Winning Probability Weighted Combined Portfolio

Zhenzhen Huang∗ Pengyu Wei† Chengguo Weng‡ Tony Wirjanto§

August 18, 2023

Abstract

This paper introduces a novel framework for constructing combined portfolios using a winning probability weighted approach. It offers a viable method to integrate machine learning to incorporate external factors into the combination coefficient. Extensive empirical studies demonstrate the superiority of the proposed approach over existing analytical methods in terms of certainty equivalent return and out-of-sample Sharpe ratio across different scenarios.

Keywords: Portfolio optimization, estimation risk, winning probability, predictive models, out-of-sample performance

1 Introduction

Traditional portfolio management aims to optimize investments by considering a trade-off between reward and risk. One prominent approach is the mean-variance framework introduced by Markowitz (1952), which has become a cornerstone of modern portfolio theory. While theoretically sound, many portfolio strategies rely on accurate knowledge of unknown parameters of the asset return distribution, presenting practical challenges on multiple fronts. A simple solution is to use sample counterparts of these parameters in portfolio strategies to obtain so-called plug-in portfolios. However, since estimators based on limited historical periods may not accurately represent future patterns, the naive plug-in method can lead to substantial fluctuations in portfolio weights over time, as noted by Green and Hollifield (1992) and others. This inevitably introduces estimation risk to the optimal portfolios.

∗Department of Statistics and Actuarial Science, University of Waterloo. E-mail: z453huan@uwaterloo.ca
†Division of Banking and Finance, Nanyang Business School, Nanyang Technological University. E-mail: pengyu.wei@ntu.edu.sg
‡Department of Statistics and Actuarial Science, University of Waterloo. E-mail: chengguo.weng@uwaterloo.ca
§Department of Statistics and Actuarial Science and the School of Accounting and Finance, University of Waterloo. E-mail: twirjanto@uwaterloo.ca
Previous research has made earnest efforts to mitigate estimation risk and enhance the out-of-sample performance of portfolios. Bayesian approaches, for instance, utilize predictive distributions of asset returns to construct more robust portfolios based on observed data, as exemplified by Klein and Bawa (1976), Jorion (1986), Bodnar et al. (2017), and Bauder et al. (2021), among others. Shrinkage-type estimators have also demonstrated superior performance compared to classical sample counterparts, as shown in Ledoit and Wolf (2004, 2017), DeMiguel et al. (2013), and Bodnar et al. (2018), to name just a few. Such a shrinkage effect can further arise from proposing constraints or regularization on portfolio weights or considering transaction costs, as demonstrated by, for instance, Jagannathan and Ma (2003), DeMiguel et al. (2009a), Fan et al. (2012), and Mei et al. (2023). Furthermore, higher-moment portfolio strategies have been explored as alternatives to mean-variance portfolios (e.g., Briec et al., 2013). Lassance (2022) reconcile mean-variance portfolio theory with non-Gaussian returns under a chosen higher-moment criterion, and Lassance and Vrins (2023) propose a target-distribution framework that improves the higher moments of mean-variance-efficient portfolios.

To directly account for estimation risk in the portfolio construction, Tu and Zhou (2011) introduce a combined portfolio strategy based on the expected out-of-sample performance optimization proposed by Kan and Zhou (2007). This strategy involves creating a weighted combination of two constituent portfolios. By carefully selecting the combination coefficient, the combined portfolio can leverage the best advantages from each constituent portfolio. The combined portfolio strategy has been explored from various perspectives. For instance, DeMiguel et al. (2013) employ a nonparametric bootstrap approach to determine the combination coefficient using different optimization criteria. Additionally, DeMiguel et al. (2015) investigate multi-period combined portfolios subject to quadratic transaction costs under parameter uncertainty. Regarding specific performance criteria, Yuan and Zhou (2021) and Lassance (2021) characterize the optimal combined portfolio using the expected out-of-sample Sharpe ratio criterion. Chakrabarti (2021) focuses on building a robust combined portfolio by minimizing regret under uncertainty. Furthermore, Kan et al. (2022) optimize the combined portfolio with estimation risk in scenarios where the risk-free asset is not included. Lassance et al. (2023) study an unconstrained strategy that combines two constituent portfolios without the convexity constraint.

The majority of literature on combined portfolios determines the combination coefficient analytically to achieve desirable expected out-of-sample performance, relying on a strong assumption of independent and multivariate normal distribution of asset returns. However, such assumptions often do not align with real financial data. In addition, due to the complexity of the explicit form of the expected out-of-sample performance measure for a combined portfolio, the analytic method is limited to specific pairs of constituent portfolios. While some studies like DeMiguel et al. (2013) and Chakrabarti (2021) relax the Gaussian distribution assumption of asset returns, they make other assumptions such as assuming asset returns are
independent and identically distributed (i.i.d.) or requiring the specification of an uncertainty set for unknown parameters like the mean and covariance of asset returns. Furthermore, despite a wealth of research demonstrating that financial market factors can significantly impact asset/portfolio value (e.g., Nti et al. (2020), Neely et al. (2014)), none of the existing literature incorporates exogenous factors like cross-sectional financial market factors to determine combination coefficients.

This paper proposes a novel winning probability weighted (WPW) framework to construct a combined portfolio, which involves a sophisticated portfolio and a 1/N portfolio as described in Tu and Zhou (2011). The 1/N rule allocates equal weight to each asset and involves neither optimization nor estimation, yet it can outperform most sophisticated portfolios from investment theory in out-of-sample performance (DeMiguel et al., 2009b). In contrast to previous studies that rely on stringent distribution assumptions for asset return vectors, our framework only requires each constituent portfolio to have weakly stationary out-of-sample portfolio returns. While stationarity is often assumed for asset returns in existing literature (Györfi et al., 2008), we empirically validate it for the out-of-sample portfolio returns using statistical hypothesis tests on multiple datasets. Furthermore, with the WPW framework, the combination coefficient, characterized as a winning probability, can be determined using various machine learning techniques. In the era of big data, there has been increasing focus on employing machine learning techniques for finance problems in the recent literature, as seen in works such as Krauss et al. (2017), Fischer and Krauss (2018), Huck (2019), Chen et al. (2022), and Flori and Regoli (2021), among others.

Our WPW framework consists of three steps. Firstly, we adjust the sophisticated portfolio to ensure that its out-of-sample returns have the same long-run variance as those of the 1/N portfolio, signifying that the associated constituent portfolios carry the same long-term risk (variance). Secondly, we determine the combination coefficient as the winning probability of the adjusted sophisticated portfolio outperforming the 1/N portfolio in terms of the out-of-sample returns. The winning probability represents the expectation of the oracle coefficient which allocates all wealth to the constituent portfolio with a higher out-of-sample return but is not practically implementable. To predict the winning probability for an implementable combined portfolio, various statistical predictive models and exogenous factors can be conveniently incorporated. Lastly, we further enhance the combined portfolio by scaling it when targeting the expected out-of-sample mean-variance utility optimization. The resulting combined portfolio, constructed after these three steps, is referred to as a WPW combined portfolio.

There are several advantages offered by the WPW framework. Firstly, by treating the prediction of the winning probability as a classification problem, it harnesses the capabilities of various machine learning methods, such as logistic regression and random forest, to leverage financial market data structures and movements in determining the combination coefficient. Secondly, the framework provides flexibility in incorporating exogenous factors, such as Fama-French factors, alongside historical asset returns, to determine
the combination coefficients. These exogenous factors may contain valuable information and can be included as covariates in the classification procedure to improve the prediction of the winning probability. Lastly, while this paper focuses on the combined portfolio setting, which involves the sophisticated portfolio and the 1/N portfolio as components, the framework can be extended to any pair of constituent portfolios without introducing additional computational complexity.

In the empirical study, we utilized logistic regression and random forests as predictive models to estimate the winning probability, employing technical and fundamental features as input factors. The technical features comprised historical out-of-sample returns of the constituent portfolios, while the fundamental features encompassed widely recognized financial market factors from the existing literature. To evaluate the out-of-sample performance of the WPW combined portfolio, we combined the 1/N portfolio with the plug-in version of each sophisticated portfolio: mean-variance portfolio, three-fund portfolio (Kan and Zhou, 2007), and Bayes-Stein shrinkage portfolio. The empirical study, conducted across various datasets, consistently demonstrated that our WPW combined portfolios outperformed both the analytical method-based portfolio (Tu and Zhou, 2011) and corresponding constituent portfolios in terms of both the certainty equivalent return (CER) and the out-of-sample Sharpe ratio. There were only a few exceptions where we observed slight underperformance. The superiority of the WPW portfolios can be attributed to their data-driven nature, as they directly extract information from historical realizations without imposing strong assumptions like the normal distribution assumption made in the analytical method. Furthermore, the WPW combined portfolios that incorporated both technical and fundamental features consistently outperformed those with only technical factors. This indicates that including fundamental factors as input features in the predictive models contributed to improving the out-of-sample performance.

The remainder of the paper is organized as follows. Section 2 provides necessary preliminary concepts, such as estimation risk, combined portfolios, and stationarity. Section 3 presents the details of the WPW combined portfolio. Section 4 introduces the prediction of the winning probability, including predictive models and feature exploration. Section 5 presents the empirical results based on various real datasets. Finally, Section 6 concludes the paper.

