DALLASFED

Occasional Paper

Hedge Fund Return Prediction and Fund Selection: A Machine-Learning Approach

Jiaqi Chen, Wenbo Wu, and Michael L. Tindall

Federal Reserve Bank of Dallas
Financial Industry Studies Department
Occasional Paper 16-04
November 2016

Hedge Fund Return Prediction and Fund Selection: A Machine-Learning Approach

by

Jiaqi Chen[1], Wenbo Wu[2], and Michael L. Tindall[1]

Abstract:

A machine-learning approach is employed to forecast hedge fund returns and perform individual hedge fund selection within major hedge fund style categories. Hedge fund selection is treated as a cross-sectional supervised learning process based on direct forecasts of future returns. The inputs to the machine-learning models are observed hedge fund characteristics. Various learning processes including the lasso, random forest methods, gradient boosting methods, and deep neural networks are applied to predict fund performance. They all outperform the corresponding style index as well as a benchmark model, which forecasts hedge fund returns using macroeconomic variables. The best results are obtained from machine-learning processes that utilize model averaging, model shrinkage, and nonlinear interactions among the factors.

Key words: Hedge fund selection, hedge fund return prediction, machine learning, the lasso, random forest, gradient boosting, deep neural networks.

[1] Financial Industry Studies Department, Federal Reserve Bank of Dallas. The views expressed herein are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System.

[2] Department of Operations and Business Analytics, University of Oregon.

1. Introduction

In recent years, hedge funds have found it increasingly difficult to produce high returns. This makes hedge fund selection a critical issue for investors. Selection is difficult because hedge funds are heterogeneous using many distinctive strategies and investing in a variety of asset classes with vastly different return characteristics. Various ways have been proposed to improve hedge fund selection; a complete survey of those methods is beyond the scope of this paper. An investor could look at historical hedge fund returns and try to forecast future returns based on macroeconomic conditions. For instance, Jagannathan, Malakhov, and Novikov (2010) and Avramov, Barras, and Kosowski (2013) implement conditional strategies to forecast returns using macroeconomic variables, and they find such strategies generate very good performance.

Another more intuitive and easily implemented approach is to rank the funds based on certain risk factors and to construct different portfolios of funds from the rankings of these factors (De Souza and Gokcan, 2004; Pedersen, 2013; Molyboga, Beak, and Bilson, 2014). This approach has been employed to test fundamental theories in asset pricing, to study pricing anomalies, and to identify profitable investment strategies. The underlying premise is that the expected returns of an asset class are related to a certain factor, and the asset group ranked highest with regard to that factor has the highest likelihood of outperforming other groups. Using this framework, there is no forecasting equation, but the procedure can be casted as a nonparametric estimator (Cattaneo, Crump, Farrell, and Schaumburg, 2016).

Our paper differs from both methods because we build cross-sectional models and select funds based on forecasts of return from those models. It relies to some extent on models like the Fama-French-Carhart (FFC hereafter) factor model (Carhart, 1997). Instead of forecasting returns of an individual fund using macroeconomic variables, we focus on forecasting returns using characteristic factors using cross-sectional methods. We treat forecasting as a supervised machine-learning problem and utilize both linear and nonlinear machine-learning techniques to construct our forecasts of returns. Portfolios are formed according to the rankings of the forecasted returns, and we find that these portfolios generate robust and superior performance results across major hedge fund style categories. Our forecast methods deviate from traditional econometric approaches and are more in line with the suggestions made in Varian (2014) who argues that data analysis in statistics and econometrics can be divided into four categories: 1) prediction, 2) summarization, 3) estimation, and 4) hypothesis testing. The machine-learning methods we employ are mainly concerned with the first category: the accuracy of out-of-sample prediction.

The next section discusses the data used in the paper. Following that, we present our machine-learning portfolio-based methodology and an alternative method using macroeconomic variables, which we use as a benchmark for comparison. After that, we present a description of the factors used in our analysis, and then we present the alternative machine-learning methods used in our analysis. Empirical results are presented in the following section. The final section presents conclusions.

2. Hedge Fund Return Data

Hedge fund monthly return data come from the Hedge Fund Research, Inc. (HFR) database. There are four major HFR styles: equity, event-driven, macro, and relative value. We evaluate individual funds in the HFR database grouped within the four styles. The period for evaluating portfolio performance is January 1996 to December 2015. There are 7,324 live funds in the HFR database and 16,304 dead, or graveyard, funds. There are considerable survivorship and backfill biases in the data. To mitigate those problems, we combine both the live and graveyard databases as follows:

1. We combine the funds from both the graveyard and live database and sort them into equity, event-driven, macro, and relative-value style classifications.
2. At each time t we find hedge funds with returns reported at that particular date. They are our live funds at time t regardless of whether the reported returns come from the live or graveyard database. That is, funds in the graveyard data that were live at time t are treated as live as of that date.
3. Our analytical method requires that each fund has a continuous reported history, which we set at 24 months. The 24-month requirement could be flexible but it has to be longer than the sum of look-back period M and the look-forward period K.
4. Once we have built the forecast model at time t, the funds are sorted by their forecast returns for the next K months into 10 groups with an equal number of funds in each group. We refer to these as our "decile portfolios."
5. We hold the equally weighted portfolios until time $t + K$. If a fund is dropped out of the database, we assume the return for that portion of asset allocation to be 0 for the remaining part of the K months.

Steps 1–5 greatly reduce unwanted survivorship and backfill biases. For example, for macro funds in January 2000, there are only 74 funds in the live database, but adding funds from the graveyard there are 416 funds in the month. Of these, 294 of them had a continuous reporting history of no fewer than 24 months as of January 2000. So our model uses those 294 funds in that month. Step 5 gives us a more conservative and realistic estimation of the portfolio returns as it would take time for an investor to reinvest funds if a hedge fund has returned proceeds to investors.

The hedge fund indexes used in the paper are the four HFR "HFRI" style indexes (the HFR indexes hereafter) for equity, event-driven, macro, and relative-value hedge funds published by Hedge Fund Research, Inc.

3. Methodology

Portfolio sorting based on certain characteristics is an important tool in empirical finance. In practice, a set of securities is typically ranked according to a risk factor under scrutiny over a certain look-back window, and long and short positions in securities are taken based on the ranking. Numerous factors have been studied in the academic research. Harvey, Liu, and Zhu (2016) have identified more than 200 such factors in the literature.

There are important issues with regard to factor-ranking procedures. One of these is the question of whether the risk premium of a particular factor is stable over time. For instance, Fama and French (1992) found that their value factor had a positive risk premium over the long term, but we find that the value premium estimated as the annual return difference between the Russell 1000 Value Index and Russell 1000 Growth Index has been negative in all five-year rolling windows from 2009 to 2015.

A natural extension of the method is to include multiple risk factors in the portfolio formation process, but this raises another issue. Typically, we may want to identify securities ranked in the top group in all the risk factors. However, with a large number of factors, we run into the high-dimensional problem that the securities/funds would be scattered sparsely in an N-dimensional space where N is the number of factors we want to combine. It may be impractical to find enough securities in the top rankings of all the factors to form a diversified portfolio. One can devise an ad hoc method where a composite score is used to combine the factors into a ranking. However, there may be no statistical principle or financial theory dictating how to form that composite score and build the best portfolio because the rank combination approach ignores interactions between the factors. A study by Bali, Brown, Murray, and Tang (2016) on "lottery" stocks, i.e., stocks that have a high return in one day in the previous month, and high-beta stocks illustrates the problem. High-beta stocks underperform low-beta stocks, and lottery stocks underperform their counterparts. One would expect then that low-beta nonlottery stocks would offer the best performance. However, Bali et al find out that high-beta nonlottery stocks outperform low-beta nonlottery stocks by about 0.6 percent per month when sorted by lottery and beta properties sequentially. Thus, an underlying interaction between the factors plays a crucial role in determining performance.

Given these difficulties, it is useful to examine this approach from a machine-learning perspective. Essentially, the approach described above falls into the realm of "unsupervised learning." This is a type of machine-learning algorithm used to draw inferences from data sets without access to the outcome response to provide correct answers for each observation (Hastie, Tibshirani, and Friedman, 2009). It is difficult to ascertain the correctness of the inferences drawn from the algorithms, and one typically relies on heuristic argument to motivate the algorithm and evaluate the quality of results. In the case of factor-sorted cross-sectional analysis, the heuristic argument is that the factor carries a significant positive risk premium in either a statistical or economic sense. Then, if we cluster all the securities into similar groups based on their rankings, we could effectively evaluate the risk premium hypothesis and the clustering algorithm simultaneously. For most of the studies dealing with a single factor, the grouping by rank is just a one-dimensional clustering algorithm where the distance between data points is the difference between their ranks. The number of the clusters is fixed beforehand with equal number of members in each cluster. The multi-factor case elevates the clustering problem into high dimensions. We can use readily available clustering algorithms like K-means or hierarchical clustering to form the clusters/groups, and we can also use methods like principal component analysis and independent component analysis to reduce the number of dimensions to a manageable extent. However, after applying unsupervised learning techniques, we are still facing the heuristic problem of finding the cluster with the best performers. The clustering algorithms could only tell us how to group the securities based on their similarities in terms of the factors, and it may be difficult to predefine which group is the leading group.

There is another branch in machine learning called "supervised learning," which directly targets the prediction of one or more responses for a given set of input or predictor variables. Regression is a familiar, if simple, example of supervised learning, and we can use the FFC four-factor model to examine how our fund selection machine-learning framework is related to more traditional cross-sectional studies. Based on the four-factor model, a fund's excess return can be decomposed into the excess return of four factors: market return, value, size, and momentum as in the following equation:

$$R_t = \alpha + \beta_M R_{M,t} + \beta_B R_{B,t} + \beta_C R_{C,t} + \beta_W R_{W,t} + \varepsilon_t \tag{1}$$

where R_t is the monthly excess return of a particular fund, $R_{M,t}$ is the monthly excess return of a market index, $R_{B,t}$ is the monthly excess return of a portfolio that is long (short) high (low) book-to-market stocks, $R_{C,t}$ is the monthly excess return of a portfolio that is long (short) small (large) capitalization stocks, and $R_{W,t}$ is the monthly excess return of a portfolio that is long the stocks with the highest returns (i.e., the returns in the top three deciles) over the previous 12 months and short the stocks with the lowest returns over the previous 12 months.

With a slight modification, we can transform equation (1) into a forecasting model for a fund's returns. In the forecast model, a fund's return at time $t+1$ can be expressed as a conditional expectation $E(R^i_{t+1}|\emptyset^i_t)$ where $R^i_{t+1}$is the return at time $t+1$ for fund i, and $\emptyset^i_t$ is all the information pertaining to fund i at time t. Following Haugen and Baker (1996), the FFC four-factor model can be rewritten as:

$$E\left(R^i_{t+1}\middle|\emptyset^i_t\right) = \alpha + \beta^i_{M,t} R_{M,t} + \beta^i_{B,t} R_{B,t} + \beta^i_{C,t} R_{C,t} + \beta^i_{W,t} R_{W,t} \tag{2}$$

where $\emptyset_t^i$ is the information represented by the fund-specific betas. The model can be estimated through either a cross-sectional regression or a Fama-MacBeth process. Using this model at a point in time t across the hedge funds in our database, we can generate cross-sectional return predictions for the funds in our database and pick up the top performing funds.

Besides linear regression, there are many machine-learning techniques widely employed to incorporate nonlinear effects to reduce forecast error by introducing bias and to utilize model averaging methods. Our goal is to use alternative machine-learning methods to build a flexible framework that is capable of combining the strengths of all the factors and to construct our dynamic forecast models without relying on heuristic arguments to define the weights in the composite score. This holds the potential to improve model performance and to build a general framework that could easily be modified to incorporate new risk factors. Our method can be viewed as a supervised learning problem based on the following equation:

$$E\left(R_{i,t+K}^K \middle| \emptyset_t^i\right) = f_t(R_t^i, R_{t-1}^i \ldots, x_{1,t}^i, x_{2,t}^i \ldots) \tag{3}$$

where $1 + R_{i,t+K}^K = \prod_{j=1}^K (1 + R_{t+j}^i)$ is the cumulative return for fund i from time $t+1$ to $t+K$ and $f_t(R_t^i, R_{t-1}^i \ldots, x_{1,t}, x_{2,t} \ldots)$ is the linear/nonlinear forecast function using the past returns (R_t^i, R_{t-1}^i ...) up to time t and quantitative factors ($x_{1,t}^i, x_{2,t}^i$ ...) describing fund i at time t.

We use a rolling window to build our forecast model and evaluate its performance. Assuming a look-back period of M and a look-forward period of K, at time t the machine-learning algorithm is applied to the fund returns information from $t - M - K + 1$ to t. The output variable is the cumulative return from $t - K + 1$ to t while the input variables are only evaluated from $t - M - K + 1$ to $t - K$ in building each model. Basically, we are constructing a supervised-learning model using past information to forecast the future returns in the next K months. Applying the model on data from $t - M + 1$ to t, we have the forecast of return for each fund. Based on the forecast, we rank the funds into 10 equal-sized groups to form 10 equally weighted portfolios, i.e., the decile portfolios, and hold the 10 portfolios for the next K months at which time we rebalance. When rebalancing at the end of K months, we retrain the forecast model following the same procedures and form a new set of 10 portfolios. We set $M = 12$ so that such a short window can follow market dynamics more closely. K is set at 3, 6, and 12 months to simulate typical hedge fund redemption restrictions. Doing so, we establish 10 out-of-sample dynamic portfolios. Following this procedure, we only rebalance the portfolios every K months, generating for each trading strategy a single time series of monthly returns. We generate this single time series of returns following the approach used by Moskowitz, Ooi, and Pedersen (2010) in which the return at time t measures the average return across all portfolios at that time.

As a benchmark for comparison to our results, we use the macroeconomic hedge fund return forecast model proposed by Avramov, Barras and Kosowski (2013) (the ABK model hereafter) to construct decile portfolios. They constructed four univariate time-series models of illiquidity-adjusted hedge fund returns regressed, respectively, on the VIX, hedge fund industry net dollar inflow, the dividend yield of the S&P 500 index, and the yield differential between Moody's BAA- and AAA-rated bonds. For each of the four regressions, the conditional t-statistic, $t(\hat{u}_{i,t+1}(\mathrm{j}))$ is estimated for the predictor $z_{j,t}$ separately at time t as:

$$t(\hat{u}_{i,t+1}(\mathrm{j})) = \frac{\hat{u}_{i,t+1}(\mathrm{j})}{se(\hat{u}_{i,t+1}(\mathrm{j}))} \tag{4}$$

where $\hat{u}_{i,t+1}$(j) is the conditional mean and $se\left(\hat{u}_{i,t+1}(\mathrm{j})\right)$is the estimated standard error from the regression equation of $\hat{u}_{i,t+1}(\mathrm{j}) = \hat{u}_i + \hat{b}_{i,j} z_{j,t}$. Their indicator is the average of the four individual $t(\hat{u}_{i,t+1}(\mathrm{j}))$, and their investment strategy is to hold the top decile of funds with the highest value of the indicator. In our estimation of equation (4), we use a 24-month rolling window.

4. Risk Factors

Our risk factors include a variety of measures including cumulative monthly returns, Sharpe ratios, skewness of returns, and kurtosis of returns. Hedge funds returns may exhibit strong serial correlation, and Getmansky, Lo, and Makarov (2004) argue this is caused by illiquidity in asset markets rather than unexploited profit opportunities. Thus, we include autocorrelation factors, which offer information about the effect of a fund's illiquidity on its future returns. We also use the Ω ratio proposed by De Souza and Gokcan (2004) as a risk factor. It is defined as:

$$\Omega \text{ ratio } = \int_L^b (1 - F(r))dr \Big/ \int_a^L F(r)dr \quad (5)$$

where $F(r)$ is the cumulative distribution function of the returns, a and b are lower and upper limits of the range of return, and L is the target return threshold defining what is considered a gain versus a loss. We set $L = 0$ in our study.

Additionally, we employ the current drawdown of a fund measured as the percentage difference between the fund's most recent monthly unit value and its trailing high watermark. We reason that this factor could bear useful information; a fund may have incentive to increase portfolio risk if the fund is far below the high watermark given that the 2-20 fee structure resembles an out-of-money call option for the manager (Goetzmann, Ingersoll and Ross, 2003). The machine-learning algorithm determines dynamically whether such action would be beneficial for the fund.

