DEPARTMENT OF ECONOMICS AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

CREATES
Center for Research in Econometric Analysis of Time Series

The Forecasting Power of the Yield Curve, a Supervised Factor Model Approach

Lorenzo Boldrini and Eric Hillebrand

CREATES Research Paper 2015-39

Department of Economics and Business Economics
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

The Forecasting Power of the Yield Curve, a Supervised Factor Model Approach

Lorenzo Boldrini* and Eric T. Hillebrand†

August 24, 2015

Abstract

We study the forecast power of the yield curve for macroeconomic time series, such as consumer price index, personal consumption expenditures, producer price index, real disposable income, unemployment rate, and industrial production. We employ a state-space model in which the forecasting objective is included in the state vector. This amounts to an augmented dynamic factor model in which the factors (level, slope, and curvature of the yield curve) are supervised for the macroeconomic forecast target. In other words, the factors are informed about the dynamics of the forecast objective. The factor loadings have the Nelson and Siegel (1987) structure and we consider one forecast target at a time. We compare the forecasting performance of our specification to benchmark models such as principal components regression, partial least squares, and ARMA(p,q) processes. We use the yield curve data from Gürkaynak, Sack, and Wright (2006) and Diebold and Li (2006) and macroeconomic data from FRED. We compare the models by means of the conditional predictive ability test of Giacomini and White (2006). We find that the yield curve has more forecast power for real variables compared to inflation measures and that supervising the factor extraction for the forecast target can improve forecast performance.

Keywords: state-space system, Kalman filter, factor model, supervision, forecasting, yield curve.
JEL classification: C32, C38, E43.

---

*Aarhus University, Department of Economics and Business Economics, CREATES - Center for Research in Econometric Analysis of Time Series, Fuglesangs Allé 4, 8210 Aarhus V, Denmark. Email: lboldrini@econ.au.dk. The author acknowledges support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation.

†Aarhus University, Department of Economics and Business Economics, CREATES - Center for Research in Econometric Analysis of Time Series, Fuglesangs Allé 4, 8210 Aarhus V, Denmark. Email: ehillebrand@econ.au.dk. The author acknowledges support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation.

1 Introduction

The forecasting power of the yield curve for macroeconomic variables has been documented in many papers, see among others Harvey (1988), Stock and Watson (1989), Estrella and Hardouvelis (1991), and Chinn and Kucko (2010). However, the predictive power of the yield curve has changed through the years, see for instance Giacomini and Rossi (2006), Rudebusch and Williams (2009), and Stock and Watson (1999a) raising doubt about its reliability as a predictor.

The aim of the present paper is twofold. The first one is to analyse the forecast power of the yield curve for macroeconomic variables. This is carried out by comparing forecasting models that make use of the yield curve information to ones that do not. We also study the stability of the yield curve as a predictor, by considering different time spans. The second objective is to assess the forecast performance of a supervised factor model, as proposed in Boldrini and Hillebrand (2015). This model is a particular specification of a factor model in which the factors are extracted conditionally on the forecast target.

There exists an extensive literature on the forecasting power of the yield curve. One of the first works testifying the forecasting power of the yield curve for macroeconomic variables was Harvey (1988), who within the framework of the consumption based asset pricing model found that the real term structure of interest rates is a good predictor for consumption growth. Within the framework of dynamic factor models, Stock and Watson (1989) find that two interest rate spreads, namely the difference between the six-month commercial paper and the six-month Treasury bill rates, and the difference between the ten-year and one-year Treasury bond rates, are good predictors of real activity. In related papers, Bernanke (1990) and Friedman and Kuttner (1993), using linear regressions, find that the spread between the commercial paper rate and the Treasury bill rate is a particularly good predictor for real activity indicators and inflation. In Estrella and Hardouvelis (1991) the authors conclude that a positive slope of the yield curve is associated with a future increase in real economic activity. They find that it outperforms both in-sample and out-of-sample other variables, such as the index of leading indicators, real short-term interest rates, lagged growth in economic activity, and lagged rates of inflation as well as survey forecasts. Still in the framework of linear regressions Kozicki (1997) and Hamilton and Kim (2000) confirm the predictive power of the spread for real growth and inflation. Ang et al. (2006) build a dynamic model for GDP growth and yields that does not allow arbitrage and completely characterizes expectations of GDP. Contrary to previous findings, they find that the short rate has more predictive power than any term spread.

Diebold and Li (2006) focus on forecasting the yield curve by means of the Nelson and Siegel (1987) (NS) model. They interpret the dynamically moving parameters as level, slope, and curvature. Diebold et al. (2006) cast the NS model in state-space form and analyse the correlations between the extracted factors and macroeconomic variables. They find a strong correlation between the level factor and inflation and between the slope factor and capacity utilization.

Giacomini and Rossi (2006) examine the stability of the forecasting power of the yield curve for economic growth in the US economy, using forecast breakdown tests. They find a forecast breakdown during the periods 1974-76 and in 1979-87, the Burns-Miller and the Volcker monetary regimes respectively, whereas during 1987-2006, corresponding to when Alan Greenspan was chairman of the FED, the yield curve proved to be a more reliable forecaster for real growth. Similarly, Stock and Watson (1999a) found some evidence of structural breaks in the relationship between the slope of the yield curve and real activity during the past years.

The ability to extract information from large datasets has made factor models an appealing tool in forecasting. Stock and Watson (1999b) and Stock and Watson (2002a), for instance, investigate forecasts of output growth and inflation using a large number of economic indicators, including many interest rates and yield spreads. The advantage of factor models is that the information contained in a (potentially) large number of predictors can be summarized in a few factors. Comprehensive surveys on factor models can be found in Bai and Ng (2008b), Breitung and Eickmeier (2006), and Stock and Watson (2011).

In the standard approach to factor models, the extracted factors are the same for all the forecast targets. One of the directions the literature has taken for improving on this approach is to select factors based on their ability to forecast a specific target. Different methods have been proposed in the literature that address this problem. The method of partial least squares regression (PLSR), for instance, constructs a set of linear combinations of the inputs (predictors and forecast target) for regression, for more details see for instance Friedman et al. (2001). Bai and Ng (2008a) proposed performing PCA on a subset of the original predictors, selected using thresholding rules. This approach is close to the supervised PCA method proposed in Bair et al. (2006), that aims at finding linear combinations of the predictors that have high correlation with the target. In particular, first a subset of the predictors is selected, based on the correlation with the target (i.e. the regression coefficient exceeds a given threshold), then PCA is applied on the resulting subset of variables. Bai and Ng (2009) consider 'boosting' (a procedure that performs subset variable selection and coefficient shrinkage) as a methodology for selecting the predictors in factor-augmented autoregressions. Finally, Giovannelli and Proietti (2014) propose an operational supervised method that selects factors based on their significance in the regression of the forecast target on the predictors.

Hillebrand et al. (2012) propose a method to exploit the yield curve information in forecasting macroeconomic variables. The model is a modified NS factor model, where the new NS yield curve factors are supervised for a specific variable to forecast. They show that it outperforms the conventional (non-supervised) NS factor model in out-of-sample forecasting of monthly US output growth and inflation.

In this paper we assess the forecast performance of the yield curve, using a supervised factor model, as presented in Boldrini and Hillebrand (2015). In the supervised framework, the factors are informed of the forecast target (supervised) and the model has a linear, state-space representation to which Kalman filtering techniques apply.
In particular, we select the Nelson and Siegel (1987) factor structure for the yield curve. The latent factors in this specification are three and represent the level, slope, and curvature of the yield curve. We consider also time variation in the factor loadings using a specification similar to the one used in Koopman et al. (2010). We include one forecast target at a time in the state vector, together with the three NS factors. This allows us to estimate the latent factors using information also in the forecast target, through the Kalman filter recursions. The factor extraction is thus conditional on the forecast target.

We compare the forecasting performance of the proposed specification to that of models that make use of the yield curve information (principal components regression, partial least squares, two-step forecasting procedures as in Stock and Watson (2002a)) and models that do not ($AR(p)$ and $MA(q)$ processes). We use a rolling windows scheme and consider both direct and indirect h-step ahead forecasts. We compare the forecast performance of the different models by means of the Giacomini and White (2006) test, considering different time spans. We use yield curve data from Gürkaynak et al. (2007) and Diebold et al. (2006), and macroeconomic variables from FRED. All the data is relative to the US economy. The selected forecast objectives are consumer price index (CPI), personal consumption expenditures (PCE), producer price index (PPI), real disposable income (RDI), unemployment rate (UR), and industrial production (IP).

The paper is organized as follows: in Section 2 we introduce the supervised factor model and relate it with other forecasting methods based on factor models; in Section 3 we provide some details on the computational aspects of the analysis; in Sections 4 we describe the empirical application; finally, Section 5 concludes.

2 Dynamic factor models and supervision

Let x_t be the forecast objective, $\mathbf{y}_t = [y_t^1, \ldots, y_t^N]$ an N-dimensional vector of predictors, h the forecast lead, and T the last available observation in the estimation window. Throughout we indicate with a "caret", estimates of scalars, vectors, or matrices.

2.1 Supervised factor model

In this section we present the supervised factor model as in Boldrini and Hillebrand (2015), that we use to assess the forecasting power of the yield curve. Consider the system

$$\begin{aligned} \begin{bmatrix} \mathbf{y}_t \\ x_t \end{bmatrix} &= \begin{bmatrix} \boldsymbol{\Lambda} & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{f}_t \\ x_t \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_t \\ 0 \end{bmatrix}, \qquad \boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \mathbf{H}), \\ \begin{bmatrix} \mathbf{f}_{t+1} \\ x_{t+1} \end{bmatrix} &= \mathbf{c} + \mathbf{T} \begin{bmatrix} \mathbf{f}_t \\ x_t \end{bmatrix} + \boldsymbol{\eta}_t, \qquad \boldsymbol{\eta}_t \sim N(\mathbf{0}, \mathbf{Q}), \end{aligned} \tag{1}$$

where $\mathbf{f}_t \in \mathbb{R}^k$ are latent factors, $\boldsymbol{\Lambda}$ is a matrix of factor loadings, $\mathbf{T}$ and $\mathbf{c}$ are a matrix and a vector of coefficients, respectively, of suitable dimensions, $\boldsymbol{\epsilon}_t \in \mathbb{R}^N$ and $\boldsymbol{\eta}_t \in \mathbb{R}^{k+1}$ are vectors of disturbances and $\mathbf{H}$ and $\mathbf{Q}$ are their respective variance-covariance matrices.

In this supervised framework, the forecast objective is placed in the state equation together with the latent factors and the predictors are modelled in the measurement equation. The intuition behind the model is that if the forecast objective is correlated with the factors then their estimation by means of the Kalman filter will benefit from the inclusion of the forecast target in the state vector. This follows because for a general linear state-space system, the Kalman filter delivers the best linear predictions of the latent states at time t, given the information from all the observables entering the measurement equation up to and including time t. In the particular case of Gaussian innovations, as in the system (1), the best linear prediction coincides with the conditional expectation. For more details on the optimality properties of the Kalman filter see for instance Brockwell and Davis (2009).

The forecasting scheme for this model is the following:

(i) the system parameters are estimated by maximizing the likelihood function, delivered by the Kalman filter recursions;

(ii) given the parameter estimates, the Kalman filter is run on the data;

(iii) indicating the state vector with $\boldsymbol{\alpha}_t = [\mathbf{f}_t', x_t]$, the forecast $\hat{x}_{T+h}$ is then

$$\hat{x}_{T+h} = [\mathbf{0}', 1] \left[\hat{\mathbf{T}}^h \hat{\boldsymbol{\alpha}}_T + \sum_{i=0}^{h-1} \hat{\mathbf{T}}^i \hat{\mathbf{c}} \right],$$

where $\hat{\boldsymbol{\alpha}}_T$ represents the filtered $\boldsymbol{\alpha}_T$.

Note that the last element of the filtered state vector at time t corresponds to x_t.

To be able to compute direct h-period ahead forecasts, the state-space form (1) can be modified extending the state vector by including $h-1$ lags of the factors. The system (1) thus becomes

$$\begin{aligned}
\begin{bmatrix} \mathbf{y}_t \\ x_t \end{bmatrix} &= \begin{bmatrix} \boldsymbol{\Lambda}^{(N)\times(K-1)} & \mathbf{0}^{N\times 1} & \mathbf{0}^{N\times K(h-1)} \\ \mathbf{0}^{1\times(K-1)} & 1 & \mathbf{0}^{1\times K(h-1)} \end{bmatrix} \begin{bmatrix} \mathbf{f}_t \\ x_t \\ \vdots \\ \mathbf{f}_{t-(h-1)} \\ x_{t-(h-1)} \end{bmatrix} + \begin{bmatrix} \boldsymbol{\epsilon}_t \\ 0 \end{bmatrix}, \\
\begin{bmatrix} \mathbf{f}_{t+1} \\ x_{t+1} \\ \vdots \\ \mathbf{f}_{t-h} \\ x_{t-h} \end{bmatrix} &= \begin{bmatrix} \mathbf{c}^f \\ \mathbf{0}^{K\times(h-1)} \end{bmatrix} + \begin{bmatrix} \mathbf{0}^{K\times(h-1)} & \mathbf{T}^f \\ \mathbf{I}^{(h-1)\times(h-1)} & \mathbf{0}^{(h-1)\times K} \end{bmatrix} \begin{bmatrix} \mathbf{f}_t \\ x_t \\ \vdots \\ \mathbf{f}_{t-(h-1)} \\ x_{t-(h-1)} \end{bmatrix} + \begin{bmatrix} \mathbf{I}^{K\times K} \\ \mathbf{0}^{(h-1)\times K} \end{bmatrix} \boldsymbol{\eta}_t,
\end{aligned} \tag{2}$$

where $\mathbf{T}^f$ is a matrix of parameters relating the state vector at times t and $t+h$, $\mathbf{c}^f$ is a vector of parameters, and $\boldsymbol{\eta}_t$ is the same as in equation (1). The state vector can be recognized to be a $VAR(1)$ representation of a restricted $VAR(h-1)$. For the representation of $VAR(p)$ processes in state-space form see for instance Aoki (1990). The forecasting scheme for the direct approach is analogous to the previous one.

