econometrics

MDPI

Article

Forecasting Asset Returns Using Nelson–Siegel Factors Estimated from the US Yield Curve

Massimo Guidolin * and **Serena Ionta**

Department of Finance, Bocconi University, 20136 Milan, Italy; serena.ionta2@unibocconi.it
* Correspondence: massimo.guidolin@unibocconi.it

Abstract: This paper explores the hypothesis that the returns of asset classes can be predicted using common, systematic risk factors represented by the level, slope, and curvature of the US interest rate term structure. These are extracted using the Nelson–Siegel model, which effectively captures the three dimensions of the yield curve. To forecast the factors, we applied autoregressive (AR) and vector autoregressive (VAR) models. Using their forecasts, we predict the returns of government and corporate bonds, equities, REITs, and commodity futures. Our predictions were compared against two benchmarks: the historical mean, and an AR(1) model based on past returns. We employed the Diebold–Mariano test and the Model Confidence Set procedure to assess the comparative forecast accuracy. We found that Nelson–Siegel factors had significant predictive power for one-month-ahead returns of bonds, equities, and REITs, but not for commodity futures. However, for 6-month and 12-month-ahead forecasts, neither the AR(1) nor VAR(1) models based on Nelson–Siegel factors outperformed the benchmarks. These results suggest that the Nelson–Siegel factors affect the aggregate stochastic discount factor for pricing all assets traded in the US economy.

Keywords: Nelson–Siegel model; forecasting; asset return prediction; yield curve dynamics; systematic risk factors; Diebold–Mariano test; model confidence set

JEL Classification: G10; G12; C53; E43; E44

Academic Editor: Julien Chevallier

Received: 1 February 2025
Revised: 30 March 2025
Accepted: 8 April 2025
Published: 11 April 2025

Citation: Guidolin, M., & Ionta, S. (2025). Forecasting Asset Returns Using Nelson–Siegel Factors Estimated from the US Yield Curve. *Econometrics*, *13*(2), 17. https://doi.org/10.3390/econometrics13020017

1. Introduction

Modeling and forecasting the term structure of interest rates is crucial for policymakers and market participants, as the yield curve reflects market expectations about future monetary policy, output growth, and inflation across a range of time horizons. Additionally, used as a risk-free benchmark, the AA(A)-rated government bond yield curve often represents the floor for borrowing costs in the economy. This paper aims to extend the literature by examining the informational content of the US yield curve beyond the US government bond market. Specifically, we test the out-of-sample explanatory power of the yield curve's shape—captured through the classical Nelson and Siegel (1987) level, slope, and curvature factors—in forecasting the returns of a range of financial assets besides bonds that, and riskless and not, encompass the equity, real estate, and commodity domains. We accomplish this task by deploying the insights derived from both simple, pairwise Diebold–Mariano (one-sided) tests of equal predictive accuracy and from more systematic model confidence set tests.

There are two primary approaches to fitting the default risk-free yield curve: non-parametric spline-based methods and parametric models. Spline-based approaches excel at providing an accurate in-sample fit, due to their flexibility; however, parsimonious parametric models typically offer superior forecasting performance, achieving an effective balance between model simplicity and predictive accuracy (Nymand-Andersen, 2018). Given that the primary focus of this paper is to test the out-of-sample predictive power of the yield curve's information content, rather than merely fitting the curve in-sample, we adopt a parsimonious parametric approach. Specifically, we employ the Nelson–Siegel (NS) model, which offers simplicity and interpretability by requiring the estimation of only three (or four, see below for a discussion) parameters to characterize the entire yield curve. According to the specification in Diebold and Li (2006), these parameters can be interpreted as capturing the level, slope, and curvature of the yield curve.

Using data on the US yield curve from January 1990 to December 2020, we assess the forecasting performance of several models during the pseudo out-of-sample period January 2011–December 2020. Our analysis focuses on AR(1) and VAR(1) models applied to forecast the NS model factors at a certain horizon h. Such predicted factors are then used to predict asset returns at the same horizon h. To forecast asset returns, we implement a two-step approach. The predictive accuracy of these models is evaluated against benchmarks, including a historical average and an AR(1) model applied to asset returns. Additionally, we investigate whether the VAR(1) model applied to the NS outputs, and which accounts for interactions among the NS factors, can provide enhanced forecasting performance.

Contrary to Diebold and Li (2006) and the literature on forecasting yields using NS factors that has followed (see, e.g., Christensen et al. (2011); Favero et al. (2012); Fernandes and Vieira (2019); Guidolin and Pedio (2019); Toczydlowska and Peters (2018); Xiang and Zhu (2013); Yu and Zivot (2011) and the references in Diebold and Rudebusch (2013); Duffee (2013)) on more recent data, we find that NS factors do not always consistently outperform a set of benchmarks in the prediction of interest rates, even though the (dynamic) NS approach remains a competitive one. We then investigate the ability of these factors to explain returns across various asset classes, such as equities, corporate bonds, commodities, and REITs. This exploration is motivated by the potential economic significance of the Nelson–Siegel factors, which reflect bond investors' expectations regarding the economic outlook and may reflect information about macroeconomic conditions. Given their ability to capture such crucial economic insights, these factors may also serve as systematic drivers of the stochastic discount factor (henceforth, SDF) that prices all financial assets.[1] Because of these properties, the NS factors may themselves represent predictors of the SDF, particularly in the context of time-varying risk premia influenced by macroeconomic shifts and investor behavior throughout business cycles (Drobetz et al., 2002). In other words, while a naive null hypothesis would imply that the dynamic factors extracted from the yields (i.e., ex ante returns over a holding period) in one market should not help forecast yields or ex post, realized returns in other asset markets, the integration of NS factors into asset pricing models may help reconcile observed return patterns with macroeconomic fundamentals, refining our understanding of the stochastic discount factor and creating cross-asset common predictability linkages that we explore and document in this paper.

Our key empirical result suggests that the Nelson–Siegel factors exhibit a predictive power extending beyond the government bond market, though their performance varies significantly across forecast horizons and asset classes. Notably, these factors outperformed standard benchmarks in forecasting equity and REIT returns, while their predictive ability for commodity futures was more limited. As expected, direct regression on the Nelson–Siegel factors proved effective for forecasting Treasury securities returns. However, the comparison between VAR(1) and AR(1) models for the Nelson–Siegel factors yielded mixed results, with their relative performance depending on the specific forecast horizon and asset class considered.

Although we are not aware of any similar paper in the literature, in some respects our empirical goal is related to Hillebrand et al. (2018), who interpreted the yield curve's NS factors as special, economically motivated substitutes of standard principal components that efficiently combine the information contained in the raw data, the yield curve. However, the standard NS factors were not "supervised" for a specific forecast target, in that they were constructed using only the predictors and not using a particular forecast target. Therefore, they proposed computing NS factors not of the original data, but of the many alternative forecasts that a range of models may offer, with each of the many forecasts being computed using one predictor at a time. These combined forecasts are shown to outperform classical benchmarks in the out-of-sample (henceforth, OOS) forecasting of monthly US output growth and inflation, especially at longer forecast horizons.

The rest of this paper is organized as follows: Section 2 describes the methodology, Section 3 outlines the data sources, Section 4 presents the key empirical results, and Section 5 concludes.

2. Methodology

This section presents the methodological framework informing the analysis. Section 2.1 introduces the parsimonious parametric model used to fit the US yield curve. Section 2.2 describes the models employed to forecast the yield curve. In Section 2.3, the focus shifts to the models used to predict the financial returns of non-government bond assets. Lastly, Section 2.4 briefly overviews the two statistical tests applied to evaluate the differential forecasting accuracy in this paper.

2.1. Fitting the Nelson–Siegel Model to the US Yield Curve

Our analysis begins by fitting a standard NS model—a parsimonious parametric framework—to the U.S. yield curve. The goal is to extract the level, slope, and curvature factors that capture the key dynamics of the term structure. Using data spanning a January 1990–December 2020 sample, these factors were computed to test whether they contain useful information for forecasting the returns of other assets, hence establishing a potential connection to the SDF(s). The functional form of the NS model for the yield y at time t, with maturity m is

$$y_t(m) = \beta_{0,t} + \beta_{1,t}\left(\frac{1-e^{-\lambda m}}{\lambda m}\right) + \beta_{2,t}\left(\frac{1-e^{-\lambda m}}{\lambda m} - e^{-\lambda m}\right) \tag{1}$$

To mitigate multicollinearity and facilitate an intuitive interpretation of the factors, we adopt the following specification (Diebold & Li, 2006):

$$y_t(m) = \beta_{0,t} + \beta_{1,t}\left(\frac{1-e^{-\lambda m}}{\lambda m}\right) + \beta_{2,t}e^{-\lambda m} \tag{2}$$

The two specifications are identical for $\lambda = 0$, when $\beta_{0,t} = \beta_{1,t}$ and $\beta_{2,t} = 0$ at a given time t. The three potentially time-varying coefficients $\beta_{0,t}$, $\beta_{1,t}$, and $\beta_{2,t}$ determine the behavior of the yield curve over the long, short, and medium terms, respectively. In the presence of time variation, if we interpret $\beta_{0,t}$, $\beta_{1,t}$, and $\beta_{2,t}$ as factors, it is easy to see that the loading on $\beta_{0,t}$ is constant and equal to one, not decaying to zero with increasing maturity, which makes $\beta_{0,t}$ the factor governing the curve level in the long term. The loading on $\beta_{1,t}$ equals one at zero maturity and decays exponentially as the maturity increases, making $\beta_{1,t}$ the factor associated with short-term dynamics. In contrast, the loading on $\beta_{2,t}$ is zero at a zero maturity, increases initially, and then decays to zero as the maturity grows, making $\beta_{2,t}$ the factor capturing the medium-term behavior of the yield curve.

Interestingly, the three factors $\beta_{0,t}$, $\beta_{1,t}$, and $\beta_{2,t}$ can also be interpreted as linear transformations of the classical, latent level, slope, and curvature features of the riskless bond yield curve. Specifically, if we define the level of the yield curve as $y(120)$, the slope as $y(120) - y(3)$, and the curvature as $2y(24) - y(120) + y(3)$, we observe a close correspondence with $\beta_{0,t}$, $-\beta_{1,t}$, and $\beta_{2,t}$ and β_2, as illustrated in Figures 1–3.

The estimated correlations between the level, slope, and curvature extracted from the observed yield curves and the fitted factors are nearly perfect, with values of 0.973, 0.995, and 0.998 for β_0, $-\beta_1$, and β_2, respectively. The sample time-series cross-factor correlations are estimated as -0.229, 0.266, and 0.620 for β_0, β_1, and β_2, respectively.

The coefficient λ governs the speed of the exponential decay of the loadings of the factors. While this may be estimated along with the factors $\beta_{0,t}$, $\beta_{1,t}$, and $\beta_{2,t}$ at any time, we fix it at $\lambda = 0.0609$, as suggested in Diebold and Li (2006). This value maximizes the loading on the medium-term, or curvature, factor $\beta_{2,t}$ at the 30-month maturity.[2]

At each point in time t, using monthly data, we recursively estimate the NS model in (2) and extract the factors best fitting the cross-section of U.S. Treasury constant maturity rates. These rates, which are available for maturities of 3 months; 6 months; and 1-, 2-, 3-, 5-, 7-, and 10 years, provide a reliable representation of the overall yield curve. By employing a recursive estimation procedure, we thus obtain a 372×3 estimated matrix of factors.

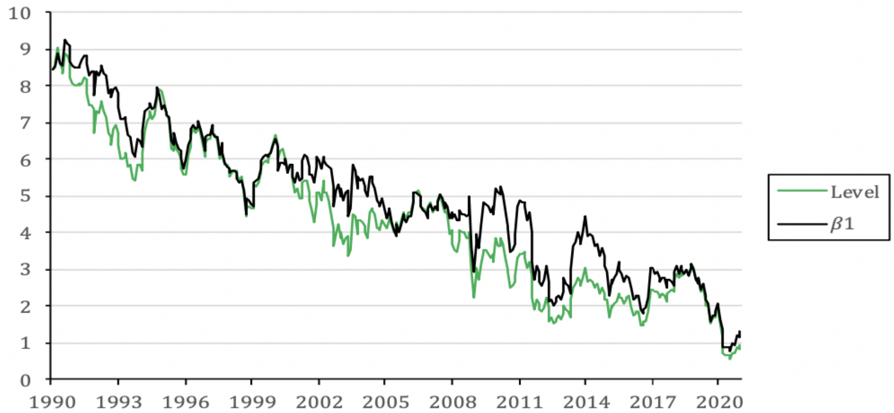

Figure 1. US yield curve level and Nelson–Siegel long-term factor. This figure illustrates the long-term level factor ($\beta_{0,t}$) derived from the Nelson–Siegel model. The data span the sample January 1990–December 2020.

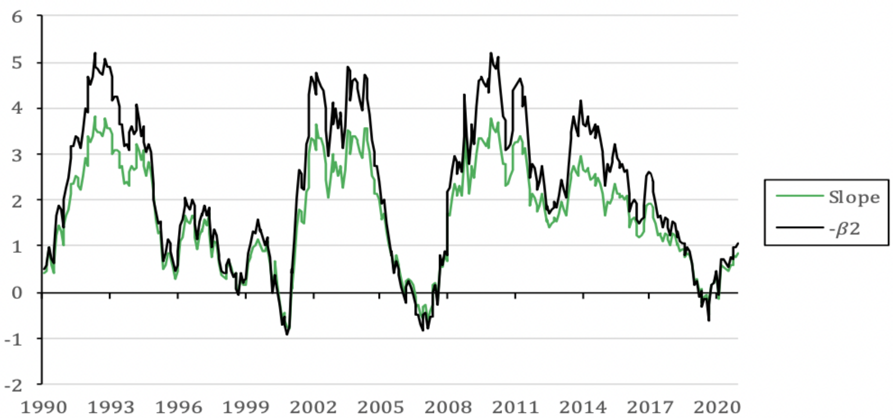

Figure 2. US yield curve slope and Nelson–Siegel short-term factor. This figure illustrates the short-term slope factor ($-\beta_{1,t}$) derived from the Nelson–Siegel model. The data span the sample January 1990–December 2020.

Figure 3. US yield curve curvature and Nelson–Siegel medium-term factor. This figure illustrates the medium-term curvature factor ($\beta_{2,t}$) from the Nelson–Siegel model. The data span the sample January 1990–December 2020.

2.2. Yield Curve Forecasting

The out-of-sample forecast period used in this paper spans January 2011–December 2020. An expanding window approach is employed, whereby the model estimation at each time t benefits from an additional data point relative to the previous estimation. The forecasting performance of the models is assessed at horizons of 1-, 6-, and 12-month steps. This is done with reference to maturities of 3 months, 1 year, 5 years, and 10 years. Forecasts for the 6 and 12 steps ahead are computed by iterating on the dynamic conditional mean models for the NS factors specified below. Specifically, for each period from t to $t+h-1$, the model uses the predicted values as inputs for computing the time $t+h$ forecast, rather than the actual observed values.

The factor prediction models are listed below.

2.2.1. AR(1) Model

We start with an AR(1) process applied to each of the time series of the three NS factors $\hat{\beta}_{1,t}$, $\hat{\beta}_{2,t}$, and $\hat{\beta}_{3,t}$ recursively estimated previously for a given maturity and horizon:

$$\hat{y}_{(t+h|t)}(m) = \hat{\beta}_{1,t+h|t} + \hat{\beta}_{2,t+h|t}\left(\frac{1-e^{-\lambda m}}{\lambda m}\right) + \hat{\beta}_{3,t+h|t}\left(\frac{1-e^{-\lambda m}}{\lambda m} - e^{-\lambda m}\right) \tag{3}$$

where the model parameters $\hat{\beta}_{i,t+h/t}$ are predicted from

$$\hat{\beta}_{i,t+h|t} = \tilde{\alpha}_{i,t,h} + \tilde{\gamma}_{i,t,h}\hat{\beta}_{i,t} \quad \text{for } i = 1,2,3 \tag{4}$$

and $\tilde{\alpha}_{i,t,h}$, $\tilde{\gamma}_{i,t,h}$ are estimated by OLS using data up to time t on the factors obtained in the first step.[3]

Equation (3) is derived directly from Equation (2), maintaining the refinement proposed by Diebold and Li (2006) to improve estimation stability and interpretability. The key distinction lies in the factor representation, whereby the medium-term component $\beta_{2,t}$ is defined with an alternative decay structure that ensures a better empirical fit without affecting the fundamental properties of the model. The factor $\beta_{2,t}$, often referred to as the "curvature factor", exhibits an initial increase before decaying to zero as the maturity grows. While its decay property differs from a strict second-derivative curvature measure, it retains its role in capturing medium-term fluctuations in the yield curve.

2.2.2. Joint VAR(1) Model

We consider a VAR(1) process applied to each of the three NS model factors $\beta_{1,t}$, $\beta_{2,t}$, and $\beta_{3,t}$ as estimated previously:

$$\hat{\boldsymbol{\beta}}_{t+h|t} = \tilde{\boldsymbol{\alpha}}_{t,h} + \tilde{\boldsymbol{\Gamma}}_{t,h}\hat{\boldsymbol{\beta}}_t \quad \text{for } i = 1,2,3 \tag{5}$$

where $\tilde{\boldsymbol{\alpha}}_{t,h}$ and $\tilde{\boldsymbol{\Gamma}}_{t,h}$ are estimated by OLS using data up to time t on the vector of factors estimated in the earlier stage.

