Charting By Machines*

Scott Murray† Houping Xiao‡ Yusen Xia§

*This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1548562.

†Assistant Professor of Finance, J. Mack Robinson College of Business, Georgia State University, 35 Broad St NW, Atlanta, GA 30303, Phone: 404-413-7351, Email: smurray19@gsu.edu.

‡Assistant Professor, Institute for Insight, J. Mack Robinson College of Business, Georgia State University, 35 Broad St NW, Atlanta, GA 30303, Phone: 404-413-7705, Email: hxiao@gsu.edu.

§Bradford and Patricia Ferrer Professorship in Analytics and Director, Institute for Insight, J. Mack Robinson College of Business, Georgia State University, 35 Broad St NW, Atlanta, GA 30303, Phone: 404-413-7704, Email: ysxia@gsu.edu.

Charting By Machines

Abstract

We test the efficient markets hypothesis by using machine learning to forecast future stock returns from historical price plots. These forecasts strongly predict the cross section of future stock returns. The predictive power holds in most subperiods, is strong among the largest 500 stocks, and is distinct from momentum and reversal. The forecasting relation is highly non-linear and remarkably stable through time. Our research design ensures that our findings are not a result of data snooping. We conclude that the efficient markets hypothesis does not hold and that investment strategies based on technical analysis and charting may have merit.

1 Introduction

The weak form of the efficient markets hypothesis (EMH hereafter) stipulates that the construction of a profitable portfolio based only on information discernable from plots depicting the historical performance of stocks (price plots hereafter) should not be possible. As such, technical analysis, or charting, should be a fruitless investment technique. Academic research on technical analysis has broadly supported this prediction, and as a result, recent work on this topic has been sparse. Despite the broad dismissal of technical analysis in the academic literature, it has remained widely used by investment managers. The continued widespread use of technical analysis suggests that its merit may not be fully discovered in the academic literature, and that further investigation is warranted.

In this paper, we test the weak form of the EMH by examining whether forecasts produced by machine learning (ML hereafter) predict the cross section of future stock returns. The forecasts are based only on the cumulative monthly returns of the stock over the past year, which are easily discernable from price plots. We find strong evidence that the ML-based forecasts have economically important and highly statistically significant predictive power. This predictive power is strong in most subperiods of our focal 196307-201912 test period, including the most recent subperiod covering 201501-201912, persists for at least four months after the forecast is generated, and remains strong among the largest 500 stocks. The relation between past and future returns is highly stable through time and highly non-linear. Finally, the predictive power of our ML-based is not explained by the well-known momentum (Jegadeesh and Titman (1993)) and reversal (Jegadeesh (1990)) effects.

Execution of the ML process requires us to make several implementation decisions. To ensure that our results are not subject to concerns related to data mining (Harvey, Liu, and Zhu (2016)) or out-of-sample forecasting power (McLean and Pontiff (2016), Green, Hand, and Zhang (2017)), we use data from 192701-196306, which we refer to as the "optimization period", to determine the optimal implementation choices. Our analyses lead us to use a convolutional neural network with long-short term memory as the ML architecture, mean-squared error as the loss function, a weighting scheme that assigns the same total weight to observations from each month and equal weight to each stock within a month, and a normalized measure of future return as the dependent variable.[1] Our use of data from the

[1] To avoid the curse of dimensionality, we rely on the machine learning literature to motivate our choices

only the optimization period to make these implementation decisions ensures that our main test results truly reflect out-of-sample predictive power.

We use the optimized ML process to generate forecasts of future stock returns during the 196307-201912 period, which we refer to as the "test period". Our main analyses examine the performance of portfolios formed by sorting stocks on the ML-based forecasts, which we denote $MLER$. We find that $MLER$ is a strong predictor of the cross section of future stock returns. The average excess returns of value-weighted decile portfolios constructed using breakpoints calculated from only NYSE-listed stocks increase from -0.13% per month for the decile one portfolio to 0.93% per month for the decile 10 portfolio. The 1.06% per month generated by the portfolio that is long the decile 10 portfolio and short an equal dollar amount of the decile one portfolio is not only economically large, but highly statistically significant, with a t-statistic of 5.38. We find no evidence that variation in average returns of the $MLER$-sorted portfolios is attributable to risk. Alphas of the decile portfolios with respect to several established factor models exhibit similar patterns to the average excess returns. Additional risk metrics such as volatility, skewness, value at risk, and expected shortfall do not support a risk-based explanation for the patterns in average returns.

Having shown that $MLER$ strongly predicts the cross section of future stock returns, we next attempt to characterize this predictive power. We find that the predictive power is strong during most subperiods of our 196307-201912 test period. The portfolio that is long (short) stocks with high (low) values of $MLER$ generates a large and highly significant average excess return of more than 1% per month during most subperiods, the exception is the 200501-201412 subperiod, during which the average excess return is close to zero due to a small number of very large monthly losses during 2009. During the most recent 201501-201912 subperiod, the portfolio once again generates more than 1% per month. The forecasting power continues to hold for several months after the calculation of forecast. The average monthly excess return of the portfolio in the fourth month after $MLER$ is calculated is 0.39% with a corresponding t-statistic of 2.01. In subsequent months, the average excess return remains economically large but it not significant at conventional levels. The predictive power of $MLER$ remains remarkably strong when the sample is restricted to only large stocks. Most notably, when using only the top 500 stocks by market capitalization to form the portfolios, the average excess return of the long-short portfolio is 0.69% (t-statistic = 4.11).

The objective of the remainder of our tests is to understand the source of predictive

of more technical parameters, such as the number of layers in the neural network, etc.

power of the ML-based forecast. The main challenge in this is that the ML process does not enable us to observe the actual function used to generate the forecast. Our tests related to this objective therefore examine specific aspects of the predictive power. The first such tests examine whether the function generated by the ML process is stable through time. We find that this is indeed the case. Forecasts based on fits of recent data are no better than forecasts based on data from the distant past. For example, when forecasting returns during the 199501-200412 period, forecasts based on fits of the data from 196307-199412 perform no better than forecasts based on fits of the data from 192701-196306. Furthermore, a forecast based only on a fit of the data from 192701-196306 performs nearly as well as our focal expanding-window measure. We then investigate the degree to which the forecasting function produced by the ML process is linear, and find that it is not. Cross-sectional OLS regressions of $MLER$ on the historical cumulative returns used to generate the forecast have an average adjusted R-squared of only 37%. We also find that the non-linearities in the forecasting function are important for prediction. Regressions of future excess returns on $MLER$ and the historical cumulative returns indicate that the predictive power of $MLER$ remains strong after accounting for linear relations between future and past returns.

Finally, we examine whether the predictive power of $MLER$ can be explained by the momentum (Jegadeesh and Titman (1993)) and/or reversal (Jegadeesh (1990), Lehmann (1990)) effects, both of which, like the effect we document, are relations between past and future returns. We find that while the ML-based forecast does include components related to momentum and reversal, a large portion of the forecasts' predictive power is unrelated to these effects. Specifically, portfolios that are constructed to be neutral to momentum and reversal while having long (short) positions in stocks with high (low) $MLER$ generate large positive average excess returns, and Fama and MacBeth (1973) regressions show that the predictive power of $MLER$ persists when controlling for momentum and reversal.

Our work contributes to three broad strands of research. First, we add the literature examining the whether past returns contain information useful for predicting future returns, which is tantamount to the literature on the weak form of the efficient markets hypothesis. The most prominent findings in this literature are the aforementioned momentum and reversal effect. A subset of this literature explicitly investigates the validity of technical analysis and charting. Several papers examine the ability of technical signals to predict the performance of broad market indices or diversified portfolios. Brock, Lakonishok, and LeBaron (1992) find that simple technical signals predict the future returns of the Dow Jones Index. However, subsequent work attributes this finding to data snooping (Sullivan, Timmermann, and White (1999), Ready (2002), and Bajgrowicz and Scaillet (2012)) and nonsynchronous trading (Bessembinder and Chan (1998)). Allen and Karjalainen (1999) use genetic algorithms (a form of machine learning) to identify technical rules for trading the S&P 500 index and find that they do not work out of sample. More recently, Zhu and Zhou (2009) find that technical analysis can be useful for making asset allocation decisions, Moskowitz, Ooi, and Pedersen (2012) find time-series momentum in a large number of asset classes, Neely, Rapach, Tu, and Zhou (2014) show that technical indicators are useful for predicting the market risk premium, and Han, Yang, and Zhou (2013) find evidence that moving average strategies work well for timing investment in volatility-sorted portfolios.[2] Barra et al. (2020) show that an S&P 500 index market timing strategy based on forecasts generated by convolutional neural networks outperforms a buy-and-hold strategy.

Despite evidence that technical analysis is widely-used in practice (Menkhoff (2010)), research on the cross-sectional predictive power of technical signals is relatively sparse. Our paper is most similar to Lo, Mamaysky, and Wang (2000), who use smoothing estimators to extract nonlinear relations between historical price patterns and future stock returns. This finding has been the subject of much scrutiny, most notably by Jegadeesh (2000), who argues that Lo et al. (2000)'s findings are not robust to variation in the bandwidth parameter used for smoothing, a subjective empirical decision, and that the profitability of trading strategies based on the patterns detected by Lo et al. (2000) are close to zero. Our paper is similar in spirit to Lo et al. (2000), except we use a more modern technology (ML) than they do (smoothing estimators) to detect the relations between past price patterns and future returns. Importantly, our paper overcomes the robustness critique of Jegadeesh (2000) by using a systematic procedure to optimize our ML process using only data from the period prior to our main test period. Our use of portfolio analyses as our main empirical methodology addresses Jegadeesh (2000)'s profitability critique. The main contribution of our paper, therefore, is to demonstrate the merit of technical analysis in a manner that overcomes the critiques of previous work in this area.

Second, we add to the recent and growing literature that uses ML to understand the cross section of expected stock returns.[3] Moritz and Zimmermann (2016), Messmer (2017),

[2]Sweeney (1986), Levich and Thomas III (1993), Neely, Weller, and Dittmar (1997), Chang and Osler (1999), Gehrig and Menkhoff (2006) find evidence that investment strategies based on technical analysis are profitable in foreign exchange markets, a finding that Osler (2003) attributes to the clustering of stop-loss and take-profit orders. LeBaron (1999) argues that the profitability of such strategies is due to central bank intervention in currency markets, but Neely (2002) concludes the opposite.

[3]Earlier work includes Allen and Karjalainen (1999), who use genetic programming to search for profitable technical patterns in the S&P 500 index, and find none. In related work, Feng, He, and Polson (2018) and

Messmer and Audrino (2017), and Freyberger, Neuhierl, and Weber (2020) use ML to identify relations between stock-level characteristics and expected stock returns. Bryzgalova, Pelger, and Zhu (2020), Chen, Pelger, and Zhu (2020), Gu, Kelly, and Xiu (2019), Kelly, Pruitt, and Su (2019), Feng, Polson, and Xu (2020), Gu, Kelly, and Xiu (2020), and Kozak, Nagel, and Santosh (2020) use ML to extract latent factors, factor exposures, and risk premia from characteristics. A consistent theme in this work is the use of ML to synthesize the information in a broad set of firm-level variables that previous work has already found to be related to expected stock returns. Our work differs in that we use ML to identify previously undiscovered relations between historical price patterns and future stock returns.

Finally, we contribute to the recent literature that broadly questions the conclusions of much of the empircal asset pricing research. Harvey et al. (2016) suggest that mining of the CRSP data has resulted in a large number of falsely significant relations between predictive variables and the cross section of future stock returns, a concern previously raised by Lo and MacKinlay (1990) and Fama (1991). McLean and Pontiff (2016) find that the predictive power established by many academic studies declines in the period subsequent to that examined by the original papers documenting the pricing effects.[4] Green, Hand, and Zhang (2017) find that only 12 of 94 characteristics examined provide independent predictive power, and only two do so during the post-2003 period. We overcome these concerns by using data from 192701-196306 to optimize our ML process, and then running our asset pricing tests *once* on data from 196307-201912. Doing so ensures that our results truly reflect the out-of-sample predictive power of the ML-based forecasts. Our research design provides a blueprint for future ML-based research proposing new determinants of the cross section of expected stock returns to overcome the concerns about empirical asset pricing research. It is worth noting that even after taking these precautions to ensure that our results measure out-of-sample performance, the t-statistics in our focal tests far surpass the benchmarks proposed by Harvey et al. (2016).

The remainder of this paper is organized as follows. Section 2 describes the construction of our sample. Section 3 describes how we optimize our ML process and the calculation of the ML-based return forecasts. Section 4 presents our main evidence that the ML-based forecasts successfully predict the cross section of future stock returns. Section 5 characterizes the nature of the predictive power of the ML-based forecasts. Section 6 concludes.

Rossi (2018) use ML to predict the returns of the S&P 500 index, and Bianchi, Buchner, and Tamoni (2020) use ML to predict bond returns.

[4]Schwert (2003) finds similar results for the size and value effects.

2 Sample and Variables

In this section we describe our sample and the variables used in our tests.

2.1 Sample

The stock data used in our study come from CRSP. Our sample contains stock, month observations for months t from 192701-201912 (inclusive). Specifically, in each month t, the sample contains all stocks that, on the last trading day of month $t-1$, are common shares of US-based firms listed on the NYSE, AMEX, or NASDAQ. In addition, to ensure that the construction of a one-year historical price plot for each stock in the sample is been feasible, we require that each included stock have a non-missing return for each of months $t-12$ through $t-1$, inclusive. Finally, because our focal analyses weight stocks by market capitalization (value-weighted hereafter), we require that the market capitalization of each stock in the sample as of the end of month $t-1$, which we define as the number of shares outstanding times the price of each share, can be calculated. Our focal tests examine the ability of ML-based forecasts calculated as of the end of month $t-1$ to predict the cross section of month t stock returns.

