




Article

Forecasting U.S. Aggregate Stock Market Excess Return: Do Functional Data Analysis Add Economic Value?

João F. Caldeira ¹, Rangan Gupta ^{2,*} and Hudson S. Torrent ³

¹ Department of Economics, Universidade Federal de Santa Catarina & CNPq, Florianópolis 88040-970, Brazil; emailocaldeira@gmail.com

² Department of Economics, University of Pretoria, Pretoria 0002, South Africa

³ Department of Statistics, Universidade Federal do Rio Grande do Sul, Porto Alegre 91509-900, Brazil; hudson.torrent@ufrgs.br

* Correspondence: Rangan.Gupta@up.ac.za

Received: 16 September 2020; Accepted: 12 November 2020; Published: 16 November 2020



Abstract: This paper analyzes the forecast performance of historical S&P500 and Dow Jones Industrial Average (DJIA) excess returns while using nonparametric functional data analysis (NP-FDA). The empirical results show that the NP-FDA forecasting strategy outperforms not only the prevailing-mean model, but also the traditional univariate predictive regressions with standard predictors used in the literature and, most cases, also combination approaches that use all predictors jointly. In addition, our results clearly have important implications for investors, from an asset allocation perspective, a mean-variance investor realizes substantial economic gains. Indeed, our results show that NP-FDA is the only one individual model that can overcome the historical average forecasts for excess returns in statistically and economically significant manners for both S&P500 and DJIA during the entire period, NBER recession, and expansions periods.

Keywords: return forecast; nonparametric functional data analysis; performance evaluation; predictive regression; classical financial mathematics.

JEL Classification: C14; G11; G17

1. Introduction

Stock return predictability has been of considerable interest to practitioners and academics in finance. Practitioners attempt to improve asset allocation and risk management performance using real-time prediction of stock returns. Further, academics use information regarding stock return forecasting to generate more realistic asset-pricing models, to test the market efficiency hypothesis, and other financial problems. Therefore, it is not surprising that the search for accurate and reliable return forecasts has attracted great interest from both finance professionals and academics.

In the early 1970s, it was widely accepted that stock markets were efficient. However, over the last two decades, academics have found evidence of stock return predictability and highlighted the potential benefits for actively managed investment strategies (see [1], for a recent survey). While many studies report evidence of in-sample return predictability, out-of-sample return predictability remains controversial. In an influential article, Ref. [2] show that it is difficult to find models that can surpass even the most naive benchmarks out-of-sample. They observe that most individual predictor variables cannot provide statistically significant forecast improvements over a historical average benchmark.

Recent studies report evidence of excess return predictability that is based on macroeconomic variables, technical indicators, short interest rates, investor sentiment, and so on. On the other hand, several studies demonstrate that the out-of-sample forecasting ability of commonly used predictors can be improved while using various strategies: imposing economically motivated restrictions on the model [3–7], diffusion indices [8–11], using regime switching models [12–15], or combining forecasts from individual predictive models [1,5,7,16–18]. Particularly, ref. [16] find that a combination forecast approach is able to deliver consistently superior out-of-sample US equity premium prediction. Refs. [19,20] use partial least squares methodology in order to create a sentiment index for predicting future market returns and finding results that are both statistically and economically significant.

In this paper, we investigate the capacity of nonparametric functional data analysis (NP-FDA) to directly forecast the S&P500 and Dow Jones Industrial Average (DJIA) excess returns and compare their performance to that of the traditional regression model. We model monthly returns as curves in a functional space. Specifically, we assume that daily cumulative returns can be interpreted as curves describing the cumulative return during a given month. The daily cumulative return curves give more relevant information, because they show how returns evolve during the month. The NP-FDA methodology can be seen as a functional regression, in which the regressor is the curve for a given month, t . As our main objective is the forecast of monthly return 1-step-ahead, the regressand, in this case, is the last day of the next month, $t + 1$. Moreover, the NP-FDA estimator that is described in Section 2.1 may be viewed as a weighted mean of monthly returns in which weights are a function of a measure of proximity between a given curve and the other curves presented in the sample. An interesting relation between our methodology and the prevailing-mean emerges. With appropriate choices for bandwidth, b , and kernel function, $K(\cdot)$, our estimates coincide with those of the prevailing-mean. Nevertheless, since the bandwidth is automatically chosen by a cross-validation procedure and the kernel is a density with smooth decay, following the nonparametric literature, this coincidence will not happen in general. (For the NP-FDA estimator to coincide with the prevailing-mean the kernel function would have to be the density of the Uniform distribution.) In other words, we attempt, in this article, to forecast monthly returns while using a flexible nonparametric estimator that uses the information presented in the sample of curves of daily cumulative returns. For every fixed month, these curves exhibit a specific pattern, typically with some upward or downward momentum, disturbed by some noise. Therefore, it is tempting to study their statistical behaviour using the framework of NP-FDA.

We employ monthly S&P500 and DJIA excess returns and the 12 monthly economic fundamentals of [2] for a sample period ranging from January 1927 to December 2019 for S&P500 and January 1889 to June 2020 for DJIA, respectively. We also consider a set of economic variables originally proposed by [2] in the predictive regression models. The analysis is performed purely out-of-sample to inform real-time investment decisions. A detailed performance assessment of forecast combinations strategies in comparison to individual models is provided in terms of both statistical accuracy and economic relevance. Because the final goal of returns forecasting is to improve economic and financial decision making, we propose evaluating the accuracy and economic relevance of excess returns forecast based on a mean-variance investor with quadratic utility, as in [1].

Our main empirical finding is that the NP-FDA delivers a statistically significant monthly out-of-sample R^2 of 0.57% and 0.37% for DJIA and S&P500, respectively. This is considerably higher than the R^2 values for the best individual model and for the standard forecast combination based on all 12 predictors. In addition, the NP-FDA also delivers high economic value in the context of a dynamic mean-variance strategy. The average utility gain (certainty equivalent return) is approximately 0.98% and 0.28% per year over and above the historical mean benchmark for the full sample for DJIA and S&P500, respectively. When we consider the economic cycle, the results are also the most impressive, 0.70% per year in expansions, and 2.68% per year in recessions for the S&P500 dataset.

The remaining sessions of the paper are organized, as follows. In Section 2, the predictive regression framework is presented and the NP-FDA estimation is detailed. Section 3 describes the datasets and reports out-of-sample results for statistical and economic evaluation. Section 4 contains concluding remarks.

2. Excess Return Forecast Models

In this section, we outline our empirical framework to predict DJIA and S&P500 excess returns, which we apply to our data in subsequent sections. In addition, we describe the methods for statistical and economic evaluation of the accuracy of the return forecast. Now, let r_{t+1} denote the excess return of the index held from period t to period $t + 1$, in excess of the risk-free rate. We use the continuously compounded returns that were obtained by subtracting the short T-bill rate from the DJIA and S&P500 returns, including dividends, which is given by:

$$r_t = \log(1 + R_t) - \log(1 + R_{t-1}), \quad (1)$$

where R_t is the return on a broad stock market index and R_{t-1} denotes the risk-free rate from period $t - 1$ to t . The unit of time can be day, month, or year.

