DEPARTMENT OF ECONOMICS AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

CREATES
Center for Research in Econometric Analysis of Time Series

Supervision in Factor Models Using a Large Number of Predictors

Lorenzo Boldrini and Eric Hillebrand

CREATES Research Paper 2015-38

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

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

Supervision in Factor Models Using a Large Number of Predictors

Lorenzo Boldrini* and Eric T. Hillebrand†

August 24, 2015

Abstract

In this paper we investigate the forecasting performance of a particular factor model (FM) in which the factors are extracted from a large number of predictors. We use a semi-parametric state-space representation of the FM in which the forecast objective, as well as the factors, is included in the state vector. The factors are informed of the forecast target (supervised) through the state equation dynamics. We propose a way to assess the contribution of the forecast objective on the extracted factors that exploits the Kalman filter recursions. We forecast one target at a time based on the filtered states and estimated parameters of the state-space system. We assess the out-of-sample forecast performance of the proposed method in a simulation study and in an empirical application, comparing its forecasts to the ones delivered by other popular multivariate and univariate approaches, e.g. a standard dynamic factor model with separate forecast and state equations.

Keywords: state-space system, Kalman filter, factor model, supervision, forecasting.
JEL classification: C32, C38, C55.

*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 availability of large datasets, the increase in computational power, and the ease of implementation have made factor models an appealing tool in forecasting. Factor models offer several advantages over other forecasting methods. For example, they do not require the choice of the variables to include in the forecasting scheme (as structural models do), they make use of a large information set, they allow to concentrate the information in all the candidate predictors in a relatively small number of factors, and they can be estimated with simple and fast methods. Using many predictors also allows to avoid the structural instability typical of low-dimensional systems. As argued for instance in Stock and Watson (2006) and Stock and Watson (2002a), also practitioners typically examine a large number of variables when making forecasts.

Forecasting using factor models is usually carried out in a two-step procedure, as suggested for instance by Stock and Watson (2002b). In the first step the factors are estimated using a set of predictors (that may include the lags of the forecast target) and in a second step the estimated factors are used to forecast the target by means of a forecast equation. In the two-step forecasting procedure suggested in Stock and Watson (2002b) however, the same factors are used to forecast different targets. That is, the selection of the factors is not supervised by the forecast target. In this paper we study a method to supervise the factor extraction for the forecast objective in order to improve on the predictive power of factor models. In the supervised framework, the factors are informed of the forecast target (supervised) through the state equation dynamics. Furthermore, we propose a way to assess the contribution of the forecast objective on the extracted factors that exploits the Kalman filter recursions.

The forecasting properties of static, restricted, and general dynamic factor models have been widely studied in the literature. Some examples are Boivin and Ng (2005) and d'Agostino and Giannone (2012), who study the predictive power of different approaches belonging to the class of general dynamic factor models. Alessi et al. (2007), Stock and Watson (2002b), Stock and Watson (2002a), and Stock and Watson (2006) compare the forecasting performance of factor models to different univariate and multivariate approaches. The evidence regarding the relative merits of factor models in forecasting, compared to other methods, differs between works. Stock and Watson (1999) and Stock and Watson (2002b) find a better forecast performance of factor models compared to univariate methods for inflation and industrial production, whereas Schumacher and Dreger (2002), Banerjee et al. (2005), and Engel et al. (2012) find mixed evidence.

The latent factors in a FM can be estimated using principal components analysis (PCA), as in Stock and Watson (2002a), by dynamic principal components analysis, using frequency domain methods, as proposed by Forni et al. (2000), or by Kalman filtering techniques. 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.

The supervised dynamic factor model we study in this paper is based on a Gaussian, factor-augmented, approximate, dynamic factor model in which the forecast objective is modelled jointly with the factors. In this paper, by dynamic factor model we mean a factor model in which the factors follow a dynamic equation. The system has a linear state-space representation and we estimate it using maximum likelihood. The likelihood function is delivered by the Kalman filter. Under this setup, we propose a way to measure the contribution of the forecast objective on the extracted factors that exploits the Kalman filter recursions. In particular, we compute the contribution of the forecast target to the variance of the filtered factors and find a positive correspondence between this quantity and the forecast performance of the supervised scheme.

We assess the out-of-sample forecast performance of the supervised scheme by means of a simulation study and in an empirical application. In the simulation study, we vary the degree of correlation between the factors and forecast objective. We compare the forecasts from the supervised model to two unsupervised FM specifications. We find that the higher the correlation between factors and forecast target, the better the forecasts of the supervised scheme. In the empirical application, we forecast selected macroeconomic time series and compare the forecast performance of the supervised FM to two unsupervised FM specifications and other multivariate and univariate methods. We use the dataset from Jurado et al. (2015), adding two more variables: real disposable personal income and personal consumption expenditure, excluding food and energy, and removing the index of aggregate weekly Hours (BLS), because this series starts later than the others. The resulting dataset comprises 132 variables. We forecast consumer price index (CPI), federal funds rate (FFR), personal consumption expenditures deflator (PCEd), producer price index (PPI), personal income (PEI), unemployment rate (UR), industrial production (IP), real disposable income (RDI), and personal consumption expenditures (PCE). The observations range from January 1960 to December 2011 and all variables refer to the US economy.

The paper is organized as follows: in Section 2 we introduce the supervised factor model and compare with other forecasting methods based on factor models; in Section 3 we show how supervision can be measured using the Kalman filter recursions; in Section 4 we provide some details on the computational aspects of the analysis; in Sections 5 and 6 we describe the empirical application and the simulation setup, respectively; finally, Section 7 concludes.

2 Forecasting with dynamic factor models

Let y_t be the forecast objective, $\mathbf{x}_t$ an N-dimensional vector of predictors (that may or not include lags of the forecast objective), h the forecast horizon and T the last available time-point in the estimation window.

2.1 Supervised factor model

We propose the following forecasting model. Consider the state-space system:

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

where $\mathbf{f}_t \in \mathbb{R}^k$ are latent factors, $\mathbf{\Lambda}$ 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 uncorrelated vectors of disturbances and $\mathbf{H}$ and $\mathbf{Q}$ are their respective variance-covariance matrices. The forecast objective is placed in the state equation together with the latent factors and the predictors are modelled in the measurement equation. We consider joint estimation of the factors using the Kalman filter recursions and maximum likelihood estimation for the parameters. The intuition behind the model is that if the forecast objective is correlated with the factors, modelling factors and forecast objective jointly should deliver a better estimate of the factors. We define supervision to be the contribution of the forecast target to the estimation of the latent factors. In the next section we derive the analytical expression of this contribution and present a measure of supervision based on it.

The state equation can be understood as a factor augmented VAR (FAVAR), introduced in Bernanke et al. (2005), in which factors are included together with observables in a VAR model. A specification similar to this one was used also in Diebold et al. (2006) to analyse the correlation between the Nelson-Siegel factors and some macroeconomic variables.

We wish to extract factors from a large number of predictors and model them jointly with the forecast objective. In order to find a parsimonious specification of the factor model we select as factor loadings basis functions of $\mathbb{R}^N$. This corresponds to taking a low order approximation of the vector of predictors at each point in time. Virtually any basis of $\mathbb{R}^N$ can be used. We choose discrete cosine basis for their ease of implementation. Mallat (1999, Theorem 8.12) shows that a random vector in $\mathbb{C}^N$ can be decomposed into discrete cosine basis. In particular, any $\mathbf{g} \in \mathbb{C}^N$ can be decomposed into

$$g_n = \frac{2}{N} \sum_{k=0}^{N-1} f_n \lambda_k cos \left[\frac{k\pi}{N} \left(n + \frac{1}{2} \right) \right],$$

for $0 \leq n < N$, where g_n is the $n-th$ component of $\mathbf{g}$,

$$\lambda_k = \begin{cases} 2^{-1/2} & \text{if } k = 0 \text{ and} \\ 1 & \text{otherwise} \end{cases}$$

and

$$f_n = \left\langle g_n, \lambda_k cos\left[\frac{k\pi}{N}\left(n+\frac{1}{2}\right)\right]\right\rangle = \lambda_k \sum_{n=0}^{N-1} g_n cos\left[\frac{k\pi}{N}\left(n+\frac{1}{2}\right)\right],$$

are the discrete cosine transform of type I.

In our specification $\mathbf{x}_t = \mathbf{g}_t + \boldsymbol{\epsilon}_t$ for each $t = 1, ..., T$ and $x_{t,n} = g_{t,n} + \epsilon_{t,n}$ with $n = 1, ..., N$, where with $\mathbf{x}_t$ we denote a vector of predictors. For each point in time we then have $g_{t,n} = \frac{2}{N}\sum_{k=0}^{N-1} f_{t,k}\lambda_k cos\left[\frac{k\pi}{N}\left(n+\frac{1}{2}\right)\right]$. The weights $f_{t,k}$ are estimated via Kalman filter/smoother recursions. The cosine basis functions are then contained in the factor loading matrix

$$\boldsymbol{\Lambda} \;=\; \begin{bmatrix} \frac{\sqrt{2}}{N} & \frac{2}{N}cos\left[\frac{\pi}{N}\left(1+\frac{1}{2}\right)\right] & \cdots & \frac{2}{N}cos\left[\frac{(k-1)\pi}{N}\left(1+\frac{1}{2}\right)\right] \\ \vdots & \vdots & & \vdots \\ \frac{\sqrt{2}}{N} & \frac{2}{N}cos\left[\frac{\pi}{N}\left(N+\frac{1}{2}\right)\right] & \cdots & \frac{2}{N}cos\left[\frac{(k-1)\pi}{N}\left(N+\frac{1}{2}\right)\right] \end{bmatrix}. \tag{2}$$

The supervised factor model is then comprised of equations (1) and (2). The forecasting scheme for this model is:

(i) estimation of the system parameters using maximum likelihood;

(ii) extraction of the factors using the Kalman filter;

(iii) the forecast $\hat{y}_{T+h}$ is obtained as the last element of the vector

$$\begin{bmatrix} \hat{\mathbf{f}}_{T+h|T} \\ \hat{y}_{T+h|T} \end{bmatrix} \;=\; \hat{\mathbf{T}}^h \begin{bmatrix} \hat{\mathbf{f}}_{T|T} \\ y_T \end{bmatrix} + \sum_{i=0}^{h-1} \hat{\mathbf{T}}^i \hat{\mathbf{c}}, \tag{3}$$

where $\hat{\mathbf{f}}_{T|T}$ is the vector of filtered factors, h is the forecast lead, and $\hat{\mathbf{T}}$ and $\hat{\mathbf{c}}$ are estimated parameters.

Note that the filtered and smoothed estimates for $\mathbf{f}_T$ are the same.

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

$$y_{t+h} \;=\; \boldsymbol{\beta}(L)'\mathbf{f}_t + \gamma(L)y_t + \epsilon_{t+h}, \tag{4}$$

$$x_{t,i} \;=\; \boldsymbol{\lambda}_i(L)\mathbf{f}_t + \eta_{t,i}, \tag{5}$$

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 (4)-(5) as a static factor model, i.e. a factor model in which the factors do not appear in lags:

$$\begin{aligned} y_{t+h} \;&=\; c + \boldsymbol{\beta}'\mathbf{F}_t + \gamma(L)y_t + \epsilon_{t+h}, \\ \mathbf{x}_t \;&=\; \boldsymbol{\Lambda}\mathbf{F}_t + \boldsymbol{\eta}_t, \end{aligned} \tag{6}$$

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 (5), using either principal components analysis or the Kalman filter;

(ii) regression of the forecast objective on the lagged estimated factors and on its lags, according to the forecasting equation (4);

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

$$\hat{y}_{T+h} = \hat{c} + \hat{\boldsymbol{\beta}}'\hat{\mathbf{F}}_t + \hat{\gamma}(L)y_T.$$

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.

The difference between the supervised DFM and the two-step forecasting procedure is that in the former model the factors are extracted conditionally on the forecast target. In the supervised framework the filtered/smoothed factors are tailored to the forecast objective. Note that for a linear state-space system the Kalman filter delivers the best linear predictions of the state vector, conditionally on the observations. Moreover, if the innovations are Gaussian, the filtered states coincide with conditional expectations, for more details on the optimality properties of the Kalman filter see Brockwell and Davis (2009).

3 Quantifying supervision

In this section we propose a statistic to quantify supervision. We are interested in quantifying the influence of the forecast target on the filtered factors. To accomplish this, we develop some results that hold for a general linear, state-space system with non-random coefficient matrices. Consider the following state-space system:

$$\begin{aligned} \boldsymbol{y}_t &= \mathbf{Z}_t\boldsymbol{\alpha}_t + \boldsymbol{\epsilon}_t, \\ \boldsymbol{\alpha}_{t+1} &= \mathbf{T}_t\boldsymbol{\alpha}_t + \mathbf{R}_t\boldsymbol{\eta}_t, \end{aligned} \tag{7}$$

where $\boldsymbol{\epsilon}_t \sim WN(\mathbf{0}, \mathbf{H}_t)$ and $\boldsymbol{\eta}_t \sim WN(\mathbf{0}, \mathbf{Q}_t)$ are uncorrelated random vectors, $\mathbf{y}_t \in \mathbb{R}^N$, $\boldsymbol{\alpha}_t \in \mathbb{R}^k$, and $\boldsymbol{\eta}_t \in \mathbb{R}^q$, and the matrices $\mathbf{Z}_t$, $\mathbf{T}_t$, and $\mathbf{R}_t$ are of suitable dimensions. Note that in this context we assume white noise innovations.

In model (1), we include the observable forecast target as last element, both in the measurement and in the state equation. In the notation of model (7), we are therefore ultimately interested in the influence of the forecast target, the last element in $\mathbf{y}_t$, denoted $y_{t,N}$ (or y_t in the notation of model (1)), on the filtered factors $\hat{\mathbf{f}}_t$. To be more precise, the objective is to quantify the influence of the sequence $\{y_{i,N}\}_{i=1,\ldots,t}$ (or $\{y_i\}_{i=1,\ldots,t}$ in the notation of model (1)) on $\hat{\mathbf{f}}_t$, the filtered factors at time t.