2.2.3. Random Walk Model for Yields

We also use as a benchmark a simple random walk model for yields:

$$\hat{y}_{(t+h|t)}(m) = y_t(m) \tag{6}$$

2.2.4. AR(1) Model for Yields

As a parametric generalization of (6), we study an AR(1) process for yields:

$$\hat{y}_{(t+h|t)}(m) = \tilde{\mu}_{t,h,m} + \tilde{\kappa}_{t,h,m} y_t(m) \tag{7}$$

where $\tilde{\mu}_{t,h,m}$ and $\tilde{\kappa}_{t,h,m}$ are estimated by OLS using data up to time t.

2.2.5. Slope Regression Model

Following, for instance, Steeley (2014) and Chen and Niu (2014), we estimate a slope regression model specified as follows:

$$\hat{y}_{(t+h/t)}(m) - y_t(m) = \tilde{\alpha}(m) + \tilde{\gamma}(m)(y_t(m) - y_t(3)) \tag{8}$$

where the forecast is generated using a regression of yield changes over the prediction horizon h on a constant and the difference between the yield at maturity m and the 3-month yield. This approach captures the slope of the yield curve across different maturities. However, it is important to note that this model does not allow the prediction of the 3-month yield.

2.2.6. VAR(1) Model for Yields

We extend the AR(1) model by incorporating a VAR(1) process applied to all available yield maturities:

$$\hat{\boldsymbol{y}}_{(t+h|t)} = \tilde{\boldsymbol{\mu}}_{t,h} + \tilde{\boldsymbol{\Gamma}}_{t,h}\boldsymbol{y}_t \tag{9}$$

where the estimation of the matrices of coefficients is straightforward but horizon h-specific.[4]

2.2.7. VAR(1) Model for Yield Changes

Because the unit root nature of US riskless yields has been debated in the traditional finance literature (see, e.g., Pesando (1979)), we also analyze a VAR(1) model for yield changes:

$$\Delta\hat{\boldsymbol{y}}_{t+h|t} = \tilde{\boldsymbol{a}}_{t,h} + \tilde{\boldsymbol{A}}_{t,h}\Delta\hat{\boldsymbol{y}}_t, \tag{10}$$

which removes by differentiation any stochastic trends but does not test or exploit the existence of cointegration among interest rates.[5]

2.2.8. AR(1) Model for Principal Components

In addition, following Diebold and Li (2006), we decompose the time t sample covariance matrix C of the constant maturity rates into $C = QWQ'$ using principal component analysis. As is well known, the eigenvalues and eigenvectors of C are represented by the diagonal matrix W and columns of Q, respectively. We consider the three largest eigenvalues w_1, w_2, w_3 and their corresponding eigenvectors $\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q}_3$. Each of the eigenvectors contains a number of elements equal to the number of the available maturities, indexed by m. We then employ univariate AR(1) models to separately forecast each of the principal components h steps ahead:

$$\hat{x}_{i,t+h|t} = \tilde{\eta}_{i,t} + \tilde{\zeta}_i x_{i,t} \quad \text{for } i = 1,2,3 \text{ and } h = 1,6,12 \tag{11}$$

where the principal components $x_{i,t}$ are defined as

$$x_{i,t} = \boldsymbol{q}_i' \boldsymbol{y}_t \tag{12}$$

and $\tilde{\eta}_{i,t}$ and $\tilde{\zeta}_i$ are recursively estimated by OLS. The predicted yields h-steps ahead are then obtained from

$$\hat{y}_{(t+ht)}(m) = q_1(m)\tilde{x}_{1,t+h|t} + q_2(m)\tilde{x}_{2,t+h|t} + q_3(m)\tilde{x}_{3,t+h|t} \tag{13}$$

where $q_i(m)$ is the element of the eigenvector $\boldsymbol{q_i}$ corresponding to maturity m.

2.3. Returns Forecasting

In this section, we introduce simple methods to evaluate the predictive power of the NS factors $\beta_{1,t}$, $\beta_{2,t}$, and $\beta_{3,t}$ for returns across various asset classes. In contrast to the riskless bond yield forecasting framework outlined in Section 2.2, this analysis requires us to use a distinct set of models. For corporate and Treasury bonds, equity, REIT, and commodity returns, we evaluate the forecasting performance of the three predictive models. First, we forecast returns using direct, h-step-ahead regression on the NS model factors estimated earlier:[6]

$$\hat{r}_{i,t+h|t} = \tilde{\alpha}_{i,t} + \tilde{\gamma}_{1,i,t}\tilde{\beta}_{1,t+h|t} + \tilde{\gamma}_{2,i,t}\tilde{\beta}_{2,t+h|t} + \tilde{\gamma}_{3,i,t}\tilde{\beta}_{3,t+h|t} \tag{14}$$

In this context, the coefficients $\tilde{\alpha}_{i,t}$, $\tilde{\gamma}_{1,i,t}$, $\tilde{\gamma}_{2,i,t}$, and $\tilde{\gamma}_{3,i,t}$ are estimated through recursive OLS. The forecasts for the factors $\tilde{\beta}_{1,t+h|t}$, $\tilde{\beta}_{2,t+h|t}$, and $\tilde{\beta}_{3,t+h|t}$, which are used as inputs for return predictions, are generated using either an AR(1) or VAR(1) process, as outlined in Section 2.2. The regression model in Equation (14) represents our workhorse model and provides direct information on whether, how, and for which assets/portfolios the NS factors may be structural features of the SDF(s), to the point of allowing us to obtain accurate predictions of the returns of other assets.

As additional benchmarks that fail to depend on the NS fixed income factors, we use a simple AR(1) model applied directly to returns:

$$\hat{r}_{i,t+h|t} = \tilde{\alpha}_{i,t} + \tilde{\gamma}_{i,t} r_{i,t} \tag{15}$$

where the coefficients are estimated according to the usual logic. A third and final benchmark is obtained by imposing the restriction $\gamma_{i,t} = 0$ for all portfolios and times, which features a classical simple historical average (HA) model, as featured in recent work on the predictability of asset returns (see, e.g., Rapach et al. (2010), Guidolin et al. (2013)), and the review in Rapach and Zhou (2013)).[7]

2.4. Measuring and Testing (Equal) Predictive Accuracy

To evaluate the predictive accuracy of alternative models, we apply two statistical tests: Diebold–Mariano (henceforth, DM) tests of pairwise equal predictive accuracy (Diebold & Mariano, 2002) and the Model Confidence Set (henceforth, MCS) procedure to perform tests and a multiple-model assessment (Hansen et al., 2011). Detailed formulas of these tests/algorithms are reported in Appendix A for completeness.

The DM test requires the computation of forecast errors for each model and an assumption on the functional form of the loss measuring the loss of utility to investors derived from the forecast errors. For yields with maturity m or asset returns indexed by i, the forecast error at time t and horizon h is defined as the difference between the forecast and the observed value. The loss functions are then assumed to be of the squared type. Although rather different loss functions might be assumed, the classical squared loss makes sense within the empirical finance literature because of the well-known connections between squared loss and measures of risk-adjusted performance, such as the Sharpe ratio (see Campbell and Thompson (2008)). At this point, the loss differentials between competing models considered in pairs are tested for statistical significance, see Appendix A for details.

The MCS procedure enables the comparison of all competing models within a multivariate framework. Starting with an initial set of models M, the procedure identifies a subset $M^* \subseteq M$ that, at a predetermined confidence level, contains the best forecasting models, in the sense that these are not dominated by any other model in the set outside M^*). Dominance is conceptually identified with rejection of the null of equal predictive accuracy (EPA) in favor of a model, in a pairwise comparison. Depending on the relative performance of the models, M^* may consist of a single model outperforming all others or include a rich set of models, in principle even all of M, when no dominance relationships can be established. Depending on the relative performance of the different models, the procedure sequentially eliminates models from the superior set until the EPA hypothesis is no longer rejected. Appendix A provides details.

2.5. Combination Forecasts With and Without NS Factors

To further evaluate the predictive power of the NS factors, we conduct a robustness analysis by implementing combination forecasts based on individual predictive regression models. Following Rapach et al. (2010), we construct forecasts using a broad set of macro-financial predictors originally introduced by Welch and Goyal (2008). The predictive regression model follows a standard formulation. Our forecasts are generated using an expanding window estimation approach with a six-month horizon forecast.

The predictive variables included in our analysis consist of valuation ratios, market risk indicators, interest rate measures, and credit risk variables. Specifically, we use the following GW predictors:[8]

- Dividend–Price Ratio (log), D/P: The logarithmic difference between total dividends paid on the S&P 500 index and the index's price level, where dividends are aggregated over a rolling one-year period.
- Dividend Yield (log), D/Y: The log of dividends minus the log of lagged S&P 500 prices.
- Earnings–Price Ratio (log), E/P: The difference between the logarithm of S&P 500 earnings (measured over a trailing 12-month period) and the log of stock prices.
- Dividend–Earnings Ratio (log), D/E: Logarithmic gap between dividend payments and earnings on the S&P 500 index.
- Stock Variance, SVAR: The cumulative sum of squared daily returns on the S&P 500 index over a given month.
- Book-to-Market Ratio, B/M: The ratio between book value and market capitalization of the Dow Jones Industrial Average firms.
- Net Equity Issuance, NTIS: The ratio of 12-month cumulative net stock issuance by NYSE-listed firms to their total year-end market value.

- Treasury Bill Rate, TBL: Yield on three-month U.S. Treasury bills traded in the secondary market.
- Long-Term Yield, LTY: Yield on long-term government bonds.
- Long-Term Return, LTR: Realized return from investing in long-term government bonds.
- Term Spread, TMS: The difference between long-term government bond yields and short-term Treasury bill rates.
- Default Yield Spread, DFY: The yield spread between BAA-rated and AAA-rated corporate bonds.
- Default Return Spread, DFR: The difference in returns between long-term corporate and long-term government bonds.
- Inflation, INFL: Computed using the Consumer Price Index (CPI) for urban consumers. Following Welch and Goyal (2008), inflation is lagged by one month to account for release timing and used as $x_{i,t-1}$ in the regressions.

To assess the impact of the NS factors, we compare two models: one that includes the NS factors in the combination forecasts and one that excludes them. This approach allows us to isolate the contribution of the NS factors to the predictive accuracy and determine whether they improve the forecasting performance relative to a benchmark set of macro-financial predictors. The out-of-sample period for the analysis spans from January 2011 to December 2020, ensuring consistency with the evaluation of Diebold–Mariano test statistics and Model Confidence Set results.

3. Data Description

For all the assets and portfolios listed below, our sample spans the period January 1990–December 2020, comprising daily US yield curve rates sourced from the US Treasury Department; monthly returns on Treasury bills, notes, and bonds across various maturities; US equity returns spanning 113 distinct portfolios; sector-specific returns for REITs; corporate bond returns from investment-grade issuers in the US; and realized, chained return series for a diverse range of commodity futures.

3.1. US Treasury Yields

We used end-of-month data on the yield curve published by the US Treasury, commonly referred to as Constant Maturity Treasury (CMT) rates. These rates were derived through the interpolation of closing market bids from actively traded on-the-run US Treasury securities in the over-the-counter market. The interpolation employs a cubic spline model with fixed maturities at 1, 2, 3, and 6 months, as well as 1, 2, 3, 5, 7, 10, 20, and 30 years.[9] The sample period January 1990–December 2020 was used to construct the monthly yield curve. The analysis was limited to maturities for which the full sample was available, specifically the 3-month, 6-month, and 1-, 2-, 3-, 5-, 7-, and 10-year maturities.[10] The maturities used to estimate the NS factors covered a broader range than those utilized for yield curve forecasting. Forecasts were generated for the 3-month, 1-year, 5-year, and 10-year maturities, while the NS factors were estimated using yields from eight distinct maturities.

3.2. US Treasury Returns

We computed monthly returns based on the constant maturity rates published by the US Treasury discussed in the previous section. These returns were derived from holding 3-month, 6-month, and 1-, 2-, 3-, 5-, 7-, and 10-year T-bills and T-notes. The calculations were performed under two different sets of assumptions, depending on the maturity of the securities analyzed.

For coupon Treasury notes with maturities of one year or longer, we assumed the coupon bond was purchased at par at the end of the previous month, with a flat yield curve corresponding to the specified maturity. Monthly returns were then calculated based on changes in the bond's valuation. Coupons are paid semiannually, on the last day of the month, beginning six months after the purchase date. The total return from holding the security for one month, comprising both the coupon income return and capital appreciation, was calculated as

$$\frac{P(t_2, y_2(m-1), m-1)}{P(t_1, y_1(m), m)} - 1 \tag{16}$$

where $P(t, y(m), m)$ is the price of the bond with maturity m (expressed in years of fractions of years) and yield $y(m)$ at time t, computed as

$$P(t, y(m), m) = \frac{100}{(1 + y_t(m)/2)^{2N-1+D/S}} \left\{ (1 + y_t(m)/2)^{1-D/S} \left[1 - (1 + y_t(m)/2)^{-2N} \right] \right\} \tag{17}$$

where D/S indicates the fraction of time remaining until the next coupon payment in terms of coupon frequency.

For maturities shorter than a year, the approximate total return consists only of capital appreciation. In this case, the T-bill does not pay coupons, and at the beginning of each month, a T-bill is purchased at the prior month's end price. Monthly, realized returns reflect the change in pricing of the bill under the assumption that the yield curve is flat for the maturity considered, which leads to a simple discounting formula for the residual life of the bill.

3.3. US Equity Returns

We studied the predictability of returns on 113 different US-based equity portfolios. These portfolios, made available through Kenneth French's website,[11] included the market portfolio, 17 industry portfolios, and various other portfolios obtained by sorting the CRSP stock universe (which currently includes all NYSE, NASDAQ, and AMEX-listed shares of stock) by a range of measurable stock characteristics: market capitalization (size) and book-to-market (value); bivariate combinations of size with other characteristics, including investments, operating profitability, cashflows, the dividend yield, the earnings-price ratio, and price momentum.

3.4. US REIT Returns

We also analyzed indices for closed-end fund-style returns derived from investments in real estate equity and debt instruments. REITs are companies that own, operate, or finance income-generating real estate across various sectors and regions, allowing investors to gain indirect exposure to the real estate market. These data were sourced from NAREIT (National Association of Real Estate Investment Trusts) and consist of the FTSE NAREIT U.S. Real Estate Index Series. Our analysis focused on monthly total returns, including those for all REITs combined (equity and mortgage), all equity REITs, all mortgage REITs, and six sector-specific REIT indices (office, industrial, retail, lodging and resorts, and self-storage). The total return series represent the sum of income and capital appreciation returns, calculated before taxes and commissions.

3.5. US Corporate Bond Returns

We assessed the predictability of four ICE Bank of America (BofA) US corporate bond series. The series correspond to portfolios of securities carrying investment-grade ratings of AAA, AA, A, and BBB as of the end of each month in the sample.[12] The data, sourced from the St. Louis FED, track total returns for US dollar-denominated investment-grade corporate bonds publicly issued in the US domestic market.

3.6. Commodity Returns

Lastly, we examined the predictability of commodity returns across 17 series, encompassing precious metals, industrial metals, agricultural products, and livestock. The commodities were gold, silver, platinum, copper, Brent crude, gasoil, light crude oil, natural gas, cotton, cocoa, coffee, corn, lumber, soybean oil, soybeans, wheat, and live cattle. The monthly data, all sourced from Refinitiv, represent continuous futures contracts, with monthly returns calculated based on changes in the series levels.

3.7. Summary Statistics

The six panels of Table A1 (collected in Appendix B, to save space, labeled A through F) provide information on the US yield curve rates and the five sets of return series outlined in Sections 3.1–3.6.

Most return series exhibited left skewness, with the exception of returns on commodity futures and on the US Treasury securities. This negative skewness can be attributed to the various market crashes that occurred in our sample. These sample skewness values emphasized the occurrence of extreme negative price events in our sample and need to be accounted for in what follows.

Table 1 presents summary statistics for the first-stage, estimated NS factors. The 1-, 12-, and 30-month sample autocorrelations ($\hat{\rho}_k$ for $k = 1, 12, 30$) reported in the table indicate that the level factor, $\hat{\beta}_0$, is the most persistent, as it seems to befit the level of a stationary time series. Its sample mean indeed agrees with a positive annualized risk-free rate of approximately 4.9 percent. In addition, unsurprisingly, the slope factor was negative on average (−2.21 percent), which captured a term structure that tends to be upward sloping in our sample. Finally, the curvature factor tended to be negative on average, which points to a concave yield curve shape. According to the augmented Dickey–Fuller test, neither the slope nor the level factors exhibited unit roots at any standard confidence level on average. However, the curvature factor, β_2, was found to possibly display a unit root at both the 5% and 10% confidence levels.[13]

Table 1. Summary statistics for the fitted Nelson–Siegel factors.

Factor	Mean	Std. Dev.	Minimum	Maximum	$\hat{\rho_1}$	$\hat{\rho_{12}}$	$\hat{\rho_{30}}$	ADF
$\hat{\beta_1}$	4.893	1.955	0.747	9.267	0.980	0.776	0.616	−1.398
$\hat{\beta_2}$	−2.205	1.582	−5.214	0.980	0.973	0.523	−0.163	−2.235
$\hat{\beta_3}$	−1.801	2.226	−6.796	4.704	0.954	0.582	0.200	−2.984

The table presents summary statistics for the estimated Nelson–Siegel factors, using data for a January 1990–December 2020 sample. The augmented Dickey–Fuller (ADF) test was computed using the Schwarz information criterion (SIC) with up to 16 lags. The critical values to reject the null hypothesis of no unit roots were −3.452, −2.870, and −2.571 at the 1%, 5%, and 10% size levels, respectively.