2.1. Functional Data Methodology and Nonparametric Estimation

In this section, we explain how the functional data analysis methodology can be applied to our problem (For further details about the method see [21,22]). Following [23], let \mathbf{X} be a functional random variable taking values in a semi-metric space (E, d) , where E is an infinite dimensional space and d is a semi-metric (as described below). Moreover, let X be an observation of \mathbf{X} . In particular, when \mathbf{X} denotes a random curve, it is convenient to identify $\mathbf{X} = \{x(\tau); \tau \in \mathcal{D}\}$ (respectively, $X = \{x(\tau); \tau \in \mathcal{D}\}$), where \mathcal{D} is the set of points at which the curve is evaluated. Now, consider the problem of predicting a dependent scalar random variable \mathbf{y} as a function of a functional regressor \mathbf{X} . Let

$$r_{NP}(X) = E(\mathbf{y} | \mathbf{X} = X) \quad (2)$$

be the nonlinear regression operator and $(\mathbf{X}_t, \mathbf{y}_t)_{t=1, \dots, T}$ be a sequence of random pairs taking values in $E \times \mathbb{R}$. Given a fixed element $X_s \in E$, the nonparametric estimator for $r_{NP}(X_s)$ is defined by

$$\hat{r}_{NP}(X_s) = \frac{\sum_{t=1}^T \mathbf{y}_t K(b^{-1}d(X_s, \mathbf{X}_t))}{\sum_{t=1}^T K(b^{-1}d(X_s, \mathbf{X}_t))}, \quad (3)$$

where K is an asymmetric kernel and $b > 0$ is a bandwidth. The estimator that is presented in (3) may be viewed as a weighted mean, in which the weights are ultimately determined by the kernel density. The argument of this kernel depends on the semi-metric, d , and the bandwidth, b . The proper selection of the bandwidth is of fundamental importance. Given an observed sample $(X_t, y_t)_{t=1, \dots, T}$, the optimal bandwidth, b_{opt} , may be selected by the following cross-validation procedure

$$b_{opt} = \arg \min_b CV(b), \quad (4)$$

where

$$CV(b) = T^{-1} \sum_{t=1}^T \left(y_t - \hat{r}_{NP}^{(-t)}(X_t) \right)^2, \quad (5)$$

and

$$\hat{r}_{NP}^{(-t)}(X_t) = \frac{\sum_{j=1, j \neq t}^T y_j K(d(X_t, X_j)/b)}{\sum_{j=1, j \neq t}^T K(d(X_t, X_j)/b)}. \quad (6)$$

The estimator that is presented in Equation (6) is well known in nonparametric literature as the *leave-one-out* estimator and it is also suggested by Ferraty et al. [24] and Ferraty and Vieu [21].

Now, we make the definition that a semi-metric d is a metric, but such that $d(x, y) = 0 \nRightarrow x = y$, where $x, y \in E$. This characteristic is suitable when dealing with infinite dimensional space, since, in such spaces, there is no equivalence between norms. Many semi-metrics are available in the literature, each one with its strength and weakness. (Here, we focus on the PCA semimetric. For other types of semimetrics, the reader can refer to [21]) As pointed in [21], the PCA semi-metric is adequate for computing proximities between curves in a reduced dimensional space and it has the advantage of being usable even if the curves are rough. On the other hand, it only applies to balanced data. Technically, as stated in [21] and in Ferraty et al. [24], the PCA semi-metric may be built when assuming that $E \int \mathbf{X}^2(s) ds < \infty$. Under this assumption, the following expansion holds

$$\mathbf{X} = \sum_{k=1}^{\infty} \left(\int \mathbf{X}(s) v_k(s) ds \right) v_k, \quad (7)$$

where v_1, v_2, \dots are orthonormal eigenfunctions of the covariance operator $\Gamma_{\mathbf{X}}(t, s) = E(\mathbf{X}(t)\mathbf{X}(s))$. Regarding its empirical counterpart, let

$$\tilde{\mathbf{X}}^{(q)} = \sum_{k=1}^q \left(\int \mathbf{X}(s) v_k(s) ds \right) v_k, \quad (8)$$

be a truncated version of \mathbf{X} (Equation (7)). Based on the L^2 -norm, we have, for all $(X_1, X_2) \in E^2$, the following parametrized family of semi-metrics

$$d_q^{PCA}(X_1, X_2) = \sqrt{\sum_{k=1}^q \left(\int (X_1(s) - X_2(s)) v_k(s) ds \right)^2}. \quad (9)$$

It is worth noting that q is a tuning parameter that needs to be estimated. One possibility is choosing q via cross-validation. In the next paragraph, we detail how to estimate the non-observable quantities that are presented in Equation (9).

Estimation Details

In practice, it is only possible to observe a discretized version of $X = \{x(\tau); \tau \in \mathcal{D}\}$, which we denote by $X = \{x(\tau); \tau = 1, 2, \dots, D\}$, implying that $\mathcal{D} = \{1, 2, \dots, D\}$. Specific to our problem, a particular month is viewed as a function (curve) that links days to cumulative excess returns. Therefore, we observe a sequence $\{X_1, X_2, \dots, X_T\}$ of realizations of \mathbf{X} , where $X_t = \{x_t(\tau); \tau = 1, 2, \dots, D\}$ corresponds to the observed excess return curve at months $t = 1, \dots, T$, and D represents the number of days each month. It is worth noting that balanced data are needed in order to use the PCA semimetric. In Section 3, we detail how to adjust the data to make D constant over t .

The main goal of this article is to forecast, one step ahead, the excess return on the last day of a given month, when considering the information available before the forecast is made. In other words, given a

sequence $\{X_1, X_2, \dots, X_T\}$ of realizations of \mathbf{X} , we want to estimate

$$r_{NP}(X_T) = E(\mathbf{y}_{T+1} | \mathbf{X} = X_T), \quad (10)$$

where $\mathbf{y}_t := \mathbf{x}_t(D)$, which means that the regressor is the excess return curve of a particular month and the regressand is the last day excess return of the next month. Now, Equation (3) may be adapted to our problem in the following way

$$\hat{r}_{NP}(X_T) = \frac{\sum_{t=2}^T y_t K(b^{-1}d(X_T, X_{t-1}))}{\sum_{t=2}^T K(b^{-1}d(X_T, X_{t-1}))}, \quad (11)$$

where K, b and d are defined as before. Roughly speaking, Equation (11) states that the one step ahead forecast for the excess return on the last day of the month is a weighted mean, in which the weights depend on the proximity, in terms of the semi-metric d , between the curve X_T and the other curves in the sample.