The standard Kalman filter recursions (see for instance Durbin and Koopman (2012)) for system (7) are:

$$\begin{aligned} \mathbf{v}_t &= \mathbf{y}_t - \mathbf{Z}_t\mathbf{a}_t, \\ \mathbf{F}_t &= \mathbf{Z}_t\mathbf{P}_t\mathbf{Z}_t' + \mathbf{H}_t, \\ \mathbf{M}_t &= \mathbf{P}_t\mathbf{Z}_t', \end{aligned} \tag{8}$$

$$
\begin{aligned}
\mathbf{a}_{t|t} &= \mathbf{a}_t + \mathbf{M}_t\mathbf{F}_t^{-1}\mathbf{v}_t, & \mathbf{P}_{t|t} &= \mathbf{P}_t - \mathbf{M}_t\mathbf{F}_t^{-1}\mathbf{M}_t', \\
\mathbf{a}_{t+1} &= \mathbf{T}_t\mathbf{a}_{t|t}, & \mathbf{P}_{t+1} &= \mathbf{T}_t\mathbf{P}_{t|t}\mathbf{T}_t' + \mathbf{R}_t\mathbf{Q}_t\mathbf{R}_t', \qquad (9)
\end{aligned}
$$

for $t = 1, \ldots, T$, where $\mathbf{P}_t = E[(\boldsymbol{\alpha}_t - \mathbf{a}_t)(\boldsymbol{\alpha}_t - \mathbf{a}_t)]$, $\mathbf{P}_{t|t} = E\left[[\boldsymbol{\alpha}_t - \mathbf{a}_{t|t}][\boldsymbol{\alpha}_t - \mathbf{a}_{t|t}]'\right]$, $\mathbf{F}_t = E[\mathbf{v}_t\mathbf{v}_t']$, and $\mathbf{a}_{t|t} = P_t(\boldsymbol{\alpha}_t) = P(\boldsymbol{\alpha}_t|\mathbf{y}_0, \ldots, \mathbf{y}_t)$ and $\mathbf{a}_t = P_{t-1}(\boldsymbol{\alpha}_t) = P(\boldsymbol{\alpha}_t|\mathbf{y}_0, \ldots, \mathbf{y}_{t-1})$ are the filtered state and one-step-ahead prediction of the state vector, respectively, and $P_t(\cdot)$ is the best linear predictor operator, see Brockwell and Davis (2002) for more details on the definition and properties of the best linear predictor operator. In the particular case of Gaussian innovations in both the state and measurement equations, the best linear predictor coincides with the conditional expectation. The forecasting step in the Kalman recursions (9) can be written as

$$
\begin{aligned}
\mathbf{a}_{t+1} &= \mathbf{T}_t\mathbf{a}_t + \mathbf{K}_t\mathbf{v}_t \\
&= \mathbf{T}_t\mathbf{a}_t + \mathbf{K}_t(\mathbf{y}_t - \mathbf{Z}_t\mathbf{a}_t) \\
&= \mathbf{S}_t\mathbf{a}_t + \mathbf{K}_t\mathbf{y}_t, \qquad (10)
\end{aligned}
$$

with $\mathbf{K}_t = \mathbf{T}_t\mathbf{P}_t\mathbf{Z}_t'\mathbf{F}_t^{-1}$ and $\mathbf{S}_t = \mathbf{T}_t - \mathbf{K}_t\mathbf{Z}_t$. Iterating backwards on the one-step-ahead prediction of the state, the filtered state can be written as

$$
\begin{aligned}
\mathbf{a}_{t|t} &= \mathbf{a}_t + \tilde{\mathbf{K}}_t\mathbf{v}_t \\
&= \mathbf{N}_t\mathbf{a}_t + \tilde{\mathbf{K}}_t\mathbf{y}_t \\
&= \mathbf{N}_t\left(\mathbf{S}_{t-1}\mathbf{a}_{t-1} + \mathbf{K}_{t-1}\mathbf{y}_{t-1}\right) + \tilde{\mathbf{K}}_t\mathbf{y}_t \\
&= \mathbf{N}_t\left[\mathbf{S}_{t-1}\left(\mathbf{S}_{t-2}\mathbf{a}_{t-2} + \mathbf{K}_{t-2}\mathbf{y}_{t-2}\right) + \mathbf{K}_{t-1}\mathbf{y}_{t-1}\right] + \tilde{\mathbf{K}}_t\mathbf{y}_t \\
&= \mathbf{N}_t\left[\mathbf{S}_{t-1}\mathbf{S}_{t-2}\mathbf{a}_{t-2} + \mathbf{S}_{t-1}\mathbf{K}_{t-2}\mathbf{y}_{t-2} + \mathbf{K}_{t-1}\mathbf{y}_{t-1}\right] + \tilde{\mathbf{K}}_t\mathbf{y}_t \\
&= \ldots \\
&= \mathbf{N}_t\left[\prod_{i=1}^{t-1}\mathbf{S}_{t-i}\mathbf{a}_1 + \sum_{i=1}^{t-1}\left(\prod_{\ell=1}^{i-1}\mathbf{S}_{t-\ell}\right)\mathbf{K}_{t-i}\mathbf{y}_{t-i}\right] + \tilde{\mathbf{K}}_t\mathbf{y}_t, \qquad (11)
\end{aligned}
$$

with $\mathbf{N}_t = (\mathbf{I} - \tilde{\mathbf{K}}_t\mathbf{Z}_t)$, $\tilde{\mathbf{K}}_t = \mathbf{P}_t\mathbf{Z}_t'\mathbf{F}_t^{-1}$ and the convention $\prod_{\ell=1}^{i-1}\mathbf{S}_{t-\ell} = \mathbf{I}_k$ if $i - 1 < 1$. The contribution of the n-th observable on the filtered state at time t can be isolated from the previous expression in the following way

$$
\begin{aligned}
\mathbf{a}_{t|t} &= \mathbf{N}_t\left[\prod_{i=1}^{t-1}\mathbf{S}_{t-i}\mathbf{a}_1 + \sum_{i=1}^{t-1}\left(\prod_{\ell=1}^{i-1}\mathbf{S}_{t-\ell}\right)\mathbf{K}_{t-i}\sum_{j=1}^{N}\mathbf{e}_j\mathbf{e}_j'\mathbf{y}_{t-i}\right] + \tilde{\mathbf{K}}_t\sum_{j=1}^{N}\mathbf{e}_j\mathbf{e}_j'\mathbf{y}_t \\
&= \mathbf{N}_t\left[\prod_{i=1}^{t-1}\mathbf{S}_{t-i}\mathbf{a}_1 + \sum_{i=1}^{t-1}\left(\prod_{\ell=1}^{i-1}\mathbf{S}_{t-\ell}\right)\mathbf{K}_{t-i}\sum_{j=1,j\neq n}^{N}\mathbf{e}_j\mathbf{e}_j'\mathbf{y}_{t-i}\right] + \tilde{\mathbf{K}}_t\sum_{j=1,j\neq n}^{N}\mathbf{e}_j\mathbf{e}_j'\mathbf{y}_t \\
&+ \mathbf{N}_t\sum_{i=1}^{t-1}\left(\prod_{\ell=1}^{i-1}\mathbf{S}_{t-\ell}\right)\mathbf{k}_{t-i,\cdot n}y_{t-i,n} \\
&+ \tilde{\mathbf{k}}_{t,\cdot n}y_{t,n}, \qquad (12)
\end{aligned}
$$

where with $\mathbf{b}_{t,\cdot n}$ and $\mathbf{b}_{t,n\cdot}$ we denote the n-th column and row of the matrix $\mathbf{B}_t$, respectively, and $y_{t,n}$ is the n-th component of $\mathbf{y}_t$; $\mathbf{e}_j$ with $j = 1, \ldots, N$ are the canonical basis vectors of $\mathbb{R}^N$. In (12) we made use of the identity $\sum_{j=1}^{N} \mathbf{e}_j \mathbf{e}_j' = \mathbf{I}_N$. The contribution of the n-th observable $\{y_{i,N}\}_{i=1,\dots,t}$ on the filtered state $\mathbf{a}_{t|t}$ is given by

$$\mathbf{s}_t^n \;=\; \mathbf{N}_t \sum_{i=1}^{t-1} \left(\prod_{\ell=1}^{i-1} \mathbf{S}_{t-\ell} \right) \mathbf{k}_{t-i,\cdot n} y_{t-i,n} + \tilde{\mathbf{k}}_{t,\cdot n} y_{t,n}. \tag{13}$$

The first moment and the variance of $\mathbf{s}_t^n$ are

$$\begin{aligned} E[\mathbf{s}_t^n] &= \mathbf{N}_t \sum_{i=1}^{t-1} \left(\prod_{\ell=1}^{i-1} \mathbf{S}_{t-\ell} \right) \mathbf{k}_{t-i,\cdot n} E[y_{t-i,n}] + \tilde{\mathbf{k}}_{t,\cdot n} E[y_{t,n}], \\ var[\mathbf{s}_t^n] &= \mathbf{N}_t \sum_{i=1,i'=1}^{t-1} \left(\prod_{\ell=1}^{i-1} \mathbf{S}_{t-\ell} \right) \mathbf{k}_{t-i,\cdot n} cov[y_{t-i,n}, y_{t-i',n}] \mathbf{k}_{t-i',\cdot n}' \left(\prod_{\ell=1}^{i'-1} \mathbf{S}_{t-\ell}' \right) \mathbf{N}_t' \\ &+ \tilde{\mathbf{k}}_{t,\cdot n} var[y_{t,n}] \tilde{\mathbf{k}}_{t,\cdot n}' \\ &+ 2\mathbf{N}_t \sum_{i=1}^{t-1} \left(\prod_{\ell=1}^{i-1} \mathbf{S}_{t-\ell} \right) \mathbf{k}_{t-i,\cdot n} cov[y_{t-i,n}, y_{t,n}] \tilde{\mathbf{k}}_{t,\cdot n}'. \end{aligned} \tag{14}$$

Note that $\mathbf{F}_t = E[\mathbf{v}_t \mathbf{v}_t']$, $\mathbf{P}_t = E[(\boldsymbol{\alpha}_t - \mathbf{a}_t)(\boldsymbol{\alpha}_t - \mathbf{a}_t)']$, and $\tilde{\mathbf{K}}_t = \mathbf{P}_t \mathbf{Z}_t' \mathbf{F}_t^{-1}$ are non-random matrices. If $y_{t,n}$ is stationary with mean $E[y_{t,n}] = \mu_{y.,n}$ and autocovariance function $cov[y_{t,n}, y_{t-h,n}] = \gamma_n(h)$, we can rewrite (14) as

$$\begin{aligned} E[\mathbf{s}_t^n] &= \left[\mathbf{N}_t \sum_{i=1}^{t-1} \left(\prod_{\ell=1}^{i-1} \mathbf{S}_{t-\ell} \right) \mathbf{k}_{t-i,\cdot n} + \tilde{\mathbf{k}}_{t,\cdot n} \right] \mu_{y.,n}, \\ var[\mathbf{s}_t^n] &= \mathbf{N}_t \sum_{i=1,i'=1}^{t-1} \left(\prod_{\ell=1}^{i-1} \mathbf{S}_{t-\ell} \right) \mathbf{k}_{t-i,\cdot n} \gamma_n(i-i') \mathbf{k}_{t-i',\cdot n}' \left(\prod_{\ell=1}^{i'-1} \mathbf{S}_{t-\ell}' \right) \mathbf{N}_t' \\ &+ \tilde{\mathbf{k}}_{t,\cdot n} \gamma_n(0) \tilde{\mathbf{k}}_{t,\cdot n}' \\ &+ 2\mathbf{N}_t \sum_{i=1}^{t-1} \left(\prod_{\ell=1}^{i-1} \mathbf{S}_{t-\ell} \right) \mathbf{k}_{t-i,\cdot n} \gamma_n(i) \tilde{\mathbf{k}}_{t,\cdot n}', \end{aligned} \tag{15}$$

or, in a more compact form

$$\begin{aligned} E[\mathbf{s}_t^n] &= \left(\mathbf{W}_{t-1} \boldsymbol{\iota}_{t-1} + \tilde{\mathbf{k}}_{t,\cdot n} \right) \mu_{y.,n}, \\ var[\mathbf{s}_t^n] &= \mathbf{W}_{t-1} \boldsymbol{\Gamma}_{t-2}^n \mathbf{W}_{t-1}' \\ &+ \tilde{\mathbf{k}}_{t,\cdot n} \gamma_n(0) \tilde{\mathbf{k}}_{t,\cdot n}' \\ &+ 2\mathbf{W}_{t-1} \boldsymbol{\gamma}_{t-1}^n \tilde{\mathbf{k}}_{t,\cdot n}', \end{aligned}$$

where $\mathbf{W}_{t-1} = \left[\mathbf{N}_t \mathbf{k}_{t-1,\cdot n}, \mathbf{N}_t \left(\prod_{\ell=1}^{1} \mathbf{S}_{t-\ell} \right) \mathbf{k}_{t-2,\cdot n}, \dots, \mathbf{N}_t \left(\prod_{\ell=1}^{t-2} \mathbf{S}_{t-\ell} \right) \mathbf{k}_{1,\cdot n} \right]$, $\boldsymbol{\iota}_{t-1}$ is a vector of ones of length $t-1$, $\boldsymbol{\gamma}_{t-1}^n = [\gamma_n(1), \dots, \gamma_n(t-1)]'$, and

$$\boldsymbol{\Gamma}_{t-2}^n \;=\; \begin{bmatrix} \gamma_n(0) & \gamma_n(1) & \cdots & \gamma_n(t-2) \\ \gamma_n(1) & \gamma_n(0) & \cdots & \gamma_n(t-3) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_n(t-2) & \gamma_n(t-3) & \cdots & \gamma_n(0) \end{bmatrix}. \tag{16}$$

The distribution of the contribution of the n-th observable on the filtered state at time t is

$$\mathbf{s}_t^n \sim p\left(E\left[\mathbf{s}_t^n\right], var\left[\mathbf{s}_t^n\right]\right), \tag{17}$$

where $p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the distribution function of $y_{t,n}$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The contribution of observable n on the j-th filtered state at time t is given by $s_{t,j}^n = \tilde{\mathbf{e}}_j' \mathbf{s}_t^n$ with $\tilde{\mathbf{e}}_j$ the j-th canonical basis vector of $\mathbb{R}^k$.