2.2 Two-step procedure

Forecasting using dynamic factor models (DFM hereafter) is often carried out in a two-step procedure as in Stock and Watson (2002a). Consider the model

$$x_{t+h} = \boldsymbol{\beta}(L)'\mathbf{f}_t + \gamma(L)x_t + \epsilon_{t+h}, \tag{3}$$

$$y_{t,i} = \boldsymbol{\lambda}_i(L)\mathbf{f}_t + \eta_{t,i}, \tag{4}$$

with $i = 1, \ldots, N$ and where $\mathbf{f}_t = (f_{t,1}, \ldots, f_{t,k})$ are k latent factors, $\boldsymbol{\eta}_t = [\eta_{t,i}, \ldots, \eta_{t,N}]'$ and ϵ_t are idiosyncratic disturbances, $\boldsymbol{\beta}(L) = \sum_{j=0}^{q} \boldsymbol{\beta}_{j+1} L^j$, $\boldsymbol{\lambda}_i(L) = \sum_{j=0}^{p} \lambda_{i(j+1)} L^j$, and $\gamma(L) = \sum_{j=0}^{s} \gamma_{j+1} L^j$ are finite lag polynomials in the lag operator L; $\boldsymbol{\beta}_j \in \mathbb{R}^k$, $\gamma_j \in \mathbb{R}$, and $\lambda_{ij} \in \mathbb{R}$ are parameters and $q, p, s \in \mathbb{N}_0$ are indices. The assumption on the finiteness of the lag polynomials allows us to rewrite (3)-(4) as a static factor model, i.e. a factor model in which the factors do not appear in lags:

$$\begin{aligned} x_{t+h} &= c + \boldsymbol{\beta}'\mathbf{F}_t + \gamma(L)x_t + \epsilon_{t+h}, \\ \mathbf{y}_t &= \boldsymbol{\Lambda}\mathbf{F}_t + \boldsymbol{\eta}_t, \end{aligned} \tag{5}$$

with $\mathbf{F}_t = [\mathbf{f}_t', \ldots, \mathbf{f}_{t-r}']'$, $r = \max(q,p)$, the i-th row of $\boldsymbol{\Lambda}$ is $[\lambda_{i,1}, \ldots, \lambda_{i,r+1}]$, and $\boldsymbol{\beta} = [\boldsymbol{\beta}_1', \ldots, \boldsymbol{\beta}_{r+1}']'$. The forecasting scheme is the following:

(i) extraction of the factors $\mathbf{f}_t$ from the predictors $\mathbf{x}_t$ modelled in equation (4) using either principal components, as in Stock and Watson (2002b), or the Kalman filter;

(ii) regression of the forecast objective on the lagged estimated factors and on its lags according to the forecasting equation (3) with $t = 1, ..., T-h$;

(iii) the forecast is obtained from the estimated factors and regression coefficients as

$$\hat{x}_{T+h} = \hat{c} + \hat{\boldsymbol{\beta}}'\hat{\mathbf{F}}_t + \hat{\gamma}(L)x_T. \tag{6}$$

Stock and Watson (2002a) developed theoretical results for this two-step procedure, in the case of principal components estimation. In particular, they show the asymptotic efficiency of the feasible forecasts and the consistency of the factor estimates.

In our empirical application we estimate this models using the NS factor loadings and extract the factors using the Kalman filter, we then use an auxiliary equation to compute the forecasts.

The difference between the supervised model (1) and model (5) is that in the latter the extraction of the factors is independent of the forecast objective whereas in the former one the extracted factors are informed (supervised) of the specific forecast target considered.

2.3 Nelson-Siegel factor model

We consider here the forecast of macroeconomic variables using the term structure of interest rates as predictor. In the 1990s, factor models have gained popularity in modelling the yield curve, for instance with the works of Litterman and Scheinkman (1991) and Knez et al. (1994), who used factor analysis to extract common features from yield curves in different countries and periods. They concluded that three factors explained the greater part of the variation in the yield curve. Nelson and Siegel (1987) proposed a way to model the yield curve based on three functions describing level, slope, and curvature, of the term structure of interest rates.

Following Diebold et al. (2006) we cast the Nelson and Siegel (1987) term structure model in state-space form and extract three factors whose interpretation is that of level, slope, and curvature (in the following they are labelled LV_t, SL_t, and CV_t respectively). The latent factors corresponding to level, slope, and curvature are unrestricted whereas the factor loadings are restricted to have the Nelson-Siegel structure. This guarantees positive forward rates at all horizons and a discount factor that approaches zero as maturity increases. The supervised factor model (1) becomes then

$$
\begin{aligned}
\begin{bmatrix} y_t^{\tau_1} \\ \vdots \\ y_t^{\tau_N} \\ x_t \end{bmatrix} &= \begin{bmatrix} \underset{N\times 3}{\boldsymbol{\Lambda}}(\lambda) & \underset{N\times 1}{\mathbf{0}} \\ \underset{1\times 3}{\mathbf{0}} & 1 \end{bmatrix} \begin{bmatrix} LV_t \\ SL_t \\ CV_t \\ x_t \end{bmatrix} + \begin{bmatrix} \epsilon_t^{\tau_1} \\ \vdots \\ \epsilon_t^{\tau_N} \\ 0 \end{bmatrix}, \\
\begin{bmatrix} LV_{t+1} \\ SL_{t+1} \\ CV_{t+1} \\ x_{t+1} \end{bmatrix} &= \mathbf{c} + \mathbf{T} \begin{bmatrix} LV_t \\ SL_t \\ CV_t \\ x_t \end{bmatrix} + \boldsymbol{\eta}_t,
\end{aligned} \tag{7}
$$

with

$$
\boldsymbol{\Lambda}(\lambda) = \begin{bmatrix} 1 & \frac{1-e^{-\tau_1\lambda}}{\tau_1\lambda} & \frac{1-e^{-\tau_1\lambda}}{\tau_1\lambda} - e^{-\tau_1\lambda} \\ 1 & \frac{1-e^{-\tau_2\lambda}}{\tau_2\lambda} & \frac{1-e^{-\tau_2\lambda}}{\tau_2\lambda} - e^{-\tau_2\lambda} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\tau_N\lambda}}{\tau_N\lambda} & \frac{1-e^{-\tau_N\lambda}}{\tau_N\lambda} - e^{-\tau_N\lambda} \end{bmatrix}, \tag{8}
$$

where $\{y_t^{\tau_1}, \ldots, y_t^{\tau_N}\}$ are the yields for maturities $\{\tau_1, \ldots, \tau_N\}$, $x_t(h)$ is the forecast objective, and h the forecast lead, $\{LV_t, SL_t, CV_t\}$ are latent factors, $\boldsymbol{\Lambda}(\lambda)$ is a matrix of factor loadings, $\mathbf{T}$ and $\mathbf{c}$ are a matrix and a vector of coefficients, respectively, of suitable dimensions, $\boldsymbol{\epsilon}_t$ and $\boldsymbol{\eta}_t$ are vectors of disturbances with $\mathbf{H}$ and $\mathbf{Q}$ as respective variance-covariance matrices. The forecast objective $x_t(h)$ is a function of the forecast horizon, see 4.3.1 for more details.

Koopman et al. (2010) suggested to make the parameter λ time varying in order to get a better fit of the yield curve. The parameter λ, or rather its logarithm, is then included in the state vector and follows joint dynamics with the slope, level, and curvature factors. We also consider a variation of (8) in which λ is time-varying, but in a more parsimonious way by letting λ_t follow an AR(1). The supervised factor model is then modified to have $\Lambda = \Lambda(\lambda_t)$. We parameterize the model in the following way

$$
\begin{bmatrix} y_t^{\tau_1} \\ \vdots \\ y_t^{\tau_N} \\ x_t(h) \end{bmatrix} = \begin{bmatrix} \underset{N\times 3}{\boldsymbol{\Lambda}}(\lambda_t) & \underset{N\times 1}{\mathbf{0}} \\ \underset{1\times 3}{\mathbf{0}} & 1 \end{bmatrix} \begin{bmatrix} LV_t \\ SL_t \\ CV_t \\ x_t(h) \end{bmatrix} + \begin{bmatrix} \epsilon_t^{\tau_1} \\ \vdots \\ \epsilon_t^{\tau_N} \\ 0 \end{bmatrix}, \tag{9}
$$

$$
\begin{bmatrix} LV_{t+1} \\ SL_{t+1} \\ CV_{t+1} \\ x_{t+1} \\ log(\lambda_{t+1}) \end{bmatrix} = \mathbf{c} + \tilde{\mathbf{T}} \begin{bmatrix} LV_t \\ SL_t \\ CV_t \\ x_t \\ log(\lambda_t) \end{bmatrix} + \tilde{\boldsymbol{\eta}}_t, \tag{10}
$$

where $\tilde{\mathbf{T}}$ is now the block matrix

$$\tilde{\mathbf{T}} = \begin{bmatrix} \mathbf{T} & \mathbf{0} \\ \mathbf{0} & \phi \end{bmatrix}, \tag{11}$$

and $\tilde{\boldsymbol{\eta}}_t \sim N(\mathbf{0}, \tilde{\mathbf{Q}})$ with

$$\tilde{\mathbf{Q}} = \begin{bmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \sigma^2_\lambda \end{bmatrix}. \tag{12}$$

The system is now non-linear in the state vector and can be estimated via the extended Kalman filter[1], see for instance Durbin and Koopman (2012). Indicate with $\boldsymbol{\alpha}_t = [LV_t, SL_t, CV_t, x_t, log(\lambda_t)]'$ the state vector. The measurement equation has then the form

$$\begin{bmatrix} \mathbf{y}_t \\ x_t \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Lambda}(\lambda_t) & \mathbf{0} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} LV_t \\ SL_t \\ CV_t \\ x_t(h) \end{bmatrix} + \begin{bmatrix} \epsilon_t^{\tau_1} \\ \vdots \\ \epsilon_t^{\tau_N} \\ 0 \end{bmatrix} = \mathbf{Z}_t(\boldsymbol{\alpha}_t) + \begin{bmatrix} \boldsymbol{\epsilon}_t \\ 0 \end{bmatrix}, \tag{13}$$

where the state vector $\boldsymbol{\alpha}_t$ follows the dynamics specified in equations (9)-(12). The extended Kalman filter is based on a local linearization of $\mathbf{Z}_t(\boldsymbol{\alpha}_t)$ at $\mathbf{a}_{t|t-1}$, an estimate of $\boldsymbol{\alpha}_t$ based on the past observations $y_1, \ldots, y_{t-1}$ and $x_1(h), \ldots, x_{t-1}(h)$. The linearized model is thus

$$\begin{aligned} \begin{bmatrix} \mathbf{y}_t \\ x_t \end{bmatrix} &= \mathbf{Z}_t(\mathbf{a}_{t|t-1}) + \dot{\mathbf{Z}}_t(\mathbf{a}_{t|t-1})(\boldsymbol{\alpha}_t - \mathbf{a}_{t|t-1}) + \begin{bmatrix} \boldsymbol{\epsilon}_t \\ 0 \end{bmatrix} \\ &= \mathbf{d}_t + \dot{\mathbf{Z}}_t(\mathbf{a}_{t|t-1})\boldsymbol{\alpha}_t + \begin{bmatrix} \boldsymbol{\epsilon}_t \\ 0 \end{bmatrix}, \end{aligned} \tag{14}$$

where

$$\mathbf{d}_t = \mathbf{Z}_t(\mathbf{a}_{t|t-1}) - \dot{\mathbf{Z}}_t(\mathbf{a}_{t|t-1})\mathbf{a}_{t|t-1}, \tag{15}$$

and

$$\dot{\mathbf{Z}}_t(\mathbf{a}_{t|t-1}) = \left. \frac{\partial \mathbf{Z}_t(\boldsymbol{\alpha}_t)}{\partial \boldsymbol{\alpha}_t'} \right|_{\boldsymbol{\alpha}_t = \mathbf{a}_{t|t-1}}, \tag{16}$$

$$\dot{\mathbf{Z}}_t(\boldsymbol{\alpha}_t) = \begin{bmatrix} \boldsymbol{\Lambda}(\lambda_t) & \mathbf{0} & \frac{\partial \mathbf{Z}_t(\boldsymbol{\alpha}_t)}{\partial log(\lambda_t)} \\ \mathbf{0} & 1 & 0 \end{bmatrix}, \tag{17}$$

with

$$\frac{\partial \mathbf{Z}_t(\boldsymbol{\alpha}_t)}{\partial log(\lambda_t)} = \begin{bmatrix} \frac{e^{-\lambda_t \tau_1}(\lambda_t \tau_1 - e^{\lambda_t \tau_1} + 1)}{\lambda_t^2 \tau_1} \lambda_t SL_t + \frac{e^{-\lambda_t \tau_1}(\tau_1^2 \lambda_t^2 + \lambda_t \tau_1 - e^{\lambda_t \tau_1} + 1)}{\lambda_t^2 \tau_1} \lambda_t CV_t \\ \vdots \\ \frac{e^{-\lambda_t \tau_N}(\lambda_t \tau_N - e^{\lambda_t \tau_N} + 1)}{\lambda_t^2 \tau_N} \lambda_t SL_t + \frac{e^{-\lambda_t \tau_N}(\tau_N^2 \lambda_t^2 + \lambda_t \tau_N - e^{\lambda_t \tau_N} + 1)}{\lambda_t^2 \tau_N} \lambda_t CV_t \end{bmatrix}.$$

The state-space system is then made up of the measurement equation (14) and state equation (10).