2 Preliminaries

2.1 Estimation Risk

Consider a portfolio construction problem of allocating wealth across \( N \) risky assets and a risk-free asset. The rates of return for the risky assets and the risk-free asset are denoted as \( \mathbf{R} = [R_1, \ldots, R_N]^\top \) and \( R_f \), respectively. The excess return vector on the \( N \) risky assets over the risk-free asset is given by \( \mathbf{r} = \mathbf{R} - R_f \mathbf{1}_N \), where \( \mathbf{1}_N \) represents the \( N \)-dimensional vector of ones. Let \( \mathbf{\mu} \) and \( \mathbf{\Sigma} \) be the \( N \)-dimensional mean vector
and $N \times N$ covariance matrix of the full rank of the excess return vector $r$. In the traditional Markowitz’s portfolio theory (Markowitz, 1952), the investor seeks to find the optimal portfolio strategy $w$ revealing a tradeoff between a portfolio mean and a portfolio variance with a mean-variance objective:

$$U(w) = w^\top \mu - \frac{\gamma}{2} w^\top \Sigma w,$$

where $\gamma$ represents the risk aversion coefficient of the decision-maker. The theoretically optimal portfolio that maximizes the mean-variance objective is given by:

$$w_{mv} = \frac{1}{\gamma} \Sigma^{-1} \mu.$$

However, the theoretically optimal mean-variance portfolio $w_{mv}$ is not directly implementable in practice because it requires the knowledge of the unknown mean vector $\mu$ and the covariance matrix $\Sigma$ of asset returns. A naive solution is to replace these parameters with their sample counterparts, leading to so-called plug-in strategies. Specifically, given the historical excess returns of the $N$ risky assets from $T$ periods, denoted as $\{r_1, \ldots, r_T\}$, the plug-in strategies estimate $\mu$ and $\Sigma$, respectively, by

$$\hat{\mu} = \frac{1}{T} \sum_{t=1}^{T} r_t \quad \text{and} \quad \hat{\Sigma} = \frac{1}{T} \sum_{t=1}^{T} (r_t - \hat{\mu})(r_t - \hat{\mu})^\top.$$

This leads to the plug-in mean-variance portfolio weight vector:

$$\hat{w}_{mv} = \frac{1}{\gamma} \hat{\Sigma}^{-1} \hat{\mu}.$$

The plug-in portfolio rules can substantially deviate from their theoretically optimal counterparts and are notoriously known to have a deteriorated out-of-sample performance due to the estimation risk associated with the sample estimators of the return mean vector and the covariance matrix (see Michaud, 1989; Frankfurter et al., 1971; DeMiguel et al., 2009b; Kan and Zhou, 2007, for instance).

### 2.2 Combined Portfolios

To diversify the estimation risk and improve the out-of-sample performance, Tu and Zhou (2011) propose a weighted combination of the $1/N$ rule with a sophisticated portfolio. The weights of the combined portfolio take the form:

$$\tilde{w}_c = \delta \hat{w}_s + (1 - \delta) w_e,$$
where $\hat{w}_s$ denotes the weights of the plug-in sophisticated portfolio, such as the plug-in mean-variance portfolio (3), and $w_e$ represents the weights of the naive $1/N$ portfolio, i.e., $w_e = (1/N, \ldots, 1/N)^\top$. The combination coefficient $\delta \in [0, 1]$ is the weight assigned to the sophisticated portfolio in the combined portfolio. The expected out-of-sample portfolio mean and portfolio variance of the out-of-sample combined portfolio are, respectively, given by

$$
\mathbb{E}[\hat{w}_c^\top \mu] = \mathbb{E}[\delta \hat{w}_s^\top \mu + (1 - \delta) w_e^\top \mu],
$$

$$
\mathbb{E}[\hat{w}_c^\top \Sigma \hat{w}_c] = \mathbb{E}[\delta^2 \hat{w}_s^\top \Sigma \hat{w}_s + (1 - \delta)^2 w_e^\top \Sigma w_e + 2\delta(1 - \delta) \hat{w}_c^\top \Sigma \hat{w}_e].
$$

(5)

In general, the combined portfolio strategy can be applied to any pair of constituent portfolio strategies, and the key lies in the determination of the combination coefficient for desirable portfolio performance. A common assumption for any meaningful derivation of the optimal combination coefficient in the literature is the independent and multivariate normal assumption for asset returns. Under this assumption, Tu and Zhou (2011) achieve the optimal combined portfolios for several pairs of constituent portfolios by maximizing the expected out-of-sample mean-variance utility, where the combination coefficient $\delta$ is treated as a constant and thus can be moved out from the expectations displayed in (5).

2.3 Stationarity

Instead of making any assumptions on the individual asset returns, we work with a setting where the out-of-sample returns of each constituent portfolio satisfy a weakly stationary and ergodic process. The weakly stationary process is defined as follows:

**Definition 1.** (Hamilton, 2020) A time series $\mathbf{r} = \{r_t\}_{t=-\infty}^{\infty}$ is (weakly) stationary if it has time-invariant first and cross moments:

$$
\mathbb{E}[r_t] = \mu \text{ for all } t,
$$

$$
\mathbb{E}[(r_t - \mu)(r_{t-j} - \mu)] = \gamma_j \text{ for all } t \text{ and any integer } j,
$$

(6)

where $\mu$ and $\gamma_j$ are constants, and the latter is called the $j$th order autocovariance of $\mathbf{r}$.

The stationary process is ergodic when the autocovariances of the covariance stationary process satisfy the absolute summability condition, $\sum_{j=0}^{\infty} |\gamma_j| < \infty$. Let $\mathbf{r}_T = \{r_t\}_{t=1}^{T}$ be a stationary and ergodic time series with size $T$. The central limit theorem states that the sample mean $\bar{r} = \sum_{t=1}^{T} r_t/T$ satisfies (Anderson, 1994)

$$
\sqrt{T}(\bar{r} - \mu) \xrightarrow{d} N \left(0, \sum_{j=-\infty}^{\infty} \gamma_j \right),
$$

(7)
where \( N(\mu, \sigma^2) \) denotes a normal distribution with mean \( \mu \) and variance \( \sigma^2 \) by convention. The long-run variance of \( r_T \) is given by

\[
LV(r_T) = \lim_{T \to \infty} Var[\sqrt{T}(\bar{r} - \mu)] = \sum_{j=-\infty}^{\infty} \gamma_j = \gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j,
\]

where the last equation is due to a symmetry property: \( \gamma_{-j} = \gamma_j \). The long-run variance is an important concept for econometric models in the literature. It is commonly estimated consistently by a weighted sum of sample autocovariances, with the weight given by a kernel function and a bandwidth parameter to ensure positive semi-definiteness of the matrix. Early contributions to the robust estimate of the long-run variance include Newey and West (1987), Andrews and Monahan (1992), Müller (2007), among others.

### 3 Winning Probability Weighted Combined Portfolio

As mentioned earlier, the analytical method of determining the combination coefficient requires stringent assumptions in the literature (e.g., Tu and Zhou, 2011; Yuan and Zhou, 2021; Lassance, 2021, to name a few). Specifically, it requires the return vector of individual assets to follow a multivariate normal assumption and remains independent from one period to another. Such requirements often do not hold in practice when dealing with empirical data. Furthermore, the combination coefficient is universally taken as a constant and moved out from the expectations of expected out-of-sample portfolio mean and variance, despite the fact that it can be related to the sample mean and covariance matrix of asset returns.

While our proposed framework can be generally applied to any pair of constituent portfolios, we focus on the combination of the sophisticated portfolio \( \hat{w}_s \) and the 1/N portfolio \( w_e \) as formulated in (4). When we consider the combination coefficient at time \( t \) to construct the portfolio to deploy over the \((t + 1)\)th period, the sophisticated portfolio should be constructed from data over a certain number of past periods and therefore could be time-varying. Thus, we use \( \hat{w}_{s,t} \) to denote the portfolio weight vector constructed based on the historical information up to time \( t \) and to be implemented in the \((t + 1)\)th period. By definition, the 1/N portfolio weight \( w_e \) remains the same over time. Let \( r_{w_t} \) denote the out-of-sample return of a portfolio \( w \) over the \( t \)-th period for every \( t \) and \( r_w \) denote the time series of out-of-sample returns for the portfolio rule \( w \). We assume both the sophisticated and the 1/N portfolios have stationary out-of-sample returns over time. The stationarity assumption is verified for all the datasets in the empirical study by statistical hypothesis tests.

As previously introduced, the construction of our WPW combined portfolio consists of three steps, each of which is described in the next three subsections, respectively.
3.1 Long-run Variance Adjustment

To construct the WPW combined portfolio strategy, we first introduce the long-run variance adjustment parameter denoted by

\[ \xi = \sqrt{\frac{LV(r_{w_1})}{LV(r_{w_e})}}. \]  

(9)

The adjusted sophisticated portfolio is achieved by multiplying the long-run variance adjustment parameter with the sophisticated portfolio weight, namely, \( \tilde{\mathbf{u}}_s = \xi \hat{\mathbf{u}}_s \). Through this adjustment, we expect that the out-of-sample return series \( r_{\tilde{\mathbf{u}}_s} \) and \( r_{w_e} \) from the adjusted sophisticated portfolio and the 1/N portfolio have the same long-run variance. It is trivial to check

\[ LV(r_{\tilde{\mathbf{u}}_s}) = LV(r_{w_e}). \]

From the central limit theorem (7), the long-run variance plays a role similar to the variance. The same value in long-run variance for both \( r_{\tilde{\mathbf{u}}_s} \) and \( r_{w_e} \) indicate that the adjusted sophisticated portfolio and the 1/N portfolio have the same risk (measured by the variance of portfolio return) in the long run. Hence, heuristically it is desirable to assign a higher weight to the constituent portfolio when it has a higher chance to beat the other constituent in their return. This motivates us to adopt the winning probability as the combination coefficient in our combined portfolio.

As a relevant note, DeMiguel et al. (2013) also include an adjustment for the sophisticated portfolio in the combined portfolio to mitigate the bias of their target portfolio. They observe that introducing the adjustment generally improves the out-of-sample performance of their combined portfolios. In addition, our method does not consider normalizing the constituent portfolios to have the same conditional variance, because a dynamic estimation of the moments of portfolio returns is notoriously challenging.

