

Predicting Interest Rates Using Shrinkage Methods, Real-Time Diffusion Indexes, and Model Combinations *

Norman R. Swanson¹ and Weiqi Xiong²

^{1,2}Rutgers University and ²Blackrock, Inc.

May 2018

Abstract

In the context of predicting the term structure of interest rates, we explore the marginal predictive content of real-time diffusion indexes extracted from a “data rich” real-time dataset, when used in dynamic Nelson-Siegel (NS) models of the variety discussed in Diebold and Li (DNS: 2007) and Svensson (NSS: 1994). We find that the indexes have significant predictive content for sample periods ranging from 2001 through 2010. Additionally, DNS and NSS type models that include these indexes are the mean square forecast error (MSFE) “best” performers, when compared with various other econometric specifications, for the same period. In our top ranked models, diffusion indexes are sometimes optimally constructed used un-targeted principal component analysis (PCA), and sometimes using targeted PCA in conjunction with elastic net and least absolute shrinkage operators. The news is not all good for models utilizing indexes constructed from real-time datasets, however. In particular, after 2010 relatively few “data-rich” prediction models “beat” a random walk benchmark. Also, forecast combinations that utilize models that exclude real-time diffusion indexes yield the lowest overall MSFEs, dominating all other combination and individual models, across all sample periods, forecast horizons and bond maturities. Two key conclusions from our analysis are the following. First, fully revised data may have an important confounding effect upon results obtained when instead carrying out real-time prediction experiments. Second, real-time diffusion indexes matter when comparing the predictive performance of individual models, indicating the presence of unspanned macroeconomic risks in the term structure of interest rates. However, there appear to be two different ways to “capture” these unspanned risks. One is to use data rich real-time diffusion indexes, and another is to simply combine predictions from many non-data rich models.

JEL Classification: C12, C22, C53.

Keywords: Real-time Forecasting, Dynamic Nelson-Siegel Model, Term Structure of Interest Rates, Real-time Dataset, FRED-MD, Real-Time Diffusion Index.

* Corresponding Author: Norman R. Swanson, Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA, nswanson@econ.rutgers.edu. Weiqi Xiong, Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ 08901, USA, wx46@economics.rutgers.edu. The authors are grateful to Mingmian Cheng, Valentina Corradi, Eric Ghysels, Yuan Liao, Massimiliano Marcellino, Christian Schumacher, Greg Tkacz, and Xiye Yang for useful comments.

1 Introduction

The term structure of interest rates contains crucial information for forecasting macroeconomic variables and pricing interest rate contingent assets. As a consequence, forecasts of U.S. Treasury bill and bond yields remain important inputs in models used in industry and government. In this paper, we add to the literature on interest rate prediction by carrying out an extensive set of forecast experiments in order to explore the marginal predictive content of so-called “data rich” or real-time latent macroeconomic factors (i.e., real-time diffusion indexes) in Nelson-Siegel (NS) type models. The particular data-rich environment that we examine is one in which real-time data that include the entire revision history for each variable are analyzed. For example, real-time GDP observations for calendar date December 2000 include the first “reading” on 4th quarter 2000 GDP that was available in March 2001, and well as the 1st revised version of this datum that became available in June 2001, and so on, up until the present date. Thus, real-time datasets include entire sequence of revisions for each calendar dated observation. Data such as these allow researchers to simulate “truly” real-time forecasting environments, which differs from the common practice of using so-called fully revised data in forecasting experiments. This is important, as “fully revised” data consist of observations that were not actually available to market participants in real-time.

We carry out our prediction experiments for various sub-samples between 2001 and 2018, and results are evaluated using a number of benchmark linear models. In particular, we assess the following classes of models: (i) DNS type models of the variety recently examined by Diebold and Li (2007), (ii) dynamic Nelson Siegel Svensson (NSS) type models (see Svensson (1994)), and (iii) various benchmark models, including vector autoregressive (VAR) and autoregressive (AR) models. The real-time diffusion indexes that we utilize are extracted using principle component analysis (PCA) of 130 U.S macro-variables for which McCracken and Ng (2016) have constructed a real-time dataset.¹ When incorporating information from this dataset in our analysis, we use both un-targeted and targeted methods. Our un-targeted method is based on application of PCA to the entire dataset, while our targeted method involves using either the well-known elastic net or the least absolute shrinkage operator (LASSO) to “pre-select” a subset of variables for subsequent use in PCA.² For a discussion of targeted and un-targeted diffusion index construction methods, see Bai and Ng (2007,2008), Carrasco and Rossi (2016), Ghysels and Marcellino (2018), Kim and Swanson (2014), and Schumacher (2007,2009). For a discussion of forecasting using factor augmented Nelson-Siegel models, see Exterkate, van Dijk, Heij, and Groenen (2013).

Although many sophisticated models of the term structure have been examined in the literature, sim-

¹The models utilized in this paper do not include a separate and important class called mixed-frequency models. For an interesting discussion of mixed frequency modeling and diffusion indexes, see Andreou, Gagliardini, Ghysels and Rubin (2018).

²Refer to Kim and Swanson (2014), Stock and Watson (2012), and the references cited therein for a discussion of shrinkage methods for forecasting with many predictors.

pler regression-based approaches to forecasting treasury yields have the best track record for minimizing out-of-sample mean squared forecast error (MSFE). The most popular of these models is currently the Dynamic Nelson-Siegel (DNS) model, as discussed in Diebold and Li (2006). The DNS model is a dynamic version of the NS model introduced by Nelson and Siegel (1987), where cross-section movements in the term structure are summarized by dynamic level, slope, and curvature factors, assumed to follow AR(1) (or VAR(1)) processes. DNS type models have become the leading method for yield curve forecasting at many policy institutions (see BIS (2005)), and although this development is largely due to the successful empirical performance of these models, findings in the recent literature suggest that DNS model performance has deteriorated in recent (post credit crisis) years (see Altavilla, Giacomini and Ragusa (2017), Diebold, and Rudebusch (2012), and Mönch (2008)). This might be explained by regime changes or structural breaks, for example. Other potential causes include generic model misspecification, model over-fitting, and measurement error. One possible solution to this problem has centered around the introduction of new variants of DNS type models. A candidate model, which we examine in this paper is the so-called dynamic NSS model mentioned above.

Another possible solution to the problem mentioned above centers around the recent general consensus that has emerged in the literature stating that one should look beyond the cross section of yields when pinning down the dynamic behavior of interest rates (Duffee (2011)). Along these lines, modeling the co-movements of the underlying economy by specifying diffusion indexes (Stock and Watson, 2002a), or using key macroeconomic indicators, has proven useful in predicting yields. For example, using a dynamic factor model, Coroneo, Giannone and Modugno (2016) find that real economic activity and real interest rates contain predictive content for government bond yields that are not spanned by the cross-section of yields. Ang and Piazzesi (2003) and Mönch (2008) also report improved forecasts using affine models, when using principal component-based macroeconomic factors (i.e., diffusion indexes). Additional recent studies consider enlarging the information set used in prediction with either observable macroeconomic factors (Diebold, Rudebusch and Aruoba (2006) and Rudebusch and Wu (2008)) or surveys (Altavilla, Giacomini and Ragusa (2017)). Ludvigson and Ng (2009) find that adding macroeconomic factors helps when forecasting bond risk premia. As discussed above, our empirical analysis adds to this literature by assessing the marginal predictive content of real-time diffusion indexes constructed in a data-rich environment. These indexes are used alone and as inputs into other models including DNS, NSS, and benchmark linear models; and predictions of yields at various maturities and for various forecast horizons are constructed. Additionally, data shrinkage methods are implemented when estimating indexes, and a number of forecast combinations are examined. Finally, results are tabulated for various sub-samples between 2001 and 2018 in order to assess whether model rankings are dependent upon sample period.

A number of clear-cut findings emerge upon examination of the results of our prediction experiments. First, real-time diffusion indexes have marginal predictive content when added to all of the individual

models considered in our analysis, for sample periods ranging from 2001 through 2010. Additionally, DNS and NSS models are the mean square forecast error (MSFE) best performers, when compared with all other econometric specifications, for the same period. Thus, there are unspanned macroeconomic risks, and they can be captured using real-time prediction methods. Second, elastic net and LASSO shrinkage operators are used for constructing indexes in many of our MSFE-best models. However, this is not a universal finding, as there are some cases where simple un-targeted use of our entire real-time dataset yields the MSFE-best models. Third, after 2010 “data-rich” prediction models perform much worse, with RW forecasts often yielding lower MSFEs.³ It is argued that this is due at least in part to the low variability associated with the current zero-lower bound interest rate regime after 2010. Moreover, forecast combinations based on all models that exclude diffusion indexes (i.e., based on a combination of DNS, NSS, and various simple benchmark models) yield the lowest overall MSFEs, dominating all other models, across all sample periods, forecast horizons and bond maturities. This result is in contrast to the findings of Swanson and Xiong (2017), where the inclusion of diffusion indexes always leads to predictive improvement. However, only fully revised macroeconomic data are utilized in that paper.⁴ Two key conclusions can be drawn from these findings. First, fully revised data may have an important confounding effect upon results obtained when instead carrying out real-time prediction experiments. Second, real-time diffusion indexes matter when comparing the predictive performance of individual models, indicating the presence of unspanned macroeconomic risks in the term structure of interest rates. This finding is particularly true when the economy is not in an economic regime associated with a zero-lower bound and low interest rate variability. However, there appear to be two different ways to “capture” these unspanned risks. One is to use data rich real-time diffusion indexes in individual models, and another is to simply combine predictions from many non-data rich models.

The rest of the paper is organized as follows. Section 2 describes the dynamic Nelson Siegel models that we analyze, and Section 3 discusses yield curve prediction with macroeconomic diffusion indexes. Section 4 includes details describing our empirical setup, Section 5 details the data used in our analysis, and Section 6 contains a summary of our empirical findings. Concluding remarks are gathered in Section 7.

³The same results hold when the benchmark is an AR(SIC), with lags selected using the Schwarz information criterion.

⁴The methodology used in the survey paper of Swanson and Xiong (2017) also differs from that used here in a number of other dimensions. For example, those authors do not use targeted methods, such as the elastic net and LASSO, for constructing diffusion indexes. Also, as they do not use real-time data in their analysis. Finally, they consider only a small subset of the models examined here.

2 Dynamic Nelson Siegel Models

2.1 Three-factor Dynamic Nelson Siegel Model

Motivated by rational expectation theory, Nelson and Siegel (1985) express spot interest rates in terms of instantaneous forward rates. Namely, the instantaneous forward interest rate of a bond with maturity m is denoted as $f(m)$, and the spot interest rate of a bond with maturity τ as $y(\tau)$. Then, the yield to maturity of a bond can be written as the average of forward rates

$$y(\tau) = \frac{1}{\tau} \int_0^\tau f(m) dm.$$

Nelson and Siegel (1985) motivate the use of the following model of the forward rate that can generate monotonically increasing, humped, and occasionally S-shaped yield curves, a range of shapes for yield curves:

$$f(m) = \beta_1 + \beta_2 \cdot \exp\left(\frac{m}{\theta_t}\right) + \beta_3 \cdot \left[\left(\frac{m}{\theta_t}\right) \exp\left(\frac{m}{\theta_t}\right)\right],$$

where $\lambda_t = \frac{1}{\theta_t}$ is the so-called decay parameter, which must be estimated, is assumed fixed in this model, and is time varying in the dynamic version of the model discussed below. It is then easy to derive the following model for bond yields:

$$y(\tau) = \beta_1 + \beta_2 \cdot \left[\frac{1 - \exp(-\frac{\tau}{\theta_t})}{\frac{\tau}{\theta_t}}\right] + \beta_3 \cdot \left[\frac{1 - \exp(-\frac{\tau}{\theta_t})}{\frac{\tau}{\theta_t}} - \exp(-\frac{\tau}{\theta_t})\right].$$

In the above model, the latent factors (i.e., the “betas”) are fixed. Diebold and Li (2006) generalize this model to allow for time-varying betas: $\beta_{1,t}$, $\beta_{2,t}$ and $\beta_{3,t}$. Their so-called Dynamic Nelson-Siegel (DNS) model is estimated using a two-step procedure. First, the rate of decay λ_t is set to a constant. Next, at each point in time, t , the yield cross section is linearly projected onto the set of factor loadings $(1, \frac{1 - \exp(-\lambda_t \tau)}{\lambda_t \tau}, \frac{1 - \exp(-\lambda_t \tau)}{\lambda_t \tau} - \exp(-\lambda_t \tau))$. In our experiments, various different dimensions are considered when specifying the yield cross section. Namely, we consider yield cross sections using 10, 12, and 30 different yield maturities. For example, with our 12-dimensional cross section, we estimate the latent factors by fitting the following regression:

$$\begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ y_t(\tau_3) \\ \vdots \\ y_t(\tau_{12}) \end{pmatrix}_{12 \times 1} = \begin{pmatrix} 1 & \frac{1 - \exp(-\lambda_t \tau_1)}{\lambda_t \tau_1} & \frac{1 - \exp(-\lambda_t \tau_1)}{\lambda_t \tau_1} - \exp(-\lambda_t \tau_1) \\ 1 & \frac{1 - \exp(-\lambda_t \tau_2)}{\lambda_t \tau_2} & \frac{1 - \exp(-\lambda_t \tau_2)}{\lambda_t \tau_2} - \exp(-\lambda_t \tau_2) \\ 1 & \frac{1 - \exp(-\lambda_t \tau_3)}{\lambda_t \tau_3} & \frac{1 - \exp(-\lambda_t \tau_3)}{\lambda_t \tau_3} - \exp(-\lambda_t \tau_3) \\ \vdots & \vdots & \vdots \\ 1 & \frac{1 - \exp(-\lambda_t \tau_{12})}{\lambda_t \tau_{12}} & \frac{1 - \exp(-\lambda_t \tau_{12})}{\lambda_t \tau_{12}} - \exp(-\lambda_t \tau_{12}) \end{pmatrix}_{12 \times 3} \begin{pmatrix} \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \end{pmatrix}_{3 \times 1}$$

The betas (i.e., $\hat{\beta}_{1,t}$, $\hat{\beta}_{2,t}$, and $\hat{\beta}_{3,t}$) are called the “level”, “slope”, and “curvature” factors. In particular, note that the loading on $\hat{\beta}_{1,t}$ is one, which is naturally interpreted as the “level” factor. The loading on $\hat{\beta}_{2,t}$ decreases as bond maturity increases, resulting in an increase of the “slope” of bond yield curve.

Finally, the loading on the third latent factor, $\hat{\beta}_{3,t}$, starts from zero on the short end of yield curve, reaches its peak at some maturity in the middle, and gradually decays to zero as maturity goes to infinity. Figures 3 exhibits the three NS factors estimated with ordinary least squares for sample period 1988:8 - 2017:10.⁵ In summary, the DNS model can be written as follows:

$$\hat{y}_t(\tau) = \hat{\beta}_{1,t} + \hat{\beta}_{2,t} \cdot \left[\frac{1 - \exp(-\lambda_t \tau)}{\lambda_t \tau} \right] + \hat{\beta}_{3,t} \cdot \left[\frac{1 - \exp(-\lambda_t \tau)}{\lambda_t \tau} - \exp(-\lambda_t \tau) \right]. \quad (2.1)$$

In order to construct predictions using the DNS model, we fit estimated factors to AR and VAR models, as follows.

$$\hat{\beta}_{i,t+1} = c_i + \gamma_i \hat{\beta}_{i,t} + \epsilon_t \quad i = 1, 2, 3 \quad \text{or}, \quad (2.2)$$

$$\hat{\beta}_{t+1} = \mathbf{c} + \mathbf{\Gamma} \hat{\beta}_t + \boldsymbol{\epsilon}_t, \quad (2.3)$$

where ϵ_t is a scalar stochastic disturbance term, $\boldsymbol{\epsilon}_t$ is a 3×1 vector of stochastic disturbance terms, and c_i , \mathbf{c} , γ_i , and $\mathbf{\Gamma}$, $i = 1, \dots, 3$, are conformably defined constants, constant vectors and constant matrices. With these last two models, one can construct predictions of the $\hat{\beta}_{i,t}$, for $i = 1, \dots, 3$, which can in turn be inserted into the above model of $\hat{y}_t(\tau)$ in order to generate predictions thereof. In all experiments in the sequel, rolling estimation is carried out when estimating the above models (and all other models), using windows of length 120 months, so that “real-time” predictions are constructed in all cases. Additionally, we consider two types of prediction models. In one, the decay parameter is fixed. In the other, the decay parameter is re-estimated prior to the construction of each new prediction. For further details, including a recent review of Treasury yield curve modeling using DNS models, see Diebold and Rudebusch (2013) and De Pooter (2007). For further discussions comparing arbitrage free dynamic latent factor and DNS models, see Ang and Piazzesi (2003), Diebold, Rudebusch and Aruoba (2006), Christensen, Diebold, and Rudebusch (2011), Duffie (2011), and the references cited therein.