[1] Exact estimation procedures for non-linear systems require a major computational effort as opposed to the extended Kalman filter.

3 Computational aspects

The objective of the study is to determine the forecasting power of the yield curve and the supervised factor model (1). The forecast performance is based on out-of-sample forecasts for which a rolling window of fixed size is used for the estimation of the parameters. The log-likelihood is maximized for each estimation window.

3.1 Estimation method

The parameters of the state-space model are estimated by maximum likelihood. The likelihood is delivered by the Kalman filter. We employ the univariate Kalman filter as derived in Koopman and Durbin (2000) as we assume a diagonal covariance matrix for the innovations in the measurement equation. The maximum of the likelihood function has no explicit form solution and numerical methods have to be employed. We make use of the following two algorithms.

- **CMA-ES**. Covariance Matrix Adaptation Evolution Strategy, see Hansen and Ostermeier (1996)[2]. This is a genetic algorithm that samples the parameter space according to a Gaussian search distribution which changes according to where the best solutions are found in the parameter space;

- **BFGS**. Broyden-Fletcher-Goldfarb-Shanno, see for instance Press et al. (2002). This algorithm belongs to the class of quasi-Newton methods. The algorithm needs the computation of the gradient of the function to be minimized.

The CMA-ES algorithm is used for the first estimation window for each forecast target. This algorithm is particularly useful, in this context, if no good guess of initial values is available. We then use the BFGS algorithm for the rest of the estimation windows as this method is substantially faster than the CMA-ES but more dependent on initial values. We use algorithmic (or automatic) differentiation[3] to compute gradients. We make use of the ADEPT library C++ library, see Hogan (2013)[4]. The advantage of using algorithmic differentiation over finite differences is twofold: increased speed and elimination of approximation errors.

3.2 Speed improvements

To gain speed we chose C++ as the programming language, using routines from the Numerical Recipes, Press et al. (2002) [5]. We compile and run the executables on a Linux 64-bit operating system using GCC [6]. We use Open MPI 1.6.4 (Message Passing Interface) with the Open MPI C++ wrapper compiler `mpic++` to parallelise the maximum

[2]See `https://www.lri.fr/~hansen/cmaesintro.html` for references and source codes. The authors provide C source code for the algorithm which can be easily converted into C++ code.

[3]See for instance Verma (2000) for an introduction to algorithmic differentiation.

[4]For a user guide see `http://www.cloud-net.org/~clouds/adept/adept_documentation.pdf`.

[5]See Aruoba and Fernández-Villaverde (2014) for a comparison of different programming languages in economics and Fog (2006) for many suggestions on how to optimize software in C++.

[6]See `http://gcc.gnu.org/onlinedocs/` for more information on the Gnu Compiler Collection, GCC.

likelihood estimations [7]. We compute gradients using the ADEPT library for algorithmic differentiation, see Hogan (2013).

4 Empirical application

4.1 Data

The macroeconomic variables selected as forecast objectives are: consumer price index (CPI), personal consumer expenditures (PCE), producer price index (PPI), real disposable income (RDI), unemployment rate (UR), and industrial production (IP). The macroeconomic data have been taken from FRED (Federal Reserve Economic Data)[8].

4.1.1 Yield curve

We use two datasets for the yield-curve data. The first one is from Gürkaynak et al. (2007)[9]. We skip-sample the data in order to avoid inducing artificial persistence in the time series. We take the first yield registered in the month as the yield for that month. This is consistent with the macroeconomic variables whose values refer to the first days of the month. The second dataset is the yield-curve data from Diebold and Li (2006). All data refer to the US economy.

4.1.2 Macroeconomic variables

- **CPI.** Series ID: CPIAUCSL, Title: Consumer Price Index for All Urban Consumers: All Items, Source: U.S. Department of Labor: Bureau of Labor Statistics, Release: Consumer Price Index, Units: Index 1982-84=100, Frequency: Monthly, Seasonal Adjustment: Seasonally Adjusted, Notes: Handbook of Methods (`http://www.bls.gov/opub/hom/pdf/homch17.pdf`).
- **PPI.** Series ID: PPIFGS, Title: Producer Price Index: Finished Goods, Source: U.S. Department of Labor: Bureau of Labor Statistics, Release: Producer Price Index, Units: Index 1982=100, Frequency: Monthly, Seasonal Adjustment: Seasonally Adjusted, Series ID: UNRATE, Title: Civilian Unemployment Rate, Source: U.S. Department of Labor: Bureau of Labor Statistics, Release: Employment Situation, Units: Percent, Frequency: Monthly, Seasonal Adjustment: Seasonally Adjusted.
- **RDI.** Series ID: DSPIC96, Title: Real Disposable Personal Income, Source: U.S. Department of Commerce: Bureau of Economic Analysis, Release: Personal Income and Outlays, Units: Billions of Chained 2009 Dollars, Frequency: Monthly, Seasonal Adjustment: Seasonally Adjusted Annual Rate, Notes: BEA Account, Code: A067RX1, A Guide to the National Income and Product Accounts of the United States (NIPA) - (`http://www.bea.gov/national/pdf/nipaguid.pdf`).

[7] See `http://www.open-mpi.org/` for more details on Open MPI and Karniadakis (2003) for a review of parallel scientific computing in C++ and MPI.

[8] The data can be downloaded from the website of the Federal Reserve Bank of St. Louis: `http://research.stlouisfed.org/fred2`.

[9] The Gürkaynak, Sack and Wright dataset can be downloaded from `http://www.federalreserve.gov/pubs/feds/2006`.

- **UR.** Series ID: UNRATE, Title: Civilian Unemployment Rate, Source: U.S. Department of Labor: Bureau of Labor Statistics, Release: Employment Situation, Units: Percent, Frequency: Monthly, Seasonal Adjustment: Seasonally Adjusted.

- **IP.** Series ID: INDPRO, Title: Industrial Production Index, Source: Federal Reserve Economic Data, Units: Levels, Frequency: Monthly, Index: 2007=100, Seasonal Adjustment: Seasonally Adjusted, Link: http://research.stlouisfed.org/fred2.

4.2 Competing models

We choose different competing models diffusely used in the forecasting literature, in order to assess the relative forecasting performance of the supervised factor model. We divide these models into direct multi-step and indirect (recursive) forecasting models.

4.2.1 Direct forecasting models

The first model is the following restricted $AR(p)$ process

$$x_{t+h}(h) = c + \phi_1 x_t(h) + \dots \phi_p x_{t-p}(h) + \epsilon_{t+h}. \tag{18}$$

The second model is a restricted $MA(q)$ process

$$x_{t+h}(h) = c + \theta_1 \epsilon_t + \dots \theta_q \epsilon_{t-q} + \epsilon_{t+h}. \tag{19}$$

Both models are estimated by maximum likelihood. The lags p and q are selected for each estimation sample as the values that minimize the Bayesian information criterion. In particular, we consider $p, q \in \{1, 2, 3\}$.

The third model is principal component regression (PCR). In the first step, principal components are extracted from the regressors $\mathbf{Y}_t = [y_t^{\tau_1}, \dots, y_t^{\tau_N}, x_t(h)]$; $x_{t+h}(h)$ is then regressed on them to obtain $\hat{\beta}^{PCR}$ for time indexes $1 \leq t \leq T_i - h$. In the second step, the principal components are projected at time T_i and then multiplied by $\hat{\beta}^{PCR}$ to obtain the h-period ahead forecast. We estimate three factors.

The fourth model considered is partial least squares regression (PLSR). In the first step, the partial least squares components $\hat{x}_t^m$ are computed using the forecast target $\{x_t(h) : h \leq t \leq T_i\}$ and the predictors $\mathbf{Y}_t = [y_t^{\tau_1}, \dots, y_t^{\tau_N}, x_t(h)]$ with $1 \leq t \leq T_i - h$ where $M \leq (N+1)$ is the number of partial least squares components and $N+1$ is the number of predictors including the lagged value of the forecast objective. In the second step, the partial least squares components $\hat{x}_t^m$ are regressed on the predictors $\mathbf{Y}_t$ to recover the coefficient vector $\hat{\beta}^{PLS}$. Note that as the partial least squares components are a linear combination of the regressors, the relation is exact, i.e. the residuals from this regression are (algebraically) null. In the third step, the partial least squares components are projected at time T_i by multiplying $\mathbf{Y}_{T_i}$ by $\hat{\beta}^{PLS}$. The projected PLS components at time T_i are then summed to obtain the h-period ahead forecast $\hat{x}_{T_i+h}(h) = \sum_{m=1}^{M} \hat{x}_{T_i}(h)^m$. We estimate three PLS directions.

The fifth direct forecasting method considered is a two-step procedure as described in equations (5) in which the factors are extracted using the Kalman filter and the factor loadings have the NS structure. We refer to these factor models as unsupervised.

The sixth model is a two-step procedure where the factors are first extracted using the semi-parametric factor model implemented in Härdle et al. (2012) and then used to forecast with an auxiliary equation as in eqn. (3).

The seventh competing method is the forecast combination (CF-NS) model derived in Hillebrand et al. (2012).

Finally, we compare the forecast performance of the supervised models with fixed λ (7), (8), and time varying λ (14), (10) to their unsupervised counterparts. In these specifications the factors are first extracted using the Kalman filter, the forecasts are then obtained using the forecast equation

$$x_{t+h}(h) \quad = \quad c + \hat{\mathbf{f}}_t \boldsymbol{\beta} + \gamma x_t(h) + u_t, \tag{20}$$

where c, $\boldsymbol{\beta}$, and γ are parameters to be estimated, u_t is the error term, and $\hat{\mathbf{f}}_t$ is the vector of filtered factors.

4.2.2 Indirect forecasting models

The first model is the following $AR(p)$ process

$$x_{t+h}(h) = c + \phi_1 x_{t+h-1}(h) + \ldots \phi_p x_{t+h-p}(h) + \epsilon_{t+h}. \tag{21}$$

The second model is an $MA(q)$ process

$$x_{t+h}(h) = c + \theta_1 \epsilon_{t+h-1} + \ldots \theta_q \epsilon_{t+h-q} + \epsilon_{t+h}. \tag{22}$$

Both models are estimated using maximum likelihood. The lags p and q are selected for each estimation sample as the values that minimize the Bayesian information criterion. In particular, we consider $p, q \in \{1, 2, 3\}$.

4.3 Forecasting

4.3.1 Forecast objective

We stationarize the forecast objectives by treating them as either $I(2)$ or $I(1)$ variables. In particular we follow Stock and Watson (2002b) and treat price indexes as $I(2)$ processes and real variables as $I(1)$ and take the following transformations

$$x_t(h) = \begin{cases} k(log(X_t) - log(X_{t-h}))/h & \text{if } X_t \text{ is I(1)}, \\ k(log(X_t) - log(X_{t-h}))/h - k(log(X_{t-h} - log(X_{t-h-1})) & \text{if } X_t \text{ is I(2)}, \end{cases} \tag{23}$$

where k is a scaling factor. In particular, we have $k = 1200$, $h \in \{1, 3, 6, 9, 12\}$, $X_t \in \{CPI_t, PCE_t, PPI_t, RDI_t, UR_t\}$.

4.3.2 Forecasting scheme

The aim is to compute the forecast of the macro variable $x_t(h)$ at time $t + h$, i.e. $x_{t+h}(h)$ where h is the forecast lead. We consider a rolling windows scheme. The reason is that one of the requirements for the application of the Giacomini and White (2006) test, in case of nested models, is to use rolling windows. We build series of forecast errors of length S for all forecast objectives/leads. The complete time series is indexed $\{\mathbf{Y}_t : t \in \mathbb{N}_{>0}, t \leq T\}$ where T is the sample length of the complete dataset and $\mathbf{Y}_t = \{y_t^1, ..., y_t^N, x_t(h)\}$, where $y_t^i, i = 1, 2, ..., N$. The estimation sample takes into account observations indexed $\{\mathbf{Y}_t : t \in \mathbb{N}_{>0}, T_i - R + 1 \leq t \leq T_i\}$ for $i \in \mathbb{N}_{>0}, i \leq S$ with $T_1 = R = T - S - h^{max} + 1$ the index of the last observation of the first estimation sample, which coincides with the size of the rolling window, and $T_i = T_1 + i$ for $i \in \mathbb{N}_{>0}, i \leq S$ and h^{max} is the maximum forecast lead. In particular, we chose different values of S, and R depending on the complete sample length of the yield curve data, T. We chose $h_{max} = 12$ for all applications. The forecasting strategy for the indirect h-step ahead forecasts for the supervised factor model, and indirect forecasts (equations (7) and (8)) is the following:

(i) estimate the system parameters of the dynamic factor model using information from time $T_i - R + 1$ up to time T_i by maximizing the log-likelihood function delivered by the Kalman filter;

(ii) indicating with $\boldsymbol{\alpha}_t$ the state vector containing the latent factors and the forecast target, compute the smoothed estimate at time T_i of the state vector, i.e. $\hat{\boldsymbol{\alpha}}_{T_i}$;

(iii) iterate h times on the filtered state at time T_i, $\hat{\boldsymbol{\alpha}}_{T_i}$, using the estimated parameters to obtain the forecast:

$$\hat{x}_{T+h|T} = [\mathbf{0}^{1\times k} : 1] \left[\hat{\mathbf{T}}^h \hat{\boldsymbol{\alpha}}_{T_i} + \sum_{j=0}^{h-1} \hat{\mathbf{T}}^j \hat{\mathbf{c}} \right]; \tag{24}$$

for $i = 1, \ldots, S$.