All three factors only showed significant partial autocorrelations at the first lag, justifying their modeling with an AR(1) and VAR(1) processes as described earlier, in Section 2.2.

4. Empirical Results

4.1. Yield Curve Forecasting

The NS factors were forecast using either the AR(1) or the VAR(1) model, with the latter model able to account for cross-factor serial correlations. These factor forecasts were then applied to compute yield forecasts for the 3-month, 1-year, 5-year, and 10-year maturities, as implied by Equation (2). While eight maturities were employed to estimate the NS factors, only the four listed maturities were predicted. Table 2 presents the results of the Diebold–Mariano test applied to forecast the bond yields.

Table 2. Diebold–Mariano tests concerning alternative forecasts of US treasury rates.

Yield Maturity	*h* = 1	*h* = 6	*h* = 12
		VAR(1) on NS	
3 Months	**−2.052**	−0.552	0.059
1 Y	0.835	0.156	0.210
5 Y	−1.538	0.631	0.249
10 Y	0.501	0.389	0.137
		AR(1) on Yields	
3 Months	2.504	0.733	0.273
1 Y	4.614	0.306	0.080
5 Y	2.801	1.435	0.926
10 Y	1.335	1.424	1.176
		VAR(1) on Yield Changes	
3 Months	2.623	0.430	−0.081
1 Y	3.431	0.111	−0.188
5 Y	2.135	1.325	0.839
10 Y	2.062	1.424	1.176
		Random Walk	
3 Months	2.221	0.440	−0.050
1 Y	4.314	0.125	−0.156
5 Y	2.527	1.351	0.850
10 Y	1.535	1.444	1.188
		Slope Regression	
3 Months	n.a.	n.a.	n.a.
1 Y	5.408	−0.411	−1.364
5 Y	2.268	0.931	0.440
10 Y	1.967	0.878	0.695
		VAR(1) on Yields	
3 Months	3.692	**−4.433**	**−3.406**
1 Y	4.358	−1.976	−1.057
5 Y	2.963	**−3.564**	−1.825
10 Y	1.458	**−7.165**	**−3.938**
		Predictive Regression on Principal Components	
3 Months	**−6.097**	−1.564	−0.975
1 Y	**−6.379**	**−2.379**	**−2.780**
5 Y	**−14.848**	**−5.825**	**−3.533**
10 Y	**−20.328**	**−7.557**	**−2.278**

The table shows the results of the Diebold–Mariano tests. In particular, the table shows the DM test statistic for the null of identical predictive accuracy applied to seven models compared to the AR(1) model applied to the Nelson–Siegel factors. The pseudo out-of-sample period spans January 2011–December 2020. Boldfaced values indicate combinations of maturities, horizon, and tested models for which we could reject the null hypothesis of equal predictive accuracy vs. the AR(1) model for NS factors in a 5%-sized, one-tailed test.

We compared the predictive accuracy of the AR(1) for the NS factors against seven benchmark models, including the VAR(1) for NS factors; the random walk for the NS factors; predictive slope regressions, AR(1), and VAR(1) for yields; AR(1) for yield changes; and a direct regression on the first three AR(1) principal components. The comparison spanned three forecast horizons (one-, six-, and twelve-step-ahead) and four yield maturities. The results from the DM one-sided test, implemented at a 5% size, show that the null hypothesis of equal predictive accuracy was only rejected for a limited set of combinations, even though the negative statistics always indicated that using AR(1) for the factors predicted more accurately.

In particular, we found that using the AR(1) model for the NS factors outperformed the direct regression on principal components across most maturities and horizons, except for 3-month yields at the six- and twelve-step horizons. It also performed better than the VAR(1) for NS factors for the one-step-ahead forecasts of 3-month yields. In the case of the VAR(1) for yields, the AR(1) for NS factors performed better at a six-month horizon in the case of the 3-month, 5-year, and 10-year yields, and at a twelve-step horizon for 3-month and 10-year yields. The DM test statistics were positive and large when comparing the AR(1) for NS factors to benchmarks like the AR(1) and VAR(1) for yields, slope regression, and random walk, for one- and six-month horizons. However, despite the poor overall performance of the AR(1) model compared to the benchmarks, the AR(1) turned out not to imply a significant difference in predictive power.[14]

While the DM tests failed to establish the clear outperformance of the AR(1) for NS factors relative to the benchmarks, an interesting pattern emerges in Table 2: the relative forecasting performance improved as the horizon lengthens. This improvement was more evident for shorter maturities (3-month and 1-year yields) than for longer ones. Overall, the AR(1) for NS factors performed relatively well for long horizons and short maturities, as previously noted by Diebold and Li (2006). However, in contrast to their findings, our results do not indicate significant outperformance against the random walk model at conventional confidence levels. This aligns with Duffee (2013), who also found that the AR(1) process for NS factors failed to outperform the random walk in terms of RMSE, using a more recent sample, as we did in our work.

The MCS procedure yielded results similar to those of the Diebold–Mariano test. Table 3 shows that the outcomes were consistent across both the TR and TMax test statistics (see Appendix A for the meaning of these statistics). The finding that the AR(1) applied to the NS factors performed better for short-term maturities and longer forecasting horizons is confirmed by these tests. The AR and VAR models were instead excluded from the superior set (M^*) when forecasting one-step-ahead yields, except for the 10-year yield. However, the models for NS factors were included in the superior set for the 6- and 12-step-ahead forecast horizons, ranking among the top three models for various maturities.

Table 3. Model confidence set result for US Treasury rates.

Panel A—MCS test results for 1-step-ahead forecasts using the TMax test statistic									
	Rankings				**Summary**				
Model	**3 Mo**	**1 Y**	**5 Y**	**10 Y**	**Model**	**Excluded**	**1st to 3rd**	**4th to 6th**	**7th to 8th**
AR(1) on NS	*excluded*	*excluded*	*excluded*	7	AR(1) on NS	3	-	-	1
VAR(1) on NS	*excluded*	*excluded*	*excluded*	6	VAR(1) on NS	3	-	1	1
Slope Regression	*excluded*	3	2	1	Slope Regression	1	3	-	-
AR on Yields	3	4	4	*excluded*	AR on Yields	-	1	3	-
VAR(1) on Yields	2	1	1	4	VAR(1) on Yields	-	3	1	-
VAR(1) on Yield Changes	1	2	3	2	VAR(1) on Yield Changes	-	4	-	-
Principal Components	*excluded*	*excluded*	*excluded*	*excluded*	Principal Components	4	-	-	-
Random Walk	4	5	5	3	Random Walk	-	1	3	-
Panel B—MCS test results for 6-step-ahead forecasts using the TMax test statistic									
Model	**3 Mo**	**1 Y**	**5 Y**	**10 Y**	**Model**	**Excluded**	**1st to 3rd**	**4th to 6th**	**7th to 8th**
AR(1) on NS	4	5	5	6	AR(1) on NS	-	-	4	-
VAR(1) on NS	5	2	5	4	VAR(1) on NS	-	1	3	-
Slope Regression	*excluded*	6	4	3	Slope Regression	1	1	2	-
AR on Yields	1	1	1	5	AR on Yields	-	3	1	-
VAR(1) on Yields	2	1	1	4	VAR(1) on Yields	-	3	1	-
VAR(1) on Yield Changes	3	4	3	2	VAR(1) on Yield Changes	-	3	1	-
Principal Components	*excluded*	*excluded*	*excluded*	*excluded*	Principal Components	4	-	-	-
Random Walk	2	3	2	1	Random Walk	-	4	-	-

Table 3. *Cont.*

Panel C—MCS test results for 12-step-ahead forecasts using the TMax test statistic									
Model	**3 Mo**	**1 Y**	**5 Y**	**10 Y**	**Model**	**Excluded**	**1st to 3rd**	**4th to 6th**	**7th to 8th**
AR(1) on NS	5	2	6	6	AR(1) on NS	-	1	3	-
VAR(1) on NS	4	3	5	5	VAR(1) on NS	-	1	3	-
Slope Regression	*excluded*	6	4	3	Slope Regression	1	1	2	-
AR on Yields	1	1	1	4	AR on Yields	-	3	1	-
VAR(1) on Yields	*excluded*	7	*excluded*	*excluded*	VAR(1) on Yields	3	-	-	1
VAR(1) on Yield Changes	3	4	3	2	VAR(1) on Yield Changes	-	4	1	-
Principal Components	6	*excluded*	*excluded*	*excluded*	Principal Components	3	-	1	-
Random Walk	2	5	2	1	Random Walk	-	2	1	-
Panel D—MCS test results for 1-step-ahead forecasts using the TR test statistic									
Model	**3 Mo**	**1 Y**	**5 Y**	**10 Y**	**Model**	**Excluded**	**1st to 3rd**	**4th to 6th**	**7th to 8th**
AR(1) on NS	*excluded*	*excluded*	*excluded*	7	AR(1) on NS	3	-	-	1
VAR(1) on NS	*excluded*	*excluded*	*excluded*	6	VAR(1) on NS	3	-	1	-
Slope Regression	*excluded*	3	2	1	Slope Regression	1	3	-	-
AR on Yields	3	4	4	5	AR on Yields	-	2	2	-
VAR(1) on Yields	2	1	1	4	VAR(1) on Yields	-	3	1	-
VAR(1) on Yield Changes	1	2	3	2	VAR(1) on Yield Changes	-	4	-	-
Principal Components	*excluded*	*excluded*	*excluded*	*excluded*	Principal Components	4	-	-	-
Random Walk	4	5	5	3	Random Walk	-	-	4	-
Panel E—MCS test results for 6-step-ahead forecasts using the TR test statistic									
Model	**3 Mo**	**1 Y**	**5 Y**	**10 Y**	**Model**	**Excluded**	**1st to 3rd**	**4th to 6th**	**7th to 8th**
AR(1) on NS	4	4	6	6	AR(1) on NS	-	2	2	-
VAR(1) on NS	5	5	5	4	VAR(1) on NS	-	1	3	-
Slope Regression	*excluded*	6	4	3	Slope Regression	1	2	1	-
AR on Yields	1	1	1	5	AR on Yields	-	2	2	-
VAR(1) on Yields	*excluded*	7	*excluded*	*excluded*	VAR(1) on Yields	3	-	1	-
VAR(1) on Yield Changes	3	2	3	2	VAR(1) on Yield Changes	-	2		
Principal Components	6	8	*excluded*	*excluded*	Principal Components	2	-	1	1
Random Walk	2	3	2	1	Random Walk	-	3	1	-
Panel F—MCS test results for 12-step-ahead forecasts using the TR test statistic									
Model	**3 Mo**	**1 Y**	**5 Y**	**10 Y**	**Model**	**Excluded**	**1st to 3rd**	**4th to 6th**	**7th to 8th**
AR(1) on NS	5	2	6	6	AR(1) on NS	-	2	2	-
VAR(1) on NS	4	4	5	5	VAR(1) on NS	-	2	2	-
Slope Regression	*excluded*	6	4	3	Slope Regression	1	1	2	-
AR on Yields	1	1	1	4	AR on Yields	-	3	1	-
VAR(1) on Yields	*excluded*	7	*excluded*	*excluded*	VAR(1) on Yields	3	-	1	-
VAR(1) on Yield Changes	3	3	3	2	VAR(1) on Yield Changes	-	2	1	1
Principal Components	6	*excluded*	*excluded*	*excluded*	Principal Components	-	2	1	-
Random Walk	2	5	2	1	Random Walk	-	2	1	-

The table shows results for the Model Confidence Set tests concerning the eight models considered for forecasting the US Treasury rates on a monthly basis in the pseudo out-of-sample period from January 2011 to December 2020. Panels A through C refer to the results obtained under the TMax test statistic. Panels D through F refer to the results obtained under the TR test statistic. Each panel shows, on the left-hand side, the rankings of the MCS test, where excluded means that the model was not part of the superior set M^*; on the right-hand side, the table displays a summary of the MCS tests for the panel, counting model rankings across the four different yield maturities forecast (3 month, 1 year, 5 year, and 10 year). Panels A and D show the MCS test results for one-month-ahead forecasts, Panels B and E show results for six-month-ahead forecasts, and Panels C and F show results for 12-month-ahead forecasts.

The superior set increased in size as the forecast horizon lengthened, indicating a convergence in forecasting accuracy among the models for longer-term forecasts. In particular, the AR(1) model for yields appeared to be one of the best models across all maturities and forecasting horizons. Despite being the most accurate model for one-step-ahead forecasts, the VAR(1) for yields was excluded from the superior set for six-step and 12-step-ahead forecasts.

Additionally, although the TMax and TR test statistics yielded similar results, they differed in the number of models excluded from the superior set. For the six-month-ahead forecasts, the TMax statistic identified a narrower superior set, excluding more models than the TR statistic. This difference arose from the distinct comparison methods employed by the two test statistics: TMax evaluated each model's predictive accuracy against the average of all other models, while TR compared models pairwise. Consequently, we expected TMax to be more selective in excluding models.

All in all, the evidence presented in Tables 2 and 3 is consistent with the idea that autoregressive modeling of the NS factors is as accurate as a number of benchmarks. However, at least for horizons of 6 and 12 months, the equal predictive accuracy tests indicate that such models failed to be uniformly dominated by any of the benchmarks, at least with reference to the prediction of the very data from which the NS factors were extracted. This justifies our continued interest in their forecasting performance for asset returns in the reminder of this paper.

Finally, as a robustness check, we compared the combination forecasts following Rapach et al. (2010) for the factors of Welch and Goyal (2008) both with and without the NS factor. Table A4 in Appendix D presents this comparison. This additional analysis allowed us to evaluate the general usefulness of NS predictors in excess asset return forecasting, assessing whether their inclusion enhanced the predictive power of combination models beyond traditional macro-financial factors. The gains were particularly noticeable for short-term maturities, with an increase from 3.24% to 4.77% for 3-month yields and from 3.32% to 4.88% for 1-year yields. This suggests that the NS factor captured valuable information for forecasting short-term interest rates. While the improvement was smaller for longer maturities, such as 5-year and 10-year yields, the inclusion of the NS factor still enhanced the predictive accuracy, albeit to a lesser extent. This pattern aligns with the idea that the NS factor is particularly effective in explaining short-term yield variations, while providing moderate benefits for longer-term forecasts.

4.2. Forecasting Treasury Returns

Table A2 (in Appendix C) shows the results of the Diebold–Mariano test for return series; specifically, Panel A presents the results of the pairwise DM test for the government bond return series.[15] The comparison of the predictive performance of the NS factors—themselves predicted from either an AR(1) or a VAR(1) model—with other specifications was performed across three forecasting horizons (one-, six-, and 12-step-ahead) and eight return series from monthly holdings of 3-month, 6-month, and 1-, 2-, 3-, 5-, 7-, and 10-year US Treasuries. The results indicate that we can reject the null hypothesis of equal predictive accuracy for only a limited set of combinations. Specifically, the AR(1) model for NS factors outperformed the AR(1) model for past returns for two maturities (3 and 6 months) and two forecasting horizons (six- and twelve-step-ahead). Additionally, it outperformed the HA model for these two maturities across all three forecasting horizons. However, the AR(1) for NS factors never outperformed the VAR(1) applied to predict NS factors, which turned out to be better suited for forecasting these returns. The VAR(1) performed better for shorter maturities and forecasting horizons, but its performance decreased for longer maturities and forecast horizons. Similarly to bond yields, the forecasting accuracy of the AR(1) NS improved as the forecasting horizon lengthened, particularly when compared to the VAR(1) NS, when their ranking flipped around and once more favored the simpler AR(1) model. However, Table A2 reports a large majority of negative DM test statistics, illustrating that the (predictions of) the NS factors provided serious competition to the AR models and the very random walk for prices, when it comes to forecasting bond returns.

In Table 4, the MCS results, obtained applying the 90% confidence testing level, further confirm the superior performance of the VAR(1) model applied to predict the NS factors compared to the other models. The VAR(1) NS model was always included in the superior set M^* across all maturities and forecast horizons, consistently ranking the highest among the four models. The AR(1) NS also performed well but was often excluded from M^* due to the VAR(1) model's superior performance. Nonetheless, the results varied depending on the specific MCS test statistic used, whether TMax or TR. Under TMax, the AR(1) for the NS factors was excluded from the superior set only five times out of 24 tests (across three forecast horizons and eight return series). When we used the TR statistic, the AR(1) was excluded more frequently, six times for one-step-ahead predictions, five times for six-step-ahead, and four times for the 12-month-ahead forecasts. Finally, regarding the comparison of forecast combining models applied to Treasury returns, we observed that the inclusion of the NS factor significantly enhanced the predictive accuracy compared to using only the GW predictors. Specifically, the out-of-sample R^2_{OS} increased substantially for shorter maturities, such as the 3 M and 6 M Treasury securities, showing a marked improvement that suggests the additional informational value of the NS factor in short-term bond return forecasting. For longer maturities, the impact of the NS factor was more moderate but remained positive for several horizons, indicating its potential to enhance the performance of combined forecasting models.

Table 4. Model confidence set results for return series.

