Pages 1–10
2023

Factor-Based Portfolio Optimization

Jun Kyung Auh and Wonho Cho
Quantitative Finance Lab
School of Business, Yonsei University
January 2023

Keywords: Portolio Optimization, Factor Model, Algorithmic Trading

ABSTRACT

We discuss the struggles in out-of-sample (OOS) performance for mean-variance optimization using historical returns and improve the OOS performance using a single-factor model. Through this paper, we only differentiate the first moment (expected return) and keep the same covariance matrix for both portfolios to simplify comparing the two models. Using 30 stocks of the Dow Jones Index components from January 2010 through July 2022, we confirm that historical forecasts are imprecise guides for future portfolio performance and improve the performance in two stages. In the first stage, we improve future portfolio performance by adopting a factor-based model. Given this result, factor-based portfolio optimization can relieve the estimation errors from historical expected returns by using factor loadings. In the second stage, a machine-learning technique called Support Vector Regression (SVR) is proposed to predict market returns for the next period. By replacing the historical average market returns with the predicted market returns in a single-factor model, we further improve the performance of the factor-based portfolio.

1 INTRODUCTION

Portfolio optimization has always been the center of attention by investment practitioners and finance academics. Since the introduction of modern portfolio theory, different models of optimization schemes have been developed. Most notably, Sharpe ratio maximization in the mean-variance framework establishes the firm ground for many others (e.g., maximum diversification and minimum variance). Theoretically, the mean-variance optimization assumes that individual assets have well-defined expected returns, return volatility, as well as correlation structure among them. For example, in continuous-time frameworks, asset return dynamics are defined with known parameters such as drift and diffusion volatility. In other words, in spite of a certain extent of uncertainty, the expected returns of each asset and return covariance across them are known. Oftentimes, they are time-invariant: at each point in time, the asset return distribution has a constant mean and variance. Even in more flexible models with time-varying distributions, the assumption is that those pieces of information are known in a forward-looking sense.

The application of this theory, however, faces an immediate challenge. While these two pieces of information (expected return and covariance matrix) are key inputs for the mean-variance optimization, estimating them is not straightforward. In practice, many use mean and correlation of historical return. This practice can be justified only under the assumption that historical returns predict the future. We already know this is not true. Although the past returns indicate characteristics of underlying risks, using them to estimate expected return is subject to a fatal flaw (Michaud, 1989; Chopra and Ziemba, 1993; Best and Grauer, 2015). We cannot argue that an asset that experienced a steep price appreciation in the past will continue to enjoy such a pattern in the future. However, feeding expected return based on the historical pattern to the mean-variance optimization results overweight on those assets, exposing the entire portfolio to the long-term momentum. Also, should any asset in the portfolio be associated with an extremely high or low historical return, the optimization often shows a corner solution: the allocation is entirely concentrated on this particular asset. In this case, this wrong estimation can defeat the purpose of portfolio optimization designed to maximize the diversification benefit.

In this document, we discuss an approach to overcome this problem. In addition to addressing this challenge, our approach is designed to achieve the following goal: it must be applicable to general equity securities without requiring any firm- and market-specific information that would limit the scope of the application. Therefore, the merit of our approach is that a portfolio of a broader set of equities can be optimized. One can imagine its immediate benefit in the context of direct indexing, a recent financial innovation that each investor can create a customized index with his/ her own choices of assets and a tailor-made optimization objective. For instance, an institutional investor mandated to make ESG investments would want to achieve the highest attainable Sharpe ratio for chosen ESG-related stocks. Another example is an application for a retail client who wishes to conduct tax harvesting on an individual portfolio. As every individual holds different portfolios, our forward-looking optimization scheme with a large asset coverage would be a good candidate.

This paper is closely related to studies that propose robust portfolio optimization using forward-looking information. Several papers use forward-looking information, such as implied volatility extracted from option data, for the portfolio covariance matrix (Kostakis *et al.*, 2011; Kempf *et al.*, 2014; Bianchi and Tassinari, 2020). In alignment with the literature, we use well-known predictors to predict market returns. However, we differentiate from prior literature in that we apply the predicted market returns in the single-factor model to predict the future expected returns of individual stocks.

Moreover, this research has a relationship with the topics of the factor model and the machine-learning to improve the optimization process. Many prior studies focus on finding the latent factors from abundant data and comparing prediction power between the well-known factor and the factors constructed by machine learning techniques (Feng *et al.*, 2018; Gu *et al.*, 2021). However, our approach differs from the existing literature. We use a machine learning method to predict factor returns, such as market returns, and then use this factor return directly in the factor model.

Lastly, this paper is related to a hybrid investment portfolio strategy. Much literature focuses on the hybrid strategy by setting multiple objective optimization processes (Roman *et al.*, 2007; Chen and Wang, 2015). Similar to these studies, we combine a minimum variance optimization and a maximizing Sharpe ratio optimization in the factor-based model for two reasons. One is the technical reason that, for a single-factor model, if market returns are negative and all the betas of individual stocks are positive, all of the expected returns of individual

stocks are negative. In this case, the quadratic problem should not be solved; therefore, an alternative optimization process, such as minimum-variance optimization, is employed instead. The other reason is the justification (implication) of this combining strategy in that when the expected returns of individual stocks are predicted to be negative, investors should implement the minimum-variance strategy instead of the maximizing Shape ratio strategy to avoid risks. This trading strategy is similar to the Chen and Wang (2015) in that maximizing the Sharpe ratio portfolio performs the best in a bull market, while the minimum-variance portfolio performs the best in a bear market.

