International Journal of Economics, Business and Management Research

Vol. 6, No.12; 2022

ISSN: 2456-7760

Portfolio Selection Using Multiple Factors: A Machine Learning Approach

Bilal Elmsili[1], Benaceur Outtaj[2]

[1]Laboratory of Economic Analysis and Modeling (LEAM) Faculty of Law, Economics and Social Sciences Souissi, Mohammed V University in Rabat, Morocco
bilal_elmsili@um5.ac.ma

[2]Laboratory of Economic Analysis and Modeling (LEAM) Faculty of Law, Economics and Social Sciences Souissi, Mohammed V University in Rabat, Morocco
b.outtaj@um5s.net.ma

doi: 10.51505/ IJEBMR.2022.61207 URL: https://doi.org/10.51505/IJEBMR.2022.61207

Abstract

In this paper we present a novel framework for portfolio selection using deep neural networks and elastic net regularization: At the beginning of each month T, we follow a three-step methodology. First, for each stock, we use the previous seven years of data in order to compute over 36 firm-specific factors. Second, we perform features selection using elastic net regularization. Finally, we train a deep neural network in order to learn portfolio weights and hold this portfolio until the end of the month T. Compared with momentum, long-term reversal, and short-term reversal strategies, our approach demonstrates a superior performance in terms of the monthly rate of return (2% versus 1.22% for long-term reversal, 1.15% for momentum, and only 0.68% for short-term reversal), Sharpe ratio (21.67% versus 19.31% for momentum, 15.51% for long-term reversal, and 8.69% for short-term reversal), and the monthly risk-adjusted return (1.85% versus 0.74% for momentum, 0.72% for long-term reversal, and 0.31% for short-term reversal). The results of our approach are all statistically significant at 1% level.

Keywords: portfolio selection, characteristic-sorted portfolios, factors selection, machine learning, deep learning, elastic net regularization, Casablanca Stock Exchange

1. Introduction

Selecting portfolios based on one or more factors (i.e. idiosyncratic characteristics) is ubiquitous in empirical finance. The premise behind it is to discover whether the expected returns of an asset are related to a certain characteristic. A natural way to investigate this is to sort assets by the characteristic value, divide them into portfolios, and then test whether the differences in average return across these portfolios are statistically significant or not. This methodology has been widely used in order to identify a number of market anomalies and to establish profitable trading strategies(Cattaneo, Crump, Farrell, & Schaumburg, 2016).

Applications of this approach are too numerous to list, but some of the influential work includes:

- Momentum portfolios: Jegadeesh and Titman (1993)selected portfolios based on the intermediate horizon past returns (i.e. cumulative return over the previous three-to-twelve months). Holding these portfolios for the next three-to-twelve months, they observed that portfolios constructed with the winners of the recent past outperform portfolios formed with previous losers.

- Enhanced momentum: The traditional momentum effect (Jegadeesh & Titman, 1993) can be enhanced by considering the interaction of the formation period returns with certain firm-level characteristics. Some of these characteristics are the nearness to the 52-Week (George & Hwang, 2004), formation period return consistency (Grinblatt & Moskowitz, 2004), volatility (Bandarchuk & Hilscher, 2013; Jiang, Lee, & Zhang, 2005; Zhang, 2006), intermediate past performance (Novy-Marx, 2012), extreme past returns (Bandarchuk & Hilscher, 2013), information discreteness (Da, Gurun, & Warachka, 2014), continuing overreaction (Byun, Lim, & Yun, 2016), and R-squared (Hou, Xiong, & Peng, 2006).
- Long-term return sorted portfolios: In contrast to Jegadeesh and Titman (1993), Bondt and Thaler (1985) and McLean (2010) document a long-term reversal phenomenon based on a stock's three-to-five years cumulative return. Thus, portfolios of stocks that outperform in the previous three-to-five years underperform in the following three-to-five years.
- Short-term return sorted portfolios: Jegadeesh (1990), Lehmann (1990), Da, Liu, and Schaumburg (2013) document negative short-term return autocorrelations. They demonstrated that the previous short-term return (one-to-four week) tend the reverse during the following month. As a consequence, abnormal returns could be obtained by investing in stocks that performed poorly in the near past.
- Technical analysis: Studies in empirical finance literature suggest several potential technical trading rules. For example, Han, Yang, and Zhou (2013)document that a moving average timing strategy outperforms the buy and hold strategy. Specifically, stocks with a price above (below) the n-day (10-, 20-, 50-, 100-, 200-day) moving average price outperform (underperform).
- Beta sorted portfolios: On the contrary to the Capital Asset Pricing Model (i.e. CAPM) prediction, according to which, in an efficient market, investors realize above-average returns only by taking above-average risks, Baker et al. (2011), Frazzini and Pedersen (2014), Hong and Sraer (2016) document that low-beta stocks outperform high-beta stocks.
- Lottery-type portfolios: Some recent studies argue that stocks with lottery-type characteristics tend to underperform. Kumar (2009), Boyer, Mitton, and Vorkink (2010)used the idiosyncratic skewness of returns distribution as a proxy to the lottery-type feature of a stock. In contrast, Bali, Cakici, and Whitelaw (2011) proposed to rank stocks based on the maximum daily return over the past one month instead of idiosyncratic skewness.

In fact, these approaches use typically one or two factors in order to construct a portfolio. However, and more recently, evidence has been established that machine learning techniques and specifically deep neural networks are capable of identifying complex patterns in the financial market by incorporating various types of explanatory variables (i.e. factors or return predictive signals). For reference, see Krauss, Do, and Huck (2017); Abe and Nakayama (2018); as well as Fischer and Krauss (2018). However, most of these applications of learning algorithms to financial data are based on predicting the value of output variables given input variables, and then use these predictions (often noisy) in order to form a portfolio. Bengio (1997) argues that better results can be achieved by learning the model parameters that directly optimize the financial criterion of interest (e.g. maximizing the portfolio rate of return, minimizing the portfolio volatility, or combining these two criteria by maximizing the reward-to-risk ratio, i.e. Sharpe ratio).

In this paper, we primarily present a novel framework for portfolio selection using deep neural networks and elastic net regularization. Our contribution to the existing literature is two folds. First, we follow the recommendation of Bengio (1997) and we train our deep neural network to directly optimize a financial-based cost function. However, unlike Bengio (1997), our framework is designed to learn portfolio weights directly instead of using a separate trading module. Second, we present a systematic approach for choosing the best subset of return predictive signals from a large set of factors using elastic net regularization. The remainder of this paper is organized as follows. In section 2, we formally define the architecture of our deep neural network. Section 3 is devoted to describing the data used in our experiment. Experiment design is presented in section 4. Section 5reports the results and discusses the most relevant findings. Finally, section 6 concludes.