Panel A—MCS test results for 1-step-ahead forecasts using the TMax test statistic (h = 1)										
	Excl.	**1st**	**2nd**	**3rd**	**4th**	**Excl.**	**1st**	**2nd**	**3rd**	**4th**
Equities (113)										
AR(1) on NS	9	60	34	8	2	8%	53%	30%	7%	2%
VAR(1) on NS	8	43	56	4	2	7%	38%	50%	4%	2%
AR(1) for returns	69	7	4	17	16	61%	6%	4%	15%	14%
Historical Average	61	3	7	29	13	54%	3%	6%	26%	12%
Treasuries (8)										
AR(1) on NS	1	0	4	1	2	13%	0%	50%	13%	25%
VAR(1) on NS	0	5	3	0	0	0%	63%	38%	0%	0%
AR(1) for returns	1	3	0	3	1	13%	38%	0%	38%	13%
Historical Average	3	0	0	3	2	38%	0%	0%	38%	25%
REITs (9)										
AR(1) on NS	0	8	0	1	0	0%	89%	0%	11%	0%
VAR(1) on NS	1	0	7	0	1	11%	0%	78%	0%	11%
AR(1) for returns	6	1	0	2	0	67%	11%	0%	22%	0%
Historical Average	7	0	1	1	0	78%	0%	11%	11%	0%
Corporate Bonds (4)										
AR(1) on NS	0	0	4	0	0	0%	0%	100%	0%	0%
VAR(1) on NS	0	4	0	0	0	0%	100%	0%	0%	0%
AR(1) for returns	3	0	0	0	1	75%	0%	0%	0%	25%
Historical Average	1	0	0	3	0	25%	0%	0%	75%	0%
Commodities (17)										
AR(1) on NS	5	1	3	6	2	29%	6%	18%	35%	12%
VAR(1) on NS	5	0	4	5	3	29%	0%	24%	29%	18%
AR(1) for returns	2	14	0	1	0	12%	82%	0%	6%	0%
Historical Average	5	2	5	0	5	29%	12%	29%	0%	29%
Equities (113)										
AR(1) on NS	21	18	4	35	33	19%	16%	4%	31%	31%
VAR(1) on NS	14	10	9	38	42	12%	9%	8%	34%	37%
AR(1) for returns	1	28	63	11	10	1%	25%	56%	10%	9%
Historical Average	1	57	37	15	3	1%	50%	33%	13%	3%

Table 4. *Cont.*

Panel B—MCS test results for 6-step-ahead forecasts using the TMax test statistic (h = 6)										
	Excl.	**1st**	**2nd**	**3rd**	**4th**	**Excl.**	**1st**	**2nd**	**3rd**	**4th**
Treasuries (8)										
AR(1) on NS	2	2	2	0	2	25%	25%	25%	0%	25%
VAR(1) on NS	0	6	0	0	2	0%	75%	0%	0%	25%
AR(1) for returns	2	0	4	2	0	25%	0%	50%	25%	0%
Historical Average	2	0	0	4	2	25%	0%	0%	50%	25%
REITs (9)										
AR(1) on NS	0	4	1	0	4	0%	44%	11%	0%	44%
VAR(1) on NS	0	2	4	3	0	0%	22%	44%	33%	0%
AR(1) for returns	1	0	3	1	4	11%	0%	33%	11%	44%
Historical Average	0	3	1	5	0	0%	33%	11%	56%	0%
Corporate Bonds (4)										
AR(1) on NS	0	0	3	0	1	0%	0%	0%	75%	0%
VAR(1) on NS	0	4	0	0	0	0%	100%	0%	0%	0%
AR(1) for returns	0	0	0	2	2	0%	0%	0%	50%	50%
Historical Average	0	0	1	2	1	0%	0%	25%	50%	25%
Commodities (17)										
AR(1) on NS	5	2	4	5	1	29%	12%	24%	29%	6%
VAR(1) on NS	8	4	2	1	2	47%	24%	12%	6%	12%
AR(1) for returns	0	11	0	6	0	0%	65%	0%	35%	0%
Historical Average	5	0	7	0	5	29%	0%	41%	0%	29%
Panel C—MCS test results for 12-step-ahead forecasts using the TMax test statistic (h = 12)										
	Excl.	**1st**	**2nd**	**3rd**	**4th**	**Excl.**	**1st**	**2nd**	**3rd**	**4th**
Equities (113)										
AR(1) on NS	68	5	2	22	16	60%	4%	2%	19%	14%
VAR(1) on NS	62	3	3	24	21	55%	3%	3%	21%	19%
AR(1) for returns	2	51	54	4	2	2%	45%	48%	4%	2%
Historical Average	3	54	49	5	2	3%	48%	43%	4%	2%
Treasuries (8)										
AR(1) on NS	2	2	4	0	0	25%	25%	50%	0%	0%
VAR(1) on NS	0	6	1	0	1	0%	75%	13%	0%	13%
AR(1) for returns	2	0	0	2	4	25%	0%	0%	25%	50%
Historical Average	2	0	1	4	1	25%	0%	13%	50%	13%
REITs (9)										
AR(1) on NS	0	2	5	0	2	0%	22%	56%	0%	22%
VAR(1) on NS	0	6	2	1	0	0%	67%	22%	11%	0%
AR(1) for returns	2	0	2	0	5	22%	0%	22%	0%	56%
Historical Average	2	1	0	6	0	22%	11%	0%	67%	0%
Corporate Bonds (4)										
AR(1) on NS	0	0	4	0	0	0%	0%	100%	0%	0%
VAR(1) on NS	0	4	0	0	0	0%	100%	0%	0%	0%
AR(1) for returns	0	0	0	0	4	0%	0%	0%	0%	100%
Historical Average	0	0	0	4	0	0%	0%	0%	100%	0%
Commodities (17)										
AR(1) on NS	16	0	0	1	0	94%	0%	0%	6%	0%
VAR(1) on NS	16	0	0	0	1	94%	0%	0%	0%	6%
AR(1) for returns	3	7	7	0	0	18%	41%	41%	0%	0%
Historical Average	1	10	6	0	0	6%	59%	35%	0%	0%
Equities (113)										
AR(1) on NS	22	59	22	5	5	19%	52%	19%	4%	4%
VAR(1) on NS	20	44	38	7	4	18%	39%	34%	6%	4%
AR(1) for returns	73	6	13	18	3	65%	5%	12%	16%	3%
Historical Average	64	25	17	3	3	57%	4%	22%	15%	3%
Treasuries (8)										
AR(1) on NS	6	1	1	0	0	75%	13%	13%	0%	0%
VAR(1) on NS	0	4	4	0	0	0%	50%	50%	0%	0%
AR(1) for returns	2	3	0	3	0	25%	38%	0%	38%	0%
Historical Average	2	0	1	3	2	25%	0%	13%	38%	25%

Table 4. *Cont.*

Panel D—MCS test results for 1-step-ahead forecasts using the TR test statistic (*h* = 1)										
	Excl.	**1st**	**2nd**	**3rd**	**4th**	**Excl.**	**1st**	**2nd**	**3rd**	**4th**
REITs (9)										
AR(1) on NS	0	8	0	1	0	0%	89%	0%	11%	0%
VAR(1) on NS	1	0	5	2	1	11%	0%	56%	22%	11%
AR(1) for returns	6	1	2	0	0	67%	11%	22%	0%	0%
Historical Average	7	0	1	1	0	78%	0%	11%	11%	0%
Corporate Bonds (4)										
AR(1) on NS	2	0	1	0	1	50%	0%	25%	0%	25%
VAR(1) on NS	0	4	0	0	0	0%	100%	0%	0%	0%
AR(1) for returns	3	0	0	1	0	75%	0%	0%	25%	0%
Historical Average	1	0	0	0	0	25%	0%	75%	0%	0%
Commodities (17)										
AR(1) on NS	5	0	4	7	1	29%	0%	24%	41%	6%
VAR(1) on NS	6	1	5	3	2	35%	6%	29%	18%	12%
AR(1) for returns	2	14	0	1	0	12%	82%	0%	6%	0%
Historical Average	7	2	3	0	5	41%	12%	18%	0%	29%
Panel E—MCS test results for 6-step-ahead forecasts using the TR test statistic (*h* = 6)										
	Excl.	**1st**	**2nd**	**3rd**	**4th**	**Excl.**	**1st**	**2nd**	**3rd**	**4th**
Equities (113)										
AR(1) on NS	21	14	20	28	30	19%	12%	18%	25%	27%
VAR(1) on NS	11	12	13	48	29	10%	11%	12%	42%	26%
AR(1) for returns	2	33	41	16	21	2%	29%	36%	14%	19%
Historical Average	2	54	38	14	5	2%	48%	34%	12%	4%
Treasuries (8)										
AR(1) on NS	5	2	1	0	0	63%	25%	13%	0%	0%
VAR(1) on NS	0	6	0	1	1	0%	75%	0%	13%	13%
AR(1) for returns	5	0	2	1	0	63%	0%	25%	13%	0%
Historical Average	5	0	0	1	2	63%	0%	0%	13%	25%
REITs (9)										
AR(1) on NS	0	4	0	1	4	0%	44%	0%	11%	44%
VAR(1) on NS	0	2	6	1	0	0%	22%	67%	11%	0%
AR(1) for returns	0	0	1	3	5	0%	0%	11%	33%	56%
Historical Average	0	3	2	4	0	0%	33%	22%	44%	0%
Corporate Bonds (4)										
AR(1) on NS	0	0	3	1	0	0%	0%	75%	25%	0%
VAR(1) on NS	0	4	0	0	0	0%	100%	0%	0%	0%
AR(1) for returns	0	0	0	3	1	0%	0%	0%	75%	25%
Historical Average	0	0	1	0	3	0%	0%	25%	0%	75%
Commodities (17)										
AR(1) on NS	6	3	5	3	0	35%	18%	29%	18%	0%
VAR(1) on NS	8	3	4	1	1	47%	18%	24%	6%	6%
AR(1) for returns	0	11	1	5	0	0%	65%	6%	29%	0%
Historical Average	11	0	2	1	3	65%	0%	12%	6%	18%
Equities (113)										
AR(1) on NS	69	5	13	13	13	61%	4%	12%	12%	12%
VAR(1) on NS	66	3	8	27	9	58%	3%	7%	24%	8%
AR(1) for returns	2	45	49	4	13	2%	40%	43%	4%	12%
Historical Average	3	60	40	9	1	3%	53%	35%	8%	1%
Treasuries (8)										
AR(1) on NS	4	3	1	0	0	50%	38%	13%	0%	0%
VAR(1) on NS	0	5	2	0	1	0%	63%	25%	0%	13%
AR(1) for returns	5	0	0	1	2	63%	0%	0%	13%	25%
Historical Average	5	0	1	2	0	63%	0%	13%	25%	0%

Table 4. *Cont.*

Panel F—MCS test results for 12-step-ahead forecasts using the TR test statistic (h = 12)										
	Excl.	**1st**	**2nd**	**3rd**	**4th**	**Excl.**	**1st**	**2nd**	**3rd**	**4th**
REITs (9)										
AR(1) on NS	0	2	3	0	4	0%	22%	33%	0%	44%
VAR(1) on NS	0	6	3	0	0	0%	67%	33%	0%	0%
AR(1) for returns	0	0	1	5	3	0%	0%	11%	56%	33%
Historical Average	0	1	2	4	2	0%	11%	22%	44%	22%
Corporate Bonds (4)										
AR(1) on NS	0	3	1	0	0	0%	75%	25%	0%	0%
VAR(1) on NS	0	1	3	0	0	0%	25%	75%	0%	0%
AR(1) for returns	0	0	0	0	4	0%	0%	0%	0%	100%
Historical Average	0	0	0	4	0	0%	0%	0%	100%	0%
Commodities (17)										
AR(1) on NS	15	0	1	0	1	88%	0%	6%	0%	6%
VAR(1) on NS	15	0	0	1	1	88%	0%	0%	6%	6%
AR(1) for returns	3	7	6	1	0	18%	41%	35%	6%	0%
Historical Average	1	10	6	0	0	6%	59%	35%	0%	0%

This table contains results from the Model Confidence Set tests. In particular, the table shows test results for four models applied to forecast the returns on various asset classes over a pseudo out-of-sample period January 2011–September 2020. Panels A through C refer to results based on the TMax test statistic. Panels D through F refer to results based on the TR test statistic. Each panel is broken down into five sub-panels, each corresponding to a different asset class. Each panel shows summary results in absolute terms (how many times a model ranked 1st, 2nd, 3rd, and 4th or was excluded from the superior set) and in relative terms (out of the total number of return series considered for each asset class, which is shown in brackets). Panels A through D refer to one-step-ahead forecasts, Panels B and E to six-step-ahead predictions, and Panels C and F to 12-step-ahead forecasts.

More importantly for our purposes, both the AR(1) for past returns and the HA benchmarks consistently underperformed compared to the models using NS factor forecasts. These benchmarks were often excluded from the superior set and ranked low when included, especially in comparison to the VAR(1) for NS factors. Therefore, maybe unsurprisingly, predictions of NS factors extracted from Treasury yield data ranked rather highly within a rather classical set of forecast models for Treasury returns, especially for very-short- and very-long-term maturities and for six- and 12-month horizons.

4.3. Forecasting Equity Portfolio Returns

Following the same design used for Treasury returns, the DM test results for equity returns are shown in Panel B of Table A2, in Appendix C. The comparison was made across the three forecast horizons pursued in this paper and with reference to 113 portfolio return series, grouped into eight sub-categories: Fama and French factors portfolios; industry portfolios; various other portfolios sorted on the basis of stock and firm characteristics, as listed in Table A1, Panel C). In the prediction of stock returns one month ahead, the AR(1) model for the NS factors often outperformed the AR(1) applied to past returns and the HA model. Out of 113 equity return series, we rejected the null hypothesis of equal predictive accuracy between the AR(1) NS and the AR(1) for equity returns 52 times, and 49 times vs. the HA model. Such statistics indicate a superior predictive accuracy that largely exceeded the 10 percent we would expect from pure chance. However, this superior performance failed to emerge in the case of six- and 12-step-ahead forecasts, where the AR(1) NS model often underperformed compared to the benchmarks. The VAR(1) NS generally performed better than the AR(1) NS, forecasting more accurately than the latter more often than not. The difference in predictive accuracy between the two models was consistent across all three forecast horizons.

The results of the MCS approach provide a more comprehensive view of a model's relative performance. A pattern involving both NS factor-based models emerged, consistent with what was reported with reference to the DM test. According to both the TR and TMax statistics, the AR(1) and VAR(1) NS models were the most accurate for one-step-ahead forecasting, rarely being excluded from the superior set. For the one-step-ahead forecast, the AR(1) was included in the superior set with 92% and 81% probabilities for TMax and TR, respectively. In fact, the VAR(1) NS was included with slightly higher probabilities: 93% and 82% for TMax and TR. As seen for US Treasury returns, the TR statistic was more selective, excluding more models from the superior set, and so these findings were to be expected. Moreover, AR(1) and VAR(1) NS turned out to be the best performing models for 53% and 38% of the target series, respectively, being placed as the best or the second best for 83% and 88% of the target series, which is rather an impressive performance. On the contrary, the AR(1) and HA models for returns performed rather poorly and were excluded from the superior set for more than 50% of the return series. Therefore, the short-term predictive power of NS factors for a large majority of the equity portfolio returns appeared to be remarkable and, possibly surprisingly, turned out to exceed what was found in the case of Treasuries.

Finally, the inclusion of NS factors did not significantly enhance the predictive power for equity returns, as it did for yields and government returns. The only notable exception was for Fabricated Products returns, where the inclusion of NS factors led to a visible improvement in predictive performance.

4.4. Forecasting REIT Returns

In contrast, for six- and 12-month-ahead forecasts, both the AR(1) applied directly to returns and the HA outperformed the AR(1) and VAR(1) models based on NS factor predictions. Specifically, the AR(1) for stock returns was included in the superior set with probabilities of 99% and 98% for six-step-ahead forecasts (using TMax and TR statistics, respectively), and 98% in the case of the 12-step-ahead horizon case. The even simpler HA model was included with slightly lower probabilities than the AR(1) but still tended to outperform the NS-based models. Overall, the HA model demonstrated superior accuracy to the AR(1), while the VAR(1) outperformed the AR(1) NS, regardless of the forecast horizon.

The results of DM tests applied to REIT returns are presented in Panel C of Table A3, as usual, across three forecasting horizons and nine return series. Similarly to the findings for equity returns, the AR(1) based on the NS factors outperformed, at the usual one-sided 5% size, the AR(1) estimated on returns and the historical mean models for one-step-ahead forecasts. The AR(1) NS, however, failed to outperform the VAR(1) NS across all REIT return series and forecasting horizons. In the case of the one-step-ahead forecast, the AR(1) NS outperformed the standard AR(1) six times and the HA model three times, out of nine series analyzed. When we compare the performance of the two NS factor-based models, AR(1) outperformed in the case of the month horizon, while VAR(1) came out on top for the 12-step-ahead horizon; we found mixed results for the six-month horizon. As in the case of equity returns, the predictive accuracy of the AR(1) NS deteriorated relative to the benchmarks as the forecast horizon lengthened. Contrary to the equity returns case, where the performance shifted in favor of the AR(1) model on returns and HA for the 12-step-ahead forecasts, the AR(1) NS consistently outperformed the two benchmarks for most REIT indices across all horizons, including the six-step and 12-step horizons.

The findings from the DM tests were reinforced by the results from the MCS procedure, which are presented in Table 4. The AR(1) and VAR(1) driven by NS predictions were rarely excluded from the superior set across all three horizons, irrespective of the test statistic used. Only for one REIT return series and for the one-step-ahead horizon was VAR(1) NS excluded from the superior set, based on the TMax statistic. The classical AR(1) and the

HA models were instead frequently excluded from the superior set for the one-step-ahead horizon. Specifically, they were excluded six and seven times, respectively, out of the nine return series considered, regardless of the MCS test statistic used. However, both models were always included in the superior set for 6- and 12-month predictions. Even when they were included in the superior set, these models consistently ranked below the AR(1) and VAR(1) NS across the majority of the nine return series considered. Specifically, for the 6-step and 12-step-ahead horizons, the standard AR(1) never ranked first. The HA model ranked first three times for the 6-step-ahead horizon and once for the 12-step-ahead horizon, according to both MCS test statistics.