Now, we turn attention to the PCA semimetric. Because Γ_X is not observable, its empirical version is set to be $\Gamma_X^n(t, s) = n^{-1} \sum_{i=1}^n X_i(t)X_i(s)$. Moreover, the integral in (9) can be approximated by numerical quadrature. Following [21], the empirical version of the PCA semi-metric is defined for all $(X_1, X_2) \in E$, as

$$d_q(X_1, X_2) = \sqrt{\sum_{k=1}^q \left(\sum_{\tau=1}^D w_j(X_1(\tau) - X_2(\tau))[v_k]_\tau \right)^2}, \quad (12)$$

where v_1, v_2, \dots , are the W -orthonormal eigenvectors of the covariance matrix

$$\Gamma^T W = T^{-1} \left(\sum_{t=1}^T X_{1,t} X_{2,t}^T \right) W,$$

where $W = \text{diag}(w_1, w_2, \dots, w_D)$ is the diagonal matrix with the quadrature weights in the main diagonal. In order to estimate q , we apply the cross-validation procedure described in (4) for $q = 1, 2, 3$ and choose the pair (q, b_{opt}) that produces the smallest value in (5).

2.2. Predictive Regressions

Following the literature on excess return predictability, we consider the standard multivariate regression model for predicting excess returns, which can be expressed as

$$r_{t+1} = \beta_0 + \beta_1 Z_t + \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim \mathcal{N}(0, \sigma_\varepsilon^2), \quad (13)$$

where r_{t+1} is the stock market index log excess return from period t to $t+1$, Z_t is a vector of predictive variables available at the end of period t . β_0 is the conditional average excess return, β_1 is the incremental expected excess return with respect to one unit change in predictor variable Z_t , ε_{t+1} is the regression residual that is assumed to follow a standard normal distribution, and n is the number of predictors. We then divide the total sample of T observations for r_t and Z_t into an in-sample period that comprises the first P observations and an out-of-sample period comprising the last q observations ($q = T - p$).

As a simple no-predictability benchmark, we use the naive prevailing-mean and variance model, as suggested by [3], which can be expressed as

$$r_{t+1} = \beta_0 + \varepsilon_{t+1}, \quad \varepsilon_{t+1} \sim \mathcal{N}(0, \sigma_\varepsilon^2), \quad (14)$$

namely, the constant mean and volatility model. Note that the prevailing-mean forecast ignores information in any predictor variable; it is simply the historical average excess return, which is calculated over all prior observations $\mathbb{E}\{r_{t+1}\} = \frac{1}{t} \sum_{s=1}^t r_s$. This is equivalent to restricting $\beta_1 = 0$ in (13).

3. Data and Results

In this section, we first describe the predictor variables that are used in the literature and the details of their data. Subsequently, we describe the methodology used to evaluate the excess return forecasts that were obtained from the NP-FDA and regression based models. Furthermore, we present the statistically based and then the economic value assessments of the forecasts based on the mean-variance optimizing investor with quadratic utility.

3.1. Data and Traditional Predictors

We examined the performance of the NP-FDA and the competing models that are discussed in Section 2 in forecasting excess returns from the DJIA and S&P 500 indices. To evaluate our finds to the numerous literature on excess market return predictability, we compare the predictive accuracy of NP-FDA to that of twelve predictor variables suggested by [2] (The data on both stock returns and used to construct the popular predictors are downloaded from the website of Amit Goyal: <http://www.hec.unil.ch/agoyal/>). Market excess returns are computed from the DJIA and S&P500 indexes (including dividends). In order to capture excess returns, a short rate (Treasury bill) is subtracted from the returns, as in Equation (1). Below, we provide a list of the predictors that constitute the set of variables used to predict the excess return in (13):

- Dividend-price ratio (DP): the difference between the log of a twelve-month moving sum of dividends paid on the S&P500 index and the log of stock prices.
- Dividend yield (DY): the difference between the log of a twelve-month moving sum of dividends and the log of lagged stock prices.
- Earning-price ratio (EP): the log of a twelve-month moving sum of earnings on the S&P 500 index minus the log of stock prices.
- Dividend-payout ratio (DE): DP minus EP.
- Excess stock return volatility (RVOL): the sum of the squared daily returns on the S&P 500 index.
- Book-to-market ratio (BM): book value at the end of the previous year divided by the end-of-month market value of the Dow-Jones Industrial Average index.
- Net equity expansion (NITS): the 12-month moving sum of net equity issues by NYSE-listed stocks divided by the total end-of-year market capitalization of NYSE stocks.
- Treasury bill rate (TBL): three-month Treasury bill interest rate from the secondary market rate.
- Long-term yield (LTY): the long-term government bond yield.
- Term Spread (TMS): LTY minus TBL.
- Default yield spread (DFY): Moody's BAA- minus AAA-rated corporate bond yields.
- Inflation (INFL): change in the Consumer Price Index (CPI) for all urban consumers (Following common practice in literature, we lag the inflation one extra month due to the delayed release of the CPI index).

Specifically, we estimate predictive regressions while using updated monthly excess return data for DJIA and S&P500 from [2] and Kenneth French's Data Library (Available at <http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/datalibrary.html>). For the main results we start the sample in 1926:04 to account for the lagged predictors when estimating the predictive regressions. After accounting for the lagged predictors, the available estimation sample covers from 1926:04 to 2019:12 (1116 observations).

In addition, in the out-of-sample analysis with NP-FDA, we consider a longest DIJA return series, from 1885:01 to 2020:06 (1620 observations).

3.2. Dataset Used in NP-FDA Estimation

The PCA semimetric requires a balanced dataset to be usable, as pointed out in Section 2.1. On the other hand, we have a sample of daily cumulative returns. Because the number of business days varies along the months, we need to make some adjustment in the data in order to have a balanced dataset. We propose organizing these data in a matrix, denoted by \mathbb{X} , in which each row represents a month while each column represents a day. We fixed the number of days to 20. The first column gets the data of the first business day of each month in the sample, likewise the last column is filled with the data of the last business day. The second and the penult column get the second and the penult business day data, respectively. We keep doing that until the matrix is fullfilled. For some months, there is more than 20 business days. In such cases, the remaining data is dropped. For months with less than 20 days, we apply linear interpolation in order to input data. The main justification to proceed this way is that the beginning and end of each month carry more information about market movements.

It is important to highlight how that matrix \mathbb{X} relates to Equation (11) in NP-FDA estimation process. Let X_1, X_2, \dots, X_T and Y_1, Y_2, \dots, Y_{20} be the rows and the columns of \mathbb{X} , respectively. Because we are interested in one-step-ahead forecast, the NP-FDA regressor is the matrix that formed by stacking X_1, X_2, \dots, X_{T-1} and the regressand is the last column of \mathbb{X} , i.e., Y_{20} , with its first element discarded.

3.3. Forecast Combination

In addition to individual forecasting models, we consider another approach to improve the excess return forecasts based on forecast combination. The motivation for doing this is in the methodological literature on forecasting, which shows that more accurate predictions can be obtained from a linear combination of two or more forecasts in relation to the use of just one forecast [25–27]. In addition, adaptive strategies to combine predictions can also alleviate the effect of structural breaks, model uncertainty, and incorrect model specification [26,28]. In particular, there is recent evidence that the combination of nested models can significantly improve the accuracy of the forecasts when compared to predictions obtained from single model specifications [29].