2. Model Formulation

Let n denotes the number of stocks available for trading at the beginning of the month T (i.e. portfolio formation day). Each stock $x_i, 1 \leq i \leq n$ is represented as a vector of k real features or firm-specific characteristics that prior scholarly research has identified as being predictive of stock returns, i.e. $x_i \in \mathbb{R}^k$. Studies on return predictive signals are too numerous to list, but some of the seminal work is discussed in the introduction and appendix 1 of this paper.

Let y refers to a vector of n real labels y_i, $1 \leq i \leq n$. In our application, $y_i \in \mathbb{R}$ denotes the true/observed monthly return of stock i, $1 \leq i \leq n$ at month T.

We define a training example as$\varphi = (x, y) \in \mathcal{X}^n \times \mathbb{R}^n$ where x is a vector of n stocksx_i, $1 \leq i \leq n$, and $\mathcal{X}$ is the space of all stocks.

Our goal is to learn an asset allocation function$f: \mathcal{X}^n \to \mathbb{R}^n$, where$w_i = f(x)|_i$, $1 \leq i \leq n$ represents the weights of the i^{th} asset in our portfolio. That is, our function f jointly maps a group of n stocks to a vector of portfolio weights of the same size. In this paper, we do not allow short-selling thus, we have

$$\sum_{i=1}^{n} w_i = 1$$

$$w_i \geq 0, \forall i = 1..n$$

In this work, we parameterize the function f using a deep neural network. As discussed earlier, Feed-Forward neural networks have widely been applied in finance. We believe a deep neural network fits well in our application compared to other models for two reasons. First, neural networks scale well to high dimensional inputs (Ai et al., 2019). This is important because in our framework the function f takes $n$stocks as input and each stock is represented as a vector of a potentially large number of features. Second, deep neural networks are very flexible in learning a mapping from inputs to outputs. For instance, and as will be discussed later in this paper, the use of the softmax activation function in the output layer of our deep neural network allows us learning a mapping from firm-specific characteristics of n stocks to a vector $y \in \mathbb{R}^n$ representing the portfolio weights.

In this paper, we construct the input layer of our neural network by concatenating the vectors representing each stock. Specifically, let

$$h_0 = concat(x_1, x_2, \dots, x_i, \dots, x_{n-1}, x_n)$$

$$x_i \in \mathbb{R}^k$$

Given the above input layer, we build a multi-layer feed-forward network with two hidden layers as follows:

$$h_z = \sigma(w_z^T * h_{z-1} + b_z)$$

$$z = 1, 2$$

Where w_z and b_z denote the weight matrix and the bias vector in the z^{th} layer, σ is a non-linear activation function. In this paper, we use the rectified linear unit (ReLU) function: $\sigma(a) = \max(0, a)$

The output of our function is thus defined as:

$$f(x) = \mathcal{S}(w_o^T * h_2 + b_o)$$

Where w_o and b_o are the weight matrix and the bias vector in the output layer, $\mathcal{S}$ in the softmax activation function. The softmax function $\mathcal{S} : \mathbb{R}^n \to \mathbb{R}^n$ is defined by the formula:

$$\mathcal{S}(v)|_i = \frac{e^{v_i}}{\sum_{i=1}^{n} e^{v_i}}$$

$$i = 1, \dots, n$$

$$v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$$

The softmax function takes as input a vector v of n real numbers, the vector v might not sum to 1 and some of its components could be negative, or greater than 1, and normalizes it into a unit vector consisting of n real numbers proportional to the exponentials of the input numbers. That is, after applying the softmax function, each component will be in the interval $(0,1)$ and the components will sum up to 1. Thus, in our framework, the output of the softmax function can be interpreted as portfolio weights when short-selling is not allowed.

In order to learn the function f, we train our neural network by minimizing the empirical loss over the training data:

$$\mathcal{L}(f) = \frac{1}{|\psi|} \sum_{(x,y) \in \varphi} \ell(y, f(x))$$

Where ψ denotes the set of training examples and $\ell(.)$ is a local loss function defined as

$$\ell(y, f(x)) = -\sum_{i=1}^{n} w_i * y_i$$

Where $\sum_{i=1}^{n} w_i * y_i$ denotes the portfolio rate of return. Thus, minimizing $\ell(.)$is equivalent to maximizing the portfolio rate of return.

Figure 1 shows a simplified visual representation of our network when sitting $n = 4$ stocks and each stock is represented by a vector of $k = 5$ features.

In this work, we tried to keep the objective function and the design of the network architecture as simple as possible. We leave the exploration of more advanced representations (e.g. recurrent connections, convolutional layer, etc.) and cost functions (e.g. minimizing the risk as measured by the portfolio variance or maximizing the portfolio sharp ratio) as future work.

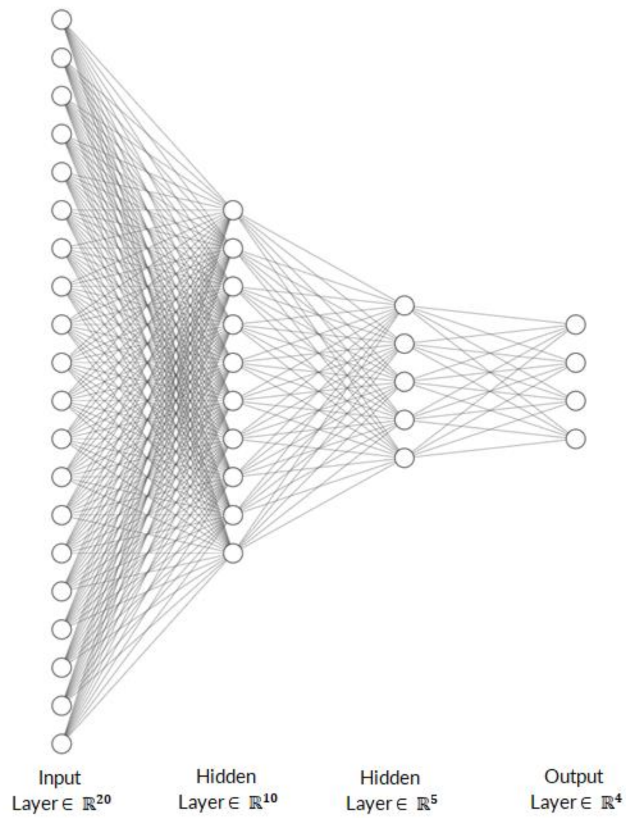

Figure 1. Architecture of our artificial neural network. For illustrative purposes, if the number of stocks $n = 4$ and each stock is represented by a feature vector of length 5, then the input layer will be a vector of length $20\ (i.e.\ 4 * 5)$ and the output layer will be a vector of length $n = 4$. In our application, we chose the first hidden layer to be half the length of the input layer, and the second hidden layer the half of the first hidden layer.

