sharpe ratio, especially in the post-cost case. For instance, the first three short-sale constrained strategies (i.e., "mv-c", "min-c", "dp-c") almost always outperform their unconstrained counterparts (i.e., "mv", "min", "dp") in terms of the post-cost out-of-sample Sharpe ratio, although it is less the case in terms of the pre-cost out-of-sample Sharpe ratio.

Fifth, although the post-cost out-of-sample Sharpe ratio is to some extent disheartening relative to the pre-cost out-of-sample Sharpe ratio, two short-sale constrained strategies (i.e., "mv-c", "dp-c") can still outperform the 1/N rule across the latter three datasets based on size and book-to-market-sorted portfolios.

Finally, the above results are robust to using a shorter estimation window of about 60 months, although the magnitude of both the pre-cost and post-cost Sharpe ratios decline with the length of estimation window, which reflects the effects of the estimation risk and confirms the results in the prior literature.

4.2 Certainty Equivalent Return (CER)

We present the Certainty Equivalent Return (CER) of our baseline models using the estimation window of 120 months ($\mathrm{M} = 120$) in Table 5 and 6, where we follow Tu and Zhou (2011) and set the risk-aversion parameter of 3 ($\gamma = 3$) and 1($\gamma = 1$), respectively. Panel A and B report the pre-cost and post-cost cases, respectively. Several interesting observations can be made.

First, in Table 5 the sample mean-variance portfolio ("mv″) has the smallest CERs across all five datasets, no matter in the pre-cost or post-cost case. This means that all the variants of the sample mean-variance portfolio ("mv″) can improve its out-of-sample CER. In fact, in the pre-cost case, "mv-c" and "dp-c" have higher CERs than equally weighted portfolio in "P25", "P20", and "FF-1 Factor", while "bs-c" has higher CERs than 1/N strategy in "P20" and "FF-1 Factor". However, in the post-cost case, only "mv-c" and "dp-c" can outperform the 1/N strategy.

Second, we present the CER of our models using the alternative estimation window of 60 months ($\mathrm{M} = 60$) in Tables 7 and 8, where we follow Tu and Zhou (2011) and set the risk-aversion parameter of 3 ($\gamma = 3$) and 1($\gamma = 1$), respectively. We find that in general shorter estimation window may lead to lower CERs in Markowitz rules, and may cause extreme value in the combination of equally weighted portfolio and mean-variance portfolio since short length of estimation window may cause larger estimation error.

Third, constraints improve the out-of-sample CER, especially in the post-cost CER. For instance, the first three short-sale constrained strategies (i.e., "mv-c", "min-c", "dp-c") almost always outperform their unconstrained counterparts (i.e., "mv", "min", "dp") in terms of the post-cost out-of-sample CER, although it is less the case in terms of the pre-cost out-of-sample CER.

Finally, although the post-cost out-of-sample CER is to some extent disheartening relative to the pre-cost out-of-sample CER, two short-sale constrained strategies (i.e., "mv-c", "dp-c") can still outperform the 1/N rule across the latter three datasets based on size and book-to-market-sorted portfolios.

4.3 Statistical significance

Although we have found that the outperformance of the two models (mv-c and dp-c) relative to the naive rule in all 6 scenarios with a large N (larger than 20) is of large economic significance, the statistical significance of the difference between the Sharpe ratio (or the CER) of each strategy from that of the simple 1/N naive rule has not been tested yet. There is a possibility that the previous results are driven by the high estimation error, due to the relatively short history of China's A-shares market. To this end, we have used two ways to compute the P-values of the difference between the Sharpe ratio of each strategy and that of the simple 1/N naive rule and selectively report the results with an estimation window of 120 months. On the one hand, we follow the 16th footnote on page 1928 of DeMiguel et al. (2009a) and compute the p-value of the difference between the Sharpe ratio of each strategy and that of the simple 1/N naive rule, using the approach suggested by Jobson and Korkie (1981) after making the correction pointed out in Memmel (2003). On the other hand, we follow we follow the 6th footnote on page 1035 of DeMiguel et al. (2014) and compute the p-values for the Sharpe ratios using the bootstrapping methodology proposed in Ledoit and Wolf (2008) that is designed for the case in which portfolio returns have fat tails. To be specific, we follow Ledoit and Wolf (2008, Remark 3.2) to generate the resulting bootstrap p-values. We use the codes available at http://www.iew.uzh.ch/chairs/wolf.html, setting B = 1,000 bootstrap resamples and an expected block size b = 5.