3.1 Variance of filtered states explained by forecast objective

In this section we derive the variance ratio used as a measure of supervision. In particular, we compute the fraction of the total variance of the filtered factors that is explained by $\mathbf{s}_t^n$, the contribution of the forecast target. According to eqn. (11) and assuming $\mathbf{a}_1$ to be a constant vector (typically $\mathbf{a}_1 = \boldsymbol{\mu} = E[\boldsymbol{\alpha}_t]$ for a stationary system), the variance of $\mathbf{a}_{t|t}$ can be written as

$$\begin{aligned} var[\mathbf{a}_{t|t}] &= var\left[\mathbf{N}_t \sum_{i=1}^{t-1} \mathbf{B}_t^i \mathbf{y}_{t-i}\right] + var\left[\tilde{\mathbf{K}}_t \mathbf{y}_t\right] + 2cov\left[\mathbf{N}_t \sum_{i=1}^{t-1} \mathbf{B}_t^i \mathbf{y}_{t-i}, \tilde{\mathbf{K}}_t \mathbf{y}_t\right] \\ &= \mathbf{N}_t \sum_{i=1,j=1}^{t-1} \mathbf{B}_t^i cov\left[\mathbf{y}_{t-i}, \mathbf{y}_{t-j}\right] (\mathbf{B}_t^j)' \mathbf{N}_t' \\ &+ \tilde{\mathbf{K}}_t var\left[\mathbf{y}_t\right] \tilde{\mathbf{K}}_t' + 2\mathbf{N}_t \sum_{i=1}^{t-1} \mathbf{B}_t^i cov\left[\mathbf{y}_{t-i}, \mathbf{y}_t\right] \tilde{\mathbf{K}}_t', \end{aligned}$$

where $\mathbf{B}_t^i = \left(\prod_{\ell=1}^{i-1} \mathbf{S}_{t-\ell}\right) \mathbf{K}_{t-i}$ and as in the previous section $\mathbf{S}_t = \mathbf{T}_t - \mathbf{K}_t \mathbf{Z}_t$. If $\mathbf{y}_t$ is stationary, we can write

$$\begin{aligned} var\left[\mathbf{a}_{t|t}\right] &= \mathbf{N}_t \sum_{i=1,j=1}^{t-1} \mathbf{B}_t^i \boldsymbol{\Sigma}(i-j) (\mathbf{B}_t^j)' \mathbf{N}_t' \\ &+ \tilde{\mathbf{K}}_t \boldsymbol{\Sigma}(0) \tilde{\mathbf{K}}_t' + 2\mathbf{N}_t \sum_{i=1}^{t-1} \mathbf{B}_t^i \boldsymbol{\Sigma}(i) \tilde{\mathbf{K}}_t', \end{aligned} \tag{18}$$

where $\boldsymbol{\Sigma}(i-j) = cov(\mathbf{y}_{t-i}, \mathbf{y}_{t-j})$.

Notice that, since the sequence of filtered states depends on the initial values of the filter, so does the sequence $\{\mathbf{s}_i^n\}_{i=1,\ldots,t}$. As a consequence, it is not a stationary and ergodic sequence. Its variance changes in time and in order to estimate it, we first need to estimate the autocovariance function of the sequence of observations $\{\mathbf{y}_i\}_{i=1,\ldots,t}$ and the parameters of the system and then evaluate expressions (16) and (18).

The variance of the j-th filtered factor explained by the n-th variable can then be assessed by means of the ratio

$$r_t^{j,n} = \frac{\tilde{\mathbf{e}}_j' var\left[\mathbf{s}_t^n\right] \tilde{\mathbf{e}}_j}{\tilde{\mathbf{e}}_j' var\left[\mathbf{a}_{t|t}\right] \tilde{\mathbf{e}}_j}, \tag{19}$$

where $\tilde{\mathbf{e}}_j$ is the j-th canonical basis vector of $\mathbb{R}^k$, as before. This quantity can be estimated by consistently replacing the data-generating parameters with consistent estimates and the autocovariances of $\mathbf{y}_t$ by their sample counterparts (under the condition of ergodic stationarity). Note that the variance ratio has the same expression also when adding a constant $\mathbf{c}_t$ to the state equation.

4 Computational aspects

The objective of this study is to determine the forecasting power of the supervised factor model (1)-(2). 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.

4.1 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 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)[1]. 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 and requires the computation of the gradient of the function to be minimized.

The CMA-ES algorithm performs very well when no good initial values are available but it is slower to converge than the BFGS routine. The BFGS algorithm, on the other hand, requires good initial values but converges considerably faster than the CMA-ES algorithm (once good initial values have been obtained). Hence, we use the CMA-ES algorithm to find good initial values and then the BFGS one to perform the minimizations with the different rolling windows of data. We use algorithmic (or automatic) differentiation[2] to compute gradients. We make use of the ADEPT C++ library, see Hogan (2013)[3]. The advantage of using algorithmic differentiation over finite differences is twofold: increased speed and elimination of approximation errors in the computation of the gradient.

[1] 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.

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

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

4.2 Speed improvements

To gain speed we chose C++ as the programming language, using routines from the Numerical Recipes, Press et al. (2002) [4]. We compile and run the executables on a Linux 64-bit operating system using the GCC compiler [5]. We use Open MPI 1.6.4 (Message Passing Interface) with the Open MPI C++ wrapper compiler `mpic++` to parallelise the maximum likelihood estimations [6]. We compute gradients using the ADEPT library for algorithmic differentiation, see Hogan (2013).

5 Empirical application

We wish to assess the forecasting performance of model (1)-(2). We fix the number of latent factors at 1, 2, and 3 for the models involving factors[7]. The complete sample size is $T = 617$, the rolling window for the parameter estimation has size $R = 306$, and the number of forecasts is $S = 300$.

5.1 Data

We use the Jurado, Ludvigson and Ng dataset as used in Jurado et al. (2015) adding two more variables, namely real disposable income (RDI) and personal consumption expenditure excluding food and energy (PCE) and removing the Index of Aggregate Weekly Hours (BLS). The resulting dataset comprises 132 variables. We have applied the same transformations as in Jurado et al. (2015) to achieve stationarity for the series in common with this dataset. For RDI and PCE we used the same transformations used for the personal income (PI) and for the personal consumption deflator (PCEd), respectively. Details on the Jurado, Ludvigson and Ng dataset used in Jurado et al. (2015) are provided by the authors at `http://www.econ.nyu.edu/user/ludvigsons/data.htm`.

The details for the time series added to the Jurado, Ludvigson and Ng dataset are the following:

- **PCE.** Series ID: DPCCRC1M027SBEA, Title: Personal consumption expenditures excluding food and energy, Source: U.S. Department of Commerce: Bureau of Economic Analysis, Release: Personal Income and Outlays, Units: Billions of Dollars, Frequency: Monthly, Seasonal Adjustment: Seasonally Adjusted Annual Rate, Notes: BEA Account Code: DPCCRC1, For more information about this series see `http://www.bea.gov/national/`.

- **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:

[4]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++.

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

[6]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.

[7]See below for more details on the choice of the number of factors.

A067RX1, A Guide to the National Income and Product Accounts of the United States (NIPA) - (`http://www.bea.gov/national/pdf/nipaguid.pdf`).

The RDI and PCE series have been taken from the FRED (Federal Reserve Economic Data) database and can be downloaded from the website of the Federal Reserve Bank of St. Louis: `http://research.stlouisfed.org/fred2`.

The macroeconomic variables selected as forecast objectives are: consumer price index (CPI), federal funds rate (FFR), personal consumption expenditures deflator (PCEd), producer price index (PPI), personal income (PEI), unemployment rate (UR), industrial production (IP), real disposable income (RDI), and personal consumer expenditures (PCE).

The observations in levels range from January 1960 to December 2011 for a total of 624 observations, and from March 1960 to December 2011 after being transformed to stationarity, for a total of 622 data points. The data refers to the US economy.

5.2 Selection of number of factors

The selection of the number of factors is a key aspect in dynamic/static factor models. A widely used information-criterion-based method for static factor models was derived by Bai and Ng (2002). Under appropriate assumptions, they show that their method can consistently identify the number of factors as both the cross-section and the sample-size tend to infinity. The method was extended to the case of restricted dynamic models by Bai and Ng (2007) and Amengual and Watson (2007). The Bai and Ng (2002) criterion was found to overestimate the true number of factors in simulation studies by e.g. Hallin and Liška (2007), who propose a new method, valid under more general assumptions, that exploits the properties of the eigenvalues of sample spectral density matrices. Alessi et al. (2010) follow the idea of Hallin and Liška (2007) to improve on Bai and Ng (2002) in the less general case of static factor models. They show using simulations that their method performs well, in particular under large idiosyncratic disturbances. Another method for selecting the number of factors in static approximate factor models and based on the eigenvalues of the variance-covariance matrix of the panel of data, was proposed by Ahn and Horenstein (2013).

We find that the Alessi et al. (2010) test is somewhat dependent on the number and sizes of the subsamples required by the test. Similarly, the number of factors selected using the Ahn and Horenstein (2013) eigenvalue ratio test, is somewhat sensitive to the choice of the maximum allowed number of factors. Motivated by this and by the empirical finding that models using a low number of factors tend to forecast better (see e.g. Stock and Watson (2002b) for the case of output and inflation) in this work we consider models with a fixed, low number of factors. In particular, we consider factor models with 1, 2, and 3 factors. Increasing the number of factors was seen not to further improve the forecasts.

5.3 Competing models

We choose different competing models widely used in the forecasting literature in order to assess the relative forecasting performance of the supervised DFM. We divide these models into direct multi-step and indirect (recursive) forecasting models. In the factor models considered as well as in the principal components regressions and partial least squares regressions we extract 1, 2, and 3 factors. In the following we denote with h the forecast horizon, y_t the forecast objective, $\mathbf{x}_t = [x_t^1, \ldots, x_t^N]$ an $(N \times 1)$ vector of predictors, ϵ_t a Gaussian white noise innovation, $\mathbf{f}_t = [f_t^1, \ldots, f_t^k]$ a $(k \times 1)$ vector of factors and $\mathbf{\Lambda}$ a matrix of factor loadings.

5.3.1 Direct forecasting models

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

$$y_{t+h}(h) = c + \phi_1 y_t(h) + \ldots \phi_p y_{t-p}(h) + \epsilon_{t+h}. \tag{20}$$

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

$$y_{t+h}(h) = c + \theta_1 \epsilon_t + \ldots \theta_q \epsilon_{t-q} + \epsilon_{t+h}. \tag{21}$$

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 $\mathbf{X}_t = [x_t^1, \ldots, x_t^N, y_t]$; y_{t+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.

The fourth model considered is partial least squares regression (PLSR). In the first step, the partial least squares components $\hat{y}_t^m$ are computed using the forecast target $\{y_t : h \leq t \leq T_i\}$ and the predictors $\mathbf{X}_t = [x_t^1, \ldots, x_t^N, y_t]$ 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 $\hat{y}_t^m$ are regressed on the predictors $\mathbf{X}_t$ to recover the coefficient vector $\hat{\beta}^{PLSR}$. 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 $\hat{\beta}^{PLSR}$. The projected PLSR components at time T_i are then summed to obtain the h-period ahead forecast $\hat{y}_{T_i+h} = \sum_{m=1}^{M} \hat{y}_{T_i}^m$.

The fifth direct forecasting method considered is a two-step procedure as described in equations (4).

The sixth direct forecasting method is the one based on the principal components estimation approach in Stock and Watson (2002a) to a specific version of model (6). In particular we take as forecasting equation

$$\begin{aligned} y_{t+h} &= c + \boldsymbol{\beta}'\mathbf{f}_t + y_t + \epsilon_{t+h}, \\ \mathbf{x}_t &= \mathbf{\Lambda}\mathbf{f}_t + \boldsymbol{\eta}_t, \end{aligned} \tag{22}$$

with $\mathbf{f}_t = [f_t^1, \ldots, f_t^k]$. The predictors are in $\mathbf{x}_t$ and are standardized for each forecasting window. As described in Stock and Watson (2002a) the factor loadings $\mathbf{\Lambda}$ and the factors $\mathbf{f}_t$ for $t = 1, \ldots, T$ can be estimated using principal components. Denote with $\mathbf{F} = [\mathbf{f}_1, \ldots, \mathbf{f}_T]$ and $\mathbf{X}' = [\mathbf{x}_1, \ldots, \mathbf{x}_T]$; the estimator of the loadings $\hat{\mathbf{\Lambda}}$ is the matrix made of the eigenvectors corresponding to the largest eigenvalues of the matrix $\mathbf{X}'\mathbf{X}$ and the factors are estimated by $\hat{\mathbf{F}} = \mathbf{X}\hat{\mathbf{\Lambda}}$. In the present paper, we are focusing on forecasting and hence any estimated rotation of the factors will suffice for the analysis.

5.3.2 Indirect forecasting models

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

$$y_{t+h} = c + \phi_1 y_{t+h-1} + \dots \phi_p y_{t+h-p} + \epsilon_{t+h}. \tag{23}$$

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

$$y_{t+h} = c + \theta_1 \epsilon_{t+h-1} + \dots \theta_q \epsilon_{t+h-q} + \epsilon_{t+h}. \tag{24}$$

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 indirect forecasting method is an alternative two-step procedure. We specify a dynamic equation for the factors and a static one between the factors and the predictors/forecast target. In particular, we allow the factors $\mathbf{F}_t \in \mathbb{R}^k$ to follow the autoregressive dynamics:

$$\mathbf{F}_{t+1} = \mathbf{c} + \mathbf{T}\mathbf{F}_t + \boldsymbol{\nu}_t, \tag{25}$$

where $\mathbf{T}$ and $\mathbf{c}$ are a matrix and a vector of coefficients, respectively, and $\boldsymbol{\nu}_t$ is a vector of disturbances with $E[\boldsymbol{\nu}_t \boldsymbol{\nu}_t'] = \boldsymbol{\Sigma}$, and a static equation is specified for the mapping between the factors and the predictors/forecast target, such as

$$\begin{bmatrix} \mathbf{x}_t \\ y_t \end{bmatrix} = \mathbf{Z}\mathbf{F}_t + \boldsymbol{\epsilon}_t, \tag{26}$$

where $\mathbf{Z}$ is a matrix of factor loadings and $\boldsymbol{\epsilon}_t$ is an innovation vector, with $E[\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t'] = \boldsymbol{\Omega}$. In particular, we use the factor loading matrix (2). Forecasts can be constructed by estimating the system, iterating on the factor equation and then mapping the factors to the forecast objective using the estimated factor loadings. Assuming Gaussianity of the idiosyncratic errors, for instance, the system (25-26) can be estimated maximizing the likelihood delivered by the Kalman filter. In this case a forecasting scheme would be of this type:

(i) estimation of the system parameters by maximum likelihood;

(ii) extraction of the factors using the Kalman filter;

(iii) forecasting of factors using the state equation

$$\hat{\mathbf{f}}_{T+h} = \left[\hat{\mathbf{T}}^h \hat{\mathbf{f}}_T + \sum_{i=0}^{h-1} \hat{\mathbf{T}}^i \hat{\mathbf{c}} \right], \tag{27}$$

where $\hat{\mathbf{f}}_t$ represent estimated factors;

(iv) the forecast is then the last element of the vector

$$\begin{bmatrix} \hat{\mathbf{x}}_{T+h} \\ \hat{y}_{T+h} \end{bmatrix} = \hat{\mathbf{Z}} \hat{\mathbf{f}}_{T+h}. \tag{28}$$

Finally, we compare the forecast performance of the supervised model (1-2) to its unsupervised counterpart. Namely, in this specification the factors are first extracted using the Kalman filter and the forecast are then obtained using the forecast equation

$$y_{t+h} = c + \hat{\mathbf{f}}_t'\boldsymbol{\beta} + \gamma y_t + u_t, \tag{29}$$

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.

5.4 Forecasting

5.4.1 Forecasting scheme

The aim is to compute the forecast of the objective variable y_t at time $t+h$, i.e. $\hat{y}_{t+h}$, 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. The variables, including the forecast target, are made stationary according to the transformations used in Jurado et al. (2015). We standardize the variables in the estimation windows by subtracting the time average and dividing by the standard deviation.