To ensure that the results of our focal tests reflect out-of-sample predictive power, we use sample months t from 192701-196306 (inclusive) to determine optimal implementation of our ML process, and therefore refer to this period as the "optimization period". The first month of the optimization period is 192701 because return data in CRSP begin in 192601, making 192701 the first month for which a full year of prior return data are available. Our focal asset pricing tests cover months from 196307 through 201912, which we refer to as the "test period". We choose 196307 as the beginning of the test period (and thus 196306 as the end of the optimization period), to conform with the start date of the sample in Fama and French (1992, 1993) and several subsequent asset pricing studies.

2.2 Variables

For each stock i in each month t, we calculate several variables. The focal dependent variable in most of our analyses is the excess stock return in month t, calculated as the delisting-adjusted month t stock return minus the month t return of the risk-free security.[5] The

[5] The month t delisting-adjusted stock return is calculated following Shumway (1997). If the stock is not delisted in month t, then the delisting-adjusted return is simply the stock return. If the stock is delisted in month t, then if a delisting return is provided by CRSP, we take the return of the stock to

independent variables, which are used to predict the cross-section of month t stock returns, are calculated using only data available as of the end of month $t-1$. The focal independent variable is the ML-based return forecast, $MLER$. Section 3 describes in detail how we calculate $MLER$. For now, we simply note that the forecast of the return of stock i in month t is generated by applying the function generated by the ML process to the cumulative returns of stock i over months $t-12$ through $t-1$, which we denote $CR_1, \ldots, CR_{12}$. Specifically, we define CR_k to be the cumulative stock return over the k-month period covering months $t-12$ through $t-12+k-1$. This notation is motivated from the point of view of an investor looking at a one-year price plot created at the end of month $t-12$, when decisions on which stock positions for month t would be made. CR_1 is the return of the stock during the first month that appears on the plot, and more generally CR_k is the cumulative return of the stock during the first k months that appear on the plot.

3 ML Process and Forecasts

In this section we describe how we use data from the optimization period to determine the optimal implementation of the ML process, and then discuss how we use the ML process to generate forecasts of future stock returns.

3.1 Optimizing the ML Process

Our objective in using ML is to generate a function f that, given values of independent variables that can be discerned from a historical price plot of a given stock, will produce a forecast of the future return of that stock. Very generally, the ML process takes a panel data set that includes one or more independent variables, a dependent variable, and a weight variable, which combined we refer to as the training data, and attempts to "learn" the function f that describes the relation between the independent variables and the expectation of the dependent variable. The ML architecture attempts to learn f by minimizing the value of a weighted loss function $\mathcal{L}$. Once the function f has been learned, it can be applied to data

be the delisting return. If no delisting return is available, then we determine the stock's return based on the delisting code in CRSP. If the delisting code is 500 (reason unavailable), 520 (went to OTC), 551-573 or 580 (various reasons), 574 (bankruptcy), or 584 (does not meet exchange financial guidelines), we take the stock's return during the delisting month to be -30%. If the delisting code has a value other than the previously-mentioned values and there is no delisting return, we take the stock's return during the delisting month to be -100%. Monthly risk-free security returns are from Ken French's website: `https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

not included in the training data to generate forecasts.

There are many ways the ML process can be implemented, and there is little evidence in the computer science or asset pricing literature to guide us in determining the optimal implementation. We therefore use data from the optimization period, which covers sample months t from 192701-196306, to determine the optimal implemenation of our ML process. Specifically, we aim to find the combination of ML architecture, loss function, loss function weight variable, and dependent variable that generates the best forecasts of the cross section of future stock returns. As independent variables, we use the monthly cumulative stock returns over the 12 months prior to the month whose return is to be forecasted, $CR_1, \ldots, CR_{12}$. We choose $CR_1, \ldots, CR_{12}$ as independent variables because they are easily discerned by someone observing a historical price plot of the stock.

To ensure that our ML process uses an architecture that is well-designed to capture non-linearities between the dependent and independent variables, following LeCun, Bengio, and Hinton (2015) and Goodfellow, Bengio, and Courville (2016), we consider four different ML architectures: feed-forward neural network (FNN), convolutional neural network (CNN), long-short term memory (LSTM), and convolutional neural network with long-short term memory (CNNLSTM). All four ML architectures are widely used representation-based deep learning methods designed to learn complex non-linear relations in the data. Here, we give a very brief comparison of the different architectures. Section I of the Internet Appendix provides a more technical discussion.

FNN is the most general of the four architectures. However, FNN does not explicity consider the grid topology or variable dependency that frequently characterize sequential and time-series data. The CNN and LSTM archictures are explicitly designed to overcome this shortfall of FNN. CNN is a specialized type of neural network designed to process data with a grid-like topology, such as our sequential cumulative returns data, which can be interpreted as a one-dimensional grid. Previous work (Hoseinzade and Haratizadeh, 2019) has shown CNN to be highly successful at time-series prediction. LSTM is a form of recurrent neural network (RNN). RNNs are explicitly designed to process one-dimensional sequential data. However, conventional RNNs have been shown to be ineffective at discerning relations at lags greater than five lags (Goodfellow et al., 2016), a drawback that LSTM overcomes. The CNNLSTM architecture combines the features of both the CNN and LSTM architectures.

ML is an optimization procedure, and the objective of this optimization procedure is to minimize a loss function $\mathcal{L}$ that aggregates forecast errors. The two loss functions we consider are mean squared error (MSE, $\mathcal{L} = \sum w_j \epsilon_j^2$) and mean absolute error (MAE, $\mathcal{L} = \sum w_j |\epsilon_j|$), where w_j is the weight assigned to observation j in the training data, ϵ_j is the error, defined as the difference between the forecast value and the actual observed value of the dependent variable, and the summation is taken over all observations in the training data. Qualitatively, the main difference between MSE and MAE is that MSE is more sensitive to outliers than MAE.

The ML process uses panel data to determine the function f that best describes the relation between the independent and dependent variables. However, our objective is to generate forecasts of the *cross section* of future stock returns. As shown in Figure 1, the number of stocks in our sample varies substantially across months. Thus, giving the same weight to each observation would cause months with more stocks to carry more total weight in the loss function. To address this concern, in addition to the equal-weighting (EW) approach, we consider two alternative weighting methodologies. The first alternative is to equally-weight each month, and within each month, equally-weight each stock (EWPM for equal weight per month). The second alternative is to equally-weight each month, but consistent with our intent to use value-weighted portfolios as our primary test methodology, weight each stock within a given month according to its market capitalization at the end of the previous month (EWPMVW for equal weight per month value weighted).

The use of panel data may also cause the function f generated by the ML structure to be determined, at least in part, by time-series patterns between historical and future returns of the aggregate market. To illustrate the concern, consider the two price plots shown in Figure 2, which depict the cumulative return of the market portfolio during the 12-month periods prior to 193304 (Panel A) and 193109 (Panel B). These are the one-year periods leading up to the highest (38.95% in April 1933) and lowest (−29.10% in September 1931) monthly market portfolio returns during our sample period, respectively. Given that there is strong covaration in returns among individual stocks, it is likely that many stocks exhibited price patterns similar to those exhibited in the plots during the corresponding periods. The high (low) market portfolio return in the subsequent month suggests that most stocks had a high (low) return during this month.[6] The ML algorithm is therefore likely to associate price patterns similar to that in Panel A (Panel B) with a high (low) future return.

To ensure that the ML process does not generate the function f based on this sort of aggregate time-series pattern, in addition to the excess stock return, r, we consider three transformations of r as potential dependent variables. The first is the standardized excess

[6] The equal-weighted average return of the stocks in our sample in April 1933 (September 1931) is 52.48% (−31.47%), and 598 out of 631 (618 out of 627) such stocks had a positive (negative) return.

return, r_{Std}, which is calculated by taking the excess return, subtracting the mean excess return, and then dividing this difference by the standard deviation of the excess stock returns, with the mean and standard deviation being calculated from the excess returns of all stocks in the given month. Using r_{Std} guarantees that the mean and standard deviation of all observations in any given month are zero and one, respectively, thereby making it unlikely that the function generated by the ML process is based largely on aggregate time-series patterns.

A potential concern with both r and r_{Std}, however, is that the cross-sectional distribution of stock returns is known to be highly leptukurtic (see e.g. Bali, Engle, and Murray (2016, Table 7.2)), making it plausible that the function f generated by the ML process is largely a manifestation of extreme observations. Furthermore, our aim is to identify stocks that are likely to perform better (or worse) than other stocks, but not necessarily forecast the magnitude of the difference in returns. We therefore consider two additional transformations of the excess stock return. The first is the normalized excess return, which is constructed to ensure that values of the dependent variable in any given month have a standard normal distribution. Specifically, we define $r_{Norm,i,t} = \Theta^{-1}[rank_{i,t}/(N_t + 1)]$ where $\Theta^{-1}[\cdot]$ is the inverse standard normal cumulative distribution function, N_t is the number of stocks in our sample in month t, and $rank_{i,t}$ is the rank of stock i's month t return among the month t returns of all stocks in our sample, with the stock that has the lowest return having a $rank_{i,t} = 1$ and the stock with the highest return having $rank_{i,t} = N_t$. The final candidate for the dependent variable is simply the percentile of the given stock's returns among all stocks' returns in the given month, calculated as $r_{Pctl} = rank_{i,t}/(N_t + 1)$. Values of r_{Pctl} in any given month have a uniform distribution on the interval from zero to one.

In sum, we consider 96 different implementations of our ML process, found by taking all combinations of four architectures (FNN, CNN, LSTM, CNNLSTM), two loss functions (MSE, MAE), three loss function weighting schemes (EW, EWPM, EWPMVW), and four dependent variables (r, r_{Std}, r_{Norm}, r_{Pctl}).

To assess the effectiveness of each implementation, we first apply each implementation to a subset of the data in the 192701-196306 optimization period. Specifically, we run each implementation 30 times on data from sample months t corresponding to even months from even years and odd months from odd years (training months hereafter) during the optimization period. Each implementation is executed 30 times because the ML process is random. Thus, if we run the exact same implementation on the exact same data twice, the two resulting functions will *not* be the same. Each of these 30 iterations, therefore, produces a different forecast function f. We then apply each of these 30 forecast functions f to each non-training month observation in the optimization period. Finally, for each such observation, we take the average of those 30 forecasts to be the forecast for the given observation based on the given ML implementation. In the end, for each non-training month observation in the optimization period, we have 96 different forecasts, one corresponding to each different implementation.

The metric we use to evaluate the different implementations is the time-series average of the monthly cross-sectional Spearman rank correlations between the forecast and the actual excess return, calculated from only non-training month observations. We focus on Spearman rank correlations, instead of Pearson product-moment correlations, because our main tests are portfolio analyses that rely solely on the ordering of stocks with respect to the forecasts. Table 1 presents the average Spearman rank correlations for each implementation. The highest average correlation is that generated by the CNNLSTM architecture using the MSE loss function with EWPM weights, and r_{Norm} as the dependent variable. Based on these results, we therefore choose to use the implementation with the highest average correlation (CNNLSTM, MSE, EWPM, and r_{Norm}) to generate our focal ML-based forecasts, $MLER$. Notably, small deviations from this implementation, such as using the MAE loss function, EW weights, or either r_{Std} or r_{Pctl} as the dependent variable, also perform well. This suggests that variation in performance across the different implementations is not spurious and that our optimization procedure is likely to have identified a near-optimal implementation.

Before generating the forecasts to be used in our focal asset pricing tests, we examine the amount of randomness that remains in our ultimate forecast calculated by taking the average forecasted value based on 30 executions on the ML process. This is important to ensure that our work is replicable by future researchers. If the number 30 is too low and a large amount of randomness remains in the ultimate forecast, then it is possible that subsequent attempts to replicate our analyses will produce different results. We test this by examining the correlation between ultimate forecasts generated by taking the average of N_F individual forecasts, for values of N_F between one and 100. Specifically, we begin by running the ML process 500 times on the training month observations in the optimization period. We then apply each of the 500 resulting functions to calculate forecasts of r_{Std} for each of the observations in non-training months. Then, for each value of N_F between one and 100, we repeat the following four steps 50 times: 1) we randomly select, without replacement, two distinct sets of N_F forecasts from the 500 available forecasts; 2) for each non-training month observation. we calculate the averages of each of these sets of N_F forecasts, and denote these averages ER_1 and ER_2; 3) we calculate the cross-sectional correlations between ER_1 and ER_2 in each non-training month; 4) we calculate the time-series average of these monthly cross-sectional correlations. After running this procedure, for each value of N_F, we have 50 average correlations. In essence, for each value of N_F, we have drawn a random sample of 50 correlations between the forecasts that would be produced by two different researchers applying our ML process to the exact same data.

Figure 3 plots the average, minimum, and maximum correlations for each value of N_F. When $N_F = 1$, the correlations range from 0.614 to 0.747, with a mean of 0.689. These values illustrate that the randomness in the fit can result in substantially different forecasts when comparing any two individual fits. When $N_F = 10$, the minimum, maximum, and mean correlations are 0.945, 0.959, and 0.952, respectively. For $N_F = 30$, these values are 0.980, 0.986, and 0.983, respectively. The benefit of increasing N_F above 30 is minimal. We therefore proceed with $N_F = 30$, which we view as a very conservative choice.

