07/RT/12

Páipéar Taighde Teicniúil
Research Technical Paper

Real-time forecasting in a data-rich environment

Joëlle Liebermann

Banc Ceannais na hÉireann
Central Bank of Ireland

Real-time forecasting in a data-rich environment

Joëlle Liebermann*

Abstract

This paper assesses the ability of different models to forecast key real and nominal U.S. monthly macroeconomic variables in a data-rich environment from the perspective of a real-time forecaster, i.e. taking into account the real-time data revisions process and data flow. We find that for the real variables predictability is confined over the recent recession/crisis period. This is in line with the findings of D'Agostino and Giannone (2012) that gains in relative performance of models using large datasets over univariate models are driven by downturn periods which are characterized by higher comovements. Regarding inflation, results are stable across time, but predictability is mainly found at the very short-term horizons. Inflation is known to be hard to forecast, but by exploiting timely information one obtains gains at nowcasting and forecasting one-month ahead, especially with Bayesian VARs. Furthermore, for both real and nominal variables, the direct pooling of information using a high dimensional model (dynamic factor model or Bayesian VAR) which takes into account the cross-correlation between the variables and efficiently deals with the "ragged edge" structure of the dataset, yields more accurate forecasts than the indirect pooling of bi-variate forecasts/models.

JEL classification: C11, C33, C53, E52.
Keywords: Real-time data, Nowcasting, Forecasting, Factor model, Bayesian VAR, Forecast pooling.

*E-mail: jlieberm@ulb.ac.be. The views expressed in this paper are those of the author and do not necessarily reflect those of the Central Bank of Ireland. This paper is a chapter of my Phd thesis. I would like to thank Domenico Giannone for his valuable comments and suggestions.

Non Technical Summary

This paper evaluates the ability of different models to forecast key real and nominal U.S. monthly macroeconomic variables in a data-rich environment from the perspective of a real-time forecaster, i.e. taking into account the real-time data revisions process and data flow. This is an issue which, to the best of our knowledge, has not yet been examined.

In the forecasting literature there is widespread empirical evidence on instabilities, whether attributed to changes in individual predictive content, in parameters and/or model, or over time. Many empirical studies evaluate different methods at forecasting key variables, using different datasets, periods and models, making the results and ranking of methods, difficulty to compare. Furthermore, and importantly for a real-time forecaster and policymaker, no studies compare the different methods in a truly real-time setting.

This study contributes to the literature in a number of ways. Firstly, we run a forecasting horse race between the different methods used to forecast using many predictors. Since using all predictors at once in traditional time series models leads to the so-called curse of dimensionality problem, specific methods have been developed that can overcome this problem. Broadly speaking, two approaches have been followed to forecast in a data-rich environment, namely pooling of bi-variate forecasts which is an indirect way to exploit large cross-section and direct pooling of information using high-dimensional models such as dynamic factor model (DFM) or Bayesian vector autoregression (BVAR) Secondly we run the forecasting horse race in a fully real-time setting. That is, we replicate the situation faced by a real-time forecaster as we take into account the preliminary nature of the data as well as the real-time data flow. Thirdly we use forecast combination schemes to pool within and across models as we seek for evidence regarding the performance of a model that is robust across specifications/combination schemes. This should also help to insure against model/specification instability. Finally, following the findings of D'Agostino and Giannone (2012) that gains from data-rich methods over univariate models mainly confine to downturn periods, we also investigate the sensitivity of the results over the (end of the) great moderation and recent recession.

We find that for the real variables, predictability is confined over the recent recession/crisis period. This in line with the findings of D'Agostino and Giannone (2012) over an earlier period, that gains in relative performance of models using large datasets over univariate models are driven by downturn periods which are characterized by higher comovements. These results are robust to the combination schemes or models used. Regarding inflation, results are stable across time, but predictability is mainly found at nowcasting and forecasting one-month ahead, with the BVAR standing out at nowcasting. The results show that the forecasting gains at these short horizons stem mainly from exploiting timely information.

Furthermore, for both real and nominal variables, the direct pooling of information using a high dimensional model (DFM or BVAR) which takes into account the cross-correlation between the variables and efficiently deals with the "ragged edge" structure of the dataset, yields more accurate forecasts than the indirect pooling of bi-variate forecasts/models.

1 Introduction

This paper evaluates the ability of different models to forecast key real and nominal U.S. monthly macroeconomic variables in a data-rich environment from the perspective of a real-time forecaster, i.e. taking into account the real-time data revisions process and data flow. This is an issue which, to the best of our knowledge, has not yet been examined.

In the forecasting literature there is widespread empirical evidence on instabilities, whether attributed to changes in individual predictive content, in parameters and/or model, or over time. Many empirical studies evaluate different methods at forecasting key variables, using different datasets, periods and models, making the results and ranking of methods, difficulty to compare. Furthermore, and importantly for a real-time forecaster and policymaker, no studies compare the different methods in a truly real-time setting.

This study contributes to the literature in a number of ways. Firstly, we run a forecasting horse race between the different methods used to forecast using many predictors. Since using all predictors at once in traditional time series models leads to the so-called curse of dimensionality problem[1], specific methods have been developed that can overcome this problem. Broadly speaking, two approaches have been followed to forecast in a data-rich environment, namely pooling of bi-variate forecasts which is an indirect way to exploit large cross-section and direct pooling of information using a high-dimensional model.

The first approach, pooling over small models' forecasts, initially emerged as a response to the instabilities found in individual predictive content. Stock and Watson (2004), using bi-variate models to forecast output growth and inflation from 1959 to 1999, find considerable instability over time and across countries of asset prices and leading indicators predictive content. More recently, Rossi and Sekhposyan (2010) re-assess the findings of Stock and Watson over the sample 1970 to 2005; they broadly confirmed the instability results of Stock and Watson and similarly to D'Agostino, Giannone and Surico (2006) found that most predictors lose their ability to forecast around the start of the great moderation or before.[2] As a more robust tool to produce forecasts in light of model uncertainty/instability, a number of studies have suggested using model averaging. Stock and Watson (2004), among others, show that using the (unweighted) average over individual predictor forecasts yields better and more stable results. Timmermann's (2006) survey put forth theoretical rationales in favor of forecast combinations such as model misspecification, structural breaks and, more generally, unknown instabilities. The combined forecast can be viewed as "integrating

[1]That is, when the size of the information set (n) is too large relative to the sample size (T), then the loss of degrees of freedom results in poor or unfeasible (if $n > T$) ordinary least squares (OLS) forecasts (see De Mol, Giannone and Reichlin, 2008).

[2]To be precise, for forecasting output, they find that financials were useful up to the mid-1970s, and that predictive content of the indicators for inflation break down around the start of the great moderation.

out"model (predictor) uncertainties, as highlighted in Bayesian model averaging. Aiolfi and Timmermann (2006), for instance, find that averaging over the top performing models (based on past forecasting performance) provides better forecasts than just relying on the first best model. This pooling over small models' forecasts helps mitigate the issue of unstable predictors and exploits the information content from many predictors, yielding in general more accurate forecasts.

The second approach, which in fact directly focusses on the particular issue of forecasting with many predictors, pools information using a high-dimensional model that can overcome the curse of dimensionality problem. Dynamic factor models (DFM) (Forni, Hallin, Lippi and Reichlin 2000 and 2005; Stock and Watson 2002a,b) have been the predominant tool used and have been found to perform well.[3] In such a framework, it is assumed that the first few factors, which capture the bulk of the comovement among the predictors, summarize all the relevant information in the dataset. Hence, the number of parameters to estimate in the forecasting equation is substantially reduced by replacing the large set of predictors by these first few factors. More recently, Bańbura, Giannone and Reichlin (2010) have considered Bayesian vector autoregression (BVAR) as an alternative for forecasting with many predictors. These authors found that by applying Bayesian shrinkage (i.e. shrinking the parameters via the imposition of priors) to deal with the curse of dimensionality problem, it is not only possible to forecast using large VAR but also that these forecasts compare favorably relative to those of the DFM.

Alternative methods to forecast using many predictors such as variables selection algorithms (e.g. Lasso and Bayesian model averaging) have not been found to yield any forecasting improvement over these two former methods and the variables selected have no clear economic interpretation as collinearity renders their weights unstable (see Stock and Watson, 2011; De Mol, Giannone and Reichlin, 2008). Hence among the methods enabling to directly forecast in a data-rich environment, we focus on the DFM and BVAR. We also retain the simple approach of pooling bi-variate forecasts as it has been reported to perform well by some authors, and it will serve as an additional benchmark against which to evaluate the former models.

Our second contribution is to run the forecasting horse race in a fully real-time setting. That is, we replicate the situation faced by a real-time forecaster as we take into account the preliminary nature of the data as well as the real-time data flow. Most of the studies forecasting key US monthly variables use revised datasets, which differ from the preliminary data available to a real-time forecaster and policy makers. Among those using use real-time data, such as Rossi and Sekhposyan (2010), Heij, van Dijk and Groenen (2011) and

[3]See Stock and Watson, 2002a,b; Bernanke and Boivin, 2003; Forni, Hallin, Lippi and Reichlin, 2005; Boivin and Ng, 2005; D'Agostino and Giannone, 2012.

Banternghansa and McCracken (2011) among others, balanced datasets are used. But since variables are released in a non-synchronous manner and with varying publication lags, taking into account this real-time data flow implies that one must use econometric approaches that allow to deal with an unbalanced panel at the end of the sample (i.e. at the forecast origin), which is commonly referred to as a "ragged edge" structure, and that the marginal predictive ability of different variables also depends on their timeliness as found in the recent literature devoted to nowcasting GDP. Furthermore the two state-of-the art techniques, BVAR and DFM have only been compared to each other using balanced revised panels (e.g. Bańbura, Giannone and Reichlin, 2010) or in a real-time setting for forecasting Euro-area inflation but benchmarked against simple univariate models (Lenza and Warmedinger, 2011; Giannone, Lenza, Momferatou and Onorante, 2010).

Our third contribution is to use forecast combination schemes to pool within and across models as we seek for evidence regarding the performance of a model that is robust across specifications/combination schemes. This should also help to insure against model/specification instability. For instance, Banternghansa and McCracken (2011) consider the real-time forecasting ability of different combining schemes to outperform information criterion based model selection for bi-variate VARs in the light of instabilities in these models. Whereas, the Bank of England[4] and the Norges Bank[5] further combine over a suite of models to generate forecasts of key macroeconomic variables.

Finally, following the findings of D'Agostino and Giannone (2012) that gains from data-rich methods over univariate models mainly confine to downturn periods, we also investigate the sensitivity of the results over the (end of the) great moderation and recent recession.

We find that for the real variables, predictability is confined over the recent recession/crisis period. This in line with the findings of D'Agostino and Giannone (2012) over an earlier period, that gains in relative performance of models using large datasets over univariate models are driven by downturn periods which are characterized by higher comovements. These results are robust to the combination schemes or models used. Regarding inflation, results are stable across time, but predictability is mainly found at nowcasting and forecasting one-month ahead, with the BVAR standing out at nowcasting. The results show that the forecasting gains at these short horizons stem mainly from exploiting timely information. Furthermore, for both real and nominal variables, the direct pooling of information using a high dimensional model (DFM or BVAR) which takes into account the cross-correlation between the variables and efficiently deals with the "ragged edge" structure of the dataset, yields more accurate forecasts than the indirect pooling of bi-variate forecasts/models.

[4]Kapetanios, Labhard and Price (2008).

[5]Aastveit, Gerdruo and Jore (2011) and Gerdrup, Jore, Smith and Thorsrud (2009).

The paper is organized as follows. In section 2 we describe the design of the real-time forecasting exercise, the dataset, models and combination schemes used to construct the forecasts. Section 3 presents the empirical results over the full sample period as well as over the pre-crisis and crisis sub-samples. This section also displays the results regarding the marginal predictive ability of the timely soft data. Section 4 concludes.

2 Real-time forecasting setting and dataset

The objective is to predict four key U.S. monthly macroeconomic variables, namely industrial production (IP) and the unemployment rate (UR) for the real-side of the economy and headline consumer price index (CPI) and personal consumption expenditures price index (PCE-P) for the nominal side.

The forecasting exercise is performed in a data-rich environment and fully real-time setting. To this end we use vintages[6] for a panel of monthly U.S. macroeconomic indicators from December 2001 to December 2011, reproducing the exact information available to a real-time forecaster. The panel consists of hard data such as industrial production, employment, retail sales, housing and prices among others, and soft data which includes surveys and financials (e.g. term and credit spreads and stock market index). These variables are released in a non-synchronous manner and with varying publication lags. As a consequence the panel is unbalanced at the end of the sample, i.e. it has a "ragged edge" structure.

To fix ideas let V_t denote the vintage for the panel available at the end of a given month t which is the forecast origin:

$$V_t = \{Z_{i,t^\star|t}, \quad i = 1...n; \quad t^\star = 1...T^\star_{i|t}\}$$

where $Z_{i,t^\star|t}$ is the month $t^\star$ value of a generic variable i available in month t and $T^\star_{i|t} \leqslant t$ because of publication lags. The bulk of the hard data are released in the month following the one they cover and a few are released with a two month delay, i.e. $T^\star_{i|t} = t-2$ or $t-1$. The financial data and commodity prices[7] and most of the surveys are very timely as they are already available at the end of the month they refer to. Hence, $T^\star_{i|t} = t$ for all the financials and commodity prices and $T^\star_{i|t} = t$ or $t-1$ for the surveys. A detailed description of the variables along with their publication lags and transformation applied to each series are reported in Appendix A.

[6]For most of the series real-time information was collected from the Federal Reserve Bank of ST.Louis ALFRED database (see Appendix A).

[7]These data are in fact observed at the daily frequency and are converted to the monthly frequency by aggregating daily price changes over the month.

Similar to a real-time forecaster constructing her forecasts at the end of each month, we generate forecasts for the series of interest for the $h = 0, 1, 3, 6$ and 12 months horizon, conditional on the information available at that point in time, using a range of models. The nowcasts, i.e. $h = 0$, are also produced since all of the predicted series are released after the close of the month they refer to, hence are not available for the month considered as the forecast origin.

3 Forecasting models and combination schemes

3.1 Forecasting models

To forecast the series of interest we consider four classes of models. Firstly, we use a random walk (RW) model. This ***naive model*** is the standard benchmark against which one evaluates more sophisticated models. Secondly, we consider autoregressive ***univariate models*** (AR), which only use past information on the targeted series to construct the forecasts. Within this class, h-steps ahead forecasts can be generated by either iterating forward a one-step ahead model (iterated or indirect approach) or by estimating a multistep model (direct approach). Since both approaches are used in the literature and that the issue of which one works better is an empirical one (see Marcellino, Stock and Watson, 2006), we construct forecasts using both approches under the AR model.

The third and fourth classes of models make use of numerous additional predictors to generate the forecasts, i.e. they aim at forecasting in a ***data-rich environment***. But, using all predictors at once in traditional time series models leads to the so-called curse of dimensionality problem. That is, when the size of the information set (n) is too large relative to the sample size (T), then the loss of degrees of freedom results in poor or unfeasible (if $n > T$) ordinary least squares (OLS) forecasts (see De Mol, Giannone and Reichlin, 2008).

One way to circumvent this problem is to pool ***bi-variate model*** forecasts. Under this approach a model is estimated using each candidate predictor one at a time in addition to the targeted series to construct the forecasts. Then, to make use of all the information, the individual predictor forecasts are combined to provide a single forecast. Similar to the AR model, these bi-variate predictions can be constructed using a direct or iterated (indirect) forecasting model. The former method entails using a standard regression model, while the latter entails the use of a VAR model; both are considered within the bi-variate (BIV) class of models.

The alternative approach is to pool information directly using a ***high-dimensional multivariate model*** that overcomes the curse of dimensionality problem. The standard model

used for that purpose which has been found to forecast well is the DFM. Recently, Bańbura, Giannone and Reichlin (2010) have considered BVAR as an alternative for forecasting with large panels. VARs are flexible models as they can accommodate rich cross-correlation and autocorrelation among variables, but as such, are heavily parametrized and run into the curse of dimension problem quickly with an increasing number of series. These authors found that by applying Bayesian shrinkage (i.e. shrinking the parameters via the imposition of priors) it is not only possible to forecast using large VARs but also that these forecasts compare favorably to those obtained by the factor model.

An important issue, governing the choice of the models, is that since the forecasts are generated in a real-time setting, they all need to be able to deal with the"ragged-edge" structure of the dataset. For the univariate models (RW and AR-direct) this is straightforward as a one month publication lag translates directly into a one month increase in the forecast horizon. The standard regression model can also directly be modified to account for publication lags. For all the other models, since they admit a state space representation, the Kalman filter and smoother algorithm with a time-varying dimension observation equation[8] can be used to compute recursively the forecasts conditional on the available information, i.e the unbalanced panel.