4.3.3 Test of forecast performance

We use the conditional predictive ability (CPA) test proposed in Giacomini and White (2006)[10] using a quadratic loss function. This test also allows to compare nested models, provided a rolling windows scheme for parameter estimation is used. The autocorrelations of the loss differentials are taken into account by computing Newey and West (1987) standard errors. We follow the "rule of thumb" in Clark and McCracken (2011) and take a sample split ratio $\pi = \frac{S}{R}$ approximately equal to one.

4.4 Results

To assess the relative forecasting performance of the supervised factor models with respect to the competing methods, we present mean squared prediction errors ratios between forecasts from model (7) and the competing models. The results are summarised in tables 1-9 in Appendix A. In bold are the ratios lower than 1 which indicate a better forecasting performance of the supervised factor model, eqn. (7), compared to the competing method. We consider three applications:

(i) Gürkaynak et al. (2007) yield curve data, 7 maturities corresponding to 1, 2, 3, 4, 5, 6, and 7 years, sample size $T = 617$, rolling window of size $R = T_1 = 306$, number

[10] At `http://www.runshare.org/CompanionSite/site.do?siteId=116` the authors provide MATLAB codes for the test.

of forecasts $S = 300$. Observations range from August 1961 to December 2012. The 1-step ahead forecasts range from February 1987 to January 2012. The 12-step ahead forecasts range from January 1988 to December 2012.

(ii) Gürkaynak et al. (2007) yield curve data, 30 maturities corresponding to 1, 2, ..., 29, and 30 years, sample size $T = 325$, rolling window of size $R = T_1 = 189$, number of forecasts $S = 125$. Observations range from December 1985 to December 2012. The 1-step ahead forecasts range from September 2001 to January 2012. The 12-step ahead forecasts range from August 2002 to December 2012.

(iii) Diebold and Li (2006) yield curve data, 17 maturities corresponding to 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108, and 120 months, sample size $T = 346$, rolling window of size $R = T_1 = 185$, number of forecasts $S = 150$. Observations range from February 1972 to November 2000. The 1-step ahead forecasts range from July 1987 to December 1999. The 12-step ahead forecasts range from June 1988 to November 2000.

In the tables we label the different forecasting models according to the following convention.

- **model 1**. Principal component regression (PCR);
- **model 2**. Partial least squares regression (PLSR);
- **model 3**. AR(p) direct, equation (18);
- **model 4**. MA(q) direct, equation (19);
- **model 5**. AR(p) indirect, equation (21);
- **model 6**. MA(q) indirect, equation (22);
- **model 7**. Supervised dynamic factor model with Nelson-Siegel factor loadings (8), and direct forecasts (2);
- **model 8**. Supervised dynamic factor model with Nelson-Siegel loadings (7), and indirect forecasts (8);
- **model 9**. Supervised dynamic factor model with Nelson-Siegel loadings with time varying $\lambda = \lambda(t)$, and indirect forecasts (10), (9);
- **model 10**. Unsupervised dynamic factor model with Nelson-Siegel loadings, (7), (8), and forecast equation 20;
- **model 11**. Unsupervised dynamic factor model with Nelson-Siegel loadings and time varying $\lambda = \lambda(t)$, (10), (9), and forecast equation 20;
- **model 12**. Factors extracted using the semiparametric factor model, Härdle et al. (2012), and forecast equation 20;
- **model 13**. Forecast combination-Nelson Siegel (CF-NS), see Hillebrand et al. (2012).

Looking at tables 1-9 in the Appendix, we can make the following remarks (divided with respect to the three applications):

(i) Concerning the first application we note that the supervised factor model (7) delivers better forecasts for output variables, namely, real disposable income (RDI), industrial production (IP), and unemployment rate (UR), relative to the different inflation variables, that is, consumer price index (CPI), personal consumption expenditures (PCE), and producer price index (PPI). With regard to the performance relative to other forecasting schemes the supervised factor model generally performs similarly to or better than principal components regression (model 1), partial least squares regression (model 2), a two-step forecasting procedure (model 12), the unsupervised scheme (model 10), and the CF-NS procedure (model 13), as can be seen from tables 1-3. For inflation related variables the direct and indirect MA(q) and AR(p) forecasts are hard to beat, in particular for forecast lead $h = 1$. The supervised factor model performs generally better than the unsupervised counterpart. Allowing for dynamics in the λ improves the forecasts only for few forecast targets. For real disposable income (RDI), the performance of model 8, relative to univariate models, in the first half of the forecasting sample (corresponding to the first 13 years of the Greenspan monetary regime, 1987-2000), is considerably worse than that in the second half of the forecasting sample (corresponding to the last 6 years of the Greenspan monetary regime and the first 6 years of the Bernanke monetary regime, 2001-2012) as can be seen from tables 2 and 3.

(ii) The results for the second application are similar to the ones for the first application, as can be seen from tables 4-6. For real disposable income (RDI), the performance of model 8, relative to univariate models, in the first half of the forecasting sample (corresponding to the last 6 years of the Greenspan monetary regime, 2001-2006), is considerably worse than that in the second half of the forecasting sample (corresponding to the first 6 years of the Bernanke monetary regime, 2006-2012) as can be seen from tables 5 and 6.

(iii) Regarding the third application the differences between inflation and output variables is less marked, as can be seen from tables 7-9. The supervised framework delivers forecasts similar to or better than the unsupervised frameworks. Unemployment rate (UR) and industrial production (IP) are the two variables for which supervision is more beneficial. Allowing for dynamics in the λ parameter does not improve the forecasts. For real disposable income (RDI), the performance of model 8, relative to univariate models, in the first half of the forecasting sample (corresponding to the first 6 years of the Greenspan monetary regime, 1987-1993), is better than that in the second half of the forecasting sample (corresponding to the Greenspan monetary regime going from 1994-1999) as can be seen from tables 8 and 9.

Overall, from tables 1-9 we can see that the yield curve has more predictive power for the periods 1987-1994 (the early Greenspan monetary regime) and 2006-2012 (the early Bernanke monetary regime) as compared to the period 1994-2006 (the late Greenspan monetary regime). The good forecasting power of the yield curve in the early Greenspan period is consistent with the findings in Giacomini and Rossi (2006). Note that in this study we cannot make use of the supervision measure proposed in Boldrini and Hillebrand (2015) as the observable variables are not guaranteed to be stationary.

5 Conclusions

In this paper we study the forecasting power of the yield curve for some macroecomic variables, for the US economy, in the framework of a supervised factor model. In this model the factors are extracted conditionally on the forecast target. The model has a linear state-space representation and standard Kalman filtering techniques apply.

We forecast macroeconomic variables using factors extracted from the yield curve. We use the yield curve data from Gürkaynak et al. (2007) and Diebold and Li (2006) and macroeconomic data from FRED. We use the Nelson and Siegel factor loadings and allow for dynamics in the factors. We forecast consumer price index (CPI), personal consumer expenditures (PCE), producer price index (PPI), real disposable income (RDI), unemployment rate (UR), and industrial production (IP).

We find that supervising the factor extraction can improve the forecasting performance of the factor model. For this dataset and specification the supervised factor model outperforms principal components regression, and partial least squares regression on most targets, in particular for UR, RDI, and IP. In forecasting inflation, both measured by consumer price index, and producer price index, $MA(q)$ and $AR(p)$ processes are difficult to beat, especially for one-step ahead forecasts. Allowing for dynamics in the Nelson and Siegel factor loadings generally does not improve the forecasts. The supervised factor model performs particularly well in forecasting unemployment rate and real disposable income. Furthermore, supervising the factor extraction leads in most cases to improved forecasts compared to unsupervised two-step forecasting schemes.

We find that the yield curve has forecast power for unemployment rate, real disposable income, and industrial production but less so for inflation measures. Similarly to Giacomini and Rossi (2006), Rudebusch and Williams (2009), and Stock and Watson (1999a) we find that the predictive ability of the yield curve is somewhat unstable and has changed through the years. In particular, we find that the yield curve has more predictive power for the periods 1987-1994 (the early Greenspan monetary regime) and 2006-2012 (the early Bernanke monetary regime) as compared to the period 1994-2006 (the late Greenspan monetary regime). The good forecasting power of the yield curve in the early Greenspan period is consistent with the findings in Giacomini and Rossi (2006).

6 Acknowledgements

This research was supported by the European Research Executive Agency in the Marie-Sklodowska-Curie program under grant number 333701-SFM.

References

Ang, A., M. Piazzesi, and M. Wei (2006). What does the yield curve tell us about gdp growth? *Journal of Econometrics 131*(1), 359–403.

Aoki, M. (1990). State space modeling of time series.

Aruoba, S. B. and J. Fernández-Villaverde (2014). A comparison of programming languages in economics.

Bai, J. and S. Ng (2008a). Forecasting economic time series using targeted predictors. *Journal of Econometrics 146*(2), 304–317.

Bai, J. and S. Ng (2008b). *Large dimensional factor analysis.* Now Publishers Inc.

Bai, J. and S. Ng (2009). Boosting diffusion indices. *Journal of Applied Econometrics 24*(4), 607–629.

Bair, E., T. Hastie, D. Paul, and R. Tibshirani (2006). Prediction by supervised principal components. *Journal of the American Statistical Association 101*(473).

Bernanke, B. (1990). On the predictive power of interest rates and interest rate spreads. Technical report, National Bureau of Economic Research.

Boldrini, L. and E. Hillebrand (2015). Supervision in factor models using a large number of predictors. Technical report, Aarhus University and CREATES.

Breitung, J. and S. Eickmeier (2006). Dynamic factor models. *Allgemeines Statistisches Archiv 90*(1), 27–42.

Brockwell, P. J. and R. A. Davis (2009). *Time series: theory and methods.* Springer.

Chinn, M. D. and K. J. Kucko (2010). The predictive power of the yield curve across countries and time. Technical report, National Bureau of Economic Research.

Clark, T. E. and M. W. McCracken (2011). Advances in forecast evaluation. *Federal Reserve Bank of St. Louis Working Paper Series*.

Diebold, F., G. Rudebusch, and S. Borağan Aruoba (2006). The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of econometrics 131*(1), 309–338.

Diebold, F. X. and C. Li (2006). Forecasting the term structure of government bond yields. *Journal of econometrics 130*(2), 337–364.

Durbin, J. and S. J. Koopman (2012). *Time series analysis by state space methods.* Oxford University Press.

Estrella, A. and G. A. Hardouvelis (1991). The term structure as a predictor of real economic activity. *The Journal of Finance 46*(2), 555–576.

Fog, A. (2006). Optimizing software in C++.

Friedman, B. M. and K. Kuttner (1993). Why does the paper-bill spread predict real economic activity? In *Business Cycles, Indicators and Forecasting*, pp. 213–254. University of Chicago Press.

Friedman, J., T. Hastie, and R. Tibshirani (2001). *The elements of statistical learning*, Volume 1. Springer series in statistics Springer, Berlin.

Giacomini, R. and B. Rossi (2006). How stable is the forecasting performance of the yield curve for output growth? *Oxford Bulletin of Economics and Statistics 68*(s1), 783–795.

Giacomini, R. and H. White (2006). Tests of conditional predictive ability. *Econometrica 74*(6), 1545–1578.

Giovannelli, A. and T. Proietti (2014). On the selection of common factors for macroeconomic forecasting.

Gürkaynak, R. S., B. Sack, and J. H. Wright (2007). The us treasury yield curve: 1961 to the present. *Journal of Monetary Economics 54*(8), 2291–2304.

Hamilton, J. D. and D. H. Kim (2000). A re-examination of the predictability of economic activity using the yield spread. Technical report, National Bureau of Economic Research.

Hansen, N. and A. Ostermeier (1996). Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. In *Evolutionary Computation, 1996., Proceedings of IEEE International Conference on*, pp. 312–317. IEEE.

Härdle, W. K., P. Majer, and M. Schienle (2012). Yield curve modeling and forecasting using semiparametric factor dynamics. Technical report, Sonderforschungsbereich 649, Humboldt University, Berlin, Germany.

Harvey, C. R. (1988). The real term structure and consumption growth. *Journal of Financial Economics 22*(2), 305–333.

Hillebrand, E. T., H. Huang, T. Lee, and C. Li (2012). Using the yield curve in forecasting output growth and inflation. Technical report, Institut for Økonomi, Aarhus Universitet, Danmark.