The long-run variances of both time series of out-of-sample returns from two constituent portfolios are unknown and need to be estimated. We adopt a Newey-West estimator (Newey and West, 1987), popular and widely used for the estimation of the long-run variance of a stationary time series, to estimate \( LV(r_{w_e}) \) and \( LV(r_{w_h}) \). Given any stationary time series \( r_t = \{r_t\}_{t=1}^{T} \), the Newey-West estimator for the long-run variance of \( r_t \) is calculated by

\[ \widehat{LV}(r_t) = \gamma_0 + 2 \sum_{j=1}^{m_T} \gamma_j K(j), \]

(10)

where \( K(\cdot) \) is a symmetric kernel function with \( K(0) = 1 \), \( m_T \) is the bandwidth parameter, and \( \gamma_i, i = \)
0, 1, …, \( m_T \), represents the lag i sample autocovariance given by

\[
\hat{\gamma}_i = \frac{1}{T} \sum_{t=i+1}^{T} (r_t - \bar{r})(r_{t-i} - \bar{r}).
\]

In our empirical studies, we use the popular Bartlett kernel \( K(j) = 1 - j/(m_T + 1) \) for \( 1 \leq j \leq m_T \) and zero otherwise. The bandwidth parameter \( m_T \) should be sample size dependent, and we set it equal to the integer part of \( 4(T/100)^{2/9} \) as suggested by Newey and West (1987). The resulting estimator of the long-run variance is asymptotically consistent and positive semi-definite in finite samples.

Let \( \hat{\Lambda}V(r_w) \) and \( \bar{\Lambda}V(\hat{r}_w) \) be the Newey-West estimators for \( \Lambda V(r_w) \) and \( \Lambda V(\hat{r}_w) \), respectively. Therefore, we obtain the empirically long-run variance-adjusted sophisticated portfolio based on the historical information up to time \( t \) and to be implemented over the \((t + 1)\)th period as

\[
\hat{u}_{s,t} = \hat{\xi} \hat{w}_{s,t},
\]

where \( \hat{\xi} \) is the estimator of the variance-adjusting parameter as \( \hat{\xi} = \sqrt{\hat{\Lambda}V(r_w)/\bar{\Lambda}V(\hat{r}_w)} \).

### 3.2 Construction for the Combination Coefficient

The WPW combined portfolio is consequently constructed from the adjusted sophisticated portfolio \( \hat{u}_s \) and \( w_e \). Based on the historical data up to time \( t \), the weight vector of the combined portfolio for the \((t + 1)\)th period is given by:

\[
\hat{w}_{c,t} = \delta_t \hat{u}_{s,t} + (1 - \delta_t)w_e = \delta_t \hat{\xi} \hat{w}_{s,t} + (1 - \delta_t)w_e,
\]

where \( \hat{u}_{s,t} = \hat{\xi} \hat{w}_{s,t} \) denotes the empirically long-run variance-adjusted sophisticated portfolio (11) and \( \delta_t \in [0, 1] \) is the combination coefficient. The out-of-sample return of the combined portfolio over the \((t + 1)\)th period is given by

\[
r_{\hat{w}_{c,t+1}} = \delta_t r_{\hat{u}_{s,t+1}} + (1 - \delta_t)r_{w_{c,t+1}},
\]

where \( r_{\hat{u}_{s,t+1}} = \hat{u}_{s,t}^\top r_{t+1} + R_{f,t+1} \) and \( r_{w_{c,t+1}} = \hat{w}_{e}^\top r_{t+1} + R_{f,t+1} \).

The oracle optimal value \( \delta^*_t \in [0, 1] \) for the combination coefficient is obviously a bang-bang solution given by the indicator variable as

\[
\delta^*_t = \mathbb{I}\{r_{\hat{u}_{s,t+1}} \geq r_{w_{c,t+1}}\},
\]

where the oracle coefficient equals to one upon the event \( \{r_{\hat{u}_{s,t+1}} \geq r_{w_{c,t+1}}\} \) and zero over its opposite event \( \{r_{\hat{u}_{s,t+1}} < r_{w_{c,t+1}}\} \). While the oracle combination coefficient is optimal to reach the highest portfolio return in every scenario, it depends on the out-of-sample return variables of constituent portfolios over the
investment period \( t + 1 \). They are unknown until the end of the \((t + 1)\)th investment period. Hence, such an oracle optimal combination coefficient is not feasible.

We propose to set the combination coefficient as

\[
\delta_t = \mathbb{E}[\{r_{u, t+1} \geq r_{w, t+1}\}|\mathcal{F}_t] = \mathbb{P}(r_{u, t+1} \geq r_{w, t+1}|\mathcal{F}_t),
\]

where \( \mathcal{F}_t \) denotes a \( \sigma \)-field of events dated up to time \( t \). In our subsequent empirical study, \( \mathcal{F}_t \) contains the historical information about the returns of the constituent portfolios and some market factors, which we respectively call the technical features and fundamental features. This proposed combination coefficient does not come from any prioritization between the expected return and the expected risk exposure associated with the resulting portfolio. Instead, it comes as the best prediction, in terms of minimizing the mean square error, for the oracle coefficient \( \mathbb{I}\{r_{u, t+1} \geq r_{w, t+1}\} \). Hence, the combination coefficient is a winning probability that the adjusted sophisticated portfolio outperforms the \( 1/N \) portfolio. In addition, \((1 - \delta_t)\) can be viewed as a concept similar to the \( 1/N \) favorability index proposed by Guo et al. (2019), and \( \delta_t \) represents an indicator of the favorability of the sophisticated portfolio strategy. Specifically, when \((1 - \delta_t)\) is bigger (i.e. when \( \delta_t \) is smaller), there are more chances for the sophisticated portfolio to have a higher return than the \( 1/N \) portfolio over the \((t + 1)\)th period.

One of the major advantages of the specification (13) for the combination coefficient is its ability to accommodate different statistical learning models (classification or regression). It also allows for the incorporation of various technical and market features in a flexible manner. In the subsequent sections, we will present two predictive models for \( \delta_t \) and investigate various feature variables, which are used for our empirical study. Given the predictive model \( f_\delta(\cdot) \) and the feature variables \( x_t \), the estimated combination coefficient can be expressed as:

\[
\hat{\delta}_t = f_\delta(x_t).
\]

Therefore, we implement the following combined portfolio weights for the period \((t + 1)\) weighted by the winning probability:

\[
\hat{w}_{c, t} = \hat{\delta}_t \hat{w}_{u, t} + (1 - \hat{\delta}_t)w_e = \hat{\delta}_t \hat{\xi} \hat{w}_{s, t} + (1 - \hat{\delta}_t)w_e. \tag{14}
\]

As an alternative procedure, one may treat the two constituent portfolios as individual assets and view the determination of the combination coefficient as a portfolio selection problem for two assets, but the resulting solution needs the estimation of the expected returns and covariance matrix as the inputs. The estimation also needs to be accomplished over each period adaptively for a dynamically rebalancing portfolio, which is known to be notoriously fragile in its precision. In principle, it is also reasonable to consider a regression problem to predict the return means and variances of the constituent portfolios for the construction of the
combined portfolio. However, as the literature suggests, stock/portfolio return as a continuous response variable is sensitive to historical data and hard to predict. Generally speaking, the classification framework performs better than the regression model in predicting the financial market; see Leung et al. (2000). Different machine learning techniques applied to financial market prediction can be found in Henrique et al. (2019).

### 3.3 Certainty Equivalent Return Optimization

The certainty equivalent return (CER) of a risk portfolio strategy represents the risk-free rate of return that an investor is willing to accept instead of undertaking such a risky portfolio approach. It also reflects the level of expected out-of-sample utility for a mean-variance investor. A higher CER indicates a better portfolio strategy. When using CER as the out-of-sample performance measure, we further adjust the combined portfolio by a constant multiplier $a$ to enhance its performance. Specifically, we consider the combined portfolio weights achieved based on the historical information up to time $t$ and to be implemented over the $(t+1)$th period as follows:

$$\tilde{v}_{c,t} := a\hat{w}_{c,t},$$

where the CER scalar parameter $a$ is determined by maximizing the CER through the following optimization problem:

$$\max_a \left( \mu_{\tilde{v}_c} - \frac{\gamma}{2} \sigma^2_{\tilde{v}_c} \right) = \max_a \left( a\mu_{\hat{w}_c} - \frac{\gamma}{2} a^2 \sigma^2_{\hat{w}_c} \right).$$

Here, we assume the combined portfolio $\hat{w}_{c,t}$ generates a weakly stationary series of returns. In the above, $\mu_{\tilde{v}_c}$ and $\sigma^2_{\tilde{v}_c}$, and $\mu_{\hat{w}_c}$ and $\sigma^2_{\hat{w}_c}$ represent the mean and variance of the out-of-sample returns from the combined portfolio $\tilde{v}_c$ and $\hat{w}_c$, respectively.

The optimal scalar parameter to maximize the CER defined in (16) can be easily derived as follows:

$$a = \frac{\mu_{\hat{w}_c}}{\gamma \sigma^2_{\hat{w}_c}}.$$  

Denote $\hat{\mu}_{\hat{w}_c}$ and $\hat{\text{LV}}(r_{\hat{w}_c})$ as the sample mean and long-run variance of the historical realized out-of-sample return sequence of $\hat{w}_c$. Then, we define the implementable WPW combined portfolio after the certain equivalent return optimization for the period $(t+1)$ by

$$\hat{v}_{c,t} = \hat{a}\hat{w}_{c,t},$$

where the empirical CER scalar parameter is given by $\hat{a} = \hat{\mu}_{\hat{w}_c} / \left( \gamma \hat{\text{LV}}(r_{\hat{w}_c}) \right)$.