Sun, Wang, and Zhen (2016) found that funds that perform well in adverse market conditions tend to be consistent winners over the long run. Thus, we include as a risk factor the average return of the fund in months when its corresponding HFR style index experiences a loss.

Another useful factor is the t-statistic of α, which measures the "significance" of fund performance. Molyboga, Beak, and Bilson (2014) propose to regress the return of a commodity trading advisor on its corresponding benchmark and define the t-statistic of alpha as:

$$t_\alpha = \frac{\alpha}{std(\alpha)} \quad (6)$$

A large positive value of t_α indicates that the manager is able to generate positive alpha consistently over time, while a large negative value of t_α indicates consistently poor performance. They report the top t_α funds significantly outperform bottom t_α funds out of sample. Following these ideas, we construct our t_α factor by regressing a fund's returns on those of its corresponding style index and calculate its alpha and the t-statistic from that regression.

With hedge fund data, we are limited in the types of factors we can employ. For instance, while we can define risk factors for stocks based on book and market values, such measures do not exist for hedge funds. In addition, there are gaps in hedge fund databases. For example, some funds do not update assets under management regularly, and so we do not incorporate this into our set of risk factors. The candidate factors on the right-hand side of equation (3) include some measures typically used to measure fund characteristics and some which are atypical. All variables are converted to normalized deviations from the mean. Exhibit 1 lists the factors used in our analysis. In the exhibit, the first nine factors are typical of those used in factor analysis while the remaining eight factors are atypical:

Exhibit 1. Factors used in our analysis

Factor	Symbol
1. One-month return	R1
2. Two-month cumulative return	R3
3. Six-month cumulative return	R6
4. Nine-month cumulative return	R9
5. 12-month average return	R12
6. Sharpe ratio calculated from the past 12 months of returns	SR
7. Standard deviation calculated from the past 12 months of returns	SD
8. Skewness calculated from the past 12 months of returns	Skew
9. Kurtosis calculated from the past 12 months of returns	Kurt
10. Ω ratio	Omega
11. Lag 1 autocorrelations calculated from the past 12 months of returns	ACF1
12. Lag 2 autocorrelations calculated from the past 12 months of returns	ACF2
13. Lag 3 autocorrelations calculated from the past 12 months of returns	ACF3
14. Fund alpha with respect to the corresponding HFR style index	Alpha
15. Alpha t-statistic	Alpha-t
16. Distance/drawdown from the high watermark in the past 12 months	DD
17. Average return of the fund in months when HFR style index experienced a loss	PerLoss

5. Machine-Learning Algorithms

Cross-Sectional OLS Regression

OLS regression for cross-sectional returns has a simple model structure as follows:

$$E\left(R_{i,t+K}^{K} \middle| \emptyset_t^i\right) = \alpha_t + \sum_{j=1}^{17} \beta_{j,t}\, x_{j,t}^i \tag{7}$$

where the superscript i indicates the fund-specific factors.

The Lasso

The lasso method is an automatic model building system based on a shrinkage method that has gained wide acceptance in the machine-learning field. The lasso method has found adherents in finance and economics (see Welsch and Zhou, 2007; Bai and Ng, 2008; and Wang and Zhu, 2008). Lasso addresses high-dimension problems by shrinking parameters that have been inflated in value by the effects of multicollinearity. It shrinks parameters by a constant factor at the limit to zero such that variables are eliminated from the initial equation. From a set of candidate variables, which may consist of all variables in the dataset, it selects a subset for inclusion in an equation, making it an automatic model building system. For example, for a linear regression $y_i = \beta_0 + \sum_{j=1}^{p} \beta_j * x_{ij} + \varepsilon_t, i = 1,2,3, \dots, n$, the lasso estimate is defined by:

$$\hat{\beta}^{lasso} = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j * x_{ij})^2 \tag{8}$$

$$\text{subject to } \sum_{j=1}^{p}|\beta_j| \leq T$$

where T is a tuning, or penalty, parameter. By making T sufficiently small, some of the coefficients are shrunk to exactly zero, which means that the lasso estimation method is also a variable selection method. The penalty parameter T is adjusted in small increments to minimize cross-validated mean square error. The lasso can also be written in the equivalent Lagrangian form (Hastie, Tibshirani, and Friedman, 2009).

$$\hat{\beta}^{lasso} = \underset{\beta}{\operatorname{argmin}}(\sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p}\beta_j * x_{ij})^2 + \lambda\sum_{j=1}^{p}|\beta_j|) \tag{9}$$

The key motivation for lasso is the bias-variance tradeoff in mean squared error (MSE). For a fitted model of $\hat{f}(x)$, where $f_0(x)$ is the true value of f(x) at point x:

$$MSE[\hat{f}(x)] = E[\hat{f}(x) - f_0(x)]^2 = Variance[\hat{f}(x)] + [E[\hat{f}(x)] - f_0(x)]^2 \tag{10}$$

By shrinking the estimator in lasso, $Variance[\hat{f}(x)]$ will be reduced. If the increase in bias, $[E[\hat{f}(x)] - f_0(x)]$, is relatively small, the forecast error could be reduced. In our study, the choice of the penalty parameter T is based on 10-fold cross-validation estimation of mean square error.

Random Forest

Random forest (RF hereafter) is an ensemble method built on a classification and regression tree (CART). The basic CART works by a recursive partitioning of the data that can be represented within a basis function framework (Berk, 2008). The basis functions are indicator variables determined by the best splits, which can be viewed as nonlinear and high-order functions of all the variables. It is a stagewise process that breaks the data into smaller and smaller pieces. The goal is to construct subsets of the data so that the values of the response variable in each subset fall within a narrow range. Defining the impurity of node as the within-node sum of squares for the response variable $i(\tau) = \sum(y_i - \overline{y(\tau)})^2$ at each split, the best split is chosen to maximize the change of impurity:

$$\Delta(s, \tau) = i(\tau) - i(\tau_L) - i(\tau_R) \tag{11}$$

where τ represents the "parent node" and τ_L and τ_R represent the two "daughter nodes." Through the whole process, each observation is placed in a terminal node and is then assigned the mean of that mode. The basic CART is prone to over-fitting, and Breiman (2001) proposes to use an ensemble approach called "random forest" to improve on the basic CART. Let N be the number of observations, and the following steps are followed to build an RF model for quantitative response variable:

1. Take a random sample of size N with replacement from the data.
2. Take a random sample without replacement of the predictors.
3. Construct the first CART partition of the data.
4. Repeat step 2 for each subsequent split until the tree is as large as desired.

5. Drop the out-of-bag data which were not selected in the first step down the tree. From these observations, a mean is calculated for each terminal node and such means would serve as the predicted values in the respective terminal nodes.
6. Repeat steps 1–5 a large number of times, e.g., we use 10,000 in our setup.
7. Using only the value assigned to each observation when that observation is not used to build the tree; we take the average of those values among all the trees as the predicted outcome from this random forest.
8. Variable importance can be computed using the shuffling approach. For a single tree, each time a given variable is used to define a splitting of the data, the reduction in the purity is recorded. Once the tree is complete, all the reductions are summed. The result is the error sum of squares reduction contributed by each predictor. Those sums are then averaged over all the trees for each predictor.

There are three major tuning parameters: 1) the terminal node size for the single tree, 2) the number of trees, and 3) the proportion of predictors to be sampled. We use a value of 5 for the terminal node size. We set the number of trees at 10,000. The proportion of the predictors to be selected for each iteration is set at 1/3.

With a random sample of input variables, the predictions made by each of the trees are more independent. As a result, averaging over a large number of independent trees can bring great improvement in prediction accuracy. Because of this, a random forest is able to work with high-dimensional data. Where predictors are highly correlated, one can use either one as the splitting variable. However, the two variables would not split the data in exactly the same way given the hierarchical structure in the regression trees. The two partitions might have large overlap while having unique content. If each of the two variables has an opportunity to be selected without competing against each other, we are able to explore a richer feature space. Overall, the predictor sampling and multiple tree averaging leads to shrink each predictor on the fitted values.

Gradient Boosting Model

The strength of the random forest model comes from two things: 1) the fitting functions are very flexible and can accommodate highly localized features of the data, and 2) averaging over the out-of-bag observations reduces the risk of over-fitting. Boosting is an alternative way of dealing with the flexibility issue. It gives observations responsible for highly localized features more weight in the fitting procedure. After a large number of iterations, which gives relatively more weight for the difficult-to-fit observation, one can combine the predictions from each iteration in a sensible way to reduce over-fitting. Basically it is able to take a "weak" learning algorithm and gradually boost it into a "strong" learning algorithm. In this paper, we will follow a specific approach called stochastic gradient boosting (GBM hereafter) to build our forecast model. For N observations and p variables, our procedure is as follows:

1. Initialize the forecast function $f_0(x)$ such that the constant c minimizes the loss function: $f_0(x) = \underset{c}{\operatorname{argmin}} \sum_{i=1}^{N} L(y_i, c)$. We use the quadratic loss function of mean square error in our setup.
2. For $m = 1,2, \ldots, M$ where M is the total number of iterations, we iterate over the steps 2a) to 2e) as follows:

   2a) Estimate the gradient for the N observations as:

   $r_{im} = -[\frac{\partial\left(L(y_i, f(x_i))\right)}{\partial f(x_i)}]_{f=f_{m-1}}$ which is just the forecast error of $y_i - f(x_i)$ in our case.

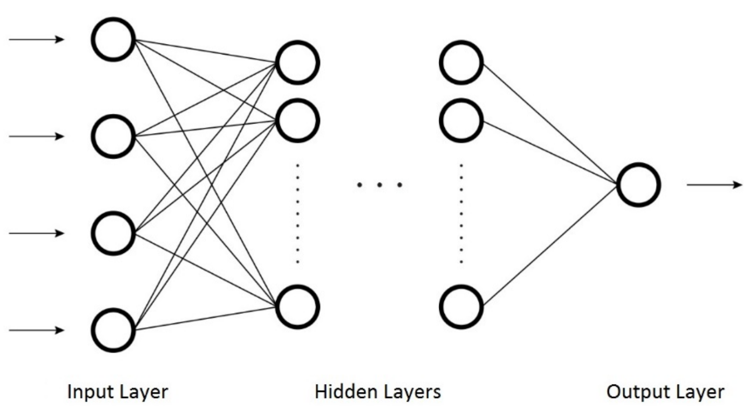
2b) Randomly select without replacement W*p cases from the whole data set where W should be smaller than the total number of observations.
2c) Fit a regression tree with J_m terminal nodes to the gradient obtained in 2a).
2d) For j= 1,2,…, J_m we calculate the optional terminal node prediction as
$\gamma_{im} = \underset{\gamma}{\operatorname{argmin}} \sum_{x_i \in R_{jm}} L(y_i, f_{m-1}(x_i) + \gamma)$ where R_{jm} is the space defining the terminal node from step 2c).
2e) Using the same sampled data, we can update $f_m(x) = f_{m-1}(x) + v * \sum_{j=1}^{J_m} \gamma_{im} I(x_i \in R_{jm})$ where v is shrinkage parameter controlling the learning rate.

3. Obtain the final output $f_M(x)$.

It has been shown (Ridgeway, 1999) that all the procedures within the generalized linear model can be properly boosted by the stochastic gradient model. The major tuning parameters of this algorithm include the sample size, the shrinkage parameter, the depth of interaction among variables, the minimum number of observations in each tree's terminal node, and the total number of iterations. We set the minimum number of observations in the terminal node to be 1 and sample all of the observations employing a shrinkage parameter of 0.02. We only optimize the interaction depth ranging from five to 15 based on five-fold cross-validation with the total number of iterations capped at 1,000.

Deep Neural Network

The deep neural network (DNN hereafter) approach, which is also known as deep learning, is composed of multiple levels of nonlinear operations. We employ feed-forward neural networks with multiple hidden layers. The diagram below shows such a network:



The computing units in neural networks are called neurons. These are represented by the cells in the diagram. The layer of cells on the left in the diagram is the input layer, and the cells within the input layer are called input neurons. The output layer contains a single output neuron, which is the forecast of the fund's future returns. The middle layer is called a hidden layer.

Let w_{jk}^l denote the weight for the connection link from the k^{th} neuron in the $(l-1)^{th}$ layer to the j^{th} neuron in the l^{th} layer. Similarly, let b_j^l denote the bias of the j^{th} neuron in the l^{th} layer and a_j^l denote the activation/output of the j^{th} neuron in the l^{th} layer. Mathematically, the neural network can be represented as:

$$u_j^l = \sum_k w_{jk}^l a_k^{l-1} \text{ and } a_j^l = \varphi(u_j^l + b_j^l) \quad (12)$$

where φ is the activation function, which can take various forms such as:

a) Threshold function: $\varphi(u) = \begin{cases} 1 \; if \; u \geq 0 \\ 0 \; if \; u < 0 \end{cases}$

b) Sigmoid function: $\varphi(u) = \frac{1}{1+e^{-\alpha u}}$ where α is the slope parameter

c) Tanh function: $\varphi(u) = \tanh(u)$

d) Rectifier function: $\varphi(u) = \max(0, u)$

The parameters are usually calculated through a back-propagation algorithm by estimating the partial derivatives of the loss function with respect to the weights and the biases. The loss function in our analysis is the quadratic loss function of mean square error. A detailed explanation of the back-propagation algorithm and its variations is beyond the scope of this paper, but Goodfellow, Bengio, and Courville (2016) provide comprehensive descriptions of the algorithms. Adding more layers to the network, we have more flexibility in approximating the true underlying response function, and in principle, we should have better performance. Typically, if one trains DNNs using the standard stochastic gradient descent approach with back-propagation, it is very unlikely that DNNs would have any advantage over simple shallow networks. This lack of performance can be attributed to the vanishing gradient problem (Glorot and Bengio, 2010). It arises from the intrinsic instability of the gradient descent approach in computing DNNs. The parameters associated with early or later layers become fixed in value during the training process, and the effective learning rate can differ by several magnitudes between layers. This intrinsic problem is part of the reason DNN methods did not become popular until work by Hinton, Osindero, and Teh (2006). They introduced Deep Belief Networks with a learning algorithm that greedily trains one layer at a time, exploiting an unsupervised learning algorithm for each layer called a Restricted Boltzmann Machine. Since then, related algorithms based on auto-encoders (Bengio, Lamblin, Popovici, and Larochelle, 2007) and other structures have been proposed. We employ the rectifier activation function because the vanishing gradient problem is less severe with it. The major tuning parameters are the number of layers and the number of neurons in each layer, the learning rates controlling parameters, the regularization parameters that add stability and improve generalization, the input drop-out ratio, which omits a fraction of the input features to improve generalization, the early stopping criteria, and the number of epochs, which determines how many times the dataset should be iterated.

There are many more tuning parameters in DNNs than there are in the previously mentioned machine-learning algorithms. We fix the learning rate to be 0.005 and stop the iterations early if the simple moving average of length 5 of the MSE does not improve. We use grid search under five-fold cross-validation with 100 epochs to choose the other tuning parameters as follows. We allow up to 10 layers with 10 neurons per layer in our DNN structures, the input dropout ratio is tuned to either 0.5 or 1, and the regularization parameters range from 10^{-4} to 10^{-7}.