3. Data

The data in this paper are collected from two main sources. Moroccan stock data are from the Casablanca Stock Exchange (CSE) website and the risk-free asset return data are from the Bank-Al-Maghreb (the central bank of Morocco) website. The risk-free rate was measured as the monthly yield on the Moroccan 13-week Treasury bills. Market returns are calculated using the MASI index.

Our sample consists of all firms listed on the CSE during the period of October 2007 (i.e. the first portfolio formation month) through July 2019 with at least eight years of data prior to the portfolio formation month T. Three years of data is used as a training set and an additional six years of data is required in order to compute some features (e.g. five-year cumulative returns, while skipping six months between the formation and holding period, is needed in order to calculate the long-term reversal factor (Bondt & Thaler, 1985)). Thus, the data used in this paper covers the whole period from 1999 through 2019. Figure 2 shows the number of stocks satisfying this criterion for each formation month during our study period.

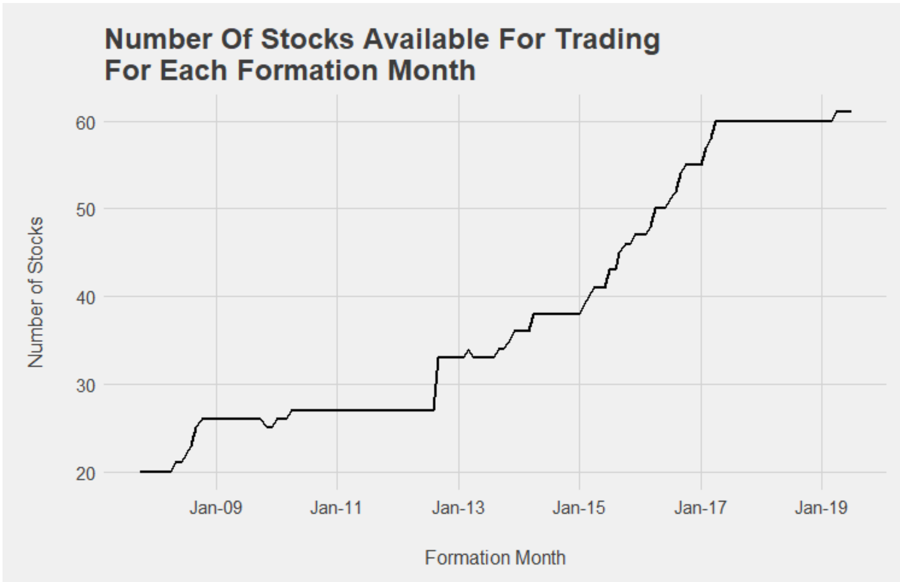

Figure 2. Number of stocks by portfolio formation month. At the beginning of each month T, starting from $T =$ October 2007, we consider all firms listed on CSE with at least eight years of historical data. Three years of data is used as a training set and an additional six years of data is required to compute some features.

4. Experiment Design

At the beginning of each month T (i.e. the portfolio formation date), starting from October 2007 through July 2019, we follow a three-step methodology. First, we use the previous seven years of data in order to compute the feature space and generate the training/validation sets. Second, we perform features selection using elastic-net regularization. Finally, we train the deep neural network described in the previous section of this paper, in order to learn portfolio weights and hold this portfolio until the end of the month T (i.e. monthly rebalancing). The remainder of this section follows the three-step methodology outlined above.

4.1. Features Generation

As discussed previously in this paper, studies in the asset pricing literature document hundreds of factors that appear to predict stock returns. Jacobs (2015) and Green, Hand, and Zhang (2013) provide an excellent review. However, due to data availability constraints, in this work, we focus on a subset of these factors that can be directly or indirectly derived from the firm's historical stock prices and trading volumes.

Let P_t^s denote the adjusted closing price of stock s at day t. Then, we define the simple daily return R_t^s for each stock s as

$$R_t^s = \frac{P_t^s}{P_{t-1}^s} - 1$$

Let $N(m)$ represent the set of trading days in month m. Following the literature, e.g. (Hou et al., 2006; Huddart, Lang, & Yetman, 2009; Lee & Swaminathan, 2000), we define the monthly returns R_m^s for each stock s as the compounded daily returns observed during the month m

$$R_m^s = \left(\prod_{t \in N(m)} R_t^s + 1 \right) - 1$$

In addition, given that the features are calculated on a daily basis, in this paper, we define a month as including 21 trading days.

In this paper, we consider about 36 return predictive signals derived from the 7 classes of factors discussed in section 1 of this paper (i.e. momentum, enhanced momentum, long-term reversal, short-term reversal, technical analysis, beta, and lottery-type). The complete list of these factors and their calculation details are provided in appendix 1 of this paper.

4.2. Features Selection

As discussed earlier in this paper, an increasingly large number of researchers document hundreds of factors that appear predictive of stock returns. However, studies in the asset pricing literature suggest that only a few of them should be sufficient to capture the dynamic of stock returns. In fact, most of these factors are likely to be proxies for the same sources of return variation (Hou, Mo, Xue, & Zhang, 2019). Another possibility is that some proposed factors are the outcome of data mining (Chordia, Goyal, & Saretto, 2018).

For this reason, among the 36 computed factors (see appendix 1), we select only a small subset of them as input to our model. More specifically, at the beginning of each month T we proceed as follow:

- First, for each stock available for trading in that month, we compute the whole set of features listed in appendix 1 of this paper;
- Second, for each stock, we perform elastic net regularization(Zou & Hastie, 2005) using the last three years of historical data (skipping the most recent month of data in order to avoid the look-ahead bias when constructing the response variable) as a train/validation set. For each trading day, the feature space (i.e. input) is represented as a vector of 36 factors and the response variable (i.e. output) is the monthly expected return;
- For each stock, the importance score of each factor is calculated as the scaled (between 0 and 100) absolute value of its corresponding coefficient in the elastic net regression.