Hogan, R. J. (2013). Fast reverse-mode automatic differentiation using expression templates in C++. *Submitted to ACM Trans. Math. Softw.*

Karniadakis, G. E. (2003). *Parallel scientific computing in C++ and MPI: a seamless approach to parallel algorithms and their implementation.* Cambridge University Press.

Knez, P. J., R. Litterman, and J. Scheinkman (1994). Explorations into factors explaining money market returns. *The Journal of Finance 49*(5), 1861–1882.

Koopman, S. J. and J. Durbin (2000). Fast filtering and smoothing for multivariate state space models. *Journal of Time Series Analysis 21*(3), 281–296.

Koopman, S. J., M. I. Mallee, and M. Van der Wel (2010). Analyzing the term structure of interest rates using the dynamic nelson–siegel model with time-varying parameters. *Journal of Business & Economic Statistics 28*(3), 329–343.

Kozicki, S. (1997). Predicting real growth and inflation with the yield spread. *Economic Review-Federal Reserve Bank of Kansas City 82*, 39–58.

Litterman, R. B. and J. Scheinkman (1991). Common factors affecting bond returns. *The Journal of Fixed Income 1*(1), 54–61.

Nelson, C. R. and A. F. Siegel (1987). Parsimonious modeling of yield curves. *Journal of business*, 473–489.

Newey, W. K. and K. D. West (1987). Hypothesis testing with efficient method of moments estimation. *International Economic Review 28*(3), 777–787.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (2002). *Numerical Recipes in C++* (Second ed.). Cambridge University Press, Cambridge.

Rudebusch, G. D. and J. C. Williams (2009). Forecasting recessions: the puzzle of the enduring power of the yield curve. *Journal of Business & Economic Statistics 27*(4).

Stock, J. H. and M. W. Watson (1989). New indexes of coincident and leading economic indicators. In *NBER Macroeconomics Annual 1989, Volume 4*, pp. 351–409. MIT Press.

Stock, J. H. and M. W. Watson (1999a). Business cycle fluctuations in us macroeconomic time series. *Handbook of macroeconomics 1*, 3–64.

Stock, J. H. and M. W. Watson (1999b). Forecasting inflation. *Journal of Monetary Economics 44*(2), 293–335.

Stock, J. H. and M. W. Watson (2002a). Forecasting using principal components from a large number of predictors. *Journal of the American statistical association 97*(460), 1167–1179.

Stock, J. H. and M. W. Watson (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics 20*(2), 147–162.

Stock, J. H. and M. W. Watson (2011). Dynamic factor models. *Oxford Handbook of Economic Forecasting*, 35–59.

Verma, A. (2000). An introduction to automatic differentiation. *CURRENT SCIENCE-BANGALORE- 78*(7), 804–807.

Appendices

A Tables

In this section we report mean square forecast errors (MSFE) ratios corresponding to the empirical application (see Section 4). The results correspond to MSFE ratios between model 8 and the competing models (see Section 4 for the description of the different models involved). We consider different subsamples of the dataset. In the tables below, three, two, and one stars refer to significance levels 0.01, 0.05, and 0.10 for the null hypothesis of equal conditional predictive ability for the Giacomini and White (2006) test.

Table 1: MSFE ratios for whole forecast sample (Gürkaynak et al. (2007) yield data, 7 maturities).

	h	*mod 1*	*mod 2*	*mod 3*	*mod 4*	*mod 5*	*mod 6*	*mod 7*	*mod 9*	*mod 10*	*mod 11*	*mod 12*	*mod 13*
	1	1,02	1,02	1,14	1,37*	1,14	1,37*	1	1,01	1	1,01	1,02	**0,99***
	3	1,16	1,16	1,12	1,23	1,12	1,31**	1,14	1,04	1,15	1,16	1,16	1,04
CPI	6	1,03	1,03	1,05	1,07	1,02	1,03	1	1,01	1,03	1,03	1,03	**0,99**
	9	**0,98**	**0,98**	1,01	1,02	1	1	**0,98****	**0,96**	**0,97****	**0,98***	**0,98**	**0,96****
	12	1	**0,99**	1	1	1,01*	1,01*	1	1,01	1,01	**0,99**	1	**0,96****
	1	1,01	1	1,33***	1,69***	1,33***	1,69***	1*	**0,81***	1*	**0,99**	1,01	**0,6****
	3	1	1	1	**0,99**	**0,95**	1,01	1	**0,88*****	1	**0,99**	1	1
PCE	6	**0,99**	**0,99**	**0,99**	**0,99**	1	1	1	**0,95***	**0,99**	**0,98**	**0,99**	1
	9	**0,98**	**0,98**	**0,98**	**0,98**	1	1	**0,99**	1	**0,97****	**0,98***	**0,98**	**0,99**
	12	1,01	1,01	1,02	1,03	1	1	1	1	1	1	1,01	**0,99**
	1	1,02	1,01	1,14	1,29*	1,14	1,29*	1	**0,13**	1	1,01	1,02	**0,82*****
	3	1,06	1,06	1,03	1,1	1,04	1,18	1,04	1,03	1,04	1,05	1,06	1,02
PPI	6	1,07	1,07	1,07	1,13*	1,02	1,02	1,03	1,01	1,05	1,07	1,07	**0,99**
	9	1,01	1	1,02	1,01	1*	1*	1	1	**0,98**	1,01	1,01	**0,96**
	12	1,05	1,05	1,07	1,08*	1	1	1,01	**0,99***	1,02	1,04	1,05	**0,96**
	1	1	1	1,01	1,02	1,01	1,02	1	**0,93**	1	1	1	**0,97**
	3	1,08	1,08	1,14*	1,18*	1,07	1,16*	1,1	1,01	1,11	1,11	1,08	1,06**
RDI	6	**0,89**	**0,89**	1,02	1,02	1,01	1,03	**0,91**	**0,79****	**0,95**	**0,92**	**0,89**	1
	9	**0,83***	**0,84***	**0,98**	**0,97**	1	1,03	**0,87****	**0,76****	**0,94**	**0,89**	**0,84***	**0,99**
	12	**0,73****	**0,75****	**0,95**	**0,96**	1,06	1,06	**0,81***	**0,79***	**0,86**	**0,82***	**0,73****	1
	1	**0,97**	**0,97**	1,12**	1,07**	1,12**	1,07**	1	**0,79*****	1	**0,99**	**0,98**	1,07***
	3	**0,95**	**0,95**	1,1*	**0,96**	1	**0,76**	**0,94****	**0,85**	**0,95****	**0,96**	**0,95**	**0,82****
UR	6	**0,9**	**0,9**	**0,98**	**0,96**	**0,99****	**0,65****	**0,8****	**0,92**	**0,84***	**0,91**	**0,9**	**0,65*****
	9	**0,81**	**0,81**	**0,83****	**0,89**	**0,81****	**0,68*****	**0,75****	1,08	**0,75****	**0,86**	**0,82**	**0,65*****
	12	**0,81****	**0,81****	**0,82*****	**0,91****	**0,74****	**0,77*****	**0,73****	**0,85**	**0,72****	**0,87***	**0,81****	**0,72****
	1	**0,95**	**0,94**	1,07	1,04	1,07	1,04	**0,83****	**0,99**	1	**0,98**	**0,95**	1
	3	**0,84****	**0,85****	**0,99**	**0,96**	**0,91**	**0,68**	**0,6****	**0,51*****	**0,9****	**0,88***	**0,85****	**0,66***
IP	6	**0,83****	**0,84****	**0,99**	1,11	**0,95**	**0,86**	**0,7****	**0,99**	**0,86*****	**0,88**	**0,84****	**0,8***
	9	**0,96****	**0,97****	1,13	1,3	1,06	1,21	**0,88****	**0,88**	**0,95*****	1,03	**0,97****	1,08
	12	1,09	1,1	1,33	1,28	1,08	1,52	1,02**	1,22	1,06	1,18	1,1	1,36

Gürkaynak et al. (2007) yield data. 7 maturities, from 1 to 7 years. MSFE ratios between model 8 and models $1-13$ for CPI, PCE, PPI, RDI, UR, and IP for all forecasting leads h. A value lower than one indicates a lower MSFE of model 8 w.r.t. models $1-13$. One, two, and three stars mean .10, .05, and .01 statistical significance, respectively, for the Giacomini and White (2006) test with quadratic loss function. The number of forecasts is $S = 300$. The 1-step ahead forecasts range from February 1987 to January 2012. The 12-step ahead forecasts range from January 1988 to December 2012. The forecast period corresponds to part of the Greenspan (1987-2006) and Bernanke (2006-2012) monetary regimes.

Table 2: MSFE ratios for first half of forecast sample (Gürkaynak et al. (2007) yield data, 7 maturities).

	h	*mod 1*	*mod 2*	*mod 3*	*mod 4*	*mod 5*	*mod 6*	*mod 7*	*mod 9*	*mod 10*	*mod 11*	*mod 12*	*mod 13*
	1	1	1,01	1,12*	1,25***	1,12*	1,25***	1	1	1	0,99	1	0,84**
	3	0,99	0,99	1	1,04	1,01	1,03	1,01	0,99	1	0,99	0,99	0,98
CPI	6	0,99	0,98	1,06	1,07	1,01	1,01	0,94	1**	0,97	0,97	0,99	0,91
	9	0,9	0,9	1,02	1,01	1,01	1,01	0,93**	1	0,9	0,89	0,9	0,86*
	12	0,85*	0,85*	1,02	1,03	1,01	1,01	0,92***	1,02	0,83*	0,85*	0,85*	0,81**
	1	1,01	1,01	1,21**	1,83***	1,21**	1,83***	1	0,88*	1	0,98	1,01	0,59*
	3	0,99	0,99	0,99	0,98	0,91	1,02	0,99	1	0,99	0,97	0,99	1*
PCE	6	0,99	0,99	1	1	1	1	1	0,9**	1	0,98	0,99	0,99
	9	0,98	0,98	0,99	0,98	1	1	0,99	1	0,98	0,97	0,98	0,99
	12	1,02	1,02	1,04**	1,05**	1,01	1,01	1,01	1,01	1,01	0,99	1,02	0,99*
	1	1	1	1,08	1,23	1,08	1,23	1	0,94	1	1*	1	0,86
	3	1,03	1,03	1,03	1,04	1	1,11*	1,01	0,98**	1,01	1,03	1,03	0,97
PPI	6	1,02	1,02	1,02	1,03	1	1	0,99	0,99*	0,99	1,02	1,02	0,96
	9	0,98	0,98	1	1	1	1	0,99	0,99	0,92	0,97	0,98	0,94
	12	0,97	0,97	1	1	1	1	0,99**	0,95	0,89	0,97	0,97	0,93
	1	0,99	0,99	1,02	1,02	1,02	1,02	1	0,96	1	0,99	0,99	0,95
	3	0,95	0,95	1,15	1,17	1,04	1,14	1,01	1,02	1,01	1,01	0,95	1,07
RDI	6	0,79**	0,8*	1,14	1,15	1,09	1,13	0,84	1,02	0,88	0,87*	0,8*	1
	9	0,75**	0,75**	1,17	1,15	1,12	1,2	0,8**	0,99	0,87*	0,87*	0,75**	1,01
	12	0,63**	0,63**	1,16*	1,11**	1,16	1,24**	0,72*	1,01	0,75*	0,79*	0,63**	1**
	1	1,02*	1,02*	1,05	1,05	1,05	1,05	1	0,9**	1	1,04**	1,02**	1,05**
	3	1	1	1,17**	1,05	1	0,96	0,92	1,06***	0,92	1,03	1,01	0,99*
UR	6	0,82	0,82	1	0,88	0,98*	0,73	0,65	1	0,65	0,84	0,83	0,79
	9	0,61	0,61	0,72*	0,77	0,82	0,59**	0,54	1,07**	0,47	0,67	0,62	0,68
	12	0,52*	0,52*	0,64*	0,72	0,83	0,56**	0,51*	1,28*	0,39	0,62	0,53*	0,66**
	1	0,95***	0,96***	1	1,05	1	1,05	0,97	1,01*	1	1***	0,96***	1,22**
	3	0,85	0,85	1,07	1,09	1	0,93	0,97	0,94	0,91	0,94	0,85	1,17
IP	6	0,74	0,74	1,05	1,05	1,01	0,92	0,95	1,04	0,77	0,86	0,74	1,31**
	9	0,62	0,62	0,94**	0,99*	0,99**	0,84	0,91	1,07	0,61	0,74	0,62	1,34***
	12	0,61*	0,61*	0,91	0,62	1,14	0,89	1,02	1,52**	0,59*	0,77*	0,61*	1,62*

Gürkaynak et al. (2007) yield data. 7 maturities, from 1 to 7 years. MSFE ratios between model 8 and models $1-13$ for CPI, PCE, PPI, RDI, UR, and IP for all forecasting leads h. A value lower than one indicates a lower MSFE of model 8 w.r.t. models $1-13$. One, two, and three stars mean .10, .05, and .01 statistical significance, respectively, for the Giacomini and White (2006) test with quadratic loss function. The number of forecasts is $S' = 150$ (the first half of the $S = 300$ complete forecast sample). The 1-step ahead forecasts range from February 1987 to July 1999. The 12-step ahead forecasts range from January 1988 to June 2000. The forecast period corresponds to part of the Greenspan (1987-2006) monetary regime.