2.2 Four-factor Nelson-Siegel-Svensson Model

Svensson (1994) extends the Nelson-Siegel Svensson (NSS) model by adding a fourth term, that allows for a second “hump” shape. In particular, he proposed the following four-factor model for fitting the instantaneous forward interest rate:

$$f(m) = \beta_1 + \beta_2 \cdot \exp\left(\frac{m}{\theta_{1,t}}\right) + \beta_3 \cdot \left[\left(\frac{m}{\theta_{1,t}}\right) \cdot \exp\left(\frac{m}{\theta_{1,t}}\right)\right] + \beta_4 \cdot \left[\left(\frac{m}{\theta_{2,t}}\right) \cdot \exp\left(\frac{m}{\theta_{2,t}}\right)\right].$$

⁵An increase in the “level” component, β_{1t} , affects all yields equally, thus it determines the level of the yield curve. Also, as maturity τ goes to infinity, $\beta_{1t} = y_t(\infty)$ by definition. An increase in “slope” component β_{2t} affects short rates more than long rates, thereby changing the slope, or the so-called “term spread” of the yield curve. Finally, an increase in β_{3t} , the “curvature” component, will increase medium-term yields and have little effect on the short and long end of the curve. Therefore, the yield curve will become more hump shaped. As demonstrated in Diebold and Li (2006), the “level” factor can be approximated with the 10-year bond yield, the “slope” factor can be approximated with 10-year - 3-month bond yield spreads, and the “curvature” factor moves closely with two times the 2-year yield minus the sum of the 3-month and 10-year yields.

Notice that in the above equation there are now two different decay parameters controlling the double-hump shape of the forward curve, called θ_1 and θ_2 . Similar to the DNS model, we consider a dynamic version of the NSS model. Thus, we utilize the following variant of the DNS model (factor estimation and prediction construction is carried out using the DNS modeling approach discussed above).

$$\begin{aligned}\hat{y}_t(\tau) = & \hat{\beta}_{1,t} + \hat{\beta}_{2,t} \cdot \left[\frac{1 - \exp(-\lambda_{1,t}\tau)}{\lambda_{1,t}\tau} \right] + \hat{\beta}_{3,t} \cdot \left[\frac{1 - \exp(-\lambda_{1,t}\tau)}{\lambda_{1,t}\tau} - \exp(-\lambda_{1,t}\tau) \right] \\ & + \hat{\beta}_{4,t} \cdot \left[\frac{1 - \exp(-\lambda_{2,t}\tau)}{\lambda_{2,t}\tau} - \exp(-\lambda_{2,t}\tau) \right],\end{aligned}$$

where we now have two decay parameters, as discussed above. These are called $\lambda_{1,t}$ and $\lambda_{2,t}$. As discussed in De Pooter (2007), the second hump in the NSS model is difficult to identify without imposing additional restrictions. We adopt his approach to solving this issue, which includes assumptions that the two humps are at least one year apart, and that the second hump reaches its maximum for a maturity which is at least twelve months shorter than the first hump. Additionally, it is assumed that $\lambda_1 \neq \lambda_2$, in order to avoid multicollinearity. Figures 4A - B plots the four NSS factors estimated with static and dynamic decay parameters, $\lambda_{1,t}$ and $\lambda_{2,t}$. Figure 4C plots estimated rates of decay used in the construction of the four Nelson-Siegel-Svensson factors, where the rates of decay ($\lambda_{1,t}, \lambda_{2,t}$) are either set to fixed numbers, or estimated recursively using nonlinear least squares. See Section 3.1.2 for details on model estimation.

3 Unspanned Risks and Diffusion Indexes

Whether or not macroeconomic, financial and other non-yield information is useful in fitting and forecasting the yield curve remains an open question. As Duffee (2013) points out, yields are usually hypothesized to follow Markov a process, which implies that all information in fundamental economic variables should already be embedded in yield cross sections. This leaves no role for so-called “unspanned risks”, as proxied for by additional economic variables and/or diffusion indexes constructed in a data rich environment. Namely, he argues that, at least theoretically, it is redundant to add current non-yield information in the above forecast equation for interest rates. On the other hand, in the empirical literature there are many examples where adding economic variables and diffusion indexes has proven to be effective in improving yield curve fitting as well as forecasting. In particular, Ang and Piazzesi (2003) find that macroeconomic variables are significant for explaining Treasury security yield dynamics, based on a VAR analysis. Mönch (2008) shows that adding estimated diffusion indexes to an affine Gaussian term structure model improves out-of-sample forecast performance. Diebold, Rudebusch, and Aruoba (2006) investigate the bidirectional causality between yield betas and macro variables and discover strong evidence in favor of linkages between macroeconomic variables and future yield curve dynamics.

Recently, focus has turned to so-called big data, and to the analysis of the usefulness of largescale datasets in the above context. As noted in Bernanke and Boivin (2003), monetary policy-makers and

academics alike are very interested in examining the (predictive) usefulness of a wide range of variables in a data-rich environments. For example, the predictive usefulness of diffusion indexes constructed using largescale datasets has been examined in countless academic papers in the past few years. The same is certainly true in industry, where the prevalence of big data and related machine learning methods is readily apparent. One important aspect of big data in our context is the use of so-called real-time data, as discussed in the introduction. Recently, McCracken and Ng (2016) and St. Louis Federal Reserve Bank’s data desk created the FRED-MD, which is a large monthly real-time database that contains over 130 macro-variables and all revisions of all of these variables. The dataset contains variables summarizing economic output and income, labor markets, consumption, money and credit, housing, and stock market, for example. Moreover, they show that diffusion indexes extracted from their FRED-MD dataset contain the same predictive content as diffusion indexes constructed using the classic Stock and Watson dataset (Stock and Watson (2002a,b)). However, the FRED-MD is a real-time database, while the Stock and Watson dataset contains only fully revised data. Several studies have revealed the importance of collecting and updating such real-time datasets including Diebold and Rudebusch (1991), Hamilton and Perez-Quiros (1996), Bernanke and Boivin (2003), and the papers cited therein.

In this paper, we ask the following question: Are diffusion indexes useful for predicting yields, when the data used to construct the indexes are purely “real-time”, rather than fully revised as in Swanson and Xiong, (2017), for example. We motivate the use of diffusion indexes by adopting a dynamic factor model framework resembling that used by Coroneo, Giannone and Modugno (2016) and many others. Namely, we assume that yields curve factors, (which are the betas in the above discussion are here called $F_{y,t}$), are driven by both past yield curve factors and macro factors, called $F_{x,t}$. Additionally, it is assumed that macroeconomic variables are driven only by $F_{x,t}$ only. Thus, we posit the following model:

$$\begin{pmatrix} F_{y,t+h} \\ x_t \end{pmatrix} = \begin{pmatrix} c_y \\ c_x \end{pmatrix} + \begin{bmatrix} \Gamma_y & \Gamma_x \\ 0 & \Gamma_{xx} \end{bmatrix} \begin{pmatrix} F_{y,t} \\ F_{x,t} \end{pmatrix} + \begin{pmatrix} e_{y,t+h} \\ e_{x,t} \end{pmatrix},$$

where c_y, c_x are vectors containing constant terms, h is the forecast horizon, Γ_y contains factor loadings on yield factors, Γ_{xx} contains factor loadings on the macro factors, and Γ_x summarizes the marginal effect of macro factors on yield factors. Additionally, $e_{y,t+h}$ and $e_{x,t}$ are idiosyncratic stochastic disturbance terms. In their paper, Coroneo, Giannone and Modugno (2016) use a so-called expectation conditional restricted maximization algorithm for model estimation, and measure the effect of “unspanned” macroeconomic variables (risks) on the yield curve. We use principal component analysis (PCA) for estimating our macro diffusion indexes (i.e., macro factors), following Stock and Watson (2002a,b), and consider various alternative models that utilize macro diffusion indexes. For instance, we examine whether adding macro diffusion indexes to our DNS and NSS models improves the predictive accuracy of these models. Of course, we also consider baseline DNS (or NSS) models that contain only yield factors. More concretely,

h -step ahead predictions for yield factors are constructed using the following model:

$$\hat{F}_{y,t+h}^f = \hat{c}_y + \hat{\Gamma}_y' \hat{F}_{y,t}, \quad (3.1)$$

where $\hat{F}_{y,t}$ is our estimated DNS (or NSS) latent factor (i.e. $\hat{F}_{y,t}$ are our betas in the above discussion), $\hat{F}_{y,t+h}^f$ is our prediction constructed by specifying simple AR(1) or VAR(1) models, \hat{c}_y is an estimate of c_y , and $\hat{\Gamma}_y$ is an estimate of Γ_y . We additionally add the first r_x principle components from a PCA analysis of our real-time dataset, denoted as $\hat{F}_{x,t}$, to the above prediction model, yielding:

$$\hat{F}_{y,t+h}^f = \hat{c}_y + \hat{\Gamma}_y' \hat{F}_{y,t} + \hat{\Gamma}_x' \hat{F}_{x,t} \quad (3.2)$$

where $\hat{\Gamma}_x$ is an estimate of Γ_x . When predicting yields, in addition to utilizing DNS and NSS models, we also examine whether adding macro diffusion indexes to benchmark AR and VAR models improves predictive accuracy. In particular, we consider the following model:

$$\begin{pmatrix} y_{t+h} \\ x_t \end{pmatrix} = \begin{pmatrix} c \\ c_x \end{pmatrix} + \begin{bmatrix} \Delta_y & \Delta_x \\ 0 & \Gamma_{xx} \end{bmatrix} \begin{pmatrix} y_t \\ F_{x,t} \end{pmatrix} + \begin{pmatrix} e_{t+h} \\ e_{x,t} \end{pmatrix},$$

where c is the vector containing constant terms, all coefficient matrices (i.e., Δ_y , Δ_x , and Γ_{xx}) are a conformably defined coefficient matrices, Δ_x summarizes the marginal effect of macro diffusion indexes on yields, and e_{t+h} is an idiosyncratic stochastic disturbance term. Summarizing, our focus of interest is on h -step ahead yield predictions constructed using the following model:

$$\hat{y}_{t+h}(\tau) = \hat{c}(\tau) + \hat{\delta}_y' y_t, \quad (3.3)$$

where $\hat{c}(\tau)$ is an estimate of $c(\tau)$, which is an element of c . Also, $\hat{\delta}_y$ is an estimate of δ_y , which is a row vector of Δ_y . y_t contains lags of $y_{t+1}(\tau)$ in autoregressive specifications, and contains lags of y_{t+1} in vector autoregressive specifications. We additionally add the macro diffusion indexes discussed above, F_t^x , to this model, yielding:

$$\hat{y}_{t+h}(\tau) = \hat{c}(\tau) + \hat{\delta}_y' y_t + \hat{\delta}_x' \hat{F}_{x,t}, \quad (3.4)$$

where $\hat{\delta}_x$ is an estimate of δ_x , which is a row vector of Δ_x . For further discussion of diffusion indexes in macroeconomic forecasting, see Banerjee, Marcellino and Marsten (2008) Boivin and Ng (2005), and Kim and Swanson (2014).

4 Empirical Setup

4.1 Predictive Accuracy Testing

When comparing the predictive performance of the models detailed below, we report the mean square forecast error (MSFE), defined as:

$$\text{MSFE}_h(\tau) = \sum_{t=1}^P (\hat{y}_{t+h}(\tau) - y_{t+h}(\tau))^2 \quad (4.1)$$

where $\hat{y}_{t+h}(\tau)$ is the h -step-ahead forecast of the Treasury bond yield, with maturity τ . P is the number of ex ante predictions used in our analysis. Additionally, all model parameters are estimated with maximum likelihood and PCA; and parameters are updated prior to the construction of each forecast using a rolling window of 120 months of historical data. For an analysis of the use of rolling versus recursive and alternative windowing techniques in the context of forecasting, see Clark and McCracken(2009) and Hansen and Timmermann (2012) and Rossi and Inoue (2012). Model MSFEs are compared using the Diebold and Mariano (1995) predictive accuracy test. The null hypothesis of the DM test is: $H_0 : E[L(\epsilon_{t+h}^{(1)})] - E[L(\epsilon_{t+h}^{(2)})] = 0$, where the $\epsilon_{t+h}^{(i)}$ is the prediction error of model i , for $i = 1, 2$. In our analysis, $L(\cdot)$ is a quadratic loss function. The DM test statistic is:

$$DM_P(\tau) = P^{-1} \sum_{t=1}^P \frac{d_{t+h}(\tau)}{\hat{\sigma}_{\bar{d}}} \quad (4.2)$$

where $d_{t+h}(\tau) = [\hat{\epsilon}_{t+h}^{(1)}(\tau)]^2 - [\hat{\epsilon}_{t+h}^{(2)}(\tau)]^2$, \bar{d} denotes the mean of $d_{t+h}(\tau)$, $\hat{\sigma}_{\bar{d}}$ is a heteroskedasticity and autocorrelation consistent estimate of the standard deviation of \bar{d} , and P denotes the number of ex ante predictions used to construct the test statistic. If the DM_P statistic has a negative value, Model 1 is preferred to Model 2. If the DM_P statistic is significantly different from zero, the difference between Model 1 and Model 2 is statistically significant. In the sequel, we assume that the DM_P test is asymptotically $N(0,1)$, although in cases where models being compared are nested, modified critical values tabulated by McCracken (2000) should be used (see Corradi and Swanson (2006) for complete details). For an interesting discussion of alternative approaches to assessing forecasting performance, see Rossi and Sekhposyan (2011).

4.2 Models Used in Forecasting Experiments

A summary of the models used in our prediction experiments is given below.

Small Data Models

Autoregressive (AR) and Vector Autoregressive (VAR) Models:

Models in this section are summarized in Table 1, and include: AR(1), VAR(1), AR(SIC), and VAR(SIC).

We utilize a number of benchmark time series models, specified as follows:

$$y_{t+h}(\tau) = c(\tau) + \delta_y' W_t + \varepsilon_{t+h}, \quad (4.3)$$

where τ denotes the maturity of a bond (bill) for which the scalar, $y_{t+h}(\tau)$, measures the annual yield. Additionally, W_t contains lags of $y_t(\tau)$ in autoregressive specifications, and contains lags of $y_t(\tau)$ and additional explanatory variables in vector autoregressive specifications, with δ_y a conformably defined

coefficient vector. $c(\tau)$ contains the constant term.⁶ In AR and VAR specifications, up to 5 lags of $y_t(\tau)$ are included, with the number of lags selected using the Schwarz information criterion (SIC). In addition to AR(SIC) and VAR(SIC) models, straw-man AR(1) and VAR(1) models are estimated. Additionally, in our unrestricted VAR models, W_t includes five bond yields with maturities 3 months, 1 year, 3 years, 5 years, and 10 years.

Dynamic Nelson Siegel (DNS) Models:

Models in this section are summarized in Table 1, and include: DNS(1), DNS(2), DNS(3), DNS(4), DNS(5), and DNS(6).