We build series of forecast errors of length S for all forecast targets. 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 $\mathbf{Y}_t = \{x_t^1, ..., x_t^N, y_t\}$. 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. The forecasting strategy for h-step ahead forecasts for the supervised factor model (1)-(2) is the following (for the competing models the forecasting scheme is analogous), for $i = 1, \ldots, S$:

(i) estimation of the system parameters 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) computation of the filtered state vector at time T_i, i.e. $\hat{\boldsymbol{\alpha}}_{T_i|T_i}$ (note that the last element of $\hat{\boldsymbol{\alpha}}_{T_i|T_i}$ is y_{T_i});

(iii) the forecast is then:

$$\hat{y}_{T_i+h|T_i} = [\mathbf{0}^{1\times L} : 1]\left[\hat{\mathbf{T}}^h \hat{\boldsymbol{\alpha}}_{T_i|T_i} + \sum_{i=0}^{h-1} \hat{\mathbf{T}}^i \hat{\mathbf{c}}\right], \tag{30}$$

where the parameter matrices are relative to equation (1).

The forecasting scheme for the competing methods is analogous.

In particular, the complete sample size is $T = 622$, the rolling window has size $R = 311$, and the number of forecasts is $S = 300$. The 1-step ahead forecasts range from February 1986 to January 2011. The 12-step ahead forecasts range from January 1987 to December 2011.

5.4.2 Test of forecast performance

We make use of the conditional predictive ability test proposed in Giacomini and White (2006)[8] to assess the forecasting performance of the supervised factor model (1)-(2), relative to the other forecasting methods. In particular, we use a quadratic loss function. This test is valid also when comparing nested models, provided a rolling 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.

5.5 Empirical application results

In this subsection we present results corresponding to the empirical application. The mean square prediction error ratios between forecasts from the supervised model and the competing models can be found in tables 1-9 in Appendix A. The supervised factor model corresponds to equations (1) with discrete cosine basis as loadings, equation (2). In the tables, 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. The different forecasting models are labelled according to the following convention:

- **model 1**. Principal component regression (PCR);
- **model 2**. Partial least squares regression (PLSR);
- **model 3**. AR(p) direct, eqn. (20);
- **model 4**. MA(q) direct, eqn. (21);
- **model 5**. AR(p) indirect, eqn. (23);
- **model 6**. MA(q) indirect, eqn. (24);
- **model 7**. Stock and Watson two-step procedure, eqn. (22);
- **model 8**. Unsupervised factor model (25), and (26) with discrete cosine basis factor loadings, eqn. (2);
- **model 9**. Unsupervised factor model as in Section 2.2 with discrete cosine basis factor loadings, eqn. (2);
- **model 10**. Supervised factor model as in eqn. (1) with discrete cosine basis factor loadings (2).

For reasons explained in Section 5.2, we estimate the supervised and unsupervised factor models using 1, 2, and 3 factors.

Looking at the tables 1-9, we can make the following remarks (divided with respect to the different number of factors used):

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

(i) **1 factor**. The supervised factor model, eqn. (1), in general delivers forecasts better than or similar to the other forecasting methods. In more than 56% of the cases the model performs better than the competing ones, in 23% equally well and in roughly 20% of the cases it performs worse. However, of the 56% cases in which the model performs better, 37% of them are statistically significant at the $\alpha = 10\%$ significance level, whereas of the 17% of cases in which it performs worse, only 9% are statistically significant at the $\alpha = 10\%$ significance level. The supervised factor model offers better forecasts relative to unsupervised ones for most targets. The model forecasts particularly well the federal funds rate (FFR). The improvements over unsupervised factor models 7, 8, and 9, are particularly marked for this variable;

(ii) **2 factors**. In most cases the supervised factor model delivers forecasts similar to or better than the other methods. In more than 51% of the cases the model performs better than the competing ones, in 36% of which the differences are statistically significant at $\alpha = 10\%$ significance level, in 23% equally well and in roughly 26% of the cases it performs worse, in 15% of which the differences are statistically significant at the $\alpha = 10\%$ significance level. The supervised factor model forecasts particularly well the federal funds rate (FFR);

(iii) **3 factors**. In most cases the supervised factor model delivers forecasts similar to or better than the other methods. In more than 60% of the cases the model performs better than the competing ones, in 33% of which the differences are statistically significant at the $\alpha = 10\%$ significance level, in 20% equally well and in roughly 20% of the cases it performs worse, in 17% of which the differences are statistically significant at the $\alpha = 10\%$ significance level. The supervised factor model forecasts particularly well the unemployment rate (UR), personal income (PEI), and real disposable income (RDI). Improvements over the unsupervised models 7, 8, and 9 are particularly clear for UR and RDI.

The indirect MA(q) process is hard to beat in forecasting inflation measures and the federal funds rate at lead $h = 1$. For the rest of variables/leads the supervised factor model performs well. The supervised factor model (model 10) performs well in forecasting unemployment rate, real disposable income, and the federal funds rate. These findings are somewhat similar to the ones in Stock and Watson (2002b), in which it was found that the factors have more predictive power for real variables rather than for inflation measures. The results are robust to the choice of sample split as can be seen from tables 1-9.

In table 10, in Appendix A, are reported the ratios between the variance of the contribution to the filtered factors of the forecast target and the total variance of the filtered factors, equation (19), for all variables. We notice a positive relation between the value of this ratio and the forecast performance of the supervised factor model. For example CPI has a much lower impact on the filtered factors compared to FFR and UR. A possible interpretation is that the forecast objectives which influence more the extraction of the unobserved factors benefit more from the supervised framework. This suggests that the supervised factors may contain additional information with respect to unsupervised ones.

From tables 1-9 it remains unclear what the best number of factors is, in terms of forecasting performance. The best number of factors seems to change with the forecast target and sample split.

6 Simulations

We perform two simulation experiments according to two different data generating processes (hereafter DGPs). We simulate a state-space system according to equations (1) with different loading coefficients:

case 1 discrete cosine basis as loadings, equation (2);

case 2 random loadings, generated as independent draws from a normal distribution $N(0, 1)$.

In both cases, the state vector follows a three dimensional, stable VAR(1). The first two components of the state vector $f_{t,1}$ and $f_{t,2}$, are treated as latent factors whereas the third component $f_{t,3}$, is regarded as the forecast objective. We simulate the system under different correlations between the factors and the forecast objective, namely $\rho_{f_1,y}$ and $\rho_{f_2,y}$, by restricting the unconditional variance-covariance matrix of the state vector. For pairs of indexes $\{i; j\} = \{1; 2\}$ and $\{i; j\} = \{2; 1\}$ we fix the correlations $\rho_{f_i,y} = 0.5$ and $\rho_{f_i,f_j} = 0.1$ and let $\rho_{f_j,y} \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$. We then compute forecasts using models 8, 9, and 10 previously defined and reported again here for convenience:

model 8. Unsupervised factor model (25) and (26) with discrete cosine basis factor loadings, eqn. (2);

model 9. Unsupervised factor model as in Section 2.2 with discrete cosine basis factor loadings, eqn. (2);

model 10. Supervised factor model as in eqn. (1) with discrete cosine basis factor loadings (2).

The complete sample size in the simulations is $T = 600$, the rolling window has size $R = 289$, and the number of forecasts is $S = 300$.

6.1 Simulations results

In this subsection we present results relative to the simulation exercise. The mean square prediction error ratios between forecasts from the supervised model and the competing models can be found in tables 11-12 in Appendix A. The supervised factor model corresponds to equations (1) with discrete cosine basis as loadings, equation (2). Looking at tables 11 and 12, we can make the following remarks divided according to the DGP and correlations.

I **discrete cosine basis loadings**

(i) **varying $\boldsymbol{\rho_{f_1,y}}$ and fixed $\boldsymbol{\rho_{f_2,y}}$.** In around 50% of the cases the supervised factor model performs better than the unsupervised counterparts (in 34% of which the difference is statistically significant at the $\alpha = 0.10$ significance level), in 30% of the cases it delivers the same forecasting performance as the other two methods and in the remaining 20% of cases it delivers slightly worse forecasts (in 11% of which the difference is statistically significant at the $\alpha = 0.10$ significance level);

(ii) **varying $\boldsymbol{\rho_{f_2,y}}$ and fixed $\boldsymbol{\rho_{f_1,y}}$.** In around 60% of the cases the supervised factor model performs better than the unsupervised counterparts (in 30% of which the difference is statistically significant at the $\alpha = 0.10$ significance level), in 25% of cases it delivers the same forecast performance as the other two methods and in the remaining 15% of cases it delivers slightly worse forecasts (in 21% of which the difference is statistically significant at the $\alpha = 0.10$ significance level).

II **random loadings**

(i) **varying $\boldsymbol{\rho_{f_1,y}}$ and fixed $\boldsymbol{\rho_{f_2,y}}$.** In around 58% of the cases the supervised factor model performs better than the unsupervised counterparts (in 34% of which the difference is statistically significant at the $\alpha = 0.10$ significance level), in 19% of cases it delivers the same forecast performance as the other two methods and in the remaining 23% of cases it delivers slightly worse forecasts (in 42% of which the difference is statistically significant at the $\alpha = 0.10$ significance level);

(ii) **varying $\boldsymbol{\rho_{f_2,y}}$ and fixed $\boldsymbol{\rho_{f_1,y}}$.** In around 75% of the cases the supervised factor model performs better than the unsupervised counterparts (in 28% of which the difference is statistically significant at the $\alpha = 0.10$ significance level), in 15% of cases it delivers the same forecast performance as the other two methods and in the remaining 10% of cases it delivers slighly worse forecasts (in 11% of which the difference is statistically significant at the $\alpha = 0.10$ significance level).

Furthermore, we notice that even with moderate levels of correlation between the forecast objective and the factors, the supervised specification delivers on average better forecasts with respect to the unsupervised counterparts.

7 Conclusions

In this paper we study the forecasting properties of a supervised factor model. In this framework the factors are extracted conditionally on the forecast target. The model has a linear state-space representation and standard Kalman filtering techniques can be used. Under this setup, we propose a way to measure the contribution of the forecast objective on the extracted factors that exploits the Kalman filter recursions. In particular, we compute the contribution of the forecast target to the variance of the filtered factors and find a positive correspondence between this quantity and the forecast performance of the supervised scheme.

We assess the forecast performance of the supervised factor model with a simulation study and an empirical application. The simulated data are generated according to different levels of correlation between the forecast objective and the factors. In the simulations experiment, we find that if the forecast objective is correlated with the factors the supervised factor model improves, on average, forecast performance compared to unsupervised schemes.

In the empirical application the supervised FM is used to forecast macroeconomic variables using factors extracted from a large number of predictors. The macroeconomic data are taken from the Jurado, Ludvigson and Ng dataset and FRED. We estimate the model considering one, two, and three factors. We forecast consumer price index (CPI), the federal funds rate (FFR), personal consumption expenditures deflator (PCEd), the producer price index (PPI), personal income (PEI), the unemployment rate (UR), industrial production (IP), the real disposable income (RDI), and personal consumption expenditures (PCE) relative to the US economy.

We find that supervising the factor extraction can improve forecasting performance compared to unsupervised factor models and other popular multivariate and univariate forecasting models. For this dataset and specification the supervised factor model outperforms partial least squares regressions and principal components regressions on most targets. In forecasting inflation, both measured by consumer price index and producer price index, $MA(q)$ processes are difficult to beat whereas the supervised factor model performs particularly well in forecasting the federal funds rate, the unemployment rate, and real disposable income. These findings are similar to the ones in Stock and Watson (2002b), in which it was found that the factors have more predictive power for real variables rather than for inflation measures.

We find that variables which contribute more to the variance of the filtered states, i.e. a higher $r_t^{j,n}$, equation (19), are the ones which benefit more from the supervised framework and vice versa. Furthermore, supervising the factor extraction leads in most cases to improved forecasts, compared to unsupervised two-step forecasting schemes.

8 Acknowledgements

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

References

Ahn, S. C. and A. R. Horenstein (2013). Eigenvalue ratio test for the number of factors. *Econometrica 81*(3), 1203–1227.

Alessi, L., M. Barigozzi, and M. Capasso (2007). Dynamic factor garch: multivariate volatility forecast for a large number of series. Technical report, LEM Working Paper Series.

Alessi, L., M. Barigozzi, and M. Capasso (2010). Improved penalization for determining the number of factors in approximate factor models. *Statistics & Probability Letters 80*(23), 1806–1813.

Amengual, D. and M. W. Watson (2007). Consistent estimation of the number of dynamic factors in a large n and t panel. *Journal of Business & Economic Statistics 25*(1), 91–96.

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

Bai, J. and S. Ng (2002). Determining the number of factors in approximate factor models. *Econometrica 70*(1), 191–221.

Bai, J. and S. Ng (2007). Determining the number of primitive shocks in factor models. *Journal of Business & Economic Statistics 25*(1).

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).

Banerjee, A., M. G. Marcellino, and I. Masten (2005). Forecasting macroeconomic variables for the new member states of the european union.

Bernanke, B. S., J. Boivin, and P. Eliasz (2005). Measuring the effects of monetary policy: a factor-augmented vector autoregressive (favar) approach. *The Quarterly Journal of Economics 120*(1), 387–422.

Boivin, J. and S. Ng (2005). Understanding and comparing factor-based forecasts. Technical report, National Bureau of Economic Research.

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

Brockwell, P. J. and R. A. Davis (2002). *Introduction to time series and forecasting*, Volume 1. Taylor & Francis.

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

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

d'Agostino, A. and D. Giannone (2012). Comparing alternative predictors based on large-panel factor models. *Oxford bulletin of economics and statistics 74*(2), 306–326.

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

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

Engel, C., N. C. Mark, and K. D. West (2012). Factor model forecasts of exchange rates. Technical report, National Bureau of Economic Research.

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

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000). The generalized dynamic factor model: One sided estimation and forecasting. *Journal of the American Statistical Association*, 830–840.

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

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.

Hallin, M. and R. Liška (2007). Determining the number of factors in the general dynamic factor model. *Journal of the American Statistical Association 102*(478), 603–617.

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.

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

Jurado, K., S. C. Ludvigson, and S. Ng (2013). Measuring uncertainty. Technical report, National Bureau of Economic Research.

Jurado, K., S. C. Ludvigson, and S. Ng (2015). Measuring uncertainty. *3*(105), 1177–1216.

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

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.

Mallat, S. (1999). *A wavelet tour of signal processing.* Academic press.

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.

Schumacher, C. and C. Dreger (2002). Estimating large-scale factor models for economic activity in germany: Do they outperform simpler models?

Stock, J. H. and M. W. Watson (1999). 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 (2006). Forecasting with many predictors. *Handbook of economic forecasting 1*, 515–554.