The summary of our findings is as follows. Firstly, we find that the portfolio optimized with a single-factor model is more diversified than the portfolio with a historical model. Remarkably, the single-factor model shows a higher diversification ratio and lower concentration (Herfindahl–Hirschman index) than the historical model. Next, the single-factor model shows the highest Sharpe ratio and the lowest maximum drawdown, while the historical model shows the worst. Given that the historical expected returns are exposed to the estimation errors such as noises and biases, factor-based portfolio optimization can relieve these errors and perform better. Lastly, we further improve a single factor-based portfolio reflecting forward-looking signals. We use high-frequency market signals that are well-known to predict negative market returns. Then, we adopt the machine learning technique to reflect market predictors and estimate expected market returns with this technique. By doing so, we show that the optimization results using machine learning outperform the previous ones.

The remainder of this paper is organized as follows. Section 2 describes the mean and variance estimation process using historical returns and a single-factor model. Section 3 explains the optimization process and defines the performance measures. Section 4 compares the performance between the factor-based model and the historical model. Section 5 shows the further improved performance adopting the machine learning technique, and Section 6 concludes.

2 FACTOR MODEL

We use the following notation throughout this paper:

- $r_{f,t}$: return for risk-free asset at date t.
- $r_{i,t}$: return for asset i at date t, stacked into $r_t := (r_{1,t}, ..., r_{N,t})^{'}$.
- $r^e_{i,t}$: excess return for asset i over risk-free return ($r_{f,t}$) at date t, stacked into $r^e_t := (r^e_{1,t}, ..., r^e_{N,t})^{'}$.
- α_i: abnormal return for asset i, unexplained by factor model.
- $f_{k,t}$: return for factor k at date t, stacked into $f_t := (f_{1,t}, ..., f_{K,t})^{'}$.
- $\varepsilon_{i,t}$: error term for asset i at date t, stacked into $\varepsilon_{i,t} := (\varepsilon_{1,t}, ..., \varepsilon_{K,t})^{'}$.
- Σ_t:= covariance matrix estimation of returns for asset i and asset j at time t, where $i = 1, ..., N$ and $j = 1, ..., N$. Its elements are composed of N variances, and $N(N-1)$ covariances.

We adopt an unconditional static factor model. For every asset $i = 1, ..., N$ and factor $k = 1, ..., K$,

$$r^e_{i,t} = \alpha_i + \beta_i^{'} f_t + \varepsilon_{i,t}, \tag{1}$$

with $\beta_i := (\beta_{i,1}, ...\beta_{i,K}^{'})$ and $E(\varepsilon_{i,t}|f_t) = 0$.

Among factor models, we adopt the single-factor model. Therefore, Equation (1) can be simplified as follows:

$$r^e_{i,t} = \alpha_i + \beta_{i,m}(r_{m,t} - r_{f,t}) + \varepsilon_{i,t}, \tag{2}$$

where $r_{m,t}$ is the market returns and $\beta_{i,m}$ is the market beta. Using Equation (2), we estimate the alpha and the market beta with a lookback window of 252 business days.

2.1 Factor-based expected returns

We calculate the model-based returns with the estimated beta. We only keep the explained terms by the factor (market) model and exclude the unexplained alpha. Furthermore, the error term also is excluded because this noise term should be technically averaged out when we calculate the expected returns. Therefore, the excess return of the factor-based model is calculated as follows:

$$r^{Model,e}_{i,t} = \hat{\beta}_i^{'} f_t, \tag{3}$$

Then, we include the risk-free rate to generate the factor-based return.

$$r^{Model}_{i,t} = r_{f,t} + \hat{\beta}_i^{'} f_t, \tag{4}$$

Finally, we calculate the model-based expected return using factor-based returns for the prior 126 business days:

$$\mu^{Model}_{i,t} = \frac{1}{126}\sum_{\tau=0}^{125} r^{Model}_{i,t-\tau}. \tag{5}$$

Since we adopt the single-factor model, Equations (4) and (5) can be written:

$$r^{factor}_{i,t} = r_{f,t} + r^{factor,e}_{i,t}, \tag{6}$$

and

$$\mu^{factor}_{i,t} = \frac{1}{126}\sum_{\tau=0}^{125} r^{factor}_{i,t-\tau}. \tag{7}$$

2.2 Historical expected returns and covariance

The historical expected return is the average of individual asset returns for the prior 126 business days:

$$\mu^{Hist}_{i,t} = \frac{1}{126}\sum_{\tau=0}^{125} r^{Hist}_{i,t-\tau}. \tag{8}$$

Furthermore, we estimate the covariance matrix using individual asset returns for the prior 126 business days:

$$\sigma^{Hist}_{ij,t} = \frac{1}{125}\sum_{\tau=0}^{125}(r^{Hist}_{i,t-\tau} - \mu^{Hist}_{i,t})(r^{Hist}_{j,t-\tau} - \mu^{Hist}_{j,t}). \tag{9}$$