The results are reported in Table 9 and Table 10, respectively. According to Table 9, the outperformance of the two models (mv-c and dp-c) relative to the naive rule in all 6 scenarios with a large N (larger than 20) is always statistically significant at conventional level (i.e., 5%), no matter whether we take transaction costs into account or not. Although it has been documented in the literature (Auer and Schuhmacher, 2013, and the papers thereafter) that it is more challenging to survive the bootstrapping methodology proposed in Ledoit and Wolf (2008), we find that the outperformance of the two models (mv-c and dp-c) relative to the naive rule in all 6 scenarios with a large N (larger than 20) is statistically significant at 10% level without transaction costs, and at 15% level with transaction costs. Our main results remain qualitatively the same when we try alternative performance measure such as CER, alternative estimation length such as 60 months, alternative black bootstrap size or alternative number of bootstrap resamples.

5. Concluding remarks

In this paper, we investigate the relative performance of the 1/N naive rule and the Markowitz mean-variance strategies in the largest emerging market (i.e., China's A-shares market) and provide several new findings.

First, we show that two mean-variance optimization strategies (i.e., "mv-c" and "dp-c") can outperform the 1/N rule in China's A-shares market, while minimum-variance strategies cannot (in consistency with Wang et al., 2015; Yan and Zhang, 2017). Using Certainty Equivalent Return (CER) instead of Sharpe ratios do not change our results qualitatively.

Second, we find an obvious advantage of the mean-variance optimization when N is large, which is different from the case of non-mispricing in DeMiguel et al. (2009) but in consistency with Huberman and Jiang (2006) and Yan and Zhang (2017).

Finally, when transaction costs are taken into account, the profitability of the unconstrained mean-variance optimizations almost vanishes, while the profitability of the mean-variance optimizations with the short-sale constraint remains. Our results are robust to using a shorter estimation window of about 60 months.

These findings have a couple of important implications. First, our results reaffirm the usefulness of some of the Markowitz mean-variance optimization strategies (i.e., "mv-c" and "dp-c") in practice, especially in less efficient EMs with large mispricing. Second, minimum-variance strategies may be attractive to the most conservative investors only, and less attractive to investors who pursue high expected returns. While the minimum-variance optimization may be more stable over time, it tends to result in more conservative portfolios than traditional mean-variance optimizations.

References

Ackermann, F., Pohl, W. and Schmedders, K. (2017). Optimal and naive diversification in currency markets. *Management Science*, 63, 3347-3360.

Auer, B. R., and Schuhmacher, F. (2013). Performance hypothesis testing with the Sharpe ratio: The case of hedge funds. *Finance Research Letters*, 10(4), 196-208.

Barry, C. B. (1974). Portfolio analysis under uncertain means, variances, and covariances. *Journal of Finance*, *29*(2), 515-522.

Bawa, V. S., S. Brown, and R. Klein. (1979). *Estimation Risk and Optimal Portfolio Choice*. Amsterdam, the Netherlands: North Holland.

Benartzi, S., and Thaler, R. (2001). Naive diversification strategies in defined contribution saving plan. *American Economic Review,* 91, 79–98.

Bessler, W., Opfer, H., and Wolff, D. (2017). Multi-asset portfolio optimization and out-of-sample performance: an evaluation of Black–Litterman, mean-variance, and naive diversification approaches. *European Journal of Finance, 23*(1), 1-30.

Best, M. J., and Grauer, R. R. (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. *Review of Financial Studies*, *4*(2), 315-342.

Best, M. J., and Grauer, R. R. (1992). Positively weighted minimum-variance portfolios and the structure of asset expected returns. *Journal of Financial and Quantitative Analysis*, *27*(4), 513-537.

Brandt, M.W. (2009). Portfolio choice problems. In: Aı¨t-Sahalia, Y., Hansen, L. (Eds.), *Handbook of Financial Econometrics*. Elsevier, North-Holland, Amsterdam, pp. 269–336.

Brown, S.J. (1979). The effect of estimation risk on capital market equilibrium, *Journal of Financial and Quantitative Analysis* 14: 215–20.