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 5) and the simulation exercise (see Section 6). The results relative to the empirical application correspond to MSFE ratios between model 10 and the competing models (see Section 5 for the description of the different models involved) and are contained in tables 1-9. We consider different subsamples of the dataset and estimate the factor models using 1, 2, and 3 factors. The results relative to the simulation exercise correspond to MSFE ratios between model 10 and models 8 and 9 and are contained in tables 11-12. We estimate the factor models using 2 factors. 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. In table 10 are reported the ratios between the variance of the contribution to the filtered factors of the forecast target and the total variance of the filtered factors, equation (19), for all variables.

Table 1: MSFE ratios for whole forecast sample (1 factor).

	h	mod 1	mod 2	mod 3	mod 4	mod 5	mod 6	mod 7	mod 8	mod 9
	1	**0,98**	1,04	1,11	1,34*	1,11	1,34*	**0,99**	**0,94**	**0,99**
	3	**0,98**	**0,9**	**0,99**	**0,99**	**0,96**	1,01	**0,99**	**0,99**	**0,99**
CPI	6	**0,99**	**0,82****	**0,99**	**0,95**	1	1	**0,99**	1	**0,99**
	9	**0,98**	**0,79****	1,01	1,01	1	1	**0,98**	1	**0,98**
	12	1,02	**0,83*****	1,01	1,01	1	1	1,02	1	1,02
	1	**0,78*****	**0,47*****	**0,95**	1	**0,95**	1	1,11	1,03	**0,98*****
	3	**0,75*****	**0,32*****	**0,68*****	**0,65*****	**0,73*****	**0,79****	**0,82**	**0,94**	**0,85****
FFR	6	**0,72*****	**0,26*****	**0,89****	**0,87*****	**0,94***	**0,92****	**0,83**	**0,97**	**0,85*****
	9	**0,86**	**0,37*****	1,02	1,03	**0,97*****	**0,98*****	**0,98**	**0,99**	1
	12	**0,75*****	**0,37*****	**0,91****	**0,92***	**0,99**	**0,99**	**0,87**	1	**0,89*****
	1	**0,94****	**0,99**	1,15	1,31*	1,15	1,31*	**0,99**	**0,91***	1
	3	1	**0,91**	**0,99**	**0,99**	**0,98**	1	**0,99**	**0,98**	**0,99**
PCEd	6	**0,99**	**0,83*****	**0,99**	**0,99**	**0,99***	1	**0,99**	1	**0,99**
	9	**0,99**	**0,87**	1,01	1	1	1	1	1	1
	12	1	**0,85*****	**0,97**	**0,97****	1	1	**0,98**	1	**0,97**
	1	**0,91****	**0,88**	1,1	1,27**	1,1	1,27**	1	**0,79*****	1
	3	**0,98**	**0,85*****	**0,97**	**0,97**	**0,98**	1	**0,98**	**0,98**	**0,98**
PPI	6	**0,99**	**0,82**	1	1	1,01	1	1	1	1
	9	**0,99**	**0,79**	**0,97**	**0,97**	1	1	**0,98**	1	**0,98**
	12	1,04	**0,9***	1,05	1,03	1	1	1	1	1
	1	1,03	**0,95**	**0,94***	**0,94****	**0,94***	**0,94****	1,04	1,03	1
	3	1,02	**0,92**	**0,96****	**0,95****	**0,97***	**0,97***	1,03	1,01	1,02
PEI	6	1,01	**0,89***	**0,99*****	**0,99*****	**0,99****	**0,99****	1,02	1	1,02
	9	1	**0,91****	**0,97**	**0,96**	1	1	1,01	1	1,01
	12	**0,99**	**0,84**	**0,97**	**0,98**	1	1	**0,98**	1	**0,97**
	1	**0,97**	1,02	**0,9**	**0,85***	**0,9**	**0,85***	1,06	**0,66*****	1
	3	1,04	1,01	**0,99**	**0,94****	**0,99**	**0,92****	1,05	**0,77*****	1,06
UR	6	1,05	**0,83**	**0,96**	**0,99**	**0,99**	**0,95****	1,05	**0,92****	1,05
	9	1,03	**0,89**	**0,95**	1	**0,99**	**0,98***	**0,99**	**0,97***	1,01
	12	**0,98*****	**0,82**	**0,96****	**0,95****	**0,99**	1**	**0,94**	**0,99****	**0,95****
	1	**0,92***	**0,99**	**0,9**	**0,87**	**0,9**	**0,87**	**0,97**	**0,94**	**0,99*****
	3	1,05	**0,98**	1,08	1,07	1,06	**0,97**	1,04	**0,96**	1,02
IP	6	1,02	**0,93**	1,01	1,02	1	**0,98**	**0,98**	**0,99**	**0,99**
	9	**0,96**	**0,89***	**0,99**	**0,98**	**0,99**	**0,99**	**0,95**	**0,99**	**0,95****
	12	**0,97**	**0,82****	**0,97****	**0,98*****	**0,99**	1	**0,95**	1	**0,95**
	1	**0,96**	**0,8*****	**0,99**	1,01	**0,99**	1,01	1,01	**0,95**	1
	3	**0,99**	**0,82*****	**0,99**	**0,99**	**0,98**	**0,99**	1	1	1*
RDI	6	**0,99**	**0,78****	**0,99**	**0,99**	1	1	1	1	1
	9	**0,99**	**0,87*****	1	1	1	1	1,01	1	1,01
	12	1	**0,85*****	1	1	1	1	1	1	1
	1	**0,63*****	1,01	1,38**	1,72***	1,38**	1,72***	**0,99**	**0,58*****	1
	3	1,05	**0,84****	1,04	1,04	**0,99**	1,06	1,03	1,05	1,03
PCE	6	**0,99**	**0,75*****	**0,99**	**0,99**	**0,99**	1	**0,99**	1	**0,99**
	9	**0,98**	**0,68*****	1,01	1	1	1	1	1	1
	12	**0,97**	**0,75*****	1,01	1,01	1	1	1,01	1	1,01

Jurado et al. (2013) dataset. MSFE ratios between model 10 and competing models for CPI, FFR, PCEd, PPI, PEI, UR, IP, RDI, and PCE for forecasting leads h. A value lower than one indicates a lower MSFE of model 10 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. Number of forecasts is $S = 300$. The number of factors in the methods involving factor models is 1. The 1-step ahead forecasts range from February 1986 to January 2011. The 12-step ahead forecasts range from January 1987 to December 2011.

Table 2: MSFE ratios for first half of forecast sample (1 factor).

	h	mod 1	mod 2	mod 3	mod 4	mod 5	mod 6	mod 7	mod 8	mod 9
	1	**0,96***	**0,92**	1,05	1,14	1,05	1,14	**0,99**	**0,92**	1
	3	**0,99***	**0,74*****	1	**0,99**	**0,98**	1,01*	**0,99**	1	1
CPI	6	1,01	**0,63****	1	1	1	1	1	1	1
	9	**0,99**	**0,66*****	1	1,01	1	1	1,01	1	1,01
	12	1,01	**0,62*****	1,01	1	1	1	1	1	1
	1	**0,84**	**0,47*****	**0,9**	**0,92**	**0,9**	**0,92**	1,09*	1,06	1
	3	**0,79****	**0,32*****	**0,74****	**0,73****	**0,77***	**0,81**	**0,91**	**0,94**	**0,94**
FFR	6	**0,74**	**0,32*****	**0,89***	**0,88***	**0,94***	**0,94***	**0,94**	**0,97**	**0,95**
	9	**0,91**	**0,51****	1,02	1,02	**0,98*****	**0,98****	1,07	**0,99*****	1,07
	12	**0,77**	**0,39*****	**0,93**	**0,93**	**0,99**	**0,99**	**0,91**	**0,99**	**0,93**
	1	**0,92***	1	1,15**	1,3***	1,15**	1,3***	**0,99**	**0,89*****	1
	3	1,01**	**0,93**	**0,98****	**0,99****	**0,98**	1,01	**0,98**	1,01	**0,98***
PCEd	6	1	**0,74*****	**0,96**	**0,96**	1	1*	**0,96**	1	**0,96**
	9	1	**0,84**	1,01	1,01	1	1	1,01	1	1,01
	12	1	**0,75*****	1	1	1	1	1	1	**0,99**
	1	**0,99**	**0,85**	1	1,2	1	1,2	1	**0,9**	1
	3	**0,99**	**0,7*****	**0,92***	**0,92**	**0,94**	**0,99**	**0,97**	**0,99***	**0,97**
PPI	6	1,01	**0,79****	**0,96**	**0,95**	1	1	**0,96**	1	**0,96**
	9	1	**0,75*****	**0,94**	**0,95**	1	1	**0,99**	1	**0,99**
	12	**0,97**	**0,7**	**0,99****	**0,99****	1*	1	**0,99**	1	**0,99***
	1	1	**0,91**	**0,96**	**0,97**	**0,96**	**0,97**	1,01	1,01	1
	3	1	**0,98**	**0,98**	**0,98**	**0,99**	**0,99**	1	1	1
PEI	6	**0,99**	**0,94*****	**0,99*****	**0,99*****	**0,99*****	**0,99*****	1	1**	1
	9	**0,98**	**0,94**	**0,96**	**0,95**	1	1	**0,98**	1	**0,99**
	12	1	**0,96*****	**0,99**	1,02	1	1	1,04	1	1,03
	1	**0,95**	**0,97**	**0,91**	**0,92***	**0,91**	**0,92***	1,02	**0,85**	1
	3	**0,98**	1	1,01	1	1,03	**0,99***	1	**0,92**	1
UR	6	1,01*	**0,91**	1	**0,99**	1,01	**0,98***	1,02**	**0,97***	1,01**
	9	1,06	**0,93**	1	1	1,01	1	1,01	1	1
	12	1,01**	**0,83**	**0,95*****	**0,92****	1	1*	**0,95*****	1**	**0,94****
	1	**0,96**	**0,95**	**0,89**	**0,93**	**0,89**	**0,93**	**0,96**	1,02	**0,97****
	3	1,07*	**0,79*****	1,04	1,02	1,01	1	1	**0,99**	**0,97**
IP	6	1,03	**0,82***	1	1	1	**0,99**	1	1	**0,96***
	9	**0,98**	**0,8****	1,02**	1,01***	1	1	1,02*	1	**0,98**
	12	1,05	**0,78*****	**0,99****	**0,98*****	1	1	1,01	1	1
	1	**0,93**	**0,86*****	**0,99**	**0,99**	**0,99**	**0,99**	1	**0,93**	1
	3	1	**0,91**	**0,99**	**0,99**	1	**0,99**	**0,99**	1	**0,99**
RDI	6	**0,99**	**0,82*****	1	1	1	1	1	1	1,01
	9	1	**0,92***	1	1	1*	1**	1	1*	1**
	12	**0,99**	**0,88*****	1	1	1	1	1	1	1
	1	**0,67**	1,15	1,32	1,72***	1,32	1,72***	**0,99***	**0,59***	1
	3	1,08	**0,97**	1,03	1,03	**0,96**	1,07	1,03	1,08	1,03
PCE	6	1	**0,85*****	**0,99**	**0,98**	1*	1	**0,99**	1	**0,99**
	9	**0,99**	**0,83****	1,02	1**	1	1	1**	1	1
	12	**0,99**	**0,84*****	1,02	1,03	1*	1	1,02	1*	1,02

Jurado et al. (2013) dataset. MSFE ratios between model 10 and competing models for CPI, FFR, PCEd, PPI, PEI, UR, IP, RDI, and PCE for forecasting leads h. A value lower than one indicates a lower MSFE of model 10 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. Number of forecasts is $S' = 150$ (the first half of the $S = 300$ out-of-sample forecasts). The number of factors in the methods involving factor models is 1. The 1-step ahead forecasts range from February 1986 to July 1998. The 12-step ahead forecasts range from January 1987 to June 1999.

Table 3: MSFE ratios for second half of forecast sample (1 factor).

	h	mod 1	mod 2	mod 3	mod 4	mod 5	mod 6	mod 7	mod 8	mod 9
	1	**0,99**	1,07	1,13	1,4	1,13	1,4	**0,99**	**0,95**	**0,99**
	3	**0,98**	**0,95**	**0,99**	**0,99**	**0,96**	1,01	**0,98**	**0,98**	**0,99**
CPI	6	**0,98**	**0,89***	**0,99**	**0,94**	1	1	**0,99**	1	**0,99**
	9	**0,98**	**0,83**	1,01	1,01	1	1	**0,98*****	1	**0,98****
	12	1,02	**0,91*****	1,01	1,02	1	1	1,03	1	1,03
	1	**0,72****	**0,47*****	1,03	1,14	1,03	1,14	1,14	**0,99**	**0,97*****
	3	**0,69*****	**0,32*****	**0,62****	**0,57****	**0,69***	**0,77**	**0,72****	**0,94**	**0,77****
FFR	6	**0,69*****	**0,21*****	**0,9**	**0,85****	**0,94**	**0,9**	**0,71****	**0,97**	**0,74****
	9	**0,81**	**0,27***	1,01	1,05	**0,97**	**0,98**	**0,9**	1	**0,92**
	12	**0,73*****	**0,36*****	**0,89**	**0,91**	1	1	**0,83****	1,01*	**0,84****
	1	**0,95**	**0,99**	1,15	1,31	1,15	1,31	**0,99***	**0,92**	**0,99**
	3	**0,99**	**0,9**	1	1	**0,98**	**0,99**	**0,99**	**0,97**	**0,99**
PCEd	6	**0,99**	**0,87***	1,01	1,01	**0,98***	1	1	1	1
	9	**0,98***	**0,88**	1,01	1	1	1	1	1	1
	12	1,01	**0,91**	**0,95**	**0,95****	1*	1	**0,97**	1	**0,97***
	1	**0,9****	**0,89**	1,13	1,29*	1,13	1,29*	**0,99**	**0,76*****	**0,99**
	3	**0,98**	**0,89***	**0,98**	**0,98**	**0,99**	1	**0,99**	**0,98**	**0,98**
PPI	6	**0,98**	**0,82**	1,01	1,01	1,01	1	1	1	1
	9	**0,99**	**0,8**	**0,97**	**0,98**	1	1	**0,98**	1	**0,97**
	12	1,06	**0,95**	1,06	1,04	1	1	1	1	1
	1	1,08	1,01	**0,91**	**0,9**	**0,91**	**0,9**	1,08	1,06	1,01
	3	1,04	**0,85**	**0,92***	**0,92***	**0,94***	**0,94***	1,08	1,03	1,05
PEI	6	1,04	**0,82**	**0,98****	**0,98****	**0,98**	**0,98**	1,05	1,01	1,04
	9	1,03	**0,87****	**0,98**	**0,98**	**0,99**	**0,99**	1,05*	1	1,04**
	12	**0,99**	**0,72**	**0,93**	**0,92**	1	1	**0,91**	1	**0,91**
	1	1	1,07	**0,88**	**0,79***	**0,88**	**0,79***	1,11	**0,53*****	1
	3	1,1	1,02	**0,98**	**0,9*****	**0,95****	**0,87****	1,1	**0,68****	1,11
UR	6	1,07	**0,78**	**0,94**	**0,99**	**0,98**	**0,93****	1,07	**0,88****	1,09
	9	1,01	**0,86**	**0,92**	**0,99**	**0,98****	**0,97***	**0,98**	**0,95***	1,02
	12	**0,95***	**0,81**	**0,97**	**0,97**	**0,99**	**0,99***	**0,94**	**0,99***	**0,95**
	1	**0,9***	1,01	**0,9**	**0,84**	**0,9**	**0,84**	**0,98**	**0,9**	1
	3	1,04	1,13	1,11	1,09	1,09	**0,96**	1,06	**0,95**	1,05
IP	6	1,01	1	1,02	1,03	1,01	**0,97**	**0,97**	**0,99**	1
	9	**0,95**	**0,94**	**0,98***	**0,97***	**0,98**	**0,99**	**0,92*****	**0,99**	**0,94****
	12	**0,94***	**0,85**	**0,97***	**0,97**	**0,98**	1	**0,93**	1	**0,93**
	1	**0,98**	**0,75****	1	1,03	1	1,03	1,01	**0,96**	1
	3	**0,99**	**0,76****	1	1	**0,97**	**0,99**	1,01	1	1,01*
RDI	6	**0,99**	**0,75**	**0,99**	**0,99**	**0,99**	1	1	1,01	1
	9	**0,98**	**0,83*****	1,01	1,01	1	1	1,02	1	1,02
	12	1	**0,83*****	1	**0,99**	1	1	**0,99**	1	1
	1	**0,56****	**0,85**	1,48**	1,71**	1,48**	1,71**	1	**0,56***	1,01
	3	1,01	**0,69****	1,05	1,05	1,03	1,04	1,04	1	1,04
PCE	6	**0,97**	**0,63*****	**0,99**	**0,99**	**0,99**	**0,99**	**0,99**	**0,99**	**0,99**
	9	**0,97**	**0,53*****	1	1	1*	1	1	1	1
	12	**0,95**	**0,67*****	1	1	1	1	1	1	**0,99**