- Factors are ranked using the average importance score across all stocks available for trading in that month. To construct input to our deep neural network model, each stock will be represented as a vector of the top five factors.

Figure 3 shows a ranking of all computed factors using the average importance score across all formation months and across all stocks. The code between parentheses refers to the ID of the corresponding factor in appendix 1.

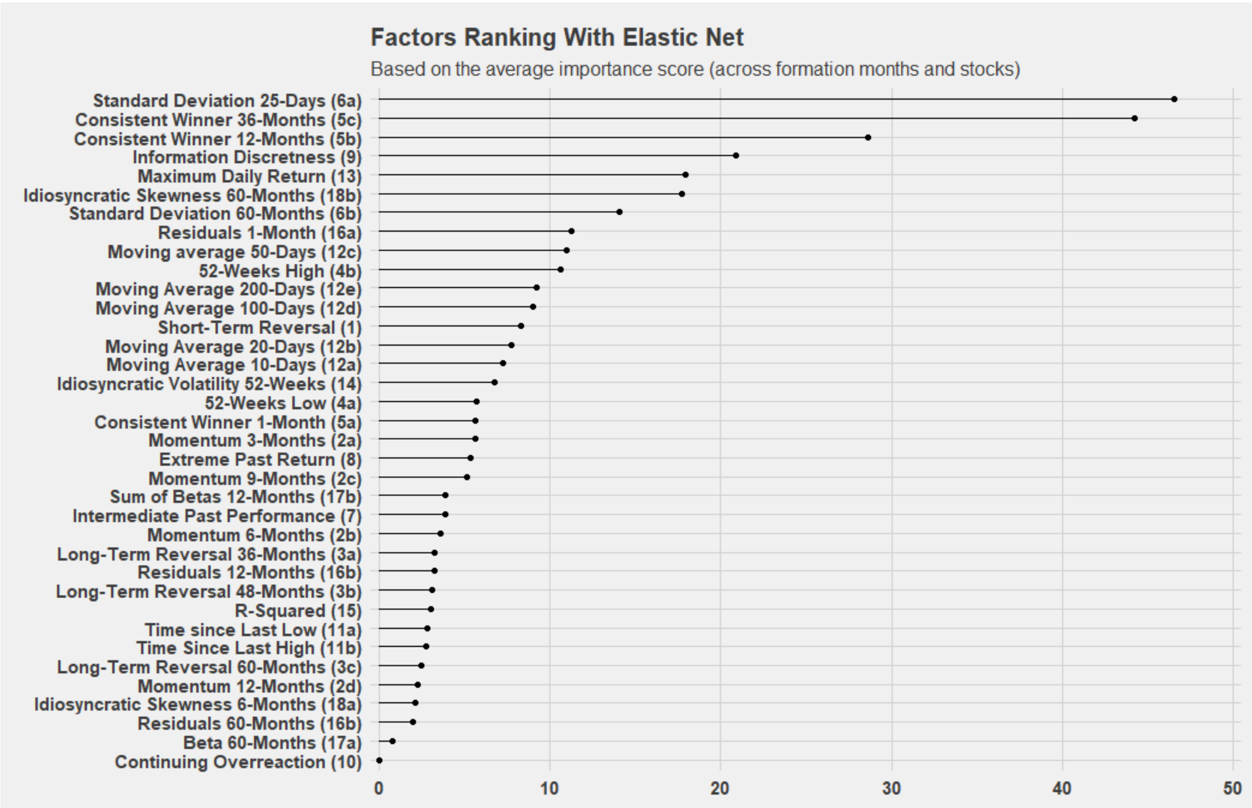

Figure 3. The importance score of each factor is calculated as the scaled (between 0 and 100) absolute value of it corresponding coefficient in the Elastic Net regression. The model is estimated for each formation month and for each stock separately. This figure reports the average importance score across formations months and across

4.3. Deep Neural Network Training and Benchmark Strategies

In order to take into account the non-stationarity of the financial time-series, and estimate performance over a variety of economic situations, multiple training experiments were performed on different training windows, at the beginning of each month T, starting from October 2007 through July 2019. Each time using the previous three years of computed features as a training set (again, as in elastic net regularization, we skip the most recent month of data in order to avoid the look-ahead bias when constructing the response variable).

The performance of our approach is compared with three benchmarks that are well documented in the literature:

- Short-term Reversal Strategy: Short-term reversal portfolios are formed based on 1-month lagged cumulative returns and held for 1 month. At the beginning of each month T, stocks are ranked in ascending order on the basis of previous month lagged cumulative returns. Based on these rankings, five quintile portfolios are formed that equally weight the stocks contained in the top quintile, the second quintile, and so on. In this paper, we report the performance of the 5th quintile portfolio.
- Momentum Strategy: Momentum portfolios are formed based on 12-months lagged cumulative returns and held for 3 months. At the beginning of each month T, stocks are ranked in ascending order on the basis of 12-months lagged cumulative returns from $T-13$ to $T-1$ (i.e. we skip a month between the portfolio formation period and the holding period). Based on these rankings, five quintile portfolios are formed that equally weight the stocks contained in the top quintile, the second quintile, and so on. In this paper, we report the performance of the top quintile portfolio.
- Long-term Reversal Strategy: Long-term reversal portfolios are formed based on 60-months lagged cumulative returns and held for 36 months. At the beginning of each month T, stocks are ranked in ascending order on the basis of 60-months lagged cumulative returns from $T-66$ to $T-6$ (i.e. we skip six months between the portfolio formation period and the holding period). Based on these rankings, five quintile portfolios are formed that equally weight the stocks contained in the top quintile, the second quintile, and so on. In this paper, we report the performance of the 5th quintile portfolio.

5. Results and Discussion

We report our findings in Table 1from both statistical and economic significance points of view. Panel A highlights some statistical properties of the monthly returns of the four portfolios. The results obtained for the three conventional strategies (short-term reversal, momentum, and long-term reversal) were in line with the expectations. They all generate a positive and statistically significant average monthly return. However, the momentum strategy is less risky (i.e. smallest standard deviation) than short-term reversal and long-term reversal. In contrast, our deep neural network approach generates the highest average monthly return (1.99% a month) but it is riskier than the conventional approaches (8% standard deviation versus 6% for long-term reversal and approximately 5% for both momentum and short-term reversal). Figure 4 highlights the uncertainty around the average monthly returns of the four investing strategies.