Chan, L. K. C., J. Karceski, and Lakonishok. J. (1999). On portfolio optimization: Forecasting covariances and choosing the risk model. *Review of Financial Studies,* 12:937–74.

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

DeMiguel, V., Nogales, F. J., & Uppal, R. (2014). Stock return serial dependence and out-of-sample portfolio performance. *Review of Financial Studies*, *27*(4), 1031-1073.

Fletcher, J. (2011). Do optimal diversification strategies outperform the 1/n strategy in UK stock returns? *International Review of Financial Analysis*, 20(5), 375-385.

Garlappi, L., Uppal, R., and Wang, T. (2007). Portfolio selection with parameter and model uncertainty: A multi-prior approach. *Review of Financial Studies*, *20*(1), 41-81.

Green, R. C., and Hollifield, B. (2002). When will mean-variance efficient portfolios be well diversified? *Journal of Finance*, 47, 1785-1809.

Huberman, G., and Jiang, W., (2006). Offering versus choice in 401(K) plans: Equity exposure and number of funds. *Journal of Finance,* 61, 763–801.

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

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

Jobson, J. D., and Korkie, B. M. (1981). Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance*, *36*(4), 889-908.

Jorion, P. (1986). Bayes-Stein estimation for portfolio analysis. *Journal of Financial and Quantitative Analysis*, *21*(3), 279-292.

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

Kirby, C., and Ostdiek, B. (2012). It's all in the timing: simple active portfolio strategies that outperform naive diversification. *Journal of Financial and Quantitative Analysis*, *47*(2), 437-467.

Klein, R. W., and Bawa, V. S. (1976). The effect of estimation risk on optimal portfolio choice. *Journal of Financial Economics*, *3*(3), 215-231.

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

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

Ledoit, O., and Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance*, *15*(5), 850-859.

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

Memmel, C. (2003). Performance hypothesis testing with the Sharpe ratio. *Finance Letters*, *1*(1):21–23.

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

Moorman, T. (2014), Optimal versus naive shrinkage: How inefficient are 50-50 combinations? *FMA Applied Finance Conference Working Paper*.

Pflug, G. C., A., Pichler, and Wozabal, D. (2012). The 1/N investment strategy is optimal under high model ambiguity. *Journal of Banking and Finance,* 36(2): 410-417.

Platanakis, E., Sutcliffe, C., and Urquhart. A. (2018). Optimal vs naive diversification in cryptocurrencies. *Economics Letters*, *171,* 93-96.

Tu, J., and Zhou, G. (2011). Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. *Journal of Financial Economics*, *99*(1), 204-215.

Wang, Z. (2005). A shrinkage approach to model uncertainty and asset allocation. *Review of Financial Studies*, 18(2), 673-705.

Wang, Y.D., C.F., Wu, and Li, Y., (2015). Hedging with futures: Does anything beat the naive hedging strategy? *Management Science*, 61. 2870 - 2889.

Yan, C., and Zhang, H. (2017). Mean-variance versus naive diversification: The role of mispricing. *Journal of International Financial Markets, Institutions and Money*, *48*, 61-81.

Table 1. An Overview of the datasets

This table lists all the datasets that we consider in this paper. The last column of the table gives the abbreviations used to refer to the datasets in the tables.

No.	Dataset	N	Source	Time period	Abbreviation
1	SMB, HML and China equity market portfolios	2+1	CSMAR	01/1996-12/2016	MKT/SMB/HML
2	Fifteen industry portfolios	15	Resset	01/1996-12/2016	Industry
3	Twenty size and book-to-market portfolios	20	Resset	01/1996-12/2016	P20
4	Twenty size and book-to-market portfolios and the MKT portfolio	20+1	Resset	01/1996-12/2016	FF+1 factor
5	Twenty-five Fama-French portfolios	25	Resset	01/1996-12/2016	P25

Table 2. An Overview of the Portfolio Rules

This table lists all the portfolio rules that we consider in this paper. The last column of the table gives the abbreviations used to refer to the strategy in the tables.