With the exception of the BVAR, in all the models series are transformed to obtain stationarity. For most of the series we take the first (i.e. month-over-month) difference of the level or log level; interest rates spreads and some of the surveys are un-transformed.[9] Thus all models, except the BVAR, produce forecasts for the month-over-month growth rate or change of the key series, whereas the BVAR for the level. These forecasts are then used to construct the target being predicted which is the cumulative growth or change over the forecast horizon for the real variables and the h-month ahead level of yearly inflation[10] for the nominal variables. Note that due to publication lags, the level of the key series are unknown for the forecast origin month t and the last available value for these series relates to $t-1$[11] (i.e. $Z_{i,T^{\star}|t} = Z_{i,t-1|t}$), hence for a h-step ahead forecast one needs in fact to forecast h+1-months.

To further set notations, let $Z_t = (Z_{1,t}...Z_{n,t})'$ denote the $n \times 1$ vector of un-transformed (in level) variables, and $X_t = (X_{1,t}...X_{n,t})'$ its transformed to stationary counterpart. We aim at forecasting some elements of Z_t (X_t) which will be subscripted by j and a candidate predictor will be be subscripted by $i^{\star}$, whereas the subscript i will be used to refer to a

[8]In practice this is implemented by using a selection matrix applied to the measurement equation (see Koopman and Durbin (2001), §4.8).

[9]See Appendix A for a description of the transformation applied to each variable.

[10]This is the standard target forecasted in the literature for price indices.

[11]Note that for PCE-P inflation, in approximatively 30% of the cases it is released with two months publication lags, hence we also need to backcast its value for $t-1$.

generic variable in the vector Z_t (X_t).

Then the target being forecasted at origin t for horizon h is defined as follows:

- $100 * (logZ_{j,t+h} - logZ_{j,t-1})$ for IP;
- $Z_{j,t+h} - Z_{j,t-1}$ for the UR;
- $100 * (logZ_{j,t+h} - logZ_{j,t+h-12})$ for CPI and PCE-P.

Naive model (RW)

At each forecast origin, the forecasts are simply set to a constant which is the recursively computed historical mean of the series of interest. For all horizons the model is defined as follows:

$$X_{j,t+h} = \alpha_j + \varepsilon^h_{j,t+h} \tag{1}$$

Given the transformations used to obtain stationarity, this naive model corresponds to a random walk (RW) with drift for the level or log level of the series.

Univariate models (AR)

The direct AR model entails projecting a dated $t + h$ variable on dated t available information. Hence for each horizon h one needs to estimate a different model as defined by equation (2):

$$X_{j,t+h} = \alpha^h_j + \sum_{l=1}^{p} \gamma^h_{j,l} X_{j,t-l} + \varepsilon^h_{j,t+h} \tag{2}$$

Note that the index in the summation in (2) starts at one and not at zero as due to publication lags, at the end of month t the available information regarding the past of the series of interest pertains only to month $t - 1$ and earlier. In the indirect approach, a single one-step ahead model is used:

$$X_{j,t} = \alpha_j + \sum_{l=1}^{p} \gamma_{j,l} X_{j,t-l} + \varepsilon_{j,t} \tag{3}$$

Then, given the estimated parameters, h-step ahead forecasts are constructed recursively by iterating on equation (3).

Bi-variate models (BIV)

For a given targeted series j and forecast horizon h, forecasts using each candidate predictor $i^\star$ one at a time will be constructed as defined by equations (4) and (5) below.

Under the direct approach, forecasts are based on a h-step ahead regression model defined

as:

$$X_{j,t+h} = \alpha^h_{ji^\star} + \sum_{l=1}^{p} \gamma^h_{ji^\star,l} X_{j,t-l} + \sum_{q=q_i^\star}^{p+q_i^\star} \beta^h_{ji^\star,q} X_{i^\star,t-q} + \varepsilon^h_{j,t+h} \tag{4}$$

This equation is the direct AR model augmented by a candidate predictor taking into account the real-time availability of the latter as q_i takes values $0, 1$ or 2 according to its publication lag.

In the case of iterated forecasts, let $X_{ji^\star,t} = (X_{j,t}\ X_{i^\star,t})'$ denote the 2×1 vector of the targeted series and a candidate predictor, then the bi-VAR is defined as:

$$X_{ji^\star,t} = A_{ji^\star,0} + \sum_{l=1}^{p} A_{ji^\star,l} X_{ji^\star,t-l} + \varepsilon_{j,t} \tag{5}$$

The model parameters are estimated on a balanced panel, i.e. truncating the panel at the last month for which both series are available. Next the model is put in the state space form and the Kalman filter and smoother are used to compute recursively the forecasts conditional on the unbalanced panel.

The parameters of equations (2) to (5) are estimated by OLS and the maximum lag length p and q are set to 6.[12]

Multivariate models (MULT)

- ***Dynamic factor model (DFM)***

The DFM framework used relies on the Giannone, Reichlin and Small (2008) framework that can deal with large and unbalanced datasets. In such a model, the $n \times 1$ vector of stationary standardized monthly variables, x_t [13], is represented as the sum of two orthogonal unobserved components: a common component which is driven by a small number of factors that account for most of the comovement among the variables and an idiosyncratic component which is driven by variable-specific shocks. This model is given by:

$$x_t = \Lambda f_t + \xi_t \tag{6}$$

where f_t is a $r \times 1$ vector of common factors, Λ is the $n \times r$ matrix of the factor loadings, and ξ_t is a $n \times 1$ vector of idiosyncratic components.

[12] Note that for equation (4), for a given value of $q_i^\star$ forecasts are constructed using $l = 0, 1, ..., 6$. Hence forecasts are also constructed without using lags of the targeted series.

[13] We use lowercase letters to denote the standardized version of the $n \times 1$ vector X_t

The factors are modeled as a stationary VAR(p):

$$f_t = A_1 f_{t-1} + ... + A_p f_{t-p} + u_t; \quad u_t \backsim i.i.d.N(0, Q) \tag{7}$$

where $A_1, ..., A_p$ are $r \times r$ matrices of autoregressive coefficients. The idiosyncratic components, ξ_t, are orthogonal to the common shocks, u_t, and are modeled as independent stationary AR(1) processes:

$$\xi_t = \psi \xi_{t-1} + e_t; \quad e_t \backsim i.i.d.N(0, R) \tag{8}$$

where $\psi = diag(\psi_1, ..., \psi_n)$ and $R = diag(\sigma_1^2, ..., \sigma_n^2)$. Modelling the serial correlation in the idiosyncratic component[14]

Substituting (8) into (6), one obtains a DFM with non-autocorrelated errors in the transformed variables $\widetilde{x_t}$:

$$\widetilde{x_t} = x_t - \psi x_{t-1} = \Lambda f_t - \psi \Lambda f_{t-1} + e_t \quad \forall t \geq 2 \tag{9}$$

The DFM (6)-(8) can be written in the state space form as:

$$\widetilde{x_t} = \widetilde{\Lambda} z_t + e_t \tag{10}$$

$$z_t = \widetilde{A} z_{t-1} + \widetilde{u_t} \tag{11}$$

where if $p \geq 2$, $\widetilde{\Lambda} = [\Lambda \quad -\psi\Lambda \quad 0_{n \times r(p-2)}]$, $\widetilde{A} = \begin{bmatrix} A_1 \dots A_{p-1} & A_p \\ I_{r(p-1)} & 0_{r(p-1)\times r} \end{bmatrix}$, $\widetilde{u_t} = \begin{bmatrix} u_t \\ 0_{r(p-1)\times 1} \end{bmatrix}$ and $z_t = \begin{bmatrix} f_t \\ f_{t-1} \\ \dots \\ f_{t-p+1} \end{bmatrix}$ is the state vector. In the case $p = 1$, $\widetilde{\Lambda} = [\Lambda \quad -\psi\Lambda]$, $\widetilde{A} = \begin{bmatrix} A_1 & 0_{r\times r} \\ I_r & 0_{r\times r} \end{bmatrix}$, $\widetilde{u_t} = \begin{bmatrix} u_t \\ 0_{r\times 1} \end{bmatrix}$ and $z_t = \begin{bmatrix} f_t \\ f_{t-1} \end{bmatrix}$.

The observation equation (10) holds $\forall t \geq 2$ and for the initial observation, since ξ_1 is

[14]To obtain optimal estimates of the unobserved state vector using the Kalman filter and smoother, the observation equation errors must be white noises. For the DFM (6)-(8), one can either first transform the model into one with non-autocorrelated errors (equation (9)) then put the transformed model into the state space form (10)-(11) or one can augment the state vector with ξ_t. Reis and Watson (2010) have followed the first formulation while Bańbura and Modugno (2010) have adopted the second. The latter authors have adapted the EM algorithm to deal with arbitrary pattern of missing observations and in such a case the first formulation is not valid. Since we estimate the parameters on a balanced panel and that the second formulation is computationally less efficient, as the dimension of the state vector increases a lot, we use the first formulation.

$N(0, (I - \psi^2)^{-1} R)$ the observation equation is:

$$x_1 = \Lambda f_1 + \xi_1 \quad where \quad \xi_1 = e_1 \tag{12}$$

The model parameters $\phi = \{\Lambda, A_1, ..., A_p, \psi, R, Q\}$ and the factors are estimated by maximum likelihood using the Expectation-Maximisation (EM) algorithm as shown in Doz, Giannone and Reichlin (2011b).[15] These authors have proved that maximum likelihood estimation of large DFM is not only consistent, as $n, T \to \infty$ along any path, but also computationally feasible as the likelihood can be maximized via the the EM[16] algorithm. Note that the DFM (6)-(8) is an exact factor model as it is assumed that there is no cross-correlation in the idiosyncratic components which might not hold in large cross-sections. However, Doz, Giannone and Reichlin (2011b) show that this estimator based on a possibly mis-specified, i.e. exact, factor model is consistent for an approximate factor model. In such a case, the estimation method is quasi maximum likelihood.

The parameters are estimated for all combinations over the range $r = 1, ..., 10$ and $p = 1, 2$ using a balanced panel as is the standard practice in the real-time forecasting literature since the unbalanced part of the panel is only at the end of the sample and concerns at most two months. Then, given the parameters estimates, the forecasts are generated by running the Kalman filter and smoother on the unbalanced part of the panel. Note that Bańbura and Modugno (2010) adapt the EM algorithm to deal with arbitrary pattern of missing observations[17] and their procedure has been further accelerated by Jungbacker, Koopman and van der Wel (2011).

- ***Bayesian VAR (BVAR)***

Bańbura, Giannone and Reichlin (2010) found that by applying Bayesian shrinkage, large VAR works well at forecasting. Bayesian estimation combines sample information with priors to yield a posterior estimate, as such shrinkage is incorporated through the priors which take the form of imposing restrictions on parameters. The standard BVAR in the forecasting literature, and used by Bańbura, Giannone and Reichlin (2010), is based on the Minnesota prior of Litterman (1986) and Doan, Litterman and Sims (1984) with the modifications proposed by Kadiyala and Karlsson (1997) and Sims and Zha (1998).

[15] An alternative consistent estimator is the two-step estimator of Doz, Giannone and Reichlin (2011a).

[16] The EM algorithm of Dempster, Laird and Rubin (1977) is a well-known approach to maximize the Gaussian likelihood function in the presence of missing data, which here are the unobserved factors. One computes the expected value of the complete data (x_t, f_t) log-likelihood and then iterate between the expectation (E) and maximization (M) steps. The procedure continues until convergence of the likelihood. In practice to speed up the computation we use the computational device of Jungbacker and Koopman (2008) and apply the Kalman filter and smoother algorithm to the collapsed observation vector.

[17] e.g. in a mixed frequency framework or in cases where some series have a shorter history such as for the euro area.

This model is defined as follow[18]:

$$z_t = A_0 + \sum_{l=1}^{p} A_l z_{t-l} + \varepsilon_t; \quad \varepsilon_t \backsim i.i.d.N(0, \Sigma) \tag{13}$$

where $z_t = (z_{1,t}...z_{n,t})'$ denotes the $n \times 1$ vector of variables in log-level or level for those series that are expressed in rates, A_0 is an $n \times 1$ vector of constants, $A_1, ..., A_p$ are $n \times n$ matrices of autoregressive coefficients. The coefficients $A_1, ..., A_p$ are assumed to be a priori independent and normally distributed random variables with mean and variances:

$$E[(A_l)_{iu}] = \begin{cases} \delta_i, & u = i, l = 1 \\ 0, & otherwise \end{cases} \quad and \quad V[(A_l)_{iu}] = \begin{cases} \frac{\lambda^2}{l^2} \\ \frac{\lambda^2}{l^2} \frac{\sigma_i^2}{\sigma_u^2} \end{cases} \tag{14}$$

where $\delta_i = 1$ for non-stationary variables and $\delta_i = 0$ for stationary ones.[19] The hyperparameter λ controls the overall tightness of the prior distribution around δ_i[20]: if $\lambda = 0$ the prior is imposed exactly and the data do not influence the estimate, whereas if $\lambda = \infty$ the posterior estimates are the OLS estimates. The prior on the intercept is diffuse and the covariance matrix of the residuals Σ is assumed to follow an inverse Wishart distribution.[21]

Furthermore, a sum of coefficients prior is also imposed which constrains $A_1 + ... + A_p$. This prior shrinks $I_n - A_1 - ... - A_p$ towards zero and is imposed exactly if this sum is zero which amounts to a VAR in first difference. A hyperparameter μ governs the degree of shrinkage of this prior.[22]

In practice, the priors are implemented by adding dummy observations. The specification ranges over which the model is estimated are: $p = 1, ..., 6$, $\lambda = 0.01{:}0.01{:}0.2$ and $\mu = 0.1\lambda, \lambda, 10\lambda$. The median of the posterior distribution of the parameters is used to compute point forecasts. Similar to the VAR, these forecasts are computed recursively using the Kalman filter and smoother to take into account the unbalanced part of the panel.

[18]For details see Bańbura, Giannone and Reichlin (2010).

[19]This corresponds to a random walk and white noise prior respectively. In practice the prior is set according to the transformations applied to the series to obtain stationarity as needed in the other models. Hence a white noise prior is used for those series which are untransformed in the X_t vector (interest spreads and some of the surveys) and a random prior is used for the other series (see Appendix A).

[20]The decay factor $1/l^2$ is the rate at which prior variance decreases with the lag length and $\sigma_i^2 \backslash \sigma_u^2$ accounts for the different scale and variability of the series. The parameters σ_i^2 is set to the estimate of the residuals variance from a univariate AR(p).

[21]For details see Kadiyala and Karlsson (1997) and Sims and Zha (1998).

[22]As $\mu \to 0$ the prior is more tightly imposed, whereas as $\mu \to \infty$ it is looser.

3.2 Forecast combination schemes

Out-of-sample forecasts are computed recursively conditional on the real-time information available at the end of each month over the period December 2001 to December 2011 and the estimation sample starts in January 1992. For each targeted variable j and each forecast horizon h we have a set of forecasts computed from the suite of models. With the exception of the RW, each model is estimated over a range of specifications as described in the previous subsection, and thereby generates multiple forecasts. To overcome the choice of model specification faced by the practitioner, e.g. which (information) criteria to use to select the parametrization of the model, and to help to insure against specification instability, two forecasts combination schemes are used to produce a single forecast for each model.

The combination schemes used assign time-varying and horizon specific weights based on past out-of-sample forecasting performance, as measured by the mean square forecast errors (MSFEs) statistic. A burning in phase of two-years starting in December 2001 is used to determine the initial weights. Firstly, we follow Aiolfi and Timmermann (2006) and take the unweighted average over the 10% first best specifications (*av.10%*) up to that point. Secondly, we consider the discounted-MSFEs (*av.d-msfe*) weighting schemes following Stock and Watson (2004) in which all specifications contribute to the average with a non-zero weight which is inversely proportional to their past forecasting performance. A discount factor equal to 0.9 is used which further assigns a higher weight to the more recent forecasting performance. Hence a single forecast is produced from combining over all the specifications for a given model (AR, BIV, DFM, BVAR) using these weighting schemes. Furthermore, pooling over all specifications from all models (ALL) and all multivariate models (MULT) using the aforementioned schemes is also considered.

4 Empirical results

4.1 Full-sample results

This section presents the forecasting results for the key series. The root mean square forecast errors (RMSFEs) statistic is used as a metric for evaluating the forecasts and the first release of the series of interest is used as actual. The evaluation period runs from December 2003 to December 2011. Since our goal is to assess the predictability at different horizons of real and nominal variables, as well as the ability of several models/combination schemes to exploit such predictability, (relative) predictability is defined as the forecasting ability of a given model/combination scheme relative to that of the RW model, i.e. the ratio between the RMSFEs of a given model/combination scheme and the naive RW model.

Figure 1 reports the relative predictive ability of the different forecasting models. Forecast horizons are displayed in rows and targeted series in colums; in each sub-figure the two consecutive same coloured bars for a given model display the results for the two combination schemes, i.e. *av.10%* and *av.d-msfe*. A bar below the solid black line (drawn at one), indicates a forecast that is more precise, on average, than the RW benchmark. To further compare the data-rich forecasting performance to those of the AR model, the dashed red line shows the relative RMSFEs of the best performing combination scheme for the AR model. Tables B.1 in Appendix B reports the actual numbers.