The high volatility characteristic of our approach is somewhat expected because our deep neural network was not explicitly trained to reduce the risk (as measured by the standard deviation of returns). Rather, it was specifically trained to maximize the portfolio rate of return (as specified in section 2 of this paper). In upcoming work, we plan to address this issue by incorporating the Sharpe ratio (reward-to-variability ratio) as an objective function. However, even though the monthly returns of our approach are more volatile than the conventional approaches, panel B of Table 1 indicates that it achieves the highest Sharpe ratio, which indicates that our method generates the best reward-to-risk ratio (21.7% per unit of risk versus 19.31% for the momentum, 15.51% for long-term reversal, and only 8.69% for short-term reversal).

Panel C of Table 1 reports the results of the Capital Asset Pricing Model (i.e. CAPM) regressions of the monthly returns of the four trading strategies

$$(r_{s,t} - r_{f,t}) = \alpha_s + \beta_s * (r_{m,t} - r_{f,t}) + \epsilon_{s,t}$$

Where r_s is the monthly return on strategy s, r_m is the monthly return on the market portfolio, and r_f is the monthly return on the risk-free asset. The alphas or risk-adjusted returns are in general smaller than unadjusted ones, ranging from 0.3% to 1.8% per month. The large and statistically significant risk-adjusted abnormal returns clearly demonstrate the profitability of our strategy compared to the conventional approaches. It generates an alpha that is about twice (0.018/0.007) as large as that generated by the momentum and long-term reversal strategies. Figure 5 illustrates the uncertainty around the monthly risk-adjusted returns of the four investing strategies.

Table 1. Performance of Deep Learning Portfolios

At the beginning of each month (i.e. the formation month), starting from October 1, 2007, until July 01, 2019, we use the previous historical data (up to approximately eight years of data) in order to compute all the features listed in Table 1, perform features selection using Elastic-Net, and train the artificial neural network (ANN) described in Section 2 of this paper. Our portfolio is rebalanced monthly according to the weights learned by the ANN. This table compares the performance of our approach with three well documented investing strategies in the literature: short-term reversal, momentum, and long-term reversal. Summary statistics of monthly returns of the four portfolios are presented in panel A of this table. Some financial metrics and the CAPM test results are presented in Panel B and C respectively; t-statistics are in parentheses. ** and * indicate significance at the 1% and 5% levels, respectively.

	Short-Term Reversal	Momentum Strategy	Long-Term Reversal	Deep Neural Network
Panel A: Summary Statistics				
Average Monthly Return	0.00676	0.01154	0.01223	0.01995
	(2.13)**	(3.74)**	(2.57)**	(2.94)**
Standard Deviation	0.04957	0.04710	0.06358	0.08087
Skewness	0.52405	0.93196	1.33165	1.06297
Kurtosis	1.09809	3.16957	3.77491	2.44280
Minimum	-0.12513	-0.12135	-0.14825	-0.14912
1st Quintile	-0.02392	-0.01823	-0.02832	-0.02412
Median	0.00151	0.00648	0.00558	0.00375
3rd Quantile	0.03531	0.03529	0.04072	0.05942
Maximum	0.20444	0.24780	0.31083	0.35364
Panel B: Financial Performance				
Average Excess Return	0.00431	0.00986	0.00610	0.01753
	(1.36)	(2.07)*	(2.22)*	(2.58)**
Sharpe Ratio	0.08690	0.19311	0.15506	0.21671
Annualized Return	8.42%	14.77%	15.70%	26.74%
Panel C: CAPM Test				
Alpha	0.00313	0.00736	0.00715	0.01845
	(1.26)	(3.10)**	(1.74)*	(2.74)**
Beta	0.71960	0.69895	0.73499	0.36744
	(12.45)**	(12.77)**	(8.03)**	(1.93)

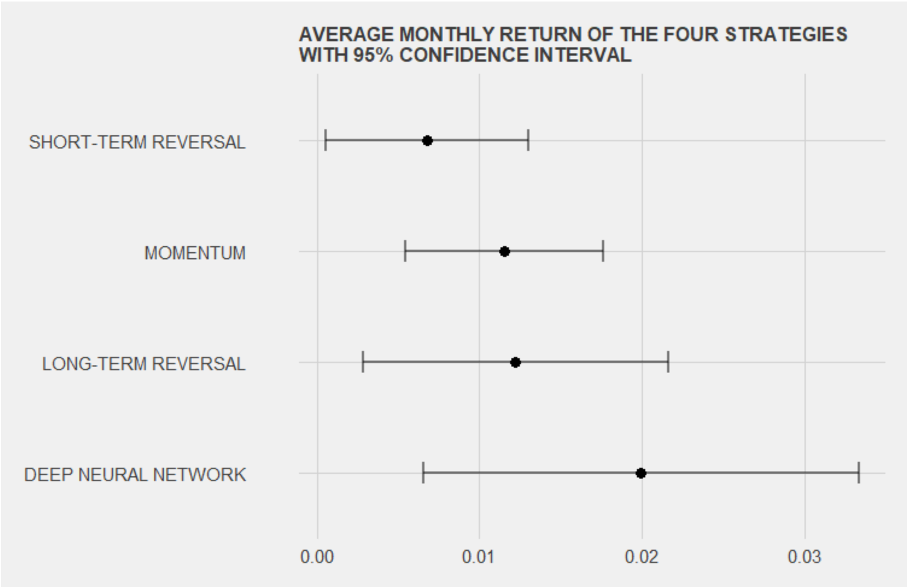

Figure 4. For each investment strategy, we report the 95% confidence interval (CI) of the corresponding mean monthly return.