As discussed above, the DNS model introduced by Li and Diebold (2006) is a dynamic version of the term structure based upon Nelson and Siegel (1987), where cross-sectional movements in the term structure are summarized by the dynamics of three underlying latent factors (betas) interpreted as “level”, “slope”, and “curvature” factors. We refer to the three betas as “Nelson-Siegel factors” (NS-factors), and in our prediction experiments, both AR(1) and VAR(1) type models are specified in order to predict these factors for subsequent use in the prediction of $y_{t+h}(\tau)$.

We estimate the latent factors by fitting the following regression:

$$y_t(\tau) = \beta_{1,t} + \beta_{2,t} \cdot \left[\frac{1 - \exp(-\lambda_t \tau)}{\lambda_t \tau} \right] + \beta_{3,t} \cdot \left[\frac{1 - \exp(-\lambda_t \tau)}{\lambda_t \tau} - \exp(-\lambda_t \tau) \right] + \varepsilon_t, \quad (4.4)$$

which is discussed in Section 2.1. Again as discussed above, we utilize yield cross sections that include 10, 12, and 30 different yield maturities. Predictions of y_{t+h} are constructed using the model:

$$\hat{y}_{t+h}(\tau) = \hat{\beta}_{1,t+h}^f + \hat{\beta}_{2,t+h}^f \cdot \left[\frac{1 - \exp(-\lambda_t \tau)}{\lambda_t \tau} \right] + \hat{\beta}_{3,t+h}^f \cdot \left[\frac{1 - \exp(-\lambda_t \tau)}{\lambda_t \tau} - \exp(-\lambda_t \tau) \right], \quad (4.5)$$

where $y_{t+h}(\tau)$ is a scalar, and $\hat{\beta}_{1,t+h}^f$, $\hat{\beta}_{2,t+h}^f$, and $\hat{\beta}_{3,t+h}^f$ are predictions constructed by specifying simple AR(1) models for $\hat{\beta}_{1,t}$, $\hat{\beta}_{2,t}$, and $\hat{\beta}_{3,t}$, including:

$$\hat{\beta}_{i,t+h}^f = \hat{c}_i + \hat{\gamma}_{y,i} \hat{\beta}_{i,t}, \quad \text{for } i = 1, 2, 3, \quad (4.6)$$

where $\hat{\beta}_{i,t+h}^f$, $\hat{\beta}_{i,t}$, \hat{c}_i and $\hat{\gamma}_i$ are scalars. Note that \hat{c}_i is an element of \hat{c}_y in equation (3.1). Also, $\hat{\gamma}_{y,i}$ is an element of $\hat{\Gamma}_y$ as defined in equation (3.1). We also construct predictions by using the following VAR(1) model:

$$\hat{\beta}_{t+h}^f = \hat{c}_y + \hat{\Gamma}_y \hat{\beta}_t, \quad (4.7)$$

where $\hat{\beta}_{t+h}^f = (\hat{\beta}_{1,t+h}^f, \hat{\beta}_{2,t+h}^f, \hat{\beta}_{3,t+h}^f)'$, \hat{c} is a 3×1 vector, and $\hat{\Gamma}_y = (\hat{\gamma}_1, \hat{\gamma}_2, \hat{\gamma}_3)$, with $\hat{\gamma}_j$ a 3×1 vector, for $j = 1, 2, 3$. In our experiments, the decay parameter is estimated both statically and dynamically

⁶When specifying VAR models, equation (4.3) constitutes only one (τ -maturity) equation in the VAR. As the same set of explanatory variables is utilized in each equation in the VAR, the SUR (seemingly unrelated regression) result ensures that consistent and efficient parameter estimates can be obtained via application of equation by equation least squares.

(prior to the construction of each new prediction). For prediction models with a static rate of decay (i.e., models DNS(1) and DNS(4) in Table 1), λ_t is set equal to 0.0609, as in Diebold and Li (2006). DNS(2), DNS(3), DNS(5), and DNS(6)) utilize a dynamically estimated decay parameter, which is estimated as follows. First, a grid search for the decay parameter ($\frac{1}{\lambda_t}$) is carried out on the domain of (6.69, 33.46), which corresponds to the domain of a “curvature hump” of one to five years. The range for the grid search is selected on the basis of bond maturities.⁷ Next, NS-factors are calculated with the selected rate of decay for the “curvature” factor that minimizes squared in-sample fitted errors. Finally, either an AR(1) or VAR(1) models are estimated in order to generate forecasts of the NS-factors, as discussed above.

Dynamic Nelson-Siegel-Svensson (NSS) Models:

Models in this section are summarized in Table 1, and include: NSS(1), NSS(2), NSS(3), NSS(4), NSS(5), and NSS(6).

The dynamic Nelson-Siegel-Svensson (NSS) model is included in our prediction experiments because it is one of the most widely used in zero-coupon yield curve construction by major central banks (see BIS (2005)). As discussed above, in the model, Svensson (1994) adds an additional factor to the classic 3-factor Nelson-Siegel model that captures a second “curvature hump”. In our experiments, the four latent factors are referred to as “Nelson-Siegel-Svensson factors” (NSS-factors). Although Svensson did not consider a dynamic version of his model in his original paper, we do so, following the approach of Diebold and Li (2006). The framework of our prediction experiments using the NSS model is, thus, analogous to that discussed above in the case of DNS model. In particular, estimates of the NSS-factors (i.e. $\beta_{1,t}$, $\beta_{2,t}$, $\beta_{3,t}$, and $\beta_{4,t}$) are constructed at each point in time by regressing $(1, \frac{1 - \exp(-\lambda_{1,t}\tau)}{\lambda_{1,t}\tau}, \frac{1 - \exp(-\lambda_{1,t}\tau)}{\lambda_{1,t}\tau} - \exp(-\lambda_{1,t}\tau), \frac{1 - \exp(-\lambda_{2,t}\tau)}{\lambda_{2,t}\tau} - \exp(-\lambda_{2,t}\tau))$ on $y_t(\tau)$. Additionally, the model is now:

$$y_t(\tau) = \beta_{1,t} + \beta_{2,t} \cdot \left[\frac{1 - \exp(-\lambda_{1,t}\tau)}{\lambda_{1,t}\tau} \right] + \beta_{3,t} \cdot \left[\frac{1 - \exp(-\lambda_{1,t}\tau)}{\lambda_{1,t}\tau} - \exp(-\lambda_{1,t}\tau) \right] + \beta_{4,t} \cdot \left[\frac{1 - \exp(-\lambda_{2,t}\tau)}{\lambda_{2,t}\tau} - \exp(-\lambda_{2,t}\tau) \right] + \varepsilon_t, \quad (4.8)$$

Resultant sequences of estimates, $\hat{\beta}_{1,t}$, $\hat{\beta}_{2,t}$, $\hat{\beta}_{3,t}$, and $\hat{\beta}_{4,t}$, for $t = 1, \dots, T$ are used to construct predictions of $y_{t+h}(\tau)$ using:

$$\hat{y}_{t+h}(\tau) = \hat{\beta}_{1,t+h}^f + \hat{\beta}_{2,t+h}^f \cdot \left[\frac{1 - \exp(-\lambda_{1,t}\tau)}{\lambda_{1,t}\tau} \right] + \hat{\beta}_{3,t+h}^f \cdot \left[\frac{1 - \exp(-\lambda_{1,t}\tau)}{\lambda_{1,t}\tau} - \exp(-\lambda_{1,t}\tau) \right] + \hat{\beta}_{4,t+h}^f \cdot \left[\frac{1 - \exp(-\lambda_{2,t}\tau)}{\lambda_{2,t}\tau} - \exp(-\lambda_{2,t}\tau) \right] \quad (4.9)$$

⁷We find that setting domains too wide results in occasional ‘extreme’ estimates for NS-factors, which in turn leads to occasional poor yield forecasts. For further discussion, see below. For an extensive discussion of decay parameter estimation, refer to De Pooter (2007).

where $y_{t+h}(\tau)$ is a scalar, and $\hat{\beta}_{1,t+h}^f$, $\hat{\beta}_{2,t+h}^f$, $\hat{\beta}_{3,t+h}^f$, and $\hat{\beta}_{4,t+h}^f$ are predictions constructed by specifying simple AR models:

$$\hat{\beta}_{i,t+h}^f = \hat{c}_i + \hat{\gamma}_{y,i} \hat{\beta}_{i,t}, \quad \text{for } i = 1, 2, 3, 4 \quad (4.10)$$

where $\hat{\beta}_{i,t+h}^f$, $\hat{\beta}_{i,t}$, \hat{c}_i , and $\hat{\gamma}_{y,i}$ are scalars. We also construct NSS-factor predictions by using the following VAR(1) model:

$$\hat{\beta}_{t+h}^f = \hat{c}_y + \hat{\Gamma}_y \hat{\beta}_t, \quad (4.11)$$

where $\hat{\beta}_{t+h}^f = (\hat{\beta}_{1,t+h}^f, \hat{\beta}_{2,t+h}^f, \hat{\beta}_{3,t+h}^f, \hat{\beta}_{4,t+h}^f)'$, $hate_y$ is 4×1 vector, and $\hat{\Gamma}_y$ is a 4×4 matrix of constants. To estimate NSS model parameters, again, two estimation methods for the decay parameters are utilized. In the case of a fixed (static) decay parameter, $\lambda_{1,t}$ is equal to 0.0609, which is the same value as that used when estimating our three-factor DNS model; and the second rate of decay, $\lambda_{2,t}$, is set equal to 0.2985, corresponding to a second curvature hump at approximately 6 months.⁸ The subsequent forecasting procedure used to construct yield predictions is the same as that discussed above for our DNS models.

Big Data Models

AR and VAR Models with Macro Diffusion Indexes:

Models in this section are summarized in Table 1, and include: AR(1)+FB1, AR(1)+FB2, AR(1)+FB3, VAR(1)+FB1, VAR(1)+FB2, and VAR(1)+FB3.

We utilize the prediction model given in equation (4.3), but with latent factors (i.e., diffusion indexes), F_t^x , estimated using PCA with a real-time macroeconomic dataset (see Section 3 for a discussion of diffusion indexes and Section 5 for a discussion of the data used in our analysis). In particular, we estimate variants of the following factor augmented forecasting model:

$$y_{t+h}(\tau) = c(\tau) + \delta_y' W_t + \delta_x' F_t^x + \varepsilon_{t+h}, \quad (4.12)$$

where F_t^x includes either 1, 2 or 3 diffusion indexes, and W_t is defined as above, yielding AR and VAR variants of these models. Here, $c(\tau)$ is a constant term, and δ_y and δ_x are conformably defined vectors of coefficients, as discussed in Section 3. Note that although multiple yield lags were tried when specifying W_t , ‘MSFE-best’ models always included only the first lag of the yield(s). For this reason all empirical results discussed in the sequel use one lag.

⁸Restrictions on the decay parameters for the NSS model are imposed to ensure that the two curvature humps are at least one year apart, for identification purposes. In addition to this restriction that $\frac{1}{\lambda_{1,t}} \geq \frac{1}{\lambda_{2,t}} + 6.69$ (see De Pooter (2007)), restrictions are imposed on the domain of the two decay parameters. Namely, the grid search for the first decay parameter $\frac{1}{\lambda_{1,t}}$ is over the domain of (6.69, 33.46); and for the second decay parameter $\frac{1}{\lambda_{2,t}}$ is on (0, 26.77). These restrictions ensure identification of two curvature factors individually, and avoids ‘extreme’ estimates for NSS-factors.

As discussed above, the real-time diffusion indexes appearing in the above equation (as well as subsequently) are constructed using both un-targeted and targeted PCA. Un-targeted PCA is carried out by simply estimating indexes with the entire real-time dataset that is available at each point in time, prior to the construction of each new prediction. Targeted PCA is the same, except that a new subset of variables from the real-time dataset is used at each point in time to construct the indexes, with the subset selected using either the elastic net or the lasso. This method is computationally intensive, as these shrinkage operators must be applied at each point in time, and for each targeted interest rate being predicted. Ten-fold cross validation is used, in real-time, to estimate the tuning parameter in the operators.⁹ For a complete description of these operators and targeted prediction in general, refer to Bai and Ng (2007,2008), Carrasco and Rossi (2016), Ghysels and Marcellino (2018), Kim and Swanson (2014), and Schumacher (2007,2009).

DNS Models with Macro Diffusion Indexes:

*Models in this section are summarized in Table 1, and include: DNS(1)+FB1, DNS(2)+FB1, DNS(3)+FB1, DNS(4)+FB1, DNS(5)+FB1, DNS(6)+FB1, DNS(1)+FB2, DNS(2)+FB2, DNS(3)+FB2, DNS(4)+FB2, DNS(5)+FB2, DNS(6)+FB2, DNS(1)+FB3, DNS(2)+FB3, DNS(3)+FB3, DNS(4)+FB3, DNS(5)+FB3, DNS(6)+FB3.*¹⁰

In this section, diffusion indexes (principle components) constructed using macro variables are augmented to DNS models discussed above. Namely, we considered DNS type predictions constructed using:

$$\hat{\beta}_{i,t+h}^f = \hat{c}_i + \hat{\gamma}'_{y,i} \hat{\beta}_{i,t} + \hat{\gamma}'_{x,i} F_t^x, \quad \text{for } i = 1, 2, 3,$$

where F_t^x again includes either 1, 2 or 3 latent factor(s). All other terms are conformably defined. We also construct predictions by using the following VAR(1) variant of this model:

$$\hat{\beta}_{t+h}^f = \hat{c}_y + \hat{\Gamma}_y \hat{\beta}_t + \hat{\Gamma}_x F_t^x,$$

where $\hat{\beta}_{t+h}^f = (\hat{\beta}_{1,t+h}^f, \hat{\beta}_{2,t+h}^f, \hat{\beta}_{3,t+h}^f)'$, \hat{c}_y is 3×1 vector, $\hat{\Gamma}_y = (\hat{\gamma}_1, \hat{\gamma}_2, \hat{\gamma}_3)$, with $\hat{\gamma}_j$ a 3×1 vector, for $j = 1, 2, 3$, and $\hat{\Gamma}_x$ is a conformably defined matrix of constants.

NSS Models with Macro Diffusion Indexes:

Models in this section are summarized in Table 1, and include: NSS(1)+FB1, NSS(2)+FB1, NSS(3)+FB1, NSS(4)+FB1, NSS(5)+FB1, NSS(6)+FB1, NSS(1)+FB2, NSS(2)+FB2, NSS(3)+FB2, NSS(4)+FB2,

⁹For the elastic net, there are two tuning parameters (rather than 1, as is the case with the lasso). These parameters control the weights placed on the L1- and L2-norm components of the loss function. While the first tuning parameter is estimated with cross validation, as mentioned above, the second parameter is fixed at various values, including 0.2, 0.4, 0.6, and 0.8.

¹⁰FB1, FB2, and FB3 denote models that have been augmented to include either 1, 2, or 3 diffusion indexes, respectively.

$NSS(5)+FB2$, $NSS(6)+FB2$, $NSS(1)+FB3$, $NSS(2)+FB3$, $NSS(3)+FB3$, $NSS(4)+FB3$, $NSS(5)+FB3$, $NSS(6)+FB3$.

Analogous to our DNS models, NSS model predictions constructed using:

$$\hat{\beta}_{i,t+h}^f = \hat{c}_i + \hat{\gamma}'_{y,i} \hat{\beta}_{i,t} + \hat{\gamma}'_{x,i} F_t^x, \quad \text{for } i = 1, 2, 3, 4$$

where F_t^x includes either 1, 2 or 3 latent factors. All other terms are conformably defined and analogous to our above discussion. We also construct predictions using the following VAR(1) variant of this model:

$$\hat{\beta}_{t+h}^f = \hat{c}_y + \hat{\Gamma}_y \hat{\beta}_t + \hat{\Gamma}_x F_t^x,$$

where $\hat{\beta}_{t+h}^f = (\hat{\beta}_{1,t+h}^f, \hat{\beta}_{2,t+h}^f, \hat{\beta}_{3,t+h}^f, \hat{\beta}_{4,t+h}^f)'$, \hat{c} is 4×1 vector, and $\hat{\Gamma}_y = (\hat{\gamma}_1, \hat{\gamma}_2, \hat{\gamma}_3, \hat{\gamma}_4)$, $\hat{\gamma}_j$ is a 4×1 vector, for $j = 1, 2, 3, 4$ and $\hat{\Gamma}_x$ is a conformably defined matrix of constants.