3 OPTIMIZATION AND PERFORMANCE MEASURES

3.1 Optimization

Among various mean-variance optimization methods, we adopt the maximizing Sharpe ratio optimization with short sales constraints, which is given by,

$$\begin{aligned} \max_w SR = & \frac{\mu^{'} w}{(w^{'}\Sigma w)^{1/2}} \\ \text{s.t.} \quad & e^{'} w = 1 \\ & w \geq 0, \end{aligned} \tag{10}$$

where $\mu \in \mathbb{R}^N$ is a vector of expected returns for N different assets, $w \in \mathbb{R}^N$ is a vector of weights for N different assets, and $\Sigma \in \mathbb{S}^N_+$ denotes the corresponding covariance matrix. The sum of weights for

assets is equal to 1 ($e^{'}w = 1$) and a short sale is constrained ($w \geq 0$) to reduce the estimation error (Jagannathan and Ma, 2003; Garlappi *et al.*, 2006).[1]

Moreover, we consider additional mean-variance optimization models for benchmarks: the minimizing variance optimization (Global Minimum Variance Portfolio; GMVP) and the maximizing diversification ratio optimization.

The minimizing variance optimization is given by,

$$\begin{aligned} \min_{w} \quad & w^{'}\Sigma w^{1/2} \\ \text{s.t.} \quad & e^{'}w = 1 \\ & w \geq 0, \end{aligned} \tag{11}$$

where $(w^{'}\Sigma w)^{1/2}$ is the portfolio's standard deviation, and constraints are the same with Equation (10).

The maximizing diversification ratio optimization is given by,

$$\begin{aligned} \max_{w} DR = & \frac{w^{'}diag(\sigma_1, ..., \sigma_N)}{(w^{'}\Sigma w)^{1/2}} \\ \text{s.t.} \quad & e^{'}w = 1 \\ & w \geq 0, \end{aligned} \tag{12}$$

where $w^{'}diag(\sigma_1, ..., \sigma_N)$ is the linear combination of the weight of the asset and its standard deviation, implying the portfolio's standard deviation without considering the diversification effect. On the other hand, $(w^{'}\Sigma w)^{1/2}$ is the portfolio's standard deviation reflecting the diversification effect. Therefore, the well-diversified portfolio should have a high value of DR because the denominator is much smaller than the numerator. Constraints are the same with Equation (10).

3.2 Performance Measures

We use several measures to compare the performance between the optimized portfolio with the historical expected returns and the factor-based expected returns. The portfolio performance measures, in general, can be classified into two categories: measures for diversification effect and financial performance. Among many alternatives for diversification, we select the Herfindahl–Hirschman Index (HHI) and the diversification ratio. These diversification measures are in-sample measures because the weights of the stocks in the portfolio, the result of mean-variance optimization with in-sample, calculate these measures. On the other hand, using ex-post portfolio returns, we measure the financial performance through the maximum drawdown (risk) and the Sharpe ratio (reward-to-risk).

We measure the degree of concentration (HHI), which is in line with Chamberlain (1983) and Green and Hollifield (1992) in that portfolio diversification is identified by the sum of squared weights, or l_2-norm. This measure converges to zero as the number of assets (N) increases to infinity. For stock i, HHI is measured as follows:

$$HHI = \sum_{i=1}^{N} w_i^2, \tag{13}$$

where w_i is the weight of stock i. HHI has a value between 0 and 1, and the high value implies a high degree of concentration. Particularly, $HHI = 1$ implies the extreme corner solution, where the portfolio is only composed of a single asset.

Then, we calculate the diversification ratio (DR) to measure the diversified effect.

$$DR = \frac{w^{'}diag(\Sigma)}{(w^{'}\Sigma w)^{1/2}}, \tag{14}$$

where diag(Σ) is the diagonal terms of the covariance matrix $(\sigma_1, ..., \sigma_N)$. $w^{'}diag(\Sigma)$ means the linear combination of the portfolio standard deviation, implying that all assets are perfectly positively correlated ($\rho = 1$). On the other hand, $(w^{'}\Sigma w)^{1/2}$ is the portfolio standard deviation reflecting the diversification effect. Therefore, as the diversification ratio is high, we expect the portfolio is well-diversified.

Next, we compute the maximum 1-year drawdown as the worst 252-day return in the sample to measure the maximum fall in the value of the investment following Grossman and Zhou (1993).

$$MDD = -min(r_{252d}). \tag{15}$$

where MDD is the maximum drawdown and r_{252d} is the cumulative return over the preceding 252 days.

Lastly, we measure the Sharpe ratio to compare the performance with respect to the risk-return context, neglecting the risk-free rate, written as follows:

$$SR = \frac{\mu^{'}w}{(w^{'}\Sigma w)^{1/2}}. \tag{16}$$

4 EMPIRICAL RESULTS

In this section, we compare the performance of the mean-variance portfolio using a single factor against several benchmarks. Throughout this paper, we use the Dow Jones Index (DJI) as a market index. In other words, we use DJI return as market return. Since the DJI comprises thirty firms, our portfolio is constructed with 30 securities.