Assuming M different models, a combined forecast for one-step-ahead return is given by

$$\hat{r}_{t+1|t} = \sum_{m=1}^M w_{t+1|t,m} \hat{r}_{t+1|t,m},$$

where $w_{t+1|t,m}$ stands for the time $-t$ weight assigned to the m^{th} model, $\hat{r}_{t+1|t,m}$. The forecast combination strategies are mostly adaptive, which means that the forecasts included in $\mathcal{M} = \{\hat{r}_{t+1|t,m}, m = 1, 2, \dots, M\}$ and/or corresponding weights $w_{t+1|t,m}$ are selected in a sub-sample of observations, based on some criteria. In this article, the following four combination strategies are considered:

1. *Equally weighted forecasts or pooled* (POOL-AVG): this forecast combination method assigns equal weights to the forecasts of all individual models, i.e., $w_{t+1|t,m} = 1/M$ for $m = 1, \dots, M$. This approach is likely to work well if the forecasting errors of different models have similar variances and are highly correlated, as explained in [30]. Therefore, in many cases, this simple average of forecasts can work well against more sophisticated weighting schemes [26,29].

2. *Thick Modeling Approach with MSFE (POOL-DMSFE)*: the second scheme consists of selecting models by means of thick modeling approach. Following [31], the weight for model m is computed as:

$$w_{t+1|t,m} = \frac{\phi_{m,t}^{-1}}{\sum_{j=1}^M \phi_{m,t}^{-1}}, \quad \text{where} \quad \phi_{m,t} = \sum_{s=j}^{t-1} \theta^{t-1-s} (r_{s+1} - \hat{r}_{s+1,m})^2,$$

where θ is a discount factor. Thus, the DMSPE approach assigns greater weights at time $t + 1$ to individual predictive models with lower MSPE values over the holdout out-of-sample period, which ranges from $j + 1$ to t .

3. *Diffusion Index*: this scheme involves the estimation of factors that are subsequently used for forecasting. The idea here is to extract a small number of common factors (often called diffusion indexes) assumed to drive the dynamics associated with a large number of potential return predictors (see, e.g., [9]). The basic intuition behind this approach is to filter out the noise present in the individual predictors, as discussed in [1]. The resulting factorial structure is more parsimonious, thus generating a more reliable signal.
4. *Sum-of-the-Parts Method*: the fourth combination scheme is the sum-of-the-parts, which is in line with the ideas that are presented in [4]. The sum-of-the-parts scheme consists of decomposing the return index into three components: the dividend yield, the earnings growth rate, and the growth rate in the price-earnings ratio. Subsequently, each of these components is predicted separately. Ref. [4] show that their sum-of-the-parts forecast scheme significantly outperforms the historical average forecast.

3.4. Forecast Evaluation

Our out-of-sample procedure mimics the situation faced by real-time forecasters. Forecasts from NP-FDA and regression models are generated while only using information available at period t . To compute the results, we use a rolling window estimation of 360 monthly observations (30 years) (To save space, we do not present the results of the in-sample analysis. The results are available upon request). Therefore, our out-of-sample period for the forecast evaluation ranges from January 1956 to December 2019. Most of our results refer to the full sample and two subsamples, one that focuses on economic expansions and the other on recessions. These periods are defined according to the National Bureau of Economic Research (NBER) business cycle dating committee methodology. Thus, a recession is the period following the peak of economic activity until the trough. Firstly, we detect the statistical and economic predictability while using individual models. Subsequently, we use the combination schemes that are presented above.

We follow the literature and compute the root mean squared forecast error (RMSFE) in order to evaluate out-of-sample forecasts. Given a sample of P out-of-sample forecasts for 1-step-ahead forecast horizon, the (RMSFE) for model m is defined, as follows:

$$\text{RMSFE}_m = \sqrt{\frac{1}{P} \sum_{t=1}^P (\hat{r}_{t+1|t,m} - r_{t+1})^2}, \quad (15)$$

where r_{t+1} is the observed return at time $t + 1$, and $\hat{r}_{t+1|t,m}$ is the corresponding forecast made at time t .

In addition, we follow the literature when using the out-of-sample R^2 (R_{oos}^2) in order to evaluate the forecasting performance. The R_{oos}^2 compares unconditional forecasts for a one-month ahead of the prevailing-mean benchmark, $\hat{r}_{t+1|t,\text{bench}}$, to the conditional forecasts, $\hat{r}_{t+1|t,m}$, of an alternative model, and it

is defined, as follows:

$$R_{\text{OOS}}^2 = 1 - \frac{\text{MSE}_{\text{m}}}{\text{MSE}_{\text{bench}}} = 1 - \frac{\sum_{t=1}^P \left(r_{t+1} - \hat{r}_{t+1|t,\text{bench}} \right)^2 - \sum_{t=1}^P \left(r_{t+1} - \hat{r}_{t+1|t,m} \right)^2}{\sum_{t=1}^P \left(r_{t+1} - \hat{r}_{t+1|t,\text{bench}} \right)^2} \quad (16)$$

A positive R_{OOS}^2 means that the alternative model presents lower MSE than the benchmark model. We implement the test proposed by [32] in order to assess the statistical significance of the R_{OOS}^2 . In this test, the null hypothesis is that the benchmark is not outperformed by the competing model in terms of forecast MSE against the complementary alternative hypothesis. The statistic is calculated by first defining,

$$\hat{f}_{t+1} = \left(r_{t+1} - \hat{r}_{t+1|t,\text{bench}} \right)^2 - \left[\left(r_{t+1} - \hat{r}_{t+1|t,m} \right)^2 - \left(\hat{r}_{t+1|t,\text{bench}} - \hat{r}_{t+1|t,m} \right)^2 \right]. \quad (17)$$

After regressing $\{\hat{f}_{t+1}\}_{t=1}^P$ on a constant, the Clark and West statistic is nothing but the t -statistic of the constant. The p -value for the test may be obtained from the standard normal distribution.

Out-of-Sample Excess Returns Predictability Results

Tables 1 and 2 report the statistical measures of the out-of-sample forecasting performance from the NP-FDA, individual regression models, and four combination schemes for S&P500 and DJIA excess returns. All of the models are evaluated relative to the prevailing-mean benchmark. The out-of-sample forecasts are generated by rolling predictive regressions. The period 1926:12-1956:12 is considered as the initial in-sample estimation period. Hence, we compute out-of-sample forecasts for 1957:01-2019:12 (756 observations). In addition to the full forecast evaluation period, we present results that were computed separately during NBER-dated business-cycle expansions and recessions.

The first column of Tables 1 and 2 report the RMSE-value relative to the benchmark prevailing-mean model, i.e., $(\text{RMFSE}_{\text{model}} - \text{RMFSE}_{\text{bench}}) / \text{RMFSE}_{\text{bench}}$. Therefore, negative entries indicate that the candidate model outperforms the benchmark model in term of RMFSE, thus generating more accurate point forecasts, while positive values indicate the opposite. The evidence found is consistent with [2]: very few of the univariate models beat the historical average benchmark in terms of RMFSE. In fact, for S&P500, only NP-FDA and SVAR models produce lower RMSE-values than the benchmark model when we consider the entire out-of-sample period and individual models. For DJIA, seven predictors and the NP-FDA produce lower RMSE-values than the prevailing-mean model for entire out-of-sample period.

