Erasmus University Rotterdam

Erasmus School of Economics

Comparing small data models versus big data models for forecasting interest rates[1]

Bachelor Thesis Econometrics and Operations Research

Martijn Riemers (445591)
martijnriemers@student.eur.nl

Supervisor:
T. van der Zwan

Second assessor:
Prof. dr. D.J.C. van Dijk

Date final version:
5th July 2020

Abstract
In this paper, we model the yield curve of government bonds using a wide range of different small and big data models. We then compare their predictive accuracy against a benchmark AR(1) model to find the best performing forecasting model. We include factor augmented models with standard principal component analysis and sparse principal component analysis and compare the respective models their predictive accuracy. We find that models predictive accuracy largely depends on the economic situation. We also find that models with factors estimated with sparse principal component analysis perform better than factors estimated with standard principal component analysis.

[1] The views stated in this thesis are those of the author and not necessarily those of the supervisor, second assessor, Erasmus School of Economics or Erasmus University Rotterdam.

Contents

1 Introduction **2**

2 Methodology **4**
2.1 Dimension reduction . . . 4
2.1.1 Principal Component Analysis (PCA) . . . 4
2.1.2 Sparse Principal Component Analysis (SPCA) . . . 5
2.2 Models . . . 6
2.2.1 Autoregressive (AR) and Vector-Autoregressive (VAR) Models . . . 6
2.2.2 Dynamic Nelson-Siegel (DNS) Models . . . 6
2.2.3 Diffusion Index Models . . . 8
2.3 Model evaluation . . . 9
2.3.1 Mean Squared Prediction Error (MSPE) . . . 9
2.3.2 Diebold-Mariano (DM) Test . . . 9

3 Data **10**
3.1 Yield data . . . 10
3.2 Macroeconomic data . . . 10
3.3 Samples . . . 10

4 Results **12**
4.1 January 1992 to December 2016 . . . 12
4.2 January 1992 to December 1999 . . . 14
4.3 January 2000 to December 2007 . . . 16
4.4 January 2008 to December 2016 . . . 18
4.5 DNS models . . . 20
4.6 DIF models . . . 20
4.7 PCA versus SPCA . . . 20

5 Conclusion **21**

A Model overview **22**

B DM-tests overview **26**
B.1 DNS Models . . . 26
B.2 DIF Models . . . 28
B.3 PCA versus SPCA . . . 31

1. Introduction

With the price per gigabyte dropping each year and Moore's law still holding up (Moore, 2006; Schaller, 1997; Waldrop, 2016), it is no surprise that companies are amassing very large sets of data. Forbes even coined the phrase: 'Data is the new oil', see Bhageshpur (2019). Consequently, a lot of research has been done the last 20 years for new method of analyzing and using these big datasets. One strand of this new research has focussed its attention to methods of penalized regression. Bell et al. (1978), Tibshirani (1996) and Zou and Hastie (2005) have contributed the ridge regressor, the lasso regressor and the elastic net, respectively. All of these methods are still used by researchers today. The amount of different shrinkage methods researchers can choose from today is immense, see Bai and Ng (2008, 2009), Hirano and Wright (2017), Kim and Swanson (2014, 2018), Schumacher (2010) and Stock and Watson (2012). Other research has focused its attention towards dimension reduction, principal component analysis being one of the best known. The question that remains to be answered is, what gains in predictive power have these new methods of data aggregation and reduction gotten us?

To compare these different models, we use them to model the yield curve of government bonds. Popular approaches of modeling the term structure of interest rates are no-arbitrage models, equilibrium models and factor models. The no-arbitrage approach focuses on fitting the term structure whilst ensuring that no arbitrage opportunities exist. Prominent publications for no-arbitrage modeling include Heath et al. (1992) and Hull and White (1990). The equilibrium approach focuses on modeling the dynamics of the instantaneous rate using affine models. Prominent publications for affine equilibrium models include Cox et al. (1985), Duffie and Kan (1996) and Vasicek (1977). No-arbitrage models and equilibrium models are notoriously bad for forecasting, see Duffee (2002). The factor model approach takes a more data driven approach to fitting the term structure. Prominent publications for factors models include Brandt and Yaron (2003), Dai and Singleton (2000), de Jong (2000), de Jong and Santa-Clara (1999), Diebold and Li (2006) and Duffee (2002). Combinations of the above mentioned approaches have also been published, see Christensen et al. (2009, 2011). We only consider the model presented by Diebold and Li (2006) in this paper.

In this paper, I replicate the findings of Swanson and Xiong (2017) by fitting numerous different models, some augmented with principal components based on a macroeconomic dataset, to yield curve data and compute its Mean Squared Prediction Error in order to find the model that is best forecasting future yields. Furthermore, I extend this research by including models augmented with sparse principal components obtained using the algorithm presented by Zou et al. (2006) and compare these models with the models augmented with principal components.

This paper is organized as follows: chapter 2 describes dimensions reduction techniques, gives a detailed explanation of the models used for forecasting and specifies the model evaluation techniques used in this paper. Chapter 3 describes how we obtained the data used for this research and the different samples we used, including a surface plot for visualizing the data. Chapter 4 gives a detailed interpretation of the our results and chapter 5 concludes this research and gives a short discussion. The code for reproduction of our results is available at https:

//github.com/ipsq/bachelorscriptie.

2. Methodology

To show the positive impact big data has had on our ability to correctly model the term structure of interest rates, we are comparing small data models with big data models. The models and predictive accuracy tests are based on those presented in Swanson and Xiong (2017). The small data models are autoregressive, vector autoregressive, and dynamic Nelson-Siegel models. The big data models are diffusion index models estimated from a macroeconomic dataset, using dimension reduction techniques described in section 2.1. Descriptive overview of the models used is given in appendix A. Detailed explanation of the models used is given in section 2.2. The models are (re-)estimated prior to construction of each forecast, using a rolling window of 120 months. Estimation of the models is done using maximum likelihood estimation and, where necessary, (sparse) principal component analysis. Monthly yield $h = 1-$step ahead forecasts are constructed for the bond maturities $\tau = 1-$, $2-$, $3-$, $5-$, and $10-$years. Performance of these models is assessed by mean square prediction error and models are compared for predictive accuracy, against the benchmark autoregressive model with a lag order of 1, using Diebold and Mariano (1994). We are also comparing the predictive accuracy of various variations of dynamic Nelson-Siegel models, diffusion index models, and we are comparing principal component analysis augmented models to sparse principal component analysis augmented models, using the test presented by Diebold and Mariano.

2.1 Dimension reduction

2.1.1 Principal Component Analysis (PCA)

Principal Component Analysis is a method of finding an orthogonal basis for your data in which different dimensions are uncorrelated. The vectors spanning this orthogonal basis are sorted beginning with the most variance explained to least variance explained. This method is often used as a dimension reduction technique by only including the first k factors to explain the data, whilst keeping most variance of the original dataset. The PCA coefficients are most often computed in one of two ways: eigenvalue decomposition on the correlation matrix, or singular value decomposition. In this paper, we only focus on singular value decomposition. Both methods result in the same set of principal components but the eigenvalue decomposition method is more restricted than the singular value decomposition. This is (part of) the reason why singular value decomposition is used for the sparse principal component analysis formulation.

Let $\mathbf{X} = \left(X_1 X_2 \cdots X_p\right)$ be our data matrix of size $n \times p$, where n is the number of observations and p the number of variables. Construct matrix $\mathbf{D}$ by centering and normalizing the matrix $\mathbf{X}$:

$$\mathbf{D} = \left(\frac{X_1 - \bar{X}_1}{\|X_1\|} \frac{X_2 - \bar{X}_2}{\|X_2\|} \cdots \frac{X_p - \bar{X}_p}{\|X_p\|} \right). \tag{2.1}$$

Construct the $p \times p$ correlation matrix $\mathbf{C} = \mathbf{D}^T \mathbf{D}/(n-1)$. Performing the singular value decomposition on the matrix $\mathbf{C}$ results in:

$$\mathbf{D} = \mathbf{U\Sigma V}^T, \tag{2.2}$$

where $\mathbf{U}$ is a unitary matrix, $\mathbf{\Sigma}$ is a diagonal matrix of singular values s_i, and $\mathbf{V}$ is a matrix of eigenvectors. The relationship between the singular values and the eigenvalues of the correlation matrix is the following:

$$\begin{aligned}\mathbf{C} &= \mathbf{D}^T\mathbf{D}/(n-1) \\ &= (\mathbf{U\Sigma V}^T)^T\mathbf{U\Sigma V}^T/(n-1) \\ &= \mathbf{V\Sigma U}^T\mathbf{U\Sigma V}^T/(n-1) \\ &= \mathbf{V\Sigma}^2\mathbf{V}^T/(n-1).\end{aligned} \tag{2.3}$$

From this, we can conclude that singular values are related to eigenvalues in the following way: $\lambda_i = s_i^2/(n-1)$. Principal components can be obtain using: $\mathbf{DV} = \mathbf{U\Sigma V}^T\mathbf{V} = \mathbf{U\Sigma}$.

2.1.2 Sparse Principal Component Analysis (SPCA)

Before we get to the definition of sparse principal component analysis, we first have to look at some regression regularization techniques. Bell et al. (1978) introduced the ridge regression. The ridge regression is a regularization method to mitigate the problem of multicollinearity in linear regression. It imposes a penalty on the squared value of the regression coefficients, effectively shrinking them towards zero.

$$\beta_{ridge} = \arg\min_{\beta} \left\{ \left\| Y - \sum_{j=1}^{p} X_j\beta_j \right\|^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \right\} \tag{2.4}$$

Tibshirani (1996) later introduced the *least absolute shrinkage and selection operator* (lasso) regression. This method imposes a penalty on the absolute value of the regression coefficients, effectively eliminating some coefficients from the model.

$$\beta_{lasso} = \arg\min_{\beta} \left\{ \left\| Y - \sum_{j=1}^{p} X_j\beta_j \right\|^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\} \tag{2.5}$$

Combining these two regularization methods resulted in what is now known as the elastic net, see Zou and Hastie (2005).

$$\beta_{elastic\ net} = (1+\lambda_2) \arg\min_{\beta} \left\{ \left\| Y - \sum_{j=1}^{p} X_j\beta_j \right\|^2 + \lambda_2 \sum_{j=1}^{p} \beta_j^2 + \lambda_1 \sum_{j=1}^{p} |\beta_j| \right\} \tag{2.6}$$

Factors estimated with Principal Component Analysis create linear combinations of the underlying predictor variables with all non-zero factor loadings. This negatively affects the parsimony of forecasting models and makes interpreting the factor loadings difficult. To overcome this, Zou et al. (2006) introduced a method of finding principal components with sparse factor loadings by formulating PCA as a regression problem and applying the elastic net.

$$\begin{aligned}\left(\widehat{\mathbf{A}}, \widehat{\mathbf{B}}\right) = \arg\min_{\mathbf{A},\mathbf{B}} & \sum_{i=1}^{n} \left\|\mathbf{x}_i - \mathbf{A}\mathbf{B}^T\mathbf{x}_i\right\|^2 + \lambda_2 \sum_{j=1}^{k} \|\beta_j\|^2 + \sum_{j=1}^{k} \lambda_{1,j} \|\beta_j\|_1 \\ & \text{subject to } \mathbf{A}^T\mathbf{A} = \mathbf{I}_{k\times k}\end{aligned} \tag{2.7}$$

One can obtain the factors as columns, after normalization, of the matrix $\widehat{\mathbf{B}}$.

2.2 Models

For a summary of the models used, see appendix A. The details of each model is given below. Each section contains a short list of the corresponding models from appendix A.

2.2.1 Autoregressive (AR) and Vector-Autoregressive (VAR) Models

Models in this section are summarized by: AR(1), VAR(1), VAR(1)+FB1(PCA), VAR(1)+FB2(PCA), VAR(1)+FB1(SPCA), VAR(1)+FB2(SPCA), AR(SIC), VAR(SIC), VAR(SIC)+FB1(PCA), VAR(SIC)+FB2(PCA), VAR(SIC)+FB1(SPCA), and VAR(SIC)+FB2(SPCA). See appendix A, panel A.

These AR(p) and VAR(p) models are formulated as follows:

$$\hat{y}_{t+1}(\tau) = \hat{c} + \hat{\beta}' W_t, \tag{2.8}$$

where τ denotes the maturity and $\hat{y}_{t+1}(\tau)$ measures the 1-step ahead annual yield forecast of the bond. For the autoregressive model, W_t contains p lags of $y_{t+1}(\tau)$. For the vector autoregressive model, W_t additionally contains yields of bonds of different maturities than τ. In both models $\hat{\beta}$ is a time-invariant coefficient vector and $\hat{c}$ is a (vector of) constant(s). We will be including AR(1) and VAR(1) models as a baseline. Furthermore, we will be including AR and VAR specifications with up to 5 lags of $y_{t+1}(\tau)$ included. The number of lags shall be selected using the Schwarz information criterion. The Schwarz information criterion, also commonly referred to as the Bayesian information criterion, published by Schwarz (1978, hereafter SIC) is a model selection criterium, closely related to the Akaike information criterion (Akaike, 1974, hereafter AIC). It is defined as follows:

$$SIC = k \ln n - 2 \ln \hat{L}, \tag{2.9}$$

where $\hat{L}$ is the likelihood of the model, n the number of observations and k the number of model parameters estimated. The difference between SIC and AIC is the higher penalty given for the addition of extra parameters to compensate for the increase in likelihood.