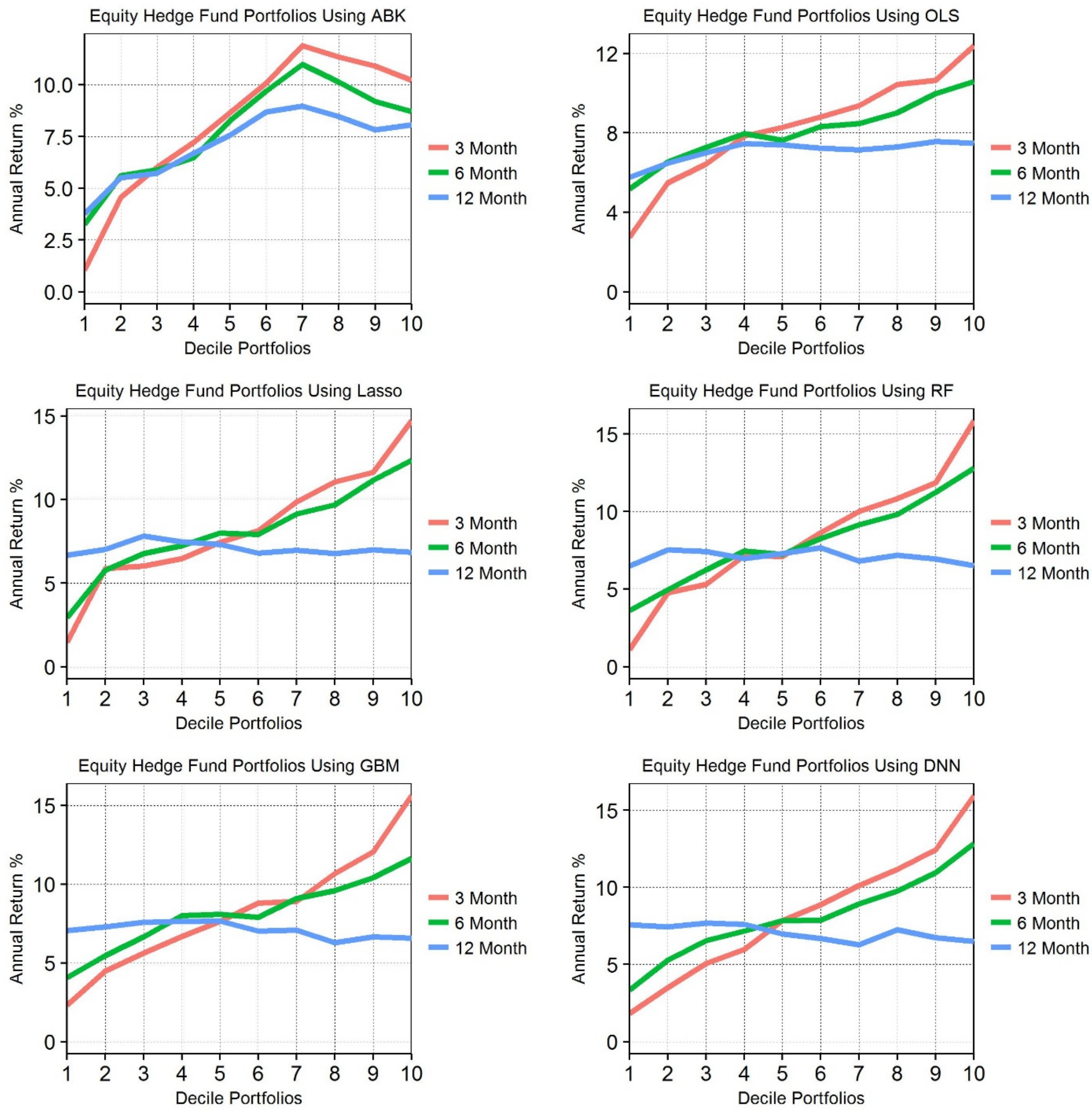
6. Empirical Results

We examine the performance of our hedge fund portfolios separately for each style.

Equity

Chart 1 shows the annual returns of our equity hedge fund portfolios ranked by the forecast from different methods and grouped by decile with a three-, six- and 12-month rebalancing cycle. Decile portfolio 10 contains the 10 percent of funds ranked highest in terms of the forecasted return, decile portfolio 9 contains the next 10 percent, and so on. The detailed summary statistics for such portfolios can be found in Appendix A.

Chart 1. Annual returns of the portfolios using ABK, OLS, Lasso, RF, GBM, and DNN forecasts



In general, the power of our machine-learning methods to select hedge funds with superior performance weakens with longer rebalancing periods and vanishes or even reverses with annual rebalancing. This is in agreement with the finding by Agarwal and Naik (2000) that persistence of return among hedge fund managers is short-term in nature. The poor performance with 12-month rebalancing may be due to this effect or because the factors in our analysis do not contain enough information to forecast returns that far in the future. With a three-month rebalancing cycle, where all the models have the best performance, the annual returns of the portfolios agree well with the decile rankings, which is a desirable feature. We note that the performance of decile portfolio 10 using the ABK model is worse than that of deciles 7–9, which means the t-statistic forecast of fund returns using macro variables is an inconsistent indicator of future fund returns.

Chart 2 and Table 1 compare the performance of the best-performing portfolio from each algorithm at a three-month rebalancing cycle with the HFR equity index to assess the benefit of our framework in fund selection. All five machine algorithms generate higher annual returns than the index and the ABK model, although some of them do not offer a better risk profile. This is understandable because the response variable in our forecast framework is the future return, and our approach was designed only to maximize returns. Among the five algorithms, DNN has the best total return. Its annual return of 15.95 percent is more than 80 percent higher than that of the HFR equity index at 8.72 percent. The DNN algorithm outperforms the RF and the GBM algorithms marginally despite its much more complicated structure. One possible reason for this is that the number of observations is very limited; DNN operates much better with larger samples.

Chart 2. The cumulative returns of the six best-performing equity hedge fund portfolios and the HFR equity index

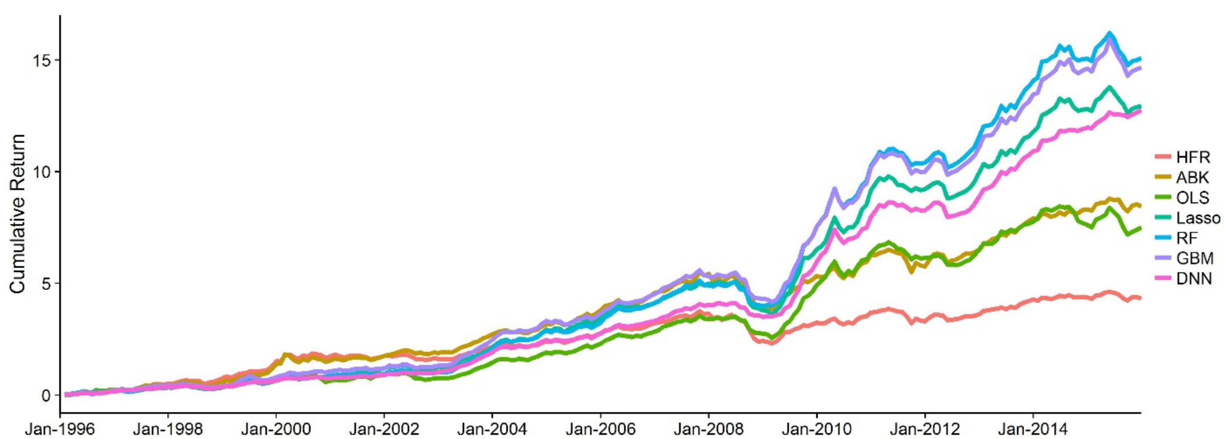


Table 1. Monthly return statistics of the six best-performing equity hedge fund portfolios and the HFR equity index

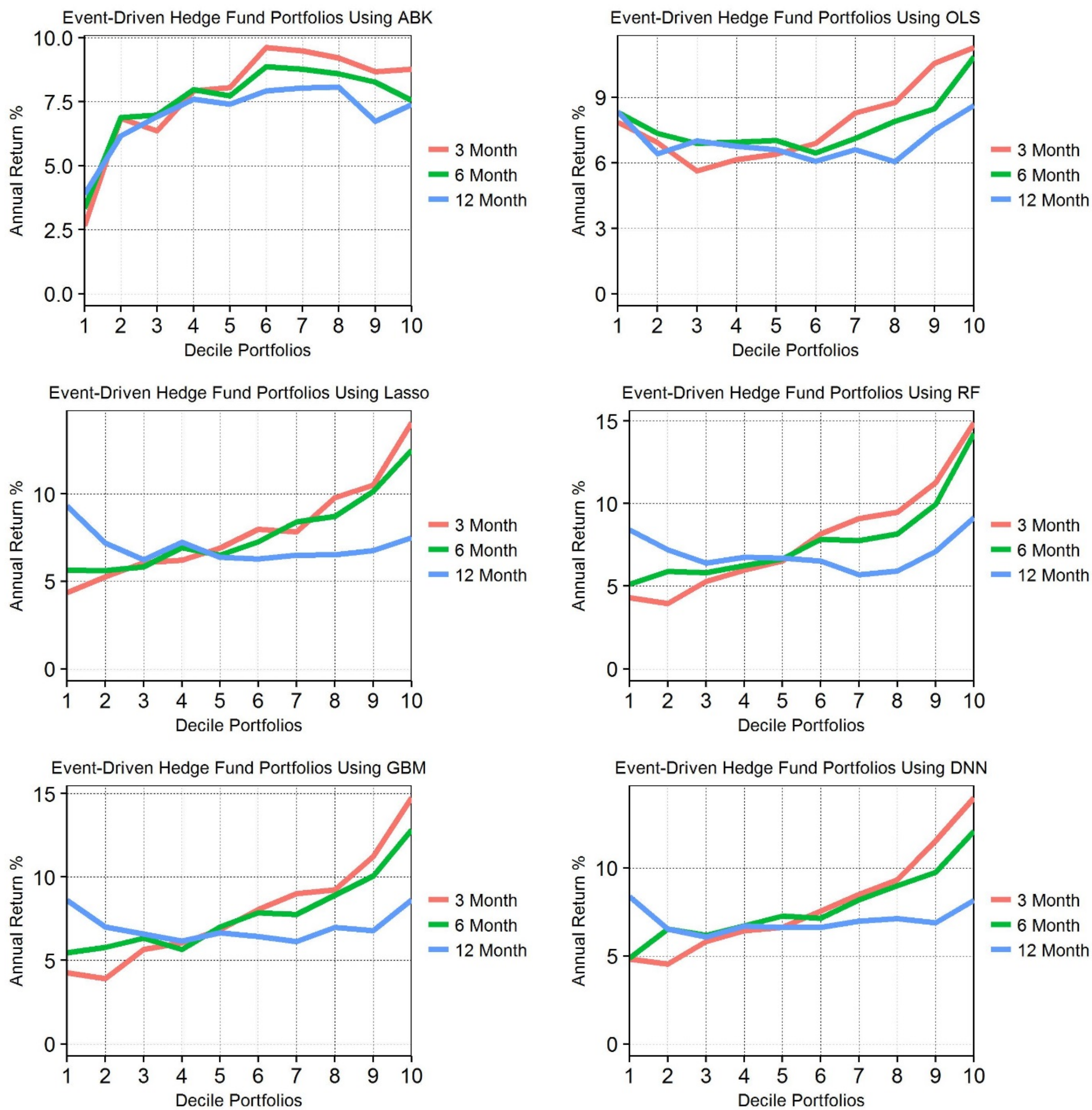
	HFR	ABK	OLS	Lasso	RF	GBM	DNN
Annual Return (%)	8.72	11.89	12.39	14.72	15.86	15.63	15.95
Annual Volatility	9.13	10.15	13.66	14.37	13.94	13.03	13.95
Sharpe Ratio	0.66	0.92	0.74	0.85	0.94	0.98	0.95

Skewness	-0.32	-0.15	0.36	0.58	0.44	0.50	0.45
Excess Kurtosis	2.17	3.89	2.78	3.76	3.87	3.15	4.28
Max Drawdown (%)	30.57	25.16	19.63	30.28	24.45	28.51	21.43

Event-Driven

Chart 3 shows the annual returns of our event-driven hedge fund portfolios ranked in the same way. The detailed summary statistics for the portfolios can be found in Appendix B.

Chart 3. Annual returns of the portfolios using ABK, OLS, Lasso, RF, GBM, and DNN forecasts



Analogous to the results with equity hedge fund, at short and medium rebalancing cycles, annual returns generally agree with the forecast ranks for the machine-learning algorithms, while the decile portfolio 10 in the ABK model still slightly lags deciles 6–8. In addition, the power to differentiate among ranks generally becomes weaker at longer rebalancing cycles.

Chart 4 and Table 2 show the performance of the best-performing portfolio from each algorithm at a three-month rebalancing cycle. All five machine algorithms generate higher annual returns than both the index and the ABK model, although some of them do not generate a better risk profile. Among the five algorithms, RF has the best total return. Its annual return of 14.90 percent is more than 70 percent higher than that of the index at 8.66 percent. All of the machine-learning algorithms generate portfolios with less negative skewness compared to the HFR event-driven index.

Chart 4. The cumulative returns of the six best-performing event-driven hedge fund portfolios and the HFR event-driven index

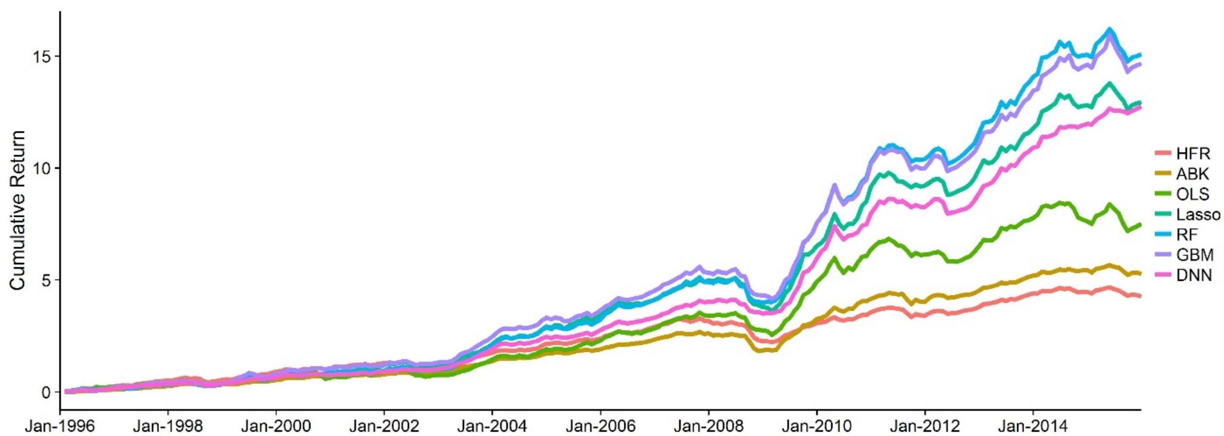


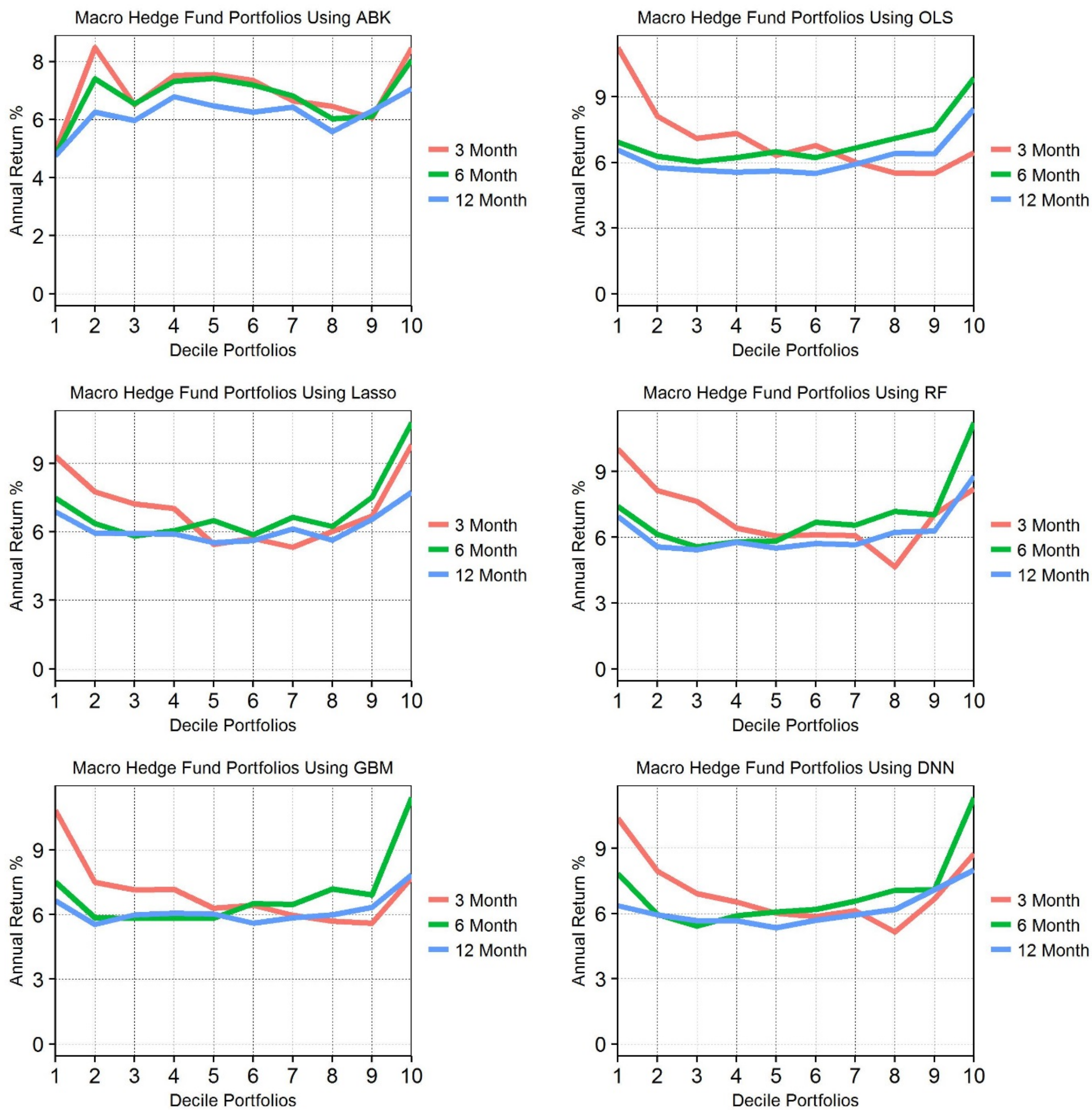
Table 2. Monthly return statistics of the six best-performing event-driven hedge fund portfolios and the HFR event-driven index

	HFR	ABK	OLS	Lasso	RF	GBM	DNN
Annual Return (%)	8.66	9.62	11.29	14.08	14.90	14.75	14.00
Annual Volatility	6.76	6.61	10.47	9.33	9.08	9.29	6.87
Sharpe Ratio	0.88	1.07	0.85	1.21	1.31	1.27	1.59
Skewness	-1.34	-0.93	-0.31	-0.31	-0.03	-0.17	0.28
Excess Kurtosis	4.48	6.94	1.03	2.27	1.18	1.14	2.47
Max Drawdown (%)	24.78	23.04	21.91	24.23	18.25	21.56	11.95

Macro

Chart 5 shows the annual returns of our macro fund portfolios ranked in the same way. The detailed summary statistics for the portfolios can be found in Appendix C.