The portfolio optimization is processed with four steps. First, we select 30 stocks from the CRSP, following the component list of the DJI. Second, given a 30-asset universe, we estimate the covariance matrix from the historical return data following Equation (9). Third, we separately calculate a set of expected returns for a single-factor model and historical model following Equations (5) and (8). Fourth, we solve the mean-variance portfolio problem using two models following Equation (10).

One thing to notice is that our maximizing Sharpe ratio optimization strategy is composed of two strategies: the maximizing Sharpe ratio optimization and the minimum variance optimization instead. For instance, if the solution for maximizing Sharpe ratio optimization cannot find any solution, we adopt the minimum variance optimization. One of the cases is when the market returns have been predicted to be negative. If so, since usual stocks have a positive market beta, all individual stocks will have negative expected returns (See Equation (3)). Since there is no solution to maximize the Sharpe ratio in this case, we minimize the variance of the portfolio instead. This strategy implies that when the expected returns of individual stocks in the future are all predicted to have negative values, it is appropriate to use a strategy that minimizes risk rather than the maximizing Sharpe Ratio.

Turning to the details, we implement the optimization at the end of the month from January 2009 through July 2022; therefore, there are 150 optimization dates during our sample period. Our computations are done on a rolling-horizon basis. Taking an optimization process as an example, for optimization on January 31 in 2010, we use the prior 126 business days (from August 1 in 2009 to January 31 in 2010) to estimate the historical expected returns and the historical covariance matrix following Equations (8) and (9). Then, we use the prior 252 business days to estimate the betas using Equation (2) and the prior 126 business days to estimate the factor-based expected returns following Equation (7). All computations were carried out in Python with the solver MOSEK, a package for specifying and solving convex problems.

[1] Note that we exclude the risk-free rate from the Sharpe ratio calculation.

We compare the result of the Sharpe ratio optimization between historical expected returns and factor-based expected returns. Since our goal is to compare the impact of the first moment (mean) between the factor-based model and the historical model, we use the second moment (covariance matrix) at the same for both optimization cases. For instance, we set $\mu^{Hist'}w$ as a numerator in Equation (10) for historical expected returns and $\mu^{factor'}w$ for single-factor-based expected returns. In contrast, the covariance matrix measured by historical returns is set as a numerator in both cases.

4.1 Diversification Performance: In-Sample Statistics

In this section, we compare the results for the historical expected returns with the factor-based expected returns.

Table 1 reports the optimized results of HHI and DR using historical expected returns and single-factor-based expected returns. Compared to the historical model (SRHist), the single-factor model (SRF1) presents a low average value of HHI (0.10 vs. 0.33) and a high average value of DR (1.59 vs. 1.44). These results imply that the optimized weight is more diversified in the factor model than in the historical model. One argument is that rather than the historical average return, the expected return estimated under CAPM's assumption shows more diversification. This means that optimization using historical returns could be failed due to estimation bias (e.g., momentum effect) or noises. In other words, the noise and bias from the historical average return can be much larger than the estimation error of market beta.

As a benchmark, we also present the results of minimizing variance optimization (MinVar) and maximizing diversification ratio optimization (MaxDR). Since the objective of these two benchmarks is to maximize the diversification effect, we observe that HHI (DR) of these two models is lower (higher) than that of the single-factor model in Table 1.

[Insert Table 1 here.]

4.2 Financial Performance: Out-of-Sample Statistics

In this section, we evaluate a portfolio's ex-post performance with respect to the benchmark. Using the optimized portfolio weight, we generate the one-month holding portfolios, which are rebalanced at the end of each month. Using ex-post portfolio returns, we calculate MDD and OOS SR measures following Equations (15) and (16). At first, we find that the order of OOS SR in Table 2 and in-sample SR in Table 1 are opposite. Notably, the in-sample SR of the factor-based model is approximately one-third of that of the historical model (1.301 vs. 3.652). However, the OOS SR of the factor-based model is higher than that of the historical model (1.243 vs. 1.143).

Moreover, we find that the result of the DJI is the worst. Among the results, DJI records the largest maximum drawdown (in Panel (b)) and lowest Sharper ratio (in Panel (c)). Since the optimized weight is more diversified for the single-factor model, we anticipate that the ex-post portfolio performance of the single-factor model will show low riskiness and a high Sharpe ratio. As expected, Table 2 shows that the portfolio based on the factor model has a low average MDD (3.4% vs. 3.6%) and a high average SR (1.21 vs. 1.12).

As a benchmark, we additionally present the results of minimizing variance optimization (MinVar), maximizing diversification (MaxDR) ratio optimization, and DJI. We observe the result of MinVar, as known as the Global Minimum Variance Portfolio (GMVP), shows better results than SRHist. This is consistent with Jorion (1985) that it is better to use GMVP than to use historical average return. In addition, we find evidence that the factor-based model improves the GMVP one step further. In Table 2, we find that SRF1 performs better (lower MDD and higher SR) than MinVar.

[Insert Table 2 here.]

5 FACTOR SIGNALS

In this section, we explain signals well-known for predicting factor returns and how to reflect these signals into the factor-based expected returns.