No.	Name of model	Abbreviation
Naïve portfolio model		
1	Equally weighted model (benchmark)	Naive, 1/N, ew
Mean-variance models and optimal combinations of portfolios		
2	Sample-based mean-variance (Markowitz model)	mv
3	Kan and Zhou's (2007) three-fund model	Kan-Zhou, kz3
4	Combination of 1/N and minimum-variance models	ew-min
5	Combination of 1/N and Markowitz models	CML, ew-mv
6	Combination of 1/N and Kan and Zhou's (2007) three-fund models	CKZ, ew-kz3
Bayesian approach to estimation error		
7	Bayesian diffuse-prior model	dp
8	Bayes-Stein Shrinkage model	Jorion, bs
Moment restrictions		
9	Minimum-variance model	min
Short-sale constrained models		
10	Sample-based mean-variance model with short-sale constraints	mv-c
11	Minimum-variance model with short-sale constraints	min-c
12	Bayesian diffuse-prior model with short-sale constraints	dp-c
13	Bayes-Stein Shrinkage model with short-sale constraints	bs-c

Table 3. Sharpe ratios of the portfolios with a window of 120 months

This table reports the monthly Sharpe ratios for the 13 portfolios formed at the estimation window of 120 months. Shaded area indicates a better performance than the 1/N rule.

Model\Dataset	MKT/SMB/HML N=3	Industry N=15	P20 N=20	FF-1 factor N=21	P25 N=25
Panel A: Without transaction costs					
ew	0.223	0.147	0.210	0.209	0.194
Mean-variance models and optimal combinations of portfolios					
mv	0.101	-0.034	0.320	0.305	0.269
kz3	0.133	0.063	0.265	0.278	0.239
ew-min	0.141	0.133	0.155	0.185	0.164
ew-mv	0.148	0.113	0.238	0.274	0.246
ew-kz3	0.147	0.066	0.264	0.225	0.207
Bayesian approach to estimation error					
dp	0.102	-0.034	0.320	0.305	0.269
bs	0.140	0.031	0.283	0.297	0.258
Moment restrictions					
min	0.128	0.088	0.096	0.158	0.122
Short-sale constrained models					
mv-c	0.122	0.106	0.266	0.244	0.233
min-c	0.128	0.139	0.167	0.165	0.130
dp-c	0.123	0.106	0.265	0.243	0.233
bs-c	0.140	0.118	0.229	0.199	0.167
Panel B: With transaction costs					
ew	0.217	0.143	0.207	0.205	0.190
Mean-variance models and optimal combinations of portfolios					
mv	0.091	-0.085	0.216	0.214	0.152
kz3	0.126	0.032	0.183	0.205	0.142
ew-min	0.133	0.118	0.120	0.143	0.124
ew-mv	0.139	0.087	-0.019	0.146	0.125
ew-kz3	0.110	0.032	0.177	0.160	0.098
Bayesian approach to estimation error					
dp	0.092	-0.085	0.216	0.214	0.152
bs	0.133	-0.004	0.196	0.218	0.156
Moment restrictions					
min	0.120	0.057	0.031	0.089	0.048
Short-sale constrained models					
mv-c	0.111	0.097	0.257	0.236	0.227
min-c	0.120	0.132	0.160	0.160	0.124
dp-c	0.112	0.097	0.257	0.236	0.227
bs-c	0.121	0.021	-0.007	-0.104	-0.193

Table 4. Sharpe ratios of the portfolios with a window of 60 months

This table reports the monthly Sharpe ratios for the 13 portfolios formed at the estimation window of 60 months. Shaded area indicates a better performance than the 1/N rule.

Model\Dataset	MKT/SMB/HML N=3	Industry N=15	P20 N=20	FF-1 factor N=21	P25 N=25
Panel A: Without transaction costs					
ew	0.152	0.066	0.110	0.109	0.098
Mean-variance models and optimal combinations of portfolios					
mv	-0.024	0.079	0.178	0.235	0.155
kz3	-0.046	0.119	0.159	0.234	0.160
ew-min	0.115	0.084	0.087	0.118	0.105
ew-mv	-0.117	0.071	0.027	0.095	0.130
ew-kz3	-0.090	0.081	0.177	0.247	0.151
Bayesian approach to estimation error					
dp	-0.024	0.079	0.178	0.235	0.155
bs	-0.036	0.110	0.165	0.240	0.168
Moment restrictions					
min	0.115	0.081	0.049	0.106	0.096
Short-sale constrained models					
mv-c	0.104	0.060	0.150	0.138	0.148
min-c	0.119	0.074	0.082	0.063	0.062
dp-c	0.105	0.059	0.150	0.137	0.148
bs-c	0.064	0.062	0.128	0.097	0.094
Panel B: With transaction costs					
ew	0.145	0.062	0.106	0.105	0.095
Mean-variance models and optimal combinations of portfolios					
mv	-0.041	-0.001	-0.008	0.048	-0.073
kz3	-0.069	0.055	-0.015	0.050	-0.059
ew-min	0.105	0.059	0.038	0.055	0.047
ew-mv	-0.134	0.049	-0.102	-0.082	-0.155
ew-kz3	-0.108	-0.190	-0.006	0.068	-0.066
Bayesian approach to estimation error					
dp	-0.041	-0.001	-0.008	0.048	-0.073
bs	-0.058	0.043	-0.005	0.063	-0.047
Moment restrictions					
min	0.106	0.023	-0.088	-0.039	-0.075
Short-sale constrained models					
mv-c	0.096	0.046	0.140	0.126	0.138
min-c	0.109	0.064	0.072	0.054	0.052
dp-c	0.096	0.045	0.140	0.126	0.138
bs-c	0.008	-0.293	-0.653	-0.824	-1.076