In Table 1 for the S&P500 excess returns forecast, none of the individual predictors generate significantly positive R_{OOS}^2 , except NP-FDA in the three sample periods. On the other hand, all of the combining methods exhibit positive R_{OOS}^2 (ranging from 0.21% to 0.48%) in the full sample period. Nevertheless, none combining approach generates significantly positive R_{OOS}^2 (significant at 10% level) in all considered sub-samples. The superiority of NP-FDA forecast is further confirmed when the NP-FDA approach methods outperform the historical average forecast in the full sample period, expansions, and recessions. Similar to Table 1, Table 2 (when we consider the DJIA excess returns) shows that all of the individual predictors also fail to deliver significantly positive R_{OOS}^2 , except TBL, LTY, and LTR in the full sample period. On the other hand, the NP-FDA method consistently generates significantly superior performance compared to the historical average premium forecast. In addition to NP-FDA, for the DJIA excess returns, all of the combination forecast strategies consistently generate performance that is significantly superior to the historical average.

Overall, we find that the NP-FDA exhibits a positive and statistically significant R^2_{OOS} for both indices. The result is particularly strong along the entire out-of-sample period, where the NP-FDA produces R^2_{OOS} of 0.37% and 0.57% for S&P500 and DJIA, respectively. Among the other predictors, we highlight the performance of TBL, LTY, and LTR. For both indices, we find that 7 of the 14 predictors result in negative R^2_{OOS} values, suggesting underperformance relative to the benchmark model. These results are consistent with the findings of [2] and more recent literature ([1,33,34], among others) that it is hard to find an individual variable that can significantly beat the historical average model. It is worth noting that predictive regressions typically have a very low R^2 . A monthly R^2 statistic of even 0.5 percent can yield an economically significant result (in terms of utility gain) in a return predictability study [3].

Table 1. Out of sample forecasting evaluation of S&P500 excess return based on nonparametric functional data analysis (NP-FDA) and regression models.

Predictor	Overall			Expansion			Recession		
	RRMSFE	R^2_{OOS}	<i>p</i> -Value	RRMSFE	R^2_{OOS}	<i>p</i> -Value	RRMSFE	R^2_{OOS}	<i>p</i> -Value
<i>Individual Models</i>									
NP-FDA	−0.262	0.374	0.023	−0.151	0.244	0.062	−0.580	0.690	0.002
log(DP)	0.244	0.106	0.101	0.334	0.169	0.025	−0.298	−0.046	0.556
log(DY)	0.366	0.144	0.081	0.690	0.203	0.021	−0.744	0.006	0.450
log(EP)	1.182	−0.535	0.567	1280	−0.121	0.297	0.617	−1.518	0.716
log(DE)	1.347	−1.779	0.636	0.469	−0.422	0.084	3.080	−5.003	0.988
SVAR	−0.351	0.104	0.096	−0.328	0.183	0.015	−0.736	−0.083	0.629
BM	0.746	−0.264	0.448	0.866	0.089	0.196	0.126	−1.103	0.665
NTIS	0.172	−0.287	0.243	−0.650	1.249	0.002	1.777	−3.937	0.973
TBL	0.152	0.270	0.018	0.243	0.522	0.001	−0.398	−0.328	0.865
LTY	0.718	0.275	0.004	1.105	0.455	0.001	−0.538	−0.153	0.738
LTR	0.207	0.213	0.124	0.233	0.381	0.094	−0.186	−0.184	0.445
TMS	0.169	−0.108	0.489	0.056	0.008	0.340	0.109	−0.382	0.873
DFY	0.234	−0.032	0.340	0.023	0.214	0.077	0.406	−0.618	0.953
DFR	0.080	−0.225	0.483	−0.041	0.227	0.121	0.038	−1.300	0.821
INFL	0.107	0.036	0.331	−0.050	0.270	0.003	0.152	−0.520	0.933
<i>Forecast Combination</i>									
POOL-AVG	−0.229	0.215	0.105	−0.547	0.559	0.002	−0.062	0.141	0.188
POOL-DMSFE	−0.283	0.221	0.085	−0.548	0.561	0.002	−0.070	−0.040	0.258
Diffusion index	−0.365	0.350	0.032	−0.485	0.434	0.025	−0.171	0.111	0.145
Sum-of-the-parts	−0.497	0.483	0.015	−0.781	1.025	0.005	0.174	−0.286	0.381

Note: This table reports out-of sample forecasting performance based on NP-FDA, univariate prediction regression, and combination schemes. RRMSFE refers to the percentage reduction in the Root Mean Square Forecast Error in relation to the prevailing-mean forecasts. A negative number means that the model with the predictor variables that are listed in the row has a lower RMSE in relation to the prevailing-mean model, while a positive number suggests the opposite. R^2_{OOS} stands for the out-of-sample R^2 defined by the percentage reduction of MSPE of the model of interest in relation to the benchmark model. *p*-value contains the *p*-values that are associated with testing $H_0 : R^2_{OOS} \leq 0$ against $H_A : R^2_{OOS} > 0$. All of the results are out-of-sample and they cover the entire forecast evaluation period (1957:01–2019:12). Separate results for NBER-dated expansions and recessions are provided. In order to calculate the results, we use a rolling window estimation of the most recent 30 years of observations.

The last four rows in each table shows the results for combination schemes. Tables 1 and 2 also reports statistics computed separately during NBER expansions and recessions, respectively. The most important finding is that NP-FDA performs well in expansions and recessions, exhibiting positive and significant statistically R^2_{OOS} for all sub-sample periods. In contrast, none of the 14 predictors perform well in both states of the economy for both indices. When we consider the other economic fundamentals, the results change considerably between periods of recession and expansion. In fact, most of the predictors produce better results in periods of expansion than in periods of recession.

Recent literature (e.g., [1,16,34]) shows that individual models have performed poorly in forecasting stock returns because of model uncertainty and instability. Ref. [16] finds that combining individual forecasts provides convincing predictability performance out-of-sample. Inspired by their paper, we examined whether forecast combinations while using NP-FDA and traditional predictor models can lead to better performance when compared to univariate counterparts.

We consider four combinations strategies, which differ depending on the weighting scheme, as discussed in Section 3.3 above. We find that for the entire out-of-sample period and for both indices the combination strategies under consideration produce R^2_{OOS} that range from 0.18% to 0.50%, and they are statistically significant at the 10% level. These results that were obtained from the forecast combination are consistent with the findings of ([7,16,34], among others). Interestingly, we find that the Sum-of-the-parts scheme results in higher R^2_{OOS} during expansion periods, but deteriorates in NBER recessions.

Although RRMSFE and R^2_{OOS} provide a statistical measure of excess return predictability, they do not take account of an investor's risk during the out-of-sample period. Interestingly, R^2_{OOS} values are typically small; nevertheless, a modest forecasting ability may yield substantial utility gains for risk-averse investors [3]. In the next section, we use a utility gain measurement to assess economic significance of return predictability.

Table 2. Out of sample forecasting evaluation of Dow Jones Industrial Average (DJIA) excess return based on NP-FDA and regression models.