This proposal of adjusting the combined portfolio with the scalar parameter $a$ (15) and resulting in the
single-variable optimization problem (16) is similar in spirit to the idea proposed by Kan and Zhou (2007). While our approach involves shrinking the combined portfolio, their approach shrinks the tangent portfolio. Additionally, we avoid the stringent multivariate normal distribution assumption on asset returns required to obtain the analytic expected out-of-sample mean-variance utility in Kan and Zhou (2007). Instead, we obtain the mean and variance of the out-of-sample returns of the combined portfolio $\hat{w}_c$ based on the historical realized information.

### 3.4 General Training-Testing Procedure

In this section, we outline the general training-testing procedure for deploying and evaluating the winning probability weighted combined portfolio, utilizing a given historical dataset for asset returns. We divide the entire dataset into two subsets: the training set, with time labels 1, 2, …, $T_1$, and the testing set, with time labels $T_1 + 1, T_1 + 2, …, T$. The training set is utilized to estimate the long-run variance adjustment parameter $\hat{\xi}$, create the predictive model for the winning probability and establish the CER scale parameter $\hat{a}$. The testing set is used to evaluate the out-of-sample performance of the resulting combined portfolio.

For the training set, which ranges from time 1 to time $T_1$, we implement a rolling window approach with a window length of $M$. Specifically, we utilize the data from the most recent $M$ days up to time $t$ to estimate the mean vector and covariance matrix of the asset returns. The plug-in sophisticated portfolio rule $\hat{w}_{s,t}$ is then obtained by substituting the unknown parameters with the corresponding sample counterparts. The weights of the $1/N$ rule $w_e$ keep constants among all periods. Given the daily excess return $r_{t+1}$ and the riskfree return $R_{f,t+1}$, we obtain the corresponding realized out-of-sample portfolio returns by applying $\hat{w}_{s,t}$ and $w_e$ over the ($t + 1$)th period, expressed as follows:

$$r_{\hat{w}_{s,t+1}} = \hat{w}_{s,t}^T r_{t+1} + R_{f,t+1}, \quad r_{w_{e,t+1}} = w_e^T r_{t+1} + R_{f,t+1}, \quad t = M, ..., T_1 - 1.$$  

After obtaining the time series $r_{\hat{w}_s} = \{r_{\hat{w}_{s,t+1}}\}_{t=M}^{T_1-1}$ and $r_{w_e} = \{r_{w_{e,t+1}}\}_{t=M}^{T_1-1}$, we calculate the estimated long-run variance of these time series $\hat{LV}(r_{\hat{w}_s})$ and $\hat{LV}(r_{w_e})$ using (10), and we obtain the estimated long-run variance adjustment parameter $\hat{\xi}$ by taking the square root of the ratio of these two estimated long-run variances.

To facilitate the training of a predictive model for the winning probability, we create a training label for each rolling window as follows:

$$Y_t = I\{r_{\hat{w}_{s,t+1}} \geq r_{w_{e,t+1}}\}, \quad t = M, ..., T_1 - 1,$$

where $\hat{u}_s = \hat{\xi} \hat{w}_s$ represents the empirically long-run variance-adjusted sophisticated portfolio, as defined
in (11). As explained in detail in Section 4 later, we consider several selected feature variables $x_t$ along with the training label $Y_t$ to establish a predictive model $f_δ(·)$ for the winning probability. In this process, we use 80 percent of the training set as training samples to build candidate models, and the remaining 20 percent serves as validation samples for model selection. Denoting $\hat{δ}_t := f_δ(x_t)$ as the output from the resulting predictive model of the winning probability, we obtain the combined portfolio strategy weighted by the winning probability:

$$\hat{w}_{c,t} = \hat{δ}_t \hat{ξ} \hat{w}_{s,t} + (1 - \hat{δ}_t)w_e.$$

Finally, the sequence of out-of-sample returns when applying the combined portfolio strategy $\hat{w}_{c,t}$ for period $t + 1$ is $r_{\hat{w}_e} = \{r_{\hat{w}_{c,t+1}}\}_{t=1}^{T_1-1}$, where $r_{\hat{w}_{c,t+1}} = \hat{w}_{c,t}^T r_{t+1} + R_{f,t+1}$. We calculate the sample mean and the long-run variance of the sequence and apply (17) to obtain $\hat{a}$ as an estimate of the CER scale parameter.

The empirical WPW combined portfolio strategy to be applied for the period $(t + 1)$ is given by:

$$\hat{v}_{c,t} = \hat{a}[\hat{δ}_t (\hat{ξ} \hat{w}_{s,t}) + (1 - \hat{δ}_t)w_{e,t}].$$

(20)

To evaluate the out-of-sample performance of the empirical WPW combined portfolio over the testing set from $T_1 + 1$ to $T$, we keep the parameter estimates $\hat{a}$, $\hat{ξ}$, and the predictive model $\hat{δ}(·)$ obtained from the training set fixed. Using the same rolling window approach with a window length $M$, we establish the sophisticated portfolio and combination coefficient for $\hat{v}_{c,t}$, where $t = T_1 + M, ..., T - 1$. The corresponding realized out-of-sample returns are obtained as follows:

$$r_{\hat{v}_e,t+1} = \hat{v}_{c,t}^T r_{t+1} + R_{f,t+1}, \quad t = T_1 + M, ..., T - 1.$$

We then calculate the sample mean $\hat{μ}_{\hat{v}_e}$ and the sample variance $\hat{σ}_{\hat{v}_e}^2$ of the series $r_{\hat{v}_e} = r_{\hat{v}_{c,t+1}} t = T_1 + M^{T-1}$. The out-of-sample performance measures, the certainty equivalent return, and the out-of-sample Sharpe ratio are respectively given by:

$$\text{CER}_{\hat{v}_e} = \hat{μ}_{\hat{v}_e} - \frac{γ}{2} \hat{σ}_{\hat{v}_e}^2, \quad \text{and} \quad \text{SR}_{\hat{v}_e} = \frac{\hat{μ}_{\hat{v}_e}}{\hat{σ}_{\hat{v}_e}}.$$

(21)

The general training-testing procedure is summarized in Algorithm 1.

### 4 Prediction of Winning Probability

In this section, we introduce two specific predictive models for the winning probability and then explore the feature variables used in these predictive models, which include technical and fundamental features. Additionally, we introduce diagnostics for detecting multicollinearity among these feature variables to facilitate
Algorithm 1: General procedure for the winning probability weighted combined portfolio

Split the dataset into the training set and the testing set;
Create rolling windows for both the training set and the testing set with a fixed length $M$;

**For the training set**
1. generate realized out-of-sample returns of the sophisticated and 1/N portfolio for each rolling window;
2. estimate the long-run variance adjustment parameter $\hat{\xi}$;
3. create the training label (19) for each rolling window;
4. use 80 percent of the training set to establish candidate predictive models and the rest 20 percent to select the best predictive model $f_{δ}(·)$ for the winning probability;
5. estimate the CER scale parameter $\hat{a}$ based on the combined portfolio weighted by the winning probability;

**For the testing set**
1. freeze the parameters $\hat{\xi}$, $\hat{a}$ and the predictive model $f_{δ}(·)$ from the training set;
2. establish the empirical winning probability weighted combined portfolio (20) for each rolling window;
3. evaluate the portfolio performance (21);

the construction of feature variables.

4.1 Predictive Models for Winning Probability

Among all of the predictive models in the literature, we focus on two of the most widely explored predictive models in our empirical studies: the logistic regression, and the random forest.

4.1.1 Logistic Regression Model

Logistic regression is a widely used classification model due to its simplicity and interpretability; for a detailed description and application examples of the model, refer to Kleinbaum et al. (2002). In a logistic regression model, a sequence of features serves as input, and a logistic function is applied to produce a value between 0 and 1, representing the forecasted probability of an outcome belonging to a specific category:

$$p(x; \beta) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + ... + \beta_n x_n)}},$$

where $x = [x_1, ..., x_n]$ is the $n$-vector of feature variables, and $\beta = [\beta_0, ..., \beta_n]$ are the explanatory coefficients. Logistic regression allows us to model the relationship between the feature variables and the probability of the winning outcome, making it suitable for predicting the winning probability in the context of our WPW framework.

The logistic regression model is typically trained using the maximum likelihood estimation approach, with the log-likelihood given by:

$$l(\beta) = \sum_i y_i \log(p(x_i)) + (1 - y_i) \log(1 - p(x_i)).$$

Modern training for logistic regression involves combining variable selection and parameter estimation in
one step. This is achieved by adding a penalty term to the log-likelihood in the estimation procedure, which helps shrink the coefficients of less influential features toward zero and, consequently reduces the model complexity. In our empirical study, we adopt the elastic net regression procedure, which introduces the following penalty term to the log-likelihood in the estimation process:

\[ L = \lambda_1 ||\beta||_1 + \lambda_2 ||\beta||^2. \] (22)

The elastic net regression combines features of both lasso regression (when \( \lambda_1 = 1 \) and \( \lambda_2 = 0 \)) and ridge regression (when \( \lambda_1 = 0 \) and \( \lambda_2 = 1 \)). It shrinks some coefficients toward zero while setting some coefficients to exactly zero. For the choice of the regularization parameters \( \lambda_1 \) and \( \lambda_2 \), we initiate a parameter grid of the two tuning parameters \( \lambda_1 \) and \( \lambda_2 \) and then apply a randomized search method to determine the best set of regularization parameters using a ten-fold validation procedure.

### 4.1.2 Random Forests

The first random forest algorithm was proposed by Ho (1995) and later expanded by Breiman (2001). Random forest is an ensemble method in which the basic building block is a regression or classification decision tree. Each decision tree is built independently based on different bootstrap samples, repeatedly drawn from the entire training set. Moreover, only a subset of input features from the complete feature set is used as a split criterion at each node while constructing an individual regression tree. By combining the predictions from all decision trees, the random forest generates a more stable and robust prediction.