3.2 ML-Based Forecasts

We generate our focal ML-based forecasts, $MLER$, by applying the ML process to expanding-window subsets of our sample, and using the functions generated by these expanding window fits to produce forecasts for subsequent periods. In general, we define $MLER_{i,t}^{t_1,t_2}$ to be the forecast of the normalized excess return of stock i in month t that results from using our ML process on data from the period between months t_1 and t_2. Specifically, for any t_1, $t_2 > t_1$, i, and t, we calculate $MLER_{i,t}^{t_1,t_2}$ as follows. First, we apply the ML process 30 times to observations from sample months between t_1 and t_2, inclusive. The result is 30 different forecasting functions. We then apply each of these 30 functions to the cumulative returns of stock i in months $t-12$ through $t-1$. The result is 30 different forecasts of the normalized excess return of stock i in month t. We then take $MLER_{i,t}^{t_1,t_2}$ to be the average of these 30 forecasts. For example, to calculate $ER_{X,200809}^{192701,196306}$, we first run the ML process 30 times using all observations in our sample from months t between 192701 and 196306, inclusive. We then apply each of these forecast functions to the monthly cumulative returns of stock X that would be observed in stock X's price plot covering the period from 200709 through 200808. The result of applying these functions to these cumulative returns is 30 forecasts of the normalized excess return of stock X in 200809. We then take $MLER_{X,200809}^{192701,196306}$ to be the average of these 30 forecast values.

The expanding windows to which we apply the ML process cover sample months t between 192701 and each of 196306, 197412, 198412, 199412, 200412, and 201412. As a result, we calculate $MLER_{i,t}^{192701,196306}$, $MLER_{i,t}^{192701,197412}$, $MLER_{i,t}^{192701,198412}$, $MLER_{i,t}^{192701,199412}$, $MLER_{i,t}^{192701,200412}$, and $MLER_{i,t}^{192701,201412}$ for each stock i and month t observation in our sample. The forecast normalized excess return of stock i in month t that we use in our focal empirical tests, $MLER_{i,t}$, is the forecast calculated from the function generated by the most recent past execution of the ML process. We therefore have:

$$MLER_{i,t} = \begin{cases} MLER_{i,t}^{192701,196306}, & \text{if } 196307 \leq t \leq 197412; \\ MLER_{i,t}^{192701,197412}, & \text{if } 197501 \leq t \leq 198412; \\ MLER_{i,t}^{192701,198412}, & \text{if } 198501 \leq t \leq 199412; \\ MLER_{i,t}^{192701,199412}, & \text{if } 199501 \leq t \leq 200412; \\ MLER_{i,t}^{192701,200412}, & \text{if } 200501 \leq t \leq 201412; \\ MLER_{i,t}^{192701,201412}, & \text{if } 201501 \leq t \leq 201912. \end{cases}$$

3.3 Summary Statistics

Table 2 presents the time-series averages of monthly cross-sectional summary statistics of $MLER$ for both the entire 196307-201912 test period and for different subperiods of the test period. In interpreting the forecasts, recall that the forecasts are for the standardized future excess return, not the excess return itself. In the average month, $MLER$ has a mean of -0.02, a median very close to zero, and a cross-sectional standard deviation of 0.11. The results indicate that extreme negative values of $MLER$ occur with higher frequency than extreme positive values, since the minimum, first percentile, and fifth percentile values are further below the mean than the 95th percentile, 99th percentile, and maximum values, respectively, are above it. The prevalence of large negative values of $MLER$ compared to large positive values should have no impact on most of our asset pricing tests, since we focus on portfolio analyses, which rely on the ordering (but not the magnitude or distribution) of $MLER$ across stocks. In the average test period month, our sample contains 4,198 stocks. The results for the subperiods indicate that the cross-sectional distribution of $MLER$ is highly consistent through time, since the salient characterstics of the cross-sectional distribution of $MLER$ do not change much across the different subperiods. The exception to this is the number of stocks in the sample, which ranges from an average of 2,245 stocks per month for the 196307-197412 subperiod to 5,792 stocks per month for the 199501-200412 subperiod.

4 Predictive Power of ML-Based Forecast

We turn now to our main objective, which is to investigate the relation between the ML-based return forecasts and the cross section of future stock returns. The EMH predicts that $MLER$ should have no ability to predict cross-sectional variation in future stock returns.

We test this hypothesis by examining the performance of portfolios formed by sorting stocks based on $MLER$. Specifically, each month t, we sort all stocks in our sample into decile portfolios based on an ascending ordering of $MLER$, which is calculated from data available as of the end of month $t-1$. The breakpoints used to determine which stocks fall into which portfolios are the deciles of $MLER$ calculated from the subset of stocks that are listed on the NYSE as of the end of month $t-1$. The month t excess return of each decile portfolio is then calculated as the weighted average excess return of all stocks in the portfolio, with the weight being proportional to $MktCap$ measured at the end of month $t-1$ (value-weighted hereafter).[7] We also calculate the excess return of a zero-cost portfolio that is long the decile 10 portfolio and short the decile one portfolio (10−1 portfolio hereafter). Our portfolio construction methodology follows Hou, Xue, and Zhang (2020) and is intended to minimize the impact of small-capitalization stocks on our analysis.

Table 3 Panel A presents the time-series averages of the monthly portfolio excess returns for each portfolio during the 196307-201912 test period. The average portfolio excess returns increase nearly monotonically from −0.13% per month for decile portfolio one to 0.93% per month for decile portfolio 10. The average excess return of the $MLER$ $10-1$ portfolio of 1.06% per month is highly statistically significant, with a Newey and West (1987) t-statistic of 5.38. The patterns in average excess returns are strong evidence contrary to the baseline prediction of the EMH that technical analysis should be fruitless.

A more refined version of the EMH allows for profitable technical strategies if the associated average returns are compensation for risk. We test whether the dispersion in the average excess returns of the $MLER$-sorted portfolios are compensation for risk in two ways. First, we use factor analysis to estimate the average risk-adjusted excess return (alpha) of each portfolio. Specifically, the portfolio alpha is calculated as the intercept coefficient from a time-series regression of excess portfolio returns on excess factor returns. A non-zero (zero) alpha indicates that exposure to the systematic risk factors captured by the factor model does not explain (explains) the average excess return of the portfolio. We conduct the factor

[7] In Section II and Table A1 (Section III and Table A2) of the Internet Appendix, we demonstrate that our results hold when using equal-weighted portfolios (portfolios constructed using breakpoints calculated from all stocks).

analyses using six previously-established factors models: a one-factor market model (CAPM model), the three-factor model of Fama and French (1993, FF model), the four-factor model of Carhart (1997, FFC model), the FFC model augmented with the aggregate liquidity factor of Pastor and Stambaugh (2003, FFCLIQ model), the five-factor model of Fama and French (2015, FF5 model), and the four-factor model of Hou et al. (2015, Q model).[8] Table 3 shows that the alphas of the portfolios relative to each factor model exhibit similar patterns to the average excess returns. Regardless of which factor model is used, the alpha of the $MLER$ $10-1$ portfolio is positive, economically large, and highly statistically significant. Furthermore, for each model, the alpha of the decile 10 (decile one) portfolio is the highest (lowest) of all portfolios.

Second, we examine whether there are patterns in the total risk of the decile portfolios that may suggest a risk-based explanation for the strong dispersion in average excess portfolio returns. Specifically, Table 3 reports the standard deviation, skewness, one-percent and five-percent values at risk, and one-percent and five-percent expected shortfalls for each of the decile portfolios. If the dispersion in average excess returns of the decile portfolios is due to risk, we should observe a positive, negative, negative, and negative relation between the average excess returns and the standard deviation, skewness, values at risk, and expected shortfalls, respectively. The results provide no evidence of a risk-based explanation. The relations between average excess return and each of standard deviation, value at risk, and expected shortfall are in the opposite direction to what would be predicted by a risk-based explanation, and the results indicate very little difference in the skewness of the decile 10 and decile one portfolios. In sum, the results of the factor and total risk analyses provide no evidence of a risk-based explanation for differntial performance of the $MLER$-sorted portfolios, and thus contradict the refined prediction of the EMH.

5 Characterization of ML-Based Forecasts

Having demonstrated that the ML-based forecasts have strong ability to predict the cross section of future stock returns, we proceed now to further characterize this predictive power.

[8] Monthly excess returns for factors in the CAPM, FF, FFC, and FF5 models are from Ken French's website. Monthly excess returns for the Pastor and Stambaugh (2003) aggregate liquidity factor are from Lubos Pastor's website: `https://faculty.chicagobooth.edu/lubos-pastor`. Monthly excess returns for the factors in the Q model are from Chen Xue's website: `http://global-q.org/factors.html`.

5.1 Predictive Power in Subperiods

We begin by examining whether the ability of the ML-based forecast to predict the cross section of future stock returns is specific to subsets of our test period. The cumulative performance of the $MLER$ $10-1$ portfolio is plotted in Figure 4. The plot shows that the predictive power of $MLER$ is highly stable through time. With the exception of a few large losses in 2009, the cumulative excess returns of the $MLER$ $10-1$ portfolio has a strong positive trajectory. To more rigourously examine the stability of the predictive power of $MLER$ through time, Table 4 reports the average excess returns of each of the $MLER$ decile portfolios during several subperiods. As expected based on Figure 4, with the exception of the January 2005-December 2014 period, the $MLER$ $10-1$ portfolio generates a large positive average excess return of greater than 1% per month during each subperiod. In each of the subperiods prior to 2005, the average excess return of this portfolio is highly significant. During the most recent period, from January 2015-December 2019, this portfolio generates an average excess return of 1.06% per month (t-statistic = 1.61). The decreased statistical significance of the average excess return during this period is likely a manifestation of the fact that this period covers only five years, whereas each of the other periods examined covers at least 10 years. We statistically test whether the performance of the $MLER$ $10-1$ portfolio is stable through time by regressing the average excess returns on a constant and indicators for each subperiod except for July 1963-December 1974. The coefficients on the indicators from this regression (untabulated) are all statistically insignificant, with the exception of the indicator for January 2005-December 2014, which is negative and marginally significant.

Examination of the performance of the individual decile portfolios during each of the subperiods shows that, for each subperiod prior to 2005, the decile 10 and decile one portfolios generate the highest and lowest, respectively, average excess returns of any of the decile portfolios. For the January 2015-December 2019 subperiod, while the average excess return of the decile one portfolio is the lowest of all portfolios, the decile 10 portfolio generates the second highest average excess return, with the highest average excess return coming from the decile nine portfolio. The results demonstrate that with the exception of the subperiod affected by losses in 2009, the predictive power of $MLER$ is highly stable through time.

5.2 Predictive Power Beyond One Month

Our next tests investigate whether predictive power of the ML-based forecast persists beyond one month. We do so by examining the month $t+k$, $k>0$ performance of portfolios formed by sorting on $MLER$. Specifically, $MLER$ is calculated at the end of month $t-1$ using only data available as of that point in time. We then wait k months before constructing the portfolios, and examine the one-month return of the portfolios during month $t+k$. Aside from the timing, all other aspects of the portfolio construction are the same as in previous analyses. Table 5 presents the average excess returns of these portfolios for $k \in \{1, 2, \dots, 6\}$. For $k \in \{1, 2, 3\}$, the average excess return of the $MLER$ $10-1$ portfolio is positive and highly significant, and the first (10th) decile portfolio generates the lowest (highest) average excess return. For $k \in \{4, 5, 6\}$, the average excess returns of the $MLER$ $10-1$ portfolio remain positive and of substantial economic magnitude, but they are statistically insignficant at the 5% level. The results therefore suggest that the predictive power persists for at least four months (when $k = 3$ the portfolios are held four months after calculating $MLER$) after the calculation of the ML-based forecast.

5.3 Predictive Power for Large Stocks

The portfolios we have examined to this point are constructed by sorting all stocks into decile portfolios using breakpoints calculated from only NYSE stocks. Previous work (Hou et al. (2020)) has shown that this approach minimizes the influence of small stocks on the results of the analyses. Nonetheless, to ensure that our results persist when small stocks are excluded from the portfolios, we repeat the portfolio analysis using subsets of our sample that exclude small stocks both as the set of stocks held in the portfolios and as the set of stocks used to calculate the breakpoints. We define the $Size > P_{20}^{\text{NYSE}}$ (NYSE/$Size > P_{20}^{\text{NYSE}}$) and $Size > P_{50}^{\text{NYSE}}$ (NYSE/$Size > P_{50}^{\text{NYSE}}$) subsets to be the sets of (NYSE-listed) stocks with market capitalizations greater than the 20th and 50th percentile values, respectively, among NYSE-listed stocks as of the end of month $t-1$, and the Top 500 subset to include only the top 500 stocks by market capitalization. We continue to use value-weighted portfolios for all of these tests.

The results of these analyses are shown in Table 6. When we use the $Size > P_{20}^{\text{NYSE}}$ subset both to calculate the breakpoints and as the set of stocks held in the portfolios, the $MLER$ $10-1$ portfolio generates an average excess return of 1.05% per month (t-statistic $= 5.63$). This result is extremely similar to that of our main test whose results are shown in Table 3. Indeed, the correlation between the excess returns of these two portfolios is 0.96 (untabulated). These results strongly suggest that, consistent with Hou et al. (2020), the influence of small stocks on our focal tests is minimal. As we impose more stringent restrictions of the size of the stocks used in the analyses, the average excess returns of the

$MLER$ $10-1$ portfolio and associated t-statistics become slightly smaller. However, even when we use only the largest 500 stocks both to calculate the breakpoints and as the set of stocks held in the portfolios, the average excess return of the $MLER$ $10-1$ of 0.69% per month (t-statistic $= 4.11$) remains economically large and highly statistically significant. Furthermore, regardless of the set of stocks used to calculate the breakpoints or held in the portfolios, the decile one (10) portfolio generates the lowest (highest) average excess return of all portfolios. These tests clearly show that the predictive power of the ML-based forecast is strong among large stocks.

5.4 Stability of Forecasting Relation

We proceed now examine the stability of the relation between price patterns and forecasts of future stock returns (forecasting relation hereafter). Specifically, we investigate whether the price patterns associated with high (low) future stock returns in the early part of our test period continue to be associated with high (low) stock returns for the duration of our test period. If this relation is highly stable, it would indicate that the predictive patterns could be learned over a prolonged period of time, thereby further suggesting the viability of technical analysis as an investment methodology.