Table 3: MSFE ratios for second half of forecast sample (Gürkaynak et al. (2007) yield data, 7 maturities).

	h	*mod 1*	*mod 2*	*mod 3*	*mod 4*	*mod 5*	*mod 6*	*mod 7*	*mod 9*	*mod 10*	*mod 11*	*mod 12*	*mod 13*
	1	1,02	1,02	1,14	1,39	1,14	1,39	1	1,02	1	1,02	1,02	1,03
	3	1,19	1,19	1,15	1,27	1,14	1,37**	1,17	1,05	1,18	1,19	1,19	1,06
CPI	6	1,04	1,04	1,05	1,07	1,02	1,03	1,02	1,01	1,04	1,04	1,04	1,01
	9	1	**0,99**	1,01	1,02	1*	1	**0,98****	**0,95***	**0,98***	**0,99***	1	**0,98**
	12	1,03	1,02	**0,99**	**0,99**	1,01	1,01*	1,01	1,01	1,05	1,02*	1,03	**0,99**
	1	1	1	1,46**	1,57	1,46**	1,57	1	**0,76**	1	1	1	**0,61***
	3	1,02	1,01	1,01	1,01	**0,99**	1,01	1,01	**0,77*****	1,01	1,01	1,02	**0,99**
PCE	6	**0,99**	**0,99**	**0,99**	**0,99**	1,01	1,01	1	1,01	**0,98**	**0,98**	**0,99**	1
	9	**0,98**	**0,98**	**0,98***	**0,98**	1	1	**0,99**	**0,99**	**0,97****	**0,98**	**0,98**	**0,99**
	12	**0,99**	**0,99**	1	1	**0,99**	**0,99**	**0,99**	1	**0,98**	1	**0,99**	**0,98**
	1	1,02	1,02	1,16	1,31	1,16	1,31	1	**0,11**	1	1,02	1,02	**0,82****
	3	1,07	1,06	1,03	1,11	1,05	1,2	1,04	1,03	1,05	1,06	1,07	1,03
PPI	6	1,09	1,08	1,08	1,15*	1,02	1,02	1,04	1,02	1,07	1,08	1,09	1
	9	1,01	1,01	1,02	1,01	1**	1**	1	1	**0,99**	1,01	1,01	**0,97**
	12	1,06	1,06	1,09	1,1**	1	1	1,01	1	1,04	1,06	1,06	**0,97**
	1	1,01	1,01	1	1,03	1	1,03	1	**0,91**	1	1,01	1,01	**0,99**
	3	1,19	1,19	1,14	1,18	1,08	1,17	1,16	1,01	1,19	1,19	1,19	1,05*
RDI	6	1	1	**0,93**	**0,92**	**0,95**	**0,95**	**0,99**	**0,65*****	1,02	**0,98**	1	1
	9	**0,94**	**0,95**	**0,83**	**0,83**	**0,89***	**0,9**	**0,95**	**0,61*****	1,01	**0,91**	**0,94**	**0,96***
	12	**0,89**	**0,93**	**0,79**	**0,84**	**0,97**	**0,91**	**0,92**	**0,63*****	1,03	**0,86**	**0,89**	1,01
	1	**0,94**	**0,93**	1,2**	1,08**	1,2**	1,08**	1	**0,72*****	1	**0,95**	**0,94**	1,09***
	3	**0,93**	**0,93**	1,07	**0,91**	1	**0,68**	**0,95****	**0,77***	**0,96****	**0,93**	**0,93**	**0,74****
UR	6	**0,93**	**0,94**	**0,97**	1	1	**0,62**	**0,89**	**0,88**	**0,95****	**0,95**	**0,94**	**0,6*****
	9	**0,89**	**0,89**	**0,87**	**0,93**	**0,81****	**0,71****	**0,84***	1,08	**0,89***	**0,93**	**0,89**	**0,65****
	12	**0,91**	**0,92**	**0,87****	**0,97****	**0,73***	**0,85****	**0,81***	**0,79****	**0,88**	**0,95**	**0,92**	**0,74****
	1	**0,94**	**0,94**	1,12	1,03	1,12	1,03	**0,77****	**0,98**	1	**0,96**	**0,94**	**0,91**
	3	**0,84***	**0,85***	**0,95***	**0,89**	**0,87**	**0,6**	**0,5*****	**0,41*****	**0,89****	**0,86**	**0,84***	**0,54****
IP	6	**0,87**	**0,87**	**0,97**	1,14	**0,93**	**0,84**	**0,65****	**0,98**	**0,9****	**0,89**	**0,87**	**0,72*****
	9	1,07*	1,07	1,17	1,38	1,07	1,31	**0,88****	**0,86**	1,06**	1,1	1,07	1,04**
	12	1,25	1,25	1,44	1,52	1,07	1,7	1,02***	1,19**	1,2	1,29	1,25	1,32

Gürkaynak et al. (2007) yield data. 7 maturities, from 1 to 7 years. MSFE ratios between model 8 and models $1-13$ for CPI, PCE, PPI, RDI, UR, and IP for all forecasting leads h. A value lower than one indicates a lower MSFE of model 8 w.r.t. models $1-13$. One, two, and three stars mean .10, .05, and .01 statistical significance, respectively, for the Giacomini and White (2006) test with quadratic loss function. The number of forecasts is $S' = 150$ (the second half of the $S = 300$ complete forecast sample). The 1-step ahead forecasts range from August 1999 to January 2012. The 12-step ahead forecasts range from July 2000 to December 2012. The forecast period corresponds to part of the Greenspan (1987-2006) and Bernanke (2006-2012) monetary regimes.

Table 4: MSFE ratios for whole forecast sample (Gürkaynak et al. (2007) yield data, 30 maturities).

	h	*mod 1*	*mod 2*	*mod 3*	*mod 4*	*mod 5*	*mod 6*	*mod 7*	*mod 9*	*mod 10*	*mod 11*	*mod 12*	*mod 13*
	1	1,02	1	1,11	1,39	1,11	1,39	1	1	1	**0,99**	1,02	1,09
	3	1,2	1,18	1,17	1,28	1,15	1,44**	**0,71**	1	1,21	1	1,2	1,1
CPI	6	1,08	1,06	1,08	1,09	1,06	1,06	**0,81**	**0,99**	1,08	**0,99**	1,08	1,07
	9	1,03	1,02	**0,95**	**0,92**	1,01*	1,01	**0,69**	1,01	1*	1	1,03	1,02
	12	1,04**	1,03**	**0,95**	**0,89**	1,02	1,02	**0,81**	1,01	1,08*	**0,99***	1,05**	1,02
	1	1,01	1,01	1,57*	1,86	1,57*	1,86	1	**0,98**	1,01	**0,98**	1,01	**0,61**
	3	1,01	1,01	1,01	1,01	**0,99**	1	**0,96***	**0,87**	1,01	**0,91**	1,01	1
PCE	6	**0,97**	**0,97**	**0,97**	**0,96**	1,01	1,01	**0,98**	**0,98**	**0,95**	1	**0,97**	1,01
	9	**0,99**	1	1	1	1	1	**0,95***	**0,95**	**0,96**	**0,97**	**0,99**	1
	12	**0,97**	**0,97**	**0,88**	**0,96**	1	1	**0,92**	1,04	**0,95**	**0,99**	**0,97**	1
	1	1,03	1,02	1,09	1,25	1,09	1,25	1	1	1	**0,99**	1,02	**0,83***
	3	1,07	1,06	1,03	1,13	1,02	1,22	**0,68****	**0,99**	1,06	**0,98**	1,07	1,06
PPI	6	1,09	1,09	1,07	1,17	1,05	1,04	**0,87**	**0,98**	1,05	**0,97**	1,09	1,05
	9	1,02	1,03	**0,98**	1,01	1,01	1,01	**0,76**	1	**0,96**	1	1,02	1,01
	12	1,07	1,08	1,06*	1,14**	1,02	1,02	**0,84**	**0,99**	1,03	1,02	1,07	1,02
	1	1,02	1,02	**0,99**	1,08	**0,99**	1,08	1	**0,99**	1	1,06	1,02	1,01
	3	1,22*	1,24*	1,18	1,21	1,13	1,3	1,04	**0,98**	1,21	1	1,22*	**0,97*****
RDI	6	**0,96**	**0,98**	**0,94**	**0,95**	**0,96**	**0,99**	**0,75****	1,02	1,01	1,01	**0,97**	**0,67*****
	9	**0,79***	**0,83**	**0,75**	**0,79**	**0,86**	**0,87**	**0,29***	1,01	**0,8****	1	**0,8***	**0,54****
	12	**0,66****	**0,69****	**0,78**	**0,83**	**0,88**	**0,89**	**0,99*****	**0,99**	**0,68*****	**0,99**	**0,66****	**0,48**
	1	**0,95**	1,06	1,14*	**0,98**	1,14*	**0,98**	1	1	**0,99****	**0,93**	**0,95**	**0,99***
	3	**0,93**	**0,98**	**0,95**	**0,78**	**0,88**	**0,62**	**0,37*****	**0,96**	**0,94**	1,04	**0,9**	**0,67**
UR	6	**0,97**	**0,97**	**0,93**	1,22	**0,9**	**0,64**	**0,34****	**0,92**	**0,94**	**0,94****	**0,96**	**0,7**
	9	**0,94***	**0,93**	**0,89**	**0,99**	**0,81**	**0,78**	**0,24***	1,01	**0,88**	**0,96**	**0,95***	**0,87**
	12	1,06*	1,05*	**0,96**	1,18	**0,77**	1,01*	**0,44**	**0,93**	**0,91**	**0,93**	1,07	1,1
	1	**0,99**	1,06	1,11	**0,96**	1,11	**0,96**	1	1,04	1	1	**0,98**	**0,98**
	3	**0,96**	**0,95**	**0,99**	**0,92**	**0,96**	**0,65**	**0,8***	**0,98**	**0,96**	**0,98**	**0,97**	**0,66**
IP	6	**0,89**	**0,87**	**0,93**	1,31	**0,93**	**0,93**	**0,78**	1,02	**0,88**	**0,97**	**0,9**	1,02
	9	1,14	1,1	1,15	1,56	1,06	1,46	**0,79**	1,02	**0,98**	**0,97**	1,15*	1,65
	12	1,37	1,27	1,44	1,66	1,04	1,84	**0,9**	1,03	1,04	**0,96**	1,36	2,13

Gürkaynak et al. (2007) yield data. 30 maturities, from 1 to 30 years. MSFE ratios between model 8 and competing models for CPI, PCE, PPI, RDI, UR, and IP for all forecasting leads h. A value lower than one indicates a lower MSFE of model 8 w.r.t. models $1-13$. One, two, and three stars mean .10, .05, and .01 statistical significance, respectively, for the Giacomini and White (2006) test with quadratic loss function. The number of forecasts is $S = 125$. The 1-step ahead forecasts range from September 2001 to January 2012. The 12-step ahead forecasts range from August 2002 to December 2012. The forecast period corresponds to part of the Greenspan (1987-2006) and Bernanke (2006-2012) monetary regimes.

Table 5: MSFE ratios for first half of forecast sample (Gürkaynak et al. (2007) yield data, 30 maturities).