5.1 Support Vector Regression for Building a Factor Return Prediction Model

Among the various machine learning algorithms, we select the Support Vector Regression (SVR) algorithm, as proposed by Cortes and Vapnik (1995), to predict market returns because of three reasons. First, SVR is the supervised machine learning model, where the machine learning operates based on example input-output pairs. Therefore, we expect to find the best combination of signals to predict market returns. Second, SVR is built based on the concept of the Support Vector Machine (SVM), where the relationship between predictive signals and outcome factor returns is not necessary to be set as a linear relationship. If $\{(x_1, y_1), (x_2, y_2), ..., (x_t, y_t)\}$, where $x_t \in X \subset R, y_t \in Y \subset R,\ t = 1, ..., T$ are the training data and T is the total number of training samples, the SVR function can be specified as

$$f(x) = a^{'}\phi(x) + b, \tag{17}$$

where $\phi(x)$ is the nonlinear kernel function that maps the input data vector x into a feature space y. Specifically, we use the Gaussian radial basis (RBF) kernel (a set of mathematical functions that takes data as input and transforms it into the required form) to allow non-linearity. In order to obtain a and b, the following regularized risk function must be minimized:

$$\begin{aligned} \min R(f) &= C\frac{1}{T}\sum_{t=1}^{T} L_\varepsilon(y_t, f(x_t)) + \frac{1}{2}||a||^2 \\ \text{s.t.} \quad & |y_t - f(x_t)| \le \varepsilon \\ & \varepsilon \ge 0, \end{aligned} \tag{18}$$

where $L_\varepsilon = \begin{cases} 0\ if\ |y_t - f(x_t)| < \varepsilon \\ |y_t - f(x_t)| - \varepsilon\ otherwise. \end{cases}$

C is the regularization parameter and ε specifies the epsilon-tube within which no penalty is associated in the training loss function with points predicted within a distance epsilon from the actual value. We set C equal to 1 and ε equal to 0.1 for the SVR model. y_t is the actual value at time t, $f(x_t)$ is the predicted value at the same period, and L_ε is ε-sensitive loss function, which identifies the predicted values within a distance epsilon from the actual value y_t.

Last but not least, SVR is unlikely to occur overfitting problems since the solution SVR may be global optimum while other neural network models may fall into a local optimal solution Kim (2003).

5.2 Exploiting Predictive Signals in Factor-based Optimization

From the large number of predictors introduced in the prior literature, we select 20 signals from the market and macro information. Since we consider the single-factor model, all of the signals are predictive measures for future market returns, especially downside market crashes. The list of selected signals is described in Table 3 with references.

[Insert Table 3 here.]

Using predictive signals, we predict the market returns using the SVR algorithm. The usual data sampling method in machine learning is to randomly split entire samples into training data (in-sample) and test data (OOS) without considering time order. Instead, our training sample is a very specific sample, which preserves the order of realized observations. Our training sample increases over time because we predict market returns by accumulating signals from January 2000 to the date of estimation. Therefore, these are essential distinctions from other possible sampling methods to closely meet investors' interests.

We then put the predicted market returns to calculate the expected returns of individual stocks. as follows:

$$\mu_{i,t+1}^{factor,ML} = r_{f,t} + \hat{\beta}_i \hat{r}_{m,t+1}, \tag{19}$$

where $\hat{r}_{m,t+1}$ is the future market returns at time $t+1$ predicted at time t.

Therefore, the difference between the corresponding section is that we predict the expected return of individual stocks by estimating not only the beta but also the market rate return, while the previous section predicts the expected return of individual stocks using the estimated beta and historical market return.

5.3 Performance Comparison

In this section, we compare the portfolio performance within factor-based models. In Table 1, SRML refers to a single-factor model with predicted market returns via the SVR algorithm. In this table, we cannot see the big difference between SRML and SRF1 models with respect to the diversification aspect. Furthermore, in-sample Sharpe ratio of SRF1 is 23% higher than that of SRML (1.301 vs. 1.058). However, when we see the ex-post statistics in Table 2, the results are exactly the opposite. In panel (a), the mean of SRML is higher, and the standard deviation of SRML is lower than the counterpart of SRF1. Therefore, we observe that the ex-post Sharpe ratio of SRML is 13% higher than that of SRF1. Furthermore, the MDD of SRML is lower than the MDD of SRF1. These results imply that the machine learning algorithms provide better guidance to investors from an asset allocation perspective.

6 CONCLUSION

We devise a method using a factor-based model to improve the mean-variance optimization model's poor performance. In addition, we point out the problem of using historical returns (low OOS performance) and suggest a factor-based model approach to address this problem.

This study uses the single-factor model, the basic model among factor models, and shows that the performance is improved by applying two methods. First, since the average historical returns have estimation errors, we minimize this error by estimating the expected returns of individual stocks using market beta estimated by a single-factor model. At this time, the alpha that the factor model does not explain is excluded from the expected return. Second, we re-estimate expected returns using estimated expected market returns via the machine-learning algorithm as well as beta.

The portfolio, constructed in this way, shows better performance in OOS than the historical model (MaxHist) and other benchmarks (DJI, MinVar, or MaxDr) with a small maximum drawdown and a large Sharpe ratio.