Table 5. CERs of the portfolios with a window of 120 months and a gamma of 3

This table reports the monthly Certainty Equivalent Return (CER) in percentage points for the 13 portfolios formed at the estimation window of 120 months and the gamma of 3. Shaded area indicates a better performance than the 1/N rule.

Model\Dataset	MKT/SMB/HML N=3	Industry N=15	P20 N=20	FF-1 factor N=21	P25 N=25
Panel A: Without transaction costs					
ew	0.60%	0.03%	0.44%	0.44%	0.32%
Mean-variance models and optimal combinations of portfolios					
mv	0.07%	-5.67%	-2.33%	-6.57%	-6.83%
kz3	0.28%	-0.84%	0.56%	0.15%	0.13%
ew-min	0.24%	-0.06%	-0.11%	0.37%	0.14%
ew-mv	0.33%	-0.36%	-11.98%	-2.02%	-0.17%
ew-kz3	0.31%	-0.74%	0.45%	-4.82%	-0.73%
Bayesian approach to estimation error					
dp	0.08%	-5.60%	-2.24%	-6.39%	-6.66%
bs	0.29%	-1.34%	0.63%	0.30%	0.08%
Moment restrictions					
min	0.21%	-0.61%	-0.82%	0.13%	-0.30%
Short-sale constrained models					
mv-c	0.25%	-0.51%	1.01%	0.83%	0.70%
min-c	0.21%	-0.03%	0.02%	0.23%	-0.18%
dp-c	0.25%	-0.51%	1.00%	0.82%	0.70%
bs-c	0.29%	-0.26%	0.59%	0.50%	0.11%
Panel B: With transaction costs					
ew	0.58%	0.00%	0.39%	0.40%	0.28%
Mean-variance models and optimal combinations of portfolios					
mv	0.01%	-6.63%	-5.07%	-9.83%	-10.66%
kz3	0.25%	-1.14%	-0.66%	-1.21%	-1.36%
ew-min	0.22%	-0.20%	-0.48%	-0.03%	-0.26%
ew-mv	0.30%	-0.61%	-17.60%	-5.20%	-2.38%
ew-kz3	0.20%	-1.05%	-0.91%	-6.79%	-2.54%
Bayesian approach to estimation error					
dp	0.02%	-6.55%	-4.97%	-9.62%	-10.46%
bs	0.27%	-1.71%	-0.76%	-1.23%	-1.63%
Moment restrictions					
min	0.19%	-0.89%	-1.52%	-0.51%	-1.04%
Short-sale constrained models					
mv-c	0.20%	-0.60%	0.90%	0.74%	0.63%
min-c	0.19%	-0.10%	-0.05%	0.19%	-0.23%
dp-c	0.20%	-0.60%	0.89%	0.74%	0.62%
bs-c	0.23%	-1.18%	-2.17%	-2.44%	-3.60%

Table 6. CERs of the portfolios with a window of 120 months and a gamma of 1

This table reports the monthly Certainty Equivalent Return (CER) in percentage points for the 13 portfolios formed at the estimation window of 120 months and the gamma of 1. Shaded area indicates a better performance than the 1/N rule.