	h	*mod 1*	*mod 2*	*mod 3*	*mod 4*	*mod 5*	*mod 6*	*mod 7*	*mod 9*	*mod 10*	*mod 11*	*mod 12*	*mod 13*
	1	1	1	1,26	1,56	1,26	1,56	1	1,01	1	**0,99**	1	1,1
	3	1,15	1,15	1,15	1,19	1,1	1,25*	1,1	1	1,14	1	1,15	1,02
CPI	6	**0,98***	**0,98***	**0,97***	**0,96**	1	1	1	1,01	**0,97**	1	**0,98***	1
	9	1,01	1,01	1,09	1,08**	1,01	1,01	**0,97**	1,01	**0,99**	1,01	1,01	1,01
	12	1,17***	1,17***	1,3*	1,16*	1,05**	1,05**	1,06	1	1,15**	**0,99**	1,17***	1,06**
	1	1,01	1,01	1,66	2,27*	1,66	2,27*	1	**0,99**	1,01	1	1,01	**0,58**
	3	1,02	1,02	1,02	1,02	1,01	1	1	**0,85**	1,02	**0,91**	1,02	1
PCE	6	**0,97**	**0,97**	**0,96**	**0,95**	1,01	1	**0,97**	**0,98**	**0,96**	1	**0,97**	1,01
	9	1	1,01	1,01	1,01	1	1	**0,96**	**0,94**	**0,99**	**0,98**	1	1
	12	**0,98**	**0,98**	**0,86**	**0,96**	1	1*	**0,9***	1,06*	**0,98**	1,01	**0,98**	1
	1	1	**0,99**	1,34**	1,49***	1,34**	1,49***	1	1,02	1	**0,98**	1	**0,91****
	3	1	1	**0,96**	**0,98**	**0,93**	1,01	**0,93**	**0,99**	1	1	1	1,01
PPI	6	**0,98**	**0,98**	**0,95**	**0,97**	1	1	**0,97**	**0,99**	**0,97**	**0,98**	**0,97**	1
	9	1,03	1,03	1,03	1,04	1,02	1,02	1,05	**0,97**	1,02	1,02	1,03	1,02
	12	1,07	1,07	1,14	1,08	1,05	1,05	1,02	1	1,05	1	1,07	1,05
	1	1,01*	1,01**	1,04	1,11	1,04	1,11	1	**0,99**	1	1,09	1,01**	**0,96**
	3	1,27*	1,28	1,34	1,31	1,36*	1,39	1,38	**0,93****	1,26	**0,98**	1,27	**0,82*****
RDI	6	1,12	1,06	1,28**	1,31*	1,13	1,23*	1,13	1,05	1,13	1,04	1,12	**0,45*****
	9	1,04	1,01	1,47***	1,46***	1,15***	1,27**	**0,21**	1,02	1,04	1,05	1,04	**0,3*****
	12	**0,76**	**0,75**	1,64***	1,54***	1,13**	1,36***	**0,99*****	**0,99**	**0,77**	**0,98*****	**0,76**	**0,23*****
	1	**0,99**	**0,98**	1,05	**0,98****	1,05	**0,98****	1	1	1	**0,99**	**0,99**	1,08
	3	**0,94****	**0,93*****	**0,96**	**0,95**	1,02***	**0,83**	**0,63****	**0,98**	**0,97***	1,09*	**0,95****	**0,89**
UR	6	**0,82*****	**0,81*****	**0,85**	1,06	**0,76**	**0,87**	**0,84****	**0,95***	**0,89****	**0,92***	**0,83*****	**0,94**
	9	**0,69*****	**0,69*****	**0,79**	**0,63*****	**0,51***	**0,75****	**0,64***	**0,94**	**0,91****	**0,95**	**0,7*****	**0,8*****
	12	**0,58****	**0,58****	**0,7**	**0,5***	**0,27**	**0,61****	**0,77****	**0,98****	**0,88****	**0,93***	**0,59*****	**0,62*****
	1	1,03	1,02	1,07	1,06	1,07	1,06	1	**0,98**	1,02*	1,02*	1,03	1,07
	3	1,04	1,04	1,2	1,21*	1,04	1,35*	1,11	**0,96**	1,08	1,02	1,04	1,12
IP	6	**0,93**	**0,93**	1,36**	1,6**	1,09	2,09*	1,32***	1	1,06	**0,94***	**0,92**	1,09
	9	**0,82**	**0,84**	1,82	1,85***	1,16	3,09*	1,22**	1	1,03	1,02	**0,82**	1,05
	12	**0,54**	**0,61**	2,94	2,48**	**0,72**	3,53**	**0,94*****	1	**0,65**	1	**0,55**	**0,83**

Gürkaynak et al. (2007) yield data. 30 maturities, from 1 to 30 years. MSFE ratios between model 8 and competing models for CPI, PCE, PPI, RDI, UR, and IP for all forecasting leads h. A value lower than one indicates a lower MSFE of model 8 w.r.t. models $1-13$. One, two, and three stars mean .10, .05, and .01 statistical significance, respectively, for the Giacomini and White (2006) test with quadratic loss function. The number of forecasts is $S' = 62$ (the first half of the $S = 125$ complete forecast sample). The 1-step ahead forecasts range from September 2001 to October 2006. The 12-step ahead forecasts range from August 2002 to September 2007. The forecast period corresponds to part of the Greenspan (1987-2006) and Bernanke (2006-2012) monetary regimes.

Table 6: MSFE ratios for second half of forecast sample (Gürkaynak et al. (2007) yield data, 30 maturities).

	h	*mod 1*	*mod 2*	*mod 3*	*mod 4*	*mod 5*	*mod 6*	*mod 7*	*mod 9*	*mod 10*	*mod 11*	*mod 12*	*mod 13*
CPI	1	1,04	**0,99**	1	1,27	1	1,27	1	1	1	1	1,03	1,09
	3	1,24	1,21	1,18	1,35	1,2	1,62	**0,57***	1	1,26	**0,99**	1,23	1,15
	6	1,12	1,1	1,14	1,16	1,09	1,1	**0,75**	**0,99**	1,13	**0,99**	1,12	1,1
	9	1,04	1,02	**0,89**	**0,85**	1,01**	1,01	**0,6**	1	1***	1	1,04	1,02
	12	**0,99***	**0,98****	**0,84**	**0,79**	1	1	**0,73**	1,01	1,05	**0,99**	**0,99***	1
PCE	1	1	1	1,3	1,07	1,3	1,07	1	**0,95**	1	**0,92**	1	**0,74**
	3	**0,97**	**0,97**	**0,97**	**0,98**	**0,92****	**0,99**	**0,81****	**0,96**	**0,96**	**0,94**	**0,97**	**0,99**
	6	**0,98**	**0,97**	1	**0,99**	1,02*	1,02	1,02	**0,97**	**0,92**	1,04	**0,98**	1,03*
	9	**0,95**	**0,96****	**0,97**	**0,97**	1	1	**0,9***	**0,99**	**0,88**	**0,95**	**0,95**	1,01*
	12	**0,94**	**0,95**	**0,95**	**0,96**	**0,99**	**0,99**	1	**0,99**	**0,85***	**0,94**	**0,94**	1***
PPI	1	1,05	1,04	**0,98**	1,13	**0,98**	1,13	1	**0,99**	1	**0,99**	1,04	**0,78***
	3	1,12	1,1	1,07	1,22	1,08	1,37	**0,59*****	**0,99**	1,09	**0,97**	1,12	1,09
	6	1,15	1,14	1,13**	1,28*	1,07	1,06	**0,83**	**0,98**	1,09	**0,97**	1,15	1,07
	9	1,01	1,04	**0,97**	**0,99**	1,01	1,01	**0,68***	1,01	**0,93**	1	1,01	1,01
	12	1,07	1,08	1,03	1,16***	1,01	1,01	**0,78**	**0,98**	1,01	1,03	1,07	1,01
RDI	1	1,02	1,02	**0,94**	1,05	**0,94**	1,05	1	**0,99**	1	1,04	1,03	1,06
	3	1,17	1,21	1,08	1,15	**0,99**	1,24	**0,87**	1,03	1,17	1,01	1,19	1,12
	6	**0,86****	**0,92**	**0,78**	**0,78**	**0,87**	**0,86**	**0,6****	1	**0,93**	**0,98****	**0,88****	1,08
	9	**0,69****	**0,75**	**0,58**	**0,62*****	**0,75*****	**0,74*****	**0,38****	1	**0,7*****	**0,97**	**0,7****	1,03*
	12	**0,61****	**0,67**	**0,62****	**0,67****	**0,79*****	**0,75*****	**0,99*****	**0,99**	**0,64*****	**0,99**	**0,61****	1,06
UR	1	**0,94**	1,11	1,2	**0,98**	1,2	**0,98**	1	1	**0,98*****	**0,91***	**0,92**	**0,94***
	3	**0,92**	1	**0,95**	**0,72**	**0,83**	**0,56**	**0,32*****	**0,96**	**0,92***	1,02	**0,88**	**0,6**
	6	1,02	1,02	**0,96**	1,26	**0,94**	**0,6**	**0,3*****	**0,91**	**0,96**	**0,94**	1	**0,66**
	9	**0,98**	**0,97**	**0,91**	1,06	**0,87***	**0,79***	**0,23****	1,02*	**0,87**	**0,96**	**0,99**	**0,87**
	12	1,11**	1,1**	**0,98**	1,27	**0,86***	1,05**	**0,43***	**0,93**	**0,91*****	**0,93***	1,12	1,16*
IP	1	**0,97**	1,07	1,13	**0,92**	1,13	**0,92**	1	1,06	**0,99*****	1	**0,96**	**0,95**
	3	**0,93**	**0,92****	**0,93**	**0,85**	**0,94**	**0,55**	**0,73****	**0,99**	**0,93**	**0,97**	**0,94**	**0,58**
	6	**0,89**	**0,87**	**0,89**	1,27	**0,91**	**0,86**	**0,73**	1,02	**0,86***	**0,98**	**0,9**	1,01**
	9	1,17*	1,13*	1,12**	1,54	1,06	1,4	**0,77**	1,02	**0,98**	**0,96**	1,18**	1,72
	12	1,47	1,34	1,41	1,64	1,06	1,81	**0,9**	1,04	1,07	**0,95**	1,46	2,29

Gürkaynak et al. (2007) yield data. 30 maturities, from 1 to 30 years. MSFE ratios between model 8 and competing models for CPI, PCE, PPI, RDI, UR, and IP for all forecasting leads h. A value lower than one indicates a lower MSFE of model 8 w.r.t. models $1-13$. One, two, and three stars mean .10, .05, and .01 statistical significance, respectively, for the Giacomini and White (2006) test with quadratic loss function. The number of forecasts is $S' = 63$ (the second half of the $S = 125$ complete forecast sample). The 1-step ahead forecasts range from November 2006 to January 2012. The 12-step ahead forecasts range from October 2007 to December 2012. The forecast period corresponds to part of the Bernanke (2006-2012) monetary regime.

Table 7: MSFE ratios for whole forecast sample (Diebold and Li (2006) yield data, 17 maturities).

	h	*mod 1*	*mod 2*	*mod 3*	*mod 4*	*mod 5*	*mod 6*	*mod 7*	*mod 9*	*mod 10*	*mod 11*	*mod 12*	*mod 13*
	1	1	0,98	1,12*	1,28***	1,12*	1,28***	1	1	1	1	1,01	0,84***
	3	1	0,99	1,01*	1,01	1,02	1,02	0,98	0,99	0,99	0,99	1,01*	1
CPI	6	0,96	0,94	1,06	1,1	1,02	1,02	0,98*	1,02*	0,98	0,96	1,07	0,96
	9	0,87	0,88**	0,99	1,03	1,02	1,02	0,96***	1,01**	0,91***	0,87*	1,02	0,89***
	12	0,78	0,85**	1,04	1,06	1,01**	1,01**	0,98	0,97	0,88**	0,81**	1,02	0,85**
	1	1,01	1,01	1,32*	1,69***	1,32*	1,69***	1	1,03	1	1,01	0,96	0,8
	3	1,06	1,05	1,06	1,06	1,03	1,02	0,99	1,01*	1,04	1,06	1,01	1
PCE	6	0,96	0,96	0,97	0,98	1	1	1,01	1	0,94	0,97	0,92	1
	9	0,95	0,94*	0,95	0,94	1	1	0,96**	1	0,92**	0,95	0,9**	0,99
	12	1	0,97	1,03*	1,02*	1,01	1,01	1	1	0,95*	1	0,97***	0,97*
	1	1	0,99**	1,08	1,27*	1,08	1,27*	1	1,13	1	1,01	1	0,87
	3	1,03	1,03	1,03	1,03	1,01	1,1	1	0,91**	1,02	1,02	1,03	0,98
PPI	6	1,07	1,06	1,08	1,07	1,01**	1,01**	1	1,01	1,07	1,07	1,08	1
	9	0,97	0,96	0,96	1	1,01	1,01	0,99	1,01	0,97	0,98	0,98	0,98
	12	0,93	0,94	0,97	0,98*	0,99	0,99	0,99	1	0,94*	0,94*	0,96	0,96
	1	0,99	0,98*	0,97	1,03	0,97	1,03	1	0,77	1	0,98	1,03	0,94
	3	0,94	0,96	1,05	1,05	0,97	1,1	0,99	0,98	0,99	0,9	1,05	1,04
RDI	6	0,9	0,92	1,12**	1,15*	1,1	1,14*	1,04	0,97	0,88	0,8	1,12	1
	9	0,88**	0,9	1,19*	1,18*	1,14	1,24*	1,02	0,97	0,83	0,79***	1,18*	1,03
	12	0,77	0,79	1,32**	1,2**	1,22	1,41**	1,03	0,97	0,77*	0,73**	1,29**	1,02
	1	1	1	1,05	1,05	1,05	1,05	1	0,99*	1	0,98*	0,99	1,04**
	3	0,93	0,91	1,11	1,05	1,02	1,03	0,99	0,84**	0,86*	0,85*	0,98**	0,85**
UR	6	0,88	0,86	1,1	1,04	1,03**	0,95	1	1,07*	0,82	0,82	0,92	0,75
	9	0,66	0,65	0,93	0,92	0,89	0,84	1	0,92	0,61	0,62	0,77	0,65
	12	0,52	0,5	0,78	0,8	0,79	0,76	1,01	0,89	0,5	0,5	0,67**	0,63
	1	1,01	1,02	1,08	1,11*	1,08	1,11*	1	0,99	1	1	1,01	1,07**
	3	0,98	0,96	1,09*	1,12*	1,04*	1,02	0,98*	1,05	0,87	0,87	1,16**	0,93
IP	6	0,98	0,96	1,24	1,19	1,15	1,14	0,97	0,54**	0,9	0,88	1,23	1,06
	9	0,84	0,82	1	1,13	1,06	1,02	0,95	1,07	0,79	0,79	1,08	1,03
	12	0,62**	0,64***	0,67	0,65	0,99	0,87	0,95***	1,02	0,63**	0,63***	0,8	0,96

Diebold and Li (2006) yield data. 17 maturities corresponding to 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108, and 120 months. MSFE ratios between model 8 and models $1-13$ for CPI, PCE, PPI, RDI, UR, and IP for all forecasting leads h. A value lower than one indicates a lower MSFE of model 8 w.r.t. the competing models. One, two, and three stars mean .10, .05, and .01 statistical significance, respectively, for the Giacomini and White (2006) test with quadratic loss function. The number of forecasts is $S = 150$. The 1-step ahead forecasts range from July 1987 to December 1999. The 12-step ahead forecasts range from June 1988 to November 2000. The forecast period corresponds to part of the Greenspan (1987-2006) monetary regimes.