Chart 5. Annual returns of the portfolios using ABK, OLS, Lasso, RF, GBM, and DNN forecasts



Macro hedge funds behave very differently than equity and event-driven funds. The best portfolios are formed at a medium six-month rebalancing cycle instead of a three-month cycle for the machine-learning algorithms, and they are much more difficult to forecast compared to the other styles. Although decile portfolio 10 at the six-month rebalancing scheme outperforms other deciles, there is little differentiating power from decile 1 to decile 9.

Chart 6 and Table 3 show the performance of the top-ranked portfolio from each machine-learning algorithm at a six-month rebalancing cycle and the best-performing portfolio from ABK model at a three-month rebalancing cycle. All five machine algorithms generate higher annual returns than both the index and the ABK model, although none of them offer a better Sharpe ratio. Among the five algorithms, GBM has the best total return. Its annual return of 11.42 percent is more than 70 percent higher than that of the index at 6.67 percent.

Chart 6. The cumulative returns of the six best-performing macro hedge fund portfolios and the HFR macro index

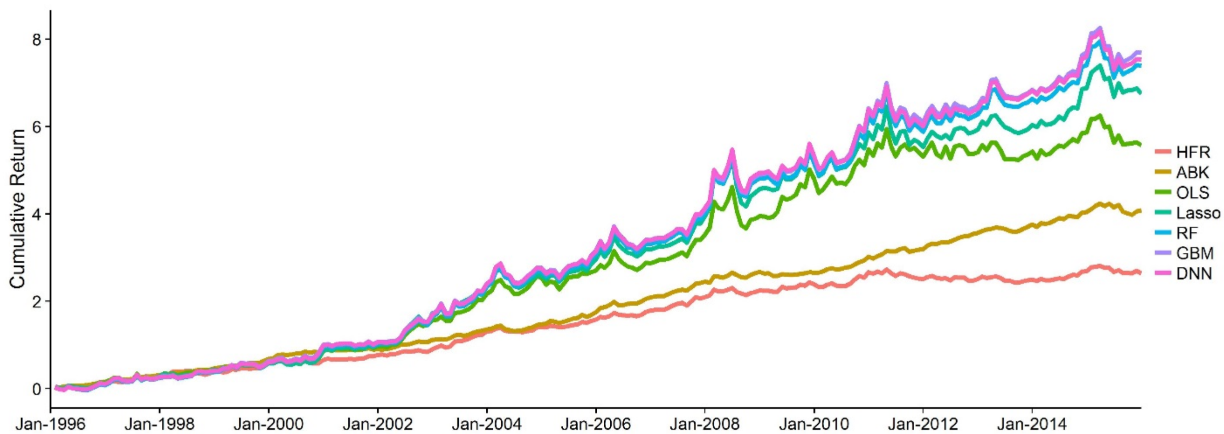


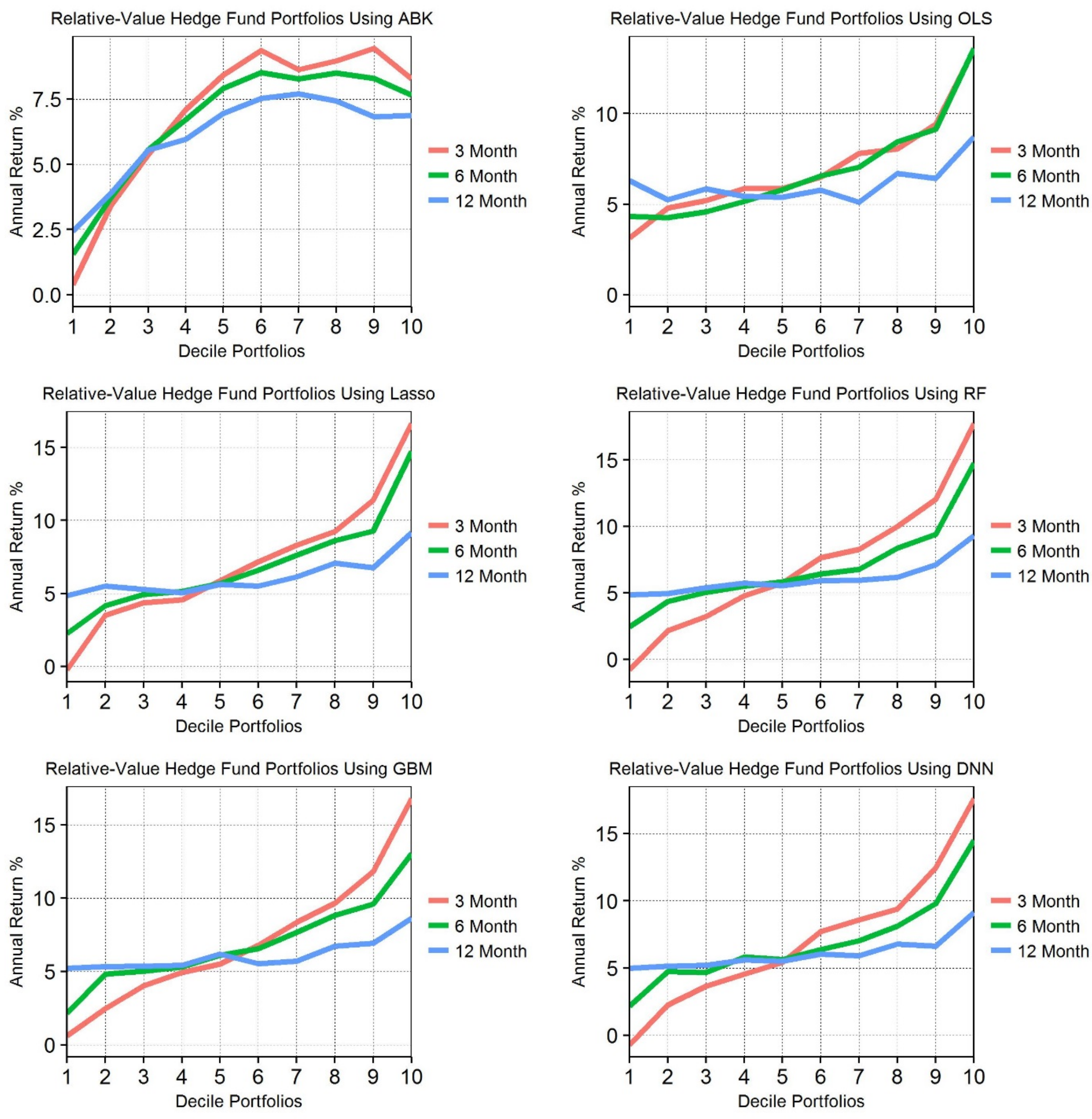
Table 3. Monthly return statistics of the six best-performing macro hedge fund portfolios and the HFR macro index

	HFR	ABK	OLS	Lasso	RF	GBM	DNN
Annual Return (%)	6.67	8.47	9.87	10.78	11.23	11.42	11.32
Annual Volatility	5.97	3.91	11.15	11.33	11.63	11.64	11.68
Sharpe Ratio	0.69	1.50	0.69	0.75	0.77	0.78	0.77
Skewness	0.47	-0.20	0.07	0.09	0.16	0.15	0.16
Excess Kurtosis	0.82	0.94	1.38	1.79	0.94	0.88	0.91
Max Drawdown (%)	8.01	5.20	16.88	17.63	14.97	14.88	14.97

Relative Value

Chart 7 shows the annual returns of our relative-value fund portfolios ranked in the same way. The detailed summary statistics for the portfolios can be found in Appendix D.

Chart 7. Annual returns of the portfolios using ABK, OLS, Lasso, RF, GBM, and DNN forecasts



Relative-value hedge funds are relatively easier to forecast. Forecasting power generally decreases with forecast horizon. Meanwhile, machine learning still generates good results at 12-month rebalancing.

Chart 8 and Table 4 show the performance of the best-performing portfolio from each algorithm at a three-month rebalancing cycle. All five machine-learning algorithms generate higher annual returns than the HFR relative value index and the ABK model. Most of them produce more than double the annual return of the index. RF has the highest return. Its annual return of 17.74% percent is more than 130 percent higher than that of the index at 7.61 percent.

Chart 8. The cumulative returns of the six best-performing relative-value hedge fund portfolios and the HFR relative-value index

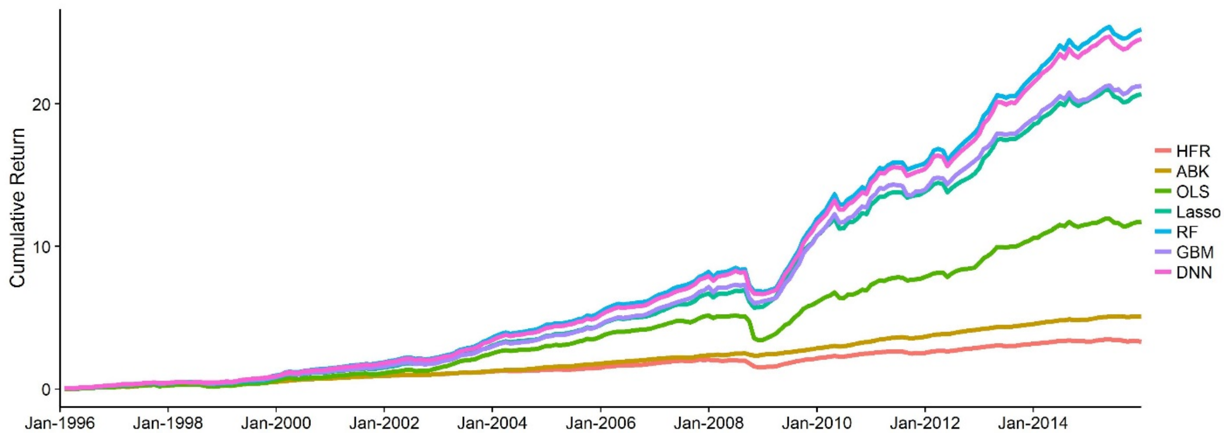


Table 4. Monthly return statistics of the six best-performing relative-value hedge fund portfolios and the HFR relative value index

	HFR	ABK	OLS	Lasso	RF	GBM	DNN
Annual Return (%)	7.61	9.45	13.56	16.63	17.74	16.78	17.59
Annual Volatility	4.31	2.50	8.74	7.49	7.21	7.09	7.12
Sharpe Ratio	1.16	2.69	1.28	1.79	1.99	1.90	2.00
Skewness	-2.91	-1.00	-3.89	-1.02	-1.23	-1.18	-1.30
Excess Kurtosis	16.56	4.10	32.88	5.84	8.29	7.23	9.15
Max Drawdown (%)	18.03	4.92	28.02	15.52	17.62	15.64	17.39

Throughout our analysis of the four trading styles, the RF, GBM and DNN models are generally the best, with very similar performance results, followed by the lasso, with OLS producing the worst results. RF and GBM both incorporate model-averaging methods, which may give them an advantage in handling inherently noisy data like hedge fund returns. DNN operates from another extreme by offering highly flexible functions designed to extract local features in the data. However, despite its simplicity, the OLS-based method still outperforms the style indexes consistently. That adds weight to our basic approach. As we have seen throughout the empirical results, if predictive signals exist in a dataset, various models are able to take advantage of such signals regardless of the techniques used in developing the model. Simple models can, therefore, be effective, but the more versatile models coupled with specific domain knowledge can produce much more accurate predictions and, thus, better portfolio performance. Such a framework can be extended to incorporate additional risk factors and alternative definitions of the response variable.

In terms of fund-selection performance, we find the greatest improvement in relative-value funds and the least improvement using macro funds. This could mean that there is more opportunity in the relative-value sector for the successful application of our fund-selection methods.

Although some of the machine-learning algorithms cannot provide straightforward information about how each factor is related with fund performance, it would be beneficial to examine how a factor contributes to performance. At each point in time, we trained the risk factor models using the RF approach employing a six-month holding horizon for macro funds and a three-month holding horizon for the other three styles. We then ranked the risk factors according to their variable importance, defined earlier, in the corresponding risk factor model. The final score is the average of such ranks for the 240 months in our study. There are 17 risk factors in our analysis, and Table 5 shows the outcome where we convert the ranks so that the highest possible rank in the exhibit is 1, and the lowest possible rank is 17. Among the 17 input variables, the higher-order moments and auto-correlation play the most important roles. The importance of the past returns is relatively low.

Table 5. Variable importance rank for each fund category based on RF models

	Equity	Event Driven	Macro	Relative Value
R1	16	14	13	15
R3	15	15	16	17
R6	13	13	15	16
R9	12	12	12	14
R12	11	11	10	11
SR	3	6	7	6
SD	17	17	17	12
Skew	2	4	2	5
Kurt	1	1	1	1
Omega	4	8	6	9
ACF1	5	5	5	4
ACF2	8	3	4	2
ACF3	7	2	3	3
Alpha	6	9	8	8
Alpha.t	10	10	11	10
DD	9	7	9	7
PerfLoss	14	16	14	13

Finally, we use the factor model proposed by Fung-Hsieh (2004) to evaluate the alpha and the factor loadings of our best machine-learning portfolios. They use seven factors: an equity-market factor measured by the S&P 500 index monthly total return, a size-spread factor defined as the difference between the Russell 2000 index monthly total return and the S&P 500 monthly total return, a bond-market factor defined as the monthly change in the 10-year Treasury constant maturity yield, a credit-spread factor defined as the monthly change in the Moody's BAA yield less the 10-year Treasury constant-maturity yield, and three trend-following (TF) factors, one each for bond, currency, and commodity markets. Table 6 shows the regression results for decile portfolio 10 for the four styles.