Our first tests of the stability of the forecasting relation examine whether the predictive power of the ML-based forecasts decreases as we increase the amount of time between the period covered by the data used in the ML process (fit period hereafter) and the period whose returns are examined. Specifically, we define $MLER_{i,t}^{192701,196306}$, $MLER_{i,t}^{196307,197412}$, $MLER_{i,t}^{197501,198412}$, $MLER_{i,t}^{198501,199412}$, $MLER_{i,t}^{199501,200412}$, and $MLER_{i,t}^{201501,201412}$ to be the forecasts for stock i in month t based on fit periods indicated by the months between and inclusive of those in the superscripts. We then create decile portfolios by sorting on each of these forecasts using the same portfolio construction procedure described in Section 4. Finally, we examine the performance of these portfolios during different subperiods. If the forecasting relation is highly stable (unstable), then we expect the predictive power to remain constant (diminish or disappear) as we increase the amount of time between the fit periodand the period whose returns are examined.

Table 7 presents the average excess retuns of the $10-1$ portfolios formed by sorting on each of these variables for each several subperiods. Note that while analyses that use a forecast based on a fit period subsequent to the period whose returns are examined are not reflective of obtainable trading profits, such analyses are valid tests of the stability of the forecasting relation. Overall, the results reveal no strong patterns relating the strength of return predictability to the amount of time between the subperiods whose returns are examined and the fit period. For example, the $MLER^{192701,196306}$ $10-1$ portfolio performs better during the 199501-200412 subperiod than during any other subperiod. This strong performance is not because this period is particularly good for portfolios formed by all of the forecast variables. In fact, the $MLER^{192701,196306}$ $10-1$ portfolio is the only portfolio that has its largest average excess return during 199501-200412. As another example, the best-performing portfolio during the 201501-201912 period is that formed by sorting on $MLER^{198501,199412}$. More broadly, for only two of the six subperiods examined is the best-performing portfolio that formed by sorting on the forecast from an adjacent fit period. Finally, it is noteable that the $MLER^{192701,196306}$ $10-1$ portfolio earns an average excess return of 0.85% per month (t-statistic = 2.09) during the 201501-201912 subperiod. Thus, forecasts based on a fit period ending in 196306 contain strong predictive power more than 50 years later.

As a second test of the stability of the forecasting relation, we examine the performance of portfolios formed by sorting on rolling-window fit periods ($MLER^{Roll}$) and a static forecast function based on the 192701-196306 fit period ($MLER^{192701-196306}$). $MLER^{Roll}$ for stock i in month t is defined as $MLER^{192701-196306}$ for months t between 196307 and 197412, $MLER^{196307-197412}$ for months t between 197501 and 198412, $MLER^{197501-198412}$ for months t between 198501 and 199412, $MLER^{198501-199412}$ for months t between 199501 and 200412, $MLER^{199501-200412}$ for months t between 200501 and 201412, and $MLER^{200501-201412}$ for months t between 201501 and 201912. If the forecasting relation is stable through time, then economic reasoning would suggest that the predictive power of all three measures ($MLER$, $MLER^{Roll}$, and $MLER^{192701,196306}$) is similar. However, if the ML process is more effective at detecting true patterns when applied to larger data sets, then we might expect a stable forecasting relation to result in the predictive power of $MLER$ exceeding that of $MLER^{Roll}$ and $MLER^{192701,196306}$, and the difference in predictive power would be indicative of the benefits of the use of additional data. If the forecasting relation is time-varying, we expect the predictive power of the variables based on fit periods that include the distant past to have relatively low predictive power. In this case, the predictive power of $MLER^{Roll}$ should be stronger than that of $MLER$, which in turn should be stronger than that $MLER^{192701,196306}$.

The performance of the decile portfolios formed by sorting on $MLER$, $MLER^{Roll}$, and $MLER^{192701,196306}$ is described in Table 8. The average excess returns of the $MLER^{Roll}$ $10-1$ of 0.98% per month and $MLER^{192701,196306}$ $10-1$ portfolio of 0.96% per month are extremely similar, which strongly suggests a stable forecasting relation. Furthermore, the

$MLER$ $10-1$ portfolio's average excess return of 1.06% per month (repeated from Table 3) is similar to, albeit slightly greater than, that of portfolios formed by sorting on $MLER^{Roll}$ and $MLER^{192701,196306}$. This suggests not only that the forecasting relation is highly stable, but that the benefits of the use of additional data by the ML process are moderate at best.

Taken together, the results in Tables 7 and 8 are consistent with a highly stable forecasting function, which in turn suggests that technical analysis may be a more viable investment strategy than understood by previous research.

5.5 Non-Linearity of Forecasting Function

The results in the previous subsection demonstrate that the function relating historical cumulative returns, given by $CR_1, \ldots, CR_{12}$, and the cross section of future stock returns, is highly stable through time, but gives no indication of its functional form. The main benefit of using ML to generate forecasts is that ML is capable of discerning highly-complex and non-linear relations. If the forecasting function generated by the ML process turns out to be linear in $CR_1, \ldots, CR_{12}$, it would suggest that the use of ML in this context is unnecessary and that more traditional techniques, such as linear regression, should suffice. Our next tests, therefore, examine the degree to which the forecasting function is linear in $CR_1, \ldots,$ CR_{12} and the degree to which any linearity in this relation accounts for the predictive power of $MLER$.

We begin with a Fama and MacBeth (1973, FM hereafter) regression analysis of the relation between $MLER$ and $CR_1, \ldots, CR_{12}$. Specifically, each month t, we run a cross-sectional OLS regression of $MLER$ on $CR_1, \ldots, CR_{12}$. Panel A of Table 9 presents the time-series averages of the monthly cross-sectional regression coefficients and adjusted R^2 values. The average adjusted R^2 of these regressions is only 37.46%, indicating that only slightly more than a third of the total variation in $MLER$ can be explained by linear a linear combination of $CR_1, \ldots, CR_{12}$, and that the forecast function generated by the ML process is highly non-linear.

While the results in Panel A demonstrate that the forecast function is highly non-linear, it is possible that the linear components drive the majority of the predictive power of $MLER$. To test the importance of the non-linearities in the forecasting functionin predicting the cross section of future stock returns, we once again conduct FM regressions, this time using the future excess stock return as the dependent variable and combinations of $MLER$ and $CR_1, \ldots, CR_{12}$ as independent variables. The results in Panel B show that when $MLER$ is the only independent variable in the regression, the average coefficient on $MLER$ of

6.80 is highly statistically significant (t-statistic = 10.05). When we add CR_1, ..., CR_{12} as independent variables to the regressions, the average coefficient on $MLER$ of 6.07 and associated t-statistic of 9.91 are very similar to those from the specification with $MLER$ as the only independent variable, indicating that no linear combination of CR_1, ..., CR_{12} explains the ability of $MLER$ to predict the cross section of future stock returns.

Since the OLS regressions whose results are shown in Panel B give equal weight to each observation, to ensure that the conclusions drawn from these results are not driven by the large number of relatively small stocks that constitute the sample of NYSE, AMEX, and NASDAQ stocks, we repeat these analyses using weighted least squared (WLS hereafter) regressions, with market capitalization as the weight variable. The results of these tests, shown in Panel C, lead to similar conclusions. The average coefficient on $MLER$ from the analysis that includes CR_1, ..., CR_{12} is very similar, both in magnitude and statistical significance, to the corresponding value from the analysis that uses only $MLER$ as an independent variable.

In sum, the results in Table 9 clearly demonstrate that the forecasting function generated by the ML process is highly non-linear in the historical cumulative returns, and that the non-linear aspects of the forecasts play a very important role in the ability of the ML-based forecast to predict the cross section of future stock returns.

5.6 Momentum and Reversal

The final set of analyses intended to characterize the ML-based forecasts examine the interactions between these forecasts and well-documented momentum and reversal effects. Our measure of momentum, Mom, is the cumulative stock return during the 11-month period covering months $t-12$ though $t-2$, inclusive (skipping month $t-1$). We measure reversal with Rev, defined as the stock return in month $t-1$. While the asset pricing literature has documented hundreds of variables that predict the cross section of future stock returns, these two effects are of particular relevance in our context because both Mom and Rev can be perfectly discerned from historical price plots.[9] In fact, both Mom and Rev are very simple functions of the cumulative returns $CR_1, \ldots, CR_{12}$ that are used to generate $MLER$. Specifically, $Mom = CR_{11}$ and $Rev = 100[(CR_{12}/100 + 1)/(CR_{11}/100 + 1) - 1]$.[10] It is therefore highly plausible that the predictive power of $MLER$ is, at least in part, driven by

[9] See Hou et al. (2015), Harvey et al. (2016), McLean and Pontiff (2016), and Linnainmaa and Roberts (2018) for extensive lists of variables that have been documented to predict the cross section of future stock returns.

[10] The multiplication and divisions by 100 are because Rev, CR_{11}, and CR_{12} are recorded in percent.

the predictive power of Mom and Rev.

5.6.1 Relations between ML-Based Forecast, Momentum, and Reversal

We begin by illustrating the momentum and reversal effects in our sample. Table 10 presents the average monthly excess returns of decile portfolios formed by sorting on each of Mom and Rev. With the exception of the sort variable, the portfolio construction methodology is identical to that used to construct the $MLER$ decile portfolios examined in Table 3. The average monthly excess return of the Mom $10-1$ portfolio is 1.26% (t-statistic = 4.70) and that of the Rev $10-1$ portfolio is -0.41% (t-statistic = 2.31). The results therefore show that both the momentum and reversal effects are strong in our sample.

We next examine the strength of the relations between the ML-based forecasts, momentum, and reversal by running FM regression analyses with $MLER$ as the dependent variable and combinations of Mom and Rev as the independent variables. Table 11 shows that when Mom (Rev) is the only independent variable, the average R^2 from the cross-sectional regressions is 9.87% (27.14%). When both Mom and Rev are included as independent variables, the average R^2 is 36.91%. The results indicate that, as expected, the ML-based forecasts have substantial components related to both the momentum and reversal effects. However, the forecasts also have a very substantial component that is unrelated to these effects.

5.6.2 Predictive Power Controlling for Momentum and Reversal

The main focus of the remainder of this section is to examine whether the ability of the ML-based forecast to predict the cross section of future stock returns is, at least in part, distinct from the moment and reversal effects. Our first tests with this objective are multivariate portfolio analyses that examine the performance of portfolios of stocks with similar levels of the control variable(s), Mom and/or Rev, but different levels of $MLER$. We construct both bivariate portfolios, which control for Mom or Rev one at a time, and trivariate portfolios that simultaneously control for Mom and Rev. If the predictive power of $MLER$ is unexplained (explained) by the control variable, we expect to find a strong pattern (no pattern) in average excess returns across portfolios containing stocks with different levels of $MLER$ but similar levels of the control variables.

To construct the bivariate portfolios, each month t we sort the stocks in our sample into five groups based on an ascending ordering of the control variable. We then sort all stocks within each of these control variable-based groups into five portfolios based on an ascending ordering of $MLER$. We take the month t excess return for each of the resulting

25 portfolios to be the value-weighted average month t excess return of all stocks in the portfolio. To create a single portfolio for each quintile of $MLER$, we define the month t excess return of the bivariate $MLER$ quintile k portfolio to be the equal-weighted average of the month t excess returns of the five portfolios (one for each quintile of the control variable) corresponding to the kth $MLER$ quintile. Finally, we define month t bivariate $MLER\,5-1$ portfolio excess return to be the month t excess return of the bivariate $MLER$ quintile five portfolio minus that of the bivariate $MLER$ quintile one portfolio.

We construct the trivariate portfolios in a similar manner, except before sorting on $MLER$ we sort on both Mom and Rev. Specifically, each month t we sort the stocks in our sample into 25 groups based on the intersections of 5 Mom groups and 5 Rev groups constructed from independent sorts on Mom and Rev. Using stocks within each of the 25 groups, we then sort stocks into five $MLER$ portfolios. The result is 125 portfolios, for which we calculate the month t value-weighted excess returns. We then define the excess return of the trivariate quintile $MLER$ quintile k portfolio to be the equal-weighted average of the month t excess returns of the 25 portfolios (one for each intersection of Mom and Rev groups) corresponding to the kth $MLER$ quintile. Finally, we define month t trivariate $MLER\,5-1$ portfolio excess return to be the month t excess return of the trivariate $MLER$ quintile five portfolio minus that of the trivariate $MLER$ quintile one portfolio.

We use quintile portfolios, instead of decile portfolios, for these analyses to ensure that each portfolio is populated in each month. To enable comparison with a benchmark that does not control for either Mom or Rev, we also construct univariate $MLER$ quintile portfolios using exactly the same procedure as was used to construct the $MLER$ decile portfolios examined in Section 4, except with five, instead of 10, portfolios. The values of Mom, Rev, and $MLER$ used to construct the portfolios are measured as of the end of month $t-1$ and the breakpoints used in all sorts are calculated using only NYSE-listed stocks.

The time-series averages of the monthly univariate, bivariate, and trivariate $MLER$ quintile portfolios are shown in Table 12. The univariate $MLER\ 5-1$ generates an average excess return of 0.72% per month (t-statistic = 4.58) and the average excess returns of the individual quintile portfolios are monotically increasing across the $MLER$ quintiles. The average excess return of the bivariate $MLER\,5-1$ constructed to control for Mom is 0.75% per month (t-statistic = 6.46), which is even higher, both in magnitude and statsistical significance, than that of the univariate $MLER\ 5-1$ portfolio. The bivariate $MLER\ 5-1$ that controls for Rev produces an average excess return of 0.57% per month (t-statistic = 3.86). Finally, the trivariate $MLER\,5-1$ portfolio has an average excess return of 0.41% per month (t-statistic = 4.38). As with the univariate portfolios, both sets of bivariate portfolios and the trivariate portfolios exhibit monontically increasing average excess returns across the $MLER$ quintiles.