Model\Dataset	MKT/SMB/HML N=3	Industry N=15	P20 N=20	FF-1 factor N=21	P25 N=25
Panel A: Without transaction costs					
ew	0.73%	0.95%	1.76%	1.72%	1.52%
Mean-variance models and optimal combinations of portfolios					
mv	0.42%	-2.31%	4.99%	4.60%	3.49%
kz3	0.39%	0.14%	2.86%	3.39%	2.50%
ew-min	0.29%	0.81%	1.10%	1.34%	1.14%
ew-mv	0.42%	0.63%	-2679.31%	-1339.49%	-668.38%
ew-kz3	-2.80%	-6547.43%	-1097.60%	-599.12%	-90.82%
Bayesian approach to estimation error					
dp	0.43%	-2.28%	4.98%	4.61%	3.50%
bs	0.38%	-0.23%	3.28%	3.82%	2.94%
Moment restrictions					
min	0.26%	0.38%	0.45%	1.05%	0.73%
Short-sale constrained models					
mv-c	0.43%	0.56%	2.50%	2.13%	2.02%
min-c	0.26%	0.87%	1.23%	1.10%	0.80%
dp-c	0.43%	0.56%	2.50%	2.13%	2.02%
bs-c	0.38%	0.68%	2.02%	1.47%	1.20%
Panel B: With transaction costs					
ew	0.71%	0.91%	1.71%	1.68%	1.48%
Mean-variance models and optimal combinations of portfolios					
mv	0.37%	-3.25%	2.19%	1.53%	-0.29%
kz3	0.37%	-0.17%	1.62%	2.07%	1.00%
ew-min	0.27%	0.66%	0.72%	0.92%	0.73%
ew-mv	0.39%	0.40%	-2642.77%	-2875.68%	-446.14%
ew-kz3	-5.37%	-55585.16%	-1034.18%	-854.27%	-42379.87%
Bayesian approach to estimation error					
dp	0.37%	-3.21%	2.20%	1.56%	-0.24%
bs	0.36%	-0.60%	1.87%	2.33%	1.21%
Moment restrictions					
min	0.24%	0.08%	-0.28%	0.40%	-0.03%
Short-sale constrained models					
mv-c	0.39%	0.47%	2.39%	2.05%	1.95%
min-c	0.24%	0.80%	1.16%	1.05%	0.73%
dp-c	0.39%	0.47%	2.39%	2.04%	1.94%
bs-c	0.32%	-0.26%	-0.78%	-1.49%	-2.53%

Table 7. CERs of the portfolios with a window of 60 months and a gamma of 3

This table reports the monthly Certainty Equivalent Return (CER) in percentage points for the 13 portfolios formed at the estimation window of 60 months and the gamma of 3. Shaded area indicates a better performance than the 1/N rule.

Model\Dataset	MKT/SMB/HML N=3	Industry N=15	P20 N=20	FF-1 factor N=21	P25 N=25
Panel A: Without transaction costs					
ew	0.44%	0.20%	0.60%	0.59%	0.48%
Mean-variance models and its extensions					
mv	-0.85%	-4.35%	-3.48%	-2.20%	-9.01%
kz3	-0.43%	0.64%	0.97%	2.74%	0.88%
ew-min	0.21%	0.36%	0.38%	0.67%	0.54%
ew-mv	-94.12%	-696.17%	-15.28%	-20.86%	-0.42%
ew-kz3	-12.03%	-69.12%	1.44%	3.04%	1.02%
Bayesian approach to estimation error					
dp	-0.82%	-4.16%	-3.25%	-1.92%	-8.58%
bs	-0.30%	0.46%	1.24%	2.87%	1.32%
Moment restrictions					
min	0.20%	0.32%	-0.04%	0.56%	0.46%
Short-sale constrained models					
mv-c	0.38%	0.10%	1.05%	0.88%	0.99%
min-c	0.21%	0.27%	0.33%	0.18%	0.16%
dp-c	0.38%	0.10%	1.05%	0.88%	0.99%
bs-c	0.16%	0.17%	0.80%	0.47%	0.44%
Panel B: With transaction costs					
ew	0.42%	0.16%	0.56%	0.55%	0.45%
Mean-variance models and its extensions					
mv	-1.04%	-7.26%	-12.66%	-12.57%	-22.12%
kz3	-0.57%	-0.34%	-3.21%	-1.79%	-4.47%
ew-min	0.19%	0.14%	-0.10%	0.08%	0.02%
ew-mv	-114.36%	-590.86%	-31.51%	-41.38%	-10.98%
ew-kz3	-19.66%	-151.09%	-2.67%	-1.02%	-3.27%
Bayesian approach to estimation error					
dp	-1.00%	-7.03%	-12.28%	-12.13%	-21.48%
bs	-0.41%	-0.63%	-2.43%	-1.01%	-3.15%
Moment restrictions					
min	0.19%	-0.21%	-1.48%	-0.81%	-1.15%
Short-sale constrained models					
mv-c	0.34%	-0.04%	0.94%	0.76%	0.89%
min-c	0.19%	0.19%	0.23%	0.10%	0.08%
dp-c	0.34%	-0.04%	0.94%	0.76%	0.89%
bs-c	-0.03%	-2.92%	-8.59%	-10.02%	-13.66%