2.2.2 Dynamic Nelson-Siegel (DNS) Models

Models in this section are summarized by: DNS(1-6), DNS(1-6)+FB1(PCA), DNS(1-6)+FB2(PCA), DNS(1-6)+FB1(SPCA), DNS(1-6)+FB2(SPCA), and DNS(1-6)+MAC. See appendix A, panel B.

The Nelson-Siegel model, published by Nelson and Siegel (1987), models the cross-sectional movement of the term structure using three underlying latent factors interpreted as 'level', 'slope', and 'curvature'. These factors are commonly known as the 'Nelson-Siegel factors'. The dynamic version of the Nelson-Siegel model was published by Diebold and Li (2006, hereafter DNS). The DNS model is formulated as follows:

$$\hat{y}_{t+1}(\tau) = \hat{\beta}^f_{1,t+1} + \hat{\beta}^f_{2,t+1}\left[\frac{1-\exp(-\lambda_t\tau)}{\lambda_t\tau}\right] + \hat{\beta}^f_{3,t+1}\left[\frac{1-\exp(-\lambda_t\tau)}{\lambda_t\tau} - \exp(-\lambda_t\tau)\right]. \tag{2.10}$$

To estimate the DNS model, we make use of the two-step formulation, presented by Diebold and Li (2006). The first step is obtaining the parameters $\hat{\beta}^f_{1,t}, \hat{\beta}^f_{2,t}, \hat{\beta}^f_{3,t}$ by regressing

$$\left\{1, \left[\frac{1-\exp(-\lambda_t\tau)}{\lambda_t\tau}\right], \left[\frac{1-\exp(-\lambda_t\tau)}{\lambda_t\tau} - \exp(-\lambda_t\tau)\right]\right\}$$

on $\mathbf{y}^f_t(\tau)$. We consider 3 variants of $\mathbf{y}^f_t(\tau)$ in this paper, namely:

$$\begin{aligned}
\mathbf{y}^{10}_t(\tau) &= [y_t(12)\ y_t(24)\ y_t(36)\ y_t(48)\ y_t(60)\ y_t(72)\ y_t(84)\ y_t(96)\ y_t(108)\ y_t(120)]', \\
\mathbf{y}^{6}_t(\tau) &= [y_t(12)\ y_t(24)\ y_t(36)\ y_t(60)\ y_t(84)\ y_t(120)]', \\
\mathbf{y}^{4}_t(\tau) &= [y_t(12)\ y_t(36)\ y_t(60)\ y_t(120)]'.
\end{aligned}$$

Diebold and Li (2006) fix λ_t to 0.0609. This maximizes the loading on the curvature factor at a 30 month maturity. The second step is estimating $\hat{\beta}^f_{1,t+1}$, $\hat{\beta}^f_{2,t+1}$, and $\hat{\beta}^f_{3,t+1}$. We estimate these parameters using both an AR(1) and a VAR(1) model. The AR(1) model is formulated as follows:

$$\hat{\beta}^f_{i,t+1} = \hat{c}_i + \hat{\gamma}_{ii}\hat{\beta}^f_{i,t}, \quad \text{for } i = 1,2,3, \tag{2.11}$$

where $\hat{c}_i$, $\hat{\gamma}_{ii}$, and $\hat{\beta}^f_{i,t}$ are scalars. The VAR(1) model is formulated as follows:

$$\hat{\beta}^f_{t+1} = \hat{c} + \hat{\Gamma}\hat{\beta}^f_t, \tag{2.12}$$

where $\hat{c}$ is a 3×1 vector, $\hat{\beta}^f_t = \left(\hat{\beta}^f_{1,t}, \hat{\beta}^f_{2,t}, \hat{\beta}^f_{3,t}\right)'$ a 3×1 vector, and $\hat{\Gamma} = (\hat{\gamma}_1, \hat{\gamma}_2, \hat{\gamma}_3)$, with $\hat{\gamma}_j$ a 3×1 vector, for $j = 1,2,3$.

DNS Models with Macroeconomic Variables

The DNS model, described in section 2.2.2, can be extended by including macroeconomic variables in the AR(1) model, given in eq. (2.11), as follows:

$$\hat{\beta}^f_{i,t+1} = \hat{c}_i + \hat{\gamma}_{ii}\hat{\beta}^f_{i,t} + \hat{\alpha}'_i M_t, \quad \text{for } i = 1,2,3, \tag{2.13}$$

where $\hat{\alpha}_i$ is a 3×1 vector, and M_t is a 3×1 vector containing the macroeconomic variables 'manufacturing capacity utilization', 'the federal funds rate', and 'the annual personal consumption expenditures price deflator'. All other terms are conformably defined. Analogous to the AR(1) model, the VAR(1) given in eq. (2.12), can also be extended by including macroeconomic variables:

$$\hat{\beta}^f_{t+1} = \hat{c} + \hat{\Gamma}\hat{\beta}^f_t + \hat{\mathrm{A}} M_t, \tag{2.14}$$

where $\hat{\mathrm{A}} = (\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3)$, with $\hat{\alpha}_j$ a 3×1 vector, for $j = 1, 2, 3$. All other terms are conformably defined.

DNS Models with Diffusion Indexes

The DNS model given in section 2.2.2, can also be extended by including diffusion indexes. The extended version of the AR(1) model is formulated as follows:

$$\hat{\beta}^f_{i,t+1} = \hat{c}_i + \hat{\gamma}_{ii}\hat{\beta}^f_{i,t} + \hat{\xi}' F^b_t, \quad \text{for } i = 1, 2, 3, \tag{2.15}$$

where F^b_t is defined as in section 2.2.3 and $\hat{\xi}$ is a conformably sized (depending on the amount of factors included) vector of coefficients. All other terms are conformably defined. Analogous to the AR(1) model, the VAR(1) given in eq. (2.12), can also be extended by including diffusion indexes:

$$\hat{\beta}^f_{t+1} = \hat{c} + \hat{\Gamma}\hat{\beta}^f_t + \hat{\Xi} F^b_t, \tag{2.16}$$

where $\hat{\Xi}$ is a conformably sized (depending on the amount of factors included) matrix of coefficients. All other terms are conformably defined.

2.2.3 Diffusion Index Models

Models in this section are summarized by: DIF(1-9), DIF(1-3)+FB1(PCA), DIF(1-3)+FB2(PCA), DIF(1-3)+FB1(SPCA), and DIF(1-3)+FB2(SPCA). See appendix A, panel C.

Diffusion Index Models, also commonly referred to as 'factor augmented forecast models', as published by Stock and Watson (2002a, 2002b, hereafter DIF), is an extension to regular forecast models, mostly (vector-)autoregressive models, supplemented by a vector of latent factors. This model is formulated as follows:

$$\hat{y}_{t+1}(\tau) = \hat{c} + \hat{\beta}' W_t + \hat{\xi}' F^b_t + \hat{\eta}' F^s_t, \tag{2.17}$$

where F^b_t contains either 1, 2, or 3 latent factors, estimated from the set of macroeconomic variables, using either PCA or SPCA (see section 2.1.2), F^s_t contains either 1, 2, or 3 latent factors, estimated with PCA, from a set of 10 yields given by $\mathbf{y}^{10}_t(\tau)$, see section 2.2.2, and $\hat{\xi}$ and $\hat{\eta}$ are conformably sized (depending on the amount of factors included) vectors of coefficients.

2.3 Model evaluation

2.3.1 Mean Squared Prediction Error (MSPE)

The Mean Square Prediction Error, hereafter MSPE, is a metric for evaluating models in forecasting. It is defined as follows:

$$MSPE = \frac{1}{q} \sum_{i=n+1}^{n+q} \left(Y_i - \hat{Y}_i \right)^2, \tag{2.18}$$

where n is the training sample and q is the forecasting sample.

2.3.2 Diebold-Mariano (DM) Test

The Diebold-Mariano Test, published by Diebold and Mariano (1994, hereafter DM), is a predictive accuracy test between two models. The hypotheses are defined as follows:

$$H_0 : E\left[u_{0,t+h}^2 - u_{1,t+h}^2\right] = 0,$$
$$H_A : E\left[u_{0,t+h}^2 - u_{1,t+h}^2\right] \neq 0,$$

where $u_{0,t+h}$ and $u_{1,t+h}$ are h-step ahead forecast errors. In practice, we only observe estimates of these out-of-sample h-step ahead forecasts, i.e. $\hat{u}_{0,t+h}$ and $\hat{u}_{1,t+h}$. The standard version of the DM predictive accuracy test is formulated as follows:

$$DM_p = \frac{\bar{d}_t}{\hat{\sigma}_{\bar{d}_t}} \xrightarrow{d} N(0,1), \tag{2.19}$$

where

$$\bar{d}_t = \frac{1}{q} \sum_{t=n+1}^{n+q} d_t, \; d_t = \hat{u}_{0,t+h}^2 - \hat{u}_{1,t+h}^2, \text{ and } \hat{\sigma}_{\bar{d}_t} = \frac{\hat{\sigma}_{d_t}}{\sqrt{q}}. \tag{2.20}$$

In the formulation above, q is the number of forecasts, n is the number of observations in the training sample, and $\hat{\sigma}_{d_t}$ is the standard deviation of the forecast errors.

3. Data

3.1 Yield data

For our yield data, we make use of the dataset published by Gürkaynak et al. (2007). This dataset consists of United States Department of Treasury zero-coupon, continuously compounded, yield curve data. The dataset contains daily yield curve data from the 14th of June 1961 to the 1st of May 2020, at the time of writing. This dataset contains maturities from 1 year up to 30 years. For our research only the maturities from 1 year up to and including 10 years are used. Because yield curve estimates only extended out to 7 years until the 16th of August 1971 (see Gürkaynak et al., 2007, p. 19), our sample will start from August 1971 to December 2016.

3.2 Macroeconomic data

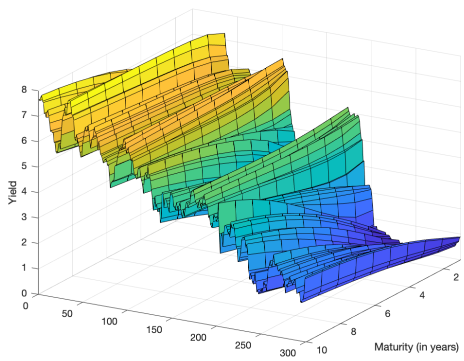
For our macroeconomic data, used for constructing the (macro-)factors, we make use of the FRED-MD dataset, published by McCracken and Ng (2016). This dataset is developed by the United States Federal Reserve Bank of St. Louis and consists of 134 monthly U.S. macroeconomic variables from January 1959 to March 2020, at the time of writing. By deleting all variables containing blank cells, only 105 variables remain. We make use of the same sample as the yield data.

3.3 Samples

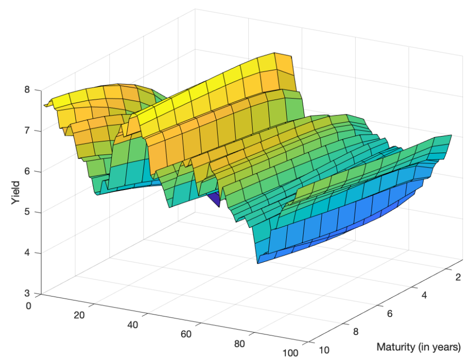
For our research, we use the following samples to compare our various different models:

1. Full sample, January 1992 to December 2016,
2. Subsample 1, January 1992 to December 1999, this is our pre-dot-com bubble crash sample;
3. Subsample 2, January 2000 to December 2007, this is our dot-com bubble crash and 9/11 sample leading up to the great recession;
4. Subsample 3, January 2008 to December 2016, this is our (post-)great recession sample.

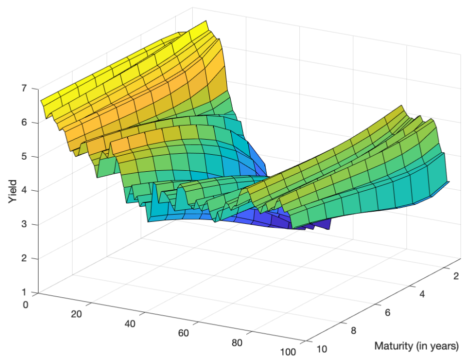
These samples roughly correspond to the samples used by Swanson and Xiong (2017). Surface plots of these samples are shown in fig. 3.1.