Therefore, one of our contributions is finding a new approach to apply machine learning algorithms to the portfolio optimization process. In addition, this study is meaningful in finding other empirical evidence of the problems with estimating the future return from historical return.

Future research can be expanded in two aspects. In the first aspect, we will verify that the results are consistent by looking at various asset classes. For example, we expand our samples with the S&P 500 or verify them with assets in other countries. In the second aspect, we will use the multi-factor model to examine whether the portfolio's performance can improve further. For example, since it is well-known that the multi-factor model, such as the Fama-French 5-factor model, can better predict stock returns than the single-factor model, we will test whether it performs better. Lastly, we can use a multi-factor model to predict factor returns such as SMB or HML by reflecting firm-specific information to the machine learning algorithm. Since the information set for a market return prediction in a single-factor model is constrained to market and macro variables, as illustrated in Table 3, we can expand our information set for predicting other factor returns, leading to better performance.

Table 1: **In-Sample Statistics**

This table shows the in-sample statistics of each portfolio. Panel (a) shows the in-sample Sharpe ratio as defined in Equation (16), Panel (b) shows the concentration measure, Herfindahl-Hirschman Index, as defined in Equation (13), and Panel (c) shows a diversification ratio, as defined in Equation (14). On the other hand, column names indicate the optimization model. SRHist (SRF1) refers to Sharpe ratio maximization using the historical model (factor-based model). SRML refers to Sharpe ratio maximization using the factor-based model, where the market returns are predicted by a machine-learning algorithm called the Support Vector Regression (SVR). MinVar refers to the minimum-variance optimization, and MaxDR refers to the maximum-diversification ratio optimization.

	SRHist	SRF1	SRML	MinVar	MaxDR
count	150	150	150	150	150
mean	3.652	1.301	1.058	1.526	1.352
std	1.378	1.197	1.909	1.326	1.386
min	0.936	-1.248	-8.259	-1.606	-1.716
25%	2.602	0.382	0.522	0.637	0.347
50%	3.688	1.224	1.502	1.434	1.361
75%	4.320	2.215	2.217	2.281	2.203
max	7.647	4.145	4.218	5.062	4.767

(a) In-Sample Sharpe Ratio

	SRHist	SRF1	SRML	MinVar	MaxDR
count	150	150	150	150	150
mean	0.333	0.147	0.143	0.199	0.120
std	0.196	0.100	0.085	0.119	0.047
min	0.094	0.066	0.061	0.076	0.061
25%	0.202	0.099	0.096	0.131	0.096
50%	0.276	0.113	0.117	0.160	0.108
75%	0.370	0.152	0.158	0.229	0.130
max	1.000	0.769	0.660	0.728	0.348

(b) Herfindahl–Hirschman Index

	SRHist	SRF1	SRML	MinVar	MaxDR
count	150	150	150	150	150
mean	1.443	1.589	1.591	1.639	1.867
std	0.283	0.288	0.298	0.335	0.342
min	1.000	1.066	1.088	1.064	1.269
25%	1.266	1.385	1.376	1.429	1.662
50%	1.415	1.574	1.582	1.609	1.849
75%	1.595	1.758	1.735	1.822	2.031
max	2.376	2.425	2.654	2.698	2.967

(c) Diversification Ratio

Table 2: **Out-of-Sample Statistics**

This table shows the out-of-sample (OOS) statistics of each portfolio. Panel (a) shows the statistics of ex-post portfolio returns, Panel (b) shows the maximum drawdown, as defined in Equation (15), and Panel (c) shows a OOS Sharpe ratio, as defined in Equation (16). On the other hand, column names indicate the optimization model. SRHist (SRF1) refers to Sharpe ratio maximization using the historical model (factor-based model). SRML refers to Sharpe ratio maximization using the factor-based model, where the market returns are predicted by a machine-learning algorithm called the Support Vector Regression (SVR). MinVar refers to the minimum-variance optimization, MaxDR refers to the maximum-diversification ratio optimization, and DJI refers to the Dow Jones Index.

	SRHist	SRF1	SRML	MinVar	MaxDR	DJI
count	3,146	3,146	3,146	3,146	3,146	3,146
mean	0.066	0.055	0.067	0.042	0.049	0.043
std	1.182	1.013	0.997	0.824	0.995	1.082
min	-10.961	-8.756	-7.643	-7.643	-9.173	-12.927
25%	-0.461	-0.370	-0.356	-0.332	-0.387	-0.369
50%	0.092	0.087	0.090	0.063	0.075	0.060
75%	0.622	0.554	0.556	0.449	0.529	0.534
max	10.773	7.291	8.681	7.312	8.665	11.365

(a) Ex-Post Returns

	SRHist	SRF1	SRML	MinVar	MaxDR	DJI
count	2,896	2,896	2,896	2,896	2,896	2,896
mean	0.035	0.032	0.027	0.026	0.034	0.037
std	0.036	0.036	0.033	0.028	0.039	0.046
min	0.000	0.000	0.000	0.000	0.000	0.000
25%	0.006	0.004	0.003	0.004	0.004	0.005
50%	0.024	0.020	0.015	0.016	0.020	0.021
75%	0.055	0.050	0.039	0.039	0.051	0.053
max	0.285	0.278	0.209	0.191	0.294	0.371