Table 8. CERs of the portfolios with a window of 60 months and a gamma of 1

This table reports the monthly Certainty Equivalent Return (CER) in percentage points for the 13 portfolios formed at the estimation window of 60 months and the gamma of 1. Shaded area indicates a better performance than the 1/N rule.

Model\Dataset	MKT/SMB/HML N=3	Industry N=15	P20 N=20	FF-1 factor N=21	P25 N=25
Panel A: Without transaction costs					
ew	0.33%	-0.56%	-0.48%	-0.46%	-0.49%
Mean-variance models and its extensions					
mv	-2.02%	-19.13%	-28.08%	-32.36%	-45.89%
kz3	-0.76%	-1.80%	-4.66%	-3.30%	-5.35%
ew-min	0.17%	-0.34%	-0.58%	-0.18%	-0.28%
ew-mv	-0.91%	-2.41%	-74990.35%	-12047.28%	-1167.68%
ew-kz3	-9.10%	-5.25%	-3.37%	-2.02%	-3.19%
Bayesian approach to estimation error					
dp	-1.95%	-18.50%	-27.11%	-31.15%	-44.32%
bs	-0.53%	-2.26%	-3.38%	-1.93%	-3.19%
Moment restrictions					
min	0.17%	-0.53%	-1.17%	-0.35%	-0.45%
Short-sale constrained models					
mv-c	0.16%	-0.87%	-0.18%	-0.14%	-0.06%
min-c	0.17%	-0.40%	-0.59%	-0.54%	-0.57%
dp-c	0.16%	-0.86%	-0.18%	-0.14%	-0.06%
bs-c	0.05%	-0.57%	-0.30%	-0.36%	-0.37%
Panel B: With transaction costs					
ew	0.31%	-0.59%	-0.52%	-0.50%	-0.53%
Mean-variance models and its extensions					
mv	-2.22%	-21.72%	-37.20%	-42.98%	-57.66%
kz3	-0.90%	-2.71%	-8.91%	-7.87%	-10.50%
ew-min	0.15%	-0.55%	-1.06%	-0.75%	-0.80%
ew-mv	-1.30%	-3.32%	-83147.18%	-119166.30%	-4428.13%
ew-kz3	-9.99%	-6.32%	-7.07%	-5.81%	-7.29%
Bayesian approach to estimation error					
dp	-2.15%	-21.05%	-36.08%	-41.59%	-55.93%
bs	-0.65%	-3.28%	-7.05%	-5.80%	-7.50%
Moment restrictions					
min	0.15%	-1.06%	-2.59%	-1.70%	-2.03%
Short-sale constrained models					
mv-c	0.12%	-1.00%	-0.28%	-0.26%	-0.16%
min-c	0.16%	-0.48%	-0.69%	-0.61%	-0.65%
dp-c	0.12%	-1.00%	-0.28%	-0.26%	-0.16%
bs-c	-0.14%	-3.67%	-10.04%	-11.31%	-15.10%

Table 9. Comparing the Sharpe ratio of each strategy with that of the 1/N rule using the DeMiguel et al. (2009) methodology

This table reports the p-values of the difference between the Sharpe ratio of each strategy from that of the 1/N rule formed at the estimation window of 120 months, which is computed using the DeMiguel et al. (2009) methodology.