As a final note, it is worth mentioning that all macroeconomic variables are standardized to zero mean and unit variance series before principle component analysis is utilized to construct diffusion indexes.

Forecast Combination:

In our experiments, we also construct and analyze various forecast combination models. The particular combinations are detailed in Table 8. Although the focus of this paper is not forecast combination, there are two reasons why we include combinations in our analysis. First, it is well known that forecast combination is useful in time series prediction. As shown in Kim and Swanson (2014), Carrasco and Rossi (2016), and Hirano and Wright (2017), much can be gained by combining forecasts from models estimated using a wide range of methods ranging from least squares estimation to machine learning.¹¹ More importantly, it turns out that while combination does not play an important role when comparing DNS and NSS type models with and without diffusion indexes if fully revised data are used in model and prediction construction, as discussed in Xiong and Swanson (2017), the same is not true when real-time data are used in our data rich environment. Indeed, we shall see that various forecast combinations dominate all of the models discussed above, when real-time data are utilized. This is important because it suggests that the use of fully revised data may be quite misleading in the types of experiments carried out in this paper.

Finally, note that in all experiments, models are estimated using rolling windows of 120 monthly observations, as discussed above. Thus, all models are re-estimated prior to the construction of each new h -step ahead forecast. Additionally, the first observation in our dataset is August 1988, and experiments are carried out for 4 different prediction periods, as discussed in the next section.

¹¹For a discussion of forecast combination using the types of factor augmented regressions discussed in this paper, see Cheng and Hansen (2015).

5 Data

Yield Data:

Our term structure data are monthly U.S. zero-coupon (end of month) yield curve data reported by the Federal Reserve Board (see <https://www.quandl.com/data/FED/SVENY-US-Treasury-Zero-Coupon-Yield-Curve> and Gürkaynak, Sack and Wright (GSW: 2006)). In particular, we utilize GSW monthly data for the August 1988 through October 2017, which contains data on 1 to 30 years maturity bond yields. In addition to GSW zero-yields, 3- and 6-months T-bill yields¹² are utilized in order to “fill-out” the short end of the yield curve. Hence, we analyze a panel of dataset containing $N = 32$ dimensional yields and $T = 351$ monthly observations. When constructing betas, we consider three variants of this data. In one variant, we utilize 12 yields (i.e., 3- and 6-months, 1 year, 2 year, ..., and 10 year yields); in a second variant, we utilize 10 yields, as done in Xiong and Swanson (2017) (i.e., 1 year, 2 year, ..., and 10 year yields); and in a third variant we utilize 30 yields (i.e., 1 year, 2 year, ..., and 30 year yields). Plots of the 12-dimensional yield dataset, which is used in the majority of our MSFE-best models when utilized for constructing DNS and NSS factors, are contained in Figures 2-4. In Figure 2, these yield data are plotted against real-time diffusion indexes constructed using un-targeted PCA. Of note is that the indexes increase markedly in volatility around the Great Recession, and exhibit the lowest historical volatility levels thereafter. Figure 3 and 4 plot the same yield data against DNS and NSS factors, constructed using the 12-dimensional yield dataset. While the first (key) factor constructed using the DNS and NSS models are quite similar (compare Figure 3 and Figure 4), the other factors are evidently quite different. These differences might account in large part for the relative forecasting performance differences between the two models.

While Dickey-Fuller tests cannot reject the null hypothesis of a unit root in yields, preliminary forecast experiments using both yields and first-difference yields resulted in little differences when comparing MSFEs of yield predictions. Moreover, finance theory is not consistent with the presence of a unit root in yield processes. For these reasons, we use only yield “levels” data in our experiments.

Macroeconomics Variables:

As discussed above, macroeconomic factors (i.e., real-time diffusion indexes) are constructed using PCA. The dataset used is the FRED-MD dataset, which is a real-time monthly database of over 130 macroeconomic time series that covers categories ranging from output and income, to labor market,

¹²3- and 6-months T-bill yields are constant-maturity, as reported in the FRED database of the Federal Reserve Bank of St. Louis. This “hybrid” zero-yield dataset is widely utilized in yield curve estimation, see Gürkaynak and Wright (2012), Hamilton and Wu (2012).

prices, and interest rates. The FRED-MD dataset is developed and maintained by the Federal Reserve Bank of St. Louis. For details, see McCracken and Ng (2016), and for access to the dataset, visit <https://research.stlouisfed.org/econ/mccracken/fred-databases/>. In their paper, McCracken and Ng (2016) conduct an empirical research which shows that diffusion indexes extracted from the dataset share the same predictive content as those based on the classic (non-real-time) Stock and Watson (2002a) dataset used so frequently in analyses such as ours. As discussed above, one advantage of FRED-MD is that all time series are updated monthly by the Federal Reserve Bank of St. Louis. Thus, researchers have truly real-time data available for conducting forecasting experiments, in which all vintages (revisions) of all variables are available. Use of such data ensures that future information cannot inadvertently be used to revise data from prior periods, which is a serious potential problem with non-real-time or fully revised data. Refer to Stark (2010) for further discussion of real-time datasets. Moreover, fully revised datasets “mix” vintages of observations, in the sense that the most recent observation in a fully revised dataset is a so-called “first release”, while earlier calendar dated observations have possibly been revised and re-released many times. Interestingly, we find that real-time data does matter, as findings from Xiong and Swanson (2017) supporting the use of macro diffusion indexes in DNS type models are reversed when real-time instead of fully revised data are used in index construction. Finally, note that only the use of real-time datasets makes it possible to replicate truly real-time modeling and forecasting of U.S. Treasury yields, when using macroeconomic data that are subject to revision. In order to illustrate the breadth of information contained in the FRED-MD, Figure one plots all 130 macroeconomic series against the 1-year Treasury yield. It is interesting to note that periods of increased volatility in the yield appear to coincide roughly with an increase in the spread and volatility of the extreme values associated with the aggregate FRED-MD dataset.

6 Empirical Findings

Our empirical investigation utilizes the models and methods discussed in Section 4. Our objective is to predict U.S. Treasury yields at various maturities. Predictions are made using “small data” models (i.e. models which only use historical yield cross sections for calibrating models), as well as “big data” models that include real-time diffusion indexes constructed from the real-time FRED-MD dataset discussed in Section 5. Our small data models include autoregressive, vector autoregressive, DNS and NSS models, and our big data models are specified as pure diffusion index models or as hybrids that combine our small data models with diffusion indexes. Finally, diffusion indexes are constructed using targeted (machine learning based) PCA as well as un-targeted PCA, as discussed in Section 4.

A number of clear-cut conclusions emerge upon examination of the results collected in Tables 1-4.

Let us first discuss the results summarized in Tables 2A-C. These tables include: (i) the three “MSFE-best” (i.e., lowest MSFE model) models for each yield maturity/forecast horizon permutation, in de-

scending order from 1st to 3rd (see Table 2A); (ii) MSFE for the three models listed in Table 2A (see Table 2B); and relative MSFEs (relative to the RW benchmark) for the 3 models (see Table 2C).¹³ All of these tables report results for DNS and NSS type models with betas constructed either using our 10, 12-, or 30-dimensional historical yield dataset, as denoted by subscripts 1, 2, and 3, respectively, in Table 2A. Finally, results are presented for 3 forecast horizons ($h = 1, 3, 12$), for 6 yield maturities (3- and 6-month, 1 year, 3 years, 5 years, and 10 years), and for 4 different forecasting periods, including: 2001:1-2005:12 (Subsample 1), 2006:1-2010:12 (Subsample 2), 2011:1-2017:10 (Subsample 3), and 2001:1-2017:10 (Subsample 4). Illustrating the layout of the tables, we see in Table 2A that for 3-month maturity yields, at the 1-step ahead forecast horizon, model NSS(4)+FB₁ yields the lowest MSFE, for Subsample 1. This means that real-time diffusion indexes constructed FRED-MD add marginal predictive content to the NSS model, and this combination is the MSFE-best combination. In this model, the subscript “2” indicates that the 12-dimensional yield dataset produced NSS factors that results in the model “win”. Additionally, notice that this model is superscripted with a “***”, indicating rejection of the Diebold-Mariano null hypothesis of equal predictive accuracy, when compared with the RW benchmark, at the 1% level. For example, for the 1st subsample, at the $h = 1$ -step ahead horizon, all MSFE models include real-time diffusion indexes, except for the model that “wins” for 5-year maturity forecasts. In addition, for 4 of 5 maturities, the 12-dimensional yield dataset is most useful for estimating NS factors, and the 10-dimensional dataset never yields the MSFE-best model. This is perhaps not surprising, as our 12- and 30-dimensional yield datasets augment the GW dataset by adding short-end yields (i.e., 3- and 6-month yields), while the 10-dimensional dataset only includes annual yields, from 1- to 10-years in maturity.

Notice that many of the models that are in the top three MSFE-best in Tables 2A-C include diffusion indexes (i.e., models with FB1, FB2, or FB3, in their names, corresponding to the addition of 1, 2, or 3 real-time diffusion indexes). Additionally, many models are of the DNS and NSS variety. This pattern is quite prevalent across all forecast horizons and bond maturities, for our first two subsamples. The pattern is less prevalent, during the third subsample, from 2011-2017. Still, we have rather strong evidence that DNS and NSS models are useful, as previously found by many authors, and that the usefulness of all models including DNS type models can often be improved by including real-time diffusion indexes. As just stated, this result is not ubiquitous, however. For example, in the third subsample (post Great Recession), diffusion indexes are only in one MSFE-best model, indicating a deterioration in predictive gains associated with using diffusion indexes, post Great Recession. However, they are in the 2nd and

¹³Results relative to AR(1) and AR(SIC) benchmarks were also tabulated, but due to the poor performance of these models, relative to the RW, in recent years, tabulated results are not as clear when these alternative numeraires are used when reporting relative MSFEs. For discussion of the zero lower bound issue affecting predictions from term structure models in this context, Hamilton and Wu (2012). In addition, note that all results reported in these tables have been compiled for each of the models analyzed. However, the number of tables is overwhelming when including all models, maturities and forecast horizons, and hence these results have been omitted for the sake of brevity. Complete results are available upon request from the authors.

3rd “best” models, for numerous yield maturities and forecast horizons. Summarizing, the usefulness of diffusion indexes appears to be somewhat sample dependent. This is not surprising, and suggests, for example, a possible use for hybrid models in which the inclusion of diffusion indexes is triggered by variables such as predicted probabilities of recessions, economic variability, or the range of yields over some pre-defined prior period of time. Needless to say, one reason why this result arises is associated with the zero-lower bound issue discussed earlier in this paper, and in many other recent papers. Evidently, in the current low interest rate regime, the RW model is a difficult univariate model to “beat”. However, it is interesting to note that even in such times, DNS and NSS models perform very well when compared with the RW model, albeit without the need for real-time diffusion indexes. Thus, a key take-away from Table 2A (compare also Tables 2B-C) is that real-time diffusion indexes are not as useful in the current low-interest rate environment as they were in the pre-Great recession environment.

As stated above, inspection of Tables 2A-C indicates that DNS or NSS type models are often the MSFE-best models. In Table 2A DNS or NSS type models are MSFE-best in 13 of 15 maturity/horizon permutations for Subsample 1, and are top three MSFE performers in 37 of 45 maturity/horizon permutations, when counting all 1st, 2nd, and 3rd place “winners”. In Subsample 2, results are similar, where analogous “wins” are 11 of 15 and 36 of 45. Finally, in Subsample 3, DNS or NSS type models “win” in only 3 of 15 cases and 22 of 45 cases, respectively. These findings are consistent with findings in the recent literature suggesting that DNS type model performance has deteriorated in recent post credit crisis years (see, e.g. Altavilla, Giacomini and Ragusa (2017), Diebold, and Rudebusch (2013), and Mönch (2008)).

Turning to our analysis of forecast combinations, note that Table 3 lists the combination models that were examined in our analysis.¹⁴ Relative MSFEs for all combinations are given in Table 4A ($h = 1$ -step ahead forecasts), Table 4B ($h = 3$ -step ahead forecasts), and Table 4C ($h = 12$ -step ahead forecasts). Results in these tables are quite interesting. For instance, models with real-time diffusion indexes (e.g., FB1, FB2, and FB3) often significantly outperform the RW benchmark, for all five maturities and all three forecast horizons in Subsamples 1, 2, and 4. However, results are mixed for Subsample 3 (2011:1-2017:10), again indicating a deterioration in predictive gains associated with using diffusion indexes, in low interest rate environments. Essentially, the mechanism that we find strong evidence of pre-2010, which ties unspanned risks to the term-structure, seems to break down in extremely low interest rate regimes. Another conclusion emerges when comparing NS(AR) and NS(VAR) forecast combination models. As discussed previously, NS-factors and NSS-factors can be better predicted by modeling their cross-correlation dynamics. This is borne out in the data. For example, in Subsample 1, 3, and 4, forecasts generated by combination model NS(VAR) have lower point MSFEs than NS(AR) models in 43/45 cases, across all three forecast horizons, and across all five bond maturities.

A key result to emerge from our combination experiments is that the MSFE-best models are almost

¹⁴Various other forecast combinations were also analyzed, but yielded poorer predictive accuracy than that based on the methods reported in here, and are not reported, for the sake of brevity.

always “FS” type forecast combinations (compare the bolded entries in each column for each subsample in Tables 4A-C). As noted in Table 3, FS forecast combination models utilize the average of all non-diffusion index type models (i.e., AR(1), AR(SIC), VAR(1), VAR(SIC), DNS(1) - DNS(6), NSS(1) - NSS(6)). Indeed, FS models “win” in 16 of 20 cases, for $h = 1$, across all five bond maturities and all four subsamples (see Table 4A). The MSFE-best combination in 3 of the other 4 cases is an NSS model. For $h = 3$, the FS combination model “wins” in 19 of 20 cases (see Table 4B); and for $h = 12$, the FS combination model “wins” in 20 of 20 cases (see Table 4C). Furthermore, all of the “best” point MSFEs in these tables are much lower than point MSFEs associated with the best individual models. Thus, combination dominates under our real-time setup, and the best combinations do not utilize macro diffusion indexes. This result differs markedly from that reported in Xiong and Swanson (2017), where big data matters. However, Xiong and Swanson (2017) carry out experiments using a fully revised macroeconomic dataset rather than a real-time macroeconomic dataset and do not consider shrinkage methods. Two conclusions can thus be drawn from our combination experiments. First, fully revised data may have an important confounding effect upon results obtained when instead carrying out real-time prediction experiments. Second, real-time diffusion indexes do matter when comparing the predictive performance of individual models, indicating the presence of unspanned risks. However, there appear to be two different ways to “capture” these unspanned risks. One is to use data rich real-time diffusion indexes, and another is to simply combine predictions from many non-data rich models.

7 Concluding Remarks

In this paper, we examine the usefulness of real-time macroeconomic diffusion indexes when using dynamic Nelson-Siegel (DNS), dynamic Nelson-Siegel Svensson (NSS), and various econometric models for forecasting the term structure of interest rates. We find that the marginal predictive content of real-time diffusion indexes is significant for many of the models that we examine. We also find that model performance, across the board, is much worse post Great Recession. Indeed, not only does the predictive performance of DNS and NSS models worsen, in accord with the findings of various recent authors, but the performance of all of our models, including ones that utilize real-time diffusion indexes also worsens. Two key conclusions that we make are the following. First, fully revised data may have an important confounding effect upon results obtained when instead carrying out real-time prediction experiments. Second, real-time diffusion indexes matter when comparing the predictive performance of individual models, indicating the presence of unspanned macroeconomic risks in the term structure of interest rates. However, there appear to be two different ways to “capture” these unspanned risks. One is to use data rich real-time diffusion indexes, and another is to simply combine predictions from many non-data rich models.