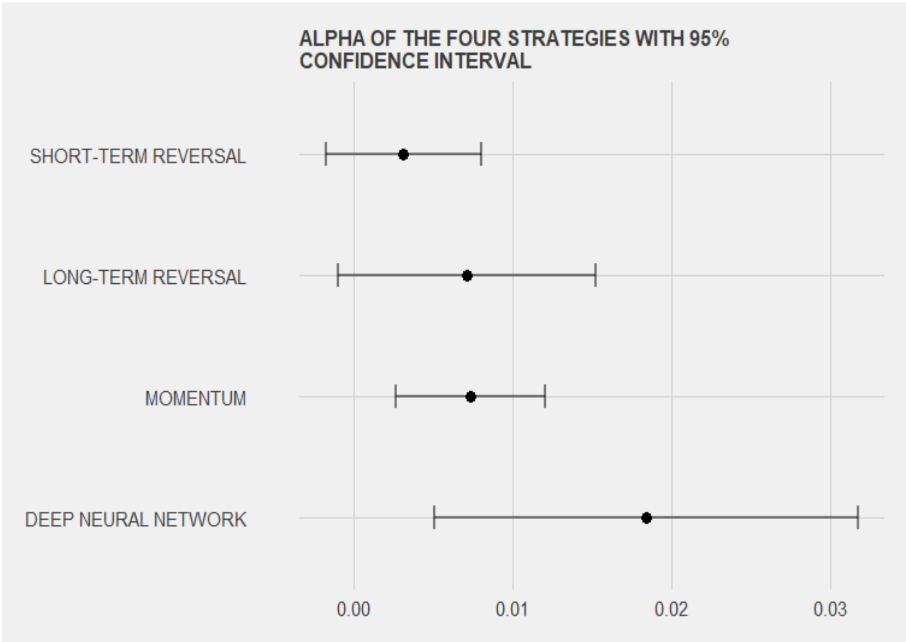

Figure 5. For each investment strategy, we report the 95% confidence interval of the parameter alpha of the regression of the monthly excess returns on the market factor:

$$(r_{s,t} - r_{f,t}) = \alpha_s + \beta_s * (r_{m,t} - r_{f,t}) + \epsilon_{s,t}.$$

Vol. 6, No.12; 2022

ISSN: 2456-7760

6. Conclusion

The results presented in this paper demonstrate an interesting application of machine learning in asset allocation. Our approach combines elastic net regularization for factors selection and novel deep neural network architecture designed specifically for learning portfolio weights. Compared with conventional approaches such as momentum, long-term reversal, and short-term reversal, our framework demonstrates a superior and statistically significant financial performance (in terms of the average monthly return, average excess return, annualized return, Sharpe ratio, and alpha).

In future work, we plan to extend our study in two ways. First, by including a larger set of factors or return predictive signals. In this paper, we only covered a small subset of these factors (36 factors in total) that could be directly or indirectly derived from the firm's historical stock prices and trading volumes. However, the literature suggests hundreds of other factors such as accruals, dividends, earnings surprises, fundamental analysis etc. Jacobs (2015) and Green, Hand, and Zhang (2013) provide an excellent review. Second, and as discussed in section 5 of this paper, the monthly returns generated by our approach are more volatile than those generated by conventional approaches such as momentum, long-term reversal, and short-term reversal strategies. We plan to address this issue by incorporating the Sharpe ratio (the reward-to-volatility ratio) as the training criterion (i.e. objective function) instead of the solo portfolio rate of return used in this version of our paper.

APPENDIX 1

List of Factors and Calculation Details

IDs	Signal	Selected Reference Papers	Description	Computational Details
1	Short-Term Return Reversal	Jegadeesh (1990); Lehmann (1990); Da, Liu, and Schaumbu rg (2013).	Short-term (one-to-four week) past returns tend to reverse in the following month.	We compute the cumulative return over the previous month.
2a, 2b, 2c, 2d	Momentum	Jegadeesh and Titman (1993).	Intermediate horizon past returns (three-to-twelve months) are positively related to future average returns.	Similar to Jegadeesh and Titman (1993), we compute four proxies to the momentum signal. The cumulative return over the previous three, six, nine, and twelve months. We also impose the standard 1-month lag between the formation and holding period in order to reduce the effect of the short-term reversal anomaly documented by Jegadeesh(1990).
3a, 3b, 3c	Long-term Return Reversal	Bondt and Thaler (1985); McLean (2010).	Long-term past returns (three-to-five years) are inversely related to future average returns and thus exhibit a long-term reversal phenomenon.	Similar to Bondt and Thaler (1985), we compute the cumulative return over the previous three, four, and five years. In addition, we fellow McLean (2010) and impose a 6-month lag between the formation and holding period in order to reduce the effect of the momentum phenomenon observed by Jegadeesh and Titman (1993).

Vol. 6, No.12; 2022

ISSN: 2456-7760

4a, 4b	Nearness to the 52-Week Low (High)	George and Hwang (2004); Huddart et al. (2009).	Nearness to the 52-week low (high) is a better predictor of future returns than are past returns.	The nearness to the 52-week low (high) is defined as the ratio between the last observed closing price and the minimum (maximum) price observed during the last twelve months. Similar to the literature, we include the standard 1-month lag between the formation and holding periods.
5a, 5b, 5c	Consistent Winner	Grinblatt and Moskowitz (2004).	Achieving a high past return with a series of steady positive returns appear to generate higher expected returns than a high past return achieved with just a few extraordinary months.	We compute the number of times formation period returns were positive divided by the number of trading days in that same period. Following Grinblatt and Moskowitz (2004), we compute the consistency over the previous twelve months (with a 1-month lag) and over the previous three years (with a 6-month lag). We also included a measure for short-term consistency (over the previous month).
6a, 6b	Volatility	Jiang, Lee, and Zhang (2005); Baker, Bradley, and Wurgler (2011).	Higher information uncertainty (measured by returns volatility) firms earn lower future returns than low volatility stocks.	We compute two proxies. The standard deviation of returns observed in the previous twenty-five trading days (Jiang et al., 2005) and the standard deviation of monthly returns observed during the last five years (Baker et al., 2011).

7	Intermediate Past Performance	Novy-Marx (2012).	Intermediate horizon past performances measured over the period from twelve to seven months prior, better predict average returns than does the recent past performance.	Following Novy-Marx (2012), we compute the intermediate past performance as the cumulative return over the period from twelve to seven months prior to the portfolio formation date.
8	Extreme Past Returns	Bandarchuk and Hilscher (2013).	Momentum profits documented by Jegadeesh and Titman (1993) are higher for stocks with extreme past returns.	Similar to Bandarchuk and Hilscher (2013), we compute the extreme past returns indicator as follow: $exp\left(\left\lvert r^{s}_{t-6\,;\,t-1} - \widetilde{r^{s}}\right\rvert\right) - 1$ Where $r^{s}_{t-6\,;\,t-1}$ denotes the logarithmic return over the previous six months and $\widetilde{r^{s}}$ denotes the median return over the same period.
9	Information Discreteness	Da, Gurun, and Warachka (2014).	Momentum profits documented by Jegadeesh and Titman (1993) are higher among firms with information arriving continuously in small amounts.	Defined as the sign (i.e. +1 or -1) of the cumulative return over the past twelve months (after skipping the most recent month) multiplied by the difference between the fraction of days with a negative return and the fraction of days with positive return (Da et al., 2014).