Our second tests examining whether the ML-based forecast has predictive power distinct from momentum and reversal are FM regression analyses with the month t excess stock return as the dependent variable and combinations of $MLER$, Mom, and Rev measured at the end of month $t-1$ as independent variables. We conduct the analyses using both OLS and WLS regressions, with the weight in the WLS regressions being proportional to market capitalization measured at the end of month $t-1$. The results of these analyses are shown in Table 13. Focussing on the WLS results (columns labeled "VW"), when $MLER$ is the only independent variable, the average coefficient on MLER of 3.24 (t-statistic = 5.47) is highly statistically significant. The average coefficients on $MLER$ from specifications that include only Mom, only Rev, and both Mom and Rev as additional independent variables are 2.73 (t-statistic = 4.92), 3.30 (t-statistic = 4.49), and 2.15 (t-statistic = 3.17), respectively. The moderate decrease in the average coefficient on $MLER$ and associated statistical significance in the specification that includes both Mom and Rev demonstrates that a portion of the predictive power of $MLER$ is associated with the momentum and reversal effects. However, that this coefficient remains highly statistically significant indicates that the ML-based forecast has predictive power that is not a manifestation of momentum or reversal. Results using OLS regressions (columns labeled "EW") are qualitatively similar, although the coefficients on $MLER$ and t-statistics associated are higher in the OLS regressions compared to those in the WLS regressions.

Taken together, the results in Tables 12 and 13 show that while the predictive power of the ML-based forecast has a non-trival component that is associated with the momentum and reversal effects, it also has an economically important component that is unrelated to these pricing effects.

6 Conclusion

In this paper, we test the weak form of the EMH by examining whether ML-based forecasts based only on return data easily observable in historical price plots can predict the cross section of future stock returns. We begin by using the 192701-196306 period to determine the optimal implementation of our ML process. We find that using a convolutional neural network with long short-term memory architecture, the mean-squared-error as the loss function, weighting observations in the loss function in a manner that gives the same total weight to observations in each month and equal-weights each stock within each month, and a normalized measure of the future stock return as the dependent variable, optimizes the performance of the ML process during the 192701-196306 optimization period. By using data from the early part of our sample period to determine the optimal implementation of the ML process, we overcome concerns about the out-of-sample validity of our results.

We implement the ML process to generate forecasts of future stock returns during our 196307-201912 test period. Portfolio and regression analyses provide convincing evidence that the ML-based forecast is a strong predictor of the cross section of future stock returns. Factor analyses and other risk metrics provide no indication that the variation in average returns associated with the ML-based forecast reflects compensation for risk. Further tests demonstrate that the predictive power of the ML-based forecast is strong during most subperiods of our main test period, including the most recent 2015-2019 period. We also find that the predictive power remains strong even among the largest 500 stocks in our sample.

The forecasting function generated by the ML process is highly stable through time. Indeed, forecasts based on a fit of data from 192701-196306 only, perform nearly as well as our focal expanding window-based measure. Linear regressions show that the forecasting function is highly non-linear and that the non-linear aspects of this function account for a very substantial portion of the predictive power. Finally, while the ML-based forecast contains non-negligible components related to the momentum and reversal effects, the majority of the predictive power is unrelated to these effects.

Our results are strong evidence contrary to the main prediction of the EMH that profitable portfolios cannot be constructed from only information contained in historical returns. While one might reasonably argue that the momentum and reversal effects are already strong evidence contradicting the EMH, the complexity of the relations between past price patterns and future returns indicate that violations of the EMH are much more intricate than previously understood. Our findings also suggest that technical analysis, or charting, has greater merit than acknowledged in academic work, and potentially sheds light on why this investment technique remains prevalent among investment practitioners.

References

Allen, F. and Karjalainen, R. 1999. Using genetic algorithms to find technical trading rules. Journal of financial Economics, 51(2), 245–271.

Bajgrowicz, P. and Scaillet, O. 2012. Technical trading revisited: False discoveries, persistence tests, and transaction costs. Journal of Financial Economics, 106(3), 473–491.

Bali, T. G., Engle, R. F., and Murray, S. *Empirical Asset Pricing: The Cross Section of Stock Returns*. John Wiley & Sons, 2016.

Barra, S., Carta, S. M., Corriga, A., Podda, A. S., and Recupero, D. R. 2020. Deep learning and time series-to-image encoding for financial forecasting. IEEE/CAA Journal of Automatica SINCA, 7(3), 683–692.

Bessembinder, H. and Chan, K. 1998. Market efficiency and the returns to technical analysis. Financial management, pages 5–17.

Bianchi, D., Buchner, M., and Tamoni, A. 2020. Bond risk premia with machine learning. Available at SSRN: https://ssrn.com/abstract=3232721.

Brock, W., Lakonishok, J., and LeBaron, B. 1992. Simple technical trading rules and the stochastic properties of stock returns. The Journal of finance, 47(5), 1731–1764.

Bryzgalova, S., Pelger, M., and Zhu, J. 2020. Forest through the trees: Building cross-sections of stock returns. Available at SSRN: https://ssrn.com/abstract=3493458.

Carhart, M. M. 1997. On persistence in mutual fund performance. Journal of Finance, 52(1), 57–82.

Chang, P. K. and Osler, C. L. 1999. Methodical madness: Technical analysis and the irrationality of exchange-rate forecasts. The Economic Journal, 109(458), 636–661.

Chen, L., Pelger, M., and Zhu, J. 2020. Deep learning in asset pricing. Available at SSRN: https://ssrn.com/abstract=3350138.

Fama, E. F. 1991. Efficient capital markets: Ii. The journal of finance, 46(5), 1575–1617.

Fama, E. F. and French, K. R. 1992. The cross-section of expected stock returns. Journal of Finance, 47(2), 427–465.

Fama, E. F. and French, K. R. 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33(1), 3–56.

Fama, E. F. and French, K. R. 2015. A five-factor asset pricing model. Journal of Financial Economics, 116(1), 1–22.

Fama, E. F. and MacBeth, J. D. 1973. Risk, return, and equilibrium: Empirical tests. Journal of Political Economy, 81(3), 607–636.

Feng, G., He, J., and Polson, N. G. 2018. Deep learning for predicting asset returns. arXiv preprint: https://arxiv.org/pdf/1804.09314.pdf.

Feng, G., Polson, N. G., and Xu, J. 2020. Deep learning in characteristics-sorted factor models. Available at SSRN: https://ssrn.com/abstract=3243683.

Freyberger, J., Neuhierl, A., and Weber, M. 2020. Dissecting characteristics nonparametrically. The Review of Financial Studies, 33(5), 2326–2377.

Gehrig, T. and Menkhoff, L. 2006. Extended evidence on the use of technical analysis in foreign exchange. International Journal of Finance & Economics, 11(4), 327–338.

Goodfellow, I., Bengio, Y., and Courville, A. *Deep learning*. MIT press, 2016.

Green, J., Hand, J. R., and Zhang, X. F. 2017. The characteristics that provide independent information about average us monthly stock returns. The Review of Financial Studies, 30(12), 4389–4436.

Gu, S., Kelly, B., and Xiu, D. 2020. Empirical asset pricing via machine learning. The Review of Financial Studies, 33(5), 2223–2273.

Gu, S., Kelly, B. T., and Xiu, D. 2019. Autoencoder asset pricing models. Available at SSRN: https://ssrn.com/abstract=3335536.

Han, Y., Yang, K., and Zhou, G. 2013. A new anomaly: The cross-sectional profitability of technical analysis. Journal of Financial and Quantitative Analysis, 48(5), 1433–1461.

Harvey, C. R., Liu, Y., and Zhu, H. 2016. ... and the cross-section of expected returns. Review of Financial Studies, 29(1), 5–68.

Hoseinzade, E. and Haratizadeh, S. 2019. Cnnpred: Cnn-based stock market prediction using a diverse set of variables. Expert Systems with Applications, 129(9), 273–285.

Hou, K., Xue, C., and Zhang, L. 2015. Digesting anomalies: An investment approach. Review of Financial Studies, 28(3), 650–705.

Hou, K., Xue, C., and Zhang, L. 2020. Replicating anomalies. The Review of Financial Studies, 33(5), 2019–2133.

Jegadeesh, N. 1990. Evidence of predictable behavior of security returns. Journal of Finance, 45(3), 881–898.

Jegadeesh, N. 2000. Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation: Discussion. The Journal of Finance, 55(4), 1765–1770.

Jegadeesh, N. and Titman, S. 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. Journal of Finance, 48(1), 65–91.

Kelly, B. T., Pruitt, S., and Su, Y. 2019. Characteristics are covariances: A unified model of risk and return. Journal of Financial Economics, 134(3), 501–524.

Kozak, S., Nagel, S., and Santosh, S. 2020. Shrinking the cross-section. Journal of Financial Economics, 135(2), 271–292.

LeBaron, B. 1999. Technical trading rule profitability and foreign exchange intervention. Journal of international economics, 49(1), 125–143.

LeCun, Y., Bengio, Y., and Hinton, G. 2015. Deep learning. nature, 521(7553), 436–444.

Lehmann, B. N. 1990. Fads, martingales, and market efficiency. Quarterly Journal of Economics, 105(1), 1–28.

Levich, R. M. and Thomas III, L. R. 1993. The significance of technical trading-rule profits in the foreign exchange market: a bootstrap approach. Journal of international Money and Finance, 12(5), 451–474.

Linnainmaa, J. T. and Roberts, M. R. 2018. The history of the cross-section of stock returns. The Review of Financial Studies, 31(7), 2606–2649.

Lo, A. W. and MacKinlay, A. C. 1990. Data-snooping biases in tests of financial asset pricing models. The Review of Financial Studies, 3(3), 431–467.

Lo, A. W., Mamaysky, H., and Wang, J. 2000. Foundations of technical analysis: Computational algorithms, statistical inference, and empirical implementation. The journal of finance, 55(4), 1705–1765.

McLean, R. D. and Pontiff, J. 2016. Does academic research destroy stock return predictability? Journal of Finance, 71(1), 5–32.

Menkhoff, L. 2010. The use of technical analysis by fund managers: International evidence. Journal of Banking & Finance, 34(11), 2573–2586.

Messmer, M. 2017. Deep learning and the cross-section of expected stock returns. Available at SSRN: https://ssrn.com/abstract=3081555.

Messmer, M. and Audrino, F. 2017. The (adaptive) lasso in the zoo - firm characteristic selection in the cross-section of expected returns. Available at SSRN: https://ssrn.com/abstract=2930436.

Moritz, B. and Zimmermann, T. 2016. Tree-based conditional portfolio sorts: The relation between past and future stock returns. Available at SSRN: https://ssrn.com/abstract=2740751.

Moskowitz, T. J., Ooi, Y. H., and Pedersen, L. H. 2012. Time series momentum. Journal of financial economics, 104(2), 228–250.

Neely, C., Weller, P., and Dittmar, R. 1997. Is technical analysis in the foreign exchange market profitable? a genetic programming approach. Journal of financial and Quantitative Analysis, 32(4), 405–426.

Neely, C. J. 2002. The temporal pattern of trading rule returns and exchange rate intervention: intervention does not generate technical trading profits. Journal of International Economics, 58(1), 211–232.

Neely, C. J., Rapach, D. E., Tu, J., and Zhou, G. 2014. Forecasting the equity risk premium: the role of technical indicators. Management science, 60(7), 1772–1791.

Newey, W. K. and West, K. D. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica, 55(3), 703–708.

Osler, C. L. 2003. Currency orders and exchange rate dynamics: An explanation for the predictive success of technical analysis. The Journal of Finance, 58(5), 1791–1819.

Pastor, L. and Stambaugh, R. F. 2003. Liquidity risk and expected stock returns. Journal of Political Economy, 111(3), 642–685.

Ready, M. J. 2002. Profits from technical trading rules. Financial Management, pages 43–61.

Rossi, A. G. 2018. Predicting stock market returns with machine learning. Working Paper.

Schwert, G. W. Anomalies and market efficiency. In G. Constantinides, M. Harris, and R. M. Stulz, editors, *Handbook of the Economics of Finance*, chapter 15, pages 939–974. Elsevier Science, B.V., Netherlands, 2003.

Shumway, T. 1997. The delisting bias in CRSP data. Journal of Finance, 52(1), 327–340.

Sullivan, R., Timmermann, A., and White, H. 1999. Data-snooping, technical trading rule performance, and the bootstrap. The journal of Finance, 54(5), 1647–1691.

Sweeney, R. J. 1986. Beating the foreign exchange market. The Journal of Finance, 41(1), 163–182.

Zhu, Y. and Zhou, G. 2009. Technical analysis: An asset allocation perspective on the use of moving averages. Journal of Financial Economics, 92(3), 519–544.

Figure 1: Plot of Number of Stocks in the Sample
This figure plots the number of stocks in our sample for each month during our sample period.