Given the impressive predictive performance of models that include real-time diffusion indexes prior

to 2010, and in particular of DNS and NSS models augmented to include such indexes, we argue that new models for low interest variability and zero lower-bound regimes need to be developed. Examples of models that might be useful include hybrid models in which the inclusion of diffusion indexes is triggered by variables such as predicted probabilities of recessions, economic variability, or the range of yields over some pre-defined prior period of time.

8 References

- Altavilla, C., R. Giacomini and G. Ragusa (2017), Anchoring the Yield Curve Using Survey Expectations, *Journal of Applied Econometrics*, 32, 1055-1068.
- Andreou, E., P. Gagliardini, E. Ghysels, and M. Rubin (2018), Is Industrial Production Still the Dominant Factor for the US Economy? Evidence from a New Class of Mixed Frequency (Group) Factor Models, Swiss Finance Institute Research Paper No. 16-11.
- Ang, A. and M. Piazzesi (2003), A No-Arbitrage Vector Autoregression of Term Structure Dynamics With Macroeconomic and Latent Variables, *Journal of Monetary Economics*, 50, 745-787.
- Bai, J. and S. Ng (2008), Forecasting Economic Time Series Using Targeted Predictors, *Journal of Econometrics*, 146, 304-317.
- Bai, J. and S. Ng (2009), Boosting Diffusion Indices, *Journal of Applied Econometrics*, 24, 607-629.
- Banerjee, A., M. Marcellino and I. Marsten (2008), Forecasting Macroeconomic Variables Using Diffusion Indexes in Short Samples with Structural Change, in **Forecasting in the Presence of Structural Breaks and Model Uncertainty**, pp. 149-194, Emerald Group Publishing, New York.
- Bernanke, B. S. and Boivin, J. (2003), Monetary Policy in a Data-rich Environment, *Journal of Monetary Economics*, 50, 525-546.
- BIS (2005), Zero-Coupon Yield Curves: Technical Documentation, Working Paper, Bank of International Settlements.
- Boivin, J. and S. Ng (2005), Understanding and Comparing Factor Based Macroeconomic Forecasts, *International Journal of Central Banking*, 1, 117-152.
- Carrasco, M. and B. Rossi, (2016), In-Sample Inference and Forecasting in Misspecified Factor Models, *Journal of Business and Economic Statistics*, 34, 313-338.
- Cheng, X. and Hansen, B. E. (2015), Forecasting with Factor-Augmented Regression: A Frequentist Model Averaging Approach, *Journal of Econometrics*, 186, 280-293.
- Christensen, J.H.E., F.X. Diebold, and G.D. Rudebusch (2011), The Affine Arbitrage Free Class of Nelson-Siegel Term Structure Models, *Journal of Econometrics*, 164, 4-20.
- Clark, T. E. and McCracken, M. W. (2009), Improving Forecast Accuracy by Combining Recursive and Rolling Forecast, *International Economic Review*, 50, 363-395.
- Corradi, V. and N.R. Swanson (2006), Predictive Density Evaluation, in G. Elliot, C. W. J. Granger, and A. Timmermann, (eds.), **Handbook of Economic Forecasting**, Volume 1, pp. 197-284, Elsevier, Amsterdam.
- Coroneo, L., D. Giannone and M. Modugno (2016), Unspanned Macroeconomic Factors in the Yield Curve, *Journal of Business and Economic Statistics*, 34, 472-485.
- De Pooter, M., (2007), Examining the Nelson-Siegel Class of Term Structure Models: In-Sample Fit versus Out-of-Sample Forecasting Performance, *Tinbergen Institute Discussion Paper*, No. 07-043/4.

- Diebold, F.X. and Li, C. (2006), Forecasting the Term Structure of Government Bond Yields, *Journal of Econometrics*, 130, 337-364.
- Diebold, F.X. and R.S. Mariano (1995), Comparing Predictive Accuracy, *Journal of Business and Economic Statistics* 13, 253-263.
- Diebold, F. X. and Rudebusch, G. D. (1991), Forecasting Output with the Composite Leading Index: A Real-Time Analysis, *Journal of the American Statistical Association*, 86(415), 603-610.
- Diebold, F.X., G.D. Rudebusch, and S. B. Aruoba (2006), The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach, *Journal of Econometrics*, 131, 309-338.
- Diebold, F.X. and G.D. Rudebusch (2013), **Yield Curve Modeling and Forecasting: The Dynamic Nelson-Siegel Approach**, Princeton University Press: Princeton.
- Duffee, G.R. (2011), Information In (and Not In) the Term Structure, *Review of Financial Studies*, 24, 2895-2934.
- Duffee, G. R. (2013), Forecasting Interest Rates, in **Handbook of Economic Forecasting**, Volume 2, pp. 385-426, Elsevier, Amsterdam.
- Exterkate, P., van Dijk, D., Heij, C., and Groenen, P.J. (2013), Forecasting the Yield Curve in a DataRich Environment Using the FactorAugmented NelsonSiegel Model, *Journal of Forecasting*, 32(3), 193-214.
- Ghysels, E. and M. Marcellino (2018), **Applied Economic Forecasting using Time Series Models**, Oxford University Press, London.
- Gürkaynak, R.S., B. Sack, and J.H. Wright (2006), The U.S. Treasury Yield Curve: 1961 to the Present, Working Paper, Federal Reserve Bank - Finance and Economics Discussion Series 2006-28.
- Gürkaynak, R.S. and J.H. Wright (2012), Macroeconomics and the Term Structure, *Journal of Economic Literature*, 50, 331-67.
- Hamilton, J. D. and G. Perez-Quiros (1996), What Do the Leading Indicators Lead?, *Journal of Business*, 69, 27-49.
- Hamilton, J.D. and J.C. Wu (2012), The Effectiveness of Alternative Monetary Policy Tools in a Zero Lower Bound Environment, *Journal of Money, Credit and Banking*, 44, 3-46.
- Hansen, P.R. and A. Timmermann (2012), Choice of Sample Split in Out-of-Sample Forecast Evaluation, Working Paper, Rady School of Management, University of California, San Diego.
- Hirano, K. and J.H. Wright (2017), Forecasting With Model Uncertainty: Representations and Risk Reduction, *Econometrica*, 85, 617-643.
- Kim, H.H. and N.R. Swanson, (2014), Forecasting Financial and Macroeconomic Variables Using Data Reduction Methods: New Empirical Evidence, *Journal of Econometrics*, 178, 352-367.
- Ludvigson, S. C. and S. Ng (2009), Macro Factors in Bond Risk Premia, *Review of Financial Studies*, 22, 5027-5067.
- McCracken, M.W. (2000), Robust Out-of-Sample Inference, *Journal of Econometrics*, 99, 195-223.

- McCracken, M.W. and S. Ng (2016), Fred-MD: A Monthly Database for Macroeconomic Research, *Journal of Business and Economic Statistics*, 34, 574-589.
- Mönch, E. (2008), Forecasting the Yield Curve in a Data-Rich Environment: A No-Arbitrage Factor-Augmented VAR Approach, *Journal of Econometrics*, 146, 26-43.
- Nelson, C. and A. Siegel (1985), Parsimonious Modeling of Yield Curves for US Treasury Bills, Working Paper 1594, National Bureau of Economic Statistics.
- Nelson, C. and A. Siegel (1987), Parsimonious Modeling of Yield Curves, *Journal of Business*, 60, 473-489.
- Rossi, B. and S. Sekhposyan (2011), Understanding Models' Forecasting Performance, *Journal of Econometrics*, 164, 158-172.
- Rossi, B. and A. Inoue (2012), Out-of-Sample Forecast Tests Robust to the Choice of Window Size, *Journal of Business and Economic Statistics*, 30, 432-453.
- Rudebusch, G. D. and T. Wu (2008), A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy, *The Economic Journal*, 118, 906-926.
- Schumacher, C. (2007), Forecasting German GDP Using Alternative Factor Models Based on Large Datasets, *Journal of Forecasting*, 26, 271-302.
- Schumacher, C. (2009), Factor Forecasting Using International Targeted Predictors: The Case of German GDP, *Economics Letters*, 107, 95-98.
- Stark, T. (2010), Realistic evaluation of real-time forecasts in the Survey of Professional Forecasters, Special Report, Federal Reserve Bank of Philadelphia Research Division.
- Stock, J.H. and M.W. Watson (2002a), Macroeconomic Forecasting Using Diffusion Indexes, *Journal of Business and Economic Statistics*, 20, 147-162.
- Stock, J.H. and M.W. Watson (2002b), Forecasting Using Principal Components from a Large Number of Predictors, *Journal of the American Statistical Association*, 97, 1167-1179.
- Stock, J. H. and M.W. Watson (2012), Generalized Shrinkage Methods for Forecasting Using Many Predictors, *Journal of Business and Economic Statistics* 30, 481-493.
- Svensson, L. E. (1994), Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994, Working Paper 4871, National Bureau of Economic Research.
- Swanson, N.R. and Xiong, W. (2017), Big Data Analytics In Economics: What Have We Learned So Far, And Where Should We Go From Here?, *Canadian Journal of Economics*, forthcoming.

Table 1: Models Used in Forecast Experiments^{*}

Model	Description
RW	Random walk model
AR(1)	Autoregressive model with one lag
AR(SIC)	Autoregressive model with lag(s) selected by the Schwarz information criterion
AR(1)+FB1	AR(1) model with diffusion index added, principle component analysis based on real-time macroeconomic dataset
AR(1)+FB2	AR(1) model with two diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
AR(1)+FB3	AR(1) model with three diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
VAR(1)	Five-dimensional vector autoregressive model with one lag
VAR(SIC)	Five-dimensional vector autoregressive model with lag(s) selected by the Schwarz information criterion
VAR(1)+FB1	VAR(1) model with one diffusion index added, principle component analysis based on real-time macroeconomic dataset
VAR(1)+FB2	VAR(1) model with two diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
VAR(1)+FB3	VAR(1) model with three diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
DNS(1)	Dynamic Nelson-Siegel (DNS) model with underlying AR(1) factor specifications fitted with twelve-dimensional yields: maturity $\tau = 12, 24, 36, 48, 60, 72, 84, 96, 108, 120$ months, with a static rate of decay parameter $\lambda = 0.0609$
DNS(2)	Dynamic Nelson-Siegel (DNS) model with underlying AR(1) factor specifications fitted with twelve-dimensional yields, with a dynamic rate of decay parameter λ_t (most recent λ_t are selected in generating predictions)
DNS(3)	Dynamic Nelson-Siegel (DNS) model with underlying AR(1) factor specifications fitted with twelve-dimensional yields, with a dynamic rate of decay parameter λ_t (median λ_t are selected in generating predictions)
DNS(4)	Dynamic Nelson-Siegel (DNS) model with underlying VAR(1) factor specifications fitted with twelve-dimensional yields, with a static rate of decay parameter $\lambda = 0.0609$
DNS(5)	Dynamic Nelson-Siegel (DNS) model with underlying VAR(1) factor specifications fitted with twelve-dimensional yields, with a dynamic rate of decay parameter λ_t (most recent λ_t are selected in generating predictions)
DNS(6)	Dynamic Nelson-Siegel (DNS) model with underlying VAR(1) factor specifications fitted with twelve-dimensional yields, with a dynamic rate of decay parameter λ_t (median λ_t are selected in generating predictions)
DNS(1)+FB1	DNS(1) model with one diffusion index added, principle component analysis based on real-time macroeconomic dataset
DNS(2)+FB1	DNS(2) model with one diffusion index added, principle component analysis based on real-time macroeconomic dataset
DNS(3)+FB1	DNS(3) model with one diffusion index added, principle component analysis based on real-time macroeconomic dataset
DNS(4)+FB1	DNS(4) model with one diffusion index added, principle component analysis based on real-time macroeconomic dataset
DNS(5)+FB1	DNS(5) model with one diffusion index added, principle component analysis based on real-time macroeconomic dataset
DNS(6)+FB1	DNS(6) model with one diffusion index added, principle component analysis based on real-time macroeconomic dataset
DNS(1)+FB2	DNS(1) model with two diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
DNS(2)+FB2	DNS(2) model with two diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
DNS(3)+FB2	DNS(3) model with two diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
DNS(4)+FB2	DNS(4) model with two diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
DNS(5)+FB2	DNS(5) model with two diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
DNS(6)+FB2	DNS(6) model with two diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
DNS(1)+FB3	DNS(1) model with three diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
DNS(2)+FB3	DNS(2) model with three diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
DNS(3)+FB3	DNS(3) model with three diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
DNS(4)+FB3	DNS(4) model with three diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
DNS(5)+FB3	DNS(5) model with three diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
DNS(6)+FB3	DNS(6) model with three diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
NSS(1)	Dynamic Nelson-Siegel-Svensson (NSS) model with underlying AR(1) factor specifications fitted with twelve-dimensional yields, with a static rate of decay parameter $\lambda = [0.0609, 0.2985]$
NSS(2)	Dynamic Nelson-Siegel-Svensson (NSS) model with underlying AR(1) factor specifications fitted with twelve-dimensional yields, with a dynamic rate of decay parameter λ_t (most recent λ_t are selected in generating predictions)
NSS(3)	Dynamic Nelson-Siegel-Svensson (NSS) model with underlying AR(1) factor specifications fitted with twelve-dimensional yields, with a dynamic rate of decay parameter λ_t (median λ_t are selected in generating predictions)
NSS(4)	Dynamic Nelson-Siegel-Svensson (NSS) model with underlying VAR(1) factor specifications fitted with twelve-dimensional yields, with a static rate of decay parameter $\lambda = 0.0609$
NSS(5)	Dynamic Nelson-Siegel-Svensson (NSS) model with underlying VAR(1) factor specifications fitted with twelve-dimensional yields, with a dynamic rate of decay parameter λ_t (most recent λ_t are selected in generating predictions)
NSS(6)	Dynamic Nelson-Siegel-Svensson (NSS) model with underlying VAR(1) factor specifications fitted with twelve-dimensional yields, with a dynamic rate of decay parameter λ_t (median λ_t are selected in generating predictions)
NSS(1)+FB1	NSS(1) model with one principle component added, principle component analysis based on real-time macroeconomic dataset
NSS(2)+FB1	NSS(2) model with one principle component added, principle component analysis based on real-time macroeconomic dataset
NSS(3)+FB1	NSS(3) model with one principle component added, principle component analysis based on real-time macroeconomic dataset
NSS(4)+FB1	NSS(4) model with one principle component added, principle component analysis based on real-time macroeconomic dataset
NSS(5)+FB1	NSS(5) model with one principle component added, principle component analysis based on real-time macroeconomic dataset
NSS(6)+FB1	NSS(6) model with one principle component added, principle component analysis based on real-time macroeconomic dataset
NSS(1)+FB2	NSS(1) model with two diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
NSS(2)+FB2	NSS(2) model with two diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
NSS(3)+FB2	NSS(3) model with two diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
NSS(4)+FB2	NSS(4) model with two diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
NSS(5)+FB2	NSS(5) model with two diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
NSS(6)+FB2	NSS(6) model with two diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
NSS(1)+FB3	NSS(1) model with three diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
NSS(2)+FB3	NSS(2) model with three diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
NSS(3)+FB3	NSS(3) model with three diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
NSS(4)+FB3	NSS(4) model with three diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
NSS(5)+FB3	NSS(5) model with three diffusion indexes added, principle component analysis based on real-time macroeconomic dataset
NSS(6)+FB3	NSS(6) model with three diffusion indexes added, principle component analysis based on real-time macroeconomic dataset

^{*} Notes: Entries in this table describe the models utilized in the forecasting experiments for which results are summarized in subsequent tables. For models with real-time diffusion indexes, 5 additional variants of each model are analyzed. In these models, a subset of variables from the real-time dataset is used when constructing diffusion indexes, with the subset selected using either the elastic net (EN) or the LASSO. Ten-fold cross validation is used, in real-time, to estimate the first tuning parameter in the operators. For the elastic net, there are two tuning parameters (rather than 1, as is the case with the lasso), and the second parameter is fixed at various values, including 0.2, 0.4, 0.6, and 0.8. Model mnemonics are correspondingly denoted as LASSO, EN02, EN04, EN06, and EN08. For complete details, refer to Section 4.