The REIT returns were predictable not only at the one-step-ahead horizon but also at the 6-step and 12-month horizons. Both the DM test and the MCS procedure showed a convergence in forecasting performance between the two NS factor-based models and the two benchmarks for the 6- and 12-step-ahead prediction horizons. The forecasts of the dynamics of the yield curve proved to contain more precise information on future REIT returns than for equity returns, indicating a surprisingly high integration between the REIT and U.S. government bond markets. This aligns with the fact that interest rate changes tend to have longer-lasting effects on real estate investments.

4.5. Forecasting Corporate Bond Returns

We analyzed the predictability of four corporate bond return series corresponding to the highest investment grade ratings. In our sample, based on the DM test statistic implemented at a 5% test size level, we failed to reject the null hypothesis of equal predictive accuracy across the three benchmarks assumed against the AR(1) NS factors model, for all forecast horizons (Panel D of Table A2). As observed for both equity and REIT returns (but not for Treasury returns), the forecasting accuracy of the AR(1) NS model decreased as the horizon lengthened, compared to the HA and the AR(1) model applied to corporate bond returns. In contrast, the AR(1) NS model performed better than the VAR(1) NS when forecasting corporate bond returns, although this turned out not to be the case for equity and REIT returns. Overall, while the AR(1) NS tended to lag behind the VAR(1), it outperformed the other two benchmark models.

The results from the MCS procedure, shown in Table 4, Panels A through F, align with those reported for REITs. In most experiments, both the AR(1) and VAR(1) NS models remained inside the superior set across all horizons, for the specific portfolio return series, and regardless of the test statistic used. The only exception concerned the AA-rated corporate bond series, for which the AR(1) NS model happened to be excluded from the superior set in the case of the one-month-ahead predictions, according to the TR statistic. On the contrary, the classical AR(1) and the HA model were typically excluded from the superior set for the one-step-ahead forecasts, but were included for the six- and twelve-month-ahead forecasts. Both models consistently ranked lower than AR(1) and VAR(1) NS, never ranking first across any of the horizons. For the six-step-ahead forecast, the HA model ranked second once and fourth three times, while the standard AR(1) for returns placed third three times and fourth once. For the 12-step-ahead forecasts, the HA model ranked third consistently, with the AR(1) for returns always ranking fourth.

Both the DM test and the MCS procedure confirmed that NS factors can predict U.S. corporate bond returns, though the predictive strength was weaker when compared to REIT returns and less pronounced than for equity returns at the one-step-ahead forecast horizon. Across the four bond return series, we could not reject the null hypothesis of equal predictive accuracy using the DM test. However, for both the 6- and 12-step-ahead predictions, the AR(1) and VAR(1) NS models outperformed the HA and the classical AR(1) and ranked higher according to the MCS procedure. The MCS results were consistent across the different test statistics, with rankings and superior set compositions remaining unchanged under both TMax and TR tests.

4.6. Forecasting Commodity Returns

With reference to the 17 return series for commodity futures, the test statistics presented in Table A2, Panel E, for various comparisons across commodities and three horizons, show that the NS factors performed poorly in predicting commodity futures returns. Similarly to the findings for corporate bond, the AR(1) NS failed to outperform the benchmarks, i.e., the AR(1) for commodity returns and the HA model. On the contrary, AR(1) NS performed similarly to the VAR(1) NS model. As in the case of other asset classes, the forecasting accuracy of the NS-driven models declined as the forecast horizon lengthened, in comparison to the benchmarks. Additionally, no clear trend was observed in the predictive performance across the four commodity subgroups (metals, energy, agricultural, and livestock).

Although this may appear unsurprising in light of the DM tests, according to the MCS procedure, the results of which are presented in Table 4, Panels A through F, forecasting commodity futures returns using NS factors led to partial success at best. The AR(1) and VAR(1) NS models were frequently excluded from the superior set M^*; this occurred for approximately 50% of forecasting experiments, for both one- and six-month-ahead forecasts. However, the performance of the two benchmarks was comparable to the HA model, which was indeed often excluded from the superior set as well. The classical AR(1) for returns had the highest predictive accuracy among all models considered, ranking first 14 times for the one-step-ahead horizon and 11 times for the six-step-ahead horizon, out of the 17 series considered. As expected, in the case of the 12-step-ahead horizon, the MCS test excluded the AR(1) and VAR(1) NS from the superior set more frequently than for the shorter prediction horizons. At 12 months, both models were only included in M^* once, according to the TMax statistic, and twice according to the TR statistic. When compared to the benchmarks, AR(1) NS outperformed HA for the one- and six-step-ahead forecasts but underperformed for the 12-step-ahead horizon. Therefore, and contrary expectations based on the influence of interest rates on commodity futures prices (Frankel, 2008), both th DM and MCS tests failed to produce significant results in forecasting commodity returns across all horizons.[16]

5. Discussion and Conclusions

The main goal of this paper was to assess the hypothesis that the three NS factors (level, slope, and curvature) derived from the US yield curve can be used to forecast returns across a variety of financial assets. To test this over-arching hypothesis, we analyzed the predictability of monthly return series covering a range of asset classes, i.e., equity portfolios, REITs, corporate bonds, commodity futures, and US Treasury notes. Given a baseline null hypothesis that, being extracted from *yield* curve data, the NS factors should contain no (or not much) information for forecasting *returns* on asset classes that remain essentially different from default risk-free bonds, our results suggest that the three Nelson–Siegel factors do carry some unsuspected predictive power well beyond the government bond market. Although the performance of AR(1) and VAR(1) models incorporating predictions of the NS factors relative to common benchmarks was not always overwhelming and it varied across forecast horizons and asset classes, the results were generally statistically significant. Specifically, we found that, in the case of one-step-ahead forecasts, the AR(1) and VAR(1) NS models frequently outperformed the benchmarks, which included a classical AR(1) model estimated on past returns and the historical mean model (derived from the random walk hypothesis for log-asset prices). However, the predictive accuracy of the AR(1) and VAR(1) NS models tended to decline as the forecast horizon lengthened, as the predictive performance at the 6- and 12-month-ahead horizons turned out to be inferior vs. the short-term horizons. These results indicate that models based on predictions of NS factors may have been more appropriate for short-term forecasting, while the benchmarks became increasingly competitive when applied within medium- to long-term applications.

As one would expect, in our recursive, pseudo OOS experiments, the strongest predictive accuracy was obtained with reference to the prediction of US Treasury returns, when both the AR(1) and VAR(1) NS models performed well across most forecast horizons. In light of the baseline null hypothesis, the results also revealed notable pockets of non-negligible predictive power in the case of corporate bonds, REITs (which are generally linked to the dynamics of interest rates through the mortgage market), and even equities. However, the informational content of the NS factors was weaker in the case of commodity futures returns, where they failed to outperform even simple benchmarks.[17]

In conclusion, our results underscore the importance of considering the historical dynamics of the yield curve's level, slope, and curvature when forecasting financial returns, and opens up the possibility that the NS factors expressed by US interest rate data may appear in a structural way in the SDF pricing and hence display predictive power for the returns on all assets. Indeed, the NS factors may be conjectured to describe the yield curve's dynamics and, in turn, capture shifts in macroeconomic conditions that affect investment opportunities and hence, risk premia. These factors, if sufficiently systematic and pervasive, could be drivers of the SDF, which governs asset pricing through the fundamental condition

$$P_t = \mathbb{E}_t[m_{t+1} r_{t+1}], \tag{18}$$

where m_{t+1} is the SDF and r_{t+1} is the return on any traded asset. If the NS factors serve as state variables, because they may affect consumption growth, investor sentiment, or financial frictions, they may be embedded into the SDF representation as

$$m_{t+1} = a + b_0\beta_{0,t} + b_1\beta_{1,t} + b_2\beta_{2,t} + \varepsilon_{t+1} \tag{19}$$

Setting $r_{t+1} = 1$ in Equation (18) implies that P_t represents the short-term yield, linking the expectation of m_{t+1} directly to the short-term rate. Since this rate is largely determined by the sum of the level and slope factors, it follows that the stochastic discount factor is primarily influenced by β_0 and β_1, rather than the curvature factor β_2. This observation suggests that the stronger predictive power observed for financial asset returns likely stems from the combined effect of level and slope factors. Empirically, our results support this interpretation, as models incorporating all three NS factors outperformed those based on individual factors. This indicates that return predictability is not solely driven by curvature effects but rather emerges from broader interactions within the term structure.

This formulation implies that any movement in the yield curve factors propagates into expected returns across asset classes. As argued since Diebold and Li (2006), the level factor (β_0) captures the general risk-free rate and long-term discount rates, the slope factor (β_1) reflects changes in short-term interest rate expectations and monetary policy, and the curvature factor (β_2) accounts for medium-term risk adjustments and liquidity considerations. Consequently, fluctuations in these factors alter discount rates and expected excess returns, influencing not only bonds but also equities, real estate, currencies, and, in principle, alternative assets.

Our findings establish a potential link between NS factors and traditional asset pricing models, such as the Fama–French factor model and the consumption-based asset pricing model (C-CAPM). While the Fama–French framework attributes equity return variations to systematic, aggregate market risk exposure, and the C-CAPM links asset prices to consumption risk, our results suggest that NS factors may similarly encapsulate macroeconomic conditions shaping expected returns across different asset classes. This implies that the NS factors could either complement existing pricing factors and serve as alternative proxies for systematic risk or simply provide instruments that can predict business cycles and hence aggregate asset market valuations. For example, the role of the level factor (β_0) in capturing variations in long-term discount rates aligns with the role played by the market risk premium in the Fama–French model. The slope factor (β_1) is a well-known predictor of business cycle fluctuations, and therefore may help explain the cross-sectional variation in stock returns, particularly in relation to small-cap (SMB) and value (HML) stocks. Meanwhile, the curvature factor (β_2) may be linked to liquidity conditions and short-term risk dynamics, concepts widely explored in liquidity-based asset pricing models (see, e.g., Holmström and Tirole (2001)). From a C-CAPM perspective, the NS factors could act as proxies for time-varying discount rates embedded in the SDF. If yield curve capture conveys shifts in consumption growth expectations and intertemporal risk aversion, then variations in (β_0), (β_1), and (β_2) should help predict excess returns across asset classes, which is consistent with the bulk of our empirical findings.

This predictive ability of NS factors aligns with theoretical constraints on return forecastability. For instance, Potì (2018) established a tight bound on the predictability of portfolio returns using features of the SDF, and showed that out-of-sample predictability is constrained by the correlation between the SDF and the returns of any portfolio. In this framework, the upper bound on return predictability is given by

$$R^2_{OOS} \leq \rho^2(r_{t+1}, m_t)\sigma^2(m_{t+1}) \tag{20}$$

where $\rho(r_{t+1}, m_{t+1})$ denotes the correlation between the SDF m_{t+1} and the returns on some portfolio r_{t+1}, while $\sigma^2(m_{t+1})$ represents the unconditional variance in the SDF. If the NS factors significantly contribute to the predicted variation in m_{t+1}, then their ability to forecast asset returns is dictated by their empirical correlation with the pricing kernel.

Our empirical evidence shows that NS factors often exhibit high predictive power for short-term returns, particularly in fixed-income and equity markets, indicating that they effectively capture fluctuations in risk premia and discount rates. The observed return predictability, if within the derived theoretical bounds, would support the idea that NS factors serve as valid proxies for economic conditions embedded in the SDF. By influencing discount rates and expected excess returns across multiple markets, the NS factors may serve as state variables that should be explicitly incorporated into asset pricing models.

Author Contributions: M.G.: conceptualization, formal analysis, writing—review and editing, supervision, funding acquisition, investigation; S.I.: software, writing—original draft preparation, formal analysis, investigation. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the European Union—NextGenerationEU, in the framework of the GRINS—Growing Resilient, Inclusive and Sustainable project (GRINS PE000018—CUP B43C22000760006). The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

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

Appendix A. Diebold–Mariano and Model Confidence Set Tests

For yields with maturity m, the forecast error for model i at time t and forecast horizon h is defined as

$$e_{i,t}(m) = \hat{y}_{i,t+h/t}(m) - y_{t+h}(m) \tag{A1}$$

For returns of asset j, the forecast error for model i is given by

$$e_{j,i,t} = \hat{r}_{j,i,t+h/t} - r_{j,i,t+h} \tag{A2}$$

Using the squared forecast error as the loss function, the loss variables for yields and returns are respectively defined as

$$l_{i,t}(m) = (e_{i,t}(m))^2, \quad l_{j,i,t} = \left(e_{j,i,t}\right)^2 \tag{A3}$$

The loss differential between competing models i and k is defined as

$$d_{ik,t}(m) = l_{i,t}(m) - l_{k,t}(m), \quad d_{j,ik,t} = l_{j,i,t} - l_{j,k,t} \tag{A4}$$

The null hypothesis of equal predictive accuracy between models i and k is tested as follows

$$H_0 : \mathbb{E}[d_{ik,t}] = 0 \; \forall t \tag{A5}$$

against the alternative:

$$H_1 : \mathbb{E}[d_{ik,t}] \neq 0 \; \forall t \tag{A6}$$

Under the assumption that the loss differential series d_t is covariance stationary and short in memory, the DM test statistic follows a standard normal distribution:

$$DM = \frac{\bar{d}}{\sqrt{\frac{2\pi \hat{f}_0(0)}{T}}} \sim \mathcal{N}(0,1) \tag{A7}$$

where $\bar{d}$ is the average loss differential. The spectral density $\hat{f}_0(0)$ is estimated as follows:

For $h = 1$,

$$\hat{f}_0(0) = \frac{\hat{\gamma}_d(0)}{2\pi} \tag{A8}$$

For $h = 6$,

$$\hat{f}_0(0) = \frac{1}{2\pi}\left[\hat{\gamma}_d(0) + 2\sum_{k=1}^{5} \hat{\gamma}_d(k)\right] \tag{A9}$$

For $h = 12$,

$$\hat{f}_0(0) = \frac{1}{2\pi}\left[\hat{\gamma}_d(0) + 2\sum_{k=1}^{11} \hat{\gamma}_d(k)\right] \tag{A10}$$

where $\hat{\gamma}_d(k)$ is the k-th autocorrelation of the loss differential series d_t. The null hypothesis is tested for the models defined in Section 2.2 at a 5% significance level using a one-tailed test. A DM test statistic below -1.645 rejects the null hypothesis.

As regards the MSC procedure, starting from an initial set of models M, the procedure identifies a subset $M^* \subseteq M$ that, with a predetermined confidence level, includes the best forecasting model. Let $d_{ij,t}$ represent the loss differential between two models i and j, as defined above. Additionally, define the average loss differential between model i and all other models as

$$d_{ij,t} = \frac{\sum_{j \in M} d_{ij,t}}{m-1} \quad \text{for} \quad i,j = 1,\ldots,m \tag{A11}$$

The null hypothesis of equal predictive accuracy (EPA) can be tested in two ways. First, we can test that all pairwise loss differentials $d_{ij,t}$ are zero across all models:

$$H_{0,M} : \mathbb{E}(d_{ij,t}) = 0 \quad \text{for all} \quad i, j = 1, 2, \ldots, m \tag{A12}$$

against the alternative hypothesis that there is at least one non-zero pairwise loss differential:

$$H_{1,M} : \mathbb{E}(d_{ij,t}) \neq 0 \quad \text{for some} \quad i, j = 1, 2, \ldots, m \tag{A13}$$

Alternatively, we can test that the average loss differentials are zero for all models:

$$H_{0,M} : \mathbb{E}(d_{i\cdot,t}) = 0 \quad \text{for all} \quad i = 1, 2, \ldots, m \tag{A14}$$

against the alternative hypothesis that at least one average loss differential is non-zero:

$$H_{1,M} : \mathbb{E}(d_{i\cdot,t}) \neq 0 \quad \text{for some} \quad i = 1, 2, \ldots, m \tag{A15}$$

The following statistics are used to test these hypotheses:

$$t_{ij} = \frac{\bar{d}_{ij}}{\sqrt{\hat{\sigma}^2(d_{ij})}} \quad \text{and} \quad t_{i\cdot} = \frac{\bar{d}_{i\cdot}}{\sqrt{\hat{\sigma}^2(d_{i\cdot})}} \tag{A16}$$

where $\bar{d}_{ij} = \frac{\sum_{t=1}^{T} d_{ij,t}}{T}$ and $\bar{d}_{i\cdot} = \frac{\sum_{j \neq i} \bar{d}_{ij}}{m-1}$, while $\hat{\sigma}^2(d_{ij})$ and $\hat{\sigma}^2(d_{i\cdot})$ are bootstrapped estimates of the variances $\sigma^2(d_{ij})$ and $\sigma^2(d_{i\cdot})$, respectively.

The two test statistics for these hypotheses are

$$T_{R,M} = \max_{i,j \in M} |t_{ij}| \quad \text{and} \quad T_{\max,M} = \max_{i \in M} t_{i\cdot} \tag{A17}$$

These test statistics are non-normally distributed, and their asymptotic distributions are estimated using the bootstrap. The procedure was carried out sequentially, eliminating models one at a time from the superior set until the EPA hypothesis was not rejected.