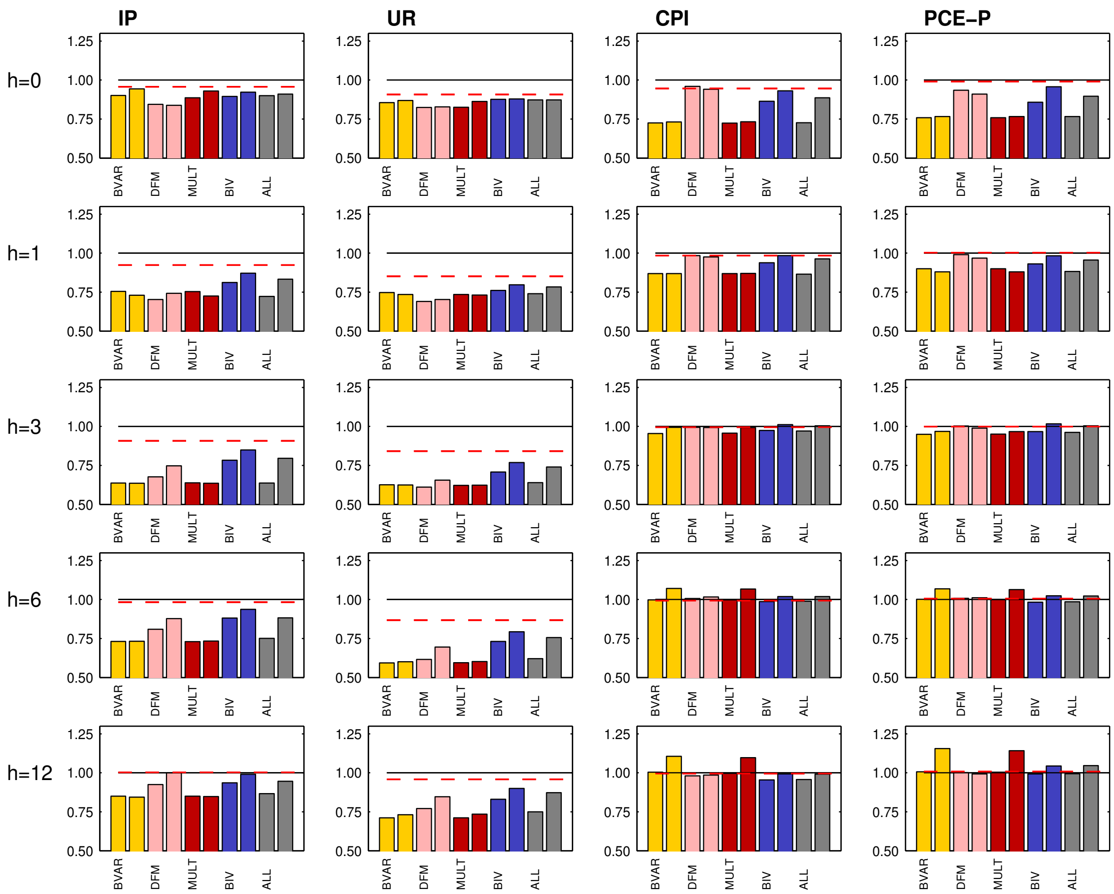
The following comments are made:

- For the **real variables** there is considerable predictability as all data-rich forecasts, irrespective of the model, combination schemes and forecast horizon, are always more accurate than the RW and AR models. The only exception is for one-year ahead forecasts for IP where one of the weighting scheme of the DFM does not yield lower RMSFEs than the RW. The DFM is the most accurate for nowcasting and pure forecasting at the short-term horizons up to $h = 1$ for IP and up to $h = 3$ for the UR. At longer horizons the BVAR and pooling over both multivariate models perform best.

- Regarding **inflation**, only at the very short-term horizons ($h = 0$ and $h = 1$) do all the data-rich forecasts beat the naive RW benchmark but do not always improve upon the best AR model. This is not surprising since inflation is known to be hard to forecast over the recent sample (see, among others, Stock and Watson, 2011 and D'Agostino, Giannone and Surico, 2006). The BVAR (and MULT) clearly stands out for both inflation series as it performs much better than the other models with reductions in RMSFEs of the order of 25% for nowcasts and 10% for one-month ahead forecasts. This is in line with the findings of Giannone, Lenza, Momferatou and Onorante (2010) who forecast euro area inflation over the period 2000 to 2009 with a real-time and "ragged-edge" dataset and also found strong improvement of the BVAR over the RW at the short-term horizons.

Notes: The Figure shows the relative (versus the RW) RMSFEs for each model over the full sample evaluation period which runs from December 2003 to December 2011. In each sub-figure the two consecutive (same coloured) bars for a given model display the results for the forecasts combination schemes used, i.e. av.10% and av.d-msfe. The dashed red line is the relative RMSFEs of the best AR model for the corresponding horizon and the black line is drawn at one (relative RMSFEs of RW).

Figure 1: Relative RMSFEs - full sample



4.2 Sub-sample results

The results presented above are for the full evaluation period which includes the two rather distinct sub-periods, namely the so-called (end of the) great moderation (pre-crisis) and the great recession (crisis), as such we further investigate the **stability/instability** of the results over these **pre-crisis** and **crisis** sub-samples. This is motivated by the findings of D'Agostino, Giannone and Surico (2006) that relative forecasting performance is related to macroeconomic volatility. These authors find that there is a sizable decline in the predictive accuracy of forecasts based on large datasets relative to univariate models in the post 1985 (great moderation) period, compared to the 70s and early 80s, which also coincides with a marked fall in macroeconomic volatility. D'Agostino and Giannone (2012) further show that most of the forecasting gains of the data-rich models over the univariate ones for the sample 1970 to 1998 are in fact confined to recession/downturn periods which are known to be characterized by higher volatility and comovements, especially among the real variables.

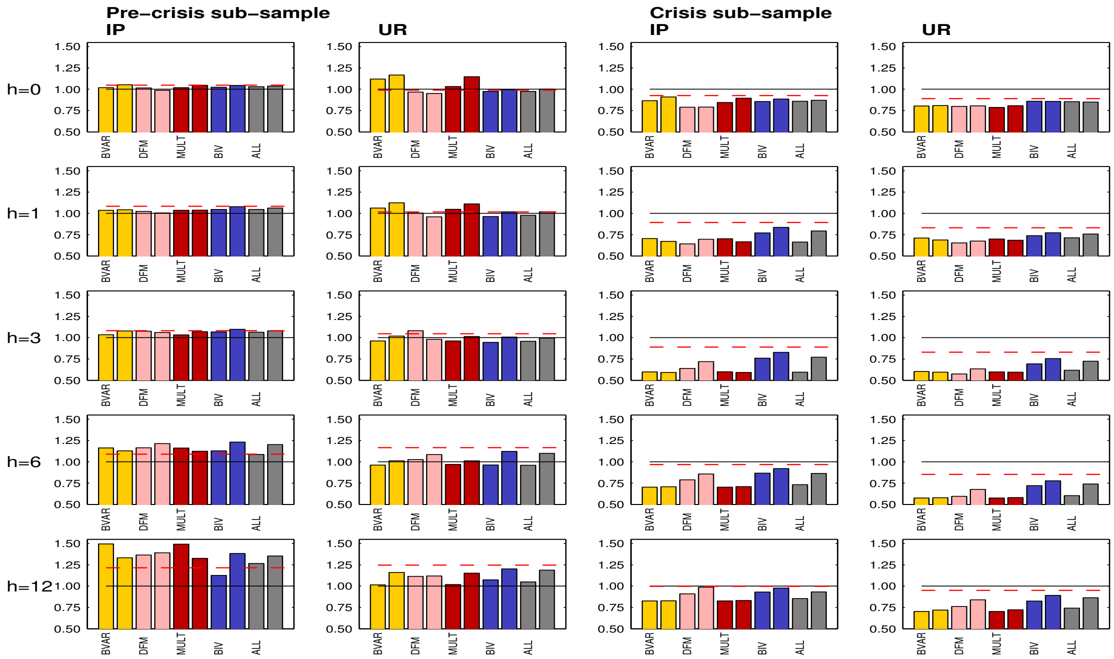
The sub-sample results for the real variables are displayed in the upper part of Figure 2 and for the nominal variables in the lower part; the following comments are made:

- The striking feature is that for the **real variables** indeed most of the full-sample predictability comes from the great recession period. Over the pre-crisis sample (left panel), there are only very marginal improvements over the univariate AR and RW models at nowcasting for IP and up to $h = 6$ for the UR. The picture that emerges when looking at the results over the crisis period (right panel) is quite different as evidenced by the fact that all relative RMSFEs are much smaller than one. The comments that can be made for the results over this period are in fact similar to those made previously over the full sample period. Broadly speaking, there is substantial predictability of the real variables, with this finding being robust across models and combination schemes and multivariate models yield more accurate forecasts that pooling of bi-variate models. For both variables, the DFM performs best at the very short-term horizons and the BVAR and MULT rank first at longer forecast horizons.

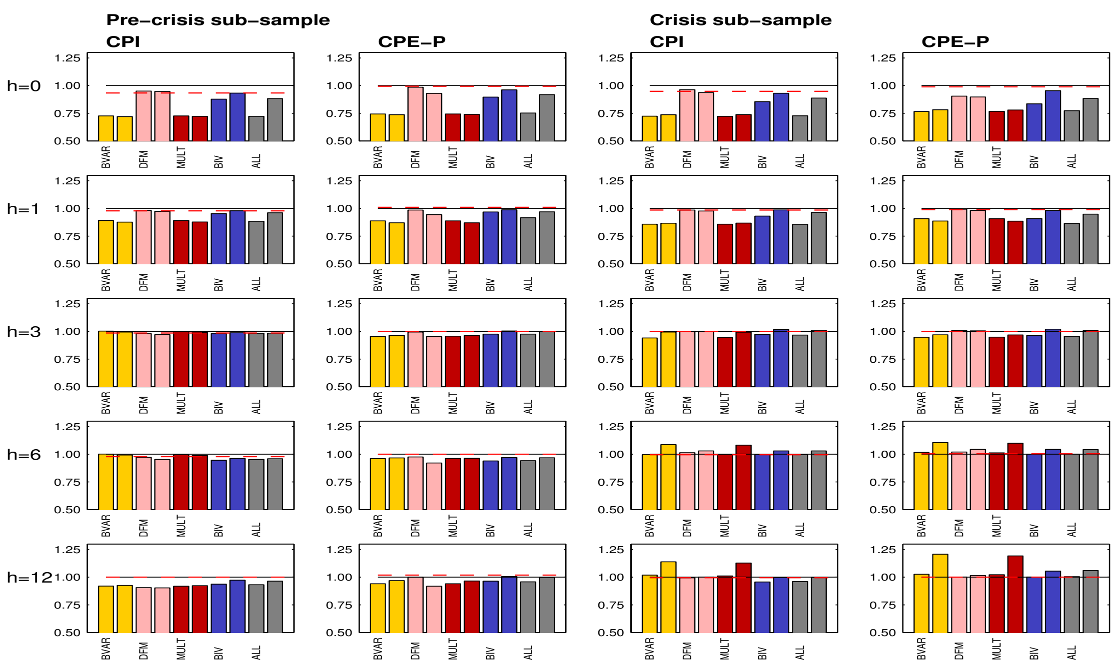
- For **inflation**, on the other hand, the results are quite stable across sub-periods. This is in line with Faust and Wright (2011) findings that forecasting results for inflation are not affected by inclusion or exclusion of the recent crisis, as inflation behavior has not been as extreme as that of the real variables. There is considerable predictability at horizon $h = 0$ and the BVAR stands out by far as the most precise nowcasting model for both inflation series. Also over both sub-samples, all models beat the RW up to horizon $h = 1$ and for longer horizons results are more mixed across models and combination schemes.

Figure 2: Relative RMSFEs - sub-samples

(a) Real variables



(b) Nominal variables



Notes: The figure shows the relative (versus the RW) RMSFEs for each model over the pre-crisis period December 2003-November 2007 (left panel) and crisis period December 2007-December 2011 (right panel). In each sub-figure the four consecutive (same coloured) bars for a given model display the results for the forecasts combination schemes used, i.e. av.10% and av.d-msfe. The dashed red line is the relative RMSFEs of the best AR model for the corresponding horizon and the black line is drawn at one (relative RMSFEs of RW).

To sum up, the results show that, when there is predictability, also in real-time does cross-sectional information helps at forecasting since data-rich forecasts not only beat the RW but also the AR model. For the real variables, predictability is confined over the recent recession/crisis period. This in line with the findings of D'Agostino and Giannone (2012) over an earlier period, that gains in relative performance of models using large datasets over univariate models are driven by downturn periods which are characterized by higher volatility and comovements. To further gauge the link between comovement and predictability, the upper part of Table 1 reports the percentage of the total panel variance accounted for by the first ten static and dynamic principal components whereas the lower part of the table displays the fraction of predictors that have information content for the targeted series.[23] Results show that over the crisis sample, when the forecasting gains of data-rich models is high, comovement is higher as well as the proportion of indicator which has information content for the targeted series.

Table 1: Comovement and predictability

Comovement:											
period / no. of PCs:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	
	Static PCs										
pre-crisis	0.16	0.28	0.38	0.45	0.51	0.56	0.61	0.65	0.69	0.73	
crisis	0.31	0.45	0.53	0.60	0.67	0.71	0.75	0.78	0.81	0.84	
	Dynamic PCs										
pre-crisis	0.38	0.59	0.73	0.83	0.89	0.93	0.95	0.96	0.97	0.98	
crisis	0.48	0.69	0.80	0.88	0.93	0.95	0.97	0.98	0.99	0.99	
Individual predictive content:											
period / for.hor.:	**h=0**	**h=1**	**h=3**	**h=6**	**h=12**		**h=0**	**h=1**	**h=3**	**h=6**	**h=12**
	IP						*UR*				
pre-crisis	0.10	0.13	0.13	0.13	0.06		0.10	0.10	0.06	0.04	0.03
crisis	0.67	0.73	0.72	0.60	0.60		0.70	0.67	0.70	70.0	0.67
	CPI						*PCE-P*				
pre-crisis	0.34	0.37	0.34	0.22	0.24		0.43	0.49	0.39	0.36	0.13
crisis	0.45	0.30	0.13	0.09	0.27		0.24	0.25	0.15	0.07	0.12

Notes: The upper part of the Table shows the percentage of total panel variance explained by the first ten static and dynamic principal components (PCs). The lower part of the Table shows the fraction of variables for each target series that have individual predictive content at a given horizon. The results are based on forecasts constructed using equation (4) without lags of the targeted series. A variable is considered to have predictive content if its RMSFEs are smaller than those of the RW model for all the forecasting combination schemes. All results are displayed over the pre-crisis (Dec.2003-Nov.2007) and crisis (Dec.2007-Dec.2011) samples

In such a situation when faced with many informative predictors that are highly collinear, i.e. admit a (approximate) factor structure, a forecaster is better off using all the information. De Mol, Giannone and Reichlin (2008) further show that in such a case, factor model

[23]For each targeted series, using each candidate predictor we generate forecasts at a given horizon with equation (4) but without lags of the targeted series over both sub-samples. The lower part of Table 1 reports the % of predictors which are found to have information content at a given horizon. This is the % of predictors that yield more accurate forecasts than the RW model, with a predictor being considered to be helpful if results are robust across combination schemes.

(principal components) forecasts and Bayesian forecasts under normal prior (with the degree of shrinkage chosen in relation to the cross-sectional dimension) yield similar results. In both models, regressors used to construct the forecasts are a linear combination of all variables (i.e. linear shrinkage) in the panel[24]; principal components forecasts put unit weight on the first (r) dominant eigenvalues of the covariance matrix of the data and zero on the others, while Bayesian forecasts assign decreasing weight to all the ordered eigenvalues.

The fact that the BVAR assigns non-zero weights to less important eigenvalues whereas the DFM gives them a zero weight also helps explain why the former model stands out for now-casting inflation in real-time. Our conjecture is that a commodity price component, which is the main driver of short-run inflation dynamics, is captured by these minor eigenvalues as commodity/oil prices and prices survey data for the reference month are known at the forecast origin. We will come back to this point in the sequel by looking at the marginal predictive ability of these timely indicators.

The results further show that the direct pooling of information using a high dimensional model (DFM or BVAR) which takes into account the cross-correlation between the variables, yields more accurate forecasts than pooling of bi-variate models, which Timmermann (2003) and Stock and Watson (2004), using a balanced datasets, report as a difficult benchmark to beat. These authors used the simple average over all specifications as a combination scheme for the bi-variate models, which in fact performs worse than the two combination schemes used in this study and hence does not change the ranking of the best performing models. This holds true for real and nominal variables. A further reason put forward for the better performance of direct pooling of information is that in a real-time setting the missing months at the forecast origin (or before), due to publication lags, are efficiently estimated taking into account all the available information via the Kalman filter and smoother.