To apply the random forest method, three parameters need to be specified: the number of trees to grow, the maximum depth for each tree, and the number of randomly selected variables at each split while building individual trees. These parameters respectively control the size of the forest, the size of each individual tree, and a form of within-tree randomness. In our empirical study, we use randomized search cross-validation to determine these hyperparameters.

The random forest method has become an important benchmark model in machine learning for prediction; see Fischer and Krauss (2018) and Krauss et al. (2020), to name a few. Notably, in the context of portfolio construction, Krauss et al. (2017) concludes that the random forest outperforms other machine learning algorithms, including deep neural networks, gradient-boosted trees, and other ensemble methods, when used for prediction to build portfolios.

### 4.2 Feature Exploration

Our framework distinguishes itself from existing combined portfolio strategies due to its adaptability in incorporating valuable financial market information to predict the combination coefficient. As a portfolio
strategy comprises a collection of assets, any economic market factors that influence asset values are likely to impact the portfolio value as well. Extensive research has been conducted on predicting financial markets using either technical or fundamental analysis, as extensively reviewed by Nti et al. (2020). In this context, technical analysis involves studying past asset trends, while fundamental analysis utilizes cross-sectional financial market data. Previous research, such as the work by Neely et al. (2014), has demonstrated that combining information from both technical and fundamental variables significantly enhances forecasts of equity risk premium. Building on these findings, we explore the use of technical and fundamental features as input variables in our predictive models for the combination coefficient. By integrating such relevant financial market information, our framework aims to improve the accuracy of predicting the winning probability and, consequently, enhance the performance of the resulting WPW combined portfolio.

4.2.1 Technical Features

The combination coefficient (13) is inherently tied to the out-of-sample returns of the constituent portfolios. As a result, we incorporate the historical realized out-of-sample returns of these portfolios as technical features. To determine which historical out-of-sample portfolio returns should be included in the predictive model for the combination coefficient, we adopt a variable importance approach, which measures the relative predictive capacity or explanatory power of each feature. Our empirical findings consistently suggest that, regardless of the dataset or complex portfolio chosen, the most recent five trading days’ realized out-of-sample returns exhibit superior predictive capabilities and yield higher explanatory power. This observation aligns with the conclusions drawn by Krauss et al. (2017), although their focus was on forecasting individual assets to outperform the overall market.

Additionally, we explore the potential efficacy of including the historical returns of individual assets in the predictive model for the combination coefficient. However, this investigation involves a large number of feature variables due to the many individual assets in the constituent portfolios. Dealing with a high-dimensional regression and variable selection problem, we employ feature selection methods like the sure independence screening method proposed by (Fan and Lv, 2008). Empirical results reveal that incorporating the historical excess returns of individual assets does not lead to an improvement in the performance of the resulting combined portfolio.

Therefore, we choose to use the historical realized out-of-sample returns of the two constituent portfolios (i.e., the sophisticated portfolio and the 1/N portfolio) from the preceding five trading days as the technical features in our predictive model. Specifically, for any trading period \( t \), the technical features included in the predictive model are represented by:

\[
x_t = [r_{u,s,t-4}, \ldots, r_{u,s,t}, r_{w_e,t-4}, r_{w_e,t}] (23)
\]
where \( r_{u,s,t} \) and \( r_{w,e,t} \) denote the out-of-sample return of the long-run variance-adjusted sophisticated portfolio and the 1/N rule over the trading period \( t \), respectively.

### 4.2.2 Fundamental Features

The literature presents a plethora of fundamental features for asset pricing, as demonstrated in Feng et al. (2020) for example. Given the wide array of available fundamental features, investing all of them would be overly ambitious. Therefore, we narrow our focus to the market factors listed in Table 1.

<table>
<thead>
<tr>
<th>No.</th>
<th>Abbreviation</th>
<th>Market factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ERM</td>
<td>Fama-French five factors: excess return on the market (Fama and French, 2015)</td>
</tr>
<tr>
<td>2</td>
<td>SMB</td>
<td>Fama-French five factors: small minus big</td>
</tr>
<tr>
<td>3</td>
<td>HML</td>
<td>Fama-French five factors: high minus low</td>
</tr>
<tr>
<td>4</td>
<td>RMW</td>
<td>Fama-French five factors: robust minus weak</td>
</tr>
<tr>
<td>5</td>
<td>CMA</td>
<td>Fama-French five factors: conservative minus aggressive</td>
</tr>
<tr>
<td>6</td>
<td>LTR</td>
<td>Long term reversal</td>
</tr>
<tr>
<td>7</td>
<td>STR</td>
<td>Short term reversal (Gatev et al., 2006)</td>
</tr>
<tr>
<td>8</td>
<td>Mom</td>
<td>Momentum (Carhart, 1997)</td>
</tr>
<tr>
<td>9</td>
<td>Inidx</td>
<td>1/n favorable index (Guo et al., 2019)</td>
</tr>
</tbody>
</table>

The first eight market factors are obtained from Kenneth French data library, while the 1/N favorable index is calculated based on the definition in Guo et al. (2019).

The Fama-French five factors are widely acknowledged in the literature for explaining asset returns. They include the market factor, the return differential between small and large-cap stocks (small minus big), the return differential between high and low book-to-market stocks (high minus low), the disparity between robust and weak operating profitability companies (robust minus weak), and the contrast between conservatively and aggressively invested companies (conservative minus aggressive). Additionally, we include the long-term reversal, short-term reversal, and momentum factors as proposed by Fischer and Krauss (2018). Lastly, the 1/N favorable index is also included as it was introduced in Guo et al. (2019) and has shown the ability to outperform specific sophisticated portfolios when applied to the 1/N portfolio.

### 4.2.3 Multicollinearity Diagnostics

Given the multitude of technical and fundamental features discussed above, we utilize multicollinearity diagnostics to select the appropriate set of feature variables for the predictive model. We initially normalize all feature variables to mitigate the impact of varying scales. This normalization is achieved by removing the mean and scaling to unit variance. Multicollinearity refers to the presence of strong linear relationships among the feature variables. As a linear relation can encompass numerous feature variables, simple correlation coefficients are inadequate for identifying such a relationship. Chatterjee and Hadi (2006) delve into the effects of multicollinearity on statistical inference and review several criteria for detecting multicollinearity.
One comprehensive diagnostic of multicollinearity is to use the variance inflation factor (VIF):

\[ \text{VIF}_{x_i} = \frac{1}{1 - R^2_i}, \quad i = 1, ..., I, \]

where \( I \) represents the total number of feature variables, and \( R^2_i \) is the square of the multiple correlation coefficient when the feature variable \( x_i \) is regressed against all other feature variables. The VIF value ranges from 1 to infinity. A large VIF \( x_i \) indicates a strong linear relationship between \( x_i \) and other feature variables, as \( R^2_i \) approaches 1. A VIF value exceeding ten is generally considered a sign of severe collinearity.

To mitigate the adverse effects of multicollinearity, we compute the VIFs for both the technical features \( x_1^t \) defined in (23) and the fundamental features listed in Table 1 for each dataset in our empirical study. The results reveal severe multicollinearity between the ERM factor and the most recent 1/N portfolio return (i.e., \( r_{w_{e,t-1}} \)). This observation aligns well with the findings of Guo et al. (2019) that the performance of the 1/N portfolio is tied to market conditions. To alleviate the collinearity effect, we exclude the ERM factor and recalculate the VIFs. The VIFs of the remaining variables are all under ten, indicating no severe collinearity among the remaining feature variables. Therefore, in the later empirical study, we exclude the ERM factor from the list of fundamental features in Table 1 and include the following vector of fundamental feature variables in the predictive model for the winning probability:

\[ x_2^t = [\text{SMB}, \text{HML}, \text{RMW}, \text{CMA}, \text{LTR}, \text{STR}, \text{Mon}, 1\text{ndx}], \quad (24) \]

5 Empirical Study

In this section, we will first introduce the real datasets and constituent portfolios that are considered in the empirical study. We will then examine the stationarity of the time series of the out-of-sample portfolio returns. We finally assess the out-of-sample performances of various portfolio strategies on these real datasets, including the certainty equivalent return and out-of-sample Sharpe ratio.

5.1 Datasets, Constituent Portfolios, and Combined Portfolios

Table 2 presents six datasets employed in our empirical study, providing an acronym, a succinct explanation, and the number of assets included in each dataset. Every dataset comprises a varying number of portfolios made up of stocks from NYSE, AMEX, and NASDAQ. We obtained the average value-weighted returns in daily frequency for each dataset\(^1\). Daily excess returns are calculated by subtracting the simple daily returns of a one-month treasury bill from the corresponding daily returns. All datasets are sourced from

\(^1\)Here, the average value-weighted returns refer to the portfolio returns, where the portfolios are constructed by the value-weighted method according to features or industries of the underlying stocks. Further details can be found in the Kenneth R. French data library website.
the Kenneth R. French data library. Our study utilizes historical data ranging from January 1, 1970, to December 31, 2017, unless otherwise stated. The start date, January 1, 1970, is chosen as it is approximately the earliest available date across all datasets without missing values. We split the dataset into a training set (1970-01-01 to 2012-12-31) and a testing set (the last five years). A rolling window length of \( M = 240 \) is adopted for the estimation of the parameters in sophisticated portfolios, which corresponds to one year of data.