Models were eliminated according to the following rules:

$$e_{R,M} = \arg\max_{i} \left\{ \sup_{j \in M} t_{ij} \right\} \quad \text{and} \quad e_{\max,M} = \arg\max_{i \in M} t_{i\cdot} \tag{A18}$$

The number of bootstrapped samples was set to 1000, the quadratic loss function was used, and the confidence level was $1 - \alpha = 0.9$. The MCS procedure was applied for models forecasting yields (models I to VIII in Section 2.2), for each forecasting horizon (1-step-ahead, 6-step-ahead, and 12-step-ahead), and for each yield analyzed (3 months, 1 year, 5 years, and 10 years). The procedure was also applied for models forecasting returns (models I to IV in Section 2.3), for each of the three forecasting horizons (1-step-ahead, 6-step-ahead, and 12-step-ahead), and for each portfolio analyzed across various asset classes (equity, US Treasury securities, corporate bonds, REITs, and commodity futures contracts).

Appendix B. Summary Statistics

Table A1. Summary statistics for target return series.

Panel A: Treasury Rates	Observations	Mean	Median	Minimum	Maximum	St. Dev	Skewness
3 Months	372	2.692	2.225	0.000	8.070	2.286	0.341
6 Months	372	2.818	2.360	0.030	8.440	2.320	0.324
1 Year	372	2.943	2.475	0.090	8.580	2.326	0.303

Table A1. *Cont.*

Panel A: Treasury Rates	Observations	Mean	Median	Minimum	Maximum	St. Dev	Skewness
2 Years	372	3.230	2.865	0.110	8.960	2.354	0.289
3 Years	372	3.441	3.130	0.110	9.050	2.312	0.276
5 Years	372	3.831	3.710	0.210	9.040	2.203	0.248
7 Years	372	4.135	4.015	0.390	9.060	2.115	0.243
10 Years	372	4.370	4.335	0.550	9.040	2.012	0.238
Panel B: Treasury Returns	**Observations**	**Mean**	**Median**	**Minimum**	**Maximum**	**St. Dev**	**Skewness**
3 Months	371	0.226	0.193	−0.017	0.698	0.195	0.377
6 Months	371	0.242	0.197	−0.060	0.876	0.217	0.556
1 Year	371	0.445	0.205	−5.024	12.214	1.985	1.326
2 Years	371	0.483	0.336	−6.957	9.110	2.327	0.355
3 Years	371	0.505	0.398	−8.362	8.550	2.480	0.065
5 Years	371	0.534	0.485	−8.637	9.317	2.542	−0.025
7 Years	371	0.551	0.480	−8.611	9.892	2.481	−0.064
10 Years	371	0.562	0.413	−8.346	11.054	2.380	0.085
Panel C	**Observations**	**Mean**	**Median**	**Minimum**	**Maximum**	**St. Dev**	**Skewness**
Fama and French Factor Portfolios							
SMB	370	0.092	0.115	−16.720	21.130	3.143	0.633
HML	370	0.191	−0.050	−9.050	14.110	3.074	0.761
Mkt	370	0.891	1.385	−17.150	13.650	4.352	−0.641
Industry Portfolios							
Food	371	0.913	1.020	−14.560	15.140	3.962	−0.286
Mines	371	0.846	0.740	−32.740	22.450	7.919	−0.284
Oil	371	0.701	0.670	−34.810	32.820	6.280	0.003
Textiles	371	1.040	1.280	−22.140	23.270	6.150	−0.216
Consumer Durables	371	0.808	1.070	−25.770	29.240	5.777	−0.217
Chemicals	371	0.897	1.240	−21.930	22.300	5.886	−0.193
Consumer	371	1.001	1.380	−10.130	15.630	4.156	−0.179
Construction	371	1.137	1.510	−20.200	17.920	5.871	−0.248
Steel	371	0.669	0.710	−32.910	30.670	8.333	−0.230
Fabricated Products	371	1.033	1.360	−23.030	18.820	5.473	−0.496
Machinery	371	1.178	1.670	−28.320	19.320	6.938	−0.514
Cars	371	1.097	0.980	−28.430	38.710	7.194	0.536
Transportation	371	1.023	1.510	−22.680	19.700	5.196	−0.621
Utilities	371	0.799	1.230	−13.020	11.720	3.965	−0.645
Retail Stores	371	1.080	1.010	−14.590	18.240	4.776	−0.070
Financials	371	0.973	1.590	−22.100	17.100	5.647	−0.623
Other	371	0.918	1.600	−17.790	15.310	4.941	−0.489
Size and Value							
Ptf 1	370	0.500	0.937	−27.845	38.321	7.982	0.209
Ptf 2	370	1.080	1.454	−21.521	42.421	7.013	0.442
Ptf 3	370	0.982	1.370	−21.929	21.057	5.740	−0.344
Ptf 4	370	1.159	1.470	−26.786	26.369	5.615	−0.344
Ptf 5	370	1.166	1.330	−27.820	16.381	5.801	−0.721
Ptf 6	370	0.875	1.527	−25.775	28.115	7.024	−0.308
Ptf 7	370	1.045	1.581	−23.116	18.785	5.775	−0.518
Ptf 8	370	1.042	1.529	−21.257	16.062	5.191	−0.661
Ptf 9	370	0.974	1.788	−23.772	16.193	5.295	−0.822
Ptf 10	370	1.025	1.771	−32.144	26.019	6.333	−0.776
Ptf 11	370	0.929	1.721	−23.097	23.130	6.485	−0.436
Ptf 12	370	1.091	1.454	−21.509	18.969	5.335	−0.418
Ptf 13	370	0.978	1.465	−17.847	16.757	4.879	−0.555
Ptf 14	370	1.077	1.522	−26.854	16.936	5.168	−0.752
Ptf 15	370	1.055	1.382	−31.175	19.058	5.885	−0.815
Ptf 16	370	1.116	1.380	−20.840	26.044	5.816	−0.175
Ptf 17	370	1.054	1.364	−20.165	15.899	4.873	−0.713
Ptf 18	370	0.931	1.462	−25.327	15.515	5.010	−0.867
Ptf 19	370	1.002	1.604	−31.427	16.233	5.025	−1.128
Ptf 20	370	0.899	1.832	−32.802	18.336	5.830	−0.961
Ptf 21	370	1.013	1.078	−14.776	15.356	4.437	−0.298
Ptf 22	370	0.911	1.313	−17.049	13.691	4.168	−0.540
Ptf 23	370	0.931	1.469	−17.308	12.570	4.246	−0.623
Ptf 24	370	0.605	1.151	−27.164	15.878	5.025	−1.209
Ptf 25	370	0.846	1.330	−28.401	21.494	6.218	−0.572
Operating Profits and Investments							
Ptf 26	371	0.929	1.305	−25.258	27.048	6.967	−0.085
Ptf 27	371	0.629	1.009	−25.572	15.998	5.420	−0.599
Ptf 28	371	0.970	1.266	−31.152	18.713	5.712	−0.764
Ptf 29	371	0.924	1.884	−26.366	18.620	6.106	−0.711

Table A1. *Cont.*

Panel C	Observations	Mean	Median	Minimum	Maximum	St. Dev	Skewness
Operating Profits and Investments							
Ptf 30	371	0.485	1.201	−27.414	18.763	6.832	−0.598
Ptf 31	371	0.945	1.098	−19.474	21.352	5.180	−0.238
Ptf 32	371	1.010	1.486	−20.912	17.886	4.983	−0.383
Ptf 33	371	0.824	1.244	−24.368	18.783	5.243	−0.920
Ptf 34	371	1.004	1.255	−20.014	23.141	5.258	−0.163
Ptf 35	371	0.822	1.189	−20.749	27.256	5.612	−0.180
Ptf 36	371	0.932	1.393	−23.239	15.767	5.105	−0.572
Ptf 37	371	1.032	1.269	−22.021	16.520	4.746	−0.522
Ptf 38	371	0.990	1.433	−20.939	13.858	4.505	−0.781
Ptf 39	371	0.975	1.296	−21.704	16.741	4.745	−0.643
Ptf 40	371	0.863	1.700	−24.746	16.446	5.825	−0.714
Ptf 41	371	1.241	1.599	−21.235	15.446	4.563	−0.449
Ptf 42	371	1.034	1.227	−13.247	15.350	3.826	−0.171
Ptf 43	371	0.947	1.201	−14.485	12.952	4.242	−0.413
Ptf 44	371	1.077	1.569	−17.855	18.125	4.613	−0.447
Ptf 45	371	1.103	1.342	−23.784	18.896	5.565	−0.412
Ptf 46	371	1.184	1.534	−17.225	15.081	4.588	−0.258
Ptf 47	371	0.973	1.292	−19.068	17.008	4.364	−0.356
Ptf 48	371	0.914	1.142	−17.249	17.431	4.253	−0.354
Ptf 49	371	1.103	1.233	−16.588	12.851	4.537	−0.352
Size and Cash flow							
Ptf 50	371	1.170	1.349	−20.308	20.017	5.884	−0.332
Ptf 51	371	0.992	1.690	−21.650	19.520	5.876	−0.375
Ptf 52	371	1.125	1.720	−22.920	16.670	4.961	−0.709
Ptf 53	371	1.122	1.690	−31.130	21.000	5.663	−0.816
Ptf 54	371	1.024	1.390	−15.670	14.360	4.522	−0.355
Size and Dividend yield							
Ptf 55	371	0.898	1.350	−15.380	12.570	4.139	−0.573
Ptf 56	371	0.935	1.350	−23.610	15.190	4.594	−0.816
Ptf 57	371	1.103	1.800	−20.890	15.930	5.180	−0.646
Ptf 58	371	1.010	1.560	−20.440	16.210	4.821	−0.633
Ptf 59	371	0.992	1.320	−23.460	21.840	4.644	−0.608
Size and Price to Earnings ratio							
Ptf 60	371	0.894	1.210	−20.310	16.100	4.861	−0.492
Ptf 61	371	0.968	1.230	−15.550	12.900	4.063	−0.555
Ptf 62	371	0.868	1.030	−16.750	14.960	4.072	−0.661
Ptf 63	371	1.000	1.700	−22.510	18.860	5.928	−0.415
Ptf 64	371	1.097	1.660	−22.890	17.030	4.873	−0.679
Ptf 65	371	1.138	1.640	−30.890	20.050	5.536	−0.844
Ptf 66	371	0.955	1.230	−15.940	13.800	4.451	−0.375
Ptf 67	371	0.930	1.250	−15.470	13.100	4.104	−0.511
Ptf 68	371	1.046	1.580	−23.590	15.950	4.788	−0.829
Ptf 69	371	0.573	0.610	−29.210	47.520	8.846	0.858
Ptf 70	371	0.897	1.280	−23.550	27.550	5.758	−0.308
Ptf 71	371	1.137	1.670	−23.430	22.480	5.190	−0.538
Ptf 72	371	1.333	1.690	−24.340	20.730	5.157	−0.548
Ptf 73	371	1.606	2.190	−22.830	31.850	6.619	−0.132
Ptf 74	371	0.737	0.760	−27.730	55.850	8.728	0.607
Ptf 75	371	1.037	1.480	−25.210	33.740	6.095	−0.156
Ptf 76	371	1.122	1.620	−22.280	18.670	5.156	−0.607
Ptf 77	371	1.218	1.630	−23.070	17.670	5.183	−0.565
Ptf 78	371	1.404	1.780	−25.600	30.160	6.799	−0.132
Ptf 79	371	0.790	0.890	−27.530	44.280	8.438	0.417
Ptf 80	371	0.970	1.440	−25.020	27.060	5.820	−0.350
Ptf 81	371	1.047	1.350	−22.200	23.520	5.012	−0.482
Ptf 82	371	1.035	1.490	−22.110	16.620	4.808	−0.637
Size and Momentum							
Ptf 83	371	1.297	1.640	−21.460	26.390	6.215	−0.267
Ptf 84	371	0.649	0.620	−38.490	46.420	8.554	0.133
Ptf 85	371	1.016	1.420	−23.550	28.880	5.738	−0.161
Ptf 86	371	1.092	1.610	−22.240	17.480	4.802	−0.580
Ptf 87	371	1.075	1.410	−19.390	12.980	4.378	−0.719
Ptf 88	371	1.245	1.430	−22.170	25.080	5.631	−0.338
Ptf 89	371	0.600	0.620	−31.450	30.950	7.720	0.163
Ptf 90	371	0.871	0.930	−20.640	24.370	5.144	−0.003
Ptf 91	371	0.899	1.130	−16.890	17.540	4.211	−0.329
Ptf 92	371	0.964	1.210	−14.190	15.800	4.019	−0.263
Ptf 93	371	1.111	1.630	−19.610	18.520	4.920	−0.381

Table A1. *Cont.*

Panel D	**Observations**	**Mean**	**Median**	**Minimum**	**Maximum**	**St. Dev**	**Skewness**
Diversified	321	0.699	1.262	−31.960	39.687	6.181	−0.575
Industrial	321	1.232	1.528	−56.188	70.483	8.254	0.264
Lodging/Resorts	321	0.570	0.590	−36.555	67.525	8.710	0.638
Office	321	0.895	1.480	−31.796	32.458	6.077	−0.487
Residential	321	1.013	1.362	−26.656	22.242	5.501	−0.833
Retail	321	0.846	1.086	−42.678	43.516	6.701	−0.985
Self Storage	321	1.326	1.616	−22.244	21.928	5.584	−0.364
All REITs	321	0.888	1.194	−30.226	27.975	5.296	−0.959
Mortgage REITs	321	0.642	1.595	−53.753	19.411	6.551	−2.455
Panel E	**Observations**	**Mean**	**Median**	**Minimum**	**Maximum**	**St. Dev**	**Skewness**
AAA	372	0.536	0.586	−6.544	6.391	1.473	−0.185
AA	372	0.586	0.565	−5.774	5.357	1.339	−0.258
BBB	372	0.531	0.671	−10.996	6.286	1.668	−1.575
A	372	0.565	0.649	−9.672	6.750	1.524	−1.063
Panel F	**Observations**	**Mean**	**Median**	**Minimum**	**Maximum**	**St. Dev**	**Skewness**
Gold	372	0.533	0.115	−18.005	16.458	4.430	0.174
Silver	372	0.866	0.235	−27.958	30.462	8.371	0.199
Platinum	372	0.417	0.309	−42.004	24.112	6.286	−0.854
Copper	372	0.582	0.421	−36.150	34.064	7.181	0.044
Brent Crude	372	0.660	0.829	−54.988	40.074	9.540	−0.361
Gasoil	372	0.557	0.566	−33.617	31.172	9.286	−0.092
Light crude oil	372	0.703	0.924	−54.245	88.376	10.485	1.198
Natural Gas	372	1.132	0.512	−41.616	62.613	15.197	0.568
Cotton	372	0.385	0.511	−36.116	24.749	8.492	−0.222
Cocoa	372	0.585	0.201	−28.083	34.565	8.921	0.457
Coffee	372	0.638	−0.727	−30.391	50.595	10.622	1.007
Corn	372	0.565	0.081	−31.380	22.190	7.976	−0.185
Lumber	372	1.091	0.485	−34.030	58.416	11.205	0.674
Soybean Oil	372	0.444	0.242	−24.461	26.862	6.933	0.039
Soybeans	372	0.485	0.434	−32.867	19.573	7.010	−0.491
Wheat	372	0.637	0.204	−25.248	42.335	8.608	0.516
Live Cattle	372	0.264	0.017	−25.565	40.342	5.652	0.437

The table contains summary statistics for the sample of monthly data used in the analysis. The sample was from January 1990 to December 2020. Each panel represents a different asset class. Panel A shows descriptive statistics of US Constant Maturity Treasury rates, taken to be representative of the US yield curve, at different maturities. Panel B shows statistics for the returns on US Treasury securities. Panels C, D, E, and F show statistics for equities, REIT, corporate bonds, and commodity returns, respectively. In particular, Panel C is split into eight subsections, showing statistics for the returns on the original Fama–French factor portfolios, the 17 industry portfolios, and various bivariate types on size, book-to-market, investments, operating profitability, cash flows, dividend yield, earnings price ratio, and momentum. Values are reported in percentage points.

Appendix C. Diebold–Mariano Test on Asset Return Series

Table A2. Diebold–Mariano test on return series.