Lastly, our findings also show that results are robust to forecasts combination schemes and that pooling over different models works. In instances when the BVAR or DFM performs less well, the pooling over these models always ranks among the best. The same can be said about pooling over all models as even if it only seldom gives the lowest RMSFEs for a given forecast horizon, it never ranks last. Although, when there are big differences between the different models RMSFEs, then generally it is only the combination scheme (av.10%) which stands out as it assigns weights only to the top 10% best past peformers. These results suggest that pooling across models is useful from a real-time forecaster perspective

[24]An alternative method is to use non-linear shrinkage such as Lasso regression, which simultaneously does shrinkage and variable selection. However, when variables are highly collinear, these forecasts should produce similar results to those obtained with DFM and Bayesian regression under normal prior, and the weights attached to individual predictors would be unstable. See De Mol, Giannone and Reichlin (2008). Stock and Watson (2011) further find that these alternative shrinkage estimators do not perform better than DFM.

who does not know ex-ante which model to choose among the competing ones.

4.3 Real-time versus revised data forecasts

In this part we briefly evaluate the sensitivity of the forecasting results to the data used to generate the forecasts. We compare forecasts constructed in real-time targeting the real-time outturn of the series of interest to revised data forecasts targeting the revised outturns. The revised data forecasts are based on pseudo real-time vintages which reproduce the pattern of real-time missing observations but using the latest released values for all series in the panel.

The full set of results are reported in Appendix B in Tables B.1 and B.2 and Table 2 below summarizes the salient feature. In each column, for a given model, results are reported for the real-time (rtd) and revised (rev) data forecasts. Firstly we investigate if whether or not the predictive ability of data-rich models are sensitive to the data used. These results are displayed in the upper part of the table which displays a triangle ($\triangle$) if in real-time a data-rich model beats the RW and if it further beats the AR too the triangle if filled-in ($\blacktriangle$). Similar results are shown for revised data using circles and filled-in circles (($\circ$) and ($\bullet$)). As evident from the table the overall results are robust to the data used to generate the forecasts as in only very few cases are the conclusions regarding predictability sensitive to the data used. Secondly, we investigate if the models ranking is sensitive to the real-time versus revised data issue. Hence in the lower part of the table, columns four to five show the % gains of the best performing data-rich model with respect to the RW and columns six and seven gives the winner name. For the real variables, irrespective of the data used to generate the forecast the DFM performs best at the short-term horizons and the BVAR for longer horizons. For inflation the BVAR strong outperformance at horizon $h = 0$ and $h = 1$ is also robust to the data used to generate the forecasts. The main difference that emerges is that, especially at the shorter horizons, the % gains of the best performer relative to the RW are higher using revised data for both real variables and for the PCE price index.

Table 2: Real-time versus revised data forecasts.

Pre-crisis sub-sample

	AR rtd	AR rev	BVAR rtd	BVAR rev	DFM rtd	DFM rev	MULT rtd	MULT rev	BIV rtd	BIV rev	ALL rtd	ALL rev
IP												
h=0	– –	– –	– –	– –	–●	– –	– –	– –	– –	– –	– –	– –
h=1	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –
h=3	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –
h=6	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –
h=12	– –	– –	– –	-▲	– –	▲▲	– –	-▲	– –	– –	– –	▲-
UR												
h=0	●–	▲▲	– –	– –	●●	▲▲	– –	△-	●○	▲▲	●-	▲▲
h=1	– –	– –	– –	– –	-●	▲▲	– –	– –	●-	▲▲	●-	▲▲
h=3	– –	– –	●-	▲▲	-●	▲▲	●-	▲▲	●-	▲▲	●-	▲▲
h=6	– –	– –	●-	▲▲	– –	– –	●-	▲▲	●-	– –	●-	▲-
h=12	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –
CPI												
h=0	●●	▲▲	●●	▲▲	○○	△△	●●	▲▲	●●	▲△	●●	▲▲
h=1	●●	▲▲	●●	▲▲	○●	△△	●●	▲▲	●○	▲△	●●	▲▲
h=3	●●	▲▲	-○	▲▲	●●	▲▲	-○	▲▲	●○	▲▲	●●	▲▲
h=6	●●	▲▲	-○	▲▲	●●	▲▲	-○	▲▲	●●	▲▲	●●	▲▲
h=12	– –	▲-	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲
PCE-P												
h=0	–●	▲▲	●●	▲▲	○●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲
h=1	– –	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲
h=3	– –	–▲	●●	▲▲	●●	▲▲	●●	▲▲	●–	▲▲	●–	▲▲
h=6	– –	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲
h=12	– –	– –	●●	▲▲	●●	▲▲	●●	▲▲	●–	▲▲	●–	▲▲

Crisis sub-sample

	AR rtd	AR rev	BVAR rtd	BVAR rev	DFM rtd	DFM rev	MULT rtd	MULT rev	BIV rtd	BIV rev	ALL rtd	ALL rev
IP												
h=0	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲
h=1	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲
h=3	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲
h=6	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲
h=12	– –	▲▲	●●	▲▲	●●	– –	●●	▲▲	●●	▲▲	●●	▲▲
UR												
h=0	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲
h=1	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲
h=3	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲
h=6	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲
h=12	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲
CPI												
h=0	●●	– –	●●	▲▲	○●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲
h=1	●-	– –	●●	▲▲	○●	▲-	●●	▲▲	●○	▲-	●●	▲▲
h=3	– –	– –	●●	▲-	– –	– –	●●	▲-	●-	▲-	●-	▲-
h=6	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –
h=12	●-	▲-	– –	– –	○-	– –	– –	– –	●-	▲-	●-	▲▲
PCE-P												
h=0	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲
h=1	●–	▲–	●●	▲▲	○●	▲▲	●●	▲▲	●●	▲▲	●●	▲▲
h=3	– –	– –	●●	▲–	– –	– –	●●	▲–	●–	▲–	●–	▲–
h=6	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –
h=12	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –	– –

% gains wrt RW (Pre-crisis sub-sample)

			best data-rich models rtd	rev	rtd	rev
IP						
h=0	-	-	1%	-	DFM	-
h=1	-	-	-	-	-	-
h=3	-	-	-	-	-	-
h=6	-	-	-	-	-	-
h=12	-	-	-	8%	-	DFM
UR						
h=0	-	-	5%	11%	DFM	DFM
h=1	-	-	4%	11%	DFM - BIV	DFM
h=3	-	-	5%	9%	BIV	DFM
h=6	-	-	4%	5%	BVAR - BIV - ALL	BVAR
h=12	-	-	-	-	-	-
CPI						
h=0	-	-	28%	27%	BVAR - MULT - ALL	BVAR
h=1	-	-	12%	14%	BVAR - MULT - ALL	BVAR
h=3	-	-	3%	6%	DFM	BVAR - MULT
h=6	-	-	5%	9%	DFM - BIV - ALL	BIV - ALL
h=12	-	-	10%	14%	DFM	DFM
PCE-P						
h=0	-	-	26%	35%	BVAR - MULT	BVAR
h=1	-	-	13%	23%	BVAR - MULT	BVAR - MULT
h=3	-	-	5%	15%	DFM	BVAR - MULT
h=6	-	-	8%	17%	DFM	BVAR - MULT
h=12	-	-	8%	12%	DFM	BVAR - MULT

% gains wrt RW (Crisis sub-sample)

			best data-rich models rtd	rev	rtd	rev
IP						
h=0	7%	10%	21%	27%	DFM	DFM
h=1	11%	18%	36%	36%	DFM	DFM
h=3	11%	19%	41%	41%	BVAR - MULT	BVAR - MULT
h=6	3%	10%	30%	32%	BVAR - MULT	BVAR - MULT
h=12	7%	2%	17%	15%	BVAR - MULT	BVAR - MULT - ALL
UR						
h=0	11%	14%	21%	33%	DFM	DFM
h=1	17%	20%	35%	46%	DFM	DFM
h=3	17%	20%	43%	49%	BVAR - MULT	DFM - MULT
h=6	15%	17%	42%	46%	BVAR - MULT	BVAR - MULT
h=12	5%	7%	30%	31%	BVAR - MULT	BVAR - MULT
CPI						
h=0	5%	-	28%	28%	BVAR - MULT	BVAR - MULT
h=1	1%	-	14%	13%	BVAR - MULT - ALL	BVAR - MULT
h=3	-	-	6%	5%	BVAR - MULT	BVAR - MULT
h=6	-	-	-	-	-	-
h=12	1%	1%	4%	4%	BIV - ALL	BIV - ALL
PCE-P						
h=0	1%	2%	23%	39%	BVAR - MULT - ALL	ALL
h=1	1%	1%	14%	19%	ALL	BVAR - MULT
h=3	-	-	5%	6%	BVAR - MULT	BVAR - MULT - ALL
h=6	-	-	-	-	-	-
h=12	-	-	-	-	-	-

4.4 Marginal predictive ability of surveys, financials and commodity prices

A related although distinct issue that we lastly investigate over these two periods is the predictive content of the timely blocks of variables, i.e. the surveys, financials and commodity prices.

The standard balanced panels literature includes mixed results regarding the predictive ability of these variables, in particular of the financials which are commonly thought of as leading indicators, using different approaches and over different samples.[25] From a real-time forecaster perspective, these blocks of data have additional value because of their timeliness. Since they are released ahead of the hard data, and are for the bulk of them already available at the end of the month they refer to, one would expect them to be more helpful at forecasting, especially at the short-term horizon. This has been extensively emphasized in the recent GDP nowcasting literature which finds that surveys help nowcasting growth especially at the beginning of the quarter when they are the only source of information on the current quarter. Their contribution to nowcasting mainly derives from their timeliness, as they are generally found to carry no information beyond and above hard data once the later have been released.[26]

To assess the marginal predictive ability for targeted series of these timely blocks, taking into account the real-time data flow, we revisit the previous results by comparing the forecasts generated with the following datasets:

- all variables;
- all variables excluding the surveys, financials and commodity prices;
- all variables excluding the surveys;
- all variables excluding the financials;
- all variables excluding the commodity prices.

Firstly we examine whether the overall results regarding the relative forecasting performance of models using large datasets over the univariate models, are dependent upon the type of information included. These results are reported in Figures B.1 to B.4 in Appendix B. Each figure shows the results for a given targeted series and sub-period; the sub-figures report the results using the different datasets for a given model and forecast horizon. At a glance, one can see that the overall finding of considerable predictability, in size and across models, for the real variables in the great-recession sub-period holds true irrespective of the dataset used to construct the forecasts. Whereas for inflation a worth mentioning pattern across both sub-samples is that forecasting performance of the BVAR in particular at the short-term

[25] See among others, Stock and Watson, 2003; Forni, Hallin, Lippi and Reichlin, 2003.

[26] See, among others, Giannone, Reichlin and Small, 2008; Liebermann, 2011; Bańbura and Runstler, 2011.

horizons, deteriorates considerably across models when excluding the soft data all together.

To further assess the marginal predictive content, above and beyond the hard data, of the timely blocks, Figure 3 reports the percentage change in RMSFEs resulting from the inclusion of all the soft data [27] in the panel, whereas the conditional (on the other blocks) marginal contribution of a given block are shown in Figures B.5 to B.8 in Appendix B. A bar below zero, means that the inclusion of all the soft data (a block) in the dataset decreases RMSFEs relative to those obtained from a dataset without them.

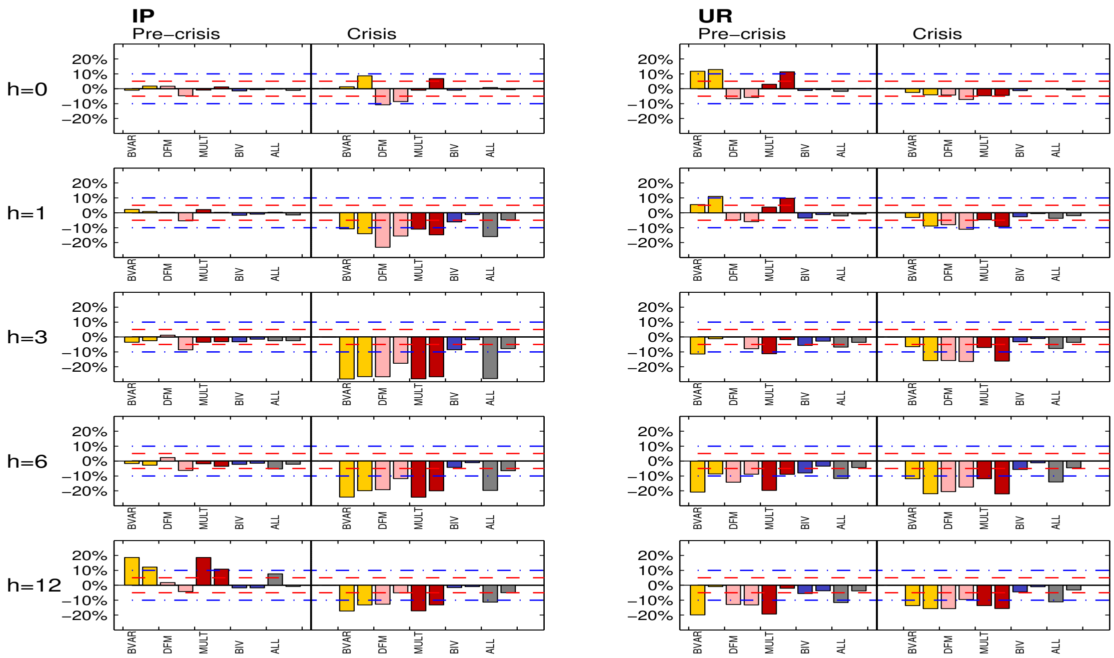
A first overall observation which helps to explain the mixed results found in the literature, is that for none of the key series, soft data (taken as a whole or looking at a specific block) uniformly, over both sub-samples and for all models (and combination schemes), have marginal predictive content over all horizons.

The evidence from Figure 3 shows that for the real variables, these timely blocks help more during the crisis period, and as such are also mostly really helpful when the hard data are. For IP the marginal cumulative reduction in MSFEs increases in absolute value up to horizon $h = 3$ and up to horizons $h = 6$ or $h = 12$ for the UR depending on the model and/or combination scheme. Also worth noting is that for IP over the more recent period the DFM and BVAR show opposite results regarding the contribution of soft data for nowcasting. Furthermore, the evidence in Appendix B shows that it is the surveys and financial block which have an impact on the real variables; and that surveys, conditional on the financials (and commodity prices) are helpful and vice versa. For both inflation series and across both sub-samples, the inclusion of the soft data block (driven by surveys and commodity prices) only helps to considerably improve the performance of the BVAR at $h = 0$ and $h = 1$, which empirically supports our conjecture made previously that this is the factor behind its better performance over the DFM for the nominal variables.

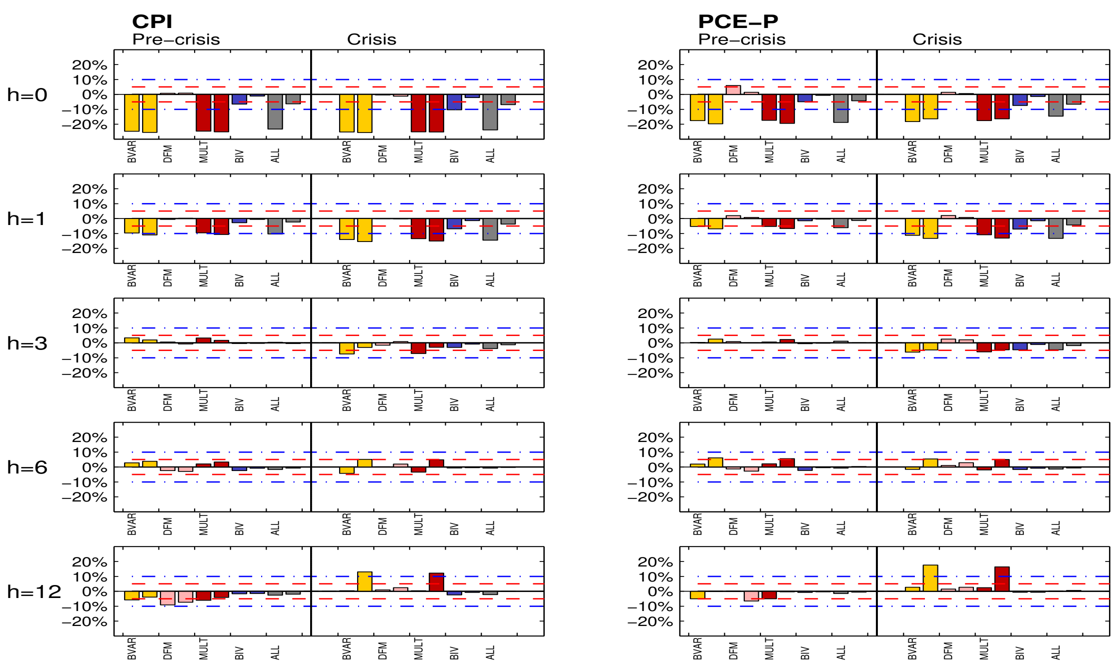
[27] For simplicity we classify the three blocks of timely variables as soft data.

Figure 3: Marginal predictive ability of soft data

(a) Real variables



(b) Inflation variables



Notes: The figure shows for each model the percentage change in RMSFEs from including the different respective block(s) of soft data. In each sub-figure the two consecutive bars for a given horizon displays the results for the four forecasts combination schemes used, i.e. av.10% and av.d-mse. The dashed and dashed dotted lines are drawn at the $+/-5\%$ and $+/-10\%$ threshold respectively.