Table 2A: Top 3 MSFE-Best Models*

Horizon		3 Months	1 Year	3 Years	5 Years	10 Years
Maturity						
2001:Jan - 2005:Dec 1st Subsample	1 Step	NSS(4)+FB1 ₂ ***	DNS(4)+FB1(EN04) ₃	DNS(1)+FB2(LASSO) ₂	NSS(5) ₂	NSS(5)+FB1(EN02) ₂
		NSS(4)+FB3 ₂ ***	AR(1)+FB2	RW	RW	NSS(2)+FB3(EN02) ₂
		NSS(4)+FB2 ₂ ***	AR(1)+FB3	DNS(1)+FB1 ₂	NSS(5)+FB1(EN06) ₂	NSS(2)+FB2(EN02) ₂
	3 Step	VAR(1)+FB1***	AR(1)+FB1*	NSS(3)+FB1 ₁ **	NSS(3)+FB1 ₁	NSS(2)+FB1(EN04) ₂ *
		VAR(1)+FB2***	NSS(4) ₁	NSS(2)+FB1 ₁	NSS(3)+FB1(LASSO) ₁	NSS(3)+FB1(EN04) ₁ **
		VAR(1)***	DNS(2)+FB1 ₁	NSS(3)+FB1(EN02) ₁	NSS(3)+FB1(EN08) ₁	DNS(2)+FB1(EN04) ₁
	12 Step	DNS(3)+FB1 ₂ ***	DNS(3)+FB1 ₂ ***	DNS(3)+FB1 ₃ ***	DNS(3)+FB1 ₃ **	DNS(3) ₃ **
		DNS(2)+FB1 ₂ ***	NSS(2)+FB1 ₁ *	DNS(6)+FB1 ₂ ***	NSS(6) ₂ ***	NSS(6) ₂ ***
		DNS(2)+FB1 ₁ ***	DNS(2)+FB1 ₂ **	NSS(2)+FB1 ₂ **	NSS(6)+FB1 ₂ ***	NSS(6)+FB1 ₂ ***
	1 Step	NSS(3) ₂	DNS(1) ₂ **	VAR(1)+FB3(EN02)	VAR(1)	NSS(1) ₃
		NSS(2) ₂	DNS(5) ₂ *	VAR(1)	VAR(1)+FB2	DNS(1) ₃
		DNS(5)+FB1(EN08) ₂	NSS(1) ₂ *	VAR(SIC)	VAR(1)+FB2(LASSO)	DNS(2) ₃
2006:Jan - 2010:Dec 2nd Subsample	3 Step	NSS(6) ₂ *	NSS(1)+FB1(LASSO) ₂	RW	DNS(6) ₁	DNS(3) ₁ **
		NSS(3) ₂	DNS(5)+FB1 ₂ **	NSS(4) ₁	DNS(6)+FB1 ₁	DNS(6) ₁
		NSS(6)+FB1 ₂	DNS(4)+FB1 ₂ *	DNS(6) ₁	NSS(3) ₂	NSS(4)+FB1 ₃
	12 Step	VAR(1)+FB1***	NSS(2)+FB1 ₂ **	NSS(1)+FB1 ₁ *	NSS(1)+FB1 ₁ *	DNS(6) ₁ **
		DNS(1)+FB1 ₁ ***	DNS(1)+FB1 ₁ **	DNS(1)+FB1 ₁ *	DNS(6) ₁	DNS(6)+FB1 ₁ *
		DNS(3)+FB1 ₃ **	NSS(3)+FB1 ₂ *	NSS(3)+FB1 ₂	AR(1)+FB1	DNS(6)+FB2 ₁
	1 Step	RW	RW	RW	RW	RW
		AR(SIC)	AR(1)	AR(1)	AR(1)	NSS(1)+FB2(EN02) ₃
		AR(1)	AR(1)+FB1	AR(1)+FB2(EN04)	AR(SIC)	NSS(1)+FB2(EN06) ₃
	3 Step	RW	RW	RW	RW	RW
		DNS(5) ₂	DNS(5) ₁	AR(1)	AR(1)	DNS(5) ₃
		NSS(4) ₂	DNS(5)+FB2 ₁	DNS(4) ₂	AR(SIC)	DNS(1)+FB2(EN06) ₂
	12 Step	DNS(5) ₂ ***	DNS(5) ₁ ***	RW	RW	NSS(6)+FB1(EN04) ₂ ***
		DNS(6) ₂ ***	DNS(5)+FB1 ₁ **	DNS(4) ₁	NSS(6) ₁	NSS(6)+FB1(EN02) ₂ ***
		DNS(5)+FB1 ₂ ***	DNS(5)+FB2 ₁ *	DNS(4)+FB1 ₁	NSS(4)+FB3 ₁	NSS(6)+FB1(EN06) ₂ **
2001:Jan - 2017:Oct 3rd Subsample	1 Step	VAR(1)+FB3(EN02)	AR(SIC)	RW	RW	RW
		VAR(1)+FB1(EN02)	DNS(5)+FB1 ₂	AR(1)	AR(1)	DNS(1) ₁
		VAR(1)+FB2	AR(1)+FB1	AR(1)+FB2(LASSO)	AR(1)+FB1	NSS(1) ₂
	3 Step	DNS(5)+FB1 ₂ ***	DNS(5)+FB1 ₂ **	RW	RW	DNS(5) ₃
		DNS(6)+FB1 ₂ ***	DNS(5)+FB2 ₂ **	NSS(4) ₁	NSS(4) ₁	RW
		DNS(5)+FB2 ₂ ***	DNS(5)+FB3 ₂ *	DNS(4)+FB1 ₁	NSS(5) ₂	DNS(5)+FB1 ₃
	12 Step	DNS(6) ₂ ***	DNS(6) ₂ ***	NSS(4) ₁	DNS(6) ₁	NSS(6) ₂ **
		DNS(6)+FB1 ₂ ***	DNS(6)+FB1 ₂ ***	DNS(6) ₁	DNS(6)+FB1 ₁	DNS(6) ₁ ***
		DNS(6)+FB2 ₂ ***	DNS(6)+FB2 ₂ ***	RW	RW	NSS(6)+FB1 ₂ **
2001:Jan - 2017:Oct Whole Sample	1 Step	VAR(1)+FB3(EN02)	AR(SIC)	RW	RW	RW
		VAR(1)+FB1(EN02)	DNS(5)+FB1 ₂	AR(1)	AR(1)	DNS(1) ₁
		VAR(1)+FB2	AR(1)+FB1	AR(1)+FB2(LASSO)	AR(1)+FB1	NSS(1) ₂
	3 Step	DNS(5)+FB1 ₂ ***	DNS(5)+FB1 ₂ **	RW	RW	DNS(5) ₃
		DNS(6)+FB1 ₂ ***	DNS(5)+FB2 ₂ **	NSS(4) ₁	NSS(4) ₁	RW
		DNS(5)+FB2 ₂ ***	DNS(5)+FB3 ₂ *	DNS(4)+FB1 ₁	NSS(5) ₂	DNS(5)+FB1 ₃
	12 Step	DNS(6) ₂ ***	DNS(6) ₂ ***	NSS(4) ₁	DNS(6) ₁	NSS(6) ₂ **
		DNS(6)+FB1 ₂ ***	DNS(6)+FB1 ₂ ***	DNS(6) ₁	DNS(6)+FB1 ₁	DNS(6) ₁ ***
		DNS(6)+FB2 ₂ ***	DNS(6)+FB2 ₂ ***	RW	RW	NSS(6)+FB1 ₂ **

* Notes: Entries in this table are the top three performing forecast models (based on MSFE), in descending order (from 1st to 3rd place), for various subsamples, horizons, and yield maturities. For a description of the model mnemonics, refer to Table 1. Entries with subscripts 1, 2, and 3 differentiate between DNS and NSS type models that are alternatively estimated using 10-, 12-, or 30-dimensional historical yield datasets, respectively. Entries with LASSO, EN02, EN04, EN06, and EN08 denote cases where targeted variables from the real-time dataset are used in construction of real-time diffusion indexes. Entries superscripted with *, **, or *** denote rejection of the DM predictive accuracy null hypothesis, in favor of the alternative that the listed model has significantly lower MSFE than the random walk benchmark, at a 10%, 5%, and 1% significance level, respectively. See Section 4 for a discussion of the tuning parameters used in these shrinkage operators, as well as for a discussion of the prediction models reported on.

Table 2B: Point MSFEs of Top 3 MSFE-Best Models*

Horizon						
Maturity		3 Months	1 Year	3 Years	5 Years	10 Years
2001:Jan - 2005:Dec 1st Subsample	1 Step	0.029***	0.057	0.108	0.110	0.082
		0.029***	0.057	0.109	0.111	0.085
		0.029***	0.059	0.112	0.112	0.085
	3 Step	0.158***	0.313*	0.258**	0.256	0.169*
		0.164***	0.324	0.337	0.276	0.171**
		0.165***	0.329	0.337	0.282	0.174
	12 Step	2.067***	2.338***	1.375***	0.869**	0.374**
		2.126***	2.542*	1.515***	0.918***	0.381***
		2.405***	2.559**	1.520**	0.931***	0.391***
2006:Jan - 2010:Dec 2nd Subsample	1 Step	0.082	0.052**	0.079	0.078	0.075
		0.086	0.053*	0.079	0.078	0.082
		0.086	0.055*	0.079	0.079	0.087
	3 Step	0.316*	0.282	0.353	0.279	0.161**
		0.316	0.284**	0.356	0.287	0.177
		0.332	0.286*	0.356	0.288	0.179
	12 Step	1.899***	1.580**	1.087*	0.752*	0.316**
		1.994***	1.613**	1.130*	0.830	0.332*
		2.149**	1.699*	1.220	0.832	0.334
2011:Jan - 2017:Oct 3rd Subsample	1 Step	0.003	0.005	0.019	0.035	0.047
		0.005	0.006	0.020	0.036	0.047
		0.005	0.007	0.021	0.037	0.049
	3 Step	0.010	0.011	0.045	0.101	0.161
		0.018	0.014	0.051	0.109	0.163
		0.018	0.014	0.052	0.111	0.165
	12 Step	0.043***	0.056***	0.134	0.316	0.337***
		0.044***	0.059**	0.140	0.317	0.338***
		0.044***	0.060*	0.140	0.324	0.340***
2001:Jan - 2017:Oct Whole Sample	1 Step	0.039	0.042	0.065	0.073	0.075
		0.041	0.042	0.069	0.076	0.075
		0.041	0.042	0.070	0.079	0.076
	3 Step	0.182***	0.190**	0.225	0.217	0.187
		0.183***	0.193**	0.235	0.234	0.189
		0.183***	0.195*	0.238	0.241	0.190
	12 Step	1.723***	1.462***	0.993	0.713	0.428**
		1.731***	1.469***	1.009	0.724	0.439***
		1.736***	1.473***	1.012	0.727	0.439**

* Notes: See notes to Table 2A. Entire are point MSFEs.

Table 2C: Relative MSFEs of Top 3 MSFE-Best Models*

Horizon						
Maturity		3 Months	1 Year	3 Years	5 Years	10 Years
2001:Jan - 2005:Dec 1st Subsample	1 Step	0.502***	0.836	0.986	0.987	0.921
		0.504***	0.839	1.000	1.000	0.952
		0.505***	0.857	1.024	1.007	0.952
	3 Step	0.442***	0.884*	0.755**	0.861	0.843*
		0.458***	0.913	0.985	0.931	0.850**
		0.461***	0.927	0.985	0.948	0.865
	12 Step	0.619***	0.726***	0.747***	0.758**	0.718**
		0.637***	0.790*	0.824***	0.800***	0.732***
		0.720***	0.795**	0.826**	0.812***	0.751***
2006:Jan - 2010:Dec 2nd Subsample	1 Step	0.883	0.768**	0.925	0.890	0.752
		0.922	0.786*	0.929	0.898	0.822
		0.926	0.808*	0.929	0.904	0.876
	3 Step	0.874*	0.815	1.000	0.942	0.743**
		0.874	0.821**	1.008	0.969	0.820
		0.918	0.828*	1.009	0.972	0.829
	12 Step	0.625***	0.687**	0.784*	0.866*	0.832**
		0.656***	0.702**	0.815*	0.955	0.873*
		0.707**	0.739*	0.880	0.958	0.878
2011:Jan - 2017:Oct 3rd Subsample	1 Step	1.000	1.000	1.000	1.000	1.000
		1.486	1.413	1.061	1.045	1.012
		1.520	1.481	1.097	1.076	1.055
	3 Step	1.000	1.000	1.000	1.000	1.000
		1.748	1.216	1.124	1.089	1.011
		1.772	1.240	1.141	1.108	1.021
	12 Step	0.676***	0.835***	1.000	1.000	0.565***
		0.688***	0.873**	1.044	1.003	0.567***
		0.691***	0.889*	1.046	1.027	0.570***
2001:Jan - 2017:Oct Whole Sample	1 Step	0.855	0.984	1.000	1.000	1.000
		0.887	0.984	1.049	1.046	1.005
		0.887	0.985	1.066	1.083	1.009
	3 Step	0.835***	0.895**	1.000	1.000	0.989
		0.841***	0.907**	1.045	1.079	1.000
		0.842***	0.919*	1.056	1.111	1.004
	12 Step	0.897***	0.877***	0.980	0.981	0.841**
		0.901***	0.882***	0.996	0.996	0.862***
		0.904***	0.884***	1.000	1.000	0.862**

* Notes: See notes to Table 2A. Entries are relative MSFEs, in which the numeraire is the random walk benchmark model.

Table 3: Forecast Combination Models*

Model	Description
All	Average of all two hundred and sixty nine forecast models
Econometrics	Average of all forty one RW, AR and VAR type models
AR	Average of all twenty AR type models
VAR	Average of all twenty VAR type models
DNS	Average of all one hundred and fourteen DNS type models
NSS	Average of all one hundred and fourteen NSS type models
NS(ar)	Average of one hundred and fourteen DNS and NSS type models with underlying AR(1) factor specifications
NS(var)	Average of one hundred and fourteen DNS and NSS type models with underlying VAR(1) factor specifications
NS(ols)	Average of seventy six DNS and NSS type models with fixed decay parameter(s), estimated with OLS
NS(nls)	Average of one hundred and fifty two DNS and NSS type models with dynamic decay parameter(s), estimated with NLS
FB	Average of two hundred and fifty two models that contain macro diffusion index(es), principle component analysis based on all macroeconomic variables
FB(pure)	Average of thirty six models that contain diffusion indexes only, namely all “AR(1)+FB” and “VAR(1)+FB” models
FB(ols)	Average of one hundred and eight models that contain macro diffusion index(es), estimated with OLS
FB(en)	Average of one hundred and sixty eight models that contain macro diffusion index(es), principle component analysis based on targeted macroeconomic variables pre-selected by the elastic net
FB(lasso)	Average of forty two models that contain macro diffusion index(es), principle component analysis based on targeted macroeconomic variables pre-selected by lasso
FS	Average of all non-FB type models

* Notes: Entries in this table describe the forecast combination models utilized in the forecasting experiments for which results are summarized in subsequent tables. Note that the 12-dimensional historical yield dataset is used in the construction of the betas utilized in all NS type models. For further details, refer to Sections 4 and 6.