Panel A—Diebold–Mariano test comparing AR(1) on NS to listed models									
	VAR(1) on NS			**AR(1) for Returns**			**Historical Average**		
	***h* = 1**	***h* = 6**	***h* = 12**	***h* = 1**	***h* = 6**	***h* = 12**	***h* = 1**	***h* = 6**	***h* = 12**
3 Months	7.283	1.736	1.517	**−1.826**	**−2.140**	**−1.875**	**−17.772**	**−5.762**	**−3.419**
6 Months	6.760	2.771	2.077	−0.786	**−4.413**	**−2.878**	**−16.321**	**−5.688**	**−3.373**
1 Year	4.887	2.489	1.046	0.940	−0.026	−0.082	−0.331	−0.054	−0.089
2 Years	4.056	2.256	1.130	0.903	0.405	−0.026	0.384	0.402	−0.021
3 Years	3.059	1.628	0.786	0.529	0.328	−0.195	0.323	0.325	−0.193
5 Years	1.391	0.374	−0.185	−0.294	−0.041	−0.677	−0.323	−0.045	−0.664
7 Years	0.559	−0.208	−0.630	−0.595	−0.177	−0.924	−0.810	−0.186	−0.896
10 Years	−0.335	−0.749	−0.967	−0.909	−0.143	−0.975	−1.353	−0.162	−0.945

Table A2. *Cont.*

Panel B—Diebold–Mariano test comparing AR(1) on NS vs. listed models									
	VAR(1) on NS			AR(1) for Returns			Historical Average		
	$h = 1$	$h = 6$	$h = 12$	$h = 1$	$h = 6$	$h = 12$	$h = 1$	$h = 6$	$h = 12$
SMB	0.126	−0.542	−1.039	−0.245	0.006	−0.113	−1.137	0.025	−0.059
HML	1.314	1.387	1.107	−1.489	−1.039	**−2.176**	**−2.247**	**−1.869**	−1.048
Mkt	0.472	1.218	1.220	**−1.922**	2.333	1.606	−1.395	0.745	2.549
Food	2.698	2.504	1.742	0.930	1.613	1.606	0.945	1.679	1.659
Mines	**−2.193**	−1.363	−1.399	1.068	2.369	2.022	1.184	2.364	2.030
Oil	−1.573	**−1.955**	−0.684	0.192	0.460	0.935	−1.081	−0.519	1.048
Textiles	2.530	2.341	1.697	−0.368	1.777	1.735	0.442	1.759	1.835
Consumer Durables	**−2.381**	−1.343	0.545	0.592	1.017	1.302	−0.038	1.077	1.366
Chemicals	1.407	1.458	1.039	−0.490	0.644	1.001	−0.351	−0.675	1.085
Steel	1.068	1.902	**−1.839**	**−2.504**	0.554	1.156	−0.902	−1.129	−0.678
Consumer	0.826	−0.061	−0.094	0.592	1.017	1.302	−0.038	1.077	1.366
Construction	0.826	−0.061	−0.094	0.592	1.017	1.302	−0.038	1.077	1.366
Steel	1.474	1.525	0.807	−0.435	1.696	2.370	−0.848	−1.662	2.197
Fabricated Products	−1.408	−1.357	−1.381	**−2.364**	1.399	1.588	−0.957	1.496	−1.670
Machinery	1.263	1.027	0.209	−0.567	1.386	1.364	−0.723	**−1.931**	−1.384
Cars	−1.498	−1.579	−1.417	−1.068	1.227	1.525	−1.566	1.249	1.544
Transportation	1.397	1.282	1.070	−1.066	1.209	1.756	−0.686	1.235	1.904
Utilities	−0.712	−1.517	**−2.243**	−2.775	1.277	1.254	**−2.444**	**−1.877**	−1.087
Retail Stores	1.269	1.430	1.833	**−2.512**	1.234	1.634	**−1.742**	1.092	1.643
Financials	2.085	2.302	2.081	0.210	1.304	1.724	0.293	1.322	1.770
Other	−0.728	−0.042	0.631	**−2.162**	−0.607	1.300	**−2.668**	−1.443	1.034
Ptf 1	−1.042	−0.875	−0.369	−1.516	0.655	2.206	**−3.251**	0.537	2.125
Ptf 2	−0.800	−0.970	−1.421	**−2.910**	0.702	1.622	**−1.624**	−0.732	1.718
Ptf 3	−0.656	−0.750	−0.769	**−2.638**	0.126	0.981	**−1.943**	0.170	1.060
Ptf 4	−1.106	−1.035	−0.711	**−2.329**	−0.243	0.221	−0.944	−0.208	0.267
Ptf 5	−1.107	−1.126	−0.897	**−1.974**	−0.305	0.234	**−1.870**	−0.259	0.292
Ptf 6	−1.461	−0.863	−0.715	**−2.311**	1.153	2.646	**−2.678**	1.091	2.213
Ptf 7	0.410	0.578	−0.046	−1.534	1.470	0.902	**−2.169**	1.388	0.773
Ptf 8	−0.857	−0.637	−0.116	**−1.807**	0.067	0.541	**−1.868**	0.145	0.608
Ptf 9	−0.988	−0.958	−0.691	**−2.272**	−0.424	0.112	−1.144	−0.375	0.175
Ptf 10	−0.353	−0.442	−0.515	**−2.196**	−0.125	0.724	−1.203	−0.047	−0.885
Ptf 11	−1.282	−1.189	−1.529	−1.493	1.198	2.775	**−1.886**	1.228	3.369
Ptf 12	−1.243	−0.533	0.027	**−1.797**	0.654	1.309	−0.895	−0.715	1.448
Ptf 13	−0.337	0.038	−0.277	**−2.253**	−0.058	0.709	**−1.799**	0.074	0.939
Ptf 14	−0.742	0.196	0.490	−1.419	0.375	0.770	−0.525	−0.438	0.845
Ptf 15	0.046	0.156	−0.111	**−2.848**	−1.122	−0.557	**−1.719**	−1.055	−0.499
Ptf 16	−1.201	−0.746	−0.997	**−1.978**	0.854	3.021	**−2.192**	0.790	3.791
Ptf 17	−0.422	−0.352	−0.067	**−1.749**	1.401	1.541	**−2.793**	1.246	3.655
Ptf 18	−0.454	0.568	0.206	**−1.764**	0.678	1.827	−1.337	0.824	2.230
Ptf 19	−0.992	−1.519	−0.867	**−2.611**	−0.730	−0.021	**−1.821**	−0.659	0.083
Ptf 20	0.626	1.153	1.022	−1.359	0.453	0.976	−0.596	0.479	1.059
Ptf 21	1.473	2.740	1.991	**−1.669**	0.307	1.786	**−1.669**	0.151	1.706
Ptf 22	1.714	2.310	1.483	0.170	1.420	1.812	−0.092	1.425	1.888
Ptf 23	0.421	0.096	0.251	**−2.148**	0.134	1.715	−1.478	−0.205	2.009
Ptf 24	1.251	0.664	0.471	−0.743	0.895	1.388	−0.368	0.930	1.471
Ptf 25	1.097	0.768	0.838	−0.521	0.824	1.455	**−1.851**	0.833	1.509
Ptf 26	−1.036	−0.979	−0.027	−1.200	1.202	1.594	**−2.065**	1.148	1.483
Ptf 27	−0.432	0.398	2.164	−0.917	0.701	1.209	−0.667	0.712	1.280
Ptf 28	2.111	1.729	1.234	−0.718	0.782	1.019	−0.104	0.798	1.033
Ptf 29	−0.572	−1.026	−1.557	**−2.588**	0.352	2.803	**−2.962**	−0.285	3.126
Ptf 30	−1.391	−0.508	−0.470	**−1.806**	−0.693	−0.189	**−2.648**	−0.842	0.287
Ptf 31	0.110	0.357	−0.005	**−2.030**	**−4.420**	−0.751	**−1.676**	**−4.702**	−0.713
Ptf 32	−1.084	**−3.046**	−1.119	−1.280	−0.560	0.632	**−1.904**	−0.519	0.783
Ptf 33	−0.158	−0.226	−0.373	−1.210	−0.088	0.400	**−2.192**	−0.157	0.329
Ptf 34	0.295	0.865	2.266	−1.154	1.115	3.148	**−1.695**	1.112	2.939
Ptf 35	−1.499	**−2.748**	**−2.188**	−1.455	0.147	1.578	−1.325	−0.009	1.453
Ptf 36	−0.679	0.379	1.114	−1.394	1.406	5.678	**−1.915**	1.463	5.460
Ptf 37	0.714	0.615	0.486	−0.812	−0.246	0.248	−0.835	−0.175	0.323
Ptf 38	−0.758	0.455	0.382	−1.869	1.210	2.144	**−2.526**	1.017	1.870
Ptf 39	3.157	2.732	2.057	1.301	2.064	2.237	0.951	2.083	2.283
Ptf 40	−0.319	−0.968	−1.168	**−1.693**	−0.805	−0.052	**−2.180**	−0.830	−0.082
Ptf 41	0.392	0.942	0.603	−0.603	1.098	1.149	−0.908	1.090	1.157
Ptf 42	−0.727	−0.543	−0.174	−0.974	0.595	2.561	**−2.044**	0.647	2.729
Ptf 43	0.655	0.174	0.421	−0.637	0.600	1.712	−1.266	0.662	1.985
Ptf 44	−0.799	−0.273	0.239	**−2.300**	0.876	3.570	**−2.645**	0.884	3.764
Ptf 45	−1.234	**−3.087**	**−2.192**	−0.644	0.666	4.135	**−2.237**	0.497	4.996
Ptf 46	3.401	2.968	2.602	1.590	2.107	2.423	1.252	2.167	2.472

Table A2. *Cont.*

Panel B—Diebold–Mariano test comparing AR(1) on NS vs. listed models									
	VAR(1) on NS			AR(1) for Returns			Historical Average		
	h = 1	*h* = 6	*h* = 12	*h* = 1	*h* = 6	*h* = 12	*h* = 1	*h* = 6	*h* = 12
Ptf 47	3.901	4.542	4.220	1.632	2.609	4.120	1.646	2.599	4.209
Ptf 48	0.423	−0.003	0.120	**−1.838**	0.876	1.364	**−1.862**	0.951	1.551
Ptf 49	−0.191	0.074	0.414	−1.104	0.831	1.579	**−2.720**	0.824	1.669
Ptf 50	0.864	1.174	1.042	−1.441	1.772	2.319	**−2.431**	1.672	2.646
Ptf 51	−0.816	−0.897	−1.052	**−2.446**	0.652	2.013	**−1.861**	−0.720	2.253
Ptf 52	−0.717	−0.617	−0.611	**−2.865**	−0.231	0.326	−1.236	−0.147	0.430
Ptf 53	−0.167	0.094	−0.064	**−2.077**	−0.404	0.179	**−1.763**	−0.310	0.343
Ptf 54	1.091	1.944	1.818	**−1.892**	0.284	1.620	**−1.835**	0.180	1.602
Ptf 55	1.830	1.968	1.374	−0.031	1.395	2.055	−0.017	1.402	2.163
Ptf 56	0.528	−0.106	−0.019	−1.314	0.217	1.081	−1.004	−0.282	1.233
Ptf 57	−0.742	−0.627	−0.521	**−2.514**	0.073	0.750	−1.260	0.169	0.909
Ptf 58	−1.143	−1.136	−0.893	**−2.518**	0.196	0.710	**−1.842**	0.247	0.803
Ptf 59	**−1.942**	−1.436	−0.650	**−1.676**	0.544	0.743	−0.240	0.588	0.804
Ptf 60	1.476	1.782	1.649	−0.716	1.099	1.806	−0.519	1.107	1.867
Ptf 61	1.224	1.042	0.967	−0.076	1.001	2.157	−0.760	0.993	2.301
Ptf 62	1.042	0.545	0.204	**−1.845**	0.227	0.948	−0.677	0.299	1.075
Ptf 63	−0.498	−0.570	−0.923	**−2.505**	0.496	1.956	**−1.849**	0.581	2.327
Ptf 64	−0.762	−0.562	−0.491	**−2.937**	−0.079	0.564	−1.298	−0.008	0.686
Ptf 65	−0.581	−0.445	−0.451	**−2.253**	−0.397	0.174	**−1.930**	−0.335	0.297
Ptf 66	1.324	2.013	1.910	−1.603	0.824	1.848	−1.250	0.780	1.887
Ptf 67	1.307	1.409	1.063	−0.253	0.990	2.175	−0.610	0.981	2.323
Ptf 68	1.101	0.817	0.638	−0.777	0.521	1.096	−0.655	0.568	1.188
Ptf 69	−0.111	−0.283	−0.398	**−1.807**	1.030	1.494	**−1.932**	0.972	1.371
Ptf 70	−1.116	−1.208	−0.735	**−2.382**	0.443	0.752	−0.774	0.443	0.793
Ptf 71	−1.498	**−1.899**	−1.178	**−2.323**	0.621	1.462	−0.584	0.641	1.559
Ptf 72	−0.960	−1.210	−1.001	**−2.172**	−0.268	0.414	−1.021	−0.199	0.490
Ptf 73	**−2.198**	−1.564	−1.096	−0.993	0.388	0.758	−0.202	−0.426	0.801
Ptf 74	1.744	1.638	0.936	−1.151	1.580	1.435	0.165	1.518	1.379
Ptf 75	1.075	1.237	0.719	**−1.727**	1.224	0.937	−0.869	1.169	0.873
Ptf 76	−0.516	−0.138	−0.254	**−2.455**	0.233	1.672	−1.551	0.312	1.951
Ptf 77	−0.949	−0.871	−0.824	**−2.308**	−0.142	0.409	**−1.901**	−0.050	0.495
Ptf 78	−1.170	−0.932	−1.037	−1.031	0.166	0.830	−0.831	0.252	0.914
Ptf 79	0.737	0.563	0.128	−1.576	1.334	1.634	−1.342	1.294	1.475
Ptf 80	−0.205	0.183	−0.011	**−2.024**	0.023	1.820	**−2.325**	−0.008	1.401
Ptf 81	1.066	1.040	1.062	**−1.871**	0.422	0.783	−0.637	−0.483	0.881
Ptf 82	−1.342	0.049	0.451	−0.290	0.493	0.730	−0.198	0.563	0.796
Ptf 83	−1.518	1.009	0.683	−0.075	0.674	1.261	−0.195	0.745	1.317
Ptf 84	−1.058	−0.955	−1.135	**−1.717**	0.655	1.411	**−1.820**	0.708	1.592
Ptf 85	−0.515	0.147	0.219	**−2.104**	0.307	1.354	−1.506	0.307	1.562
Ptf 86	0.602	0.387	0.536	−1.522	0.612	1.581	−1.261	0.674	1.862
Ptf 87	0.364	0.160	0.322	−1.500	0.592	1.944	**−1.749**	−0.687	2.414
Ptf 88	−1.199	−0.564	−0.862	−0.562	0.636	1.475	−0.749	0.707	1.565
Ptf 89	1.467	1.768	1.516	−1.580	1.298	2.332	−0.615	1.291	2.392
Ptf 90	−0.536	−0.323	0.125	**−1.901**	0.022	1.081	**−2.532**	−0.022	1.036
Ptf 91	0.895	1.284	1.055	**−2.489**	0.151	2.307	**−2.666**	−0.143	2.191
Ptf 92	0.170	0.218	0.617	−0.508	0.246	1.882	**−1.812**	−0.232	2.093
Ptf 93	1.514	2.127	1.807	0.048	1.230	1.880	−0.436	1.257	1.981

Panel C—Diebold–Mariano test comparing AR(1) on NS vs. listed models									
	VAR(1) on NS			AR(1) for Returns			Historical Average		
	h = 1	*h* = 6	*h* = 12	*h* = 1	*h* = 6	*h* = 12	*h* = 1	*h* = 6	*h* = 12
Diversified	−0.669	−0.527	1.117	**−2.594**	−0.534	−1.274	**−1.973**	−0.515	−1.269
Industrial	−0.576	0.370	1.847	−0.371	0.349	0.390	−1.499	−0.353	0.375
Lodging/Resorts	−0.994	0.204	0.710	**−1.736**	−0.142	−0.337	**−1.824**	−0.139	−0.323
Office	−0.885	−0.057	−0.064	**−2.286**	−1.595	−0.803	**−1.701**	−1.598	−0.774
Residential	−0.204	−0.116	0.449	**−1.676**	−0.698	−1.058	**−2.513**	−0.666	−1.054
Retail	**−1.580**	−0.668	−0.362	**−1.893**	−0.883	**−1.768**	−1.178	−0.877	−1.586
Self Storage	−0.363	0.963	0.036	−0.131	0.058	−0.457	**−2.379**	−0.063	−0.425
All REITs	−0.572	0.158	1.728	**−2.790**	0.506	−0.974	**−1.824**	0.539	−0.916
Mortgage REITs	0.243	0.850	1.405	0.992	0.973	0.267	0.091	1.107	0.290

Table A2. *Cont.*

Panel D—Diebold–Mariano test comparing AR(1) vs. NS to listed models									
	VAR(1) on NS			**AR(1) for Returns**			**Historical Average**		
	***h* = 1**	***h* = 6**	***h* = 12**	***h* = 1**	***h* = 6**	***h* = 12**	***h* = 1**	***h* = 6**	***h* = 12**
AAA	0.720	0.357	0.386	−1.169	−0.076	−0.314	−1.117	−0.072	−0.290
AA	1.832	0.472	−0.236	−1.113	−0.364	−0.850	−0.430	−0.354	−0.829
BBB	1.482	0.178	−0.169	−0.954	0.012	−0.350	−0.217	0.020	−0.302
A	1.505	0.219	−0.340	**−1.884**	−0.291	−0.710	−0.484	−0.289	−0.683
Panel E—Diebold–Mariano test comparing AR(1) on NS vs. listed models									
	VAR(1) on NS			**AR(1) for Returns**			**Historical Average**		
	***h* = 1**	***h* = 6**	***h* = 12**	***h* = 1**	***h* = 6**	***h* = 12**	***h* = 1**	***h* = 6**	***h* = 12**
Gold	−1.159	−1.594	**−1.839**	0.787	1.858	1.572	0.607	1.683	1.570
Silver	−0.100	−1.614	−1.032	0.152	5.875	2.313	−0.136	3.118	2.166
Platinum	2.183	2.147	2.748	1.460	2.563	5.188	1.893	2.549	4.752
Copper	−0.346	−0.214	−0.283	0.515	2.712	2.393	1.747	2.605	2.412
Brent Crude	0.785	0.171	−0.004	1.339	2.513	2.538	0.729	2.498	2.458
Gasoil	0.660	0.081	−0.104	1.745	1.892	2.567	0.907	1.725	2.438
Light crude oil	0.704	0.036	-0.063	0.863	2.292	3.414	0.238	2.386	2.919
Natural Gas	0.662	−1.216	0.131	0.205	−1.102	1.434	−0.059	−1.539	1.604
Cotton	−1.218	−1.386	−0.857	1.402	1.283	3.114	1.158	1.201	3.421
Cocoa	0.451	−0.007	−0.206	1.821	−0.763	0.437	−0.362	−0.871	0.535
Coffee	**−2.830**	**−1.829**	−0.898	0.068	0.411	1.632	−0.295	0.104	1.685
Corn	−0.061	0.470	3.181	0.325	−0.647	1.662	−1.077	−1.144	1.776
Lumber	0.123	0.540	1.536	−0.377	−0.213	2.186	−0.795	−0.544	2.015
Soybean Oil	−0.660	−0.299	−0.354	1.144	1.306	2.033	0.226	0.501	1.983
Soybeans	0.085	0.303	1.393	0.191	−0.439	1.861	−0.521	−0.683	2.062
Wheat	0.460	−0.752	−0.351	0.926	−0.848	2.093	−0.802	−0.918	2.059
Live Cattle	0.923	0.986	−0.020	1.339	0.746	1.791	0.410	0.525	1.882
Panel F—Summary of Diebold–Mariano test results									
	VAR(1) on NS			**AR(1) for Returns**			**Historical Average**		
	***h* = 1**	***h* = 6**	***h* = 12**	***h* = 1**	***h* = 6**	***h* = 12**	***h* = 1**	***h* = 6**	***h* = 12**
Equity	4	5	4	52	1	1	49	4	0
Treasuries	0	0	0	1	2	2	2	2	2
REITs	1	0	0	6	0	1	6	0	0
Corporate Bonds	0	0	0	1	0	0	0	0	0
Commodities	1	1	1	0	0	0	0	0	0

The table contains results of the Diebold–Mariano test. In particular, the table shows test statistics for the three models (VAR(1) on NS factors, AR(1) on returns, and Historical Average) considered for forecasting returns of US Treasury securities, equities, REIT, corporate bonds, and commodity against an AR(1) on Nelson–Siegel factors, in the pseudo out-of-sample period going from January 2011 to September 2020. Test statistics are shown according to three dimensions: (i) the model used to compute the forecasts, (ii) the returns series considered (one of 151 among the five asset classes analyzed), and (iii) the forecast horizon (one-step-ahead, six-step-ahead, and twelve-step-ahead forecasts). Boldfaced values represent series for which we can reject the null hypothesis of equal predictive accuracy in a 5%-sized, one-tailed test. Panels A to E show DM test statistics for each of the five asset classes considered, while Panel F shows a summary of DM tests by asset class and forecast horizon. Panel F summarizes how many times we can reject the null hypothesis of equal predictive accuracy in a 5%-sized, one-tailed test. Numbers in brackets in Panel F represent the number of series analyzed for each asset class.