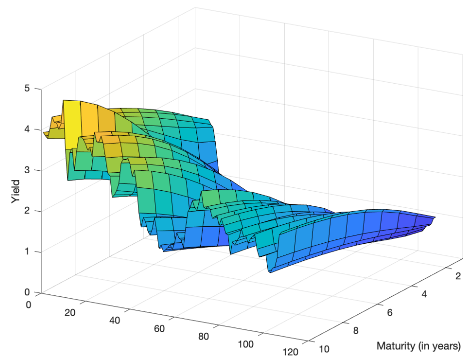
(a) Full sample, January 1992 to December 2016, 300 observations



(b) Subsample 1, January 1992 to December 1999, 96 observations



(c) Subsample 2, January 2000 to December 2007, 96 observations



(d) Subsample 3, January 2008 to December 2016, 108 observations

Figure 3.1: Surface plots of the yields for the 4 different samples

4. Results

Tables 4.1 to 4.4 contain relative MSPEs for yield forecasts constructed using the models listed in appendix A, for $h = 1$, for $\tau = 1, 2, 3, 5, 10$ year maturities, and for 4 different forecasting samples, listed in section 3.3. We used the simple AR(1) model as the benchmark for the relative MSPEs. Analyzing our results, some observations can be made. When comparing our results to those of Swanson and Xiong (2017), some similarities exist but the order of magnitude of our results differ greatly to those of Swanson and Xiong. Especially in subsample 1 and the full sample, do our models almost never beat the simple benchmark model, contrary to the results of Swanson and Xiong, where the factor augmented DNS models perform best for the lower maturities ($\tau = 1, 2, 3$).

The reason for these large deviations in results compared to Swanson and Xiong (2017) could have a couple reasons. All our models were implemented using maximum likelihood estimation (MLE). MLE formulations are most commonly solved using line search algorithms which do not provide a deterministic result and thus could differ from run to run. We also used a slightly different sample from the ones used by Swanson and Xiong. We extended subsample 1 and the full sample to include 3 more months of post-recession data and extended subsample 3 and the full sample to include additional data points Swanson and Xiong did not have access to at the time. Another possible reason for the deviations could be that Swanson and Xiong demeaned the data and removed any seasonality. There was no reference of this in Swanson and Xiong (2017) but it is commonly applied to data when working with AR and VAR models since they cannot correctly model non-stationary processes.

Moving forward, all tables contain the relative mean squared prediction errors, relative to the benchmark AR(1) model. The models, as listed in the first column, are summarized in appendix A. Entries in bold denote models with the lowest MSPE for a given maturity. Entries superscripted with *, **, and *** denote rejections of the null hypotheses of equal predictive accuracy compared with the benchmark AR(1) at 0.10, 0.05, and 0.01 significance levels, respectively, based on the application of the DM-test, see section 2.3.2.

4.1 January 1992 to December 2016

Looking at our results from the full sample in table 4.1, we can make the following observations. For the maturities 1, 2, 3, and 5 years, the AR models perform best based on raw MSPE. With the AR model with lags based on the Schwarz Information Criterion again being the best performing model for the 1 year maturity, based on raw MSPE. It is also the only model for this sample which rejects the null hypothesis for equal predictive accuracy, with a significance of 0.10. For the 10 year maturity, we observe that the dynamic Nelson-Siegel model with AR formulation, with or without factor augmentation, beats the benchmark on MSPE. From these models, the DNS model with AR formulation and 1 principal component again having the lowest MSPE.

Table 4.1: 1-step-ahead relative MSPEs of all forecasting models (Full sample: 1992:1-2016:12)

Model	rMSPE				
Maturity	1 year	2 years	3 years	5 years	10 years
AR(1)	1	**1**	**1**	**1**	1
VAR(1)	1.060	1.098	1.097	1.102	1.144
VAR(1)+FB1(PCA)	1.079	1.091	1.076	1.071	1.091
VAR(1)+FB2(PCA)	1.111	1.121	1.102	1.094	1.126
VAR(1)+FB1(SPCA)	1.069	1.090	1.079	1.075	1.094
VAR(1)+FB2(SPCA)	1.071	1.097	1.088	1.088	1.126
AR(SIC)	**0.920***	1.001	1.011	1.015	1.006
VAR(SIC)	1.060	1.098	1.097	1.102	1.144
VAR(SIC)+FB1(PCA)	1.079	1.091	1.076	1.071	1.091
VAR(SIC)+FB2(PCA)	1.111	1.121	1.102	1.094	1.126
VAR(SIC)+FB1(SPCA)	1.069	1.090	1.079	1.075	1.094
VAR(SIC)+FB2(SPCA)	1.071	1.097	1.088	1.088	1.126
DNS(1)	1.308	1.045	1.042	1.154	0.982
DNS(2)	1.270	1.049	1.061	1.193	0.979
DNS(3)	1.207	1.049	1.036	1.188	0.996
DNS(4)	1.119	1.117	1.079	1.160	1.063
DNS(5)	1.103	1.101	1.081	1.193	1.057
DNS(6)	1.090	1.157	1.083	1.182	1.066
DNS(1)+FB1(PCA)	1.186	1.167	1.063	1.008	1.016
DNS(2)+FB1(PCA)	1.172	1.138	1.044	1.022	1.008
DNS(3)+FB1(PCA)	1.167	1.222	1.077	1.020	**0.965**
DNS(4)+FB1(PCA)	1.135	1.195	1.110	1.081	1.083
DNS(5)+FB1(PCA)	1.122	1.168	1.093	1.097	1.075
DNS(6)+FB1(PCA)	1.127	1.252	1.124	1.097	1.039
DNS(1)+FB2(PCA)	1.264	1.264	1.115	1.044	1.040
DNS(2)+FB2(PCA)	1.253	1.224	1.088	1.056	1.032
DNS(3)+FB2(PCA)	1.271	1.323	1.129	1.058	0.995
DNS(4)+FB2(PCA)	1.169	1.310	1.185	1.092	1.149
DNS(5)+FB2(PCA)	1.163	1.271	1.154	1.100	1.140
DNS(6)+FB2(PCA)	1.181	1.387	1.207	1.100	1.091
DNS(1)+FB1(SPCA)	1.215	1.147	1.062	1.032	1.012
DNS(2)+FB1(SPCA)	1.197	1.125	1.050	1.050	1.004
DNS(3)+FB1(SPCA)	1.180	1.195	1.072	1.046	0.967
DNS(4)+FB1(SPCA)	1.110	1.181	1.102	1.086	1.072
DNS(5)+FB1(SPCA)	1.099	1.155	1.088	1.104	1.063
DNS(6)+FB1(SPCA)	1.105	1.237	1.115	1.102	1.034
DNS(1)+FB2(SPCA)	1.210	1.251	1.106	1.040	1.040
DNS(2)+FB2(SPCA)	1.202	1.209	1.077	1.051	1.032
DNS(3)+FB2(SPCA)	1.226	1.311	1.119	1.054	0.998
DNS(4)+FB2(SPCA)	1.097	1.278	1.159	1.083	1.134
DNS(5)+FB2(SPCA)	1.094	1.238	1.128	1.093	1.125
DNS(6)+FB2(SPCA)	1.118	1.356	1.182	1.093	1.084
DNS(1)+MAC	1.295	1.310	1.187	1.119	1.057
DNS(2)+MAC	1.281	1.275	1.165	1.135	1.048
DNS(3)+MAC	1.293	1.367	1.201	1.135	1.018
DNS(4)+MAC	1.063	1.176	1.124	1.130	1.110
DNS(5)+MAC	1.059	1.156	1.115	1.152	1.100
DNS(6)+MAC	1.051	1.231	1.138	1.146	1.072
DIF(1)	1.061	1.105	1.102	1.105	1.160
DIF(2)	1.096	1.138	1.136	1.142	1.204
DIF(3)	1.108	1.147	1.144	1.152	1.228
DIF(4)	1.079	1.091	1.076	1.071	1.091
DIF(5)	1.111	1.121	1.102	1.094	1.126
DIF(6)	1.113	1.132	1.125	1.124	1.137
DIF(7)	1.069	1.090	1.079	1.075	1.094
DIF(8)	1.071	1.097	1.088	1.088	1.126
DIF(9)	1.074	1.113	1.119	1.136	1.159
DIF(1)+FB1(PCA)	1.096	1.112	1.094	1.086	1.119
DIF(2)+FB1(PCA)	1.128	1.145	1.128	1.122	1.162
DIF(3)+FB1(PCA)	1.161	1.175	1.157	1.156	1.225
DIF(1)+FB2(PCA)	1.131	1.146	1.125	1.111	1.146
DIF(2)+FB2(PCA)	1.166	1.179	1.159	1.147	1.189
DIF(3)+FB2(PCA)	1.194	1.206	1.187	1.181	1.249
DIF(1)+FB1(SPCA)	1.086	1.111	1.096	1.088	1.125
DIF(2)+FB1(SPCA)	1.120	1.145	1.132	1.127	1.174
DIF(3)+FB1(SPCA)	1.155	1.176	1.160	1.158	1.233
DIF(1)+FB2(SPCA)	1.086	1.119	1.109	1.106	1.151
DIF(2)+FB2(SPCA)	1.123	1.154	1.145	1.146	1.202
DIF(3)+FB2(SPCA)	1.160	1.188	1.178	1.183	1.262

4.2 January 1992 to December 1999

Looking only at our results for subsample 1 in table 4.2, we can make the following observations. The AR(SIC) model ‘wins’ for all maturities except $\tau = 5$. This maturity seems to be best represented by the benchmark AR(1) model. The AR(SIC) is the only model to out perform the benchmark model twice, based on the DM-test with 0.05 significance. It appears that this period, when the economy was first recovering but later on booming, is a hard one for our models to predict. Looking at the other models, only the standard DNS models, without factor augmentation, regressed on 10 different maturity yields, performed reasonably well coming in second for the 1 year maturity. For the 2 year maturity, the second best performing model is the diffusion index model with two SPCA factors. This model appears to be performing consistently for the early maturities, after which it fails to predict the farther out maturities. For the 3, 5, and 10 year maturities, the DNS models without factors augmentation are all the second best performing models. Comparing the DNS models with AR formulations against those with VAR formulations, we can see that the AR formulations consistently out perform the VAR formulations. After performing a DM test between these models, we cannot reject the null hypothesis of equal predictive accuracy.

Table 4.2: 1-step-ahead relative MSPEs of all forecasting models (Subsample 1: 1992:1-1999:12)

Model	rMSPE				
Maturity	1 year	2 years	3 years	5 years	10 years
AR(1)	1	1	1	**1**	**1**
VAR(1)	1.101	1.097	1.091	1.094	1.156
VAR(1)+FB1(PCA)	1.143	1.120	1.098	1.085	1.116
VAR(1)+FB2(PCA)	1.095	1.096	1.090	1.084	1.108
VAR(1)+FB1(SPCA)	1.137	1.118	1.099	1.091	1.130
VAR(1)+FB2(SPCA)	1.048	1.068	1.073	1.081	1.121
AR(SIC)	**0.904****	**0.944****	**0.983**	1.006	**1**
VAR(SIC)	1.101	1.097	1.091	1.094	1.156
VAR(SIC)+FB1(PCA)	1.143	1.120	1.098	1.085	1.116
VAR(SIC)+FB2(PCA)	1.095	1.096	1.090	1.084	1.108
VAR(SIC)+FB1(SPCA)	1.137	1.118	1.099	1.091	1.130
VAR(SIC)+FB2(SPCA)	1.048	1.068	1.073	1.081	1.121
DNS(1)	1.015	1.082	1.039	1.008	1.060
DNS(2)	1.019	1.071	1.029	1.013	1.056
DNS(3)	1.025	1.109	1.044	1.012	1.023
DNS(4)	1.088	1.152	1.093	1.060	1.116
DNS(5)	1.094	1.137	1.082	1.069	1.112
DNS(6)	1.096	1.184	1.098	1.056	1.083
DNS(1)+FB1(PCA)	1.136	1.231	1.131	1.022	1.129
DNS(2)+FB1(PCA)	1.143	1.207	1.107	1.018	1.123
DNS(3)+FB1(PCA)	1.170	1.278	1.146	1.017	1.061
DNS(4)+FB1(PCA)	1.125	1.234	1.151	1.058	1.200
DNS(5)+FB1(PCA)	1.135	1.216	1.132	1.055	1.191
DNS(6)+FB1(PCA)	1.147	1.272	1.160	1.047	1.118
DNS(1)+FB2(PCA)	1.150	1.165	1.119	1.070	1.140
DNS(2)+FB2(PCA)	1.146	1.153	1.107	1.072	1.136
DNS(3)+FB2(PCA)	1.147	1.192	1.129	1.077	1.084
DNS(4)+FB2(PCA)	1.125	1.145	1.121	1.096	1.187
DNS(5)+FB2(PCA)	1.122	1.136	1.114	1.102	1.181
DNS(6)+FB2(PCA)	1.100	1.154	1.120	1.098	1.126
DNS(1)+FB1(SPCA)	1.112	1.219	1.126	1.015	1.156
DNS(2)+FB1(SPCA)	1.120	1.195	1.101	1.009	1.147
DNS(3)+FB1(SPCA)	1.151	1.269	1.143	1.010	1.074
DNS(4)+FB1(SPCA)	1.124	1.240	1.155	1.055	1.216
DNS(5)+FB1(SPCA)	1.133	1.221	1.135	1.051	1.205
DNS(6)+FB1(SPCA)	1.148	1.278	1.164	1.044	1.126
DNS(1)+FB2(SPCA)	1.080	1.128	1.090	1.047	1.145
DNS(2)+FB2(SPCA)	1.078	1.116	1.078	1.048	1.140
DNS(3)+FB2(SPCA)	1.081	1.158	1.102	1.052	1.082
DNS(4)+FB2(SPCA)	1.069	1.125	1.103	1.075	1.195
DNS(5)+FB2(SPCA)	1.067	1.114	1.093	1.079	1.187
DNS(6)+FB2(SPCA)	1.053	1.138	1.103	1.075	1.125
DNS(1)+MAC	1.158	1.374	1.245	1.072	1.246
DNS(2)+MAC	1.171	1.340	1.209	1.055	1.234
DNS(3)+MAC	1.222	1.433	1.264	1.059	1.140
DNS(4)+MAC	1.062	1.243	1.148	1.025	1.233
DNS(5)+MAC	1.089	1.228	1.129	1.021	1.225
DNS(6)+MAC	1.118	1.297	1.164	1.012	1.127
DIF(1)	1.119	1.124	1.123	1.126	1.181
DIF(2)	1.105	1.120	1.122	1.125	1.172
DIF(3)	1.105	1.113	1.113	1.116	1.170
DIF(4)	1.143	1.120	1.098	1.085	1.116
DIF(5)	1.095	1.096	1.090	1.084	1.108
DIF(6)	1.083	1.109	1.127	1.152	1.210
DIF(7)	1.137	1.118	1.099	1.091	1.130
DIF(8)	1.048	1.068	1.073	1.081	1.121
DIF(9)	1.061	1.102	1.128	1.170	1.249
DIF(1)+FB1(PCA)	1.186	1.167	1.145	1.114	1.105
DIF(2)+FB1(PCA)	1.174	1.165	1.147	1.116	1.101
DIF(3)+FB1(PCA)	1.165	1.155	1.140	1.118	1.111
DIF(1)+FB2(PCA)	1.160	1.157	1.147	1.123	1.109
DIF(2)+FB2(PCA)	1.166	1.165	1.155	1.129	1.107
DIF(3)+FB2(PCA)	1.179	1.174	1.164	1.143	1.128
DIF(1)+FB1(SPCA)	1.170	1.156	1.137	1.111	1.110
DIF(2)+FB1(SPCA)	1.158	1.154	1.139	1.112	1.104
DIF(3)+FB1(SPCA)	1.149	1.145	1.133	1.114	1.113
DIF(1)+FB2(SPCA)	1.108	1.121	1.120	1.108	1.108
DIF(2)+FB2(SPCA)	1.113	1.129	1.128	1.113	1.104
DIF(3)+FB2(SPCA)	1.129	1.139	1.139	1.127	1.122