Table 8: MSFE ratios for first half of forecast sample (Diebold and Li (2006) yield data, 17 maturities).

	h	mod 1	mod 2	mod 3	mod 4	mod 5	mod 6	mod 7	mod 9	mod 10	mod 11	mod 12	mod 13
	1	1	1*	1,07	1,21*	1,07	1,21*	1	1	1	1	1	0,84**
	3	1,02	1,03	1,05	1,06	1,05	1,08	0,99	1	1,01	1,02	1,03	1,02
CPI	6	0,95	0,97	1,09	1,13	1,02	1,02	0,98	1,01	0,99	0,96	1,08	0,99
	9	0,81**	0,89**	0,95	1,01	1,02	1,02	0,94**	1	0,89*	0,82**	1	0,95
	12	0,71**	0,87**	1,03	1,03	0,99***	0,99***	0,97	0,97	0,89***	0,76**	0,99	0,93***
	1	1,01	1,01	1,31	1,69**	1,31	1,69**	1	1,01	1	1,01	0,99	0,84
	3	1,07	1,06	1,07	1,08	1,05	1,02	0,99	1*	1,05	1,07	1,05	1*
PCE	6	0,96	0,95	0,97	0,98	1	1	0,99	1	0,94	0,96	0,94	1
	9	0,94*	0,93**	0,95	0,94	1	1	0,98	1	0,91**	0,94	0,93	0,99
	12	0,99*	0,96*	1,02	1,02	1,01	1,01	1	1,01*	0,93***	0,99**	0,99*	0,96**
	1	1	1	1,06	1,26	1,06	1,26	1	1,14	1	1,01	0,99	0,88
	3	1,04	1,04	1,04	1,04	1,01	1,15	1*	0,93*	1,02	1,03	1,03	0,99
PPI	6	1,09	1,09	1,1	1,08	1,01*	1,01*	1,01	1,01	1,09	1,1	1,08	1,02***
	9	0,97*	0,98	0,94	0,99	1	1	0,99*	1,01	0,99	0,98	0,97	1,01
	12	0,91**	0,95	0,95	0,95**	0,98	0,98	1	1	0,95	0,94**	0,93	1
	1	0,99	1	0,92	0,99	0,92	0,99	1	0,63	1	0,97	1	0,96
	3	1	1,02	0,96	0,97	0,86	1,03	0,95	0,99	1,01	0,89	0,97	1,11
RDI	6	0,92	0,91	0,96	0,99	0,97	1,01	0,97	0,99	0,84	0,73	0,96	1,05
	9	0,98	0,95	1,02	1,01	0,98	1,08	0,91***	1,01	0,8	0,76	0,99	1,12
	12	0,9	0,88	1,06	0,91	0,95	1,15*	0,92***	1,1	0,74	0,7	1,01	1,1
	1	1,03**	1,02	1,12	1,13	1,12	1,13	1	1,01	1**	0,98	1,07	1,13***
	3	0,85***	0,82***	1,22***	1,16***	1,06	0,94	1,03	0,75**	0,74**	0,75**	1,04	0,63**
UR	6	0,78	0,77	1,2**	1,11**	1,1***	1,01	1,04	1,17**	0,71	0,72*	1,05**	0,6
	9	0,55*	0,54*	0,93	0,96	0,88	0,89	1,05**	0,99**	0,51*	0,52*	0,86	0,51
	12	0,41	0,4	0,73	0,74**	0,73	0,76	1,07	0,8*	0,4	0,4*	0,69**	0,49
	1	1,04*	1,04*	1,09	1,12	1,09	1,12	1	1,01	1**	1	1,12	1,05
	3	0,91	0,89*	1,2**	1,3***	1,25***	1,23	0,96	1,07	0,77**	0,78	1,35***	0,85
IP	6	0,96	0,94	1,59***	1,57***	1,38**	1,99**	0,95**	0,53*	0,85	0,85	1,71***	1,06
	9	0,84*	0,83**	1,54**	1,6**	1,32**	2,09**	0,92*	1,11	0,76*	0,78	1,82***	0,99**
	12	0,65**	0,71**	1,26	1,24*	1,43	1,89**	0,91**	0,97	0,63**	0,65**	1,52	0,92**

Diebold and Li (2006) yield data. 17 maturities corresponding to 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108, and 120 months. MSFE ratios between model 8 and models $1-13$ for CPI, PCE, PPI, RDI, UR, and IP for all forecasting leads h. A value lower than one indicates a lower MSFE of model 8 w.r.t. the competing models. One, two, and three stars mean .10, .05, and .01 statistical significance, respectively, for the Giacomini and White (2006) test with quadratic loss function. The number of forecasts is $S' = 75$ (the first half of the $S = 150$ complete forecast sample). The 1-step ahead forecasts range from July 1987 to September 1993. The 12-step ahead forecasts range from June 1988 to August 1994. The forecast period corresponds to part of the Greenspan (1987-2006) monetary regime.

Table 9: MSFE ratios for second half of forecast sample (Diebold and Li (2006) yield data, 17 maturities).

	h	*mod 1*	*mod 2*	*mod 3*	*mod 4*	*mod 5*	*mod 6*	*mod 7*	*mod 9*	*mod 10*	*mod 11*	*mod 12*	*mod 13*
CPI	1	1,01	__0,96__	1,2	1,37**	1,2	1,37**	1	1	1	1	1,02	__0,85**__
	3	__0,96__	__0,95*__	__0,97__	__0,95__	__0,97__	__0,94__	__0,96__	__0,98*__	__0,97*__	__0,96__	__0,98__	__0,96__
	6	1	__0,89__	1,01	1,03	1,02	1,02	1*	1,03	__0,97__	__0,98__	1,05**	__0,9__
	9	1	__0,87__	1,06**	1,05**	1,02	1,02	__0,98*__	1,02***	__0,93*__	__0,97__	1,06***	__0,8***__
	12	__0,96__	__0,82__	1,08	1,13	1,06	1,06	__0,99__	__0,98__	__0,85__	__0,9__	1,06	__0,75__
PCE	1	1	1	1,37	1,68**	1,37	1,68**	1	1,06	1	1,01	__0,88*__	__0,69*__
	3	1,03	1,02	1,03*	1,04*	__0,97__	1,01	1	1,01	1,01	1,03	__0,9__	1,01
	6	__0,97__	__0,97__	__0,96__	__0,98__	1**	1**	1,07	__0,99__	__0,96**__	__0,98__	__0,86__	1
	9	__0,97__	__0,97__	__0,96__	__0,95__	1	1	__0,92**__	__0,99__	__0,95*__	__0,97__	__0,84**__	1,01
	12	1,03	1,02	1,04*	1,04*	1,01	1,01	1,01	__0,98**__	1	1,03	__0,9__	__0,99__
PPI	1	1,01	__0,96**__	1,13	1,3	1,13	1,3	1	1,1	1	1,01	1,02	__0,83**__
	3	1,01	1	1,01	__0,99__	1,01	__0,97__	__0,98__	__0,86*__	1,01	1,01	1,02	__0,96__
	6	1,02	__0,99__	1,04	1,05	1,02	1,02	__0,97__	1,02	__0,99__	1	1,05	__0,93__
	9	__0,99__	__0,92__	1,01	1,02	1,03	1,03	1	1	__0,93__	__0,95__	1,01*	__0,91__
	12	__0,99__	__0,91__	1,04	1,05	1,03	1,03	__0,97*__	1	__0,9__	__0,93__	1,03	__0,86__
RDI	1	__0,98__	__0,96***__	1,04	1,08	1,04	1,08	1	1,01	1	__0,99__	1,07	__0,93__
	3	__0,88*__	__0,88*__	1,22	1,17	1,18	1,21	1,05**	__0,96__	__0,95__	__0,92*__	1,18	__0,96**__
	6	__0,87*__	__0,92__	1,4	1,43	1,32	1,34	1,13**	__0,95__	__0,94__	__0,9*__	1,38	__0,94*__
	9	__0,79*__	__0,84__	1,46	1,42	1,37	1,49	1,15**	__0,93__	__0,87__	__0,83__	1,47	__0,94*__
	12	__0,69**__	__0,72__	1,64*	1,63*	1,57	1,72**	1,12***	__0,89*__	__0,79__	__0,75*__	1,65*	__0,96__
UR	1	__0,98__	__0,99__	__0,99__	__0,99__	__0,99__	__0,99__	1	__0,98*__	1*	__0,98**__	__0,92__	__0,96__
	3	1,04	1,04	1,01	__0,95__	__0,97**__	1,15	__0,94**__	__0,97__	1,04	__0,99__	__0,92*__	1,4***
	6	1,16*	1,14	__0,94__	__0,92__	__0,91__	__0,85__	__0,92***__	__0,91*__	1,15**	1,11	__0,73__	1,45*
	9	1,19**	1,14**	__0,94***__	__0,84__	__0,91__	__0,76__	__0,91**__	__0,79__	1,17***	1,11	__0,63__	1,54***
	12	1,06	1	__0,88***__	__0,97***__	__0,98__	__0,76__	__0,91*__	1,12	1,06	1,03	__0,63__	1,53**
IP	1	__0,98__	1,01	1,07	1,09	1,07	1,09	1	__0,98__	1,01	__0,99__	__0,91__	1,09
	3	1,15	1,14	__0,92__	__0,88__	__0,78**__	__0,77__	1,01	1	1,16	1,14	__0,9__	1,14
	6	1,07	1,01	__0,7__	__0,65__	__0,74__	__0,46*__	1,02	__0,57__	1,1	1,04	__0,63__	1,08
	9	__0,86__	__0,79__	__0,51*__	__0,63*__	__0,69*__	__0,42**__	1,04	__0,96__	__0,89__	__0,84__	__0,5*__	1,13
	12	__0,57__	__0,53**__	__0,34**__	__0,33*__	__0,61*__	__0,42*__	1,06***	1,15*	__0,62**__	__0,6*__	__0,4**__	1,08

Diebold and Li (2006) yield data. 17 maturities corresponding to 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108, and 120 months. MSFE ratios between model 8 and models $1-13$ for CPI, PCE, PPI, RDI, UR, and IP for all forecasting leads h. A value lower than one indicates a lower MSFE of model 8 w.r.t. the competing models. One, two, and three stars mean .10, .05, and .01 statistical significance, respectively, for the Giacomini and White (2006) test with quadratic loss function. The number of forecasts is $S' = 75$ (the second half of the $S = 150$ complete forecast sample). The 1-step ahead forecasts range from September 1994 to December 1999. The 12-step ahead forecasts range from September 1994 to November 2000. The forecast period corresponds to part of the Greenspan (1987-2006) monetary regime.

Research Papers 2013

CREATES
Center for Research in Econometric Analysis of Time Series

2015-22:	Girum D. Abate and Niels Haldrup: Space-time modeling of electricity spot prices
2015-23:	Eric Hillebrand, Søren Johansen and Torben Schmith: Data revisions and the statistical relation of global mean sea-level and temperature
2015-24:	Tommaso Proietti and Alessandra Luati: Generalised partial autocorrelations and the mutual information between past and future
2015-25:	Bent Jesper Christensen and Rasmus T. Varneskov: Medium Band Least Squares Estimation of Fractional Cointegration in the Presence of Low-Frequency Contamination
2015-26:	Ulrich Hounyo and Rasmus T. Varneskov: A Local Stable Bootstrap for Power Variations of Pure-Jump Semimartingales and Activity Index Estimation
2015-27:	Rasmus Søndergaard Pedersen and Anders Rahbek: Nonstationary ARCH and GARCH with t-distributed Innovations
2015-28:	Tommaso Proietti and Eric Hillebrand: Seasonal Changes in Central England Temperatures
2015-29:	Laurent Callot and Johannes Tang Kristensen: Regularized Estimation of Structural Instability in Factor Models: The US Macroeconomy and the Great Moderation
2015-30:	Davide Delle Monache, Stefano Grassi and Paolo Santucci de Magistris: Testing for Level Shifts in Fractionally Integrated Processes: a State Space Approach
2015-31:	Matias D. Cattaneo, Michael Jansson and Whitney K. Newey: Treatment Effects with Many Covariates and Heteroskedasticity
2015-32:	Jean-Guy Simonato and Lars Stentoft: Which pricing approach for options under GARCH with non-normal innovations?
2015-33:	Nina Munkholt Jakobsen and Michael Sørensen: Efficient Estimation for Diffusions Sampled at High Frequency Over a Fixed Time Interval
2015-34:	Wei Wei and Denis Pelletier: A Jump-Diffusion Model with Stochastic Volatility and Durations
2015-35:	Yunus Emre Ergemen and Carlos Velasco: Estimation of Fractionally Integrated Panels with Fixed Effects and Cross-Section Dependence
2015-36:	Markku Lanne and Henri Nyberg: Nonlinear dynamic interrelationships between real activity and stock returns
2015-37:	Markku Lanne and Jani Luoto: Estimation of DSGE Models under Diffuse Priors and Data-Driven Identification Constraints
2015-38:	Lorenzo Boldrini and Eric Hillebrand: Supervision in Factor Models Using a Large Number of Predictors
2015-39:	Lorenzo Boldrini and Eric Hillebrand: The Forecasting Power of the Yield Curve, a Supervised Factor Model Approach