### Table 2: List of empirical datasets.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Dataset</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLT6</td>
<td>6 Portfolios Formed on Size and Long-Term Reversal</td>
<td>6</td>
</tr>
<tr>
<td>LT10</td>
<td>10 Portfolios Formed on Long-Term Reversal</td>
<td>10</td>
</tr>
<tr>
<td>M10</td>
<td>10 Portfolios Formed on Momentum</td>
<td>10</td>
</tr>
<tr>
<td>SLT25</td>
<td>25 Portfolios Formed on Size and Long-Term Reversal</td>
<td>25</td>
</tr>
<tr>
<td>BMI25</td>
<td>25 Portfolios Formed on Book-to-Market and Investment</td>
<td>25</td>
</tr>
<tr>
<td>IP38</td>
<td>38 Industry Portfolios</td>
<td>38</td>
</tr>
</tbody>
</table>

The six datasets containing daily returns of assets are downloaded from the Kenneth French data library. The first column gives an acronym for each dataset, the middle column is a succinct description of each dataset, and the last column indicates the number of assets contained in each dataset.

Table 3 outlines the constituent portfolios included in our empirical analysis, each denoted by an abbreviation listed in the last column of the table. Panel A illustrates the 1/N portfolio, which equally distributes capital across all portfolio assets. Panel B introduces three sophisticated portfolios from existing literature, including the plug-in mean-variance portfolio (notated as \( \hat{w}_{mv} \)) in (3), the plug-in three-fund portfolio (notated as \( \hat{w}_{kz} \)), and the plug-in Bayes-Stein shrinkage rule (notated as \( \hat{w}_{bs} \)). Further details about the last two sophisticated portfolios can be found in Appendix A. For each of these three sophisticated portfolios, Tu and Zhou (2011) has developed an analytical combined portfolio based on normality assumption for asset return vectors. We consider these portfolios so that we can compare the performance of the combined portfolio from our winning probability based method with those from the analytical approach by Tu and Zhou (2011).

Table 4 outlines two panels of combined portfolios analyzed in our empirical study with an abbreviation displayed for each in the “Abbreviation” column. The prefix of each abbreviation signifies the sophisticated portfolio used in the combined portfolio, and the suffix indicates the approach used in determining the combination coefficient. Prefixes “cmv”, “cbs” and “ckz”, respectively, indicate that the classical mean-variance, the Bayes-Stein shrinkage portfolio, and the three-fund portfolio are used as the sophisticated portfolio. Suffixes “lg” and “rf”, respectively, mean that the logistic regression and random forest are applied in our winning probability based method for the determination of the combination coefficient. Furthermore, the suffix “tz” indicates the combination coefficient is determined by the analytical method from Tu and Zhou (2011). The combined portfolios with the suffix “tz” come from a further scaling of those with “tz”
and their weights are given by the formula:

$$\hat{\nu}_{c,tz} = \hat{a} [\delta \hat{w}_s + (1 - \delta) w_c]$$  \hspace{1cm} \text{(25)}$$

where the combination coefficient $\delta$ is analytically obtained as per Tu and Zhou (2011), and the scale constant $\hat{a}$ is determined by (17) to further enhance the certainty equivalent return.

Table 3: List of constituent portfolio strategies.

<table>
<thead>
<tr>
<th>#</th>
<th>constituent Portfolio Strategy</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Naive portfolio strategies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1/N portfolio</td>
<td>ew</td>
</tr>
<tr>
<td>Panel B: Sophisticated portfolio strategies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>classical mean-variance portfolio (Jorion, 1986, 1991)</td>
<td>mv</td>
</tr>
<tr>
<td>3</td>
<td>Bayes-Stein shrinkage portfolio (Kan and Zhou, 2007)</td>
<td>bs</td>
</tr>
<tr>
<td>4</td>
<td>three-fund portfolio</td>
<td>kz</td>
</tr>
</tbody>
</table>

This table lists the constituent portfolios considered in the empirical study. Panels A and B present the naive portfolios and sophisticated portfolios, respectively. The 1/N portfolio in Panel A is combined with each sophisticated portfolio from Panel B to yield three different combined portfolios.

Table 4: List of combined portfolio strategies.

<table>
<thead>
<tr>
<th>#</th>
<th>Methodology</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Our winning probability weighted combined portfolio strategy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>the logistic regression model</td>
<td>cmv-lg, cbs-lg, ckz-lg</td>
</tr>
<tr>
<td>2</td>
<td>the random forest model</td>
<td>cmv-rf, cbs-rf, ckz-rf</td>
</tr>
<tr>
<td>Panel B: Other traditional methods</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>the analytic method by Tu and Zhou (2011)</td>
<td>cmv-tz, cbs-tz, ckz-tz</td>
</tr>
<tr>
<td>5</td>
<td>the analytic method by Tu and Zhou (2011) with scaling (25)</td>
<td>cmv-tz”, cbs-tz”, ckz-tz”</td>
</tr>
</tbody>
</table>

This table lists the various combined portfolio strategies considered in the analysis. Panel A presents our winning probability weighted combined portfolio strategies with different predictive models. Panel B shows other traditional methods, including the analytic methods with and without scaling.

### 5.2 Stationarity Analysis of Portfolio Returns

When we propose the winning probability weighted combined portfolio in Section 3, we assume that the out-of-sample returns of each constituent portfolio constitute a weakly stationary time series. To verify the validity of the stationarity assumption, we perform both the augmented Dickey-Fuller (ADF) unit root test (Fuller, 1995) and the KPSS stationarity test (Kwiatkowski et al., 1992) for the out-of-sample returns series of each constituent portfolio. The null hypothesis of the ADF test is that the series has a unit root, while the null hypothesis of the KPSS test is that the series is trend stationary.

Table 5 presents the ADF test statistics and KPSS test statistics, along with the corresponding p-values, for each dataset listed in Tables 2 and each constituent portfolio (either the 1/N portfolio or a sophisticated portfolio) listed in Table 3. We utilized the rolling window approach to estimate the unknown parameters in portfolio strategies over the training dataset (from 1970-01-01 to 2012-12-31) with an estimation window...
Table 5: The statistical hypothesis tests for stationarity.

<table>
<thead>
<tr>
<th></th>
<th>SLT6</th>
<th>LT10</th>
<th>M10</th>
<th>SLT25</th>
<th>BMI25</th>
<th>IP38</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: the Augmented Dickey-Fuller (ADF) test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ew</td>
<td>-18.65</td>
<td>-18.82</td>
<td>-18.71</td>
<td>-18.52</td>
<td>-18.81</td>
<td>-18.58</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>mv</td>
<td>-19.32</td>
<td>-97.39</td>
<td>-32.02</td>
<td>-16.18</td>
<td>-49.76</td>
<td>-18.62</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Panel B: the KPSS test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ew</td>
<td>0.03</td>
<td>0.05</td>
<td>0.04</td>
<td>0.03</td>
<td>0.05</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>mv</td>
<td>0.12</td>
<td>0.04</td>
<td>0.04</td>
<td>0.06</td>
<td>0.08</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>kz</td>
<td>0.12</td>
<td>0.04</td>
<td>0.06</td>
<td>0.08</td>
<td>0.05</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>bs</td>
<td>0.13</td>
<td>0.04</td>
<td>0.05</td>
<td>0.08</td>
<td>0.07</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
</tr>
</tbody>
</table>

The table presents the ADF test statistics and KPSS test statistics for the realized returns of each ingredient portfolio (including both the naive and sophisticated portfolio strategies) from the training dataset. The numbers in the brackets are the corresponding p-values (the probabilities that the null hypothesis will not be rejected). In the ADF test, the p-values are significantly smaller than 0.005 and thus indicated by “0.00”. In the KPSS test, a boundary point is returned if the p-value is outside the interval [0.01, 0.1].

length of \( M = 240 \). Subsequently, we obtained a series of out-of-sample portfolio realized returns for each portfolio strategy and each dataset. The ADF test and KPSS test were then applied to each realized series individually. From the results of the ADF test, all of the p-values are significantly smaller than 0.05, leading us to reject the existence of a unit root in these time series. This implies that the returns of every constituent portfolio are stationary. On the other hand, the results of the KPSS test show that all of the p-values are significantly larger than 0.05, and thus, we cannot reject the null hypothesis. We follow Cheung and Chinn (1996) in interpreting the results from these two tests. That is, rejecting the ADF test (i.e., p-value < 0.05) and accepting the KPSS test (i.e., p-value > 0.05) indicate strong evidence of stationarity.

Table 6: The variance adjusted parameters \( \xi \).

<table>
<thead>
<tr>
<th>sophisticated portfolio</th>
<th>SLT6</th>
<th>LT10</th>
<th>M10</th>
<th>SLT25</th>
<th>BMI25</th>
<th>IP38</th>
</tr>
</thead>
<tbody>
<tr>
<td>mv</td>
<td>0.1193</td>
<td>0.1653</td>
<td>0.1409</td>
<td>0.0798</td>
<td>0.1015</td>
<td>0.0725</td>
</tr>
<tr>
<td>kz</td>
<td>0.1588</td>
<td>0.3369</td>
<td>0.2503</td>
<td>0.1653</td>
<td>0.2856</td>
<td>0.2109</td>
</tr>
<tr>
<td>bs</td>
<td>0.1474</td>
<td>0.2598</td>
<td>0.2038</td>
<td>0.1094</td>
<td>0.1625</td>
<td>0.1116</td>
</tr>
</tbody>
</table>

This table presents the long-run variance adjusted parameter \( \xi \) for each of the three sophisticated portfolios.

In view of the ADF and KPSS test results, we calculate the long-run variances for each constituent portfolio using the Bartlett (Newey-West) estimator previously introduced in equation (10), and consequently obtain the estimate of the long-run variance adjustment parameter \( \xi \) for each of the three sophisticated portfolios as shown in Table 6. With these estimates, we derive the empirical long-run variance-adjusted sophisticated portfolio using (11), denoted as \( \hat{u}_s = \hat{\xi} \hat{w}_s \), which is taken as a component in the combined
Furthermore, the implementation of the CER optimization described in Section 3.3 requires the return rate sequence from the combined portfolio \( \hat{w}_c \) (defined in (14)) to have a long-run variance. To ensure this, we also check the out-of-sample return sequence of \( \hat{w}_c \) by the ADF test and KPSS test for each dataset. The results are reported in Table 7 when the logistic regression is used as the predictive model for the estimation of the winning probability. Similar results have been observed when the random forest is adopted for the estimation of the winning probability. The findings from both the ADF and KPSS tests support the stationarity assumption for almost all the datasets. The only exception is with the dataset SLT6, where the ADF test confirms stationarity, but the KPSS test does not at the significance level of 0.05.