4.3 January 2000 to December 2007

Looking at our results from subsample 2 in table 4.3, we immediately observe that the benchmark model is no longer the best in class. For the 1 year maturity, it is the DNS model with VAR formulation, augmented with three macroeconomic variables, that performs best. We also observe that the diffusion index models are performing reasonably well for this maturity. Unfortunately, for this maturity, not a single model beat the benchmark model according to the DM test. The 2 year maturity is best represented by a diffusion index model augmented with 3 sparse principal components. The DNS models with VAR formulation and macroeconomic variables also performs better than the benchmark, based on raw MSPE. For this maturity, all models beating the benchmark on raw MSPE are either augmented with sparse principal components or the three macroeconomic variables. For the maturities 3, 5, and 10 years, the 'best' models are all DNS models augmented with sparse principal components. For the 5 year maturity, 11 DNS models augmented with (sparse) principal components, reject the null hypothesis of equal predictive accuracy, compared with the benchmark model.

Table 4.3: 1-step-ahead relative MSPEs of all forecasting models (Subsample 2: 2000:1-2007:12)

Model	rMSPE				
Maturity	1 year	2 years	3 years	5 years	10 years
AR(1)	1	1	1	1	1
VAR(1)	0.976	1.036	1.038	1.052	1.115
VAR(1)+FB1(PCA)	0.980	1.019	1.012	1.019	1.048
VAR(1)+FB2(PCA)	0.970	1.005	0.999	1.011	1.059
VAR(1)+FB1(SPCA)	0.965	1.014	1.011	1.020	1.050
VAR(1)+FB2(SPCA)	0.936	0.979	0.979	0.997	1.059
AR(SIC)	0.937	1.040	1.042	1.036	1.012
VAR(SIC)	0.976	1.036	1.038	1.052	1.115
VAR(SIC)+FB1(PCA)	0.980	1.019	1.012	1.019	1.048
VAR(SIC)+FB2(PCA)	0.970	1.005	0.999	1.011	1.059
VAR(SIC)+FB1(SPCA)	0.965	1.014	1.011	1.020	1.050
VAR(SIC)+FB2(SPCA)	0.936	0.979	0.979	0.997	1.059
DNS(1)	1.185	1.009	0.991	1.080	0.963*
DNS(2)	1.160	1.010	1.002	1.106	0.962*
DNS(3)	1.129	1.011	0.990	1.110	0.982
DNS(4)	1.019	1.075	1.039	1.089	1.033
DNS(5)	1.016	1.065	1.039	1.116	1.033
DNS(6)	1.025	1.109	1.046	1.106	1.038
DNS(1)+FB1(PCA)	1.113	1.086	0.997	0.943**	0.971
DNS(2)+FB1(PCA)	1.103	1.065	0.982	0.950*	0.963
DNS(3)+FB1(PCA)	1.119	1.123	1.009	0.954	0.924*
DNS(4)+FB1(PCA)	1.095	1.136	1.069	1.023	1.045
DNS(5)+FB1(PCA)	1.084	1.117	1.055	1.031	1.039
DNS(6)+FB1(PCA)	1.114	1.182	1.087	1.039	0.996
DNS(1)+FB2(PCA)	1.096	1.193	1.040	0.930**	0.995
DNS(2)+FB2(PCA)	1.102	1.155	1.010	0.933**	0.986
DNS(3)+FB2(PCA)	1.157	1.248	1.054	0.933**	0.941*
DNS(4)+FB2(PCA)	1.040	1.226	1.119	0.975	1.115
DNS(5)+FB2(PCA)	1.040	1.189	1.081	0.966	1.103
DNS(6)+FB2(PCA)	1.084	1.292	1.138	0.965	1.024
DNS(1)+FB1(SPCA)	1.069	1.063	0.980	0.930**	0.965
DNS(2)+FB1(SPCA)	1.060	1.042	**0.964**	0.936*	0.957
DNS(3)+FB1(SPCA)	1.078	1.101	0.992	0.941*	**0.916****
DNS(4)+FB1(SPCA)	1.051	1.112	1.051	1.022	1.031
DNS(5)+FB1(SPCA)	1.043	1.094	1.039	1.033	1.025
DNS(6)+FB1(SPCA)	1.075	1.158	1.068	1.039	0.992
DNS(1)+FB2(SPCA)	1.050	1.192	1.033	0.908**	1.000
DNS(2)+FB2(SPCA)	1.060	1.150	0.998	**0.906****	0.991
DNS(3)+FB2(SPCA)	1.122	1.254	1.050	0.907**	0.939*
DNS(4)+FB2(SPCA)	0.955	1.189	1.081	0.950	1.091
DNS(5)+FB2(SPCA)	0.963	1.151	1.043	0.943	1.080
DNS(6)+FB2(SPCA)	1.014	1.258	1.101	0.942	1.008
DNS(1)+MAC	1.127	1.081	1.051	1.100	0.992
DNS(2)+MAC	1.107	1.071	1.052	1.123	0.989
DNS(3)+MAC	1.101	1.095	1.052	1.125	0.999
DNS(4)+MAC	0.926	0.994	1.022	1.101	1.033
DNS(5)+MAC	0.910	0.989	1.026	1.124	1.028
DNS(6)+MAC	**0.890**	1.011	1.028	1.124	1.030
DIF(1)	0.974	1.045	1.052	1.075	1.149
DIF(2)	1.076	1.114	1.105	1.113	1.179
DIF(3)	1.102	1.127	1.112	1.120	1.193
DIF(4)	0.980	1.019	1.012	1.019	1.048
DIF(5)	0.970	1.005	0.999	1.011	1.059
DIF(6)	0.947	1.000	1.010	1.035	1.097
DIF(7)	0.965	1.014	1.011	1.020	1.050
DIF(8)	0.936	0.979	0.979	0.997	1.059
DIF(9)	0.923	**0.976**	0.990	1.023	1.098
DIF(1)+FB1(PCA)	1.007	1.059	1.056	1.062	1.093
DIF(2)+FB1(PCA)	1.097	1.124	1.110	1.112	1.153
DIF(3)+FB1(PCA)	1.169	1.185	1.164	1.160	1.194
DIF(1)+FB2(PCA)	0.987	1.043	1.045	1.061	1.115
DIF(2)+FB2(PCA)	1.065	1.099	1.093	1.107	1.173
DIF(3)+FB2(PCA)	1.109	1.134	1.125	1.142	1.213
DIF(1)+FB1(SPCA)	0.995	1.051	1.050	1.059	1.095
DIF(2)+FB1(SPCA)	1.086	1.118	1.106	1.112	1.164
DIF(3)+FB1(SPCA)	1.161	1.176	1.155	1.154	1.198
DIF(1)+FB2(SPCA)	0.956	1.015	1.022	1.045	1.114
DIF(2)+FB2(SPCA)	1.039	1.075	1.073	1.095	1.184
DIF(3)+FB2(SPCA)	1.097	1.118	1.109	1.130	1.217

4.4 January 2008 to December 2016

Looking at our results from subsample 3 in table 4.4, the results are similar to those of the full sample. For the maturities 1, 2, 3, and 5 years, no model manages to 'beat' the benchmark model on raw MSPE. Only the AR model with lags based on the Schwarz Information Criterion manages to beat its one lagged counter part for the 1 year maturity. With all the other, more complicated, models not even close to its raw MSPE. Only for the 10 year maturity, do we see that the DNS models start to outperform the benchmark model on raw MSPE, with the DNS model with AR formulation and 1 principal component having the lowest MSPE. No model rejects the null hypothesis of equal predictive accuracy for any maturity.

Table 4.4: 1-step-ahead relative MSPEs of all forecasting models (Subsample 3: 2008:1-2016:12)