Table 4A: 1-Step-Ahead Relative MSFEs of Forecast Combination Models^{*}

		Model	rMSFE				
		Maturity	3 months	1 year	3 years	5 years	10 years
2001:Jan - 2005:Dec 1st Subsample	All	0.391***	0.516***	0.450***	0.471***	0.442***	
	Econometrics	0.599***	0.872	1.067	1.121	1.134	
	AR	0.709***	0.876*	1.068	1.112	1.076	
	VAR	0.543***	0.902	1.107	1.176	1.240	
	DNS	0.268***	0.414***	0.392***	0.450***	0.373***	
	NSS	0.640**	0.593***	0.403***	0.369***	0.365***	
	NS(ar)	0.608**	0.654**	0.411***	0.397***	0.353***	
	NS(var)	0.251***	0.376***	0.374***	0.401***	0.374***	
	NS(ols)	0.274***	0.414***	0.373***	0.420***	0.388***	
	NS(nls)	0.455**	0.539***	0.395***	0.390***	0.373***	
	FB	0.489***	0.641***	0.553***	0.581***	0.542***	
	FB(pure)	0.591***	0.871	1.072	1.132	1.147	
	FB(ols)	0.413***	0.617***	0.650***	0.716***	0.674***	
	FB(en)	0.970	1.232	1.099	1.134	1.033	
FB(lasso)	0.910	1.435	1.062	1.197	1.141		
FS	0.284***	0.350***	0.219***	0.235***	0.177***		
2006:Jan - 2010:Dec 2nd Subsample	All	0.401**	0.526**	0.483***	0.450***	0.417***	
	Econometrics	0.945	0.912	0.960	0.965	1.021	
	AR	1.038	1.010	1.042	1.077	1.072	
	VAR	0.964	0.893	0.943	0.917	0.998	
	DNS	0.356**	0.440**	0.418***	0.461***	0.418***	
	NSS	0.385**	0.597*	0.489***	0.416***	0.382***	
	NS(ar)	0.378**	0.513**	0.445***	0.417***	0.343***	
	NS(var)	0.313**	0.499*	0.434***	0.381***	0.355***	
	NS(ols)	0.364**	0.514*	0.406***	0.426***	0.388***	
	NS(nls)	0.328**	0.502**	0.440***	0.392***	0.342***	
	FB	0.485**	0.629*	0.587***	0.550***	0.510***	
	FB(pure)	0.979	0.968	0.981	0.978	1.033	
	FB(ols)	0.602**	0.706	0.652***	0.666***	0.639***	
	FB(en)	0.923	1.174	1.126	1.076	0.991	
FB(lasso)	0.903	1.145	1.121	1.025	0.953		
FS	0.172***	0.164***	0.255***	0.252***	0.154***		
2011:Jan - 2017:Oct 3rd Subsample	All	2.000	1.328	0.691**	0.646***	0.509***	
	Econometrics	1.757	1.969	1.321	1.266	1.324	
	AR	1.516	1.511	1.101	1.120	1.229	
	VAR	3.151	3.634	1.905	1.602	1.518	
	DNS	2.847	0.992	0.674**	0.692**	0.488***	
	NSS	1.941	1.883	0.766*	0.608***	0.446***	
	NS(ar)	3.051	1.999	0.788	0.607***	0.409***	
	NS(var)	1.749	1.004	0.626***	0.637**	0.489***	
	NS(ols)	2.031	0.975	0.490***	0.638***	0.399***	
	NS(nls)	2.344	1.836	0.897	0.769*	0.563***	
	FB	2.591	1.678	0.868	0.815	0.632***	
	FB(pure)	1.804	2.049	1.344	1.288	1.352	
	FB(ols)	2.148	1.316	0.828*	0.939	0.718***	
	FB(en)	5.535	3.363	1.713	1.647	1.251	
FB(lasso)	5.799	3.717	1.949	1.713	1.205		
FS	2.377	1.104	0.492***	0.509***	0.300***		
2001:Jan - 2017:Oct 4th Subsample	All	0.445***	0.556***	0.492***	0.497***	0.449***	
	Econometrics	0.841*	0.939	1.056	1.094	1.137	
	AR	0.931	0.968	1.062	1.101	1.113	
	VAR	0.874	1.016	1.139	1.166	1.214	
	DNS	0.397***	0.452***	0.436***	0.500***	0.420***	
	NSS	0.525**	0.651**	0.479***	0.432***	0.392***	
	NS(ar)	0.542**	0.646**	0.469***	0.445***	0.363***	
	NS(var)	0.333***	0.462***	0.427***	0.439***	0.396***	
	NS(ols)	0.380***	0.486***	0.399***	0.464***	0.391***	
	NS(nls)	0.435***	0.578***	0.472***	0.464***	0.409***	
	FB	0.549***	0.680**	0.604***	0.615***	0.552***	
	FB(pure)	0.860	0.969	1.069	1.107	1.154	
	FB(ols)	0.578***	0.690**	0.672***	0.741***	0.671***	
	FB(en)	1.077	1.297	1.182	1.213	1.071	
FB(lasso)	1.051	1.396	1.190	1.236	1.083		
FS	0.279***	0.294***	0.266***	0.294***	0.199***		

^{*} Notes: See notes to Tables 2 and 3. Entries are MSFEs, relative to the benchmark random walk (RW), based on 1-step-ahead forecasts of monthly U.S. Treasury bond yields of various maturities, for forecast combination models listed in Table 3. Entries in bold denote models with lowest MSFE, for a given maturity. Entries superscripted with *, **, or *** denote rejection of the DM predictive accuracy null hypothesis, in favor of the alternative that the listed model has significantly lower MSFE than the random walk benchmark, at a 10%, 5%, and 1% significance level, respectively. A 12-dimensional historical yield dataset is used in the construction of betas in all NS type models.

Table 4B: 3-Step-Ahead Relative MSFEs of Forecast Combination Models*

Model		rMSFE				
	Maturity	3 months	1 year	3 years	5 years	10 years
2001:Jan - 2005:Dec 1st Subsample	All	0.283***	0.513***	0.553***	0.605***	0.526***
	Econometrics	0.535***	0.984	1.385	1.540	1.528
	AR	0.648***	0.997	1.374	1.501	1.384
	VAR	0.479***	1.045	1.486	1.664	1.773
	DNS	0.204***	0.401***	0.481***	0.556***	0.449***
	NSS	0.324***	0.519***	0.443***	0.455***	0.383***
	NS(ar)	0.364***	0.502***	0.458***	0.434***	0.336***
	NS(var)	0.199***	0.436***	0.473***	0.585***	0.526***
	NS(ols)	0.239***	0.444***	0.506***	0.583***	0.462***
	NS(uls)	0.263***	0.460***	0.431***	0.459***	0.390***
	FB	0.355***	0.629***	0.672***	0.733**	0.635***
	FB(pure)	0.528***	1.007	1.421	1.595	1.582
	FB(ols)	0.374***	0.691**	0.861*	0.983	0.855*
	FB(en)	0.714***	1.197	1.229	1.407	1.209
	FB(lasso)	0.927	1.378	1.533	1.295	1.421
	FS	0.238***	0.259***	0.180***	0.180***	0.156***
2006:Jan - 2010:Dec 2nd Subsample	All	0.517**	0.454**	0.506***	0.553***	0.484***
	Econometrics	1.084	0.997	1.132	1.187	1.161
	AR	1.145	1.024	1.169	1.313	1.289
	VAR	1.191	1.063	1.166	1.137	1.098
	DNS	0.451**	0.381***	0.460***	0.528***	0.439***
	NSS	0.452**	0.419***	0.402***	0.431***	0.381***
	NS(ar)	0.468**	0.387***	0.425***	0.471***	0.381***
	NS(var)	0.455**	0.433**	0.469***	0.509***	0.431***
	NS(ols)	0.456**	0.399***	0.452***	0.513***	0.414***
	NS(uls)	0.442**	0.390***	0.417***	0.459***	0.397***
	FB	0.621*	0.547**	0.619***	0.675***	0.587***
	FB(pure)	1.153	1.041	1.161	1.223	1.198
	FB(ols)	0.730	0.653*	0.748***	0.818**	0.710**
	FB(en)	1.157	1.033	1.167	1.245	1.111
	FB(lasso)	1.155	1.065	1.222	1.335	1.095
	FS	0.192***	0.163***	0.208***	0.212***	0.151***
2011:Jan - 2017:Oct 3rd Subsample	All	2.670	2.767	1.163	0.757*	0.482***
	Econometrics	2.571	3.865	2.019	1.534	1.342
	AR	2.310	2.374	1.518	1.341	1.320
	VAR	5.300	8.919	3.549	2.111	1.465
	DNS	3.214	2.212	1.134	0.741*	0.436***
	NSS	2.734	3.391	1.061	0.616***	0.342***
	NS(ar)	3.861	2.989	1.030	0.621***	0.357***
	NS(var)	2.571	3.099	1.307	0.786	0.432***
	NS(ols)	3.078	2.009	1.042	0.796	0.399***
	NS(uls)	2.849	3.290	1.179	0.672**	0.394***
	FB	3.334	3.383	1.413	0.925	0.588***
	FB(pure)	2.670	4.136	2.139	1.595	1.369
	FB(ols)	3.188	2.768	1.499	1.164	0.746**
	FB(en)	6.694	6.281	2.612	1.779	1.137
	FB(lasso)	6.942	7.142	2.680	1.638	1.129
	FS	1.661	1.173	0.372***	0.266***	0.169***
2001:Jan - 2017:Oct 4th Subsample	All	0.443***	0.534***	0.581***	0.612***	0.497***
	Econometrics	0.844	1.053	1.319	1.395	1.339
	AR	0.924	1.040	1.290	1.395	1.330
	VAR	0.921	1.225	1.506	1.534	1.438
	DNS	0.382***	0.431***	0.525***	0.580***	0.441***
	NSS	0.433***	0.534***	0.474***	0.476***	0.368***
	NS(ar)	0.481***	0.501***	0.490***	0.484***	0.358***
	NS(var)	0.370***	0.493***	0.540***	0.592***	0.462***
	NS(ols)	0.399***	0.457***	0.525***	0.594***	0.424***
	NS(uls)	0.400***	0.488***	0.486***	0.499***	0.394***
	FB	0.542***	0.649***	0.708***	0.746***	0.603***
	FB(pure)	0.877	1.091	1.359	1.444	1.378
	FB(ols)	0.602***	0.717**	0.860**	0.950	0.768***
	FB(en)	1.045	1.229	1.314	1.412	1.151
	FB(lasso)	1.153	1.352	1.482	1.376	1.210
	FS	0.242***	0.233***	0.209***	0.209***	0.159***

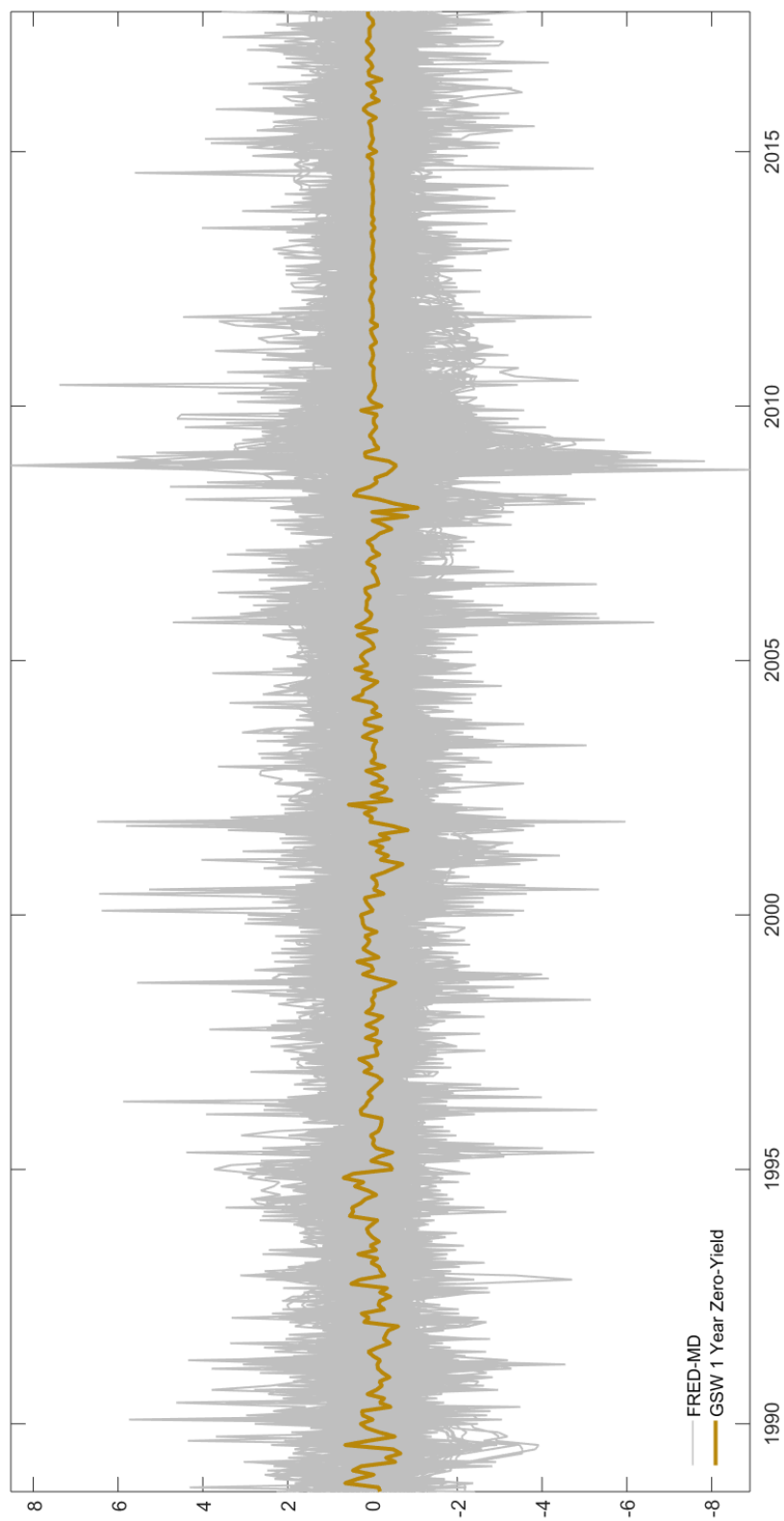
* Notes: See notes to Table 4B. Entries are MSFEs, relative to the benchmark random walk model, based on 3-step-ahead forecasts of monthly U.S. Treasury bond yields of various maturities.