10	Continuing Overreaction	Byun, Lim, and Yun (2016).	A measure of continuing overreaction is a better predictor of future returns than past returns.	We closely follow Byun et al. (2016) and compute the continuing overreaction measure as the sum of the weighted signed volumes (increasing weights to more recent days) over the past twelve months (after skipping the most recent month) divided by the average trading volume over the same period.
11a, 11b	Time since Last Low/High	Huddart et al. (2009).	Risk-adjusted stock returns following the week and the month after a stock moves beyond its previous high or its previous low are positive. This effect is more pronounced the longer the time since the previous high or low was established.	Following Huddart et al. (2009), we compute the number of days since the last time the lowest/highest price was observed during the previous 52-week period. We normalize this number by the number of trading days during the same period.
12a, 12b, 12c, 12d, 12e	Nearness to the N-Day Moving Average	Han, Yang, and Zhou (2013).	A moving average timing strategy outperforms the buy-and-hold strategy. Specifically, stocks with a price above (below) the n-day moving average price outperform (underperform).	Measured as the ratio between the stock price at day $t-1$ and the average stock price in the previous n trading days. Following Han et al. (2013), we compute the nearness to the 10-day, 20-day, 50-day, 100-day, and 200-day moving averages.

Vol. 6, No.12; 2022
ISSN: 2456-7760

13	Maximum Daily Return	Bali, Cakici, and Whitelaw (2011).	A negative and significant relation between the maximum daily return over the previous month and the expected stock returns.	Following Bali et al. (2011), we compute the maximum daily return observed in the previous trading month: $\max_{1 \le i \le 21} \{R^s_{t-i}\}$
14	Idiosyncratic Volatility	Zhang (2006); Bandarchuk and Hilscher (2013).	Momentum profits documented by Jegadeesh and Titman (1993) are higher among firms with high idiosyncratic volatility.	We estimate the market model at a rolling basis, at the end of each trading day, using weekly returns (computed by compounding daily returns) over the previous 52 weeks: $R_{s,t} = a_s + b_s * R_{m,t} + \epsilon_{s,t}$ Following Bandarchuk and Hilscher (2013), idiosyncratic volatility is measured as the annualized standard deviation of the residuals returns.
15	R-Squared	Hou et al. (2006).	Momentum profits documented by Jegadeesh and Titman (1993) are higher among firms with low R^2.	We compute R^2 from the market model: $R_{s,t} = a_s + b_s * R_{m,t} + \epsilon_{s,t}$ Following Hou et al. (2006), this model is estimated at a rolling basis, at the end of each trading day, using the weekly returns observed over the past year.

Vol. 6, No.12; 2022

ISSN: 2456-7760

16a, 16b, 16c	Residuals	Da et al. (2013).	Stock returns unexplained by fundamentals (i.e. residuals) are more likely to reverse in the short-term. In this paper, we also compute residuals-based momentum and residuals-based long-term reversal.	Residuals are computed from the capital asset pricing model (CAPM): $(R_{s,t} - R_{f,t}) = \alpha_s + \beta_s * (R_{m,t} - R_{f,t}) + \epsilon_{s,t}$ Where $R_{f,t}$ denotes the risk-free rate at time t. Following Da et al. (2013), the CAPM is estimated at a rolling basis, at the end of each trading day, using the monthly returns observed over the past five years. In this paper, we use three variants of residuals. Cumulative residuals over the previous month, 1-year (skipping one month between formation and holding periods), and 5-year (skipping six months between formation and holding periods).
17a, 17b	Beta	Baker et al. (2011); Frazzini and Pedersen (2014); Hong and Sraer (2016).	Contrary to the CAPM prediction (in an efficient market, investors realize above-average returns only by taking above-average risks), low-beta stocks beta outperform high-beta stocks.	We compute two proxies. The beta from the standard CAPM model as in Baker et al. (2011) and Frazzini and Pedersen (2014): $(R_{s,t} - R_{f,t}) = \alpha_s + \beta_s * (R_{m,t} - R_{f,t}) + \epsilon_{s,t}$ The CAPM is estimated at a rolling basis, at the end of each trading day, using the monthly returns observed over the past five years. We also follow Hong and Sraer (2016) and compute the sum of betas from the model: $(R_{s,t} - R_{f,t}) = \alpha_s + \sum_{k=0}^{5} \beta_{s,k} * (R_{m,(t-k)} - R_{f,(t-k)}) + \epsilon_{s,t}$ Similar to Hong and Sraer (2016), this model is estimated at a rolling basis, at the end of each trading day, using the daily data over the last year.

Vol. 6, No.12; 2022

ISSN: 2456-7760

18a, 18b	Idiosyncratic Skewness	Kumar (2009); Boyer, Mitton, and Vorkink (2010)	Idiosyncratic skewness and expected returns are negatively correlated. Thus, stocks with high idiosyncratic skewness (lottery-type stocks) underperform stocks with low idiosyncratic skewness.	Computed as the skewness of residuals returns obtained by fitting the CAPM: $(R_{s,t} - R_{f,t}) = \alpha_s + \beta_s * (R_{m,t} - R_{f,t}) + \epsilon_{s,t}$ The CAPM is estimated at a rolling basis, at the end of each trading day, using the daily returns observed over the past six months similar to Kumar (2009). We also compute another proxy using the previous 5-year daily returns as in Boyer et al. (2010).

References

Abe, M., & Nakayama, H. (2018). Deep Learning for Forecasting Stock Returns in the Cross-Section. In D. Phung, V. S. Tseng, G. I. Webb, B. Ho, M. Ganji, & L. Rashidi (Eds.), *Advances in Knowledge Discovery and Data Mining* (pp. 273–284). Springer International Publishing.