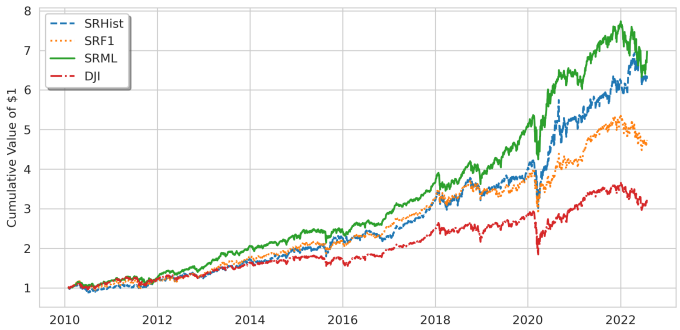
(b) Maximum Drawdown

	SRHist	SRF1	SRML	MinVar	MaxDR	DJI
count	3,021	3,021	3,021	3,021	3,021	3,021
mean	1.143	1.243	1.402	1.145	1.186	1.066
std	1.024	1.206	1.170	1.151	1.214	1.085
min	-1.376	-1.816	-1.820	-2.072	-2.378	-1.929
25%	0.528	0.303	0.651	0.275	0.428	0.290
50%	1.103	1.263	1.375	1.054	1.030	1.123
75%	1.612	1.831	2.021	1.764	1.735	1.882
max	5.929	5.916	5.844	5.443	6.119	4.446

(c) OOS Sharpe Ratio

Table 3: **List of Signals for Market Crash**

Num	Category	Paper	Variable	Object
1	Market	Lleo and Ziemba (2012)	BSEYD	BSEYD (Bond Stock Earning Yield Differential) is the spread between bond yields and stock yields. Bond Yield uses the T-Bond 10 YR Constant (DGS10), Stock Yield uses the reciprocal of CAPE10 (1/CAPE10), and CAPE10 is the PE Ratio created by Shiller, which is adjusted for inflation with the current index (P) as CPI. It is calculated by dividing by the average earnings over 10 years (E10).
2	Market	Lleo and Ziemba (2012)	Ln(BSEYD)	The logarithm of the BSEYD.
3	Market	Nyberg (2013)	ΔD/P	The first difference value of the dividend-price ratio (dividend yield) of the market (S&P 500) index.
4	Market	Nyberg (2013)	ΔE/P	The first difference value of the earning-price ratio of the market (S&P 500) index.
5	Market	Kole and van Dijk (2017)	π^{MS}	Transition probability to bear market using the Markov switching model, where a bear market is defined as the market under a high volatility regime. Since the status we are interested in is whether that market is under bear market at time t, this value is the sum of the probability that the phase transitions from bull to bear and the probability that the phase remains bear from bear.
6	Market	Chen and Vincent (2016)	MOM	12-month momentum of the market (S&P 500) index.
7	Market	Chen and Vincent (2016)	MOM_MA	Moving average of market momentum. This variable measures the momentum of the current index relative to the average market index.
8	Market	Baker and Wurgler (2006)	CEFD	Closed-end fund discount (Sentiment Components).
9	Market	Baker and Wurgler (2006)	NIPO	Number of IPO. IPO volume. (Sentiment Components)
10	Market	Baker and Wurgler (2006)	RIPO	First-day returns on IPO (Sentiment Components).
11	Market	Baker and Wurgler (2006)	PDND	Price of dividend stock to non-dividend stock (Sentiment Components). This measure is the value-weighted dividend premium following Baker and Wurgler (2004).
12	Market	Baker and Wurgler (2006)	EQIS	Equity issuance ratio (Sentiment Components). The total volume of equity issues over the prior twelve months divided by the total volume of equity and debt issues over the prior twelve months.
13	Market	Baker and Wurgler (2006)	SENT	Investor Sentiment. It is a sentiment index in Baker and Wurgler (2006) based on first principal component of five (CEFD, NIPO, RIPO, PDND, EQIS) standardized sentiment proxies.
14	Market	Baker and Wurgler (2006)	SENT_ORTH	Investor Sentiment (Orthogonalized). It is orthogonalized SENT with respect to a set of six macroeconomic indicators (industrial production index, nominal durables consumption, nominal nondurable consumption, nominal services consumption, NBER recession indicator, employment, and CPI)
15	Macro	Nyberg (2013)	ΔCPI	Rate of change of CPI (consumer price index for all urban consumer: all Items in U.S. city average; CPIAUCSL), or inflation rate.
16	Macro	Nyberg (2013)	ΔINDPRO	Rate of change of the industrial production index.
17	Macro	Nyberg (2013)	ΔUNRATE	Rate of change of of the empoloyment rate.
18	Macro	Nyberg (2013)	TS	Term spread is the difference between short-term and long-term interest rates, where the 3-Month Treasury Bill and the 10-Year Government Bond ared used for each interest rate.
19	Macro	Zouaoui *et al.* (2011)	CCI_ORTH	Orthogonalized Consumer Confidenc Index (OCCI). The Consumer Confidence Index, which is a collection of data from direct consumer surveys conducted by Michigan University every month, is used as the direct investment sentiment. The stronger the investment sentiment, the higher the market crash is expected. We use CCI, the consumption sentiment of individual investors, as a direct proxy for predicting market crash. However, since CCI information includes not only behavioral factors (such as market beliefs) that affect individual investors, but also Macroeconomic factors, Macroeconomic factors are removed. Therefore, OCCI is the residual of the regression analysis that is not explained by Macroeconomic variables. The term spread (difference between 10-year US Treasury bonds and 3-month US Treasury bonds), the private credit growth rate, the industrial production index growth rate, and the consumption growth rate of durable goods/non-durable goods/services are used as Macroeconomic variables.
20	Macro	Zouaoui *et al.* (2011)	ΔCREDIT	U.S. Gross Domestic Credit (% GDP). It is provided quarterly and is calculated as the first difference of the quarterly data.