Jurado et al. (2013) dataset. MSFE ratios between model 10 and competing models for CPI, FFR, PCEd, PPI, PEI, UR, IP, RDI, and PCE for forecasting leads h. A value lower than one indicates a lower MSFE of model 10 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. Number of forecasts is $S' = 150$ (the second half of the $S = 300$ out-of-sample forecasts). The number of factors in the methods involving factor models is 1. The 1-step ahead forecasts range from August 1998 to January 2011. The 12-step ahead forecasts range from July 1999 to December 2011.

Table 4: MSFE ratios for whole forecast sample (2 factors).

	h	mod 1	mod 2	mod 3	mod 4	mod 5	mod 6	mod 7	mod 8	mod 9
	1	**0,98**	1,04	1,12	1,34*	1,12	1,34*	**0,99**	**0,99**	1
	3	**0,98**	**0,9**	**0,99***	**0,99***	**0,96**	1	**0,97**	1	**0,98***
CPI	6	**0,99**	**0,82****	**0,99**	**0,95**	1	1	**0,98**	1	**0,99**
	9	**0,98**	**0,79****	1,01	1,01	1	1	**0,98**	1	**0,97**
	12	1,02	**0,83*****	1,01	1,01	1	1	1,03	1	1,02
	1	1,37***	**0,83**	1,66***	1,75***	1,66***	1,75***	1,88	1,68***	1
	3	**0,78*****	**0,33*****	**0,71*****	**0,67*****	**0,76*****	**0,82****	**0,81**	**0,91**	**0,84*****
FFR	6	**0,74*****	**0,27*****	**0,92***	**0,89*****	**0,97***	**0,95****	**0,83**	**0,97***	**0,84*****
	9	**0,87**	**0,37*****	1,03	1,04	**0,98*****	**0,99*****	**0,96**	**0,99*****	**0,93**
	12	**0,75*****	**0,37*****	**0,91****	**0,92***	1	1	**0,85**	1	**0,86*****
	1	**0,94**	**0,99**	1,15	1,31*	1,15	1,31*	1	**0,95**	1
	3	1	**0,91**	**0,99**	1	**0,98**	1	**0,99**	1	**0,98**
PCEd	6	**0,99**	**0,83*****	**0,99**	**0,99**	**0,99***	1	**0,99**	1	1
	9	**0,99**	**0,87**	1,01	1	1	1	1	1	**0,99**
	12	1	**0,85*****	**0,97**	**0,97****	1	1	**0,98**	1	**0,97**
	1	**0,92***	**0,89**	1,11	1,28**	1,11	1,28**	1,01	**0,84*****	**0,99**
	3	**0,98**	**0,85*****	**0,98**	**0,98**	**0,98**	1	**0,97**	1	**0,99**
PPI	6	**0,99**	**0,82**	1	1	1,01	1	**0,99**	1	**0,99**
	9	**0,99**	**0,79**	**0,97**	**0,97**	1	1	**0,98**	1	**0,98**
	12	1,04	**0,9***	1,05	1,03	1	1	1,02	1	1,02
	1	1	**0,92**	**0,91**	**0,91**	**0,91**	**0,91**	1,01	**0,96**	1
	3	1,02	**0,93**	**0,96***	**0,96****	**0,97**	**0,97**	1,02	1,03	1,02
PEI	6	1,02	**0,89**	**0,99*****	**0,99*****	**0,99***	**0,99***	1,01	1,01	1,01
	9	1	**0,91****	**0,97**	**0,97**	1	1	1,01	1	1,01
	12	1	**0,84**	**0,97**	**0,98**	1	1	**0,97**	1	**0,98**
	1	**0,98**	1,02	**0,9**	**0,85***	**0,9**	**0,85***	1,04	**0,56*****	**0,99**
	3	1,06	1,03	1,02	**0,97****	1,01	**0,94***	1,09	**0,75*****	1,07
UR	6	1,06	**0,84**	**0,98**	1,01	1,01	**0,97****	1,05	**0,93****	1,06
	9	1,04	**0,9**	**0,96**	1,01	1	**0,99****	1,02	**0,98****	**0,99**
	12	**0,98*****	**0,82**	**0,96****	**0,95****	1	1**	**0,95**	**0,99****	**0,95****
	1	**0,94**	1,01	**0,92**	**0,89**	**0,92**	**0,89**	**0,99**	1	**0,95*****
	3	1,05	**0,98**	1,08	1,06	1,06	**0,97**	1,04	**0,98**	1,01
IP	6	1,04	**0,95**	1,03	1,04	1,02	1	1,01	1,01	**0,98**
	9	**0,99**	**0,93***	1,03	1,02	1,03	1,03	**0,98**	1,03	1
	12	1,01	**0,86****	1,01	1,02	1,03	1,04	**0,99**	1,04	**0,99**
	1	**0,95**	**0,79*****	**0,99**	1,01	**0,99**	1,01	1	**0,97**	1
	3	**0,99**	**0,82*****	**0,99**	**0,99**	**0,98**	**0,99**	1	**0,99**	**0,98**
RDI	6	**0,99**	**0,78****	**0,99**	**0,99**	1	1	**0,98**	1	**0,99**
	9	**0,99**	**0,87*****	1	1	1*	1	1	1	1
	12	1	**0,85*****	1	1	1	1	**0,99**	1	1
	1	**0,63*****	1,02	1,39**	1,74***	1,39**	1,74***	1	**0,63****	1
	3	1,05	**0,83****	1,03	1,03	**0,98**	1,05	1,03	1,05	1,02
PCE	6	**0,99**	**0,75*****	**0,99**	**0,99**	**0,99***	1	**0,99**	1	**0,97**
	9	**0,98**	**0,68*****	1,01	1	1	1	**0,98**	1	1
	12	**0,97**	**0,75*****	1,01	1,02	1	1	1,01	1	**0,99**

Jurado et al. (2013) dataset. MSFE ratios between model 10 and competing models for CPI, FFR, PCEd, PPI, PEI, UR, IP, RDI, and PCE for forecasting leads h. A value lower than one indicates a lower MSFE of model 10 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. Number of forecasts is $S = 300$. The number of factors in the methods involving factor models is 2. The 1-step ahead forecasts range from February 1986 to January 2011. The 12-step ahead forecasts range from January 1987 to December 2011.

Table 5: MSFE ratios for first half of forecast sample (2 factors).

	h	mod 1	mod 2	mod 3	mod 4	mod 5	mod 6	mod 7	mod 8	mod 9
	1	**0,94**	**0,9**	1,03	1,12	1,03	1,12	**0,97**	**0,94**	1
	3	**0,99***	**0,74*****	1	**0,98**	**0,98**	1	**0,99****	1,01	**0,98**
CPI	6	1,01	**0,63****	1	1	1	1	1	1	**0,99**
	9	**0,99**	**0,66*****	1	1,01	1	1	1	1	**0,98**
	12	1,01	**0,62*****	1,01	1	1	1	1	1	1,01
	1	1,34**	**0,75***	1,43**	1,47***	1,43**	1,47***	1,75***	1,59***	1,01
	3	**0,81****	**0,33*****	**0,76****	**0,74****	**0,79****	**0,83***	**0,89**	**0,91**	**0,93**
FFR	6	**0,76**	**0,33*****	**0,91**	**0,9***	**0,96***	**0,96***	**0,92**	**0,98***	**0,93**
	9	**0,92**	**0,51****	1,03	1,03	**0,99*****	**0,99****	1,08	**0,99*****	1,01
	12	**0,77**	**0,39*****	**0,94**	**0,93**	**0,99**	**0,99**	**0,87**	**0,99**	**0,91**
	1	**0,92***	1	1,14**	1,29***	1,14**	1,29***	**0,99**	**0,9*****	1
	3	1,01**	**0,93**	**0,98****	**0,99***	**0,98**	1,01	**0,98**	1,01**	**0,98**
PCEd	6	1	**0,74*****	**0,96**	**0,96**	1	1*	**0,96**	1	**0,96**
	9	1	**0,84**	1,01	1,01	1	1	1,01	1	1,01
	12	1	**0,75*****	1	1	1	1	1	1**	**0,99**
	1	**0,99**	**0,85***	**0,99**	1,19	**0,99**	1,19	1,01	**0,94**	**0,99**
	3	1	**0,71*****	**0,93**	**0,93**	**0,96**	1	**0,98**	1	**0,97**
PPI	6	1,01	**0,79****	**0,96**	**0,95**	1	1	**0,96**	1	**0,96**
	9	1	**0,75*****	**0,94**	**0,95**	1	1	**0,99**	1	**0,99**
	12	**0,97**	**0,7**	**0,99****	**0,99****	1	1	**0,99**	1	**0,94**
	1	1	**0,91**	**0,95**	**0,96**	**0,95**	**0,96**	1	**0,98**	1
	3	1	**0,99**	**0,99**	**0,99**	1	1	1,01	1	**0,99**
PEI	6	**0,99**	**0,94*****	**0,99*****	**0,99*****	1***	1***	**0,98**	1*	1
	9	**0,98**	**0,94**	**0,96**	**0,95**	1	1	**0,99**	1	**0,98**
	12	1	**0,96*****	**0,99**	1,02	1	1	1	1	1,03
	1	**0,97**	**0,99**	**0,93**	**0,94**	**0,93**	**0,94**	1,03	**0,78***	1
	3	**0,98**	1	1,02	1	1,04	**0,99***	1	**0,9***	1,01
UR	6	1,02*	**0,91**	1,01	**0,99**	1,01	**0,98****	1,01**	**0,97****	1,01*
	9	1,06	**0,93**	1,01	1,01	1,01	1	1,09	1	**0,95**
	12	1,01**	**0,83**	**0,95*****	**0,92****	1	1	**0,98*****	1*	**0,94*****
	1	**0,95**	**0,94**	**0,88**	**0,92**	**0,88**	**0,92**	**0,95**	1,03	**0,96****
	3	1,08*	**0,8*****	1,05	1,03	1,02	1,02	1,06	1,02	**0,98**
IP	6	1,03	**0,82***	1	1	1	**0,99***	1,01	1	**0,97**
	9	**0,97**	**0,8****	1,02*	1,01**	1	1	1,01	1	**0,98**
	12	1,05	**0,78*****	**0,99***	**0,98****	1	1	1,05	1	1,02
	1	**0,93**	**0,86*****	**0,99**	**0,99**	**0,99**	**0,99**	**0,99**	**0,94**	1
	3	1	**0,91**	**0,99**	**0,99**	1	**0,99**	**0,99**	1	**0,96**
RDI	6	**0,99**	**0,82*****	1	1	1	1	1	1	1
	9	1	**0,92***	1	1	1*	1**	1	1*	1,01
	12	**0,99**	**0,88*****	1	1	1	1	**0,99**	1	1,01
	1	**0,68**	1,17	1,34	1,75***	1,34	1,75***	1,01**	**0,65**	**0,99****
	3	1,07	**0,97**	1,02	1,03	**0,96**	1,06	1,02	1,07	1,03
PCE	6	1,01	**0,85*****	**0,99**	**0,98**	1*	1	**0,99**	1	**0,99**
	9	**0,99**	**0,83****	1,02	1**	1	1	**0,99*****	1	1
	12	**0,99**	**0,84*****	1,02	1,03	1	1	1,01	1	1

Jurado et al. (2013) dataset. MSFE ratios between model 10 and competing models for CPI, FFR, PCEd, PPI, PEI, UR, IP, RDI, and PCE for forecasting leads h. A value lower than one indicates a lower MSFE of model 10 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. Number of forecasts is $S' = 150$ (first half of the $S = 300$ out-of-sample forecasts). The number of factors in the methods involving factor models is 2. The 1-step ahead forecasts range from February 1986 to July 1998. The 12-step ahead forecasts range from January 1987 to June 1999.