Table A3. Root mean squared forecast errors (RMSE) for government bond yields.

Model	**RMSE (*h* = 1)**	**RMSE (*h* = 6)**	**RMSE (*h* = 12)**
VAR(1) on NS	1.88	0.22	0.03
AR(1) on Yields	9.30	0.93	0.44
VAR(1) on Yield Changes	14.44	1.56	0.77
Random Walk	8.07	1.03	0.54
Slope Regression	12.75	0.60	0.84
VAR(1) on Yields	10.88	21.90	0.77
Regression on Principal Components	177.89	24.79	0.77

Note: The table presents root mean squared forecast errors (RMSEs) for government bond yields at different maturities (3 months, 1 year, 5 years, and 10 years) across three forecast horizons: one-step-ahead (h = 1), six-step-ahead (h = 6), and twelve-step-ahead (h = 12). The RMSEs were computed for different forecasting models, including VAR(1) on Nelson–Siegel (NS) factors, AR(1) on yields, and benchmark methods such as the Random Walk and Principal Component Regression. Lower RMSE values indicate better forecasting accuracy.

Appendix D. Combining Forecast with Goyal and Welch Predictors

Table A4. Combining out-of-sample R^2_{OS} (%) for 6-month horizon.

Yield Curve			Treasury Returns			Equity Returns		
Maturity	GW Predictors	GW + NS Factor	Maturity	GW Predictors	GW + NS Factor	Sector	GW Predictors	GW + NS Factor
3 M	3.24 ***	4.77 ***	3 M	0.71 ***	5.77 ***	Mkt	0.25 **	0.13 *
1 Y	3.32 ***	4.88 ***	6 M	3.38 ***	5.66 ***	Food	0.25	0.09
5 Y	2.54 ***	3.28 ***	1 Y	0.87 ***	0.95 ***	Mines	0.06	0.13 *
10 Y	2.25 ***	3.03 ***	2 Y	0.14	0.66 **	Oil	−0.09	−0.11
			3 Y	0.08	0.32 *	Textiles	0.25	0.06
			5 Y	−0.01	1.52 **	Consumer Durables	0.53 ***	0.31 *
			7 Y	0.04	0.08	Chemicals	0.14	0.09
			10 Y	0.14	0.24	Consumer	0.45 **	0.40 **
						Construction	0.21 *	0.14
						Steel	0.03	0.01
						Fabricated Products	0.13	0.26 *
						Machinery	0.03	0.01
						Cars	0.31 **	−0.01
						Transportation	0.16 *	0.01
						Utilities	−0.09	−0.17
						Retail Stores	0.08	0.13
						Financials	0.38 **	0.26 *
						Other	0.52 ***	0.53 **
REITs			**Corporate Bonds**			**Commodities**		
Sector	GW Predictors	GW + NS Factor	Rating	GW Predictors	GW + NS Factor	Commodity	GW Predictors	GW + NS Factor
Diversified	−0.18	−0.27	AAAA	0.12	0.16	Gold	0.13	0.19
Industrial	−0.09	−0.79	AA	0.45 *	0.55 *	Silver	0.10	0.04
Lodging/Resorts	0.33	0.32	BBB	0.46 *	0.33	Platinum	−0.34	−0.39
Office	−0.21	−0.23	A	0.85 **	0.74 *	Copper	−0.30	−0.33
Residential	−0.07	−0.23				Brent Crude	−0.12	−0.21
Retail	−0.17	−0.07				Gasoil	−0.26	−0.35
Self Storage	−0.19	−0.07				Light crude oil	−0.17	0.16
All REITs	−0.23	−0.04				Natural Gas	0.01	0.20
Mortgage REITs	0.15 *	0.03				Cotton	0.35 **	0.09
						Cocoa	−0.03	−0.02
						Coffee	−0.15	−0.18
						Corn	−0.09	−0.17
						Lumber	0.38 ***	0.24 **
						Soybean Oil	−0.06	−0.20
						Soybeans	−0.09	−0.15
						Wheat	−0.04	−0.03
						Live Cattle	−0.04	−0.11

The table reports the out-of-sample R^2_{OS} (%) for a 6-month horizon, computed as the median of the combination forecasts. The predictive variables included in the combination models consist of various macro-financial indicators. These include valuation ratios such as the dividend–price ratio (D/P), the dividend yield (D/Y), the earnings–price ratio (E/P), and the dividend–payout ratio (D/E). Additionally, measures of market risk and performance such as stock variance (SVAR), the book-to-market ratio (B/M), and net equity expansion (NTIS) are considered. The analysis also incorporates interest rate variables, including the Treasury bill rate (TBL), the long-term yield (LTY), the long-term return (LTR), and the term spread (TMS), along with credit risk indicators such as the default yield spread (DFY) and the default return spread (DFR). Finally, inflation (INFL) is included, following Welch and Goyal (2008), with the predictive regression using the previous month's inflation rate due to the lag in data availability. The table compares combination forecasts with and without NS factors to assess their impact on predictive accuracy. The out-of-sample period spans from January 2011 to December 20200. Statistical significance for the R^2_{OS} statistic is based on the p-value for the Clark and West (2007) out-of-sample MSPE-adjusted statistic. The test corresponds to a one-sided hypothesis where the null states that the competing forecasting model has an expected squared prediction error equal to that of the historical average benchmark forecasting model. The alternative hypothesis posits that the competing forecasting model exhibits a lower expected squared prediction error relative to the benchmark. Statistical significance is denoted as follows: * for the 10% level, ** for the 5% level, and *** for 1%.

Notes

1 The stochastic discount factor, also called the pricing kernel, is a random variable used in asset pricing to discount future payoffs. It ensures that the expected discounted payoffs equal current asset prices. The SDF captures time value, risk preferences, and market conditions.

2 Fixing λ allows us to avoid estimating the baseline NS specification using nonlinear least squares, enhancing the simplicity and robustness of the forecasts obtained from the model. In fact, Diebold et al. (2006) emphasized that NLS estimation can encounter issues such as local minima or convergence failures.

3 In Equation (4) and in all subsequent forecasting models, we use a caret to denote the final predictions (of factors and yields) that we assign an economic meaning to and that we use to assess the forecast accuracy of the different models. A tilde is used to indicate intermediate-step parameter estimates that are instrumental in computing the final predictions.

4 In unreported tests, we tried to perform formal model specification searches concerning the number of VAR lags, and in particular investigated the performance for the yield series of a VAR(2) model, finding a similar or uniformly worse predictive performance.

5 The use of cointegrated VAR models for interest rate forecasting does not consistently outperform other methods. The inclusion of many yields increases model complexity, which can lead to overfitting and poor out-of-sample predictions. Despite improvements like cubic splines and ECM approaches, these models often overlook the fact that bond prices (yields) reflect all available information about future interest rates, limiting their forecasting effectiveness (Duffee, 2013).

6 In the case of Treasury data, even though an obvious (yet non linear) relationship exists between (the changes in ex ante) yields and (ex post, realized) returns, even when the NS factors are used, the forecast results obtained for returns are not presumed to be the same as the ones obtainable for yields.

7 Some literature labels this third predictive benchmark as a "random walk", which obviously is at odds with the restriction $\tilde{\gamma}_{i,t} = 0$ applied to Equation (15). The label derives from the fact that restricting $\theta_{i,t} = 1$ in $lnP_{i,t+h|t} = \theta_{0,i,t} + \theta_{1,i,t} lnP_{i,t} + \epsilon_{i,t+h}$, one obtains the model in (15) with $\tilde{\gamma}_{i,t} = 0$.

8 These predictors are available at Goyal's personal site: https://sites.google.com/view/agoyal145/home?authuser=0, accessed on 20 March 2025.

9 Negative yields, which may arise due to the low interest rate environment or specific technical market factors, such as cash and repurchase agreements in US Treasury markets, were excluded from the CMT calculation. This ensures that the dataset included only positive and economically meaningful yields.

10 Maturities such as the 1-month, 2-month, 20-year, and 30-year constant maturity rates, which were discontinued during the study period, were excluded.

11 https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html, accessed on 20 March 2025.

12 To qualify for inclusion, bonds had to have an investment-grade rating (based on an average of Moody's, S&P, and Fitch) and meet specific criteria regarding maturity (a remaining maturity of at least one year at the time of inclusion in the index), a fixed coupon schedule (i.e., zero coupon bonds are excluded), and minimum amount outstanding (USD 250 million). The index was capitalization-weighted based on the outstanding amount of each security at the end of each month.

13 Nonetheless, this is considered partially unrealistic, also in light of the recent results in Kim and Choi (2017).

14 The sample autocorrelation in the forecast errors further supports these findings, with errors from the AR(1) and VAR(1) models for yields being serially correlated for most maturities and horizons, except for longer maturities (3-, 5-, 7-, and 10-year yields) at one-step-ahead horizons.

15 Because the results for MCS test are starker and allow us to draw sharper conclusions, in the main text, we present the tables documenting the findings from the MCS methodology and report in an Appendix those from the simpler, but intuitive DM tests.

16 While the two test statistics used in the MCS procedure, TMax and TR, produced similar results, they differed in the total number of models excluded from the superior set. The TMax statistic tended to exclude fewer models from M^* across all forecasting horizons compared to the TR statistic. Using the TMax statistic, we excluded 187, 62, and 181 models from M^* across all return series for the one-step-ahead, six-step-ahead, and 12-step-ahead forecasts, respectively. Using the TR statistic, we excluded 229, 76, and 188 models from M^* for the same horizons. In the case of the return series, we observed the opposite of what was found for the bond yield series, with the TMax test appearing less selective than the TR statistic. This difference was likely due to the fact that only four models were compared when forecasting returns, whereas up to eight models were compared when forecasting bond yields, making pairwise comparisons based on the TR statistic more powerful than the multivariate comparisons used by the TMax statistic.

17 As a sub-question, we also explored whether an AR(1) or a VAR(1) model based on (predicted) NS factors could be preferable in forecasting returns. Although no definitive conclusion emerged, our analysis suggested that the two types of models did not exhibit equivalent performance across different asset classes and forecast horizons. In overall terms, the VAR(1) NS tended to outperform the AR(1) NS, particularly in the case of Treasury securities, equities, and corporate bonds. However, the AR(1) NS

returned a slightly superior performance in the prediction of REIT returns. As far as the forecast horizon is concerned, the AR(1) NS performance relative to the VAR(1) improved as the horizon lengthened, especially for US Treasuries and corporate bonds. Conversely, the VAR(1) model proved more effective in forecasting REIT returns.

References

Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *The Review of Financial Studies*, *21*(4), 1509–1531. [CrossRef]

Chen, Y., & Niu, L. (2014). Adaptive dynamic Nelson–Siegel term structure model with applications. *Journal of Econometrics*, *180*(1), 98–115. [CrossRef]

Christensen, J. H. E., Diebold, F. X., & Rudebusch, G. D. (2011). The affine arbitrage-free class of Nelson–Siegel term structure models. *Journal of Econometrics*, *164*(1), 4–20. [CrossRef]

Clark, T. E., & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, *138*(1), 291–311. [CrossRef]

Diebold, F. X., & Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics*, *130*(2), 337–364. [CrossRef]

Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, *20*(1), 134–144.

Diebold, F. X., & Rudebusch, G. D. (2013). *Yield curve modeling and forecasting: The dynamic Nelson-Siegel approach*. Princeton University Press.

Diebold, F. X., Rudebusch, G. D., & Aruoba, S. B. (2006). The macroeconomy and the yield curve: A dynamic latent factor approach. *Journal of Econometrics*, *131*(1–2), 309–338.

Drobetz, W., Stürmer, S., & Zimmermann, H. (2002). Conditional asset pricing in emerging stock markets. *Swiss Journal of Economics and Statistics (SJES)*, *138*(IV), 507–526.

Duffee, G. (2013). Forecasting interest rates. In G. Elliott, & A. Timmermann (Eds.), *Handbook of economic forecasting* (Vol. 2, pp. 385–426). North Holland.

Favero, C. A., Niu, L., & Sala, L. (2012). Term structure forecasting: No-arbitrage restrictions versus large information set. *Journal of Forecasting*, *31*(2), 124–156. [CrossRef]

Fernandes, M., & Vieira, F. (2019). A dynamic Nelson–Siegel model with forward-looking macroeconomic factors for the yield curve in the US. *Journal of Economic Dynamics and Control*, *106*, 103720. [CrossRef]

Frankel, J. A. (2008). The effect of monetary policy on real commodity prices. In *Asset prices and monetary policy* (pp. 291–333). University of Chicago Press.

Guidolin, M., McMillan, D. G., & Wohar, M. E. (2013). Time varying stock return predictability: Evidence from US sectors. *Finance Research Letters*, *10*(1), 34–40. [CrossRef]

Guidolin, M., & Pedio, M. (2019). Forecasting and trading monetary policy effects on the riskless yield curve with regime switching Nelson–Siegel models. *Journal of Economic Dynamics and Control*, *107*, 103723.

Hansen, P. R., Lunde, A., & Nason, J. M. (2011). The model confidence set. *Econometrica*, *79*(2), 453–497. [CrossRef]

Hillebrand, E., Huang, H., Lee, T.-H., & Li, C. (2018). Using the entire yield curve in forecasting output and inflation. *Econometrics*, *6*(3), 40. [CrossRef]

Holmström, B., & Tirole, J. (2001). LAPM: A liquidity-based asset pricing model. *The Journal of Finance*, *56*(5), 1837–1867. [CrossRef]

Kim, J. H., & Choi, I. (2017). Unit roots in economic and financial time series: A re-evaluation at the decision-based significance levels. *Econometrics*, *5*(3), 41. [CrossRef]

Nelson, C. R., & Siegel, A. F. (1987). Parsimonious modeling of yield curves. *Journal of Business*, 473–489. [CrossRef]

Nymand-Andersen, P. (2018). Yield curve modelling and a conceptual framework for estimating yield curves: Evidence from the European Central Bank's yield curves. In *ECB statistics paper*. European Central Bank.

Pesando, J. E. (1979). On the random walk characteristics of short-and long-term interest rates in an efficient market. *Journal of Money, Credit and Banking*, *11*(4), 457–466. [CrossRef]

Potì, V. (2018). A new tight and general bound on return predictability. *Economics Letters*, *162*, 140–145.

Rapach, D., Strauss, J., & Zhou, G. (2010). Out-of-sample equity premium prediction: Consistently beating the historical average. *Review of Financial Studies*, *23*(2), 821–862. [CrossRef]

Rapach, D., & Zhou, G. (2013). Forecasting stock returns. In *Handbook of economic forecasting* (Vol. 2, pp. 328–383). Elsevier Science & Technology.

Steeley, J. M. (2014). Forecasting the term structure when short-term rates are near zero. *Journal of Forecasting*, *33*(5), 350–363. [CrossRef]

Toczydlowska, D., & Peters, G. W. (2018). Financial big data solutions for state space panel regression in interest rate dynamics. *Econometrics*, *6*(3), 34. [CrossRef]

Welch, I., & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, *21*(4), 1455–1508. [CrossRef]

Xiang, J., & Zhu, X. (2013). A regime-switching Nelson–Siegel term structure model and interest rate forecasts. *Journal of Financial Econometrics*, *11*(3), 522–555. [CrossRef]

Yu, W.-C., & Zivot, E. (2011). Forecasting the term structures of Treasury and corporate yields using dynamic Nelson-Siegel models. *International Journal of Forecasting*, *27*(2), 579–591. [CrossRef]