Model	rMSPE				
Maturity	1 year	2 years	3 years	5 years	10 years
AR(1)	1	**1**	**1**	**1**	1
VAR(1)	1.150	1.221	1.208	1.174	1.157
VAR(1)+FB1(PCA)	1.160	1.179	1.150	1.121	1.106
VAR(1)+FB2(PCA)	1.412	1.394	1.298	1.208	1.183
VAR(1)+FB1(SPCA)	1.150	1.189	1.163	1.125	1.102
VAR(1)+FB2(SPCA)	1.368	1.378	1.297	1.210	1.176
AR(SIC)	**0.917**	1.028	1.001	**1**	1.006
VAR(SIC)	1.150	1.221	1.208	1.174	1.157
VAR(SIC)+FB1(PCA)	1.160	1.179	1.150	1.121	1.106
VAR(SIC)+FB2(PCA)	1.412	1.394	1.298	1.208	1.183
VAR(SIC)+FB1(SPCA)	1.150	1.189	1.163	1.125	1.102
VAR(SIC)+FB2(SPCA)	1.368	1.378	1.297	1.210	1.176
DNS(1)	2.053	1.050	1.134	1.414	0.947
DNS(2)	1.916	1.087	1.212	1.505	0.943
DNS(3)	1.672	1.017	1.101	1.487	0.988
DNS(4)	1.364	1.139	1.126	1.362	1.050
DNS(5)	1.288	1.110	1.149	1.430	1.038
DNS(6)	1.204	1.204	1.123	1.421	1.076
DNS(1)+FB1(PCA)	1.411	1.208	1.069	1.074	0.977
DNS(2)+FB1(PCA)	1.355	1.160	1.053	1.117	0.969
DNS(3)+FB1(PCA)	1.251	1.317	1.085	1.105	**0.933**
DNS(4)+FB1(PCA)	1.228	1.240	1.116	1.181	1.037
DNS(5)+FB1(PCA)	1.174	1.182	1.098	1.227	1.028
DNS(6)+FB1(PCA)	1.115	1.352	1.133	1.225	1.019
DNS(1)+FB2(PCA)	1.785	1.579	1.239	1.157	1.009
DNS(2)+FB2(PCA)	1.730	1.485	1.191	1.191	1.000
DNS(3)+FB2(PCA)	1.708	1.703	1.257	1.192	0.976
DNS(4)+FB2(PCA)	1.493	1.771	1.398	1.233	1.149
DNS(5)+FB2(PCA)	1.470	1.674	1.340	1.263	1.140
DNS(6)+FB2(PCA)	1.510	1.991	1.463	1.270	1.117
DNS(1)+FB1(SPCA)	1.672	1.184	1.101	1.179	0.955
DNS(2)+FB1(SPCA)	1.595	1.163	1.116	1.239	0.948
DNS(3)+FB1(SPCA)	1.429	1.246	1.099	1.218	0.936
DNS(4)+FB1(SPCA)	1.201	1.213	1.107	1.200	1.012
DNS(5)+FB1(SPCA)	1.149	1.159	1.099	1.253	1.002
DNS(6)+FB1(SPCA)	1.086	1.316	1.121	1.247	1.005
DNS(1)+FB2(SPCA)	1.743	1.586	1.253	1.196	1.003
DNS(2)+FB2(SPCA)	1.691	1.493	1.211	1.235	0.995
DNS(3)+FB2(SPCA)	1.675	1.698	1.265	1.238	0.987
DNS(4)+FB2(SPCA)	1.418	1.725	1.382	1.260	1.126
DNS(5)+FB2(SPCA)	1.394	1.630	1.330	1.296	1.117
DNS(6)+FB2(SPCA)	1.434	1.940	1.443	1.300	1.111
DNS(1)+MAC	1.855	1.643	1.329	1.198	0.986
DNS(2)+MAC	1.806	1.558	1.292	1.240	0.973
DNS(3)+MAC	1.785	1.785	1.358	1.234	0.955
DNS(4)+MAC	1.327	1.411	1.262	1.287	1.086
DNS(5)+MAC	1.292	1.353	1.247	1.336	1.072
DNS(6)+MAC	1.246	1.541	1.285	1.326	1.066
DIF(1)	1.127	1.190	1.154	1.117	1.153
DIF(2)	1.117	1.218	1.211	1.196	1.242
DIF(3)	1.126	1.248	1.248	1.232	1.289
DIF(4)	1.160	1.179	1.150	1.121	1.106
DIF(5)	1.412	1.394	1.298	1.208	1.183
DIF(6)	1.485	1.432	1.317	1.202	1.121
DIF(7)	1.150	1.189	1.163	1.125	1.102
DIF(8)	1.368	1.378	1.297	1.210	1.176
DIF(9)	1.385	1.400	1.327	1.238	1.147
DIF(1)+FB1(PCA)	1.111	1.116	1.081	1.085	1.146
DIF(2)+FB1(PCA)	1.109	1.147	1.129	1.140	1.207
DIF(3)+FB1(PCA)	1.139	1.192	1.173	1.195	1.318
DIF(1)+FB2(PCA)	1.359	1.329	1.228	1.160	1.190
DIF(2)+FB2(PCA)	1.358	1.361	1.277	1.217	1.250
DIF(3)+FB2(PCA)	1.382	1.406	1.327	1.275	1.348
DIF(1)+FB1(SPCA)	1.117	1.145	1.110	1.099	1.155
DIF(2)+FB1(SPCA)	1.119	1.182	1.165	1.163	1.225
DIF(3)+FB1(SPCA)	1.153	1.230	1.211	1.214	1.332
DIF(1)+FB2(SPCA)	1.300	1.317	1.239	1.180	1.204
DIF(2)+FB2(SPCA)	1.300	1.353	1.295	1.246	1.276
DIF(3)+FB2(SPCA)	1.336	1.410	1.356	1.314	1.380

4.5 DNS models

Looking at our results from tables B.1 to B.4, we can make the following observations. With a significance of 0.05, we cannot say one formulation performs better than the other, based on subsample 1, see table B.2. The story is different with subsample 2, see table B.3. Here we can see that for the 1 year maturity, the VAR formulation is often better than the AR formulation. The story changes however, when looking at the later maturities, especially the 10 year maturity. Now the tables have turned and the AR formulation is consistently better. For subsample 3, see table B.4, the story is even more polarized than for subsample 2. The 1 year maturity is consistently dominated by the VAR formulation, whilst the 10 year maturity is better forecasted by the AR formulation. The full sample, see table B.1, has much of the same interpretation as subsample 3, with the 1 year maturity being dominated by the VAR formulation and the other maturities better represented by the AR formulation.

4.6 DIF models

Looking at our results from tables B.5 to B.8, we can make the following observations. Again, just like the results form subsample 1 for the DNS models, with a significance of 0.05, we cannot say one formulation performs better than the other, based on subsample 1, see table B.6. The same is true for subsample 2, where again, all values are around 0.5, indicating equal predictive performance, see table B.7. For subsample 3, there are some factor augmented models that perform worse than their non factor augmented counterpart, in the 1 year maturity, see table B.8. For the full sample, see table B.5, the story is again the same as for subsample 3. Some non factor augmented models perform better than their factor augmented counterpart. Still, we cannot reject the null hypothesis for a large part of the models.

4.7 PCA versus SPCA

Looking at our results from tables B.9 to B.12, we can make the following observations. For subsample 1, see table B.10, SPCA is out performing PCA a large number of times, for the 1, 2, 3, and 5 year maturities. For the 10 year maturity, the results are about even. For subsample 2, see table B.11, the result is even more polarized than for subsample 1, with SPCA outperforming PCA also in the 10 year maturities. For subsample 3, see table B.12, the results are not as clear cut as before. We see instances of SPCA grossly outperforming PCA and vice versa. For the maturities 2, 3, and 5 years, we observe that PCA greatly outperforms SPCA. Looking at the full sample, see table B.9, we see again that SPCA is flagrantly outperforming PCA, except some instances in the 5 year maturity.

5. Conclusion

The goal of this paper was to replicate the findings of Swanson and Xiong (2017) and to extend their research by including an additional dimension reduction technique in the form of sparse principal components. The model forecasts differ in predictive power very significantly based on the economic situation of the sample that was used to estimate the model. Simple models appear to perform better in forecast scenarios when both an economic recession and upturn are included in the sample. We believe this to be the affect of sudden changes in interest rates by the Federal Reserve. These changes cannot be captured accurately by our models and introduce errors. It is also apparent that when a sample only includes an economic upturn, as was the case with our second sample, the dynamic Nelson-Siegel models perform best in predictive accuracy. When comparing different DNS model formulations, there was no clear winner for all maturities. We find that VAR formulations perform better for smaller maturities and AR formulations perform better for longer maturities. For diffusion index models, we cannot draw any conclusions from our results. It appeared there is no significant difference between factor augmented and non factor augmented models. For our extension, it is clear that sparse principal component analysis is outperforming standard principal component analysis significantly. This leads us to believe that the increase in parsimony of the underlying factors manages to extract more important information out of the data than standard principal component analysis.

A. Model overview

Table A.1: Models used in forecasting

Panel A: Autoregressive (AR) and Vector-Autoregressive (VAR) Models	
Model	**Description**
AR(1)	Autoregressive model with one lag
VAR(1)	Five-dimensional vector autoregressive model with one lag
VAR(1)+FB1(PCA)	VAR(1) model with one principle component added, principle component analysis based on all 105 macroeconomic variables
VAR(1)+FB2(PCA)	VAR(1) model with two principle components added, principle component analysis based on all 105 macroeconomic variables
VAR(1)+FB1(SPCA)	VAR(1) model with one sparse principle component added, sparse principle component analysis based on all 105 macroeconomic variables
VAR(1)+FB2(SPCA)	VAR(1) model with two sparse principle components added, sparse principle component analysis based on all 105 macroeconomic variables
AR(SIC)	Autoregressive model with lag(s) selected by the Schwarz information criterion
VAR(SIC)	Five-dimensional vector autoregressive model with lag(s) selected by the Schwarz information criterion
VAR(SIC)+FB1(PCA)	VAR(SIC) model with one principle component added, principle component analysis based on all 105 macroeconomic variables
VAR(SIC)+FB2(PCA)	VAR(SIC) model with two principle components added, principle component analysis based on all 105 macroeconomic variables
VAR(SIC)+FB1(SPCA)	VAR(SIC) model with one sparse principle component added, sparse principle component analysis based on all 105 macroeconomic variables
VAR(SIC)+FB2(SPCA)	VAR(SIC) model with two sparse principle components added, sparse principle component analysis based on all 105 macroeconomic variables

Panel B: Dynamic Nelson-Siegel (DNS) Models	
Model	**Description**
DNS(1)	Dynamic Nelson-Siegel (DNS) model with underlying AR(1) factor specifications fitted with ten-dimensional yields: maturity $\tau = 12, 24, 36, 48, 60, 72, 84, 96, 108, 120$ months
DNS(2)	DNS model with underlying AR(1) factor specifications fitted with six-dimensional yields: maturity $\tau = 12, 24, 36, 60, 84, 120$ months
DNS(3)	DNS model with underlying AR(1) factor specifications fitted with four-dimensional yields: maturity $\tau = 12, 36, 60, 120$ months
DNS(4)	DNS model with underlying VAR(1) factor specifications fitted with ten-dimensional yields: maturity $\tau = 12, 24, 36, 48, 60, 72, 84, 96, 108, 120$ months
DNS(5)	DNS model with underlying VAR(1) factor specifications fitted with six-dimensional yields: maturity $\tau = 12, 24, 36, 60, 84, 120$ months
DNS(6)	DNS model with underlying VAR(1) factor specifications fitted with four-dimensional yields: maturity $\tau = 12, 36, 60, 120$ months
DNS(1)+FB1(PCA)	DNS(1) model with one principle component added, principle component analysis based on all 105 macroeconomic variables
DNS(2)+FB1(PCA)	DNS(2) model with one principle component added, principle component analysis based on all 105 macroeconomic variables

Panel B: Dynamic Nelson-Siegel (DNS) Models	
Model	**Description**
DNS(3)+FB1(PCA)	DNS(3) model with one principle component added, principle component analysis based on all 105 macroeconomic variables
DNS(4)+FB1(PCA)	DNS(4) model with one principle component added, principle component analysis based on all 105 macroeconomic variables
DNS(5)+FB1(PCA)	DNS(5) model with one principle component added, principle component analysis based on all 105 macroeconomic variables
DNS(6)+FB1(PCA)	DNS(6) model with one principle component added, principle component analysis based on all 105 macroeconomic variables
DNS(1)+FB2(PCA)	DNS(1) model with two principle components added, principle component analysis based on all 105 macroeconomic variables
DNS(2)+FB2(PCA)	DNS(2) model with two principle components added, principle component analysis based on all 105 macroeconomic variables
DNS(3)+FB2(PCA)	DNS(3) model with two principle components added, principle component analysis based on all 105 macroeconomic variables
DNS(4)+FB2(PCA)	DNS(4) model with two principle components added, principle component analysis based on all 105 macroeconomic variables
DNS(5)+FB2(PCA)	DNS(5) model with two principle components added, principle component analysis based on all 105 macroeconomic variables
DNS(6)+FB2(PCA)	DNS(6) model with two principle components added, principle component analysis based on all 105 macroeconomic variables
DNS(1)+FB1(SPCA)	DNS(1) model with one sparse principle component added, sparse principle component analysis based on all 105 macroeconomic variables
DNS(2)+FB1(SPCA)	DNS(2) model with one sparse principle component added, sparse principle component analysis based on all 105 macroeconomic variables
DNS(3)+FB1(SPCA)	DNS(3) model with one sparse principle component added, sparse principle component analysis based on all 105 macroeconomic variables
DNS(4)+FB1(SPCA)	DNS(4) model with one sparse principle component added, sparse principle component analysis based on all 105 macroeconomic variables
DNS(5)+FB1(SPCA)	DNS(5) model with one sparse principle component added, sparse principle component analysis based on all 105 macroeconomic variables
DNS(6)+FB1(SPCA)	DNS(6) model with one sparse principle component added, sparse principle component analysis based on all 105 macroeconomic variables
DNS(1)+FB2(SPCA)	DNS(1) model with two sparse principle components added, sparse principle component analysis based on all 105 macroeconomic variables
DNS(2)+FB2(SPCA)	DNS(2) model with two sparse principle components added, sparse principle component analysis based on all 105 macroeconomic variables
DNS(3)+FB2(SPCA)	DNS(3) model with two sparse principle components added, sparse principle component analysis based on all 105 macroeconomic variables
DNS(4)+FB2(SPCA)	DNS(4) model with two sparse principle components added, sparse principle component analysis based on all 105 macroeconomic variables
DNS(5)+FB2(SPCA)	DNS(5) model with two sparse principle components added, sparse principle component analysis based on all 105 macroeconomic variables
DNS(6)+FB2(SPCA)	DNS(6) model with two sparse principle components added, sparse principle component analysis based on all 105 macroeconomic variables
DNS(1)+MAC	DNS(1) model with three key macroeconomic variables added: manufacturing capacity utilization, the federal funds rate, and annual price inflation
DNS(2)+MAC	DNS(2) model with three key macroeconomic variables added: manufacturing capacity utilization, the federal funds rate, and annual price inflation
DNS(3)+MAC	DNS(3) model with three key macroeconomic variables added: manufacturing capacity utilization, the federal funds rate, and annual price inflation

Panel B: Dynamic Nelson-Siegel (DNS) Models	
Model	**Description**
DNS(4)+MAC	DNS(4) model with three key macroeconomic variables added: manufacturing capacity utilization, the federal funds rate, and annual price inflation
DNS(5)+MAC	DNS(5) model with three key macroeconomic variables added: manufacturing capacity utilization, the federal funds rate, and annual price inflation
DNS(6)+MAC	DNS(6) model with three key macroeconomic variables added: manufacturing capacity utilization, the federal funds rate, and annual price inflation