Table 6: MSFE ratios for second half of forecast sample (2 factors).

	h	mod 1	mod 2	mod 3	mod 4	mod 5	mod 6	mod 7	mod 8	mod 9
	1	1	1,08	1,14	1,41	1,14	1,41	1	1,01	**0,99**
	3	**0,97**	**0,95**	**0,99**	**0,99**	**0,96**	1	**0,96**	1	**0,98**
CPI	6	**0,98**	**0,89***	**0,99**	**0,94**	1	1	**0,98**	1	**0,99**
	9	**0,98**	**0,83**	1,01	1,01	1	1	**0,98***	1	**0,97*****
	12	1,02	**0,91*****	1,01	1,02	1*	1	1,03	1	1,02
	1	1,4**	**0,92**	2,01***	2,23***	2,01***	2,23***	2,06***	1,8***	**0,99****
	3	**0,73*****	**0,33*****	**0,65****	**0,6****	**0,73***	**0,82**	**0,73****	**0,92**	**0,76*****
FFR	6	**0,71****	**0,22*****	**0,93**	**0,88****	**0,97**	**0,93**	**0,73****	**0,97**	**0,74*****
	9	**0,82**	**0,27***	1,02	1,06	**0,98***	**0,98***	**0,84**	1	**0,85**
	12	**0,73*****	**0,36*****	**0,89**	**0,91**	1	1*	**0,82*****	1,01	**0,82*****
	1	**0,95**	**0,99**	1,15	1,31	1,15	1,31	1	**0,97**	1
	3	1	**0,9**	1	1	**0,98**	1	**0,99**	1	**0,98**
PCEd	6	**0,99**	**0,87***	1,01	1,01	**0,98***	1	1,01	1	1,01
	9	**0,98***	**0,88**	1,01	1	1	1	**0,99**	1	**0,99***
	12	1,01	**0,91**	**0,95**	**0,95****	1*	1**	**0,98**	1	**0,97**
	1	**0,91***	**0,9**	1,15	1,31*	1,15	1,31*	1,01	**0,81****	**0,99**
	3	**0,98**	**0,89***	**0,98**	**0,99**	**0,99**	1	**0,97**	1	**0,99**
PPI	6	**0,98**	**0,82**	1,01	1,01	1,01	1	1	1	1
	9	**0,99**	**0,8**	**0,97**	**0,98**	1	1	**0,98**	1	**0,97**
	12	1,06	**0,95**	1,06	1,04	1	1	1,03	1	1,04
	1	1	**0,94**	**0,84**	**0,83**	**0,84**	**0,83**	1,03	**0,95**	1
	3	1,04	**0,85**	**0,92***	**0,92***	**0,94***	**0,94***	1,05	1,06	1,07
PEI	6	1,05	**0,83**	**0,99****	**0,99****	**0,99**	**0,99**	1,05	1,02	1,03
	9	1,03	**0,87***	**0,98**	**0,98**	1	1	1,04	1,01	1,05**
	12	**0,99**	**0,72**	**0,93**	**0,92**	1	1	**0,92**	1	**0,93**
	1	**0,99**	1,06	**0,87**	**0,78***	**0,87**	**0,78***	1,06	**0,43*****	**0,99**
	3	1,15	1,06	1,02	**0,94*****	**0,99**	**0,91****	1,17	**0,67****	1,12
UR	6	1,1	**0,8**	**0,96**	1,02	1	**0,95****	1,08	**0,9****	1,1
	9	1,02	**0,87**	**0,93**	1,01	**0,99**	**0,98***	**0,97**	**0,96****	1,03
	12	**0,96**	**0,82**	**0,97**	**0,97**	**0,99**	1*	**0,93****	**0,99***	**0,96**
	1	**0,94**	1,06	**0,94**	**0,87**	**0,94**	**0,87**	1,01**	**0,98**	**0,94*****
	3	1,03	1,12	1,1	1,08	1,08	**0,95**	1,03	**0,96**	1,03
IP	6	1,04	1,03*	1,05	1,06	1,04	1	1,01	1,02	**0,99**
	9	1	1*	1,03	1,02	1,04	1,05	**0,97**	1,05	1
	12	1	**0,9**	1,03	1,03	1,04	1,06	**0,97**	1,06	**0,97**
	1	**0,98**	**0,74****	**0,99**	1,02	**0,99**	1,02	1,01	**0,99**	1
	3	**0,99**	**0,76****	**0,99**	**0,99**	**0,97**	**0,99**	1	**0,99**	1,01
RDI	6	**0,99**	**0,75**	**0,99**	**0,99**	**0,99**	1	**0,97**	1	**0,99**
	9	**0,98**	**0,83*****	1,01	1,01	1*	1	1,01	1	**0,99**
	12	1	**0,83****	1	1	1	1	1	1	**0,99**
	1	**0,56****	**0,85**	1,48**	1,71*	1,48**	1,71*	1	**0,61**	1
	3	1	**0,68****	1,04	1,04	1,03	1,04	1,04	1,04	1,01
PCE	6	**0,96**	**0,63*****	**0,99**	**0,99**	**0,98**	**0,99**	**0,99**	**0,99**	**0,95**
	9	**0,97**	**0,53*****	1	1	1	1	**0,97**	1	**0,99**
	12	**0,95**	**0,67*****	1	1	1	1	1,01	1	**0,99**

Jurado et al. (2013) dataset. MSFE ratios between model 10 and competing models for CPI, FFR, PCEd, PPI, PEI, UR, IP, RDI, and PCE for forecasting leads h. A value lower than one indicates a lower MSFE of model 10 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. Number of forecasts is $S' = 150$ (second half of the $S = 300$ out-of-sample forecasts). The number of factors in the methods involving factor models is 2. The 1-step ahead forecasts range from August 1998 to January 2011. The 12-step ahead forecasts range from July 1999 to December 2011.

Table 7: MSFE ratios for whole forecast sample (3 factors).

	h	mod 1	mod 2	mod 3	mod 4	mod 5	mod 6	mod 7	mod 8	mod 9
	1	1	1,06	1,14	1,37**	1,14	1,37**	1	1,02	1,01
	3	0,98	0,9	0,99	0,99	0,96	1	0,97	1	0,97
CPI	6	0,99	0,83**	1	0,95	1	1	0,99	1	0,99
	9	0,98	0,79**	1,01	1,01	1	1	0,96	1	0,97
	12	1,02	0,83***	1,01	1,01	1	1	1,01	1	1,02
	1	1,15	0,69**	1,39***	1,47***	1,39***	1,47***	1,32	1,61***	0,98**
	3	0,79**	0,34***	0,72**	0,69**	0,78	0,84	0,83	1,06	0,94
FFR	6	0,71**	0,26***	0,89	0,86	0,93	0,91	0,72	1,02	0,8***
	9	0,85	0,36***	1	1,02	0,96	0,96	0,92	1,02	0,92*
	12	0,76**	0,38***	0,92	0,94	1,01	1,01	0,78	1,03	0,9
	1	0,93	0,98	1,13	1,29*	1,13	1,29*	1	0,94	1
	3	1	0,91	0,99	0,99	0,98	1	0,98	1	0,98
PCEd	6	0,99	0,83***	0,99	0,99	0,99	1	1	1	1*
	9	0,99	0,87	1,01	1	1	1	0,99	1	1
	12	1	0,85***	0,97	0,97**	1*	1	0,99	1	0,98
	1	0,93*	0,9	1,13	1,3**	1,13	1,3**	1,01	0,86**	1
	3	0,98	0,85***	0,98	0,98	0,98	1	0,97	1	0,98
PPI	6	0,99	0,82	1	1	1,01	1	0,99	1	0,99
	9	0,99	0,79	0,97	0,97	1	1	0,98	1	0,97
	12	1,04	0,9*	1,05	1,03	1	1	1,06	1	1,03
	1	0,97	0,9**	0,88*	0,88*	0,88*	0,88*	0,99	0,97	1
	3	0,99	0,9***	0,93**	0,93**	0,95*	0,95*	1	1,02	0,99
PEI	6	0,99	0,87*	0,96**	0,96*	0,97**	0,97**	0,98	1,01	0,99*
	9	0,99	0,9***	0,96	0,96	0,99	0,99	0,99	1*	0,99
	12	0,99	0,84	0,97	0,98	1	1	0,97	0,99	0,97
	1	0,89***	0,93	0,82***	0,78***	0,82***	0,78***	0,94	0,49***	1,01
	3	0,98	0,95	0,94**	0,89**	0,93	0,87**	0,99	0,6***	0,99
UR	6	0,99	0,78	0,91**	0,94	0,94	0,9*	0,99	0,76**	0,98**
	9	0,99	0,86	0,92*	0,96	0,95	0,94	1	0,85	0,91**
	12	0,97***	0,82	0,95*	0,94**	0,99	0,99	0,97	0,94*	0,94**
	1	0,92***	0,99	0,9	0,87	0,9	0,87	0,96	1	0,97**
	3	1,06	0,99	1,09	1,07	1,07	0,98	1,07	1,01	1,02
IP	6	1,03	0,94	1,02	1,03	1,01	0,99	1,03	1,01	0,97
	9	0,98	0,92	1,02	1,01	1,02	1,02	0,99	1,01	0,98
	12	1	0,85**	1	1	1,02	1,03	0,99	1,01	0,96
	1	0,94	0,78***	0,98	1	0,98	1	0,99	0,96	1
	3	0,99	0,82***	0,99	0,99	0,98	0,99*	0,99	0,99	0,98
RDI	6	0,98	0,77**	0,99	0,99	0,99	0,99	0,97	0,99	0,98*
	9	0,98	0,87***	1	1	1	1	0,98	0,99	0,99
	12	1	0,85***	1	1	1	1	0,99	1	0,99
	1	0,64***	1,03	1,4***	1,75***	1,4***	1,75***	1	0,65**	0,99*
	3	1,04	0,83**	1,03	1,03	0,98	1,05	1,02	1,05	1,01
PCE	6	0,99	0,75***	0,99	0,99	0,99	1	0,97	1	0,97
	9	0,98	0,68***	1,01	1	1	1	0,98	1	0,99
	12	0,97	0,75***	1,01	1,01	1	1	0,98	1	0,98

Jurado et al. (2013) dataset. MSFE ratios between model 10 and competing models for CPI, FFR, PCEd, PPI, PEI, UR, IP, RDI, and PCE for forecasting leads h. A value lower than one indicates a lower MSFE of model 10 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. Number of forecasts is $S = 300$. The number of factors in the methods involving factor models is 3. The 1-step ahead forecasts range from February 1986 to January 2011. The 12-step ahead forecasts range from January 1987 to December 2011.

Table 8: MSFE ratios for first half of forecast sample (3 factors).

	h	mod 1	mod 2	mod 3	mod 4	mod 5	mod 6	mod 7	mod 8	mod 9
	1	**0,93**	**0,89**	1,02	1,1	1,02	1,1	**0,95***	**0,92**	**0,99**
	3	**0,99***	**0,74*******	1	**0,98**	**0,98**	1	**0,99****	1	**0,98**
CPI	6	1,01	**0,63****	1*	1*	1	1	1,02	1	**0,99**
	9	**0,99**	**0,66*****	1	1,01	1	1	1	1	**0,99**
	12	1,01	**0,62*****	1,01	1	1	1	1	1	1
	1	1,15	**0,65*****	1,23	1,27	1,23	1,27	1,24**	1,49***	**0,98**
	3	**0,8**	**0,33*****	**0,75***	**0,73***	**0,78**	**0,82**	**0,85**	**0,98**	1
FFR	6	**0,68***	**0,3*****	**0,82**	**0,81**	**0,87**	**0,87**	**0,69***	**0,94**	**0,86**
	9	**0,84**	**0,47****	**0,94**	**0,94**	**0,9***	**0,9***	**0,92**	**0,94**	**0,96**
	12	**0,73***	**0,37*****	**0,88**	**0,88**	**0,94**	**0,94**	**0,75**	**0,96**	**0,88**
	1	**0,9****	**0,98**	1,12**	1,27***	1,12**	1,27***	**0,96***	**0,88*****	**0,97****
	3	1,01**	**0,93**	**0,98***	**0,99***	**0,98**	1,01	**0,97**	1,01**	**0,98**
PCEd	6	1	**0,74*****	**0,96**	**0,96**	1	1	**0,97**	1*	**0,96**
	9	1	**0,84**	1,01	1,01	1	1	1,01	1	1,01
	12	1	**0,75*****	1	1	1	1	1	1	**0,98**
	1	1	**0,86**	1	1,2	1	1,2	1,02	**0,95**	**0,99****
	3	1,01	**0,71*****	**0,94**	**0,94**	**0,96**	1,01	**0,97**	1,01	**0,96**
PPI	6	1,01	**0,79****	**0,96**	**0,95**	1	1	**0,98**	1	**0,97**
	9	1	**0,75*****	**0,94**	**0,95**	1	1	**0,99**	1	**0,99**
	12	**0,97**	**0,7**	**0,99****	**0,99****	1	1	**0,97***	1	**0,93**
	1	**0,99**	**0,9**	**0,94**	**0,95**	**0,94**	**0,95**	1	**0,98**	1
	3	1	**0,98**	**0,98***	**0,98**	**0,99**	**0,99**	1,01	1	1
PEI	6	**0,98**	**0,93*****	**0,98*****	**0,98*****	**0,99****	**0,99****	**0,97***	1,01**	1
	9	**0,98**	**0,94**	**0,96**	**0,95**	1	1	**0,96**	1	**0,98**
	12	**0,99**	**0,96*****	**0,99**	1,02	1	1	1	1	1,02
	1	**0,91**	**0,93**	**0,87**	**0,88***	**0,87**	**0,88***	**0,97**	**0,7****	1,01
	3	**0,97**	**0,99**	1,01	**0,99**	1,03*	**0,98****	**0,98**	**0,83****	**0,98**
UR	6	1	**0,9**	**0,99**	**0,97**	1	**0,97***	1*	**0,92****	**0,99**
	9	1,04	**0,92**	**0,99**	**0,99**	1	**0,99**	1,04	**0,97***	**0,9*****
	12	1**	**0,83**	**0,94*****	**0,92*****	**0,99**	**0,99****	**0,97*****	**0,98****	**0,93*****
	1	**0,87***	**0,86**	**0,81****	**0,84***	**0,81****	**0,84***	**0,89***	**0,99**	**0,97**
	3	1,09	**0,8****	1,06	1,04	1,03	1,02	1,05	1,02	**0,98**
IP	6	1,03	**0,82****	1	1	1	**0,99**	1,01	1,01	**0,97**
	9	**0,97**	**0,8****	1,01	1,01	1	1	**0,97**	1	**0,97**
	12	1,05	**0,78*****	**0,99**	**0,98***	1	1	1,04	1	**0,99**
	1	**0,92**	**0,85*****	**0,98**	**0,98**	**0,98**	**0,98**	**0,98**	**0,93**	1
	3	1	**0,9**	**0,99**	**0,99**	**0,99**	**0,99**	**0,99**	**0,99**	**0,95**
RDI	6	**0,99**	**0,81*****	**0,99**	**0,99**	**0,99****	**0,99****	**0,99**	**0,99****	1
	9	**0,99**	**0,92***	**0,99**	**0,99**	1*	1*	**0,99**	1*	1,02
	12	**0,99**	**0,88*****	1	1	1	1	**0,99**	1	1
	1	**0,68**	1,17	1,35	1,76***	1,35	1,76***	1,01**	**0,67**	1
	3	1,07	**0,97**	1,02	1,03	**0,96**	1,06	1,03	1,06	1,02
PCE	6	1,01	**0,85*****	**0,99**	**0,98**	1*	1	**0,99**	1	**0,98**
	9	**0,99**	**0,83****	1,02	1**	1	1	**0,99**	1	1,01
	12	**0,99**	**0,84*****	1,02	1,03	1**	1*	**0,99**	1*	**0,99**

Jurado et al. (2013) dataset. MSFE ratios between model 10 and competing models for CPI, FFR, PCEd, PPI, PEI, UR, IP, RDI, and PCE for forecasting leads h. A value lower than one indicates a lower MSFE of model 10 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. Number of forecasts is $S' = 150$ (first half of $S = 300$ out-of-sample forecasts). The number of factors in the methods involving factor models is 3. The 1-step ahead forecasts range from February 1986 to July 1998. The 12-step ahead forecasts range from January 1987 to June 1999.