5 Conclusion

This paper evaluates the ability of different models to forecast key real and nominal U.S. monthly macroeconomic variables in a data-rich environment and from the perspective of a real-time forecaster, i.e. taking into account the real-time data revisions process and data flow.

Our findings show that predictability of the real variables is confined over the recent recession/crisis period. This in line with the findings of D'Agostino and Giannone (2012) over an earlier period, that gains in relative performance of models using large datasets over univariate models are driven by downturn periods which are characterized by higher comovements. These results are robust to the combination schemes or models used. Regarding inflation, results are stable across time, but predictability is mainly found at nowcasting and forecasting one-month ahead, with the BVAR standing out at nowcasting. The results show that the forecasting gains at these short horizons stem mainly from exploiting timely information.

For both real and nominal variables, the direct pooling of information using a high dimensional model (DFM or BVAR) which takes into account the cross-correlation between the variables and efficiently deals with the "ragged-edge"structure of the dataset, yields more accurate forecasts than the indirect pooling of bi-variate forecasts/models. The fact that the DFM and the BVAR yield similar results (forecasts with comparable RMSFEs and highly correlated) when there is comovement has been shown by De Mol, Giannone and Reichlin (2008). In both models, regressors used to construct the forecasts are linear combinations of all variables in the panel; factor model forecasts put unit weight on the first dominant eigenvalues of the covariance matrix of the data and zero on the others, while Bayesian forecasts assign decreasing weights to all the ordered eigenvalues. The fact that the BVAR assigns non-zero weights to less important eigenvalues whereas the DFM gives them a zero weight also helps to explain why the former model stands out for nowcasting inflation in real-time. Presumably, an oil price component, which is the main driver of short-run inflation dynamics, is captured by these minor eigenvalues as commodity prices and prices survey data for the reference month are known at the forecast origin.

References

Aiolfi, M. and Timmermann, A. (2006): "Persistence in forecasting performance and conditional combination strategies", *Journal of Econometrics*, 135, pp.31-53.

Bańbura, M. and Modugno, M. (2010): *Maximum likelihood estimation of large factor model on datasets with arbitrary pattern of missing data*, Working paper series 1189, European Central Bank.

Bańbura, M., Giannone, D. and Reichlin, L. (2010): "Large Bayesian VARs", *Journal of Applied Econometrics*, 25(1), pp.71-92.

Bańbura, M., Giannone, D. and Reichlin, L. (2011): "Nowcasting", in *Oxford Handbook on Economic Forecasting*, ed. by Clements, M. and Hendry, D., part II, pp.193-224. Oxford University press.

Bańbura, M. and Rünstler, G. (2011): "A look into the factor model black box. Publication lags and the role of hard and soft data in forecasting GDP", *International Journal of Forecasting*, 27(2), pp.333-346.

Banternghansa, C. and McCracken, M. (2011): "Real-time forecast averaging with ALFRED", *Federal Reserve bank of St.Louis Review*, January/February 2011.

Clark, T. and McCracken, M. (2010): "Averaging forecasts from VARs with uncertain instabilities", *Journal of Applied Econometrics*, 25(1), pp.5-29.

D'Agostino, A., Giannone, D. and Surico, P. (2006): *(Un)predictability and macroeconomic stability,* Working paper series 605, European Central Bank.

D'Agostino, A. and Giannone, D. (2012): "Comparing alternative predictors based on large-panel factor models stability", *Oxford Bulletin of Economics and Statistics*, 74(2), pp.306-322.

Dempster, A., Laird, N. and Rubin, D. (1977): "Maximum likelihood from incomplete data via the EM algorithm", *Journal of the Royal Statistical Society*, 39(B), pp.1-38.

De Mol, C., Giannone, D. and Reichlin, L. (2008): "Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components?", *Journal Econometrics*, 146(2), pp.318-328.

Doan, T., Litterman, R. and Sims, C. (1984): "Forecasting and conditional projection using realistic prior distributions ", *Econometric Reviews*, 3(1), pp.1-100.

Doz, C., Giannone, D. and Reichlin, L. (2011a): "A two-step estimator for large approximate dynamic factor models based on Kalman filtering", *Journal of Econometrics*, 164(1), pp.188-205.

Doz, C., Giannone, D. and Reichlin, L. (2011b): "A quasi maximum likelihood approach for large approximate dynamic factor models", *Review of Economics and Statistics*, forthcoming.

Durbin, J. and Koopman, S.J. (2001): *Time series analysis by state space methods*, Oxford University Press.

Faust, J. and Wright, J. (2011): *Forecasting inflation*, mimeo.

Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2000): "The generalized dynamic factor model: identification and estimation", *Review of Economics and Statistics*, 82(4), pp.540-554.

Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2003): "Do financial variables help forecasting inflation and real activity in the euro area?", *Journal of Monetary Economics*, 50(6), pp.1243-1255.

Forni, M., Hallin, M., Lippi, M. and Reichlin, L. (2005): "The generalized dynamic factor model: one-sided estimation and forecasting", *Journal of the American Statistical Association*, 100(471), pp.830-840.

Giannone, D., Reichlin, L. and Small, D. (2008): "Nowcasting: the real-time information content of macroeconomic data releases", *Journal of Monetary Economics*, 55(4), pp.665-676.

Giannone, D., Lenza, M., Momferatou, D. and Onorante, L. (2010): *Short-term inflation projections: a Bayesian vector autoregressive approach,* CEPR Discussion paper series 7746.

Heij, C, Dijk, D. and Groenen, P. (2011): "Real-time forecasting with leading indicators: an empirical comparison", *Journal of Business Cycle Measurement and Analysis*, 4(1), pp.73-92.

Jungbacker, B. and Koopman, S. (2008): *L*ikelihood-based analysis for dynamic factor factor models, Tinbergen Institute Discussion paper series, 08-007/4.

Jungbacker, B., Koopman, S. and van der Wel, M. (2011): "Maximum likelihood estimation for dynamic factor models with missing data", *Journal of Economic Dynamic and Control*, 35, pp.1358-1368.

Kadiyala, K. and Karlsson, S. (1997): "Numerical methods for estimation and inference in Bayesian VAR-Models", *Journal of Applied Econometrics*, 12(2), pp.99-132.

King, T. Levin, A. and Perli, R. (2007): *Financial market perceptions of recession risk*,Finance and Economic discussion series 57, Federal Reserve Board.

Lenza, M. and Warmedinger, M. (2011):"A factor model for Euro-are short-term inflation analysis", *Jahrbücher f. Nationalökonomie u. Statistik*, 231(1), pp.50-62.

Liebermann, J. (2011): *Real-time nowcasting of GDP: factor model versus professional forecasters*, Technical paper series 03/11, Central Bank of Ireland.

Litterman R. (1986): "Forecasting with Bayesian Vector Autoregressions - five years of experience", *Journal of Business and Economics Statistics*, 4(1), pp.25-38.

Marcellino, M., Stock, J. and Watson, M. (2006): "A comparison of direct and iterated

AR methods for forecasting macroeconomic series", *Journal of Econometrics*, 135(1-2), pp.499-526.

Reis, R. and Watson, M. (2010): "Relative goods' prices and pure inflation,"*Americcan Economic Journal: Macroeconomics*, 2(3), pp.128-157.

Rossi, B. and Sekhposyan, T. (2010): "Have economic models' forecasting performance for US output growth and inflation changed over time, and when?", *International Journal of Forecasting*, 26(4), pp.808-835.

Shumway, R. and Stoffer, D. (1982): "An approach to time series smoothing and forecasting using the EM algorithm", *Journal of Time Series Analysis*, 3, pp.253-264.

Sims, C. and Zha, T. (1998): "Bayesian methods for dynamic multivariate models", *International Economic Review*, 39(4), pp.949-968.

Stock, J. and Watson, M. (2002a): "Forecasting Using Principal Components From a Large Number of Predictors", *Journal of the American Statistical Association*, 97(460), pp.147-162.

Stock, J. and Watson, M. (2002b): "Macroeconomic forecasting using diffusion" *Journal of Business and Economic Statistics*, 20(2), pp.147-162.

Stock, J. and Watson, M. (2003): "Forecasting output and inflation: The role of asset prices", *Journal of Economic Literature*, 41(3), pp.788-829.

Stock, J. and Watson, M. (2004): "Combination forecasts of output growth in a seven-country data set", *Journal of Forecasting*, 23(6), pp.405-430.

Stock, J. and Watson, M. (2011): *Generalized shrinkage methods for forecasting using many predictors*, manuscript.

Timmermann, A. (2006): "Forecast combinations", in *Handbook of Economic Forecasting*, ed. by Elliot, G., Granger, C. and Timmermann, A., Volume 1, pp.135-196. Elsevier.

Waggoner, D. and Zha, T. (1999): "Conditional forecasts in dynamic multivariate models", *The Review of Economics and Statistics*, 81(4), pp.639-651.

Watson, M. and Engle, R. (1983): "Alternative algorithms for the estimation of dynamic factor, mimic and varying coefficient regression", *Journal of Econometrics*, 23(3), pp.385-400.

Appendix A: Data description

Releases/reports	Sources	Series	Publication lag (in months)	Transformation code
Business Outlook Survey	(1)(b)	diffusion index of current activity	0	0
		diffusion index of current employment	0	0
		diffusion index of current price paid	0	0
		diffusion index of current price received	0	0
Manufacturing ISM Report on Business	(2)(a)(b)	ISM manufacturing: PMI composite index	1	0
NFIB Small Business Optimism index	(10)(c)	Small Business Optimism index	1	0
The Conference Board Consumer's index	(3)(d)	Index of consumer confidence	0	1
University of Michigan Consumer Sentiment index	(4)(b)(c)	Index of consumer sentiment (preliminary (4)(d), final (4)(a))	0	1
University of Michigan Inflation expectation	(4)(b)(c)	Index of consumer sentiment	0 or 1	1
The Employment Situation	(5)(a)	Employees on nonfarm payrolls: total	1	2
		Employees on nonfarm payrolls: manufacturing	1	2
		Employees on nonfarm payrolls: construction	1	2
		Employees on nonfarm payrolls: financial activities	1	2
		Employees on nonfarm payrolls: government	1	2
		Employees on nonfarm payrolls: goods-producing industries	1	2
		Employees on nonfarm payrolls: other servives	1	2
		Employees on nonfarm payrolls: service-providing industries	1	2
		Employees on nonfarm payrolls: retail trade	1	2
		Employees on nonfarm payrolls: wholesale trade	1	2
		Employees on nonfarm payrolls: total private industries	1	2
		Average hourly earnings: total private industries	1	3
		Average weekly hours: total private industries	1	2
		Average weekly hours: manufacturing	1	2
		Average weekly hours: overtime - manuf.	1	2
		Civilian unemployment rate	1	1
		Civilian participation rate	1	1
		Mean duration of unemployment	1	2
G.17 Industrial Production and Capacity Utilization	(6)(a)	Industrial production : total	1	2
		Capacity utilization: total	1	1
Advance Monthly Sales for Retail and Food Services	(7)(a)	Retail and food services sales	1	2
Manufacturing and Trade Inventories and Sales	(7)(a)	Inventories: total business	2	2
		Inventories to sales ratio: total business	2	2
Manufacturer's Shipments, Inventories, and Orders (M3)	(7)(a)	Manufacturer's new orders: durable goods	1	2
		Manufacturer's new orders: non defense capital goods excluding aircraft	2	2
Supplemental Estimates, Motor Vehicles	(8)(a)	Light weight vehicle sales: autos & light trucks	1	2
New Residential Construction	(7)(a)	Housing starts: total	1	2
		Housing starts: 1-unit structure	1	2
		New private housing units authorized by building permit	1	2
New Residential Sales	(7)(a)	New one family houses sold: U.S.	1	2
Personal Income and Outlays	(8)(a)	Real disposable personal income	1 or 2	2
		Real personal consumption expenditures	1 or 2	2
		Real personal consumption expenditures: durable goods	1 or 2	2
		Real personal consumption expenditures: nondurable goods	1 or 2	2
		Real personal consumption expenditures: services	1 or 2	2
		Personal saving rate	1 or 2	1
		Personal consumption expenditures: chain-type price index	1 or 2	2
		Personal consumption expenditures: chain-type price index less food & energy	1 or 2	2

Releases/reports	Sources	Series	Publication lag (in months)	Transformation code
Consumer Price Index	(5)(a)	Consumer price index for all urban consumers: all items	1	2
		Consumer price index for all urban consumers: all items less food & energy	1	2
		Consumer price index for all urban consumers: food	1	2
		Consumer price index for all urban consumers: energy	1	2
Producer Price Index	(5)(a)	Produce price index: finished goods	1	2
		Produce price index: finished goods less food & energy	1	2
		Produce price index: finished consumer goods excluding foods	1	2
		Produce price index: finished energy goods	1	2
		Produce price index: crude materials for further processing	1	2
		Produce price index: intermediate materials: supplies & components	1	2
G.19. Consumer Credit	(6)(a)	Total consumer credit outstanding	2	2
H.6. Money Stock Measures	(6)(a)	M2 money stock	1	2
H.8. Assets and Liabilities of Comm. Banks in the U.S.	(6)(a)	Consumer (Individual) Loans at all commercial banks	1	2
		Commercial and Industrial Loans at all commercial banks	1	2
Selected Interest Rates	(6)	Federal funds rate	0	1
		spread of 10-year U.S. Treasury yield (constant maturity) over the Federal funds rate	0	0
		spread of Moody's BAA Corporate bonds yield over 10-year U.S. Treasury yield (constant maturity)	0	0
Exchange rate	(6)	Usd versus weighted average of foreign currencies (broad)	0	2
S&P	(c)	S&P500 Composite	0	2
Commodity price index	(c)	Commodity price index	0	2
WTI Oil price	(c)	Oil price	0	2

Notes:

The second column refers to the following sources:

(1) The Federal Reserve Bank of Philadelphia
(2) The Institute for Supply Management
(3) The Conference Board
(4) The University of Michigan
(5) U.S. Department of Labor: Bureau of Labor Statistics
(6) Board of Governors of the Federal Reserve System
(7) U.S. Department of Commerce: Census Bureau
(8) U.S. Department of Commerce: Bureau of Economic Analysis
(9) U.S. Department of Commerce: Bureau of Economic Analysis and Census Bureau
(10) NFIB Research Foundation

(a) The Federal Reserve Bank of St.Louis, ALFRED database
(b) Bloomberg
(c) Datastream

The last column refers to the following transformation codes:

code:	transformation to stationary:	transformation used in the the BVAR in level:	prior δ_i:
0	$X_{it} = Z_{it}$	$z_{it} = Z_{it}$	0
1	$X_{it} = (1 - L)Z_{it}$	$z_{it} = Z_{it}$	1
2	$X_{it} = 100 \times (1 - L)log(Z_{it})$	$z_{it} = log(Z_{it})$	1

Appendix B: Tables and figures

Table B.1: RMSFEs relative to the RW

Full-sample

	AR 10%	AR d-mse	BVAR 10%	BVAR d-mse	DFM 10%	DFM d-mse	MULT 10%	MULT d-mse	BIV 10%	BIV d-mse	ALL 10%	ALL d-mse	AR 10%	AR d-mse	BVAR 10%	BVAR d-mse	DFM 10%	DFM d-mse	MULT 10%	MULT d-mse	BIV 10%	BIV d-mse	ALL 10%	ALL d-mse
	IP												CPI											
h=0	0.98	0.96	0.90	0.94	0.84	0.84	0.89	0.93	0.90	0.92	0.90	0.91	0.95	0.95	0.72	0.73	0.96	0.94	0.72	0.73	0.86	0.93	0.73	0.89
h=1	0.97	0.92	0.75	0.73	0.70	0.74	0.75	0.73	0.81	0.87	0.72	0.83	0.98	1.00	0.87	0.87	0.98	0.98	0.87	0.87	0.94	0.98	0.87	0.96
h=3	0.95	0.91	0.64	0.64	0.68	0.75	0.64	0.64	0.78	0.85	0.64	0.80	1.00	1.01	0.96	0.99	0.99	0.99	0.96	0.99	0.98	1.01	0.97	1.00
h=6	0.99	0.98	0.73	0.73	0.81	0.88	0.73	0.73	0.88	0.94	0.75	0.88	0.99	1.01	1.00	1.07	1.01	1.02	1.00	1.07	0.99	1.02	0.99	1.02
h=12	1.00	1.03	0.85	0.84	0.92	1.00	0.85	0.85	0.94	0.99	0.87	0.95	1.00	1.00	1.00	1.11	0.98	0.99	1.00	1.10	0.95	0.99	0.96	0.99
	UR												PCE-P											
h=0	0.95	0.91	0.86	0.87	0.82	0.83	0.83	0.86	0.88	0.88	0.87	0.87	0.99	0.99	0.76	0.77	0.93	0.91	0.76	0.77	0.86	0.96	0.77	0.90
h=1	0.90	0.85	0.75	0.73	0.69	0.70	0.73	0.73	0.76	0.80	0.74	0.78	1.00	1.00	0.90	0.88	0.99	0.97	0.90	0.88	0.93	0.98	0.88	0.96
h=3	0.86	0.84	0.63	0.62	0.61	0.66	0.62	0.62	0.71	0.77	0.64	0.74	1.00	1.02	0.95	0.97	1.00	0.99	0.95	0.97	0.97	1.02	0.96	1.00
h=6	0.92	0.87	0.59	0.60	0.62	0.70	0.59	0.60	0.73	0.79	0.62	0.76	1.01	1.02	1.00	1.07	1.01	1.01	1.00	1.06	0.98	1.02	0.99	1.02
h=12	0.99	0.96	0.71	0.73	0.77	0.85	0.71	0.73	0.83	0.90	0.75	0.87	1.01	1.03	1.01	1.15	1.00	0.99	1.00	1.14	0.99	1.04	0.99	1.05