Panel C: Diffusion Index Models	
Model	**Description**
DIF(1)	Diffusion index model with one principle component estimator based on all ten-dimensional yields
DIF(2)	Diffusion index model with two principle component estimator based on all ten-dimensional yields
DIF(3)	Diffusion index model with three principle component estimator based on all ten-dimensional yields
DIF(4)	Diffusion index model with one principle component estimator based on all 105 macroeconomic variables
DIF(5)	Diffusion index model with two principle component estimator based on all 105 macroeconomic variables
DIF(6)	Diffusion index model with three principle component estimator based on all 105 macroeconomic variables
DIF(7)	Diffusion index model with one sparse principle component estimator based on all 105 macroeconomic variables
DIF(8)	Diffusion index model with two sparse principle component estimator based on all 105 macroeconomic variables
DIF(9)	Diffusion index model with three sparse principle component estimator based on all 105 macroeconomic variables
DIF(1)+FB1(PCA)	DIF(1) model with one principle component added, principle component analysis based on all 105 macroeconomic variables
DIF(2)+FB1(PCA)	DIF(2) model with one principle component added, principle component analysis based on all 105 macroeconomic variables
DIF(3)+FB1(PCA)	DIF(3) model with one principle component added, principle component analysis based on all 105 macroeconomic variables
DIF(1)+FB2(PCA)	DIF(1) model with two principle components added, principle component analysis based on all 105 macroeconomic variables
DIF(2)+FB2(PCA)	DIF(2) model with two principle components added, principle component analysis based on all 105 macroeconomic variables
DIF(3)+FB2(PCA)	DIF(3) model with two principle components added, principle component analysis based on all 105 macroeconomic variables
DIF(1)+FB1(SPCA)	DIF(1) model with one sparse principle component added, sparse principle component analysis based on all 105 macroeconomic variables
DIF(2)+FB1(SPCA)	DIF(2) model with one sparse principle component added, sparse principle component analysis based on all 105 macroeconomic variables
DIF(3)+FB1(SPCA)	DIF(3) model with one sparse principle component added, sparse principle component analysis based on all 105 macroeconomic variables
DIF(1)+FB2(SPCA)	DIF(1) model with two sparse principle components added, sparse principle component analysis based on all 105 macroeconomic variables
DIF(2)+FB2(SPCA)	DIF(2) model with two sparse principle components added, sparse principle component analysis based on all 105 macroeconomic variables

Panel C: Diffusion Index Models	
Model	**Description**
DIF(3)+FB2(SPCA)	DIF(3) model with two sparse principle components added, sparse principle component analysis based on all 105 macroeconomic variables

B. DM-tests overview

The following tables contain DM-test probabilities when comparing various models for equal predictive accuracy. The DM-test is a two-sides test, meaning when the probabilities are higher than $1-(\alpha/2)$, the latter model has higher predictive accuracy than the former. Correspondingly, if the probability is lower than $\alpha/2$, the former model has higher predictive accuracy than the latter.

B.1 DNS Models

Table B.1: DM-test probabilities between AR and VAR formulated models, respectively (Full sample: 1992:1-2016:12)

Models		DM-test (p)				
	Maturity	1 year	2 years	3 years	5 years	10 years
DNS(1)	DNS(4)	0.998	0.051	0.150	0.423	0.001
DNS(2)	DNS(5)	0.994	0.121	0.293	0.497	0.001
DNS(3)	DNS(6)	0.970	0.008	0.089	0.589	0.003
DNS(1)+FB1(PCA)	DNS(4)+FB1(PCA)	0.810	0.216	0.057	0.003	0.001
DNS(2)+FB1(PCA)	DNS(5)+FB1(PCA)	0.802	0.214	0.057	0.004	0.002
DNS(3)+FB1(PCA)	DNS(6)+FB1(PCA)	0.769	0.202	0.049	0.002	0.000
DNS(1)+FB2(PCA)	DNS(4)+FB2(PCA)	0.965	0.164	0.035	0.064	0.000
DNS(2)+FB2(PCA)	DNS(5)+FB2(PCA)	0.955	0.153	0.040	0.080	0.000
DNS(3)+FB2(PCA)	DNS(6)+FB2(PCA)	0.952	0.104	0.025	0.094	0.001
DNS(1)+FB1(SPCA)	DNS(4)+FB1(SPCA)	0.967	0.161	0.073	0.012	0.002
DNS(2)+FB1(SPCA)	DNS(5)+FB1(SPCA)	0.954	0.200	0.096	0.015	0.002
DNS(3)+FB1(SPCA)	DNS(6)+FB1(SPCA)	0.924	0.109	0.054	0.008	0.000
DNS(1)+FB2(SPCA)	DNS(4)+FB2(SPCA)	0.990	0.272	0.066	0.065	0.000
DNS(2)+FB2(SPCA)	DNS(5)+FB2(SPCA)	0.986	0.257	0.070	0.073	0.000
DNS(3)+FB2(SPCA)	DNS(6)+FB2(SPCA)	0.984	0.171	0.046	0.089	0.000
DNS(1)+MAC	DNS(4)+MAC	0.995	0.989	0.916	0.380	0.051
DNS(2)+MAC	DNS(5)+MAC	0.994	0.980	0.870	0.312	0.049
DNS(3)+MAC	DNS(6)+MAC	0.997	0.989	0.916	0.374	0.046

Table B.2: DM-test probabilities between AR and VAR formulated models, respectively (Subsample 1: 1992:1-1999:12)

Models		DM-test (p)				
	Maturity	1 year	2 years	3 years	5 years	10 years
DNS(1)	DNS(4)	0.115	0.111	0.144	0.125	0.101
DNS(2)	DNS(5)	0.118	0.126	0.146	0.105	0.104
DNS(3)	DNS(6)	0.122	0.098	0.150	0.167	0.075
DNS(1)+FB1(PCA)	DNS(4)+FB1(PCA)	0.557	0.477	0.325	0.189	0.089
DNS(2)+FB1(PCA)	DNS(5)+FB1(PCA)	0.542	0.431	0.286	0.181	0.101
DNS(3)+FB1(PCA)	DNS(6)+FB1(PCA)	0.621	0.545	0.376	0.225	0.120
DNS(1)+FB2(PCA)	DNS(4)+FB2(PCA)	0.620	0.623	0.480	0.303	0.223
DNS(2)+FB2(PCA)	DNS(5)+FB2(PCA)	0.612	0.601	0.448	0.274	0.230
DNS(3)+FB2(PCA)	DNS(6)+FB2(PCA)	0.707	0.718	0.567	0.345	0.244
DNS(1)+FB1(SPCA)	DNS(4)+FB1(SPCA)	0.441	0.338	0.242	0.150	0.110
DNS(2)+FB1(SPCA)	DNS(5)+FB1(SPCA)	0.433	0.310	0.214	0.138	0.119
DNS(3)+FB1(SPCA)	DNS(6)+FB1(SPCA)	0.512	0.426	0.307	0.181	0.123
DNS(1)+FB2(SPCA)	DNS(4)+FB2(SPCA)	0.557	0.519	0.401	0.275	0.193
DNS(2)+FB2(SPCA)	DNS(5)+FB2(SPCA)	0.555	0.513	0.386	0.256	0.202
DNS(3)+FB2(SPCA)	DNS(6)+FB2(SPCA)	0.638	0.625	0.489	0.316	0.218
DNS(1)+MAC	DNS(4)+MAC	0.875	0.956	0.923	0.792	0.578
DNS(2)+MAC	DNS(5)+MAC	0.842	0.938	0.893	0.737	0.560
DNS(3)+MAC	DNS(6)+MAC	0.891	0.956	0.922	0.791	0.586

Table B.3: DM-test probabilities between AR and VAR formulated models, respectively (Subsample 2: 2000:1-2007:12)

Models		DM-test (p)				
	Maturity	1 year	2 years	3 years	5 years	10 years
DNS(1)	DNS(4)	0.959	0.183	0.212	0.422	0.045
DNS(2)	DNS(5)	0.930	0.230	0.268	0.423	0.048
DNS(3)	DNS(6)	0.872	0.092	0.171	0.539	0.082
DNS(1)+FB1(PCA)	DNS(4)+FB1(PCA)	0.579	0.191	0.051	0.015	0.036
DNS(2)+FB1(PCA)	DNS(5)+FB1(PCA)	0.582	0.199	0.063	0.025	0.041
DNS(3)+FB1(PCA)	DNS(6)+FB1(PCA)	0.525	0.148	0.033	0.011	0.042
DNS(1)+FB2(PCA)	DNS(4)+FB2(PCA)	0.778	0.314	0.099	0.182	0.010
DNS(2)+FB2(PCA)	DNS(5)+FB2(PCA)	0.795	0.309	0.119	0.251	0.013
DNS(3)+FB2(PCA)	DNS(6)+FB2(PCA)	0.837	0.268	0.088	0.263	0.047
DNS(1)+FB1(SPCA)	DNS(4)+FB1(SPCA)	0.582	0.190	0.053	0.008	0.056
DNS(2)+FB1(SPCA)	DNS(5)+FB1(SPCA)	0.572	0.195	0.059	0.011	0.059
DNS(3)+FB1(SPCA)	DNS(6)+FB1(SPCA)	0.516	0.151	0.036	0.005	0.034
DNS(1)+FB2(SPCA)	DNS(4)+FB2(SPCA)	0.917	0.518	0.189	0.160	0.021
DNS(2)+FB2(SPCA)	DNS(5)+FB2(SPCA)	0.916	0.494	0.198	0.194	0.025
DNS(3)+FB2(SPCA)	DNS(6)+FB2(SPCA)	0.940	0.474	0.173	0.205	0.049
DNS(1)+MAC	DNS(4)+MAC	0.981	0.926	0.725	0.495	0.101
DNS(2)+MAC	DNS(5)+MAC	0.982	0.916	0.708	0.489	0.115
DNS(3)+MAC	DNS(6)+MAC	0.989	0.923	0.696	0.501	0.158

Table B.4: DM-test probabilities between AR and VAR formulated models, respectively (Subsample 3: 2008:1-2016:12)

Models		DM-test (p)				
	Maturity	1 year	2 years	3 years	5 years	10 years
DNS(1)	DNS(4)	0.999	0.224	0.542	0.821	0.011
DNS(2)	DNS(5)	0.998	0.421	0.782	0.904	0.012
DNS(3)	DNS(6)	0.989	0.053	0.386	0.884	0.026
DNS(1)+FB1(PCA)	DNS(4)+FB1(PCA)	0.878	0.360	0.270	0.049	0.016
DNS(2)+FB1(PCA)	DNS(5)+FB1(PCA)	0.883	0.401	0.282	0.051	0.019
DNS(3)+FB1(PCA)	DNS(6)+FB1(PCA)	0.829	0.341	0.253	0.029	0.002
DNS(1)+FB2(PCA)	DNS(4)+FB2(PCA)	0.982	0.090	0.047	0.121	0.001
DNS(2)+FB2(PCA)	DNS(5)+FB2(PCA)	0.970	0.080	0.049	0.131	0.001
DNS(3)+FB2(PCA)	DNS(6)+FB2(PCA)	0.913	0.034	0.021	0.116	0.001
DNS(1)+FB1(SPCA)	DNS(4)+FB1(SPCA)	0.999	0.366	0.463	0.329	0.007
DNS(2)+FB1(SPCA)	DNS(5)+FB1(SPCA)	0.999	0.518	0.600	0.388	0.010
DNS(3)+FB1(SPCA)	DNS(6)+FB1(SPCA)	0.993	0.192	0.363	0.257	0.001
DNS(1)+FB2(SPCA)	DNS(4)+FB2(SPCA)	0.995	0.151	0.082	0.155	0.001
DNS(2)+FB2(SPCA)	DNS(5)+FB2(SPCA)	0.992	0.140	0.086	0.165	0.000
DNS(3)+FB2(SPCA)	DNS(6)+FB2(SPCA)	0.967	0.052	0.035	0.158	0.001
DNS(1)+MAC	DNS(4)+MAC	0.932	0.861	0.683	0.150	0.045
DNS(2)+MAC	DNS(5)+MAC	0.931	0.832	0.627	0.139	0.044
DNS(3)+MAC	DNS(6)+MAC	0.942	0.867	0.696	0.141	0.032

B.2 DIF Models

Table B.5: DM-test probabilities between Diffusion Index models without and with factor augmentation, respectively (Full sample: 1992:1-2016:12)