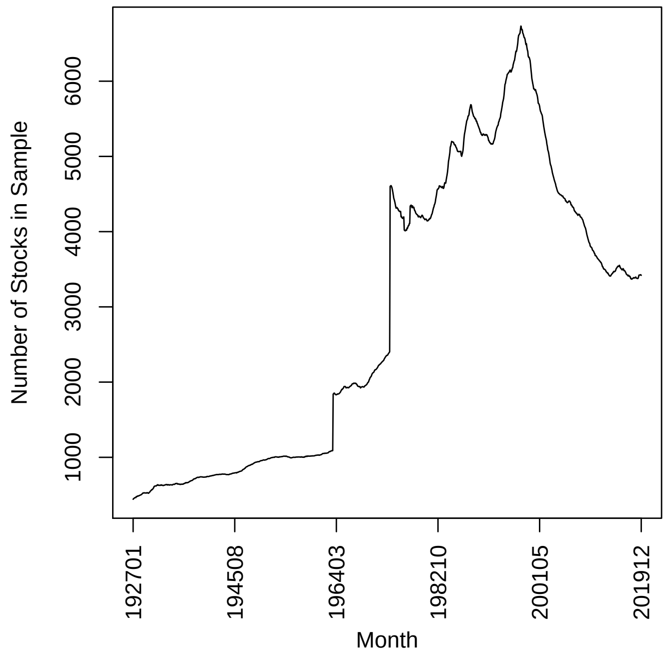
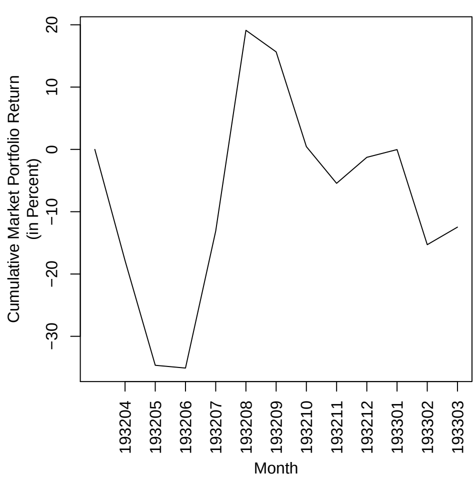


Figure 2: Market Portfolio Price Plots
This figure plots the monthly cumulative return of the market portfolio between April 1932 and March 1933 (Panel A) and between September 1930 and August 1931 (Panel B), inclusive.

Panel A: April 1932-March 1933



Panel B: September 1930-August 1931

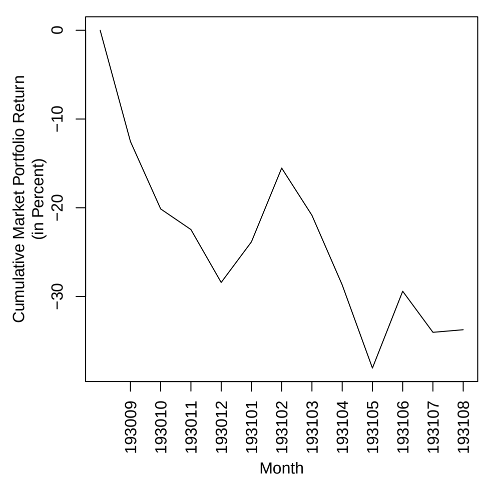


Figure 3: Plot of Correlations Between Forecasts for Different Values of N_F
This figure plots the average (solid line), minimum (dashed line), and maximum (dotted line) correlation between ML-based forecasts as a function of the number of times the ML process is executed to generate the forecast (N_F). The procedure used to generate the correlations is described in detail in Section 3.1

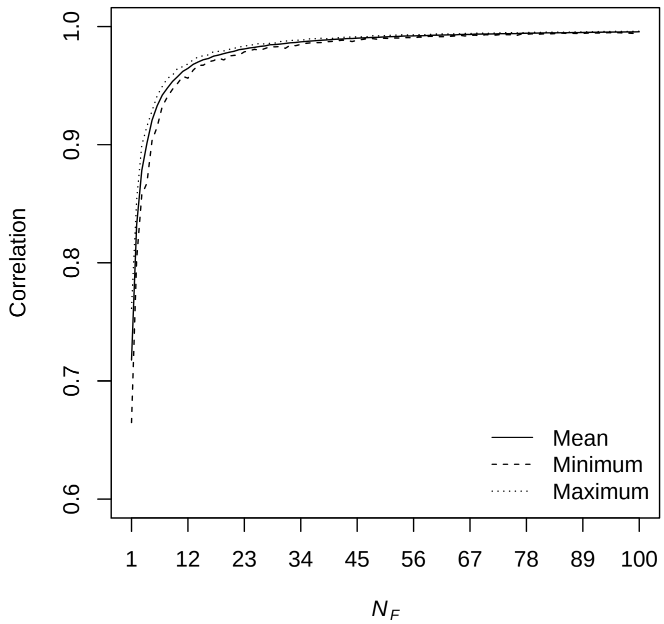


Figure 4: Plot of Performance of $MLER$ Long-Short Portfolio

This figure plots the cumulative log excess return of the $MLER$ $10-1$ portfolio. To calculate the cumulative log excess portfolio returns, we begin by adding the return of the risk-free asset to the monthly excess returns of the portfolio, giving the portfolio return in month t. We then compound these returns beginning in 196307 and ending in 201912 to generate a cumulative return series, and take the natural log of the values in that series to be the cumulative log return of the portfolio. We then calculate the compounded risk-free asset return beginning in 196307 and ending in 201912 and take the natural log of the compounded risk-free asset returns to be the cumulative log risk-free asset return. We cumulative log excess return of the $MLER$ $10-1$ portfolio in any month t is taken to be the cumulative log return of the portfolio in that month minus the cumulative log risk-free asset return for the same month.

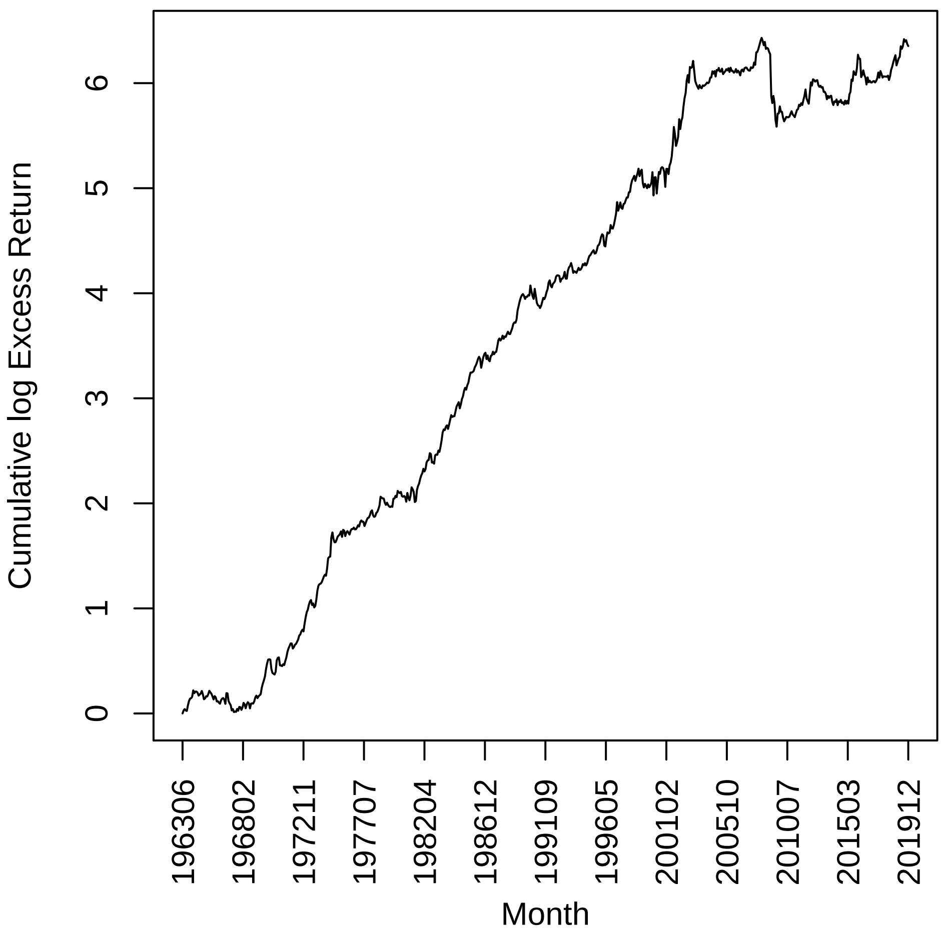


Table 1: ML Process Optimization

This table presents the results of our tests used to determine the optimal parameters for the ML process. The tests are run using only data from the 192701-196306 optimization period. We examine all combinations of four ML architectures, two loss functions, three loss function weighting methodologies, and four dependent variables. The ML architectures are feed-forward neural network (FNN), convolutional neural network (CNN), long-short term memory (LSTM), and convoluational neural network with long-short term memory (CNNLSTM). The loss functions are the mean squared error (MSE) and the mean absolute error (MAE). The loss function weighting methodologies are to equal-weight each observation (EW), to equal-weight each month, and within each month give equal weight to each observation (EWPM), and to equal-weight each month and within each month weight each observation according to its market capitalization (EWPMVW). The four dependent variables we examine are the excess stock return (r), the standardized excess stock return (r_{Std}), the normalized excess stock return (r_{Norm}), and the percentile of the stock return (r_{Pctl}). Using each of the 96 possible combinations of parameters, we execute the ML process using data from the even months in even years and odd months in odd years to generate a forecast function. We then apply the forecast function to generate return forecasts for returns in odd months in even years and even months in odd years. The table shows the time-series averages of the monthly cross-sectional Spearman rank correlations between the forecast and the actual excess return. The column labeled "Dependent Variable" indicates the dependent variable. The column labeled "Weighting Methodology" indicates the weighting methodology. The remaining column headers indicate the the ML architecture and loss function. The Spearman rank correlations are shown in percent.

Dependent Variable	Weighting Methodology	FNN MAE	FNN MSE	CNN MAE	CNN MSE	LSTM MAE	LSTM MSE	CNNLSTM MAE	CNNLSTM MSE
r	EW	4.6	1.6	4.0	1.0	5.5	−1.8	5.3	0.8
	EWPM	4.6	1.4	3.8	1.2	6.7	0.8	5.0	2.6
	EWPMVW	0.9	−0.5	0.5	−2.3	2.5	0.9	1.0	2.0
r_{Std}	EW	3.4	2.3	9.7	7.3	9.8	8.5	10.4	9.8
	EWPM	5.0	1.7	8.9	7.4	9.8	7.6	10.5	9.9
	EWPMVW	4.7	2.0	4.7	4.6	4.7	5.1	6.0	4.7
r_{Norm}	EW	7.5	1.5	9.7	8.8	10.3	10.2	10.4	10.7
	EWPM	6.9	1.4	9.1	8.0	10.1	10.2	10.6	10.8
	EWPMVW	3.7	0.4	4.3	4.3	5.2	6.3	6.9	7.5
r_{Pctl}	EW	−2.3	−1.6	7.7	−2.7	9.8	9.4	10.2	10.1
	EWPM	0.7	−3.0	7.8	−3.5	9.6	9.3	10.2	10.2
	EWPMVW	−0.1	−3.1	1.2	−1.7	4.9	5.7	7.3	7.8

Table 2: ML-Based Forecast Summary Statistics

This table presents summary statistics for the ML-based return forecast, $MLER$. $MLER$ is a forecast of the normalized excess stock return. The column labeled "Period" indicates the period used to calculate the summary statistics in the given row. The columns labeled "Mean", "S.D.", "Min.", "P_1", "P_5", "P_{25}", "Median", "P_{75}", "P_{95}", "P_{99}", and "Max." present the time-series averages of the monthly cross-sectional mean, standard deviation, and the minimum, first percentile, fifth percentile, 25th percentile, median, 75th percentile, 95th percentile, 99th percentile, and maximum, respectively, values of $MLER$. The column labeled "n" shows the time-series average of the number of observations in each month. The set of stocks included in the month t sample is all common shares of US-based firms that are listed on the NYSE, American Stock Exchange, or NASDAQ as of the end of month $t-1$. We also require that a return is available for each stock in each of months $t-12$ through $t-1$, and that the stock's market capitalization as of the end of month $t-1$ is available. Values of $MLER$ are calculated as of the end of month $t-1$.

Period	Mean	S.D.	Min.	P_1	P_5	P_{25}	Median	P_{75}	P_{95}	P_{99}	Max.	n
196307-201912	−0.02	0.11	−0.61	−0.34	−0.22	−0.07	−0.00	0.05	0.12	0.20	0.39	4198
196307-197412	−0.01	0.09	−0.53	−0.28	−0.17	−0.06	−0.00	0.04	0.12	0.20	0.41	2245
197501-198412	−0.01	0.11	−0.75	−0.35	−0.21	−0.06	0.01	0.06	0.13	0.20	0.40	4332
198501-199412	−0.02	0.12	−0.66	−0.38	−0.24	−0.07	−0.00	0.05	0.15	0.26	0.57	5337
199501-200412	−0.04	0.12	−0.61	−0.36	−0.26	−0.11	−0.02	0.04	0.12	0.20	0.38	5792
200501-201412	−0.01	0.10	−0.53	−0.33	−0.21	−0.06	0.00	0.05	0.11	0.15	0.26	3956
201501-201912	−0.01	0.10	−0.59	−0.31	−0.21	−0.06	0.01	0.06	0.11	0.14	0.20	3445

Table 3: Portfolio Analysis

This table presents the results of portfolio analyses examining the ability of $MLER$ to predict the cross section of future stock returns. At the end of each month $t-1$, all stocks in the month t sample are sorted into 10 portfolios based on an ascending ordering of $MLER$. The breakpoints used to determine which stocks are in which portfolio are the deciles of $MLER$ calculated using only stocks listed on the NYSE. The month t excess return of each portfolio is then taken to be the market capitalization-weighted average month t excess return of all stocks in the portfolio, with market capitalization calculated as of the end of month $t-1$. We also calculate the excess return of a zero-cost portfolio that is long portfolio 10 and short portfolio 1. The columns labeled "$MLER$ 1" through "$MLER$ 10" present results for decile portfolios 1 through 10. The column labeled "$MLER$ 10-1" presents results for the zero-cost long-short portfolio that is long the decile 10 portfolio and short the decile 1 portfolio. The row label "Excess Return" presents the time-series averages of the monthly portfolio excess returns for each of the portfolios. The rows labeled "α^{CAPM}", "α^{FF}", "α^{FFC}", "α^{FFCLIQ}", "α^{FF5}", and "α^{Q}" present alphas from a factor analysis using the CAPM, FF, FFC, FFCLIQ, FF5, and Q factor models, respectively. The rows labeled "S.D." and "Skewness" present the standard deviation and skewness of the monthly excess portfolio returns. The rows labeled "P_1" and "P_5" present the first and fifth percentiles of the monthly excess return, which measure value at risk. The rows labeled "ES_1" and "ES_5" present the expected shortfall using a 1% and 5% threshold, respectively. All excess returns and alphas are reported in percent per month. The values in parentheses are t-statistics, calculated following Newey and West (1987) using 12 lags, testing the null hypothesis that the average monthly excess return or alpha is equal to zero. The analysis covers return months t from July 1963 through December 2019, inclusive.