Predictor	Full Sample			Expansion			Recession		
	RRMSFE	R^2_{OOS}	p -Value	RRMSFE	R^2_{OOS}	p -Value	RRMSFE	R^2_{OOS}	p -Value
<i>Individual Models</i>									
NP-FDA	−0.287	0.574	0.010	−0.529	1.056	0.001	0.404	0.280	0.089
log(DP)	−0.002	0.005	0.382	−0.0640	0.128	0.173	0.168	0.248	0.101
log(DY)	−0.058	0.116	0.142	−0.1306	0.262	0.022	0.143	1.110	0.008
log(EP)	0.277	−0.555	0.666	0.1092	−0.219	0.433	0.737	−0.617	0.414
log(DE)	0.805	−1.618	0.684	0.2395	−0.48	0.143	2.347	−5.453	0.901
SVAR	−0.066	0.131	0.117	−0.141	0.283	0.015	0.143	1.365	0.119
BM	0.156	−0.313	0.470	0.008	−0.016	0.271	0.564	−0.116	0.371
NTIS	0.142	−0.286	0.141	−0.771	1.537	0.000	2.616	−4.782	0.950
TBL	−0.147	0.295	0.028	−0.305	0.610	0.000	0.287	0.976	0.121
LTY	−0.147	0.294	0.014	−0.263	0.525	0.000	0.171	0.736	0.099
LTR	−0.141	0.282	0.089	−0.272	0.543	0.054	0.219	0.863	0.117
TMS	0.072	−0.145	0.450	−0.014	0.028	0.293	0.308	0.286	0.266
DFY	0.015	−0.031	0.326	−0.109	0.219	0.087	0.357	−0.836	0.825
DFR	0.090	−0.182	0.519	−0.172	0.344	0.038	0.810	−0.586	0.533
INFL	−0.041	0.082	0.204	−0.114	0.228	0.039	0.160	−0.208	0.960
<i>Forecast Combination</i>									
POOL-AVG	−0.090	0.181	0.065	−0.262	0.524	0.000	−0.613	1.223	0.007
POOL-DMSFE	−0.093	0.187	0.058	−0.261	0.523	0.000	−0.657	1.311	0.010
Diffusion index	−0.093	0.187	0.060	−0.270	0.542	0.000	−1.884	3.734	0.000
Sum of the parts	−0.256	0.513	0.007	−0.797	1.590	0.000	1.968	−3.976	0.999

Note: this table reports out-of sample forecasting performance based on NP-FDA, univariate prediction regression, and combination schemes. RRMSFE refers to the percentage reduction in the Root Mean Square Forecast Error in relation to the prevailing-mean forecasts. A negative number means that the model with the predictor variables listed in the row has a lower RMSE in relation to the prevailing-mean model, while a positive number suggests the opposite. R^2_{OOS} stands for the out-of-sample R^2 that is defined by the percentage reduction of MSPE of the model of interest in relation to the benchmark model. p -value contains the p -values associated with testing $H_0 : R^2_{OOS} \leq 0$ against $H_A : R^2_{OOS} > 0$. All results are out-of-sample and cover the entire forecast evaluation period (1957:01–2019:12). Separate results for NBER-dated expansions and recessions are provided. In order to calculate the results, we use a rolling window estimation of the most recent 30 years of observations.

3.5. Economic-Based Forecast Evaluation

Although our analysis is focused on statistical measures of predictive accuracy, it is important to evaluate the extent to which apparent gains in predictive accuracy can be used in real time to improve economic utility of the investor, that is, translate into better investment performance. Given that statistical significance does not necessarily imply economic significance ([1,5,33,34], among others), we assessed the economic value of the predictive power of stock returns by investigating the utility gains for investors who exploit the predictability of excess returns over an alternative without predictability associated with the prevailing average model. The motivation here is that investors are primarily concerned with the performance of return forecasts in terms of asset allocation.

We consider a mean-variance investor with quadratic utility who split her portfolio between stock market indices and a risk-free rate [1]. At the end of t , she chooses to invest, during $t + 1$, the following share of her portfolio to stocks:

$$w_{i,t+1} = \left(\frac{1}{\gamma} \right) \left(\frac{\hat{r}_{i,t+h}}{\hat{\sigma}_{i,t+h}^2} \right) \quad (18)$$

where γ is the relative risk aversion γ , and $\hat{\sigma}_i^2$ is a forecast of the variance of the index returns.

Similar to [16], we simply assume the variance to be a 10-year rolling window of quarterly returns. Over the forecast evaluation period, the investor realizes the average utility,

$$\hat{v}_m = \hat{\mu}_m - 0.5\gamma\hat{\sigma}_m^2, \quad (19)$$

where $\hat{\mu}_m$ and $\hat{\sigma}_m^2$ are, respectively, the out-of-sample mean and variance of the return on the dynamic portfolio formed on the basis of \hat{r}_{t+h} and $\hat{\sigma}_m^2$ over the forecast evaluation period for each model m . The quantity that is defined in Equation (19) may be calculated for the prevailing-mean model as well (\hat{v}_{bench}). The difference $\hat{v}_m - \hat{v}_{bench}$ represents the utility gain of using the competitor model m to forecast the excess returns in place of the prevailing-mean forecast in the asset allocation decision. This utility gain is the certainty equivalent return, which may be viewed as the quantity an investor would be willing to pay to have access to the return forecasts generated using the model m instead of using the information in the historical average.

We compute the utility for the same investor who uses the historical average excess return forecast. We measure the utility gain (or certainty equivalent return) as the extra utility that is generated from Equation (19) relative to this benchmark. This difference is multiplied by 1200 to express the utility gain in average annualized percentage return. Table 3 reports certainty equivalent (in annualized percent return) for a mean-variance investor with $\gamma = 5$ who allocates among stocks and risk-free bills while using forecasts that are based on competitors models in place of historical average forecasts.

The results indicate that the NP-FDA is the only one that provides economic value relative to the prevailing-mean benchmark for all considered datasets and for both market indices. Specifically, when we consider the entire out-of-sample period the NP-FDA approach provides annualized gain above 0.7% and 0.3% for S&P500 and DJIA, respectively, meaning, in these cases, that the investor would be willing to pay more than 0.7% and 0.3% to have access to the information in the NP-FDA compared to the historical average forecasts, respectively. The certainty equivalent gains of the popular predictor models imply evidence of moderate economic performance relative to the prevailing-mean model. Indeed, only 5 out of the 14 univariate models produces positive Δ for S&P500 excess returns, indicating that the predictive regression forecast has a higher certainty equivalent than the historical average. The performance for the DJIA index is even worse. Overall, the average utility gains provided here supports NP-FDA approach to forecast stock indices excess returns. It is important to emphasize the need to supplement standard statistical criteria with more direct economic-based measures when analyzing out-of-sample bond return predictability.

Table 3. Performance evaluation for an Investor with Mean Variance Utility with $\gamma = 5$, 1956:01-2019:12.