Table 7: The statistical hypothesis tests for stationarity.

<table>
<thead>
<tr>
<th></th>
<th>SLT6</th>
<th>LT10</th>
<th>M10</th>
<th>SLT25</th>
<th>BMI25</th>
<th>IP38</th>
</tr>
</thead>
<tbody>
<tr>
<td>( M = 240, \gamma = 5 )</td>
<td>Panel A: the Augmented Dickey-Fuller (ADF) test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td></td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
<td>(0.00)</td>
</tr>
<tr>
<td>Panel B: the KPSS test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cmv-lg</td>
<td>0.17</td>
<td>0.08</td>
<td>0.09</td>
<td>0.10</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>ckz-lg</td>
<td>0.16</td>
<td>0.07</td>
<td>0.09</td>
<td>0.12</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>cbs-lg</td>
<td>0.17</td>
<td>0.07</td>
<td>0.10</td>
<td>0.12</td>
<td>0.03</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
</tr>
</tbody>
</table>

The table presents the ADF test statistics and KPSS test statistics for the realized returns of combined portfolios (combining the 1/N rule with one of the sophisticated portfolio strategies) from the training dataset. The winning probability is achieved by logistic regression. The numbers in the brackets are the corresponding p-values (the probabilities that the null hypothesis will not be rejected). In the ADF test, the p-values are significantly smaller than 0.005 and thus indicated by “0.00”. In the KPSS test, a boundary point is returned if the p-value is outside the interval [0.01, 0.1].

5.3 Portfolio Performance Assessment

As previously discussed, we employ both the logistic regression and the random forest as predictive models to construct the combination coefficient for the WPW combined portfolios. To examine the improvement achieved from the fundamental factors, we compare two scenarios of input features for each predictive model: one including only technical features and the other having both fundamental and technical features. Following the training-testing procedure outlined in Section 3.4, we apply a rolling window approach with the window length \( M = 240 \). The risk aversion coefficient is set to be \( \gamma = 5 \). We use the training dataset (from 1970-01-01 to 2012-12-31) to train the predictive models, and then freeze the model parameters for developing the WPW combined portfolios over the testing set (from 2013-01-01 to 2017-12-31). To compare the out-of-sample performance with traditional portfolios, the same testing set and rolling window approach
are applied to both the constituent portfolios and the analytically combined portfolios, in our empirical study.

Table 8: Certainty equivalent return on various datasets.

<table>
<thead>
<tr>
<th></th>
<th>SLT6</th>
<th>LT10</th>
<th>M10</th>
<th>SLT25</th>
<th>BMP25</th>
<th>IP38</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Constituent portfolios</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ew</td>
<td>3.94</td>
<td>4.15</td>
<td>4.24</td>
<td>4.03</td>
<td>4.19</td>
<td>3.81</td>
</tr>
<tr>
<td>mv</td>
<td>-22.68</td>
<td>-33.96</td>
<td>-58.89</td>
<td>-120.01</td>
<td>-129.46</td>
<td>-149.15</td>
</tr>
<tr>
<td>kz</td>
<td>2.35</td>
<td>4.72</td>
<td>-2.53</td>
<td>7.23</td>
<td>4.10</td>
<td>7.52</td>
</tr>
<tr>
<td>bs</td>
<td>-0.34</td>
<td>-1.73</td>
<td>-13.52</td>
<td>-23.61</td>
<td>-24.85</td>
<td>-18.17</td>
</tr>
<tr>
<td>Panel B: Our winning probability weighted combined portfolios</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) Logistic regression with both technical and fundamental features</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cmv-lg</td>
<td>8.85</td>
<td>6.98</td>
<td>2.21</td>
<td>6.67</td>
<td>6.88</td>
<td>8.93</td>
</tr>
<tr>
<td>ckz-lg</td>
<td>8.38</td>
<td>7.64</td>
<td>5.76</td>
<td>8.10</td>
<td>7.83</td>
<td>8.01</td>
</tr>
<tr>
<td>cbs-lg</td>
<td>8.57</td>
<td>7.54</td>
<td>4.60</td>
<td>7.80</td>
<td>7.62</td>
<td>8.89</td>
</tr>
<tr>
<td>(2) Logistic regression with only technical features</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cmv-lg</td>
<td>8.36</td>
<td>6.62</td>
<td>1.56</td>
<td>6.21</td>
<td>7.10</td>
<td>8.95</td>
</tr>
<tr>
<td>ckz-lg</td>
<td>7.68</td>
<td>7.32</td>
<td>5.43</td>
<td>7.75</td>
<td>7.71</td>
<td>7.66</td>
</tr>
<tr>
<td>cbs-lg</td>
<td>7.93</td>
<td>7.21</td>
<td>4.31</td>
<td>7.28</td>
<td>7.51</td>
<td>8.84</td>
</tr>
<tr>
<td>(3) Random Forest with both technical and fundamental features</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cmv-rf</td>
<td>8.82</td>
<td>7.20</td>
<td>2.09</td>
<td>6.30</td>
<td>6.94</td>
<td>8.34</td>
</tr>
<tr>
<td>ckz-rf</td>
<td>8.41</td>
<td>7.47</td>
<td>5.60</td>
<td>7.62</td>
<td>7.60</td>
<td>7.59</td>
</tr>
<tr>
<td>cbs-rf</td>
<td>8.43</td>
<td>7.71</td>
<td>4.34</td>
<td>6.99</td>
<td>7.28</td>
<td>8.66</td>
</tr>
<tr>
<td>(4) Random Forest with only technical features</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cmv-rf</td>
<td>7.78</td>
<td>7.05</td>
<td>1.51</td>
<td>5.09</td>
<td>6.66</td>
<td>8.62</td>
</tr>
<tr>
<td>ckz-rf</td>
<td>7.68</td>
<td>7.27</td>
<td>5.35</td>
<td>6.70</td>
<td>7.49</td>
<td>7.47</td>
</tr>
<tr>
<td>cbs-rf</td>
<td>7.28</td>
<td>7.33</td>
<td>4.20</td>
<td>6.39</td>
<td>7.12</td>
<td>8.68</td>
</tr>
<tr>
<td>Panel C: Combined portfolios with the analytic method</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) without scaling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cmv-tz</td>
<td>-0.07</td>
<td>3.44</td>
<td>-4.63</td>
<td>-0.54</td>
<td>0.08</td>
<td>4.39</td>
</tr>
<tr>
<td>ckz-tz</td>
<td>3.19</td>
<td>5.98</td>
<td>3.25</td>
<td>7.37</td>
<td>5.84</td>
<td>5.05</td>
</tr>
<tr>
<td>cbs-tz</td>
<td>0.91</td>
<td>4.93</td>
<td>-0.43</td>
<td>4.34</td>
<td>3.33</td>
<td>5.71</td>
</tr>
<tr>
<td>(2) with scaling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cmv-tz'</td>
<td>2.88</td>
<td>3.34</td>
<td>-0.03</td>
<td>2.61</td>
<td>2.27</td>
<td>1.94</td>
</tr>
<tr>
<td>ckz-tz'</td>
<td>2.98</td>
<td>4.23</td>
<td>2.23</td>
<td>3.69</td>
<td>4.08</td>
<td>2.49</td>
</tr>
<tr>
<td>cbs-tz'</td>
<td>3.04</td>
<td>3.71</td>
<td>1.31</td>
<td>3.09</td>
<td>2.73</td>
<td>2.01</td>
</tr>
</tbody>
</table>

This table presents the certainty equivalent return (in basis point) of various datasets in the testing period. The moving window length is 240, and the risk aversion coefficient is 5. Panel B includes our WPW combined portfolios from different predictive models and input features. Panels A and C present the constituent portfolios and the analytic combined portfolios for comparison.

Table 8 presents the CER of various datasets and portfolios. This table generally shows the superiority of the WPW combined portfolios. Several key observations can be made from this table. Firstly, a comparison between Panels A and B shows that our WPW combined portfolios significantly outperform all of the constituent portfolios in nearly all scenarios. Specifically, for the dataset M10, while the WPW combined portfolio with the mean-variance portfolio as an ingredient yields a better CER than the corresponding sophisticated portfolio, it falls behind when compared to the 1/N portfolio. Even with this dataset, the WPW combined portfolios only fall slightly short of the corresponding constituent portfolio. Nevertheless, we observe significant improvement when changing from the constituent portfolios to the corresponding WPW combined portfolios. Secondly, comparing Panels B and C reveals that the WPW combined portfolios generally produce a superior out-of-sample CER than those from the analytical method. In most scenarios, they exceed their counterparts from the analytical method by a substantial margin. Finally, Panel B also
examines the impact of different input features in the predictive model. Including both technical and fundamental features as inputs leads to better performance than using only technical features in most scenarios. This finding suggests that the inclusion of fundamental features provides valuable additional information, thereby enhancing the out-of-sample performance of the WPW combined portfolio.

Table 9: Out-of-sample Sharpe ratio on various datasets.