This figure shows the ex-post performance of each portfolio from January 2010 to June 2022. For brevity, we plot four models out of six. SRHist (SRF1) refers to Sharpe ratio maximization using the historical model (factor-based model). SRML refers to Sharpe ratio maximization using the factor-based model, where the market returns are predicted by a machine-learning algorithm called the Support Vector Regression (SVR). DJI refers to the Dow Jones Index.

Figure 1: **Ex-Post Portfolio Performance**

REFERENCES

Baker, M. and Wurgler, J. (2006) Investor sentiment and the cross-section of stock returns, *The Journal of Finance*, **61**, 1645–1680.

Best, M. J. and Grauer, R. R. (2015) On the Sensitivity of Mean-Variance-Efficient Portfolios to Changes in Asset Means: Some Analytical and Computational Results, *The Review of Financial Studies*, **4**, 315–342.

Bianchi, M. L. and Tassinari, G. L. (2020) Forward-looking portfolio selection with multivariate non-gaussian models, *Quantitative Finance*, **20**, 1645–1661.

Chamberlain, G. (1983) A characterization of the distributions that imply mean—variance utility functions, *Journal of Economic Theory*, **29**, 185–201.

Chen, Y. and Wang, X. (2015) A hybrid stock trading system using genetic network programming and mean conditional value-at-risk, *European Journal of Operational Research*, **240**, 861–871.

Chen, Y.-t. and Vincent, K. (2016) The role of momentum, sentiment, and economic fundamentals in forecasting bear stock market, *Journal of Forecasting*, **35**, 504–527.

Chopra, V. K. and Ziemba, W. T. (1993) The effect of errors in means, variances, and covariances on optimal portfolio choice, *The Journal of Portfolio Management*, **19**, 6–11.

Cortes, C. and Vapnik, V. (1995) Support-vector networks, *Machine Learning*, **20**, 273–297.

Feng, G., He, J., Polson, N. G. and Xu, J. (2018) Deep learning in characteristics-sorted factor models.

Garlappi, L., Uppal, R. and Wang, T. (2006) Portfolio Selection with Parameter and Model Uncertainty: A Multi-Prior Approach, *The Review of Financial Studies*, **20**, 41–81.

Green, R. C. and Hollifield, B. (1992) When will mean-variance efficient portfolios be well diversified?, *The Journal of Finance*, **47**, 1785–1809.

Grossman, S. J. and Zhou, Z. (1993) Optimal investment strategies for controlling drawdowns, *Mathematical Finance*, **3**, 241–276.

Gu, S., Kelly, B. and Xiu, D. (2021) Autoencoder asset pricing models, *Journal of Econometrics*, **222**, 429–450, annals Issue:Financial Econometrics in the Age of the Digital Economy.

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

Jorion, P. (1985) International portfolio diversification with estimation risk, *The Journal of Business*, **58**, 259–278.

Kempf, A., Korn, O. and Saßning, S. (2014) Portfolio Optimization Using Forward-Looking Information*, *Review of Finance*, **19**, 467–490.

Kim, K.-J. (2003) Financial time series forecasting using support vector machines, *Neurocomputing*, **55**, 307–319, support Vector Machines.

Kole, E. and van Dijk, D. (2017) How to identify and forecast bull and bear markets?, *Journal of Applied Econometrics*, **32**, 120–139.

Kostakis, A., Panigirtzoglou, N. and Skiadopoulos, G. (2011) Market timing with option-implied distributions: A forward-looking approach, *Management Science*, **57**, 1231–1249.

Lleo, S. and Ziemba, W. T. (2012) Stock market crashes in 2007–2009: were we able to predict them?, *Quantitative Finance*, **12**, 1161–1187.

Michaud, R. O. (1989) The markowitz optimization enigma: Is 'optimized' optimal?, *Financial Analysts Journal*, **45**, 31–42.

Nyberg, H. (2013) Predicting bear and bull stock markets with dynamic binary time series models, *Journal of Banking & Finance*, **37**, 3351–3363.

Roman, D., Darby-Dowman, K. and Mitra, G. (2007) Mean-risk models using two risk measures: a multi-objective approach, *Quantitative Finance*, **7**, 443–458.

Zouaoui, M., Nouyrigat, G. and Beer, F. (2011) How does investor sentiment affect stock market crises? evidence from panel data, *Financial Review*, **46**, 723–747.