Pre-crisis sub-sample

	AR 10%	AR d-mse	BVAR 10%	BVAR d-mse	DFM 10%	DFM d-mse	MULT 10%	MULT d-mse	BIV 10%	BIV d-mse	ALL 10%	ALL d-mse	AR 10%	AR d-mse	BVAR 10%	BVAR d-mse	DFM 10%	DFM d-mse	MULT 10%	MULT d-mse	BIV 10%	BIV d-mse	ALL 10%	ALL d-mse
	IP												CPI											
h=0	1.05	1.06	1.02	1.05	1.02	0.99	1.02	1.05	1.02	1.04	1.03	1.04	0.93	0.94	0.73	0.72	0.95	0.95	0.73	0.72	0.88	0.93	0.72	0.88
h=1	1.08	1.11	1.04	1.04	1.02	1.00	1.04	1.04	1.05	1.08	1.05	1.06	0.98	0.99	0.89	0.88	0.98	0.97	0.89	0.88	0.95	0.98	0.88	0.96
h=3	1.08	1.13	1.04	1.08	1.08	1.06	1.03	1.07	1.07	1.10	1.06	1.08	0.99	0.99	1.00	0.99	0.98	0.97	1.00	0.99	0.98	0.99	0.98	0.98
h=6	1.09	1.25	1.16	1.13	1.16	1.21	1.16	1.12	1.13	1.23	1.09	1.20	0.98	0.98	1.00	0.99	0.97	0.95	1.00	0.99	0.95	0.96	0.95	0.96
h=12	1.22	1.40	1.50	1.33	1.37	1.39	1.49	1.33	1.13	1.38	1.27	1.35	1.00	1.00	0.92	0.93	0.91	0.90	0.92	0.92	0.94	0.97	0.93	0.96
	UR												PCE-P											
h=0	0.99	1.01	1.12	1.17	0.97	0.95	1.03	1.15	0.98	0.99	0.97	1.00	1.00	0.99	0.74	0.74	0.99	0.93	0.74	0.74	0.90	0.96	0.75	0.92
h=1	1.02	1.05	1.06	1.12	1.00	0.96	1.05	1.11	0.96	1.02	0.98	1.02	1.03	1.01	0.89	0.87	0.99	0.95	0.89	0.87	0.97	0.99	0.92	0.97
h=3	1.06	1.05	0.96	1.02	1.08	0.98	0.96	1.01	0.95	1.01	0.96	1.00	1.00	1.02	0.95	0.96	0.99	0.95	0.96	0.96	0.98	1.00	0.98	1.00
h=6	1.17	1.19	0.96	1.01	1.03	1.09	0.97	1.01	0.96	1.12	0.96	1.10	1.01	1.00	0.96	0.97	0.98	0.92	0.96	0.96	0.94	0.97	0.94	0.97
h=12	1.25	1.27	1.01	1.16	1.11	1.12	1.02	1.15	1.07	1.20	1.05	1.19	1.03	1.02	0.94	0.97	1.00	0.92	0.94	0.97	0.96	1.01	0.96	1.00

Crisis sub-sample

	AR 10%	AR d-mse	BVAR 10%	BVAR d-mse	DFM 10%	DFM d-mse	MULT 10%	MULT d-mse	BIV 10%	BIV d-mse	ALL 10%	ALL d-mse	AR 10%	AR d-mse	BVAR 10%	BVAR d-mse	DFM 10%	DFM d-mse	MULT 10%	MULT d-mse	BIV 10%	BIV d-mse	ALL 10%	ALL d-mse
	IP												CPI											
h=0	0.96	0.93	0.86	0.91	0.79	0.79	0.84	0.89	0.85	0.88	0.86	0.87	0.95	0.95	0.72	0.74	0.96	0.94	0.72	0.74	0.86	0.93	0.73	0.89
h=1	0.96	0.89	0.70	0.67	0.64	0.70	0.70	0.67	0.77	0.84	0.66	0.79	0.99	1.00	0.86	0.87	0.99	0.98	0.86	0.87	0.93	0.99	0.86	0.97
h=3	0.94	0.89	0.60	0.59	0.64	0.72	0.60	0.59	0.76	0.83	0.60	0.77	1.00	1.02	0.94	0.99	1.00	1.00	0.94	0.99	0.97	1.02	0.97	1.01
h=6	0.98	0.97	0.70	0.71	0.79	0.86	0.70	0.71	0.87	0.92	0.73	0.86	1.00	1.02	1.00	1.09	1.01	1.03	1.00	1.08	1.00	1.03	1.00	1.03
h=12	1.00	1.02	0.83	0.83	0.91	0.99	0.83	0.83	0.93	0.98	0.85	0.93	0.99	1.01	1.02	1.14	0.99	1.00	1.01	1.13	0.96	1.00	0.96	1.00
	UR												PCE-P											
h=0	0.94	0.89	0.80	0.81	0.80	0.81	0.79	0.81	0.86	0.86	0.85	0.85	0.99	0.99	0.77	0.78	0.90	0.90	0.77	0.78	0.84	0.95	0.77	0.88
h=1	0.89	0.83	0.71	0.69	0.65	0.67	0.70	0.69	0.71	0.77	0.71	0.76	0.99	1.00	0.91	0.89	0.99	0.98	0.91	0.88	0.91	0.98	0.86	0.95
h=3	0.85	0.83	0.60	0.60	0.57	0.63	0.60	0.60	0.69	0.75	0.62	0.72	1.00	1.02	0.95	0.97	1.01	1.01	0.95	0.97	0.96	1.02	0.96	1.01
h=6	0.90	0.85	0.58	0.58	0.60	0.68	0.58	0.58	0.72	0.78	0.60	0.74	1.00	1.03	1.02	1.11	1.02	1.04	1.01	1.10	1.00	1.04	1.00	1.04
h=12	0.98	0.95	0.70	0.72	0.76	0.84	0.70	0.72	0.82	0.89	0.74	0.86	1.00	1.03	1.03	1.21	1.00	1.01	1.02	1.19	1.00	1.06	1.00	1.06

Notes: The table displays the relative RMSFEs for each model and combination scheme over the RW benchmark. The full evaluation sample runs from December 2003 to December 2011. Numbers below one are put in bold and are further shaded if the relative RMSFE is lower than that of the best AR model for the corresponding forecast horizon.

Table B.2: RMSFEs relative to the RW - revised data forecasts and targets

Full-sample

	AR 10%	AR d-mse	BVAR 10%	BVAR d-mse	DFM 10%	DFM d-mse	MULT 10%	MULT d-mse	BIV 10%	BIV d-mse	ALL 10%	ALL d-mse	AR 10%	AR d-mse	BVAR 10%	BVAR d-mse	DFM 10%	DFM d-mse	MULT 10%	MULT d-mse	BIV 10%	BIV d-mse	ALL 10%	ALL d-mse
	IP												CPI											
h=0	0.95	0.92	0.90	0.91	0.80	0.84	0.89	0.90	0.90	0.90	0.90	0.89	0.98	0.98	0.73	0.75	0.96	0.97	0.73	0.75	0.88	0.94	0.74	0.90
h=1	0.89	0.86	0.76	0.73	0.70	0.74	0.75	0.72	0.80	0.83	0.74	0.79	1.00	1.01	0.87	0.89	0.97	1.00	0.87	0.89	0.95	1.00	0.88	0.98
h=3	0.86	0.84	0.66	0.64	0.67	0.73	0.66	0.64	0.77	0.80	0.66	0.75	0.99	1.01	0.95	1.02	1.02	1.02	0.95	1.01	0.98	1.01	0.97	1.01
h=6	0.93	0.91	0.72	0.71	0.80	0.85	0.72	0.71	0.84	0.88	0.73	0.82	1.00	1.01	1.00	1.10	1.02	1.04	1.00	1.09	0.99	1.02	0.99	1.02
h=12	0.98	0.99	0.85	0.86	0.99	1.02	0.85	0.87	0.88	0.95	0.85	0.92	0.99	1.00	1.01	1.14	1.02	1.02	1.00	1.13	0.96	1.00	0.96	1.00
	UR												PCE-P											
h=0	0.89	0.88	0.79	0.80	0.71	0.73	0.75	0.79	0.81	0.83	0.80	0.82	0.96	0.97	0.65	0.64	0.89	0.88	0.65	0.64	0.79	0.92	0.64	0.83
h=1	0.83	0.82	0.66	0.66	0.58	0.61	0.64	0.66	0.70	0.75	0.67	0.73	0.99	1.00	0.80	0.83	0.96	0.96	0.80	0.83	0.89	0.97	0.82	0.93
h=3	0.83	0.81	0.55	0.56	0.54	0.60	0.55	0.56	0.67	0.73	0.58	0.70	1.01	1.01	0.93	0.97	1.01	0.99	0.93	0.97	0.94	1.00	0.93	0.99
h=6	0.88	0.85	0.56	0.58	0.61	0.69	0.56	0.58	0.72	0.77	0.61	0.74	1.02	1.01	1.00	1.09	1.03	1.03	1.00	1.08	0.97	1.01	0.97	1.01
h=12	0.95	0.94	0.70	0.72	0.84	0.88	0.70	0.73	0.83	0.89	0.75	0.86	1.01	1.02	1.01	1.16	1.03	1.03	1.01	1.15	0.98	1.03	0.98	1.03

Pre-crisis sub-sample

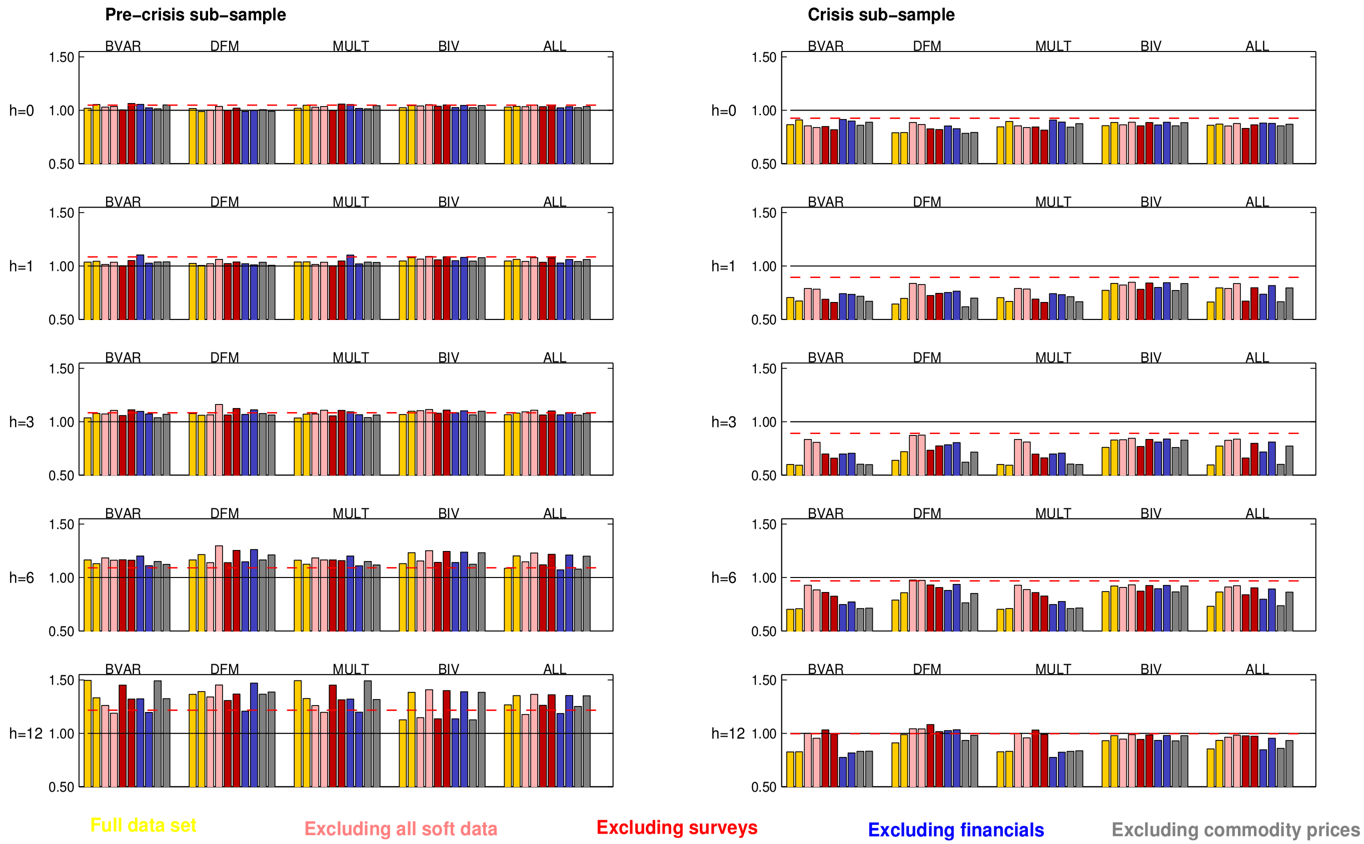
	AR 10%	AR d-mse	BVAR 10%	BVAR d-mse	DFM 10%	DFM d-mse	MULT 10%	MULT d-mse	BIV 10%	BIV d-mse	ALL 10%	ALL d-mse	AR 10%	AR d-mse	BVAR 10%	BVAR d-mse	DFM 10%	DFM d-mse	MULT 10%	MULT d-mse	BIV 10%	BIV d-mse	ALL 10%	ALL d-mse
	IP												CPI											
h=0	1.06	1.03	1.21	1.11	1.01	1.01	1.21	1.10	1.06	1.02	1.07	1.02	0.93	0.94	0.75	0.73	0.96	0.96	0.75	0.74	0.89	0.93	0.74	0.89
h=1	1.08	1.11	1.08	1.14	1.08	1.07	1.08	1.13	1.10	1.10	1.10	1.09	0.98	0.98	0.89	0.86	0.99	0.98	0.89	0.87	0.96	0.98	0.87	0.96
h=3	1.05	1.19	1.08	1.21	1.07	1.11	1.09	1.20	1.11	1.18	1.11	1.17	0.99	0.99	0.96	0.94	0.97	0.96	0.96	0.94	0.96	0.98	0.95	0.97
h=6	1.15	1.35	1.20	1.31	1.15	1.21	1.20	1.30	1.27	1.35	1.21	1.32	0.98	0.97	0.95	0.92	0.94	0.94	0.94	0.92	0.91	0.94	0.91	0.93
h=12	1.24	1.16	1.15	0.94	0.98	0.97	1.15	0.92	1.19	1.12	0.98	1.02	0.99	1.00	0.92	0.89	0.86	0.91	0.92	0.89	0.94	0.97	0.93	0.96
	UR												PCE-P											
h=0	0.98	0.98	1.09	1.11	0.91	0.89	0.98	1.09	0.97	0.96	0.96	0.97	0.92	0.95	0.69	0.65	0.88	0.87	0.68	0.66	0.83	0.89	0.68	0.83
h=1	1.04	1.02	1.07	1.12	0.96	0.89	1.01	1.10	0.96	0.98	0.98	0.98	0.98	0.98	0.79	0.77	0.92	0.90	0.79	0.77	0.90	0.94	0.79	0.91
h=3	1.05	1.03	0.96	0.96	0.99	0.91	0.96	0.96	0.92	0.98	0.94	0.96	1.01	0.99	0.88	0.85	0.91	0.89	0.88	0.85	0.91	0.95	0.90	0.94
h=6	1.19	1.14	0.95	0.97	1.07	1.02	0.97	0.96	1.00	1.06	0.98	1.04	0.99	0.97	0.87	0.83	0.88	0.86	0.87	0.83	0.86	0.91	0.86	0.90
h=12	1.27	1.23	1.03	1.11	1.08	1.12	1.03	1.10	1.08	1.15	1.04	1.14	1.02	1.00	0.90	0.88	0.90	0.90	0.90	0.88	0.91	0.96	0.90	0.95