Table 9: MSFE ratios for second half of forecast sample (3 factors).

	h	mod 1	mod 2	mod 3	mod 4	mod 5	mod 6	mod 7	mod 8	mod 9
	1	1,02	1,11	1,17	1,45**	1,17	1,45**	1,01	1,04	1,01
	3	**0,97**	**0,95**	**0,99**	**0,99**	**0,96**	1	**0,96**	1	**0,97**
CPI	6	**0,98**	**0,89***	**0,99**	**0,94**	1	1	**0,99**	1	**0,99**
	9	**0,98**	**0,83**	1,02	1,01	1	1	**0,95**	1	**0,97*****
	12	1,02	**0,91*****	1,01	1,02	1	1	1,01	1	1,02
	1	1,15	**0,75**	1,64**	1,82***	1,64**	1,82***	1,41***	1,77***	**0,98**
	3	**0,78**	**0,36*****	**0,7**	**0,65**	**0,78**	**0,87**	**0,81**	1,17	**0,89**
FFR	6	**0,74**	**0,23****	**0,97**	**0,92**	1,01	**0,97**	**0,76**	1,13	**0,74****
	9	**0,87**	**0,29***	1,08	1,12	1,03	1,04	**0,91**	1,11	**0,89***
	12	**0,8**	**0,39****	**0,97**	1	1,09	1,09	**0,82**	1,11	**0,92**
	1	**0,94**	**0,98**	1,13	1,29	1,13	1,29	1,01	**0,96**	1
	3	1	**0,9**	1	1	**0,98**	1	**0,99**	1	**0,99**
PCEd	6	**0,99**	**0,87****	1,01	1,01	**0,99**	1	1,01	1	1,02*
	9	**0,98***	**0,88**	1,01	1	1	1	**0,98**	1	**0,99**
	12	1,01	**0,91**	**0,95**	**0,95****	1**	1	**0,99**	1	**0,98**
	1	**0,92***	**0,91**	1,16	1,33*	1,16	1,33*	1	**0,84****	1
	3	**0,98**	**0,89***	**0,98**	**0,99**	**0,99**	1	**0,97**	1	**0,99**
PPI	6	**0,99**	**0,83**	1,01	1,02	1,01	1	1	1	**0,99**
	9	**0,99**	**0,8**	**0,97**	**0,98**	1	1	**0,98**	1	**0,97**
	12	1,06	**0,95**	1,06	1,04	1	1	1,08	1	1,05
	1	**0,95**	**0,9**	**0,8**	**0,79**	**0,8**	**0,79**	**0,98**	**0,96**	1,01
	3	**0,99**	**0,81****	**0,87***	**0,87***	**0,89***	**0,89***	**0,99**	1,05	**0,99**
PEI	6	1	**0,79**	**0,94**	**0,94**	**0,94**	**0,94**	**0,99**	1,02	**0,97**
	9	1,01	**0,85****	**0,96**	**0,96**	**0,98**	**0,98**	1,02	1,01*	1,01
	12	**0,99**	**0,73**	**0,93**	**0,92**	1	1	**0,94**	**0,99**	**0,91**
	1	**0,86****	**0,93**	**0,77****	**0,68*****	**0,77****	**0,68*****	**0,9****	**0,36*****	1,01
	3	**0,98**	**0,91**	**0,87*****	**0,81****	**0,85*****	**0,78****	**0,99**	**0,47****	**0,99**
UR	6	**0,98**	**0,71**	**0,85****	**0,91**	**0,89**	**0,85***	**0,99****	**0,66****	**0,97*****
	9	**0,95**	**0,81**	**0,86***	**0,93**	**0,92**	**0,91**	**0,96**	**0,78**	**0,92**
	12	**0,95****	**0,81**	**0,96**	**0,96**	**0,98**	**0,99**	**0,97**	**0,91***	**0,94**
	1	**0,94**	1,06	**0,95**	**0,88**	**0,95**	**0,88**	**0,99**	1	**0,97**
	3	1,04	1,13	1,1	1,09	1,09	**0,95**	1,08	1,01	1,04
IP	6	1,03	1,01	1,03	1,05	1,02	**0,99**	1,04	1,01	**0,97**
	9	**0,99**	**0,98**	1,02	1,01	1,03	1,03	**0,99**	1,02	**0,98**
	12	**0,98**	**0,88**	1,01	1,01	1,02	1,04	**0,97**	1,01	**0,95**
	1	**0,96**	**0,73*****	**0,98**	1,01	**0,98**	1,01	**0,99**	**0,98**	1
	3	**0,99**	**0,76*****	**0,99**	**0,99**	**0,97**	**0,99***	**0,99**	**0,98**	1
RDI	6	**0,98**	**0,74**	**0,98**	**0,98**	**0,98**	**0,99**	**0,95**	**0,98**	**0,97**
	9	**0,97**	**0,83*****	1	1	**0,99**	**0,99**	**0,98**	**0,99**	**0,97**
	12	1	**0,83*****	1	**0,99**	1	1	**0,99**	1	**0,98**
	1	**0,57****	**0,86**	1,5**	1,73*	1,5**	1,73*	**0,99**	**0,62**	**0,99***
	3	1	**0,68****	1,04	1,04	1,03	1,04	1,01	1,04	1
PCE	6	**0,96**	**0,63*****	**0,99**	**0,99**	**0,98**	**0,99**	**0,95**	**0,99**	**0,94**
	9	**0,97**	**0,53*****	1	1	1	1	**0,96**	1	**0,97**
	12	**0,95**	**0,67*****	1	1	1	1	**0,97**	1*	**0,96**

Jurado et al. (2013) dataset. MSFE ratios between model 10 and competing models for CPI, FFR, PCEd, PPI, PEI, UR, IP, RDI, and PCE for forecasting leads h. A value lower than one indicates a lower MSFE of model 10 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. Number of forecasts is $S' = 150$ (second half of $S = 300$ out-of-sample forecasts. The number of factors in the methods involving factor models is 3. The 1-step ahead forecasts range from February 1986 to July 1998. The 12-step ahead forecasts range from January 1987 to June 1999. The 1-step ahead forecasts range from August 1998 to January 2011. The 12-step ahead forecasts range from July 1999 to December 2011.

Table 10: Variance ratios $r_t^{j,n}$.

		CPI	FFR	PCEd	PPI	PEI	UR	IP	RDI	PCE
1 factor	1^{st} factor	1,346E-07	0,0063	4,26E-08	3,75E-07	1,38E-07	1,71E-02	2,61E-05	4,16E-07	1,11E-06
2 factors	1^{st} factor	8,23E-06	0,0068	2,28E-06	1,30E-05	5,34E-07	2,89E-03	7,09E-06	3,24E-07	1,04E-07
	2^{nd} factor	6,93E-05	0,02	2,33E-05	0,00015	8,07E-06	2,49E-03	7,39E-06	6,83E-06	1,08E-05
3 factors	1^{st} factor	5,63E-06	0,04	1,28E-06	8,42E-06	3,93E-07	3,30E-03	8,09E-05	4,72E-08	5,55E-07
	2^{nd} factor	0,00017	0,0097	5,37E-05	0,00017	5,82E-06	2,56E-03	3,27E-03	5,53E-06	4,17E-06
	3^{rd} factor	8,24E-06	0,09	3,56E-06	4,14E-06	2,17E-07	2,68E-03	8,80E-04	1,10E-06	1,10E-06

Values for average variance ratio $r_t^{j,n}$ (in percentage), eqn. (19), for CPI, FFR, PCEd, PPI, PEI, UR, IP, RDI, and PCE for the supervised model (1)-(2), with different numbers of estimated factors ranging from one to three. The estimates of the parameters are computed using the whole sample.

Table 11: MSFE ratios, cosine basis loadings.

$\rho_{f_2,y} = 0.5$				$\rho_{f_1,y} = 0.5$			
$\rho_{f_1,y}$	h	mod 8	mod 9	$\rho_{f_2,y}$	h	mod 8	mod 9
	1	1,01	1,02		1	0,98	0,94
	3	0,99	1		3	0,97	1,01
0.1	6	1	1,01	0.1	6	1	1,01
	9	1,01	1		9	0,96**	1
	12	0,98	1		12	0,97	0,99
	1	0,95**	0,73***		1	0,98	0,64***
	3	1	0,99		3	0,99	0,95
0.2	6	0,98	1,01*	0.2	6	0,96***	1
	9	1	1		9	0,97	1
	12	0,99	1		12	0,98	1
	1	0,98	0,91*		1	0,96	1
	3	0,98**	1		3	0,98	0,98
0.3	6	0,98	1	0.3	6	0,98	0,99
	9	1	1		9	0,97***	1
	12	0,97**	1		12	0,98**	1
	1	0,98	0,6***		1	0,92**	0,85***
	3	0,98	0,96		3	0,99	0,98
0.4	6	0,97	1,01	0.4	6	0,98	0,99
	9	0,96***	1		9	0,98	1
	12	0,98	1		12	0,98	1
	1	1,01	0,99		1	0,94**	0,34***
	3	0,99*	1		3	0,93**	0,74**
0.5	6	0,98	1	0.5	6	0,94	0,97
	9	0,98**	1		9	0,98	1,03
	12	0,99	1		12	1	1,03
	1	0,97	0,75***		1	0,92**	0,87***
	3	0,98	0,99		3	1	1,01
0.6	6	1	1,02	0.6	6	0,99	1
	9	1,01	1,02		9	0,98	1
	12	0,97**	1,02		12	0,99	1
	1	0,92**	0,86***		1	0,95**	0,99
	3	0,97	1,01		3	0,97	1,01
0.7	6	0,98	1	0.7	6	0,99	1
	9	0,99	1		9	0,99	1,01
	12	0,99	1		12	0,98**	1,01
	1	0,96	1,03		1	0,98	0,86***
	3	1,01	1,03*		3	0,96	0,99
0.8	6	0,99	1	0.8	6	0,99	1
	9	1,01	1		9	0,95	1
	12	0,99**	1		12	0,98	1
	1	0,96	0,81***		1	1,03	1,03
	3	0,98	0,97		3	1,02	1,01
0.9	6	0,98	1,03	0.9	6	1**	1,01
	9	0,99	1,02		9	1	1,01*
	12	1	1		12	0,99	1*

Simulated data with cosine basis loadings. MSFE ratios between model 10 and models 8 and 9 for different levels of correlation between factors and forecast objective. Forecasting leads $h = 1, 3, 6, 9, 12$. A value lower than one indicates a lower MSFE of model 10 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. Number of forecasts is $S = 300$. The number of latent factors is 2. The correlation between the two latent factors is $\rho_{f_1,f_2} = 0.10$.

Table 12: MSFE ratios, random loadings.

	$\rho_{f_2,y} = 0.5$				$\rho_{f_1,y} = 0.5$		
$\rho_{f_1,y}$	h	mod 8	mod 9	$\rho_{f_2,y}$	h	mod 8	mod 9
	1	1	**0,44*****		1	1	**0,96**
	3	**0,98***	**0,94**		3	**0,97**	**0,98**
0.1	6	**0,98**	1,01*	**0.1**	6	**0,98**	1
	9	**0,97**	1,01		9	**0,96**	1
	12	**0,98**	1		12	**0,96**	1
	1	1,04***	**0,9****		1	**0,99**	0,91***
	3	1,08**	1,1***		3	**0,97****	**0,97**
0.2	6	1,08*	1,08**	0.2	6	**0,96***	**0,98**
	9	1,04	1,07*		9	**0,95**	**0,98**
	12	1,03	1,05		12	**0,96**	**0,98**
	1	**0,93**	**0,9**		1	1	**0,85*****
	3	**0,97**	1		3	**0,99**	1
0.3	6	**0,98**	1	0.3	6	1	1
	9	**0,98**	1		9	**0,98**	1
	12	**0,97**	1		12	**0,98**	1
	1	1	**0,28*****		1	**0,99**	**0,89***
	3	**0,99**	**0,66*****		3	**0,97**	**0,98**
0.4	6	**0,97***	**0,84***	0.4	6	**0,97**	**0,99**
	9	**0,97****	**0,94**		9	**0,96***	**0,99**
	12	**0,96**	**0,94**		12	**0,96***	**0,99**
	1	**0,98**	**0,34*****		1	**0,97**	**0,5*****
	3	**0,98***	**0,79*****		3	1	**0,6*****
0.5	6	**0,97**	**0,99**	0.5	6	1	**0,78*****
	9	1	1,02		9	**0,99**	**0,88****
	12	1,01	1		12	**0,98**	**0,95**
	1	**0,97**	**0,87*****		1	1,01	**0,32*****
	3	1,01	1,01		3	**0,99**	**0,72*****
0.6	6	1,01	1,02	0.6	6	**0,99**	**0,9****
	9	1,01	1,01		9	**0,98**	**0,96**
	12	**0,99**	**0,99**		12	**0,98**	1
	1	1	**0,65*****		1	1,01	**0,77*****
	3	**0,99***	**0,97***		3	**0,98**	**0,96**
0.7	6	**0,99**	1	0.7	6	**0,99**	**0,98**
	9	**0,99**	1*		9	**0,98**	**0,97**
	12	**0,98**	1**		12	**0,97**	**0,99**
	1	**0,98**	**0,25*****		1	1,02	**0,65*****
	3	1,02*	**0,77*****		3	1,04	**0,86****
0.8	6	**0,98**	**0,96**	0.8	6	1,05	**0,94**
	9	**0,97**	**0,99**		9	1,03	0,99
	12	**0,97**	**0,99**		12	1,03	1,04
	1	**0,99**	**0,95**		1	1	**0,29*****
	3	**0,99**	1		3	1,01**	**0,56*****
0.9	6	1,01	1	0.9	6	**0,99**	**0,83**
	9	**0,98**	1		9	**0,97**	**0,91**
	12	**0,99****	1		12	**0,99**	**0,98**

Simulated data with random loadings. MSFE ratios between model 10 and models 8 and 9 for different levels of correlation between factors and forecast objective. Forecasting leads $h = 1, 3, 6, 9, 12$. A value lower than one indicates a lower MSFE of model 10 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. Number of forecasts is $S = 300$. The number of latent factors is 2. The correlation between the two latent factors is $\rho_{f_1,f_2} = 0.10$.

Research Papers 2013

2015-21: Ulrich Hounyo and Bezirgen Veliyev: Validity of Edgeworth expansions for realized volatility estimators

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