Models		DM-test (p)				
	Maturity	1 year	2 years	3 years	5 years	10 years
DIF(1)	DIF(1)+FB1(PCA)	0.076	0.407	0.597	0.709	0.834
DIF(2)	DIF(2)+FB1(PCA)	0.098	0.408	0.606	0.726	0.837
DIF(3)	DIF(3)+FB1(PCA)	0.011	0.153	0.329	0.456	0.521
DIF(1)	DIF(1)+FB2(PCA)	0.090	0.198	0.298	0.435	0.610
DIF(2)	DIF(2)+FB2(PCA)	0.113	0.206	0.303	0.447	0.621
DIF(3)	DIF(3)+FB2(PCA)	0.056	0.112	0.160	0.227	0.387
DIF(1)	DIF(1)+FB1(SPCA)	0.131	0.418	0.582	0.702	0.800
DIF(2)	DIF(2)+FB1(SPCA)	0.153	0.396	0.557	0.676	0.757
DIF(3)	DIF(3)+FB1(SPCA)	0.017	0.132	0.282	0.421	0.475
DIF(1)	DIF(1)+FB2(SPCA)	0.313	0.390	0.435	0.485	0.572
DIF(2)	DIF(2)+FB2(SPCA)	0.318	0.380	0.417	0.455	0.519
DIF(3)	DIF(3)+FB2(SPCA)	0.169	0.205	0.207	0.199	0.318

Table B.6: DM-test probabilities between Diffusion Index models without and with factor augmentation, respectively (Subsample 1: 1992:1-1999:12)

Models		DM-test (p)				
	Maturity	1 year	2 years	3 years	5 years	10 years
DIF(1)	DIF(1)+FB1(PCA)	0.086	0.197	0.337	0.591	0.866
DIF(2)	DIF(2)+FB1(PCA)	0.073	0.177	0.310	0.565	0.856
DIF(3)	DIF(3)+FB1(PCA)	0.069	0.143	0.248	0.487	0.849
DIF(1)	DIF(1)+FB2(PCA)	0.294	0.297	0.335	0.527	0.865
DIF(2)	DIF(2)+FB2(PCA)	0.211	0.230	0.270	0.469	0.851
DIF(3)	DIF(3)+FB2(PCA)	0.162	0.155	0.159	0.284	0.782
DIF(1)	DIF(1)+FB1(SPCA)	0.130	0.251	0.388	0.611	0.836
DIF(2)	DIF(2)+FB1(SPCA)	0.120	0.234	0.367	0.591	0.828
DIF(3)	DIF(3)+FB1(SPCA)	0.123	0.199	0.301	0.519	0.826
DIF(1)	DIF(1)+FB2(SPCA)	0.563	0.520	0.519	0.633	0.857
DIF(2)	DIF(2)+FB2(SPCA)	0.457	0.445	0.454	0.586	0.846
DIF(3)	DIF(3)+FB2(SPCA)	0.368	0.327	0.304	0.408	0.801

Table B.7: DM-test probabilities between Diffusion Index models without and with factor augmentation, respectively (Subsample 2: 2000:1-2007:12)

Models		DM-test (p)				
	Maturity	1 year	2 years	3 years	5 years	10 years
DIF(1)	DIF(1)+FB1(PCA)	0.187	0.368	0.470	0.585	0.785
DIF(2)	DIF(2)+FB1(PCA)	0.306	0.408	0.460	0.509	0.628
DIF(3)	DIF(3)+FB1(PCA)	0.047	0.103	0.156	0.241	0.498
DIF(1)	DIF(1)+FB2(PCA)	0.439	0.509	0.533	0.567	0.673
DIF(2)	DIF(2)+FB2(PCA)	0.541	0.561	0.552	0.529	0.524
DIF(3)	DIF(3)+FB2(PCA)	0.469	0.469	0.438	0.388	0.393
DIF(1)	DIF(1)+FB1(SPCA)	0.270	0.428	0.519	0.628	0.816
DIF(2)	DIF(2)+FB1(SPCA)	0.396	0.463	0.489	0.508	0.578
DIF(3)	DIF(3)+FB1(SPCA)	0.059	0.114	0.157	0.227	0.468
DIF(1)	DIF(1)+FB2(SPCA)	0.580	0.628	0.642	0.658	0.692
DIF(2)	DIF(2)+FB2(SPCA)	0.634	0.655	0.642	0.591	0.477
DIF(3)	DIF(3)+FB2(SPCA)	0.519	0.539	0.515	0.445	0.367

Table B.8: DM-test probabilities between Diffusion Index models without and with factor augmentation, respectively (Subsample 3: 2008:1-2016:12)

Models		DM-test (p)				
	Maturity	1 year	2 years	3 years	5 years	10 years
DIF(1)	DIF(1)+FB1(PCA)	0.743	0.949	0.905	0.703	0.539
DIF(2)	DIF(2)+FB1(PCA)	0.622	0.935	0.923	0.821	0.688
DIF(3)	DIF(3)+FB1(PCA)	0.299	0.825	0.863	0.681	0.422
DIF(1)	DIF(1)+FB2(PCA)	0.036	0.082	0.171	0.256	0.343
DIF(2)	DIF(2)+FB2(PCA)	0.035	0.077	0.196	0.378	0.465
DIF(3)	DIF(3)+FB2(PCA)	0.017	0.077	0.178	0.294	0.348
DIF(1)	DIF(1)+FB1(SPCA)	0.636	0.824	0.782	0.621	0.492
DIF(2)	DIF(2)+FB1(SPCA)	0.468	0.764	0.783	0.712	0.594
DIF(3)	DIF(3)+FB1(SPCA)	0.183	0.616	0.709	0.593	0.389
DIF(1)	DIF(1)+FB2(SPCA)	0.054	0.091	0.120	0.129	0.274
DIF(2)	DIF(2)+FB2(SPCA)	0.049	0.079	0.123	0.180	0.346
DIF(3)	DIF(3)+FB2(SPCA)	0.035	0.077	0.094	0.144	0.276

B.3 PCA versus SPCA

Table B.9: DM-test probabilities between PCA and SPCA factor augmented models, respectively (Full sample: 1992:1-2016:12)

Models		DM-test (p)				
	Maturity	1 year	2 years	3 years	5 years	10 years
VAR(1)+FB1(PCA)	VAR(1)+FB1(SPCA)	0.939	0.557	0.369	0.359	0.392
VAR(1)+FB2(PCA)	VAR(1)+FB2(SPCA)	0.998	0.983	0.929	0.724	0.480
VAR(SIC)+FB1(PCA)	VAR(SIC)+FB1(SPCA)	0.939	0.557	0.369	0.359	0.392
VAR(SIC)+FB2(PCA)	VAR(SIC)+FB2(SPCA)	0.998	0.983	0.929	0.724	0.480
DNS(1)+FB1(PCA)	DNS(1)+FB1(SPCA)	0.045	0.933	0.550	0.030	0.627
DNS(2)+FB1(PCA)	DNS(2)+FB1(SPCA)	0.069	0.852	0.337	0.018	0.634
DNS(3)+FB1(PCA)	DNS(3)+FB1(SPCA)	0.187	0.983	0.653	0.020	0.427
DNS(4)+FB1(PCA)	DNS(4)+FB1(SPCA)	1.000	0.954	0.830	0.313	0.859
DNS(5)+FB1(PCA)	DNS(5)+FB1(SPCA)	0.999	0.947	0.749	0.215	0.866
DNS(6)+FB1(PCA)	DNS(6)+FB1(SPCA)	0.999	0.965	0.862	0.264	0.707
DNS(1)+FB2(PCA)	DNS(1)+FB2(SPCA)	0.988	0.771	0.760	0.660	0.478
DNS(2)+FB2(PCA)	DNS(2)+FB2(SPCA)	0.985	0.794	0.784	0.662	0.483
DNS(3)+FB2(PCA)	DNS(3)+FB2(SPCA)	0.972	0.740	0.762	0.656	0.369
DNS(4)+FB2(PCA)	DNS(4)+FB2(SPCA)	1.000	0.972	0.969	0.797	0.927
DNS(5)+FB2(PCA)	DNS(5)+FB2(SPCA)	1.000	0.977	0.969	0.742	0.925
DNS(6)+FB2(PCA)	DNS(6)+FB2(SPCA)	0.999	0.966	0.971	0.757	0.773
DIF(4)	DIF(7)	0.939	0.557	0.369	0.359	0.392
DIF(5)	DIF(8)	0.998	0.983	0.929	0.724	0.480
DIF(6)	DIF(9)	0.993	0.917	0.679	0.126	0.037
DIF(1)+FB1(PCA)	DIF(1)+FB1(SPCA)	0.925	0.564	0.419	0.415	0.276
DIF(2)+FB1(PCA)	DIF(2)+FB1(SPCA)	0.906	0.476	0.309	0.268	0.128
DIF(3)+FB1(PCA)	DIF(3)+FB1(SPCA)	0.783	0.467	0.380	0.392	0.212
DIF(1)+FB2(PCA)	DIF(1)+FB2(SPCA)	0.999	0.994	0.958	0.703	0.311
DIF(2)+FB2(PCA)	DIF(2)+FB2(SPCA)	0.998	0.988	0.925	0.542	0.133
DIF(3)+FB2(PCA)	DIF(3)+FB2(SPCA)	0.996	0.978	0.863	0.414	0.085

Table B.10: DM-test probabilities between PCA and SPCA factor augmented models, respectively (Subsample 1: 1992:1-1999:12)

Models		DM-test (p)				
	Maturity	1 year	2 years	3 years	5 years	10 years
VAR(1)+FB1(PCA)	VAR(1)+FB1(SPCA)	0.683	0.564	0.481	0.359	0.231
VAR(1)+FB2(PCA)	VAR(1)+FB2(SPCA)	0.983	0.929	0.834	0.578	0.237
VAR(SIC)+FB1(PCA)	VAR(SIC)+FB1(SPCA)	0.683	0.564	0.481	0.359	0.231
VAR(SIC)+FB2(PCA)	VAR(SIC)+FB2(SPCA)	0.983	0.929	0.834	0.578	0.237
DNS(1)+FB1(PCA)	DNS(1)+FB1(SPCA)	0.877	0.714	0.607	0.649	0.110
DNS(2)+FB1(PCA)	DNS(2)+FB1(SPCA)	0.870	0.709	0.623	0.709	0.136
DNS(3)+FB1(PCA)	DNS(3)+FB1(SPCA)	0.824	0.662	0.562	0.669	0.271
DNS(4)+FB1(PCA)	DNS(4)+FB1(SPCA)	0.570	0.278	0.338	0.588	0.167
DNS(5)+FB1(PCA)	DNS(5)+FB1(SPCA)	0.567	0.324	0.397	0.642	0.197
DNS(6)+FB1(PCA)	DNS(6)+FB1(SPCA)	0.438	0.256	0.321	0.627	0.291
DNS(1)+FB2(PCA)	DNS(1)+FB2(SPCA)	0.999	0.970	0.947	0.932	0.390
DNS(2)+FB2(PCA)	DNS(2)+FB2(SPCA)	0.999	0.972	0.953	0.942	0.404
DNS(3)+FB2(PCA)	DNS(3)+FB2(SPCA)	0.998	0.956	0.940	0.946	0.552
DNS(4)+FB2(PCA)	DNS(4)+FB2(SPCA)	0.999	0.907	0.905	0.933	0.327
DNS(5)+FB2(PCA)	DNS(5)+FB2(SPCA)	0.998	0.924	0.927	0.946	0.351
DNS(6)+FB2(PCA)	DNS(6)+FB2(SPCA)	0.996	0.863	0.893	0.951	0.506
DIF(4)	DIF(7)	0.683	0.564	0.481	0.359	0.231
DIF(5)	DIF(8)	0.983	0.929	0.834	0.578	0.237
DIF(6)	DIF(9)	0.863	0.641	0.483	0.228	0.083
DIF(1)+FB1(PCA)	DIF(1)+FB1(SPCA)	0.866	0.775	0.704	0.578	0.384
DIF(2)+FB1(PCA)	DIF(2)+FB1(SPCA)	0.880	0.789	0.722	0.605	0.420
DIF(3)+FB1(PCA)	DIF(3)+FB1(SPCA)	0.882	0.772	0.698	0.595	0.454
DIF(1)+FB2(PCA)	DIF(1)+FB2(SPCA)	0.997	0.985	0.959	0.850	0.537
DIF(2)+FB2(PCA)	DIF(2)+FB2(SPCA)	0.997	0.985	0.960	0.860	0.577
DIF(3)+FB2(PCA)	DIF(3)+FB2(SPCA)	0.998	0.991	0.970	0.881	0.637

Table B.11: DM-test probabilities between PCA and SPCA factor augmented models, respectively (Subsample 2: 2000:1-2007:12)