Crisis sub-sample

	AR 10%	AR d-mse	BVAR 10%	BVAR d-mse	DFM 10%	DFM d-mse	MULT 10%	MULT d-mse	BIV 10%	BIV d-mse	ALL 10%	ALL d-mse	AR 10%	AR d-mse	BVAR 10%	BVAR d-mse	DFM 10%	DFM d-mse	MULT 10%	MULT d-mse	BIV 10%	BIV d-mse	ALL 10%	ALL d-mse
	IP												CPI											
h=0	0.92	0.90	0.81	0.85	0.73	0.79	0.80	0.84	0.86	0.87	0.85	0.85	1.00	1.00	0.72	0.76	0.96	0.98	0.72	0.76	0.88	0.94	0.74	0.90
h=1	0.86	0.82	0.71	0.66	0.64	0.69	0.70	0.66	0.76	0.79	0.68	0.75	1.01	1.02	0.87	0.90	0.97	1.01	0.87	0.90	0.95	1.01	0.88	0.99
h=3	0.85	0.81	0.63	0.59	0.64	0.70	0.63	0.59	0.74	0.77	0.62	0.71	1.00	1.01	0.95	1.03	1.03	1.03	0.95	1.03	0.98	1.02	0.98	1.02
h=6	0.92	0.90	0.70	0.68	0.79	0.83	0.70	0.68	0.82	0.86	0.71	0.80	1.00	1.02	1.00	1.12	1.03	1.06	1.01	1.12	1.00	1.03	1.00	1.04
h=12	0.98	0.98	0.85	0.86	1.00	1.02	0.85	0.87	0.87	0.95	0.85	0.92	0.99	1.00	1.02	1.18	1.04	1.04	1.02	1.17	0.96	1.00	0.96	1.00
	UR												PCE-P											
h=0	0.88	0.86	0.72	0.73	0.67	0.70	0.70	0.72	0.77	0.81	0.76	0.79	0.98	0.98	0.63	0.63	0.90	0.89	0.63	0.63	0.78	0.94	0.61	0.83
h=1	0.81	0.80	0.61	0.61	0.54	0.58	0.59	0.60	0.67	0.73	0.64	0.70	0.99	1.01	0.81	0.86	0.98	0.98	0.81	0.86	0.89	0.98	0.83	0.95
h=3	0.82	0.80	0.52	0.53	0.51	0.58	0.51	0.53	0.66	0.71	0.56	0.68	1.01	1.02	0.94	1.00	1.04	1.02	0.94	1.00	0.95	1.01	0.94	1.00
h=6	0.86	0.83	0.54	0.56	0.59	0.67	0.54	0.56	0.71	0.76	0.59	0.72	1.03	1.03	1.03	1.15	1.07	1.07	1.03	1.14	1.00	1.04	1.00	1.04
h=12	0.94	0.93	0.69	0.71	0.83	0.88	0.69	0.72	0.83	0.88	0.74	0.85	1.01	1.03	1.05	1.26	1.08	1.08	1.05	1.25	1.01	1.05	1.01	1.06

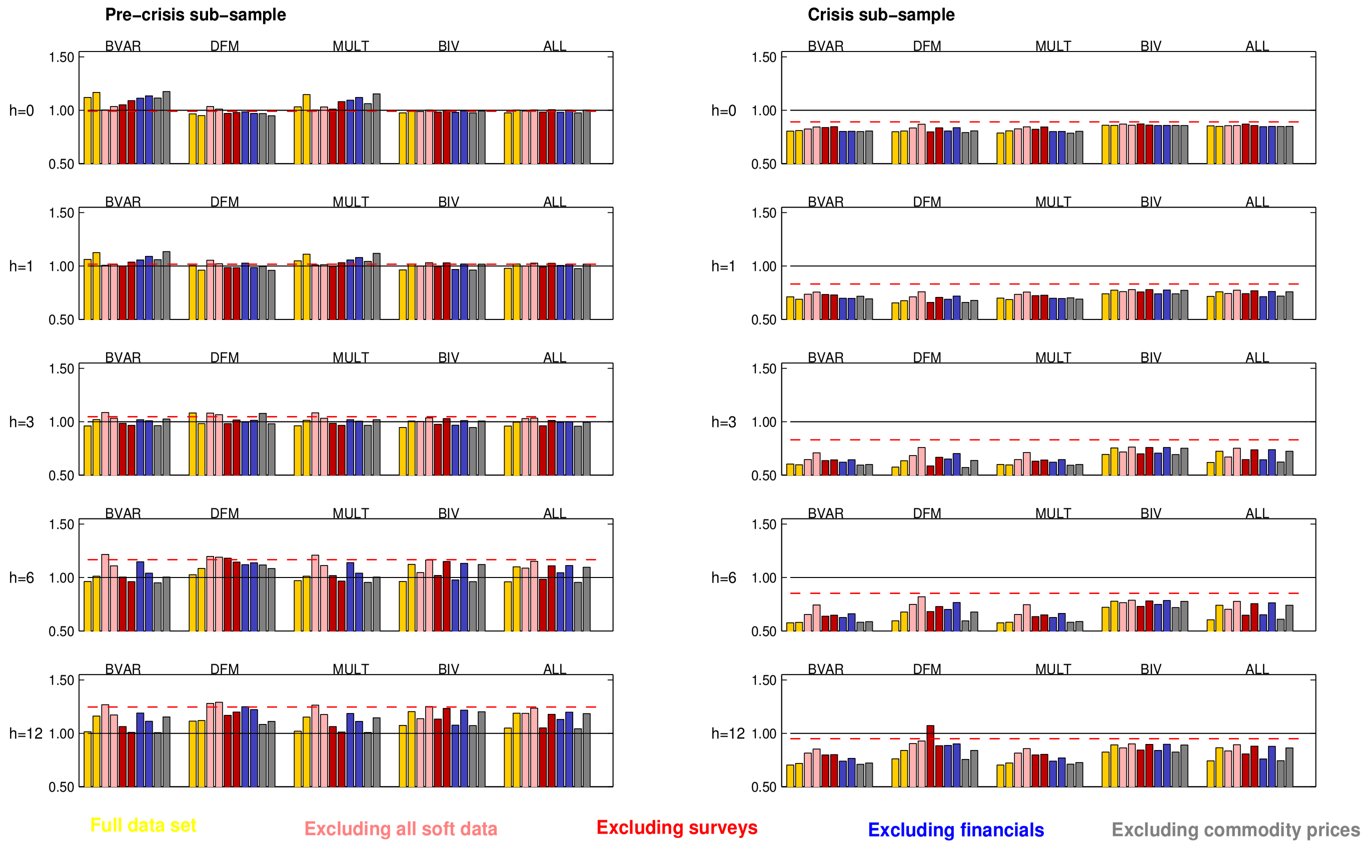
Notes: The table displays the relative RMSFEs for each model and combination scheme over the RW benchmark. The full evaluation sample runs from December 2003 to December 2011. Numbers below one are put in bold and are further shaded if the relative RMSFE is lower than that of the best AR model for the corresponding forecast horizon.

Figure B.1: Relative RMSFEs of the different panels for **industrial production**



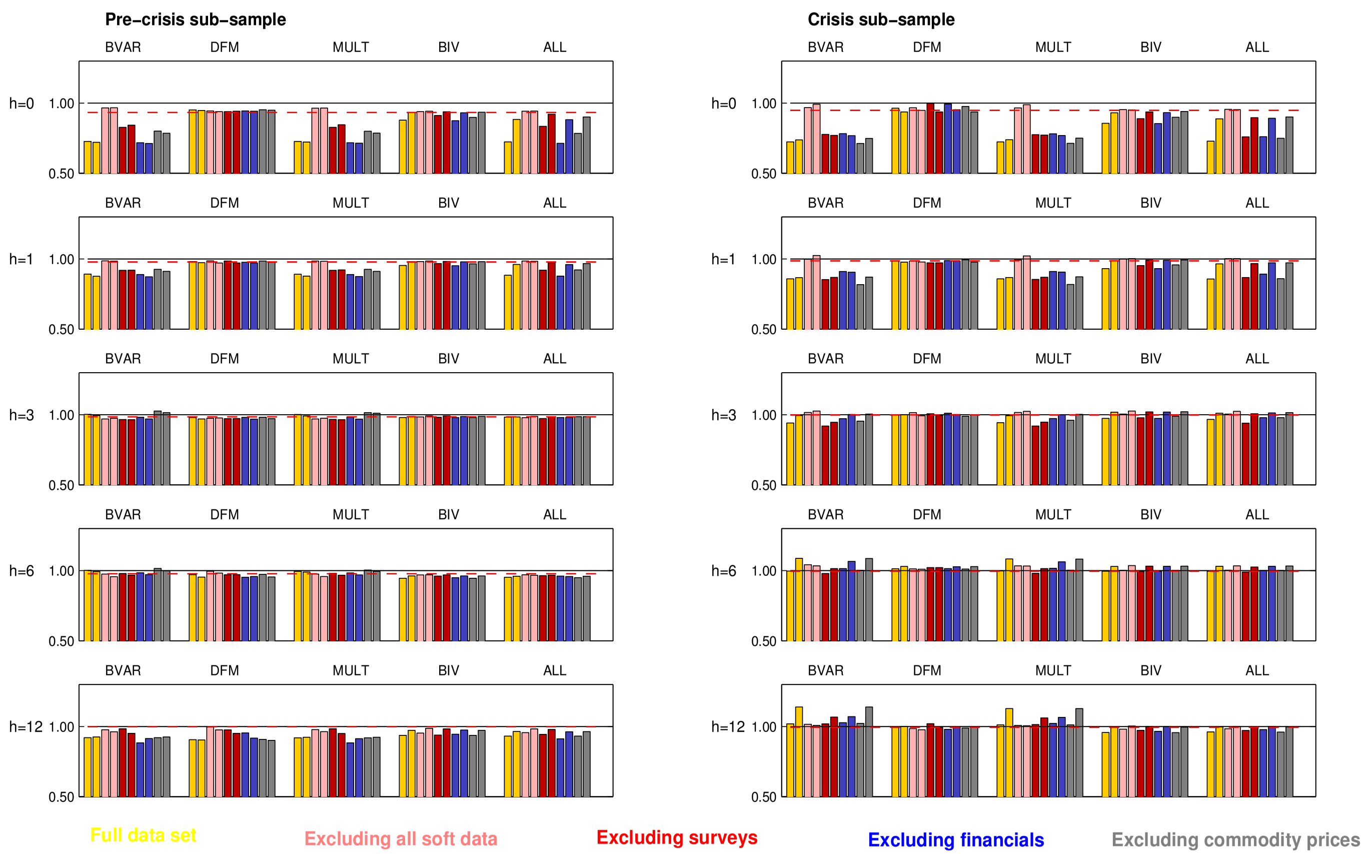
Notes: The left and right parts of the figure shows the relative (versus the RW) RMSFEs for each model over the pre-crisis sample (Dec. 2003 - Nov. 2007) and crisis sample (Dec. 2007 - Dec. 2011) respectively. For each model, the forecasts are computed using five different data sets displayed in different colours: the full data set and the data sets obtained by excluding all the soft data, the surveys, the financials and the commodity prices respectively. In each sub-figure the two consecutive (same coloured) bars for a given model and data set display the results for the forecasts combination schemes used, i.e. av.10% and av.d-msfe. The dashed red line is the relative RMSFEs of the best AR model for the corresponding horizon and the black line is drawn at one (relative RMSFEs of RW).

Figure B.2: Relative RMSFEs of the different panels for the **unemployment rate**



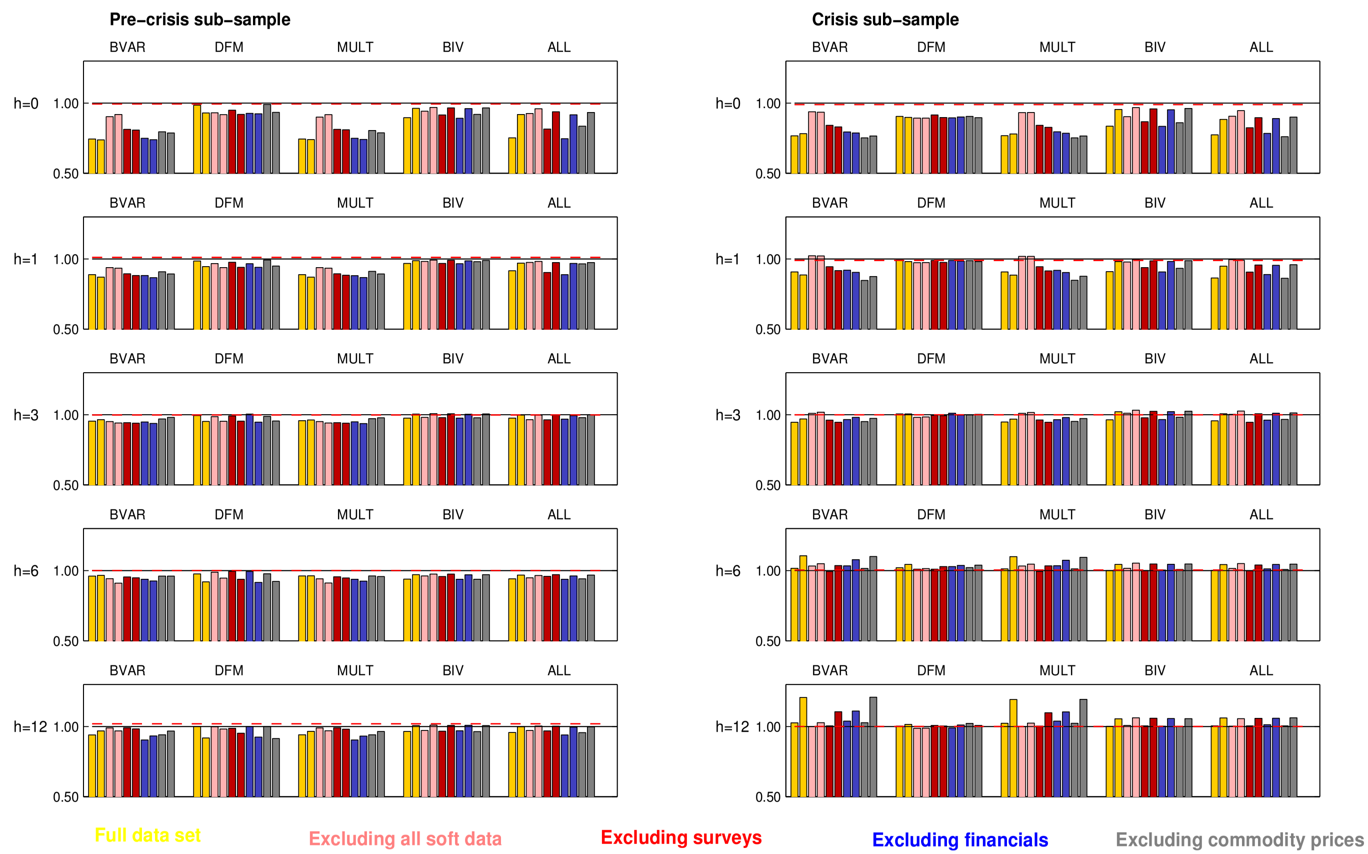
Notes: The left and right parts of the figure shows the relative (versus the RW) RMSFEs for each model over the pre-crisis sample (Dec. 2003 - Nov. 2007) and crisis sample (Dec. 2007 - Dec. 2011) respectively. For each model, the forecasts are computed using five different data sets displayed in different colours: the full data set and the data sets obtained by excluding all the soft data, the surveys, the financials and the commodity prices respectively. In each sub-figure the two consecutive (same coloured) bars for a given model and data set display the results for the forecasts combination schemes used, i.e. av.10% and av.d-msfe. The dashed red line is the relative RMSFEs of the best AR model for the corresponding horizon and the black line is drawn at one (relative RMSFEs of RW).

Figure B.3: Relative RMSFEs of the different panels for the **consumer price index**



Notes: The left and right parts of the figure shows the relative (versus the RW) RMSFEs for each model over the pre-crisis sample (Dec. 2003 - Nov. 2007) and crisis sample (Dec. 2007 - Dec. 2011) respectively. For each model, the forecasts are computed using five different data sets displayed in different colours: the full data set and the data sets obtained by excluding all the soft data, the surveys, the financials and the commodity prices respectively. In each sub-figure the two consecutive (same coloured) bars for a given model and data set display the results for the forecasts combination schemes used, i.e. av.10% and av.d-msfe. The dashed red line is the relative RMSFEs of the best AR model for the corresponding horizon and the black line is drawn at one (relative RMSFEs of RW).