Predictor	S&P500			DJIA		
	Full Sample	Expansion	Recession	Full Sample	Expansion	Recession
<i>Individual Models</i>						
NP-FDA	0.983	0.691	2.676	0.277	0.390	0.087
log(DP)	−0.005	−0.025	0.113	−0.205	−0.223	−0.096
log(DY)	0.052	0.024	0.216	−0.020	−0.023	0.000
log(EP)	−1031	−0.164	−6.172	−0.866	−0.34	−3.989
log(DE)	−3.132	−0.617	−1.773	−2.851	−0.842	−1.458
SVAR	0.000	0.000	0.000	0.000	0.000	0.000
BM	−0.666	−0.019	−4.504	−0.624	−0.189	−3.214
NTIS	−1.303	1.156	−1.566	−0.860	1.482	−1.452
TBL	0.137	0.409	−1.471	0.162	0.452	−1.556
LTY	0.239	0.352	−0.435	0.273	0.363	−0.265
LTR	0.066	0.271	−1154	0.045	0.272	−1.308
TMS	−0.446	−0.237	−1.690	−0.565	−0.350	−1.839
DFY	−0.205	0.027	−1.582	−0.309	−0.147	−1.274
DFR	−0.412	0.022	−2.983	−0.490	0.041	−3.634
INFL	−0.107	0.150	−1.633	−0.092	−0.089	−0.111
<i>Forecast Combination</i>						
POOL-AVG	2.162	2.853	1.388	0.044	0.034	−0.756
POOL-DMSFE	2.269	2.536	1.516	0.073	0.034	−0.571
Diffusion index	2.360	2.360	3.921	0.104	0.054	0.050
Sum-of-the-parts	1.947	3.149	−0.152	−0.012	1.608	−1.282

Note: This table presents the average utility gain (Δ) in the portfolio management fee (in annualized percentage return) that an investor would be willing to pay to have access to the NP-FDA forecasts and the predictive regression forecast that is based on the economic variables given in the first column relative to the prevailing-mean benchmark forecast. All results are out-of-sample and were calculated using a rolling window estimation of the most recent 30 years of observations (1957:01-2019:12).

As discussed earlier, forecast combinations improve the forecast results by weighting the individual regression forecasts resulting in a lower variance relative to the individual forecasted returns. The utility gains in the certainty equivalent, Δ (annual %) are presented in the last four rows of Table 3 for S&P500 and DJIA. The utility gains provide economic evidence for stock return predictability of combination forecast models. Specifically, we find consistent evidence that the incorporation of NP-FDA improves the forecast combination performance, from both the statistical and economic perspectives, mainly for S&P500 excess returns. The best performance is found for the diffusion index scheme, in which the utility gains are greater than 230 basis points for all of the considered time periods. This finding on the return predictability based on forecast combination is consistent with the result in ([7,16,34], among others). Interestingly, for the S&P500 index, the simple equally weighted forecast combination (POOL-AVG) delivers an Δ (annual %) of 2.62%, which is considerably higher than the Δ (annual %) of the better individual model.

Our results show that the performance of some of the predictor models, according to R^2_{OOS} and utility gain metrics, is quite similar. For example, in the S&P500 index, the NP-FDA approach has an R^2_{OOS} of 0.37 percent and a utility gain of 0.98 percent. The prior literature suggests a lack of association between utility gain and R^2_{OOS} [1] and shows that different performance metrics can lead researchers to draw different conclusions regarding the forecasting ability of the predictor variables.

3.6. Robustness Check

In addition to evaluate the forecast performance of the NP-FDA approach in both NBER expansions and recessions, we extend our analysis to a longer DJIA dataset, from 1985:01-2020:06. As in the previous analysis, we use a 30-year rolling window of the data, generating 1252 out-of-sample forecasts for the period 1915:01-2020:06. Table 4 reports the same statistics that were considered before. In this case, we compare the NP-FDA forecasts with those of the prevailing-mean model. Once again, we report results for the full sample as well as NBER-dated expansions and recessions. In the expansion periods, the NP-FDA exhibit positive and statistically significant R^2_{os} of 0.48%. However, the performance deteriorates in recession periods. For the full sample period NP-FDA forecasts beat the historical mean, statistically and economically. More importantly, taking into account that investors are primarily concerned with the performance of return forecasts in terms of asset allocation, the NP-FDA forecasts overcome the prevailing-mean model in economic terms at all time periods considered. In summary, this result corroborates previous findings that NP-FDA presents good performance when used to forecast excess stock returns.

Table 4. Performance evaluation for an Investor with Mean Variance Utility, 1956:01-2019:12.

	RRMSFE	R^2_{os}	<i>p</i> -Value	Δ (annual %)
Full-Sample	−0.081	0.163	0.067	0.141
Expansion	−0.238	0.477	0.008	0.471
Recession	0.070	−0.142	0.715	0.051

4. Conclusions

This study provides a comprehensive examination of out-of-sample predictability of excess returns for both the S&P500 and DJIA indices while using NP-FDA approach, a variety of individual regression models as well as combination forecast models. Similar to the previous literature, our results reveal that it is difficult to find an individual predictor that consistently outperforms the historical average benchmark and provides significant out-of-sample forecasts of the equity premium. Nevertheless, the NP-FDA tends to generate significantly positive R^2_{os} for both indices and different out-of-sample periods.

Assuming that daily cumulative returns can be interpreted as curves describing the cumulative return during a given month, we propose predicting the S&P500 and the DJIA excess returns out of sample based on the NP-FDA estimator. This estimator may be viewed as a weighted mean of monthly stock excess returns, in which the weights are defined based on a measure of proximity between those curves. In addition, we propose combining the forecasts that are generated by the NP-FDA and the traditional predictor variables. We find that NP-FDA produces excess returns forecasts that are statistically and economically significantly superior to the benchmark prevailing-mean model, over our out-of-sample period, 1957:01-2019:12. Indeed, our results show that NP-FDA is the only one individual model that can overcome the historical average forecasts for excess returns in statistically and economically significant manners for both S&P500 and DJIA during the entire period, NBER recession, and expansions periods. Furthermore, the NP-FDA provides additional relevant information for predicting stock returns in relation to popular predictor variables to a large extent.

Our results show that adding NP-FDA forecasts to the traditional univariate predictive regressions with a popular predictors significantly improves the forecasting combination performance with different schemes in terms of conventional statistical measures of forecast performance and economic gains term. Economic gains are of particular importance from the perspective of investors, and our NP-FDA approach is of tremendous value in this regard, when compared to the methods adopted in the existing literature so far.

As part of future analysis, it would be interesting to extend our analysis to other developed and emerging market economies in order to confirm the superiority of the NP-FDA approach in a robust manner.