<table>
<thead>
<tr>
<th></th>
<th>SLT6</th>
<th>LT10</th>
<th>M10</th>
<th>SLT25</th>
<th>BMI25</th>
<th>IP38</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Ingredient portfolios</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ew</td>
<td>6.72</td>
<td>7.19</td>
<td>7.18</td>
<td>6.75</td>
<td>7.23</td>
<td>6.67</td>
</tr>
<tr>
<td>mv</td>
<td>7.10</td>
<td>5.61</td>
<td>-0.44</td>
<td>5.79</td>
<td>4.98</td>
<td>5.77</td>
</tr>
<tr>
<td>kz</td>
<td>7.80</td>
<td>7.88</td>
<td>4.90</td>
<td>9.41</td>
<td>7.51</td>
<td>8.67</td>
</tr>
<tr>
<td>bs</td>
<td>7.78</td>
<td>7.20</td>
<td>3.07</td>
<td>7.68</td>
<td>6.00</td>
<td>7.01</td>
</tr>
<tr>
<td><strong>Panel B: Our winning probability weighted combined portfolios</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) Logistic regression with both technical and fundamental features</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cmv-lg</td>
<td>9.43</td>
<td>8.42</td>
<td>5.02</td>
<td>8.65</td>
<td>8.30</td>
<td>9.47</td>
</tr>
<tr>
<td>ckw-lg</td>
<td>9.19</td>
<td>8.90</td>
<td>7.59</td>
<td>9.29</td>
<td>8.90</td>
<td>8.96</td>
</tr>
<tr>
<td>cbs-lg</td>
<td>9.29</td>
<td>8.79</td>
<td>6.88</td>
<td>9.21</td>
<td>8.73</td>
<td>9.46</td>
</tr>
<tr>
<td>(2) Logistic regression with only technical features</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cmv-lg</td>
<td>9.19</td>
<td>8.17</td>
<td>5.00</td>
<td>8.45</td>
<td>8.42</td>
<td>9.48</td>
</tr>
<tr>
<td>ckw-lg</td>
<td>8.84</td>
<td>8.66</td>
<td>7.38</td>
<td>9.13</td>
<td>8.82</td>
<td>8.75</td>
</tr>
<tr>
<td>cbs-lg</td>
<td>8.97</td>
<td>8.56</td>
<td>6.69</td>
<td>8.99</td>
<td>8.67</td>
<td>9.43</td>
</tr>
<tr>
<td>(3) Random Forest with both technical and fundamental features</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cmv-rf</td>
<td>9.41</td>
<td>8.60</td>
<td>5.25</td>
<td>8.48</td>
<td>8.33</td>
<td>9.14</td>
</tr>
<tr>
<td>ckw-rf</td>
<td>9.20</td>
<td>8.80</td>
<td>7.48</td>
<td>9.08</td>
<td>8.78</td>
<td>8.71</td>
</tr>
<tr>
<td>cbs-rf</td>
<td>9.22</td>
<td>8.94</td>
<td>6.67</td>
<td>8.86</td>
<td>8.54</td>
<td>9.32</td>
</tr>
<tr>
<td>(4) Random forest with only technical features</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cmv-rf</td>
<td>8.89</td>
<td>8.48</td>
<td>4.87</td>
<td>7.99</td>
<td>8.17</td>
<td>9.29</td>
</tr>
<tr>
<td>ckw-rf</td>
<td>8.85</td>
<td>8.65</td>
<td>7.32</td>
<td>8.69</td>
<td>8.69</td>
<td>8.64</td>
</tr>
<tr>
<td>cbs-rf</td>
<td>8.64</td>
<td>8.66</td>
<td>6.59</td>
<td>8.61</td>
<td>8.44</td>
<td>9.33</td>
</tr>
<tr>
<td><strong>Panel C: Combined portfolios with the analytic method</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>without scaling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cmv-tz</td>
<td>6.29</td>
<td>6.49</td>
<td>0.92</td>
<td>5.82</td>
<td>5.10</td>
<td>6.63</td>
</tr>
<tr>
<td>ckw-tz</td>
<td>6.79</td>
<td>7.77</td>
<td>5.79</td>
<td>8.85</td>
<td>8.20</td>
<td>8.01</td>
</tr>
<tr>
<td>cbs-tz</td>
<td>6.64</td>
<td>7.23</td>
<td>3.85</td>
<td>7.27</td>
<td>6.18</td>
<td>7.82</td>
</tr>
<tr>
<td>with scaling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cmv-tz</td>
<td>6.40</td>
<td>6.56</td>
<td>1.05</td>
<td>5.93</td>
<td>5.19</td>
<td>6.82</td>
</tr>
<tr>
<td>ckw-tz</td>
<td>6.90</td>
<td>7.83</td>
<td>5.90</td>
<td>8.96</td>
<td>8.26</td>
<td>8.14</td>
</tr>
<tr>
<td>cbs-tz</td>
<td>6.75</td>
<td>7.30</td>
<td>3.98</td>
<td>7.39</td>
<td>6.27</td>
<td>8.05</td>
</tr>
</tbody>
</table>

This table presents the out-of-sample Sharpe ratio (in percentage) of various datasets in the testing period. The moving window length is 240, and the risk aversion coefficient is 5. Panel B includes our WPW combined portfolios from different predictive models and input features. Panels A and C present the constituent portfolios and the analytic combined portfolios for comparison.

Table 9 presents the out-of-sample Sharpe ratio of various datasets and portfolios, and provides evidence of the superior performance of the WPW combined portfolios. Several key observations can be made from this table. First, a comparison between Panels A and B shows that the WPW combined portfolios provide a higher out-of-sample Sharpe ratio than all constituent portfolios in most cases. The exceptional scenarios occur with the 1/N portfolio in M10 and the three-fund portfolio in SLT25, and the shortfall is marginal. Second, comparing Panel B with Panel C reveals that the WPW combined portfolios achieve a significantly larger out-of-sample Sharpe ratio than their counterparts from the analytical method in most cases. This result further underscores the superiority of our WPW combined portfolio approach. Finally, Panel B also
examines the impact of different input features in the predictive model. Consistent with the findings from the CER results, including both technical and fundamental features leads to a higher out-of-sample Sharpe ratio compared to using only technical features in most scenarios. This observation highlights the value of considering additional fundamental features for constructing the combined portfolio, a benefit that is not readily achieved in existing analytical methods.

6 Conclusion

In this paper, we introduced a flexible statistical framework for constructing combined portfolio rules that do not rely on the unrealistic assumption of normal distribution for asset returns. Instead, we propose a winning probability weighted approach, where the combination coefficient is characterized as the winning probability and determined using machine learning techniques. This allows for the incorporation of both technical and fundamental feature variables in the determination of the combination coefficient. While we have focused on combining the 1/N portfolio with one of three specific sophisticated portfolios in our study, the framework can be easily extended to any pair of constituent portfolios to construct a desirably combined portfolio as long as the constituent portfolios have weakly stationary return rate sequences.

Our extensive empirical studies have demonstrated the superior performance of the winning probability weighted combined portfolio over most scenarios, as evidenced by higher certainty equivalent returns and out-of-sample Sharpe ratios. Even in cases where our approach slightly underperforms compared to the analytical method, the difference is marginal. Additionally, we have observed that including fundamental factors in the determination of the combination coefficients, along with historical asset returns, improves the out-of-sample performance of the combined portfolio.

As a future research direction on the topic, it is interesting to explore more machine-learning techniques for determining the combination coefficient within the winning probability weighted framework. For instance, deep neural networks, long short-term memory networks, or ensemble models that combine several algorithms could be investigated. Moreover, there is room for exploring more cross-sectional financial market feature variables to enhance the construction of the combination coefficient and further improve the out-of-sample performance of the combined portfolio.

References


28


Appendices

A Constituent sophisticated portfolios

A.1 Three-fund rule

The three-fund rule is proposed by Kan and Zhou (2007) as

$$
\hat{w}_{kz} = \frac{1}{\gamma} (c\hat{\Sigma}^{-1}\hat{\mu} + d\hat{\Sigma}^{-1}1_N),
$$

where $c$ and $d$ are constants determined to maximize the expected out-of-sample mean-variance utility. The optimal values of $c$ and $d$ indeed depend on the true values of $\mu$ and $\Sigma$, and their estimations have also been addressed in Kan and Zhou (2007).

A.2 Jorion’s rule

Motivated by both a shrinkage consideration and a Bayesian analysis under a suitable informed prior, Jorion (1986, 1991) develops a Bayes-Stein estimator of $\mu$ as

$$
\hat{\mu}_{bs} = (1 - \hat{v})\hat{\mu} + \hat{v}\hat{\mu}_g 1_N,
$$

where $\hat{\mu}_g$ is the average excess return on the sample global minimum-variance portfolio as the shrinkage target and $\hat{v}$ is the weight given to the target. That is,

$$
\hat{\mu}_g = \frac{1^N\Sigma^{-1}\hat{\mu}}{1^N\Sigma^{-1}1_N},
$$

$$
\hat{v} = \frac{N + 2}{(N + 2) + T(\hat{\mu} - \hat{\mu}_g 1_N)^T \Sigma^{-1}(\hat{\mu} - \hat{\mu}_g 1_N)},
$$

where $\Sigma$ is the shrinkage estimator of covariance matrix as $\hat{\Sigma} = T\hat{\Sigma}/(T - N - 2)$. The Bayes-Stein estimator of $\Sigma$ as given in Kan and Zhou (2007) is

$$
\hat{\Sigma}_{bs} = \left( \frac{T + \hat{u} + 1}{T + \hat{u}} \right) \hat{\Sigma} + \frac{\hat{u}}{T(T + 1 + \hat{u})} \frac{1_N 1_N^T}{1_N^N \Sigma^{-1}1_N},
$$

where $\hat{u} = (N + 2)/[(\hat{\mu} - \hat{\mu}_g 1_N)^T \Sigma^{-1}(\hat{\mu} - \hat{\mu}_g 1_N)]$. Hence, the Bayes-Stein portfolio is given by

$$
\hat{w}_{bs} = \frac{1}{\gamma} (\hat{\Sigma}_{bs})^{-1}\hat{\mu}_{bs}.
$$