Table 6. Regression of the decile portfolio 10 returns on Fung-Hsieh factors with p-values in parentheses

Factor	OLS	LASSO	RF	GBM	DNN
Equity funds:					
Intercept	0.0063 (0.359)	0.0114 (0.128)	0.0136 (0.062)	0.0123 (0.063)	0.0116 (0.111)
Equity Market	0.4793 (0.000)	0.4603 (0.000)	0.4074 (0.000)	0.4238 (0.000)	0.4169 (0.000)
Size Spread	0.3747 (0.000)	0.4250 (0.000)	0.4594 (0.000)	0.4078 (0.000)	0.4431 (0.000)
Bond Market	0.0105 (0.268)	0.0192 (0.062)	0.0164 (0.101)	0.0171 (0.060)	0.0188 (0.061)
Credit Spread	0.0002 (0.951)	-0.0009 (0.739)	-0.0014 (0.597)	-0.0011 (0.672)	-0.0006 (0.815)
Bond TF	-0.0119 (0.426)	0.0012 (0.940)	-0.0067 (0.673)	-0.0057 (0.692)	-0.0072 (0.649)
Currency TF	0.0118 (0.336)	0.0098 (0.462)	0.0068 (0.600)	0.0101 (0.391)	0.0068 (0.601)
Commodity TF	0.0137 (0.363)	0.0034 (0.836)	0.0131 (0.409)	0.0138 (0.340)	0.0119 (0.455)
R^2	0.4094	0.3825	0.3772	0.4093	0.3775
Event-Driven Funds:					
Intercept	0.0086 (0.087)	0.0115 (0.015)	0.0100 (0.033)	0.0121 (0.009)	0.0109 (0.003)
Equity Market	0.3734 (0.000)	0.3082 (0.000)	0.2873 (0.000)	0.3189 (0.000)	0.1871 (0.000)
Size Spread	0.2653 (0.000)	0.2138 (0.000)	0.1961 (0.000)	0.2075 (0.000)	0.1315 (0.000)
Bond Market	0.0133 (0.056)	0.0119 (0.069)	0.0129 (0.048)	0.0125 (0.048)	0.0114 (0.025)
Credit Spread	-0.0009 (0.652)	-0.0010 (0.589)	-0.0002 (0.929)	-0.0011 (0.526)	-0.0006 (0.679)
Bond TF	-0.0090 (0.409)	-0.0015 (0.886)	-0.0157 (0.126)	-0.0139 (0.164)	-0.0224 (0.005)
Currency TF	0.0012 (0.898)	0.0030 (0.725)	0.0067 (0.427)	0.0086 (0.296)	0.0056 (0.389)
Commodity TF	0.0023 (0.832)	-0.0120 (0.247)	-0.0053 (0.609)	-0.0101 (0.313)	-0.0019 (0.817)
R^2	0.4503	0.3903	0.3692	0.4245	0.3324
Macro Funds:					
Intercept	0.0059 (0.364)	0.0077 (0.253)	0.0099 (0.151)	0.0097 (0.158)	0.0096 (0.166)
Equity Market	0.2050 (0.000)	0.1621 (0.001)	0.1627 (0.001)	0.1634 (0.001)	0.1681 (0.001)
Size Spread	0.0455 (0.427)	0.0736 (0.217)	0.0708 (0.245)	0.0705 (0.246)	0.0763 (0.210)
Bond Market	-0.0063 (0.482)	-0.0044 (0.639)	-0.0022 (0.820)	-0.0019 (0.839)	-0.0029 (0.761)
Credit Spread	0.0006 (0.790)	0.0003 (0.901)	-0.0004 (0.886)	-0.0002 (0.927)	-0.0002 (0.928)
Bond TF	0.0331 (0.020)	0.0311 (0.035)	0.0331 (0.028)	0.0336 (0.025)	0.0331 (0.028)
Currency TF	0.0387 (0.001)	0.0363 (0.003)	0.0406 (0.001)	0.0407 (0.001)	0.0410 (0.001)
Commodity TF	0.0467 (0.001)	0.0448 (0.003)	0.0476 (0.002)	0.0486 (0.001)	0.0483 (0.002)
R^2	0.1981	0.1624	0.1747	0.1778	0.1786
Relative-Value Funds:					
Intercept	0.0118 (0.006)	0.0076 (0.056)	0.0150 (0.000)	0.0124 (0.001)	0.0151 (0.000)
Equity Market	0.2666 (0.000)	0.2157 (0.000)	0.2068 (0.000)	0.2178 (0.000)	0.1973 (0.000)
Size Spread	0.1712 (0.000)	0.1312 (0.000)	0.1287 (0.000)	0.1354 (0.000)	0.1381 (0.000)
Bond Market	-0.0024 (0.682)	0.0042 (0.446)	0.0051 (0.332)	0.0037 (0.470)	0.0040 (0.446)
Credit Spread	-0.0013 (0.415)	0.0014 (0.352)	-0.0011 (0.431)	-0.0004 (0.750)	-0.0012 (0.400)
Bond TF	-0.0136 (0.138)	-0.0194 (0.026)	-0.0121 (0.145)	-0.0134 (0.095)	-0.0131 (0.113)
Currency TF	-0.0096 (0.202)	0.0001 (0.989)	0.0013 (0.847)	0.0017 (0.801)	0.0014 (0.833)
Commodity TF	-0.0153 (0.099)	-0.0084 (0.337)	-0.0132 (0.115)	-0.0103 (0.205)	-0.0127 (0.128)
R^2	0.4003	0.3201	0.3255	0.3513	0.3202

Except for macro hedge funds, most of the machine-learning algorithms have significant alpha of more than 1 percent per month. The loadings on the equity-market and size-spread factors are all positive, which implies that all the algorithms tend to generate long exposures to the equity market and exhibit a preference for small cap companies. Although the loadings on the equity-market and size-spread factors are positive, the coefficients are not large. Thus, it may be that the superior performance of the algorithms is not coming from overleveraging the two factors but rather from the algorithms. This is reinforced by the observation that the R^2 in all the regressions are relatively small. The trend-following factors were significant only for the macro hedge funds algorithms. That agrees with the fact that many macro funds employ trend-following strategies.

7. Conclusion

We propose a supervised machine-learning approach to forecast hedge fund returns and select hedge funds quantitatively. The framework is based on cross-sectional forecasts of hedge fund returns utilizing a set of 17 factors. The approach allows the investor to identify funds that are likely to perform well and to construct the corresponding portfolios. We find that our method is applicable across hedge fund style categories. Focusing on factors constructed from characteristics idiosyncratic to individual funds, our models offer distinctive perspectives when compared to models that are driven by macroeconomic variables. Retrospectively, when benchmarked against a traditional factor model, our machine-learning approach generates portfolios with large alphas. The relatively low explanatory power of the regressions indicates that most of the performance of the algorithm-generated portfolios is due to success in identifying funds likely to deliver good performance. Our approach is flexible enough to incorporate new developments both in risk-factor research field and in the machine-learning field.

References

Agarwal, Vikas, and Narayan Y. Naik (2000), "Multi-Period Performance Persistence Analysis of Hedge Funds," *Journal of Financial and Quantitative Analysis,* 35 (3): 327–42.

Avramov, Doron, Laurent Barras, and Robert Kosowski (2013), "Hedge Fund Return Predictability Under the Magnifying Glass," *Journal of Financial and Quantitative Analysis,* 48 (4): 1057–83.

Bai, Jushan, and Serena Ng (2008), "Forecasting Economic Time Series Using Targeted Predictors," *Journal of Econometrics,* 146 (2): 304–17.

Bali, Turan, Stephen Brown, Scott Murray and Yi Tang (2016), "A Lottery Demand-Based Explanation of the Beta Anomaly," Georgetown McDonough School of Business Research Paper No. 2408146.

Bengio, Yoshua, Pascal Lamblin, Dan Popovici, and Hugo Larochelle (2007), "Greedy Layer-Wise Training of Deep Networks," *Advances in Neural Information Processing Systems,* 19: 153–60.

Berk, Richard A. (2008), *Statistical Learning from a Regression Perspective,* New York: Springer.

Breiman, Leo (2001), "Random Forests," *Machine Learning*, 45 (1): 5–32.

Carhart, Mark M. (1997), "On Persistence in Mutual Fund Performance," *The Journal of Finance*, 52 (1): 57–82.

Cattaneo, Matias, Richard Crump, Max Farrell and Ernst Schaumburg (2016), "Characteristic-Sorted Portfolios: Estimation and Inference," Federal Reserve Bank of New York Staff Reports No.788.

De Souza, Clifford, and Suleyman Gokcan (2004), "Hedge Fund Investing: A Quantitative Approach to Hedge Fund Manager Selection and De-Selection," *Journal of Wealth Management*, 6 (4): 52–73.

Fama, Eugene F., and Kenneth R. French (1992), "The Cross-Section of Expected Stock Returns," *The Journal of Finance*, 47 (2): 427–65.

Fung, William and David Hsieh (2004), "Hedge Fund Benchmarks: A Risk-Based Approach," *Financial Analysts Journal*, 60 (5): 65–80.

Getmansky, Mila, Andrew W. Lo, and Igor Makarov (2004), "An Econometric Model of Serial Correlation and Illiquidity in Hedge Fund Returns," *Journal of Financial Economics,* 74: 529–609.

Glorot, Xavier, and Yoshua Bengio (2010), "Understanding the Difficulty of Training Deep Feedforward Neural Networks," *Proceedings of the International Conference on Artificial Intelligence and Statistics*, May 2010.

Goetzmann, William, Jonathan Ingersoll, and Stephen Ross (2003), "High-Water Marks and Hedge Fund Management Contracts," *The Journal of Finance,* 58 (4): 1685–717.

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville (2016), *Deep Learning*, Cambridge, Massachusetts: MIT Press.

Harvey, Campbell R., Yan Liu, and Heqing Zhu (2016), "… and the Cross-Section of Expected Returns," *The Review of Financial Studies*, 29 (1): 5–68.

Hastie, Trevor, Robert Tibshirani, Jerome Friedman (2009), *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*, New York: Springer.

Haugen, Robert A., and Nardin L. Baker (1996), "Commonality in the Determinants of Expected Stock Returns," *Journal of Financial Economics*, 41 (3): 401–39.

Hinton, Geoffrey E., Simon Osindero, and Yee-Whye Teh (2006), "A Fast Learning Algorithm for Deep Belief Nets," *Neural Computation*, 18 (7): 1527–54.

Jagannathan, Ravi, Alexey Malakhov, and Dmitry Novikov (2010), "Do Hot Hands Exist Among Hedge Fund Managers? An Empirical Evaluation," *The Journal of Finance*, 65 (1): 217–55.

Molyboga, Marat, Seungho Baek, and John F.O. Bilson (2014), "CTA Performance Persistence: 1994–2010," *The Journal of Alternative Investments*, 16 (4): 61–70.

Moskowitz, Tobias J., Yao H. Ooi, and Lasse H. Pedersen (2010), "Time Series Momentum," *Journal of Financial Economics*, 104 (2): 228–50.

Pedersen, Niels K. (2013), *A Quantitative Framework for Hedge Fund Manager Selection*, Quantitative Research and Analytics, Pacific Investment Management Company, LLC.

Ridgeway, Greg (1999), "The State of Boosting," *Computing Science and Statistics*, 31: 172–81.

Sun, Zheng, Ashley W. Wang, and Lu Zheng (2016), "Only Winners in Tough Times Repeat: Hedge Fund Performance Persistence Over Different Market Conditions," *Finance and Economics Discussion Series 2016-030*. Washington: Board of Governors of the Federal Reserve System.

Varian, Hal R. (2014), "Big Data: New Tricks for Econometrics," *Journal of Economic Perspectives*, 28 (2): 3–28.

Wang, Li, and Ji Zhu (2008), "Financial Market Forecasting Using a Two-Step Kernel Learning Method for the Support Vector Regression," *Annals of Operations Research*, 174 (1): 103–20.

Welsch, Roy E., and Xinfeng Zhou (2007), "Application of Robust Statistics to Asset Allocation Models," *Statistical Journal*, 5 (1): 97–114.

Appendix A. Statistics of monthly returns of the equity hedge fund portfolios

Table A1. Statistics of monthly returns of ABK portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	1.04	4.57	6.01	7.22	8.65	10.08	11.89	11.35	10.91	10.23
	Ann Volatility	9.51	11.80	11.58	10.92	11.15	10.72	10.15	9.75	8.08	6.64
	Sharpe Ratio	-0.09	0.24	0.36	0.48	0.59	0.73	0.92	0.91	1.02	1.14
	Skewness	-0.60	-0.75	-0.86	-0.99	-0.73	-0.28	-0.15	0.06	0.34	-0.14
	Excess Kurtosis	1.94	3.72	3.26	3.74	3.51	2.98	3.89	4.07	3.15	0.98
	Max Draw %	31.53	35.79	37.01	34.39	33.38	29.54	25.16	24.31	19.24	13.24
6-month	Ann Return %	3.25	5.61	5.89	6.48	8.26	9.70	10.98	10.14	9.20	8.72
	Ann Volatility	8.69	10.98	10.80	10.73	10.70	10.37	10.00	9.32	8.02	6.69
	Sharpe Ratio	0.14	0.34	0.37	0.43	0.58	0.72	0.86	0.83	0.84	0.94
	Skewness	-0.33	-0.56	-0.78	-0.95	-0.73	-0.41	-0.22	-0.01	-0.09	-0.78
	Excess Kurtosis	1.95	3.33	3.13	3.47	2.67	2.47	3.30	3.04	2.80	3.54
	Max Draw %	29.54	34.77	35.53	34.60	31.39	28.38	24.91	24.10	20.82	14.71
12-month	Ann Return %	3.78	5.51	5.73	6.69	7.57	8.69	8.97	8.46	7.82	8.07
	Ann Volatility	7.90	9.88	10.11	10.09	10.23	9.94	9.82	9.16	8.01	6.70
	Sharpe Ratio	0.21	0.36	0.37	0.46	0.54	0.66	0.69	0.68	0.69	0.84
	Skewness	-0.41	-0.14	-0.75	-0.79	-0.60	-0.54	-0.57	-0.40	-0.11	-0.33
	Excess Kurtosis	2.92	2.80	3.20	2.75	2.38	2.74	3.24	3.18	2.91	2.45
	Max Draw %	25.96	30.13	31.45	30.59	30.28	28.23	27.82	26.69	21.46	16.20

Table A2. Statistics of monthly returns of OLS portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	2.75	5.48	6.43	7.84	8.29	8.81	9.37	10.45	10.66	12.39
	Ann Volatility	16.30	11.84	10.14	9.08	8.53	8.21	8.44	8.87	10.12	13.66
	Sharpe Ratio	0.11	0.32	0.44	0.63	0.71	0.78	0.82	0.90	0.82	0.74
	Skewness	-1.43	-1.25	-1.22	-1.07	-0.85	-0.29	0.05	-0.09	0.12	0.36
	Excess Kurtosis	7.82	5.89	4.82	4.85	4.30	1.05	2.18	2.54	3.77	2.78
	Max Draw %	52.67	42.45	33.62	30.70	23.84	24.05	21.30	15.73	14.10	19.63
6-month	Ann Return %	5.18	6.53	7.27	7.96	7.62	8.33	8.47	9.02	9.98	10.59
	Ann Volatility	13.67	10.41	9.42	8.54	8.32	8.30	8.07	8.78	10.32	13.31
	Sharpe Ratio	0.27	0.44	0.55	0.67	0.65	0.73	0.76	0.76	0.74	0.64
	Skewness	-0.84	-0.95	-1.12	-1.03	-0.85	-0.75	-0.36	-0.06	0.23	0.44
	Excess Kurtosis	4.38	3.56	4.63	4.10	3.32	3.29	1.87	2.40	3.69	4.92
	Max Draw %	41.02	32.44	27.53	26.83	24.91	24.81	22.96	25.22	25.17	33.67
12-month	Ann Return %	5.77	6.48	6.99	7.47	7.40	7.24	7.15	7.29	7.57	7.48
	Ann Volatility	8.20	7.23	7.31	7.28	7.76	8.42	9.12	9.96	11.86	15.44
	Sharpe Ratio	0.44	0.58	0.64	0.71	0.66	0.60	0.55	0.52	0.48	0.39
	Skewness	0.05	-0.20	-0.48	-0.56	-0.56	-0.76	-0.64	-0.66	-0.69	-0.49
	Excess Kurtosis	2.12	2.28	4.50	3.70	3.51	3.62	2.80	2.49	2.95	2.66
	Max Draw %	19.98	18.55	18.43	20.45	22.57	26.46	28.25	30.16	35.61	47.21