Author Contributions: Conceptualization, J.F.C. and R.G.; methodology, J.F.C. and H.S.T.; software, J.F.C. and H.S.T.; validation, J.F.C., R.G. and H.S.T.; formal analysis, J.F.C.; investigation, J.F.C., R.G. and H.S.T.; resources, J.F.C., R.G. and H.S.T.; data curation, J.F.C. and R.G.; writing—original draft preparation, J.F.C. and H.S.T.; writing—review and editing, R.G. and H.S.T.; visualization, J.F.C., R.G. and H.S.T.; supervision, H.S.T.; project administration, R.G.; funding acquisition, J.F.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: We would like to thank the three anonymous referees for many helpful comments. However, any remaining errors are solely ours. João F. Caldeira gratefully acknowledges support provided by CNPq under Grants 430192/2016-9 and 306886/2018-9. The author Hudson Torrent wishes to thank MCTIC/CNPq (process number 438642/2018-0).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rapach, D.; Zhou, G. Forecasting Stock Returns. In *Handbook of Economic Forecasting*; Elliott, G., Granger, C., Timmermann, A., Eds.; Elsevier: Amsterdam, The Netherlands, 2013; Volume 2, Chapter 0, pp. 328–383. [\[CrossRef\]](#)
2. Welch, I.; Goyal, A. A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. *Rev. Financ. Stud.* **2008**, *21*, 1455–1508. [\[CrossRef\]](#)
3. Campbell, J.; Thompson, S.B. Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? *Rev. Financ. Stud.* **2008**, *21*, 1509–1531. [\[CrossRef\]](#)
4. Ferreira, M.; Santa-Clara, P. Forecasting stock market returns: The sum of the parts is more than the whole. *J. Financ. Econ.* **2011**, *100*, 514–537. [\[CrossRef\]](#)
5. Pettenuzzo, D.; Timmermann, A.; Valkanov, R. Forecasting stock returns under economic constraints. *J. Financ. Econ.* **2014**, *114*, 517–553. [\[CrossRef\]](#)
6. Pan, Z.; Pettenuzzo, D.; Wang, Y. Forecasting stock returns: A predictor-constrained approach. *J. Empir. Financ.* **2020**, *55*, 200–217. [\[CrossRef\]](#)
7. Tsiakas, I.; Li, J.; Zhang, H. Equity premium prediction and the state of the economy. *J. Empir. Financ.* **2020**, *58*, 75–95. [\[CrossRef\]](#)
8. Brown, G.W.; Cliff, M.T. Investor sentiment and the near-term stock market. *J. Empir. Financ.* **2004**, *11*, 1–27. [\[CrossRef\]](#)
9. Kelly, B.; Pruitt, S. Market Expectations in the Cross-Section of Present Values. *J. Financ.* **2013**, *68*, 1721–1756. [\[CrossRef\]](#)
10. Neely, C.J.; Rapach, D.E.; Tu, J.; Zhou, G. Forecasting the Equity Risk Premium: The Role of Technical Indicators. *Manag. Sci.* **2014**, *60*, 1772–1791. [\[CrossRef\]](#)
11. Mascio, D.A.; Fabozzi, F.J. Sentiment indices and their forecasting ability. *J. Forecast.* **2019**, *38*, 257–276. [\[CrossRef\]](#)
12. Henkel, S.; Martin, J.S.; Nardari, F. Time-varying short-horizon predictability. *J. Financ. Econ.* **2011**, *99*, 560–580. [\[CrossRef\]](#)
13. Ang, A.; Timmermann, A. Regime Changes and Financial Markets. *Annu. Rev. Financ. Econ.* **2012**, *4*, 313–337. [\[CrossRef\]](#)
14. Dangl, T.; Halling, M. Predictive regressions with time-varying coefficients. *J. Financ. Econ.* **2012**, *106*, 157–181. [\[CrossRef\]](#)
15. Chang, K.L. Does the return-state-varying relationship between risk and return matter in modeling the time series process of stock return? *Int. Rev. Econ. Financ.* **2016**, *42*, 72–87. [\[CrossRef\]](#)
16. Rapach, D.E.; Strauss, J.; Zhou, G. Out-of-Sample Equity Premium Prediction: Combination Forecasts and Links to the Real Economy. *Rev. Financ. Stud.* **2010**, *23*, 821–862. [\[CrossRef\]](#)

17. Lima, L.R.; Meng, F. Out-of-Sample Return Predictability: A Quantile Combination Approach. *J. Appl. Econom.* **2017**, *32*, 877–895. [[CrossRef](#)]
18. Mascio, D.A.; Fabozzi, F.J.; Zumwalt, J.K. Market timing using combined forecasts and machine learning. *J. Forecast.* **2020**, [[CrossRef](#)]
19. Baker, M.; Wurgler, J.; Yuan, Y. Global, Local, and Contagious Investor Sentiment. *J. Financ. Econ.* **2012**, *104*, 272–287. [[CrossRef](#)]
20. Huang, D.; Jiang, F.; Tu, J.; Zhou, G. Investor Sentiment Aligned: A Powerful Predictor of Stock Returns. *Rev. Financ. Stud.* **2015**, *28*, 791–837. [[CrossRef](#)]
21. Ferraty, F.; Vieu, P. *Nonparametric Functional Data Analysis: Theory and Practice*, 1st ed.; Springer: New York, NY, USA, 2006.
22. Ramsay, J.O.; Silverman, B.W. *Functional Data Analysis*, 2nd ed.; Springer: New York, NY, USA, 2005.
23. Caldeira, J.; Torrent, H. Forecasting the US Term Structure of Interest Rates Using Nonparametric Functional Data Analysis. *J. Forecast.* **2017**, *36*, 56–73. [[CrossRef](#)]
24. Ferraty, F.; Rabhi, A.; Vieu, P. Conditional Quantiles for Dependent Functional Data with Application to the Climatic “El Niño” Phenomenon. *Sankhyā Indian J. Stat.* **2005**, *67*, 378–398.
25. Granger, C.W. Combining Forecasts—Twenty Years Later. *J. Forecast.* **1989**, *8*, 167–173. [[CrossRef](#)]
26. Newbold, P.; Harvey, I.H. Forecast combination and encompassing. In *A Companion to Economic Forecasting*; Clements, M.P., Hendry, D.F., Eds.; Wiley-Blackwell: Hoboken, NJ, USA, 2002.
27. Aiolfi, M.; Timmermann, A. Persistence in Forecasting Performance and Conditional Combination Strategies. *J. Econom.* **2006**, *135*, 31–53. [[CrossRef](#)]
28. Pesaran, M.H.; Timmermann, A. Selection of estimation window in the presence of breaks. *J. Econom.* **2007**, *137*, 134–161. [[CrossRef](#)]
29. Clark, T.E.; McCracken, M.W. Improving Forecast Accuracy By Combining Recursive And Rolling Forecasts. *Int. Econ. Rev.* **2009**, *50*, 363–395. [[CrossRef](#)]
30. Timmermann, A. Forecast combinations. *Handb. Econ. Forecast.* **2006**, *1*, 135–196.
31. Stock, J.H.; Watson, M.W. Combination forecasts of output growth in a seven-country data set. *J. Forecast.* **2004**, *23*, 405–430. [[CrossRef](#)]
32. Clark, T.E.; West, K.D. Approximately normal tests for equal predictive accuracy in nested models. *J. Econom.* **2007**, *138*, 291–311. [[CrossRef](#)]
33. Cenesizoglu, T.; Timmermann, A. Do return prediction models add economic value? *J. Bank. Financ.* **2012**, *36*, 2974–2987. [[CrossRef](#)]
34. Wang, Y.; Pan, Z.; Wu, C.; Wu, W. Industry equi-correlation: A powerful predictor of stock returns. *J. Empir. Financ.* **2020**, *59*, 1–24. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).