Table 4C: 12-Step-Ahead Relative MSFEs of Forecast Combination Models*

Model		rMSFE				
	Maturity	3 months	1 year	3 years	5 years	10 years
2001:Jan - 2005:Dec 1st Subsample	All	0.648**	0.852	1.144	1.288	1.462
	Econometrics	1.625	2.416	3.566	4.073	4.557
	AR	1.370	1.904	2.931	3.515	4.004
	VAR	2.069	3.241	4.637	5.082	5.590
	DNS	0.528***	0.664**	0.951	1.090	1.177
	NSS	0.526***	0.666**	0.769**	0.843	1.039
	NS(ar)	0.450***	0.483***	0.474***	0.457***	0.667***
	NS(var)	0.663**	0.948	1.457	1.755	1.803
	NS(ols)	0.646**	0.855	1.191	1.349	1.378
	NS(nls)	0.475***	0.583***	0.717**	0.798**	0.985
	FB	0.764*	0.996	1.325	1.491	1.689
	FB(pure)	1.660	2.498	3.711	4.274	4.833
	FB(ols)	1.040	1.449	2.076	2.366	2.505
	FB(en)	1.335	1.699	2.250	2.509	2.756
	FB(lasso)	1.369	1.817	2.393	2.605	3.510
	FS	0.132***	0.111***	0.072***	0.067***	0.089***
2006:Jan - 2010:Dec 2nd Subsample	All	0.515***	0.632***	0.642***	0.798**	1.054
	Econometrics	0.943	1.041	1.377	1.771	2.401
	AR	1.190	1.247	1.355	1.679	2.466
	VAR	0.879	1.098	1.605	2.035	2.696
	DNS	0.420***	0.562***	0.550***	0.755**	0.926
	NSS	0.499***	0.609***	0.579***	0.643***	0.900
	NS(ar)	0.480***	0.606***	0.544***	0.601***	0.793*
	NS(var)	0.449***	0.567***	0.585***	0.797**	1.086
	NS(ols)	0.431***	0.623**	0.656***	0.851	1.078
	NS(nls)	0.471***	0.561***	0.509***	0.606***	0.808
	FB	0.618***	0.752**	0.773**	0.964	1.251
	FB(pure)	1.066	1.162	1.477	1.877	2.632
	FB(ols)	0.681**	0.867	0.998	1.292	1.666
	FB(en)	1.127	1.306	1.444	1.735	2.249
	FB(lasso)	1.183	1.534	1.389	1.835	1.971
	FS	0.200***	0.197***	0.215***	0.229***	0.180***
2011:Jan - 2017:Oct 3rd Subsample	All	6.704	7.568	3.593	1.533	0.771**
	Econometrics	10.268	11.881	6.212	2.764	1.630
	AR	9.204	9.068	5.099	2.640	1.615
	VAR	17.035	20.252	9.020	3.313	1.771
	DNS	7.433	6.952	3.482	1.547	0.759**
	NSS	5.671	7.471	3.161	1.249	0.575***
	NS(ar)	3.680	4.096	1.779	1.053	0.631***
	NS(var)	11.235	12.720	5.821	1.955	0.736**
	NS(ols)	7.089	6.899	3.467	1.576	0.686***
	NS(nls)	6.068	7.226	3.222	1.310	0.663***
	FB	7.886	8.835	4.222	1.825	0.923
	FB(pure)	11.339	13.390	7.062	3.058	1.727
	FB(ols)	9.105	9.468	4.943	2.251	1.099
	FB(en)	13.361	14.875	6.934	3.162	1.629
	FB(lasso)	16.433	15.654	7.975	3.431	1.854
	FS	1.081	0.973	0.565***	0.303***	0.173***
2001:Jan - 2017:Oct 4th Subsample	All	0.667***	0.872	1.071	1.157	1.043
	Econometrics	1.421	2.009	2.817	3.025	2.689
	AR	1.391	1.753	2.406	2.709	2.528
	VAR	1.712	2.643	3.638	3.688	3.135
	DNS	0.571***	0.726**	0.924	1.052	0.923
	NSS	0.583***	0.755**	0.820**	0.844**	0.788***
	NS(ar)	0.507***	0.593***	0.573***	0.613***	0.678***
	NS(var)	0.705**	0.986	1.336	1.450	1.137
	NS(ols)	0.632***	0.859	1.095	1.212	0.983
	NS(nls)	0.549***	0.684***	0.767***	0.820***	0.793***
	FB	0.792**	1.025	1.256	1.363	1.228
	FB(pure)	1.511	2.130	2.982	3.209	2.870
	FB(ols)	0.980	1.343	1.791	1.965	1.651
	FB(en)	1.400	1.755	2.173	2.349	2.108
	FB(lasso)	1.485	1.929	2.284	2.478	2.382
	FS	0.177***	0.161***	0.157***	0.166***	0.149***

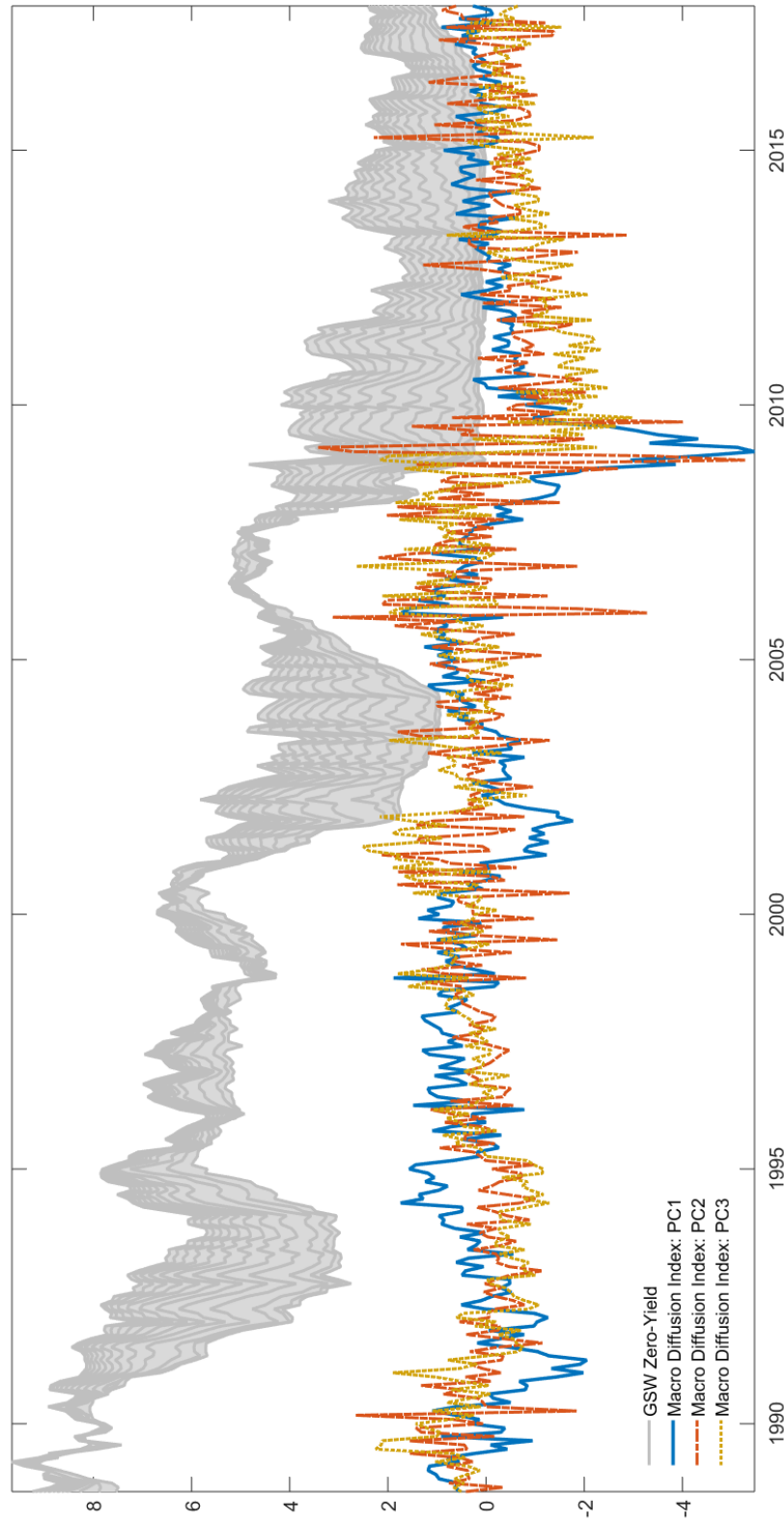
* Notes: See notes to Table 4A. Entries are MSFEs, relative to the benchmark random walk model, based on 12-step-ahead forecasts of monthly U.S. Treasury bond yields of various maturities.

Figure 1: FRED MD Dataset for Sample Period 1988:8 - 2017:10*



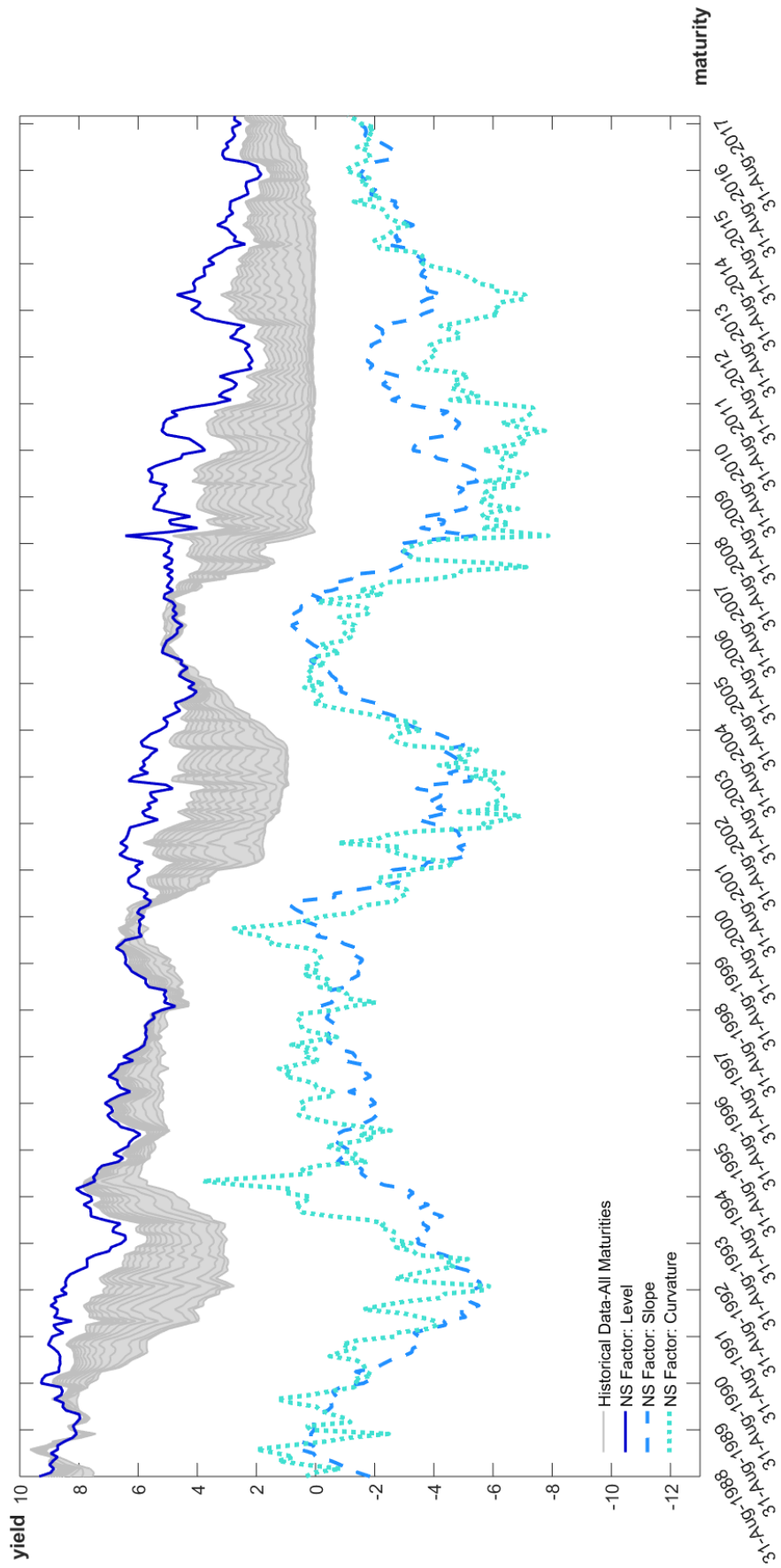
(*) Notes: This figure contains plots of the entire set of FRED-MD variables, as well as a plot of the GSW 1-year yield. All non-yield variables are transformed to ensure stationarity, and all variables are standardized to zero mean and unit variance.

Figure 2: Yields and Macroeconomic Diffusion Indexes for the Sample Period 1988:8 - 2017:10*



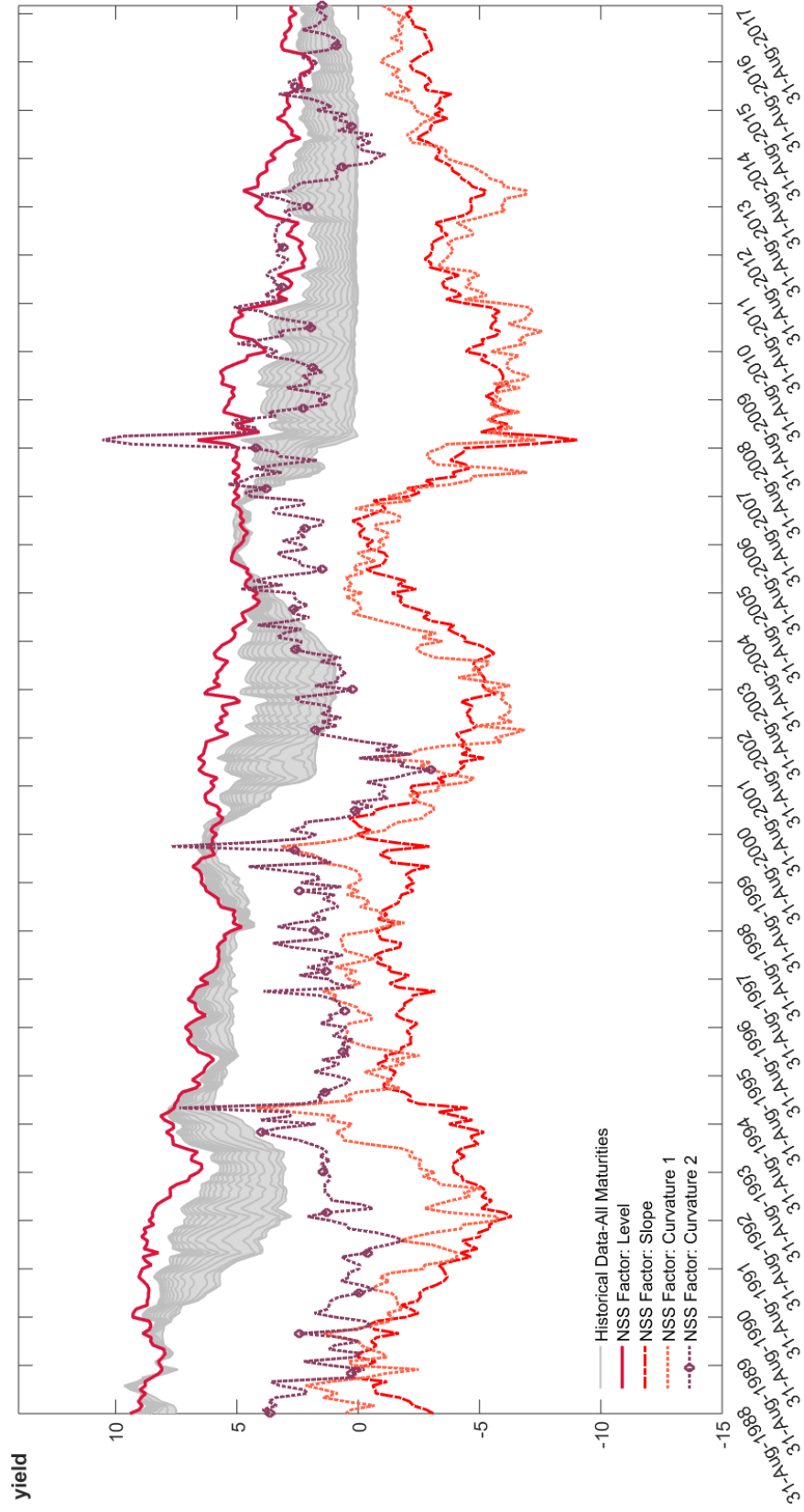
(*) Notes: This figure displays the 12-dimensional yield dataset used in the forecasting experiments summarized in Table 1-4. The yield dataset combines data from FRED H.15 GSW datasets (see Section 5 for complete details). Plots of un-targeted real-time diffusion indexes estimated using data from 1988:8-2017:10 are also displayed.

Figure 3: Yields and Dynamic Nelson Siegel Factors for the Sample Period 1988:8 - 2017:10*



(*) Notes: This figure displays the 12-dimensional yield dataset used in the forecasting experiments summarized in Table 1-4. The yield dataset combines data from FRED H.15 GSW datasets (see Section 5 for complete details). Plots three DNS factors (level, slope, and curvature), constructed with a fixed rate of decay estimated using ordinary least squares, are also displayed. See Sections 2 and 4 for further details.

Figure 4: Yields and Dynamic Nelson Siegel Svensson Factors for the Sample Period 1988:8 - 2017:10*



(*) Notes: This figure displays the 12-dimensional yield dataset used in the forecasting experiments summarized in Table 1-4. The yield dataset combines data from FRED H.15 GSW datasets (see Section 5 for complete details). Plots three NSS factors (level, slope, and curvature), constructed with a fixed rate of decay estimated using ordinary least squares, are also displayed. See Sections 2 and 4 for further details.