Table A3. Statistics of monthly returns of the lasso portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	1.46	5.87	6.02	6.46	7.44	8.15	9.85	11.06	11.62	14.72
	Ann Volatility	16.06	11.79	9.78	9.01	8.62	8.05	8.22	9.20	10.24	14.37
	Sharpe Ratio	0.02	0.35	0.41	0.48	0.61	0.72	0.90	0.92	0.89	0.85
	Skewness	-1.51	-1.15	-1.46	-0.92	-0.96	-0.41	0.07	0.35	0.32	0.58
	Excess Kurtosis	8.10	5.61	6.18	3.72	3.85	1.25	2.17	2.33	3.62	3.76
	Max Draw %	45.70	34.99	31.53	28.06	26.87	23.66	20.79	18.86	22.11	30.28
6-month	Ann Return %	2.92	5.77	6.78	7.24	8.00	7.92	9.14	9.69	11.15	12.35
	Ann Volatility	13.80	10.17	9.37	8.46	8.27	8.19	8.54	9.33	10.69	13.99
	Sharpe Ratio	0.11	0.38	0.50	0.60	0.69	0.69	0.79	0.78	0.82	0.73
	Skewness	-1.15	-1.17	-1.25	-1.06	-0.65	-0.86	-0.30	-0.01	0.23	0.20
	Excess Kurtosis	4.70	4.44	4.73	4.21	2.06	3.74	2.80	3.08	3.84	4.77
	Max Draw %	47.70	34.47	31.70	27.99	25.38	24.10	22.19	23.07	22.52	29.20
12-month	Ann Return %	6.68	7.02	7.82	7.47	7.31	6.80	6.96	6.78	6.99	6.84
	Ann Volatility	8.93	7.48	7.10	7.48	7.79	8.31	9.27	10.02	11.88	15.70
	Sharpe Ratio	0.50	0.63	0.77	0.69	0.65	0.56	0.52	0.47	0.44	0.35
	Skewness	0.16	-0.25	-0.09	-0.55	-0.55	-0.90	-0.73	-0.73	-0.78	-0.65
	Excess Kurtosis	2.34	2.54	1.65	3.88	3.53	4.45	3.40	2.72	3.05	3.30
	Max Draw %	25.77	18.41	17.67	20.70	21.98	26.12	28.94	30.28	36.34	47.23

Table A4. Statistics of monthly returns of random forest portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	1.09	4.78	5.3	7.17	7.1	8.64	10.01	10.84	11.86	15.86
	Ann Volatility	16.6	11.52	10.36	8.97	8.37	8.06	8.69	9.21	10.78	13.94
	Sharpe Ratio	0.01	0.26	0.33	0.56	0.59	0.78	0.87	0.9	0.87	0.94
	Skewness	-1.61	-1.4	-1.51	-0.83	-0.89	-0.41	0.28	0.15	0.27	0.44
	Excess Kurtosis	9.21	6.93	6.76	3.99	3.54	2.12	2.73	2.56	4.07	3.87
	Max Draw %	49.39	41	36.21	29.02	26.18	20.54	16.76	17.88	19.68	24.45
6-month	Ann Return %	3.62	4.97	6.24	7.46	7.23	8.27	9.15	9.81	11.24	12.79
	Ann Volatility	14.46	10.99	9.64	8.56	8.01	8.02	8.83	9.36	11.01	13.7
	Sharpe Ratio	0.16	0.29	0.44	0.62	0.62	0.74	0.77	0.79	0.8	0.77
	Skewness	-1.18	-1.2	-1.32	-0.92	-0.91	-0.55	-0.25	-0.07	0.18	0.18
	Excess Kurtosis	6.09	6.49	6.23	3.62	2.94	2.86	2.57	3.08	4.08	4.67
	Max Draw %	48.34	39.86	36.18	30.69	28.29	24.09	21.85	17.09	17.61	21.79
12-month	Ann Return %	6.49	7.53	7.42	6.98	7.29	7.66	6.8	7.2	6.95	6.51
	Ann Volatility	9.09	8.07	7.67	7.63	7.96	8.42	9.02	9.92	11.42	14.44
	Sharpe Ratio	0.48	0.65	0.67	0.62	0.63	0.64	0.52	0.52	0.44	0.35
	Skewness	0.1	-0.11	-0.57	-0.64	-0.62	-0.69	-0.75	-0.43	-0.58	-0.49
	Excess Kurtosis	2.28	2.45	3.02	3.09	2.95	3.89	4.11	2.69	3.68	3.77
	Max Draw %	24.07	22.36	22.61	24.05	25.21	25.46	27.83	27.66	33.19	39.08

Table A5. Statistics of monthly returns of gradient boosting portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	2.33	4.49	5.63	6.69	7.64	8.81	8.91	10.69	12.05	15.63
	Ann Volatility	15.76	11.64	10.19	9.08	8.35	8.11	8.42	9.53	10.38	13.03
	Sharpe Ratio	0.08	0.24	0.36	0.51	0.65	0.79	0.77	0.86	0.91	0.98
	Skewness	-1.38	-1.67	-1.49	-1.05	-0.78	-0.33	0.10	0.12	0.68	0.50
	Excess Kurtosis	7.50	8.32	6.94	4.07	2.34	2.26	2.38	3.04	5.47	3.15
	Max Draw %	46.00	40.76	32.87	29.89	25.06	19.14	19.25	19.43	20.44	28.51
6-month	Ann Return %	4.06	5.46	6.66	8.02	8.10	7.91	9.09	9.59	10.43	11.64
	Ann Volatility	13.77	10.66	9.11	8.42	8.07	8.26	8.72	9.59	10.51	13.27
	Sharpe Ratio	0.19	0.34	0.50	0.68	0.72	0.68	0.77	0.76	0.77	0.71
	Skewness	-1.08	-1.32	-1.33	-0.74	-0.46	-0.68	-0.40	-0.05	-0.09	0.26
	Excess Kurtosis	4.74	6.16	5.64	3.16	2.28	2.70	2.83	2.45	3.11	4.25
	Max Draw %	43.84	37.95	33.14	30.23	26.67	25.93	20.87	21.30	17.19	20.51
12-month	Ann Return %	7.07	7.30	7.59	7.63	7.66	7.03	7.09	6.30	6.67	6.58
	Ann Volatility	9.64	8.44	7.73	7.86	7.83	7.94	8.71	9.82	11.12	13.90
	Sharpe Ratio	0.51	0.60	0.68	0.68	0.69	0.60	0.56	0.44	0.43	0.36
	Skewness	0.06	-0.53	-0.29	-0.58	-0.59	-0.66	-0.51	-0.72	-0.78	-0.69
	Excess Kurtosis	2.10	3.46	1.66	2.64	2.80	3.20	2.57	4.16	3.79	4.07
	Max Draw %	26.60	25.11	22.16	23.72	23.84	23.88	25.91	25.76	32.53	38.48

Table A6. Statistics of monthly returns of deep neural network portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	1.82	3.49	5.07	5.98	7.85	8.88	10.11	11.16	12.43	15.95
	Ann Volatility	16.53	12.02	10.19	9.33	8.28	8.16	8.36	9.07	10.40	13.95
	Sharpe Ratio	0.05	0.15	0.31	0.42	0.67	0.80	0.91	0.95	0.94	0.95
	Skewness	-1.49	-1.42	-1.12	-1.07	-0.63	-0.25	-0.08	0.19	0.27	0.45
	Excess Kurtosis	7.54	6.04	4.15	4.02	3.27	2.04	2.08	3.22	4.23	4.28
	Max Draw %	47.78	39.55	35.84	30.97	25.23	23.13	19.71	19.33	18.39	21.43
6-month	Ann Return %	3.35	5.29	6.56	7.18	7.84	7.89	8.94	9.78	10.96	12.84
	Ann Volatility	14.46	10.93	9.36	8.79	8.10	8.39	8.56	9.41	11.04	14.45
	Sharpe Ratio	0.14	0.32	0.48	0.57	0.69	0.67	0.77	0.79	0.78	0.74
	Skewness	-1.33	-1.31	-1.22	-1.03	-0.85	-0.69	-0.34	-0.02	0.14	0.32
	Excess Kurtosis	7.24	6.54	5.49	3.84	3.07	3.95	3.12	3.06	4.14	5.46
	Max Draw %	48.09	41.67	36.01	32.14	27.43	24.95	19.60	17.35	17.47	26.18
12-month	Ann Return %	7.59	7.44	7.68	7.61	6.98	6.69	6.28	7.25	6.73	6.51
	Ann Volatility	9.80	8.38	7.95	7.76	7.81	8.10	8.85	9.79	11.07	14.73
	Sharpe Ratio	0.55	0.62	0.68	0.69	0.61	0.55	0.47	0.53	0.44	0.35
	Skewness	0.08	-0.35	-0.27	-0.51	-0.73	-0.56	-0.59	-0.37	-0.69	-0.71
	Excess Kurtosis	2.49	3.05	2.30	2.42	2.91	2.47	3.20	3.32	3.54	4.50
	Max Draw %	24.12	23.66	23.76	22.80	23.55	23.58	25.27	27.93	32.80	41.89

Appendix B. Statistics of monthly returns of the event-driven fund portfolios

Table B1. Statistics of monthly returns of ABK portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	2.65	6.85	6.37	7.92	8.05	9.62	9.49	9.21	8.67	8.77
	Ann Volatility	10.16	9.84	8.54	7.36	7.06	6.61	5.58	5.42	4.05	4.25
	Sharpe Ratio	0.08	0.49	0.50	0.76	0.81	1.07	1.24	1.24	1.50	1.46
	Skewness	-1.16	-1.70	-1.47	-1.03	-1.17	-0.93	-1.16	-2.23	-1.35	-1.20
	Excess Kurtosis	6.14	9.77	5.87	4.16	6.58	6.94	4.69	10.87	3.45	4.40
	Max Draw %	35.20	38.27	34.18	27.75	27.69	23.04	20.90	24.03	11.97	10.50
6-month	Ann Return %	3.34	6.88	6.99	7.98	7.73	8.87	8.77	8.60	8.27	7.55
	Ann Volatility	9.86	9.04	7.77	7.02	6.59	6.26	5.43	5.22	4.01	4.47
	Sharpe Ratio	0.15	0.53	0.61	0.80	0.81	1.02	1.15	1.17	1.43	1.14
	Skewness	-1.49	-1.83	-1.28	-1.03	-1.11	-1.24	-1.21	-1.96	-1.70	-2.25
	Excess Kurtosis	7.82	9.43	4.89	3.42	4.63	6.75	4.53	8.72	5.52	10.95
	Max Draw %	35.67	36.33	31.30	27.30	26.37	20.83	21.52	22.91	12.52	13.21
12-month	Ann Return %	3.88	6.17	6.91	7.60	7.40	7.93	8.04	8.08	6.74	7.39
	Ann Volatility	8.31	8.25	7.23	6.85	6.22	5.90	5.22	5.05	4.20	4.31
	Sharpe Ratio	0.22	0.49	0.64	0.77	0.81	0.93	1.07	1.11	1.03	1.14
	Skewness	-0.69	-1.65	-1.41	-1.24	-1.33	-1.16	-1.08	-1.38	-1.84	-1.91
	Excess Kurtosis	4.69	8.81	5.89	4.92	5.13	4.95	3.13	5.76	6.34	7.92
	Max Draw %	29.53	34.53	29.77	28.64	23.09	21.39	20.22	20.25	15.79	15.07

Table B2. Statistics of monthly returns of OLS portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	7.87	6.93	5.63	6.14	6.39	6.89	8.30	8.77	10.56	11.29
	Ann Volatility	11.47	7.76	5.97	5.91	5.51	5.11	5.34	5.77	6.66	10.47
	Sharpe Ratio	0.52	0.62	0.57	0.66	0.74	0.88	1.09	1.09	1.19	0.85
	Skewness	-1.65	-3.13	-2.29	-2.87	-2.06	-1.55	-1.35	-1.07	-0.35	-0.31
	Excess Kurtosis	11.13	18.16	10.29	15.80	8.48	5.10	5.04	3.29	2.15	1.03
	Max Draw %	43.24	35.16	28.74	25.05	25.16	22.87	19.50	18.65	15.87	21.91
6-month	Ann Return %	8.32	7.36	6.90	6.96	7.03	6.45	7.12	7.90	8.49	10.87
	Ann Volatility	11.06	6.48	5.93	5.72	5.13	5.32	5.31	5.57	6.71	9.73
	Sharpe Ratio	0.57	0.78	0.77	0.81	0.90	0.77	0.89	0.98	0.91	0.88
	Skewness	-0.56	-1.47	-1.97	-2.14	-1.78	-2.28	-2.02	-1.50	-1.56	-1.04
	Excess Kurtosis	5.12	6.22	8.84	9.74	7.26	10.13	8.43	6.27	6.10	4.52
	Max Draw %	38.03	30.84	28.56	29.09	24.41	22.60	22.15	19.21	19.76	19.95
12-month	Ann Return %	8.34	6.42	7.01	6.76	6.60	6.07	6.60	6.05	7.53	8.63
	Ann Volatility	10.26	6.49	5.55	5.37	5.34	4.86	4.73	5.30	5.67	8.62
	Sharpe Ratio	0.60	0.64	0.84	0.82	0.80	0.76	0.89	0.70	0.91	0.74
	Skewness	-0.06	-1.26	-1.49	-1.59	-1.87	-1.54	-1.69	-2.00	-1.97	-1.49
	Excess Kurtosis	2.33	6.62	6.80	6.83	8.05	5.57	6.15	9.72	8.77	7.65
	Max Draw %	35.88	29.37	27.08	24.96	25.88	18.34	18.30	17.53	19.54	21.00

Table B3. Statistics of monthly returns of the lasso portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	4.36	5.26	6.06	6.21	6.90	7.99	7.84	9.79	10.51	14.08
	Ann Volatility	13.29	7.53	6.85	6.34	5.25	5.34	5.49	5.29	6.48	9.33
	Sharpe Ratio	0.22	0.42	0.57	0.63	0.86	1.04	0.99	1.35	1.21	1.21
	Skewness	-1.54	-2.16	-2.94	-2.78	-1.93	-1.52	-1.50	-0.40	0.02	-0.31
	Excess Kurtosis	9.81	11.93	16.31	14.78	7.11	6.03	6.66	1.12	2.63	2.27
	Max Draw %	44.25	27.72	31.28	30.84	21.87	20.78	24.12	13.46	16.46	24.23
6-month	Ann Return %	5.63	5.62	5.84	6.92	6.49	7.26	8.40	8.72	10.15	12.52
	Ann Volatility	12.37	6.71	6.19	5.79	5.46	4.84	5.09	5.62	6.14	9.18
	Sharpe Ratio	0.32	0.51	0.58	0.79	0.76	1.00	1.16	1.11	1.23	1.08
	Skewness	-1.92	-1.78	-2.37	-1.88	-2.47	-1.49	-1.33	-1.36	-1.06	-0.70
	Excess Kurtosis	13.07	8.34	12.36	7.71	11.63	4.38	3.77	3.93	3.22	1.63
	Max Draw %	47.56	31.99	32.46	27.32	27.50	18.45	17.37	18.69	13.85	19.63
12-month	Ann Return %	9.31	7.18	6.24	7.23	6.39	6.28	6.51	6.52	6.76	7.51
	Ann Volatility	10.07	6.49	5.36	5.48	4.93	5.05	5.31	5.62	6.12	8.84
	Sharpe Ratio	0.70	0.75	0.73	0.89	0.81	0.78	0.78	0.75	0.73	0.61
	Skewness	0.12	-1.15	-1.30	-1.89	-1.71	-1.92	-1.64	-1.75	-1.37	-1.15
	Excess Kurtosis	6.54	9.15	5.70	10.28	6.36	7.82	5.73	7.58	4.72	3.17
	Max Draw %	34.07	29.41	25.50	26.51	23.42	23.15	19.15	18.29	20.29	19.22