Model\Dataset	MKT/SMB/HML N=3	Industry N=15	P20 N=20	FF-1 factor N=21	P25 N=25
Panel A: Without transaction costs					
ew	NA	NA	NA	NA	NA
Mean-variance models and optimal combinations of portfolios					
mv	NA	NA	NA	NA	NA
kz3	0.051	0.033	0.091	0.114	0.144
ew-min	0.067	0.153	0.008	0.114	0.084
ew-mv	0.036	0.050	0.052	0.006	0.167
ew-kz3	0.004	0.013	0.071	0.007	0.030
Bayesian approach to estimation error					
dp	0.029	0.011	0.048	0.092	0.113
bs	0.055	0.014	0.063	0.086	0.110
Moment restrictions					
min	0.059	0.070	0.004	0.080	0.053
Short-sale constrained models					
mv-c	0.031	0.075	0.000	0.019	0.029
min-c	0.059	0.180	0.005	0.061	0.011
dp-c	0.031	0.075	0.000	0.020	0.029
bs-c	0.055	0.055	0.048	0.179	0.063
Panel B: With transaction costs					
ew	NA	NA	NA	NA	NA
Mean-variance models and optimal combinations of portfolios					
mv	NA	NA	NA	NA	NA
kz3	0.049	0.012	0.172	0.250	0.139
ew-min	0.065	0.089	0.001	0.013	0.009
ew-mv	0.034	0.044	0.007	0.003	0.049
ew-kz3	0.002	0.007	0.003	0.001	0.001
Bayesian approach to estimation error					
dp	0.026	0.003	0.227	0.233	0.177
bs	0.053	0.004	0.218	0.223	0.172
Moment restrictions					
min	0.057	0.031	0.000	0.007	0.004
Short-sale constrained models					
mv-c	0.026	0.064	0.001	0.027	0.035
min-c	0.057	0.151	0.003	0.057	0.009
dp-c	0.027	0.063	0.001	0.028	0.036
bs-c	0.039	0.000	0.000	0.000	0.000

Table 10. Comparing the Sharpe ratio of each strategy with that of the 1/N rule using the Ledoit and Wolf (2008) methodology

This table reports the p-values of the difference between the Sharpe ratio of each strategy from that of the 1/N rule formed at the estimation window of 120 months, which is computed using the Ledoit and Wolf (2008) methodology. To be specific, we follow Ledoit and Wolf (2008, Remark 3.2) to generate the resulting bootstrap p-values. We use the codes available at http://www.iew.uzh.ch/chairs/wolf.html, setting B = 1,000 bootstrap resamples and an expected block size b = 5.

Model\Dataset	MKT/SMB/HML N=3	Industry N=15	P20 N=20	FF-1 factor N=21	P25 N=25
Panel A: Without transaction costs					
ew	NA	NA	NA	NA	NA
Mean-variance models and optimal combinations of portfolios					
mv	NA	NA	NA	NA	NA
kz3	0.141	0.109	0.215	0.240	0.360
ew-min	0.190	0.321	0.016	0.242	0.172
ew-mv	0.136	0.109	0.188	0.014	0.360
ew-kz3	0.095	0.175	0.249	0.046	0.095
Bayesian approach to estimation error					
dp	0.082	0.042	0.142	0.230	0.313
bs	0.153	0.070	0.145	0.184	0.291
Moment restrictions					
min	0.160	0.155	0.008	0.161	0.105
Short-sale constrained models					
mv-c	0.110	0.201	0.004	0.067	0.086
min-c	0.172	0.381	0.015	0.148	0.044
dp-c	0.113	0.189	0.005	0.067	0.087
bs-c	0.146	0.172	0.133	0.366	0.163
Panel B: With transaction costs					
ew	NA	NA	NA	NA	NA
Mean-variance models and optimal combinations of portfolios					
mv	NA	NA	NA	NA	NA
kz3	0.145	0.056	0.367	0.499	0.353
ew-min	0.176	0.193	0.001	0.035	0.015
ew-mv	0.124	0.027	0.049	0.008	0.224
ew-kz3	0.028	0.040	0.175	0.014	0.049
Bayesian approach to estimation error					
dp	0.077	0.014	0.464	0.468	0.403
bs	0.150	0.029	0.443	0.447	0.384
Moment restrictions					
min	0.173	0.085	0.001	0.009	0.011
Short-sale constrained models					
mv-c	0.102	0.161	0.005	0.080	0.103
min-c	0.164	0.328	0.011	0.147	0.045
dp-c	0.108	0.176	0.003	0.078	0.102
bs-c	0.109	0.001	0.001	0.001	0.001