Models		DM-test (p)				
	Maturity	1 year	2 years	3 years	5 years	10 years
VAR(1)+FB1(PCA)	VAR(1)+FB1(SPCA)	0.969	0.785	0.575	0.479	0.424
VAR(1)+FB2(PCA)	VAR(1)+FB2(SPCA)	0.933	0.929	0.919	0.842	0.501
VAR(SIC)+FB1(PCA)	VAR(SIC)+FB1(SPCA)	0.969	0.785	0.575	0.479	0.424
VAR(SIC)+FB2(PCA)	VAR(SIC)+FB2(SPCA)	0.933	0.929	0.919	0.842	0.501
DNS(1)+FB1(PCA)	DNS(1)+FB1(SPCA)	0.995	0.981	0.975	0.961	0.785
DNS(2)+FB1(PCA)	DNS(2)+FB1(SPCA)	0.995	0.982	0.979	0.970	0.782
DNS(3)+FB1(PCA)	DNS(3)+FB1(SPCA)	0.994	0.976	0.975	0.971	0.875
DNS(4)+FB1(PCA)	DNS(4)+FB1(SPCA)	0.999	0.995	0.992	0.587	0.959
DNS(5)+FB1(PCA)	DNS(5)+FB1(SPCA)	0.999	0.996	0.986	0.389	0.942
DNS(6)+FB1(PCA)	DNS(6)+FB1(SPCA)	0.998	0.995	0.993	0.510	0.710
DNS(1)+FB2(PCA)	DNS(1)+FB2(SPCA)	0.860	0.514	0.601	0.866	0.347
DNS(2)+FB2(PCA)	DNS(2)+FB2(SPCA)	0.844	0.562	0.673	0.903	0.364
DNS(3)+FB2(PCA)	DNS(3)+FB2(SPCA)	0.790	0.430	0.564	0.898	0.581
DNS(4)+FB2(PCA)	DNS(4)+FB2(SPCA)	0.981	0.859	0.909	0.902	0.944
DNS(5)+FB2(PCA)	DNS(5)+FB2(SPCA)	0.974	0.873	0.916	0.884	0.936
DNS(6)+FB2(PCA)	DNS(6)+FB2(SPCA)	0.964	0.839	0.908	0.883	0.863
DIF(4)	DIF(7)	0.969	0.785	0.575	0.479	0.424
DIF(5)	DIF(8)	0.933	0.929	0.919	0.842	0.501
DIF(6)	DIF(9)	0.804	0.838	0.864	0.875	0.469
DIF(1)+FB1(PCA)	DIF(1)+FB1(SPCA)	0.951	0.851	0.734	0.619	0.454
DIF(2)+FB1(PCA)	DIF(2)+FB1(SPCA)	0.927	0.815	0.661	0.487	0.251
DIF(3)+FB1(PCA)	DIF(3)+FB1(SPCA)	0.758	0.850	0.798	0.662	0.382
DIF(1)+FB2(PCA)	DIF(1)+FB2(SPCA)	0.961	0.962	0.945	0.883	0.519
DIF(2)+FB2(PCA)	DIF(2)+FB2(SPCA)	0.887	0.931	0.920	0.803	0.295
DIF(3)+FB2(PCA)	DIF(3)+FB2(SPCA)	0.764	0.888	0.900	0.811	0.406

Table B.12: DM-test probabilities between PCA and SPCA factor augmented models, respectively (Subsample 3: 2008:1-2016:12)

Models		DM-test (p)				
	Maturity	1 year	2 years	3 years	5 years	10 years
VAR(1)+FB1(PCA)	VAR(1)+FB1(SPCA)	0.772	0.336	0.291	0.426	0.598
VAR(1)+FB2(PCA)	VAR(1)+FB2(SPCA)	0.934	0.733	0.522	0.459	0.631
VAR(SIC)+FB1(PCA)	VAR(SIC)+FB1(SPCA)	0.772	0.336	0.291	0.426	0.598
VAR(SIC)+FB2(PCA)	VAR(SIC)+FB2(SPCA)	0.934	0.733	0.522	0.459	0.631
DNS(1)+FB1(PCA)	DNS(1)+FB1(SPCA)	0.000	0.718	0.210	0.002	0.808
DNS(2)+FB1(PCA)	DNS(2)+FB1(SPCA)	0.000	0.468	0.061	0.001	0.792
DNS(3)+FB1(PCA)	DNS(3)+FB1(SPCA)	0.000	0.955	0.360	0.001	0.453
DNS(4)+FB1(PCA)	DNS(4)+FB1(SPCA)	0.976	0.832	0.616	0.228	0.893
DNS(5)+FB1(PCA)	DNS(5)+FB1(SPCA)	0.965	0.793	0.495	0.159	0.897
DNS(6)+FB1(PCA)	DNS(6)+FB1(SPCA)	0.978	0.873	0.663	0.196	0.765
DNS(1)+FB2(PCA)	DNS(1)+FB2(SPCA)	0.734	0.434	0.279	0.021	0.640
DNS(2)+FB2(PCA)	DNS(2)+FB2(SPCA)	0.723	0.412	0.208	0.011	0.626
DNS(3)+FB2(PCA)	DNS(3)+FB2(SPCA)	0.697	0.551	0.370	0.012	0.262
DNS(4)+FB2(PCA)	DNS(4)+FB2(SPCA)	0.968	0.907	0.743	0.088	0.885
DNS(5)+FB2(PCA)	DNS(5)+FB2(SPCA)	0.966	0.896	0.656	0.052	0.885
DNS(6)+FB2(PCA)	DNS(6)+FB2(SPCA)	0.963	0.921	0.785	0.062	0.640
DIF(4)	DIF(7)	0.772	0.336	0.291	0.426	0.598
DIF(5)	DIF(8)	0.934	0.733	0.522	0.459	0.631
DIF(6)	DIF(9)	0.993	0.874	0.322	0.032	0.111
DIF(1)+FB1(PCA)	DIF(1)+FB1(SPCA)	0.279	0.052	0.083	0.237	0.301
DIF(2)+FB1(PCA)	DIF(2)+FB1(SPCA)	0.180	0.030	0.048	0.141	0.167
DIF(3)+FB1(PCA)	DIF(3)+FB1(SPCA)	0.159	0.048	0.066	0.183	0.208
DIF(1)+FB2(PCA)	DIF(1)+FB2(SPCA)	0.902	0.658	0.307	0.187	0.258
DIF(2)+FB2(PCA)	DIF(2)+FB2(SPCA)	0.902	0.602	0.216	0.100	0.123
DIF(3)+FB2(PCA)	DIF(3)+FB2(SPCA)	0.877	0.436	0.048	0.025	0.030

Bibliography

Akaike, H. (1974). A new look at the statistical model identification. In E. Parzen, K. Tanabe & G. Kitagawa (Eds.), *Selected Papers of Hirotugu Akaike* (pp. 215–222). Springer New York. https://doi.org/10.1007/978-1-4612-1694-0_16. (Cit. on p. 6)

Bai, J. & Ng, S. (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics*, *146*(2), 304–317. https://doi.org/10.1016/j.jeconom.2008.08.010 (cit. on p. 2)

Bai, J. & Ng, S. (2009). Boosting diffusion indices. *Journal of Applied Econometrics*, *24*(4), 607–629. https://doi.org/10.1002/jae.1063 (cit. on p. 2)

Bell, J. B., Tikhonov, A. N. & Arsenin, V. Y. (1978). Solutions of ill-posed problems. *Mathematics of Computation*, *32*(144), 1320. https://doi.org/10.2307/2006360 (cit. on pp. 2, 5)

Bhageshpur, K. (2019). Council post: Data is the new oil – and that's a good thing. Retrieved June 21, 2020, from https://www.forbes.com/sites/forbestechcouncil/2019/11/15/data-is-the-new-oil-and-thats-a-good-thing/. (Cit. on p. 2)

Brandt, M. & Yaron, A. (2003). *Time-consistent no-arbitrage models of the term structure* (tech. rep. w9458). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w9458. (Cit. on p. 2)

Christensen, J. H. E., Diebold, F. X. & Rudebusch, G. D. (2009). An arbitrage-free generalized Nelson–Siegel term structure model. *The Econometrics Journal*, *12*(3), C33–C64. https://doi.org/10.1111/j.1368-423X.2008.00267.x (cit. on p. 2)

Christensen, J. H., Diebold, F. X. & Rudebusch, G. D. (2011). The affine arbitrage-free class of Nelson–Siegel term structure models. *Journal of Econometrics*, *164*(1), 4–20. https://doi.org/10.1016/j.jeconom.2011.02.011 (cit. on p. 2)

Cox, J. C., Ingersoll, J. E. & Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, *53*(2), 385. https://doi.org/10.2307/1911242 (cit. on p. 2)

Dai, Q. & Singleton, K. J. (2000). Specification analysis of affine term structure models. *The Journal of Finance*, *55*(5), 1943–1978. https://doi.org/10.1111/0022-1082.00278 (cit. on p. 2)

de Jong, F. (2000). Time series and cross-section information in affine term-structure models. *Journal of Business & Economic Statistics*, *18*(3), 300. https://doi.org/10.2307/1392263 (cit. on p. 2)

de Jong, F. & Santa-Clara, P. (1999). The dynamics of the forward interest rate curve: A formulation with state variables. *The Journal of Financial and Quantitative Analysis*, *34*(1), 131. https://doi.org/10.2307/2676249 (cit. on p. 2)

Diebold, F. & Mariano, R. (1994). *Comparing predictive accuracy* (tech. rep. t0169). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/t0169. (Cit. on pp. 4, 9)

Diebold, F. X. & Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics*, *130*(2), 337–364. https://doi.org/10.1016/j.jeconom.2005.03.005 (cit. on pp. 2, 7)

Duffee, G. R. (2002). Term premia and interest rate forecasts in affine models. *The Journal of Finance*, *57*(1), 405–443. https://doi.org/10.1111/1540-6261.00426 (cit. on p. 2)

Duffie, D. & Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance*, *6*(4), 379–406. https://doi.org/10.1111/j.1467-9965.1996.tb00123.x (cit. on p. 2)

Gürkaynak, R. S., Sack, B. & Wright, J. H. (2007). The U.S. Treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, *54*(8), 2291–2304. https://doi.org/10.1016/j.jmoneco.2007.06.029 (cit. on p. 10)

Heath, D., Jarrow, R. & Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, *60*(1), 77. https://doi.org/10.2307/2951677 (cit. on p. 2)

Hirano, K. & Wright, J. H. (2017). Forecasting with model uncertainty: Representations and risk reduction. *Econometrica*, *85*(2), 617–643. https://doi.org/10.3982/ECTA13372 (cit. on p. 2)

Hull, J. & White, A. (1990). Pricing interest-rate-derivative securities. *Review of Financial Studies*, *3*(4), 573–592. https://doi.org/10.1093/rfs/3.4.573 (cit. on p. 2)

Kim, H. H. & Swanson, N. R. (2014). Forecasting financial and macroeconomic variables using data reduction methods: New empirical evidence. *Journal of Econometrics*, *178*, 352–367. https://doi.org/10.1016/j.jeconom.2013.08.033 (cit. on p. 2)

Kim, H. H. & Swanson, N. R. (2018). Mining big data using parsimonious factor, machine learning, variable selection and shrinkage methods. *International Journal of Forecasting*, *34*(2), 339–354. https://doi.org/10.1016/j.ijforecast.2016.02.012 (cit. on p. 2)

McCracken, M. W. & Ng, S. (2016). Fred-md: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, *34*(4), 574–589. https://doi.org/10.1080/07350015.2015.1086655 (cit. on p. 10)

Moore, G. E. (2006). Cramming more components onto integrated circuits, Reprinted from Electronics, volume 38, number 8, April 19, 1965, pp.114 ff. *IEEE Solid-State Circuits Society Newsletter*, *11*(3), 33–35. https://doi.org/10.1109/N-SSC.2006.4785860 (cit. on p. 2)

Nelson, C. R. & Siegel, A. F. (1987). Parsimonious modeling of yield curves. *The Journal of Business*, *60*(4), 473. https://doi.org/10.1086/296409 (cit. on p. 7)

Schaller, R. (1997). Moore's law: Past, present and future. *IEEE Spectrum*, *34*(6), 52–59. https://doi.org/10.1109/6.591665 (cit. on p. 2)

Schumacher, C. (2010). Factor forecasting using international targeted predictors: The case of German GDP. *Economics Letters*, *107*(2), 95–98. https://doi.org/10.1016/j.econlet.2009.12.036 (cit. on p. 2)

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136 (cit. on p. 6)

Stock, J. H. & Watson, M. W. (2002a). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, *97*(460), 1167–1179. https://doi.org/10.1198/016214502388618960 (cit. on p. 8)

Stock, J. H. & Watson, M. W. (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, *20*(2), 147–162. https://doi.org/10.1198/073500102317351921 (cit. on p. 8)

Stock, J. H. & Watson, M. W. (2012). Generalized shrinkage methods for forecasting using many predictors. *Journal of Business & Economic Statistics*, *30*(4), 481–493. https://doi.org/10.1080/07350015.2012.715956 (cit. on p. 2)

Swanson, N. R. & Xiong, W. (2017). Big data analytics in economics: What have we learned so far, and where should we go from here? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2998299 (cit. on pp. 2, 4, 10, 12, 21)

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x (cit. on pp. 2, 5)

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, *5*(2), 177–188. https://doi.org/10.1016/0304-405X(77)90016-2 (cit. on p. 2)

Waldrop, M. M. (2016). The chips are down for Moore's law. *Nature*, *530*(7589), 144–147. https://doi.org/10.1038/530144a (cit. on p. 2)

Zou, H. & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *67*(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x (cit. on pp. 2, 5)

Zou, H., Hastie, T. & Tibshirani, R. (2006). Sparse principal component analysis. *Journal of Computational and Graphical Statistics*, *15*(2), 265–286. https://doi.org/10.1198/106186006X113430 (cit. on pp. 2, 5)