Value	$MLER$ 1	$MLER$ 2	$MLER$ 3	$MLER$ 4	$MLER$ 5	$MLER$ 6	$MLER$ 7	$MLER$ 8	$MLER$ 9	$MLER$ 10	$MLER$ $10-1$
Excess Return	−0.13	0.31	0.40	0.51	0.51	0.65	0.68	0.66	0.75	0.93	1.06
	(−0.50)	(1.35)	(1.98)	(2.62)	(2.80)	(4.03)	(3.99)	(3.91)	(4.19)	(5.00)	(5.38)
α^{CAPM}	−0.80	−0.32	−0.17	−0.06	−0.02	0.13	0.17	0.16	0.25	0.38	1.18
	(−6.20)	(−3.48)	(−2.27)	(−0.97)	(−0.42)	(2.19)	(3.43)	(2.43)	(3.68)	(3.86)	(6.17)
α^{FF}	−0.86	−0.37	−0.21	−0.09	−0.05	0.10	0.15	0.12	0.25	0.36	1.22
	(−7.80)	(−4.19)	(−2.83)	(−1.63)	(−0.85)	(1.80)	(3.02)	(2.47)	(4.06)	(3.80)	(7.00)
α^{FFC}	−0.58	−0.20	−0.04	0.07	0.05	0.13	0.12	0.05	0.12	0.19	0.78
	(−6.21)	(−2.38)	(−0.44)	(1.29)	(0.84)	(2.09)	(2.26)	(0.98)	(1.89)	(2.11)	(5.03)
α^{FFCLIQ}	−0.58	−0.20	−0.05	0.08	0.04	0.14	0.11	0.06	0.12	0.21	0.78
	(−5.89)	(−2.28)	(−0.59)	(1.39)	(0.67)	(2.19)	(2.01)	(1.10)	(1.88)	(2.16)	(4.84)
α^{FF5}	−0.68	−0.25	−0.17	−0.06	−0.06	0.09	0.10	0.03	0.18	0.31	0.99
	(−6.27)	(−2.77)	(−2.07)	(−0.89)	(−0.94)	(1.67)	(2.01)	(0.63)	(2.65)	(3.56)	(6.27)
α^{Q}	−0.50	−0.15	−0.06	0.05	0.03	0.13	0.08	−0.04	0.12	0.21	0.71
	(−4.27)	(−1.70)	(−0.61)	(0.59)	(0.38)	(2.03)	(1.32)	(−0.70)	(1.57)	(1.89)	(3.66)
S.D.	6.31	5.52	5.04	4.93	4.53	4.45	4.36	4.28	4.38	4.98	
Skewness	−0.10	−0.23	−0.34	−0.21	−0.41	−0.49	−0.58	−0.49	−0.03	−0.07	
P_1	−17.95	−13.69	−13.04	−12.22	−10.87	−10.73	−12.08	−10.24	−10.63	−13.40	
P_5	−10.72	−8.89	−7.96	−7.65	−7.11	−6.56	−6.50	−6.37	−6.54	−6.96	
ES_1	−20.35	−18.22	−17.52	−15.86	−15.14	−15.40	−15.11	−14.77	−14.29	−17.13	
ES_5	−14.90	−12.45	−11.19	−10.88	−10.09	−10.02	−9.83	−9.78	−9.41	−11.21	

Table 4: Portfolio Subperiod Analysis

This table presents the results describing the performance the $MLER$ decile portfolios during different subperiods. The portfolios are the exact same portfolios whose performances were examined in Table 3. The column labeled "Subperiod" indicates the subperiod covered by each analysis. The columns labeled "$MLER$ 1" through "$MLER$ 10" and "$MLER$ 10-1" present results for the portfolio indicated in the column header. The table shows the average excess return of each portfolio in each subperiod. All excess returns are reported in percent per month. The values in parentheses are t-statistics, calculated following Newey and West (1987) using 12 lags, testing the null hypothesis that the average monthly excess return is equal to zero.

Subperiod	$MLER$ 1	$MLER$ 2	$MLER$ 3	$MLER$ 4	$MLER$ 5	$MLER$ 6	$MLER$ 7	$MLER$ 8	$MLER$ 9	$MLER$ 10	$MLER$ $10-1$
196307-197412	−0.90	−0.47	−0.26	−0.14	−0.20	−0.03	0.15	0.07	−0.03	0.25	1.15
	(−1.91)	(−1.10)	(−0.63)	(−0.34)	(−0.53)	(−0.09)	(0.40)	(0.20)	(−0.08)	(0.58)	(3.93)
197501-198412	0.16	0.27	0.42	0.65	0.65	0.66	0.85	0.81	1.03	1.47	1.32
	(0.29)	(0.55)	(0.96)	(1.48)	(1.58)	(1.59)	(2.01)	(1.96)	(2.31)	(2.95)	(3.63)
198501-199412	−0.17	0.49	0.44	0.63	0.85	0.78	0.79	0.84	0.98	1.00	1.17
	(−0.40)	(1.16)	(1.08)	(1.61)	(2.06)	(1.86)	(1.87)	(1.99)	(2.24)	(2.11)	(3.66)
199501-200412	−0.14	0.66	0.50	0.68	0.55	0.93	0.72	0.94	1.10	1.57	1.71
	(−0.19)	(1.10)	(0.94)	(1.31)	(1.27)	(2.12)	(1.85)	(2.62)	(3.03)	(3.66)	(2.85)
200501-201412	0.48	0.49	0.91	0.62	0.59	0.97	0.88	0.64	0.56	0.40	−0.08
	(0.76)	(0.86)	(1.80)	(1.31)	(1.43)	(2.39)	(2.28)	(1.63)	(1.47)	(0.86)	(−0.17)
201501-201912	−0.05	0.72	0.57	0.95	0.91	0.67	0.86	0.79	1.18	1.01	1.06
	(−0.06)	(1.17)	(0.92)	(1.61)	(1.63)	(1.47)	(1.90)	(1.69)	(2.43)	(2.36)	(1.61)

Table 5: Predictive Power Beyong One Month

This table presents the results describing the performance the $MLER$ decile portfolios beyond one month after the measurement of $MLER$. $MLER$ is calculated as of the end of month $t-1$. We then wait k months before constructing the portfolios. The portfolios are then held for the duration of month $t+k$. Aside from the timing, the construction of the portfolios is exactly the same as was used to construct the portfolios whose performances are examined in Table 3. The column labeled "k" indicates the number of months between the calculation of $MLER$ and formation of the portfolios. The columns labeled "$MLER$ 1" through "$MLER$ 10" and "$MLER$ 10-1" present results for the portfolio indicated in the column header. The table shows the average excess return of each portfolio. All excess returns are reported in percent per month. The values in parentheses are t-statistics, calculated following Newey and West (1987) using 12 lags, testing the null hypothesis that the average monthly excess return is equal to zero.

k	$MLER$ 1	$MLER$ 2	$MLER$ 3	$MLER$ 4	$MLER$ 5	$MLER$ 6	$MLER$ 7	$MLER$ 8	$MLER$ 9	$MLER$ 10	$MLER$ $10-1$
1	0.31	0.53	0.61	0.52	0.55	0.55	0.56	0.65	0.68	0.81	0.50
	(1.30)	(2.31)	(2.83)	(2.57)	(3.02)	(3.01)	(3.30)	(4.01)	(4.17)	(4.71)	(3.00)
2	0.39	0.51	0.54	0.56	0.56	0.58	0.56	0.64	0.62	0.72	0.34
	(1.49)	(2.22)	(2.44)	(2.96)	(3.03)	(3.25)	(3.39)	(3.72)	(3.67)	(3.95)	(2.04)
3	0.40	0.45	0.53	0.46	0.54	0.55	0.59	0.61	0.65	0.79	0.39
	(1.48)	(1.91)	(2.51)	(2.41)	(3.00)	(3.14)	(3.36)	(3.61)	(3.73)	(4.42)	(2.01)
4	0.52	0.49	0.51	0.43	0.50	0.61	0.56	0.65	0.63	0.79	0.27
	(1.96)	(1.96)	(2.33)	(2.11)	(2.80)	(3.59)	(3.35)	(3.75)	(3.82)	(4.07)	(1.36)
5	0.46	0.54	0.46	0.46	0.55	0.48	0.63	0.71	0.61	0.72	0.26
	(1.66)	(2.45)	(2.25)	(2.25)	(2.98)	(2.65)	(3.78)	(4.25)	(3.68)	(3.51)	(1.13)
6	0.47	0.50	0.58	0.51	0.47	0.55	0.56	0.62	0.69	0.78	0.31
	(1.84)	(2.25)	(2.81)	(2.45)	(2.62)	(3.20)	(2.98)	(3.70)	(4.35)	(4.12)	(1.75)

Table 6: Predictive Power Among Large Stocks

This table presents the results of portfolio analyses examining the ability of the ML-based forecast to predict the cross section of future stock returns among large stocks. The construction of the portfolios is exactly the same as was used to construct the portfolios whose performances are examined in Table 3, except that we vary the set of stocks used in the calculation of the portfolio breakpoints and the set of stocks that are sorted into the portfolios. The column labeled "Breakpoints" indicates the set of stocks used to calculate the breakpoints. The column labeled "Holdings" indicates the set of stocks sorted into the portfolios. The $Size > P_{20}^{\mathrm{NYSE}}$ (NYSE/$Size > P_{20}^{\mathrm{NYSE}}$) and $Size > P_{50}^{\mathrm{NYSE}}$ (NYSE/$Size > P_{50}^{\mathrm{NYSE}}$) subsets include all (NYSE-listed) stocks with market capitalizations greater than the 20th and 50th percentile values, respectively, among NYSE-listed stocks as of the end of month $t-1$. The Top 500 subset includes the top 500 stocks by market capitalization at the end of month $t-1$. The columns labeled "$MLER$ 1" through "$MLER$ 10" and "$MLER$ 10-1" present results for the portfolio indicated in the column header. The table shows the average excess return of each portfolio. All excess returns are reported in percent per month. The values in parentheses are t-statistics, calculated following Newey and West (1987) using 12 lags, testing the null hypothesis that the average monthly excess return is equal to zero. The analyses covers return months t from July 1963 through December 2019, inclusive.

Breakpoints	Holdings	$MLER$ 1	$MLER$ 2	$MLER$ 3	$MLER$ 4	$MLER$ 5	$MLER$ 6	$MLER$ 7	$MLER$ 8	$MLER$ 9	$MLER$ 10	$MLER$ 10 − 1
$Size > P_{20}^{\mathrm{NYSE}}$	$Size > P_{20}^{\mathrm{NYSE}}$	−0.13	0.34	0.44	0.50	0.51	0.57	0.74	0.61	0.73	0.92	1.05
		(−0.51)	(1.47)	(2.18)	(2.65)	(2.78)	(3.56)	(4.33)	(3.68)	(4.07)	(4.97)	(5.63)
NYSE/$Size > P_{20}^{\mathrm{NYSE}}$	$Size > P_{20}^{\mathrm{NYSE}}$	−0.00	0.36	0.42	0.58	0.48	0.61	0.74	0.62	0.73	0.90	0.91
		(−0.02)	(1.64)	(2.17)	(3.13)	(2.73)	(3.79)	(4.42)	(3.63)	(4.16)	(4.85)	(5.08)
$Size > P_{50}^{\mathrm{NYSE}}$	$Size > P_{50}^{\mathrm{NYSE}}$	0.04	0.36	0.42	0.51	0.47	0.63	0.63	0.62	0.76	0.83	0.79
		(0.17)	(1.69)	(2.15)	(2.69)	(2.76)	(3.75)	(3.64)	(3.60)	(4.37)	(4.38)	(4.46)
NYSE/$Size > P_{50}^{\mathrm{NYSE}}$	$Size > P_{50}^{\mathrm{NYSE}}$	0.11	0.37	0.42	0.58	0.52	0.62	0.65	0.65	0.74	0.85	0.75
		(0.44)	(1.81)	(2.12)	(3.32)	(3.09)	(3.62)	(3.85)	(3.80)	(4.22)	(4.59)	(4.48)
Top 500	Top 500	0.10	0.41	0.32	0.57	0.46	0.65	0.57	0.63	0.75	0.79	0.69
		(0.45)	(1.90)	(1.65)	(3.08)	(2.77)	(3.91)	(3.14)	(3.68)	(4.10)	(4.28)	(4.11)

Table 7: Stability of Forecasting Relation

This table presents the results of portfolio analyses examining the ability of ML-based forecasts generated by applying the ML process to data from one subperiod to predict the cross section of future stock returns during other subperiods. The column labeled "Return Period" indicates the period for which the excess portfolio returns are examined. The remaining column named indicate the variable used to construct the portfolios. $MLER^{t_1-t_2}$ is the ML-based forecast based on applying the ML process to data from months t between t_1 and t_2, inclusive. With the exception of the sort variable, the portfolio construction methodology is identical to that used to construct the portfolios examined in Table 3. The table presents the average monthly excess return for the $10-1$ period formed by sorting on the given variable during the given subperiod. All excess returns are reported in percent per month. The values in parentheses are t-statistics, calculated following Newey and West (1987) using 12 lags, testing the null hypothesis that the average monthly excess return is equal to zero. Each analysis covers return months from k months after July 1963 through December 2019, inclusive.