Ai, Q., Wang, X., Bruch, S., Golbandi, N., Bendersky, M., & Najork, M. (2019). Learning Groupwise Multivariate Scoring Functions Using Deep Neural Networks. *Proceedings of the 2019 ACM SIGIR International Conference on Theory of Information Retrieval*, 85–92. https://doi.org/10.1145/3341981.3344218

Baker, M., Bradley, B., & Wurgler, J. (2011). Benchmarks as Limits to Arbitrage: Understanding the Low-Volatility Anomaly. *Financial Analysts Journal*, *67*(1), 40–54. https://doi.org/10.2469/faj.v67.n1.4

Bali, T. G., Cakici, N., & Whitelaw, R. F. (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics*, *99*(2), 427–446. https://doi.org/10.1016/j.jfineco.2010.08.014

Bandarchuk, P., & Hilscher, J. (2013). Sources of Momentum Profits: Evidence on the Irrelevance of Characteristics. *Review of Finance*, *17*(2), 809–845. https://doi.org/10.1093/rof/rfr036

Bengio, Y. (1997). Using a Financial Training Criterion Rather than a Prediction Criterion. *International Journal of Neural Systems*, *08*(04), 433–443. https://doi.org/10.1142/S0129065797000422

Bondt, W. F. M. D., & Thaler, R. (1985). Does the Stock Market Overreact? *The Journal of Finance*, *40*(3), 793–805. https://doi.org/10.2307/2327804

Boyer, B., Mitton, T., & Vorkink, K. (2010). Expected Idiosyncratic Skewness. *The Review of Financial Studies*, *23*(1), 169–202. Retrieved from JSTOR.

Byun, S. J., Lim, S. S., & Yun, S. H. (2016). Continuing Overreaction and Stock Return Predictability. *Journal of Financial and Quantitative Analysis*, *51*(6), 2015–2046. https://doi.org/10.1017/S0022109016000594

Cattaneo, M., Crump, R., Farrell, M., & Schaumburg, E. (2016). *Characteristic-Sorted Portfolios: Estimation and Inference* (Staff Report No. 788). Retrieved from Federal Reserve Bank of New York website: https://www.newyorkfed.org/research/staff_reports/sr788.html

Chordia, T., Goyal, A., & Saretto, A. (2018). *p-Hacking: Evidence from Two Million Trading Strategies* (No. 17–37). Retrieved from Swiss Finance Institute website: https://ideas.repec.org/p/chf/rpseri/rp1737.html

Da, Z., Gurun, U. G., & Warachka, M. (2014). Frog in the Pan: Continuous Information and Momentum. *Review of Financial Studies*, *27*(7), 2171–2218.

Da, Z., Liu, Q., & Schaumburg, E. (2013). A Closer Look at the Short-Term Return Reversal.

Management Science, *60*(3), 658–674. https://doi.org/10.1287/mnsc.2013.1766

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European Journal of Operational Research*, *270*(2), 654–669. https://doi.org/10.1016/j.ejor.2017.11.054

Frazzini, A., & Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics*, *111*(1), 1–25. https://doi.org/10.1016/j.jfineco.2013.10.005

George, T. J., & Hwang, C.-Y. (2004). The 52-Week High and Momentum Investing. *The Journal of Finance*, *59*(5), 2145–2176. https://doi.org/10.1111/j.1540-6261.2004.00695.x

Green, J., Hand, J. R. M., & Zhang, X. F. (2013). The supraview of return predictive signals. *Review of Accounting Studies*, *18*(3), 692–730. https://doi.org/10.1007/s11142-013-9231-1

Grinblatt, M., & Moskowitz, T. J. (2004). Predicting stock price movements from past returns: The role of consistency and tax-loss selling. *Journal of Financial Economics*, *71*(3), 541–579. https://doi.org/10.1016/S0304-405X(03)00176-4

Han, Y., Yang, K., & Zhou, G. (2013). A New Anomaly: The Cross-Sectional Profitability of Technical Analysis. *Journal of Financial and Quantitative Analysis*, *48*(5), 1433–1461. https://doi.org/10.1017/S0022109013000586

Hong, H., & Sraer, D. A. (2016). Speculative Betas. *The Journal of Finance*, *71*(5), 2095–2144. https://doi.org/10.1111/jofi.12431

Hou, K., Mo, H., Xue, C., & Zhang, L. (2019). Which Factors? *Review of Finance*, *23*(1), 1–35. https://doi.org/10.1093/rof/rfy032

Hou, K., Xiong, W., & Peng, L. (2006). *R2 and Price Inefficiency* (SSRN Scholarly Paper No. ID 954559). Retrieved from Social Science Research Network website: https://papers.ssrn.com/abstract=954559

Huddart, S., Lang, M., & Yetman, M. H. (2009). Volume and Price Patterns around a Stock's 52-Week Highs and Lows: Theory and Evidence. *Management Science*, *55*(1), 16–31. Retrieved from JSTOR.

Jacobs, H. (2015). What explains the dynamics of 100 anomalies? *Journal of Banking & Finance*, *57*, 65–85. https://doi.org/10.1016/j.jbankfin.2015.03.006

Jegadeesh, N. (1990). Evidence of Predictable Behavior of Security Returns. *The Journal of Finance*, *45*(3), 881–898. https://doi.org/10.2307/2328797

Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance*, *48*(1), 65–91. https://doi.org/10.2307/2328882

Jiang, G., Lee, C. M. C., & Zhang, Y. (2005). Information Uncertainty and Expected Returns. *Review of Accounting Studies*, *10*(2), 185–221. https://doi.org/10.1007/s11142-005-1528-2

Krauss, C., Do, X. A., & Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *European Journal of Operational Research*, *259*(2), 689–702. https://doi.org/10.1016/j.ejor.2016.10.031

Kumar, A. (2009). Who Gambles in the Stock Market? *The Journal of Finance*, *64*(4), 1889–1933. https://doi.org/10.1111/j.1540-6261.2009.01483.x

Lee, C. M. C., & Swaminathan, B. (2000). Price Momentum and Trading Volume. *The Journal of Finance*, *55*(5), 2017–2069. https://doi.org/10.1111/0022-1082.00280

Lehmann, B. N. (1990). Fads, Martingales, and Market Efficiency. *The Quarterly Journal of Economics*, *105*(1), 1–28. https://doi.org/10.2307/2937816

McLean, R. D. (2010). Idiosyncratic Risk, Long-Term Reversal, and Momentum. *The Journal of Financial and Quantitative Analysis*, *45*(4), 883–906. Retrieved from JSTOR.

Novy-Marx, R. (2012). Is momentum really momentum? *Journal of Financial Economics*, *103*(3), 429–453.

Zhang, X. F. (2006). Information Uncertainty and Stock Returns. *The Journal of Finance*, *61*(1), 105–137. https://doi.org/10.1111/j.1540-6261.2006.00831.x

Zou, H., & Hastie, T. (2005). Regularization and Variable Selection via the Elastic Net. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, *67*(2), 301–320. Retrieved from JSTOR.