Table B4. Statistics of monthly returns of random forest portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	4.31	3.94	5.29	5.97	6.53	8.16	9.11	9.48	11.27	14.90
	Ann Volatility	13.34	8.17	6.29	5.98	5.49	5.36	5.79	5.86	6.59	9.08
	Sharpe Ratio	0.21	0.24	0.49	0.62	0.76	1.06	1.14	1.18	1.29	1.31
	Skewness	-1.56	-3.34	-2.80	-2.41	-2.24	-1.48	-0.81	-0.74	0.11	-0.03
	Excess Kurtosis	11.15	21.52	14.16	11.18	9.16	4.54	2.78	2.56	1.72	1.18
	Max Draw %	46.15	36.77	27.38	27.86	25.28	20.82	20.24	16.63	13.22	18.25
6-month	Ann Return %	5.10	5.89	5.82	6.26	6.66	7.83	7.77	8.16	9.96	14.25
	Ann Volatility	11.47	6.79	6.39	5.76	5.47	5.56	5.37	5.80	6.30	8.95
	Sharpe Ratio	0.29	0.54	0.56	0.69	0.79	0.97	0.99	0.98	1.17	1.27
	Skewness	-1.62	-1.61	-2.22	-2.32	-2.48	-1.52	-1.22	-1.19	-0.88	-0.42
	Excess Kurtosis	12.73	8.85	12.44	10.20	11.81	5.26	3.35	3.06	1.94	0.58
	Max Draw %	43.94	30.11	30.96	28.22	26.63	22.68	19.43	19.75	14.90	15.10
12-month	Ann Return %	8.41	7.21	6.41	6.74	6.71	6.53	5.70	5.92	7.11	9.14
	Ann Volatility	9.22	5.94	5.60	5.31	5.33	5.33	5.69	5.55	6.60	8.98
	Sharpe Ratio	0.67	0.81	0.73	0.82	0.82	0.78	0.60	0.65	0.73	0.77
	Skewness	-0.37	-0.89	-1.42	-1.52	-1.78	-1.76	-1.94	-1.68	-1.82	-1.11
	Excess Kurtosis	9.83	6.75	8.72	6.77	7.16	6.62	7.90	5.76	8.74	3.25
	Max Draw %	35.77	27.03	27.27	23.49	22.60	22.03	24.92	22.83	17.94	20.44

Table B5. Statistics of monthly returns of gradient boosting portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	4.27	3.90	5.65	6.05	6.83	8.06	9.01	9.23	11.23	14.75
	Ann Volatility	12.71	7.79	5.99	5.95	6.25	5.24	5.51	5.68	6.56	9.29
	Sharpe Ratio	0.21	0.24	0.57	0.64	0.73	1.07	1.18	1.17	1.29	1.27
	Skewness	-1.12	-3.69	-2.14	-2.78	-2.77	-1.64	-1.07	-0.67	0.06	-0.17
	Excess Kurtosis	8.03	25.35	10.58	15.24	14.87	6.38	3.02	2.22	1.57	1.14
	Max Draw %	42.02	33.87	25.86	27.73	29.13	22.47	19.96	17.27	15.49	21.56
6-month	Ann Return %	5.45	5.79	6.33	5.66	7.01	7.87	7.76	8.92	10.05	12.83
	Ann Volatility	10.92	6.53	5.94	6.03	5.33	5.54	5.44	5.87	6.43	8.94
	Sharpe Ratio	0.33	0.54	0.68	0.57	0.87	0.98	0.98	1.09	1.16	1.14
	Skewness	-1.51	-1.68	-2.58	-2.71	-1.68	-1.24	-1.46	-0.99	-0.80	-0.74
	Excess Kurtosis	9.74	9.21	12.90	15.15	6.66	4.64	4.95	3.40	1.88	1.55
	Max Draw %	41.58	30.44	28.86	31.86	26.66	21.30	19.40	15.56	17.29	17.47
12-month	Ann Return %	8.60	7.01	6.58	6.15	6.65	6.44	6.14	6.99	6.78	8.64
	Ann Volatility	9.22	5.70	5.38	5.06	5.01	5.16	5.75	5.52	6.75	9.06
	Sharpe Ratio	0.69	0.81	0.79	0.75	0.85	0.79	0.67	0.84	0.67	0.71
	Skewness	-0.37	-0.97	-1.57	-1.74	-1.51	-1.76	-2.11	-1.45	-1.83	-1.08
	Excess Kurtosis	7.98	6.77	9.61	8.92	6.19	6.22	9.15	4.54	8.86	2.67
	Max Draw %	34.09	26.68	25.36	22.82	22.24	22.45	21.76	20.55	20.09	22.12

Table B6. Statistics of monthly returns of deep neural network portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	4.82	4.55	5.82	6.44	6.63	7.56	8.51	9.33	11.55	14.00
	Ann Volatility	11.93	7.94	6.95	6.46	6.22	6.31	5.41	5.43	5.96	6.87
	Sharpe Ratio	0.26	0.31	0.52	0.65	0.70	0.83	1.11	1.25	1.47	1.59
	Skewness	-1.41	-2.09	-1.90	-2.04	-1.90	-2.09	-0.91	-0.84	-0.59	0.28
	Excess Kurtosis	8.27	11.29	8.36	10.42	7.90	9.17	2.31	2.91	2.58	2.47
	Max Draw %	44.08	35.27	30.26	30.03	26.33	26.18	19.30	17.57	14.17	11.95
6-month	Ann Return %	4.88	6.54	6.17	6.73	7.26	7.16	8.20	9.00	9.77	12.11
	Ann Volatility	10.66	7.05	6.42	6.13	5.86	5.75	5.30	5.75	6.00	8.34
	Sharpe Ratio	0.29	0.61	0.61	0.72	0.84	0.83	1.08	1.13	1.20	1.13
	Skewness	-1.85	-2.46	-1.97	-1.98	-2.18	-1.62	-1.41	-1.13	-1.15	-0.37
	Excess Kurtosis	16.23	14.33	10.50	9.05	10.45	5.45	4.32	2.92	2.89	0.36
	Max Draw %	45.93	32.51	29.90	27.96	24.27	24.27	17.47	17.84	13.60	16.52
12-month	Ann Return %	8.39	6.54	6.09	6.69	6.63	6.64	6.98	7.12	6.88	8.18
	Ann Volatility	7.39	6.08	5.38	5.22	5.43	5.26	5.32	5.77	6.39	9.11
	Sharpe Ratio	0.81	0.70	0.70	0.83	0.79	0.81	0.86	0.82	0.72	0.66
	Skewness	-0.52	-1.59	-1.55	-1.73	-1.84	-1.59	-1.45	-1.62	-1.47	-1.33
	Excess Kurtosis	6.96	9.44	7.23	7.41	8.26	5.69	4.64	5.79	5.24	4.45
	Max Draw %	27.87	28.27	25.51	23.59	24.48	22.91	21.05	21.68	21.81	21.44

Appendix C. Statistics of monthly returns of the macro fund portfolios

Table C1. Statistics of monthly returns of ABK portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	4.91	8.48	6.51	7.52	7.55	7.35	6.64	6.46	6.05	8.47
	Ann Volatility	7.42	8.78	9.22	9.11	9.06	8.81	7.76	6.77	5.48	3.91
	Sharpe Ratio	0.36	0.70	0.47	0.58	0.59	0.58	0.57	0.61	0.67	1.50
	Skewness	0.96	0.96	0.73	0.38	0.29	0.29	0.04	0.08	0.04	-0.20
	Excess Kurtosis	4.30	2.14	2.13	0.39	0.36	0.54	0.19	0.58	0.06	0.94
	Max Draw %	21.49	12.30	9.21	10.27	10.72	11.03	10.14	8.37	7.14	5.20
6-month	Ann Return %	4.77	7.41	6.55	7.32	7.42	7.19	6.82	6.03	6.11	8.05
	Ann Volatility	6.67	8.41	8.89	8.65	8.56	8.43	7.47	6.33	5.06	3.70
	Sharpe Ratio	0.38	0.61	0.49	0.59	0.60	0.59	0.61	0.58	0.73	1.48
	Skewness	1.02	0.95	0.53	0.42	0.33	0.28	0.05	0.12	0.01	-0.17
	Excess Kurtosis	3.47	2.30	1.26	0.59	0.44	0.36	0.52	0.28	0.02	0.95
	Max Draw %	17.17	9.44	9.10	10.04	9.72	13.24	9.55	8.91	6.53	4.44
12-month	Ann Return %	4.75	6.25	5.97	6.79	6.47	6.26	6.43	5.58	6.30	7.06
	Ann Volatility	5.94	7.67	8.09	8.11	8.12	8.03	6.93	6.04	4.88	3.71
	Sharpe Ratio	0.41	0.52	0.46	0.56	0.52	0.50	0.59	0.54	0.80	1.23
	Skewness	1.18	0.76	0.41	0.47	0.23	0.24	0.20	0.05	0.13	-0.01
	Excess Kurtosis	3.10	2.18	1.03	1.04	0.75	0.66	0.37	0.33	-0.12	0.96
	Max Draw %	9.26	8.01	11.35	9.98	10.12	11.04	9.84	7.83	5.16	5.15

Table C2. Statistics of monthly returns of OLS portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	11.25	8.10	7.10	7.32	6.32	6.77	6.02	5.51	5.50	6.46
	Ann Volatility	10.72	7.61	6.91	6.10	5.61	6.16	6.58	7.30	8.76	11.85
	Sharpe Ratio	0.80	0.74	0.68	0.80	0.70	0.71	0.56	0.45	0.39	0.39
	Skewness	1.71	1.08	0.94	0.73	0.46	0.47	0.18	0.14	0.16	0.09
	Excess Kurtosis	10.11	3.70	2.42	1.24	1.11	0.68	0.73	0.39	1.12	1.41
	Max Draw %	15.16	7.16	6.40	6.45	5.85	6.95	10.15	10.81	15.84	18.60
6-month	Ann Return %	6.93	6.28	6.03	6.23	6.49	6.22	6.67	7.09	7.51	9.87
	Ann Volatility	8.37	6.81	6.45	5.71	5.95	6.14	6.32	7.56	8.18	11.15
	Sharpe Ratio	0.56	0.58	0.57	0.67	0.69	0.63	0.68	0.63	0.64	0.69
	Skewness	0.73	0.64	0.64	0.50	0.63	0.31	0.39	0.22	0.29	0.07
	Excess Kurtosis	3.32	2.97	2.60	1.12	1.24	0.59	0.25	0.37	0.58	1.38
	Max Draw %	8.06	6.99	8.25	6.82	8.54	7.41	7.30	8.84	8.86	16.88
12-month	Ann Return %	6.57	5.76	5.66	5.55	5.62	5.50	5.91	6.42	6.40	8.46
	Ann Volatility	7.07	5.27	5.11	4.98	5.34	5.72	6.43	6.95	8.14	11.19
	Sharpe Ratio	0.60	0.64	0.64	0.64	0.61	0.55	0.56	0.59	0.51	0.57
	Skewness	0.30	0.52	0.47	0.85	0.65	0.58	0.28	0.04	0.05	-0.06
	Excess Kurtosis	0.89	1.56	1.22	1.85	1.63	1.60	0.77	0.56	0.81	1.29
	Max Draw %	11.23	8.38	6.59	6.40	5.57	7.32	9.01	9.73	12.30	18.43

Table C3. Statistics of monthly returns of the lasso portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	9.30	7.75	7.22	7.03	5.45	5.70	5.33	6.01	6.70	9.82
	Ann Volatility	10.60	7.88	6.77	6.38	5.91	5.89	6.61	7.08	8.74	11.91
	Sharpe Ratio	0.66	0.68	0.71	0.73	0.53	0.57	0.46	0.53	0.52	0.64
	Skewness	1.25	0.97	0.60	0.61	0.52	0.20	0.24	0.18	0.03	0.26
	Excess Kurtosis	6.56	3.01	2.27	1.57	0.88	0.52	1.00	1.07	1.43	1.86
	Max Draw %	12.10	7.10	7.98	6.96	6.26	8.36	10.16	10.68	13.67	14.48
6-month	Ann Return %	7.47	6.36	5.82	6.05	6.50	5.87	6.63	6.24	7.51	10.78
	Ann Volatility	8.79	6.67	6.22	5.74	5.92	6.00	6.88	7.66	8.46	11.33
	Sharpe Ratio	0.59	0.60	0.56	0.64	0.69	0.59	0.62	0.52	0.62	0.75
	Skewness	0.75	0.74	0.71	0.49	0.50	0.28	0.38	0.17	0.17	0.09
	Excess Kurtosis	4.27	2.68	2.95	0.96	0.66	0.44	0.68	0.88	1.07	1.79
	Max Draw %	12.67	8.37	7.17	7.65	7.38	8.63	10.29	11.05	10.94	17.63
12-month	Ann Return %	6.88	5.94	5.93	5.90	5.53	5.61	6.13	5.63	6.55	7.74
	Ann Volatility	6.90	5.74	5.50	5.52	5.22	5.62	6.22	7.04	7.82	11.09
	Sharpe Ratio	0.66	0.62	0.65	0.64	0.61	0.58	0.61	0.48	0.55	0.52
	Skewness	0.16	0.39	0.63	0.58	0.53	0.54	0.37	0.10	0.11	-0.06
	Excess Kurtosis	2.26	2.45	2.47	1.69	1.50	1.34	0.81	1.17	0.79	1.12
	Max Draw %	11.04	9.42	7.07	6.56	5.62	7.05	9.23	9.88	13.89	18.48

Table C4. Statistics of monthly returns of random forest portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	10.03	8.14	7.64	6.42	6.05	6.11	6.07	4.66	7.05	8.20
	Ann Volatility	10.56	7.30	6.22	5.97	5.94	6.22	6.81	7.37	8.64	11.88
	Sharpe Ratio	0.72	0.78	0.83	0.68	0.62	0.61	0.55	0.34	0.56	0.52
	Skewness	1.19	1.04	0.82	0.65	0.15	0.43	0.24	-0.06	-0.17	0.14
	Excess Kurtosis	3.71	2.65	2.21	1.70	1.17	0.26	0.98	1.31	1.53	1.32
	Max Draw %	13.40	7.61	5.83	7.06	6.23	8.75	11.29	9.53	13.18	13.49
6-month	Ann Return %	7.41	6.13	5.56	5.79	5.83	6.68	6.56	7.19	7.02	11.23
	Ann Volatility	7.42	6.42	6.01	5.97	6.14	6.41	6.82	7.35	8.05	11.63
	Sharpe Ratio	0.68	0.59	0.54	0.58	0.57	0.67	0.62	0.66	0.59	0.77
	Skewness	0.65	0.84	0.70	0.37	0.24	0.40	0.31	0.33	0.07	0.16
	Excess Kurtosis	1.86	2.67	2.02	0.91	0.38	0.80	0.48	0.73	0.43	0.94
	Max Draw %	8.73	6.12	7.23	6.85	7.57	7.81	10.32	9.71	11.41	14.97
12-month	Ann Return %	6.96	5.57	5.43	5.77	5.50	5.72	5.66	6.24	6.28	8.80
	Ann Volatility	6.73	5.33	5.36	5.42	5.35	5.73	6.20	7.07	8.08	10.57
	Sharpe Ratio	0.68	0.60	0.58	0.63	0.59	0.59	0.54	0.56	0.50	0.63
	Skewness	0.49	0.33	0.41	0.59	0.35	0.28	0.31	0.10	0.10	0.20
	Excess Kurtosis	2.16	1.58	1.30	1.34	0.49	0.84	1.10	0.80	1.37	1.26
	Max Draw %	8.76	6.49	6.04	5.93	6.31	7.90	10.40	12.92	14.42	14.17

Table C5. Statistics of monthly returns of gradient boosting portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	10.85	7.50	7.15	7.16	6.30	6.43	5.96	5.70	5.60	7.71
	Ann Volatility	10.40	7.20	6.84	6.24	6.23	6.24	6.75	7.49	8.54	11.77
	Sharpe Ratio	0.80	0.71	0.69	0.76	0.63	0.65	0.54	0.46	0.41	0.49
	Skewness	1.15	0.81	1.05	0.86	0.41	0.32	0.26	0.18	-0.10	0.19
	Excess Kurtosis	4.84	2.76	3.13	2.47	0.91	0.52	1.09	1.04	1.02	1.84
	Max Draw %	10.54	7.84	5.67	5.60	6.65	7.66	9.73	11.49	12.41	15.44
6-month	Ann Return %	7.51	5.86	5.83	5.84	5.83	6.51	6.48	7.18	6.91	11.42
	Ann Volatility	7.46	6.42	6.01	6.09	6.24	6.23	6.83	7.27	8.00	11.64
	Sharpe Ratio	0.69	0.55	0.58	0.58	0.56	0.67	0.61	0.66	0.58	0.78
	Skewness	0.71	0.62	0.76	0.41	0.26	0.34	0.35	0.33	0.08	0.15
	Excess Kurtosis	1.92	2.25	1.91	1.03	0.58	0.67	0.59	0.70	0.42	0.88
	Max Draw %	8.93	6.22	6.91	7.50	8.23	7.09	10.92	9.62	11.15	14.88
12-month	Ann Return %	6.65	5.55	5.98	6.06	6.03	5.61	5.85	5.98	6.34	7.84
	Ann Volatility	6.40	5.62	5.25	5.69	5.84	6.06	6.22	6.89	7.78	10.91
	Sharpe Ratio	0.67	0.57	0.68	0.65	0.63	0.54	0.57	0.54	0.53	0.53
	Skewness	0.40	0.38	0.61	0.68	0.60	0.25	0.22	0.22	0.11	-0.07
	Excess Kurtosis	1.42	2.50	2.20	1.88	1.57	0.85	0.51	0.94	0.85	1.08
	Max Draw %	9.76	8.11	6.96	6.70	7.07	9.15	8.55	10.29	11.14	17.90