Return Period	$MLER^{192701,196306}$	$MLER^{196307,197412}$	$MLER^{197501,198412}$	$MLER^{198501,199412}$	$MLER^{199501,200412}$	$MLER^{200501,201412}$
196307-197412	1.15		1.35	1.46	1.43	0.82
	(3.93)		(3.49)	(3.57)	(3.62)	(1.85)
197501-198412	1.23	1.43		1.19	1.32	0.58
	(3.01)	(4.09)		(3.18)	(3.47)	(1.41)
198501-199412	0.86	1.27	1.28		1.21	1.05
	(2.94)	(3.79)	(3.83)		(3.16)	(2.57)
199501-200412	1.57	1.44	1.08	1.31		0.52
	(3.12)	(2.42)	(1.44)	(1.95)		(0.62)
200501-201412	0.02	−0.18	0.02	0.20	−0.02	
	(0.06)	(−0.45)	(0.05)	(0.35)	(−0.04)	
201501-201912	0.85	0.63	0.87	1.23	0.78	0.48
	(2.09)	(1.06)	(1.24)	(1.54)	(0.97)	(0.63)

Table 8: Rolling-Window and Static Forecast-Based Portfolios

This table presents the results of portfolio analyses examining the ability of ML-based forecasts based on rolling-window ($MLER^{Roll}$) and static-window ($MLER^{192701,196306}$) fit periods to predict the cross section of future stock returns. $MLER^{Roll}$ for observations in months t between 1963701 and 197412, 197501 and 198412, 198501 and 199412, 199501 and 200412, 200501 and 201412, and 201501 and 201912 is the forecast generated by applying the ML process to data from 192701 through 196306, 196307 through 197412, 197501 through 198412, 198501 through 199412, 199501 through 200412, and 200501 through 201412, respectively. $MLER^{192701,196306}$ is the forecast generated by applying the ML process to data from 192701 through 196306. The column labeled "Sort Variable" indicates the variable used to sort stocks into portfolios. The columns labeled "1" through "10" and "10 − 1" present results for decile portfolios 1 through 10 and the $10-1$ portfolio, respectively. With the exception of the sort variable, the portfolio construction methodology is identical to that used to construct the portfolio examined in Table 3. The table presents the time-series averages of the monthly portfolio excess returns for each of the portfolios. All excess returns are reported in percent per month. The values in parentheses are t-statistics, calculated following Newey and West (1987) using 12 lags, testing the null hypothesis that the average monthly excess return is equal to zero. The analysis covers return months t from July 1963 through December 2019, inclusive.

Sort Variable	1	2	3	4	5	6	7	8	9	10	$10-1$
$MLER$	−0.13	0.31	0.40	0.51	0.51	0.65	0.68	0.66	0.75	0.93	1.06
	(−0.50)	(1.35)	(1.98)	(2.62)	(2.80)	(4.03)	(3.99)	(3.91)	(4.19)	(5.00)	(5.38)
$MLER^{Roll}$	−0.14	0.30	0.40	0.55	0.55	0.58	0.73	0.64	0.72	0.85	0.98
	(−0.48)	(1.28)	(1.81)	(2.70)	(3.13)	(3.29)	(4.43)	(3.94)	(4.33)	(4.89)	(4.57)
$MLER^{192701,196306}$	−0.05	0.30	0.51	0.58	0.48	0.61	0.71	0.66	0.72	0.91	0.96
	(−0.22)	(1.48)	(2.66)	(3.32)	(2.62)	(3.44)	(3.98)	(3.79)	(4.04)	(4.86)	(5.81)

Table 9: Linearity of ML-Based Forecast

This table presents the results of Fama and MacBeth (1973) regression analyses examining the linearity of the ML-based forecast $MLER$ in the cumulative returns CR_1, …, CR_{12}. Each month t we run a cross-sectional regression of the dependent variable on one or more independent variables. The results in Panel A are for regressions with $MLER$ as the dependent variable and CR_1, …, CR_{12} as independent variables. The results in Panels B and C are for regressions with the future excess stock return as the dependent variable and combinations of $MLER$ and CR_1, …, CR_{12} as independent variables. The columns labeled "$MLER$", "CR_1", …, and "CR_{12}" present the time-series averages of the monthly cross-sectional coefficients on the variables indicated in the column header, along with t-statistics, calculated following Newey and West (1987) using 12 lags, testing the null hypothesis that the average coefficient is equal to zero (in parentheses). The column labeled "Adj. R^2" in Panel A presents the time-series average of adjusted R^2 values from the monthly cross-sectional regressions. In Panels B and C, each pair of rows presents the results for an analysis using a different set of independent variables, and all independent variables are winsorized at the 0.5% level on a monthly basis. The results in Panels A and B are for OLS regressions, and the results in Panel C are for WLS regressions with weights proportional to market capitalization. The analysis covers sample months t from July 1963 through December 2019, inclusive.

Panel A: FM Regresssions of $MLER$

CR_1	CR_2	CR_3	CR_4	CR_5	CR_6	CR_7	CR_8	CR_9	CR_{10}	CR_{11}	CR_{12}	Adj. R^2
0.030	0.023	0.006	0.009	0.004	0.002	0.011	0.033	−0.067	0.135	0.255	−0.283	37.46%
(11.61)	(12.90)	(2.74)	(6.14)	(2.35)	(1.11)	(3.06)	(9.07)	(−11.39)	(17.57)	(17.09)	(−15.68)	

Panel B: FM Regresssions of Future Excess Returns - Equal-Weighted

$MLER$	CR_1	CR_2	CR_3	CR_4	CR_5	CR_6	CR_7	CR_8	CR_9	CR_{10}	CR_{11}	CR_{12}
6.80												
(10.05)												
6.07	0.01	0.00	−0.00	0.00	−0.00	−0.00	−0.00	−0.00	0.00	−0.00	0.02	−0.01
(9.91)	(1.98)	(1.32)	(−1.76)	(1.07)	(−1.12)	(−0.25)	(−0.48)	(−0.68)	(0.38)	(−0.70)	(4.37)	(−2.60)

Panel C: FM Regresssions of Future Excess Returns - Value-Weighted

$MLER$	CR_1	CR_2	CR_3	CR_4	CR_5	CR_6	CR_7	CR_8	CR_9	CR_{10}	CR_{11}	CR_{12}
3.24												
(5.37)												
2.68	0.00	0.01	−0.01	0.01	0.00	−0.01	0.01	0.00	−0.00	0.00	0.02	−0.01
(5.15)	(0.40)	(1.52)	(−3.56)	(3.93)	(0.09)	(−3.00)	(2.37)	(0.54)	(−1.38)	(0.41)	(4.16)	(−2.70)

Table 10: Momentum and Reversal Portfolios

This table presents the results of portfolio analyses examining the ability of Mom and Rev to predict the cross section of future stock returns. With the exception of the sort variable, the portfolio construction methodology is identical to that used to construct the portfolio examined in Table 3. The column labeled "Sort Variable" indicates the variable used to sort stocks into decile portfolios. The columns labeled "1", ..., "10" and "10 − 1" present the time-series averages of the monthly portfolio excess returns, along with t-statistics, calculated following Newey and West (1987) using 12 lags, testing the null hypothesis that the average monthly excess return is equal to zero (in parentheses), for each of the portfolios. All excess returns are reported in percent per month. The analysis covers return months t from July 1963 through December 2019, inclusive.

Sort Variable	1	2	3	4	5	6	7	8	9	10	10 − 1
Mom	−0.16	0.34	0.49	0.52	0.49	0.55	0.55	0.69	0.74	1.11	1.26
	(−0.47)	(1.41)	(2.32)	(2.78)	(2.93)	(3.23)	(3.45)	(4.00)	(4.03)	(4.75)	(4.70)
Rev	0.70	0.83	0.77	0.68	0.63	0.54	0.53	0.52	0.34	0.30	−0.41
	(2.64)	(4.17)	(3.85)	(3.83)	(3.77)	(3.14)	(3.05)	(2.90)	(1.79)	(1.39)	(−2.31)

Table 11: Relations between ML-Based Forecast, Momentum, and Reversal

This table presents the results of Fama and MacBeth (1973) regression analyses examining the ability of Mom and Rev to explain variation in $MLER$. The process for conducting the Fama and MacBeth (1973) regression analyses is exactly as described in Table 9. The dependent variable in the regressions is $MLER$ and the independent variables are combinations of Mom and Rev. Results for different regression specifications are shown in different columns of the table. The table presents the time-series averages of monthly cross-sectional regression coefficients, along with t-statistics, calculated following Newey and West (1987) using 12 lags, testing the null hypothesis that the average coefficient is equal to zero (in parentheses). The rows labeled "Mom" and "Rev" present results pertaining to the coefficient on the indicated variable. The row labeled "R^2" shows the average R^2 from the cross-sectional regressions. The analysis covers return months t from July 1963 through December 2019, inclusive.

	(1)	(2)	(3)
Mom	0.001		0.001
	(31.27)		(34.49)
Rev		−0.004	−0.004
		(−73.52)	(−75.95)
R^2	9.87%	27.14%	36.91%

Table 12: Multivariate Portfolio Analysis - Control for Momentum and Reversal
This table presents the results of univariate, bivariate, and trivariate portfolio analyses examining the ability of the ML-based forecast to predict the cross section of future stock returns after controlling for momentum and reversal. The procedure used to generate the univariate portfolios is identical to that used to generate the portfolios in Table 3, except we create five quintile portfolios instead of 10 decile portfolios. The bivariate portfolios are constructed by sorting all stocks into quintile based on a control variable, either Mom or Rev, and then within each quintile, into five $MLER$ portfolios. The trivariate portfolios are formed by independently sorting stocks into quintiles of Mom and Rev, and then sorting stocks in each of the 25 groups formed by the intersections of the Mom and Rev quintiles, into $MLER$ quintiles. Breakpoints for all sorts are calculated using only NYSE-listed stocks. Values of $MLER$, Mom, and Rev are calculated as of the end of month $t-1$. The month t excess return of each of the resulting portfolios is taken to be the market capitalization-weighted average month t excess return of all stocks in the given portfolio, with market capitalization calculated as of the end of month $t-1$. For the bivariate (trivariate) portfolios, the month t excess return of the $MLER$ quintile k portfolio is taken to be the equal-weighted average, across all quintiles of the control variable (across all 25 Mom and Rev groups), of the $MLER$ quintile k portfolio. Finally, the $MLER$ $5-1$ portfolio excess return is taken to be the difference between the $MLER$ quintile five and quintile one portfolio excess returns. The table presents the time-series averages of the monthly portfolio excess returns for each of the $MLER$ quintile portfolios. The column labeled "Control Variable" indicates the control variable(s) used to construct the portfolios. The columns labeled "$MLER$ 1" through "$MLER$ 5" and "$MLER$ $5-1$" present the average monthly excess portfolio returns along with t-statistics, calculated following Newey and West (1987) using 12 lags, testing the null hypothesis that the average monthly excess return of the given portfolio is equal to zero. All excess returns are reported in percent per month. The analysis covers return months t from July 1963 through December 2019, inclusive.

Control Variable(s)	$MLER$ 1	$MLER$ 2	$MLER$ 3	$MLER$ 4	$MLER$ 5	$MLER$ $5-1$
None	0.11	0.46	0.58	0.66	0.83	0.72
(Univariate)	(0.47)	(2.32)	(3.52)	(4.01)	(4.72)	(4.58)
Mom	0.10	0.37	0.61	0.75	0.84	0.75
(Bivariate)	(0.45)	(1.84)	(3.35)	(4.22)	(4.60)	(6.46)
Rev	0.24	0.53	0.63	0.63	0.81	0.57
(Bivariate)	(0.99)	(2.63)	(3.66)	(3.76)	(4.67)	(3.86)
Mom and Rev	0.37	0.49	0.61	0.73	0.78	0.41
(Trivariate)	(1.67)	(2.46)	(3.20)	(4.06)	(4.22)	(4.38)

Table 13: Fama and MacBeth Regressions with Momentum and Reversal
This table presents the results of Fama and MacBeth (1973) regression analyses examining the ability of $MLER$, Mom and Rev to predict the cross section of future stock returns. The process for conducting the Fama and MacBeth (1973) regression analyses is exactly as described in Table 9. The dependent variable in the regressions is the excess stock return in month t and the independent variables are combinations of $MLER$, Mom and Rev measured at the end of month $t-1$. Results for different regression specifications are shown in different columns of the table. Columns labeled "EW" present results for OLS regressions and columns labeled "VW" present results for weight-least squares regressions using market capitalization measured at the end of month $t-1$ as the weight variable. The table presents the time-series averages of monthly cross-sectional regression coefficients, along with t-statistics, calculated following Newey and West (1987) using 12 lags, testing the null hypothesis that the average coefficient is equal to zero (in parentheses). The rows labeled "$MLER$" , "Mom", and "Rev" present results pertaining to the coefficient on the indicated variable. The analysis covers return months t from July 1963 through December 2019, inclusive.

	EW	EW	EW	EW	VW	VW	VW	VW
$MLER$	6.804	6.603	4.536	3.883	3.237	2.731	3.299	2.147
	(13.78)	(14.14)	(7.22)	(6.87)	(5.47)	(4.92)	(4.49)	(3.17)
Mom		0.002		0.004		0.005		0.006
		(0.98)		(2.42)		(2.68)		(3.37)
Rev			−0.028	−0.033			−0.001	−0.010
			(−5.23)	(−6.61)			(−0.11)	(−1.58)