Table C6. Statistics of monthly returns of deep neural network portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	10.39	7.95	6.92	6.53	6.01	5.86	6.12	5.14	6.66	8.76
	Ann Volatility	10.80	7.79	6.95	6.23	5.84	6.06	6.58	7.28	8.71	11.57
	Sharpe Ratio	0.73	0.71	0.66	0.67	0.63	0.58	0.58	0.40	0.51	0.58
	Skewness	1.48	1.10	0.55	0.62	0.39	0.35	0.20	0.03	0.01	0.31
	Excess Kurtosis	6.63	3.46	1.63	1.29	0.47	0.62	0.63	1.16	1.59	2.13
	Max Draw %	11.63	7.75	6.71	5.75	7.25	8.27	10.29	12.26	12.73	15.05
6-month	Ann Return %	7.84	5.95	5.41	5.89	6.06	6.19	6.57	7.07	7.08	11.32
	Ann Volatility	7.56	6.22	5.97	6.18	6.21	6.19	6.76	7.43	7.99	11.68
	Sharpe Ratio	0.72	0.58	0.52	0.58	0.60	0.62	0.63	0.64	0.60	0.77
	Skewness	0.78	0.69	0.45	0.39	0.27	0.41	0.32	0.25	0.13	0.16
	Excess Kurtosis	2.31	2.06	1.58	0.81	0.76	0.79	0.75	0.52	0.57	0.91
	Max Draw %	8.24	6.76	7.30	7.36	6.39	8.53	10.27	9.48	10.78	14.97
12-month	Ann Return %	6.36	5.94	5.67	5.67	5.35	5.68	5.94	6.19	7.09	8.01
	Ann Volatility	6.88	6.08	5.68	5.50	5.40	5.82	6.20	7.00	7.77	10.60
	Sharpe Ratio	0.59	0.59	0.58	0.60	0.56	0.57	0.58	0.56	0.62	0.56
	Skewness	0.27	0.68	0.60	0.77	0.53	0.56	0.20	0.09	0.07	-0.21
	Excess Kurtosis	2.16	3.38	2.50	2.19	1.08	1.31	0.93	0.86	0.98	0.93
	Max Draw %	11.90	9.70	7.95	5.57	6.05	7.55	9.92	9.92	12.99	18.43

Appendix D. Statistics of monthly returns of the relative-value fund portfolios

Table D1. Statistics of monthly returns of ABK portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	0.38	3.38	5.37	7.10	8.43	9.38	8.63	8.98	9.45	8.29
	Ann Volatility	7.03	7.35	6.94	5.81	5.58	4.89	4.66	2.98	2.50	2.44
	Sharpe Ratio	-0.25	0.17	0.46	0.83	1.08	1.40	1.33	2.13	2.69	2.34
	Skewness	-3.47	-2.31	-3.19	-3.50	-3.33	-2.75	-4.06	-2.01	-1.00	-2.99
	Excess Kurtosis	22.63	13.38	24.16	25.79	28.56	21.46	28.29	8.24	4.10	18.18
	Max Draw %	36.78	33.28	33.91	27.31	23.45	18.77	21.56	8.74	4.92	6.69
6-month	Ann Return %	1.54	3.69	5.57	6.71	7.91	8.53	8.29	8.51	8.30	7.66
	Ann Volatility	6.21	6.68	6.47	5.74	5.27	4.55	4.21	3.22	2.55	2.37
	Sharpe Ratio	-0.10	0.23	0.52	0.77	1.04	1.32	1.38	1.85	2.23	2.15
	Skewness	-3.48	-2.84	-3.40	-3.27	-2.67	-2.20	-3.58	-2.64	-1.85	-2.42
	Excess Kurtosis	23.65	17.88	25.23	23.97	20.05	14.06	23.68	13.11	7.11	11.18
	Max Draw %	32.73	32.52	32.99	27.26	21.47	16.82	18.63	11.52	6.58	5.68
12-month	Ann Return %	2.43	3.87	5.54	5.96	6.96	7.53	7.71	7.43	6.83	6.87
	Ann Volatility	5.28	5.62	5.59	5.53	4.91	4.45	4.12	3.58	2.81	2.40
	Sharpe Ratio	0.04	0.29	0.59	0.67	0.93	1.14	1.28	1.39	1.55	1.82
	Skewness	-3.35	-2.06	-3.03	-3.14	-2.36	-2.32	-3.40	-3.15	-3.45	-2.72
	Excess Kurtosis	22.83	11.15	18.94	21.73	15.98	14.66	23.70	17.10	21.15	14.65
	Max Draw %	27.71	27.45	27.89	25.17	19.31	17.82	17.60	14.72	10.84	7.52

Table D2. Statistics of monthly returns of OLS portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	3.12	4.79	5.20	5.87	5.87	6.48	7.81	8.05	9.40	13.56
	Ann Volatility	7.18	4.51	4.16	3.83	3.97	4.23	3.84	5.19	5.61	8.74
	Sharpe Ratio	0.14	0.55	0.69	0.91	0.88	0.98	1.39	1.10	1.24	1.28
	Skewness	-1.88	-2.36	-2.62	-2.48	-3.38	-4.95	-3.26	-4.18	-3.52	-3.89
	Excess Kurtosis	14.03	14.59	14.61	12.72	22.27	44.39	25.66	34.21	25.54	32.88
	Max Draw %	32.87	22.51	21.90	18.01	19.04	18.80	14.50	20.56	17.57	28.02
6-month	Ann Return %	4.32	4.25	4.57	5.15	5.79	6.56	7.05	8.44	9.13	13.60
	Ann Volatility	6.76	4.93	4.85	4.40	4.45	3.89	3.98	3.86	4.64	6.21
	Sharpe Ratio	0.32	0.40	0.48	0.65	0.79	1.07	1.16	1.52	1.41	1.71
	Skewness	-3.12	-4.13	-4.82	-4.86	-5.30	-3.22	-2.76	-1.49	-1.59	-0.47
	Excess Kurtosis	18.81	27.48	36.39	38.01	45.13	22.42	18.15	7.45	9.21	2.50
	Max Draw %	32.30	25.88	26.71	23.36	23.49	16.71	15.41	11.25	13.69	10.75
12-month	Ann Return %	6.31	5.24	5.84	5.42	5.39	5.77	5.10	6.69	6.41	8.72
	Ann Volatility	6.37	4.94	4.24	4.31	3.74	3.69	4.01	3.70	3.78	6.66
	Sharpe Ratio	0.63	0.59	0.82	0.72	0.81	0.92	0.69	1.15	1.05	0.96
	Skewness	-1.46	-2.82	-2.50	-4.45	-3.64	-4.22	-4.01	-2.44	-1.77	-2.79
	Excess Kurtosis	11.13	19.88	18.00	38.48	26.88	34.58	27.04	11.92	6.89	14.11
	Max Draw %	20.52	21.57	17.83	21.20	18.58	17.86	19.22	15.98	13.93	24.54

Table D3. Statistics of monthly returns of the lasso portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	-0.22	3.51	4.35	4.55	5.84	7.15	8.29	9.23	11.38	16.63
	Ann Volatility	8.67	5.14	4.82	4.64	4.18	3.65	3.91	4.10	4.70	7.49
	Sharpe Ratio	-0.26	0.25	0.44	0.49	0.84	1.29	1.48	1.60	1.82	1.79
	Skewness	-3.51	-4.03	-5.12	-4.74	-5.00	-3.03	-2.61	-0.61	-0.31	-1.02
	Excess Kurtosis	23.24	30.26	45.71	36.18	40.95	21.65	16.05	6.69	4.05	5.84
	Max Draw %	46.27	29.16	24.98	23.77	22.11	15.47	15.12	11.12	6.81	15.52
6-month	Ann Return %	2.26	4.17	4.94	5.12	5.67	6.58	7.62	8.61	9.26	14.71
	Ann Volatility	7.70	5.44	5.31	4.65	4.18	3.93	3.56	3.62	4.24	6.06
	Sharpe Ratio	0.03	0.36	0.51	0.62	0.80	1.06	1.43	1.66	1.56	1.91
	Skewness	-4.00	-4.89	-5.89	-5.59	-4.04	-3.19	-2.06	-1.31	-1.14	-0.58
	Excess Kurtosis	30.26	42.06	53.50	48.83	29.56	21.16	9.85	3.76	3.19	1.70
	Max Draw %	38.44	29.09	28.38	25.01	21.60	17.15	12.42	9.73	8.62	8.80
12-month	Ann Return %	4.85	5.51	5.29	5.05	5.62	5.49	6.12	7.08	6.77	9.17
	Ann Volatility	6.99	4.88	4.56	4.19	3.96	3.72	3.67	3.44	3.85	5.87
	Sharpe Ratio	0.38	0.65	0.66	0.65	0.83	0.84	1.01	1.34	1.13	1.14
	Skewness	-1.83	-2.79	-4.14	-3.59	-4.01	-3.69	-2.76	-2.18	-2.20	-2.25
	Excess Kurtosis	14.62	22.73	37.21	28.71	31.47	26.36	16.46	9.88	9.57	9.52
	Max Draw %	30.13	24.12	24.29	22.81	18.56	17.18	13.48	10.83	13.34	17.79

Table D4. Statistics of monthly returns of random forest portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	-0.78	2.17	3.22	4.79	5.74	7.65	8.28	10.01	12.05	17.74
	Ann Volatility	8.02	6.28	5.90	5.06	4.12	3.57	3.84	3.99	4.45	7.21
	Sharpe Ratio	-0.36	0.00	0.18	0.51	0.82	1.44	1.49	1.83	2.05	1.99
	Skewness	-2.75	-6.59	-6.69	-5.47	-3.52	-2.79	-1.46	-1.01	-0.13	-1.23
	Excess Kurtosis	17.49	67.55	68.25	49.26	23.81	18.24	6.17	8.82	4.96	8.29
	Max Draw %	42.48	32.82	30.76	24.96	20.92	14.13	10.44	10.16	4.82	17.62
6-month	Ann Return %	2.45	4.34	5.04	5.51	5.83	6.43	6.78	8.36	9.40	14.76
	Ann Volatility	7.93	5.77	5.37	4.68	4.17	3.84	3.44	3.64	4.20	6.03
	Sharpe Ratio	0.05	0.37	0.52	0.69	0.84	1.05	1.26	1.59	1.62	1.93
	Skewness	-3.51	-4.44	-4.38	-4.81	-3.69	-2.77	-2.58	-1.69	-1.43	-1.05
	Excess Kurtosis	32.78	39.62	40.20	41.61	26.35	14.36	13.12	4.60	3.53	3.51
	Max Draw %	39.81	31.39	28.69	25.55	19.02	15.55	12.84	8.74	7.77	10.52
12-month	Ann Return %	4.86	4.96	5.40	5.74	5.54	5.93	5.95	6.18	7.12	9.32
	Ann Volatility	5.66	5.08	4.48	4.23	3.94	3.92	3.77	3.66	3.88	6.54
	Sharpe Ratio	0.46	0.54	0.69	0.80	0.81	0.91	0.95	1.03	1.20	1.06
	Skewness	-1.36	-4.82	-3.75	-3.22	-3.21	-3.14	-3.15	-2.14	-1.79	-2.79
	Excess Kurtosis	14.34	45.80	31.33	25.37	21.41	18.68	20.11	7.78	6.30	14.76
	Max Draw %	21.88	22.91	20.80	20.20	19.36	20.28	16.03	14.77	10.44	25.25

Table D5. Statistics of monthly returns of gradient boosting portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	0.61	2.46	4.03	4.92	5.50	6.78	8.35	9.64	11.81	16.78
	Ann Volatility	8.00	6.18	5.01	4.75	4.55	4.23	3.61	3.71	4.49	7.09
	Sharpe Ratio	-0.18	0.05	0.36	0.56	0.70	1.04	1.61	1.87	1.99	1.90
	Skewness	-2.84	-4.90	-5.17	-6.10	-4.75	-4.15	-1.87	-0.40	-0.59	-1.18
	Excess Kurtosis	20.33	41.09	44.36	60.01	39.95	32.75	9.89	5.00	3.96	7.23
	Max Draw %	41.27	32.96	27.03	23.38	22.29	19.60	13.04	8.52	6.10	15.64
6-month	Ann Return %	2.15	4.82	5.01	5.30	6.10	6.54	7.65	8.83	9.60	13.03
	Ann Volatility	8.32	5.88	5.19	4.63	3.93	3.84	3.42	3.58	4.28	5.89
	Sharpe Ratio	0.02	0.45	0.54	0.65	0.95	1.08	1.50	1.73	1.62	1.73
	Skewness	-3.40	-4.65	-5.29	-4.59	-4.17	-2.89	-2.11	-1.41	-1.48	-1.52
	Excess Kurtosis	32.14	43.12	47.47	39.96	32.46	16.43	7.31	3.15	4.24	5.25
	Max Draw %	40.98	30.14	27.83	24.49	20.17	16.73	10.39	6.51	7.54	13.96
12-month	Ann Return %	5.21	5.34	5.35	5.42	6.17	5.54	5.69	6.73	6.92	8.63
	Ann Volatility	5.77	4.76	4.48	3.93	4.02	4.06	3.68	3.67	3.68	6.56
	Sharpe Ratio	0.51	0.64	0.68	0.78	0.95	0.79	0.90	1.17	1.21	0.96
	Skewness	-1.78	-2.92	-3.98	-2.95	-3.66	-3.47	-3.08	-2.31	-1.76	-3.01
	Excess Kurtosis	14.31	26.45	33.51	21.09	27.06	22.38	17.33	9.40	6.38	16.43
	Max Draw %	22.70	20.31	21.08	18.22	18.83	21.49	18.10	13.44	12.48	24.75

Table D6. Statistics of monthly returns of deep neural network portfolios

Rebalance	Decile Portfolio	1	2	3	4	5	6	7	8	9	10
3-month	Ann Return %	-0.70	2.27	3.64	4.55	5.43	7.70	8.59	9.38	12.45	17.59
	Ann Volatility	8.11	6.37	5.81	5.08	4.02	3.72	3.65	4.17	4.45	7.12
	Sharpe Ratio	-0.35	0.02	0.25	0.46	0.77	1.40	1.64	1.62	2.13	2.00
	Skewness	-2.92	-6.18	-6.19	-5.80	-3.58	-2.65	-1.33	-1.37	-0.08	-1.30
	Excess Kurtosis	19.71	60.91	58.69	54.06	23.90	17.88	6.56	9.75	4.85	9.15
	Max Draw %	43.72	32.93	30.15	26.77	19.02	13.88	10.92	10.89	4.89	17.39
6-month	Ann Return %	2.18	4.75	4.68	5.83	5.63	6.38	7.04	8.11	9.78	14.49
	Ann Volatility	8.02	5.67	5.61	4.40	4.42	3.64	3.56	3.50	4.16	6.14
	Sharpe Ratio	0.02	0.45	0.44	0.80	0.75	1.09	1.28	1.59	1.71	1.86
	Skewness	-3.54	-3.76	-5.28	-3.99	-4.17	-2.85	-2.50	-1.72	-1.47	-1.19
	Excess Kurtosis	32.06	33.62	49.92	32.58	33.25	13.98	12.34	4.41	4.00	4.02
	Max Draw %	40.89	28.88	30.67	22.83	22.79	14.18	12.67	8.72	7.28	11.62
12-month	Ann Return %	4.99	5.15	5.19	5.62	5.52	6.05	5.91	6.78	6.64	9.13
	Ann Volatility	5.83	4.71	4.56	4.32	4.08	3.77	3.70	3.55	3.98	6.53
	Sharpe Ratio	0.47	0.61	0.63	0.76	0.78	0.97	0.95	1.22	1.06	1.03
	Skewness	-1.48	-4.23	-3.88	-3.62	-3.52	-2.74	-3.01	-2.10	-2.12	-2.60
	Excess Kurtosis	14.99	37.98	33.93	27.71	24.39	15.68	18.25	7.84	8.03	12.93
	Max Draw %	22.83	20.53	20.83	20.81	21.00	17.78	18.11	12.26	12.67	25.00