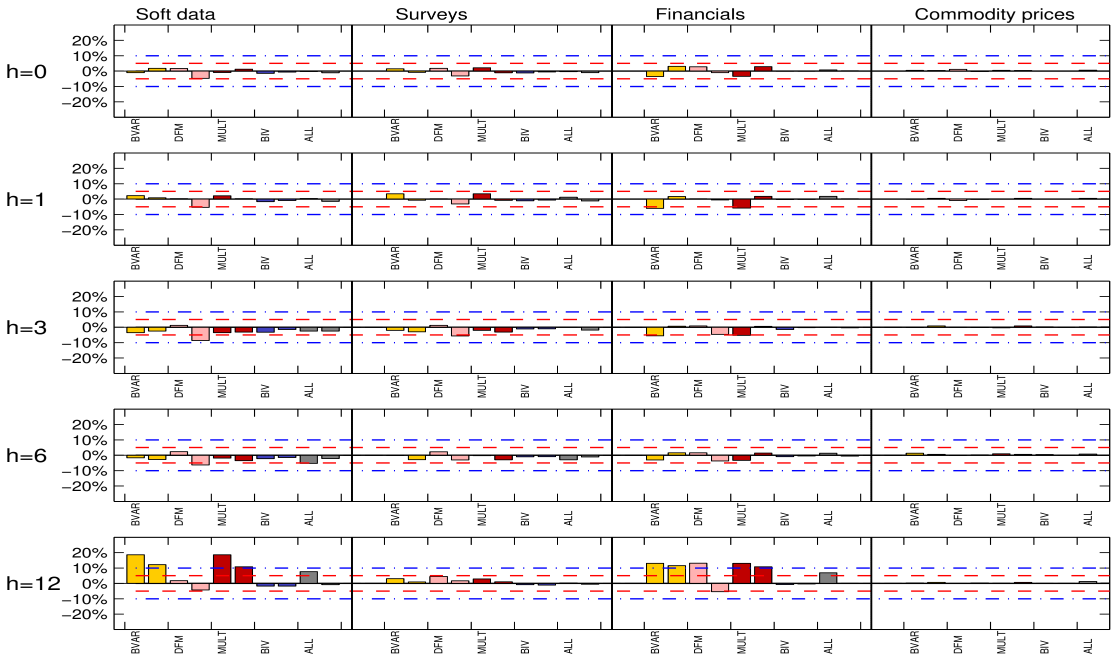
Figure B.4: Relative RMSFEs of the different panels for the **personal consumption expenditures price index**



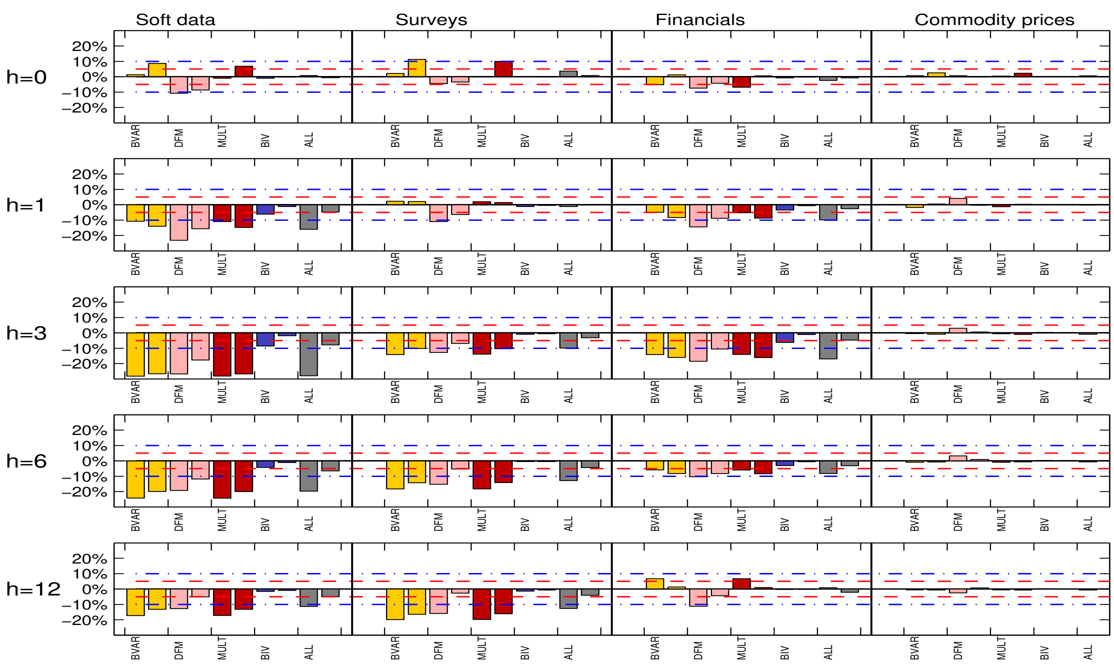
Notes: The left and right parts of the figure shows the relative (versus the RW) RMSFEs for each model over the pre-crisis sample (Dec. 2003 - Nov. 2007) and crisis sample (Dec. 2007 - Dec. 2011) respectively. For each model, the forecasts are computed using five different data sets displayed in different colours: the full data set and the data sets obtained by excluding all the soft data, the surveys, the financials and the commodity prices respectively. In each sub-figure the two consecutive (same coloured) bars for a given model and data set display the results for the forecasts combination schemes used, i.e. av.10% and av.d-msfe. The dashed red line is the relative RMSFEs of the best AR model for the corresponding horizon and the black line is drawn at one (relative RMSFEs of RW).

Figure B.5: Marginal predictive ability of soft data for industrial production

(a) Pre-crisis sample



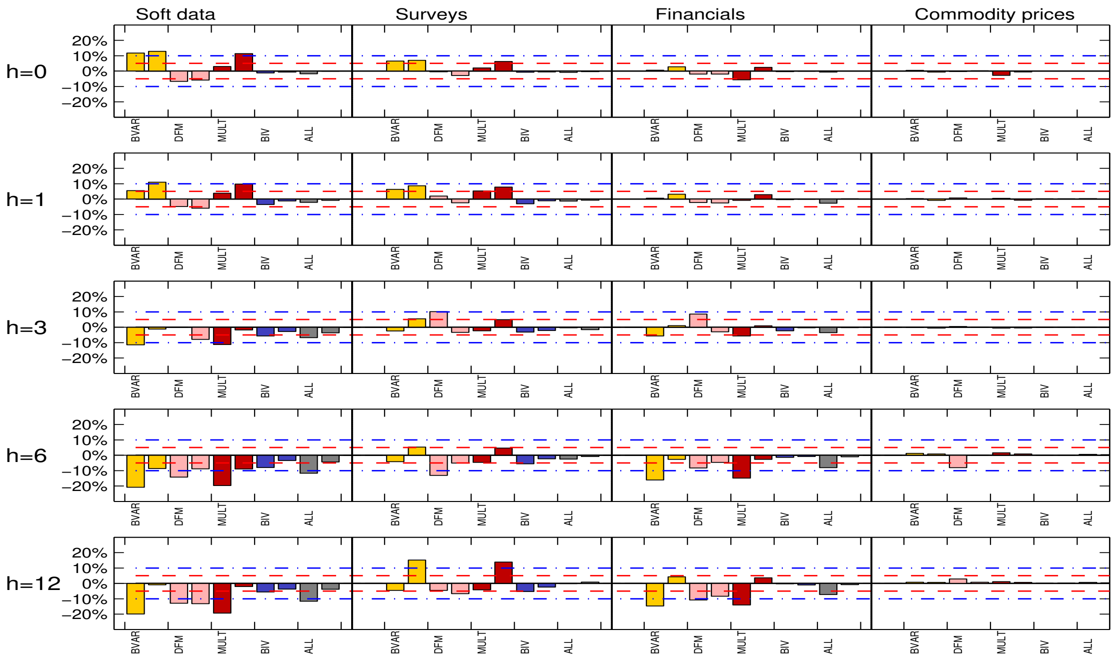
(b) Crisis sample



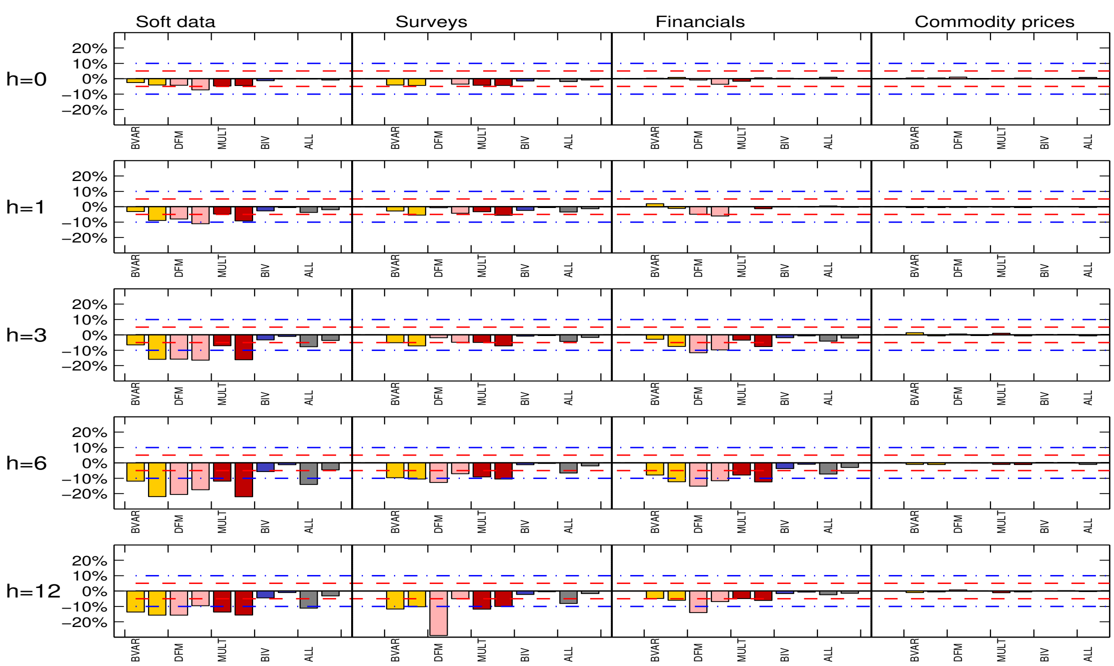
Notes: The figure shows for each model the percentage change in RMSFEs from including the different respective block(s) of soft data. In each sub-figure the two consecutive bars for a given horizon displays the results for the forecasts combination schemes used, i.e. av.10% and av.d-mse. The dashed and dashed dotted lines are drawn at the $+/-5\%$ and $+/-10\%$ threshold respectively.

Figure B.6: Marginal predictive ability of soft data for the unemployment rate

(a) Pre-crisis sample



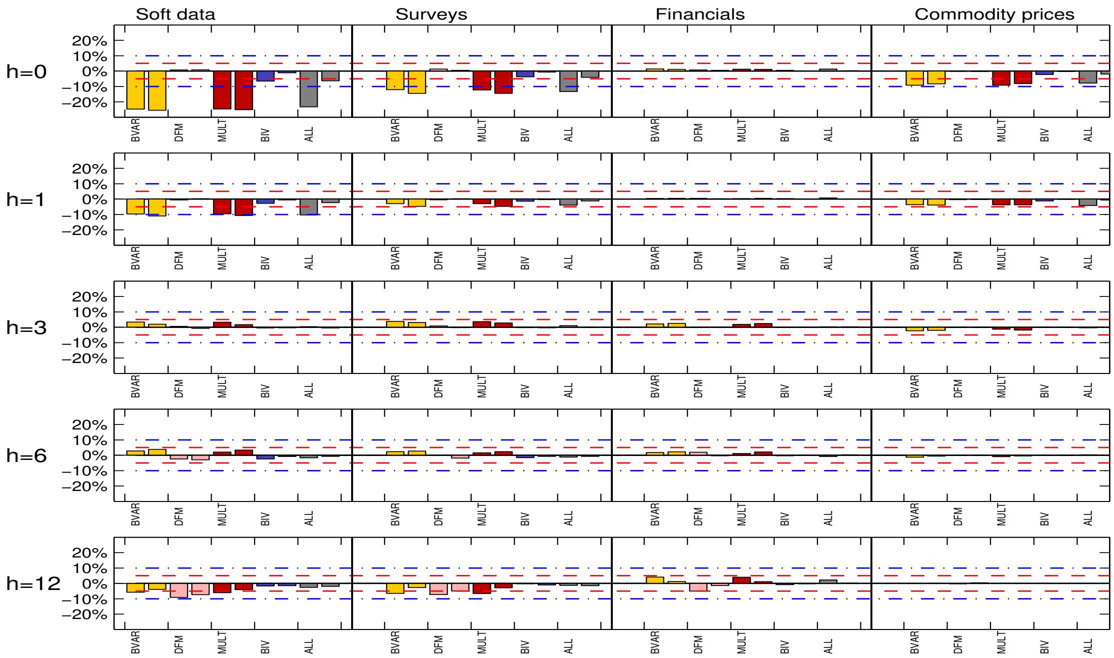
(b) Crisis sample



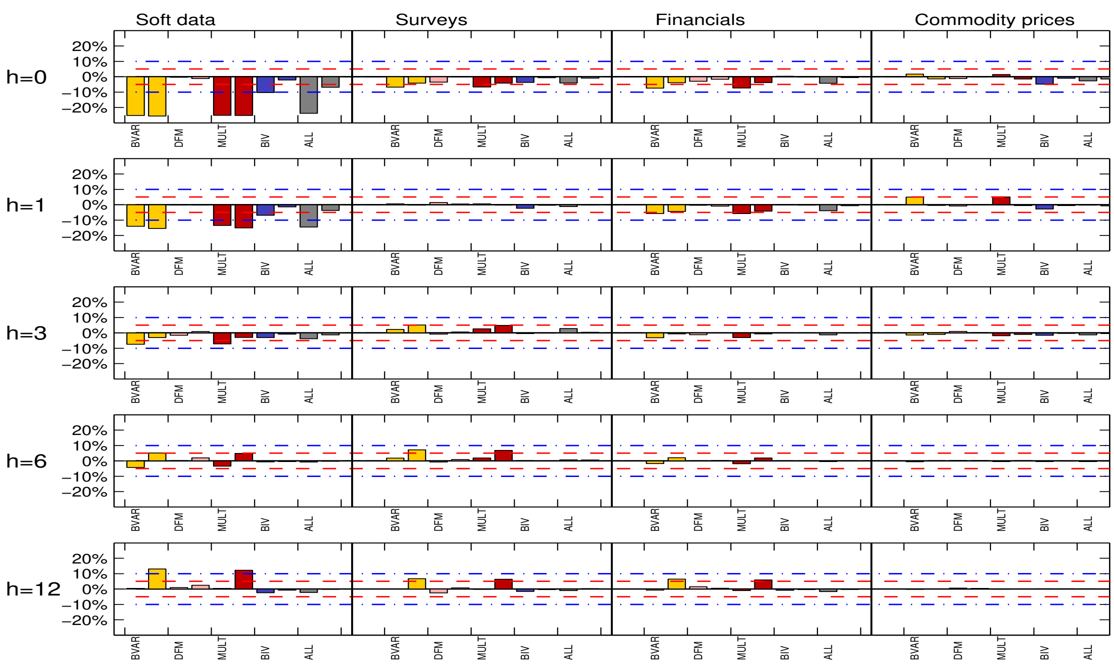
Notes: The figure shows for each model the percentage change in RMSFEs from including the different respective block(s) of soft data. In each sub-figure the two consecutive bars for a given horizon displays the results for the forecasts combination schemes used, i.e. av.10% and av.d-mse. The dashed and dashed dotted lines are drawn at the $+/-5\%$ and $+/-10\%$ threshold respectively.

Figure B.7: Marginal predictive ability of soft data for the consumer price index

(a) Pre-crisis sample



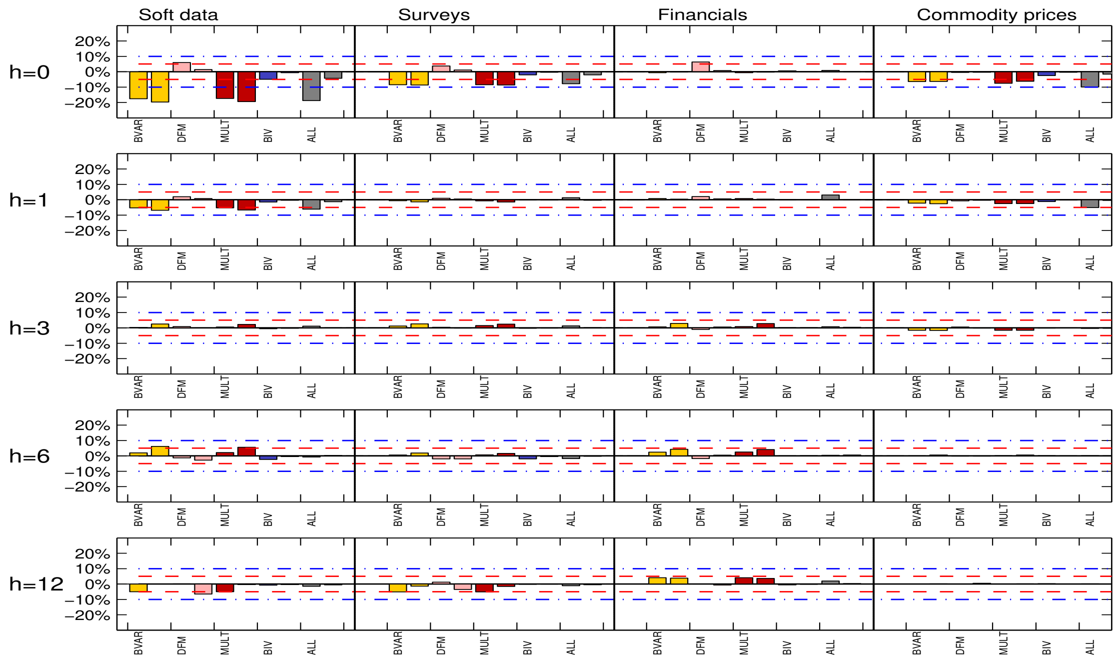
(b) Crisis sample



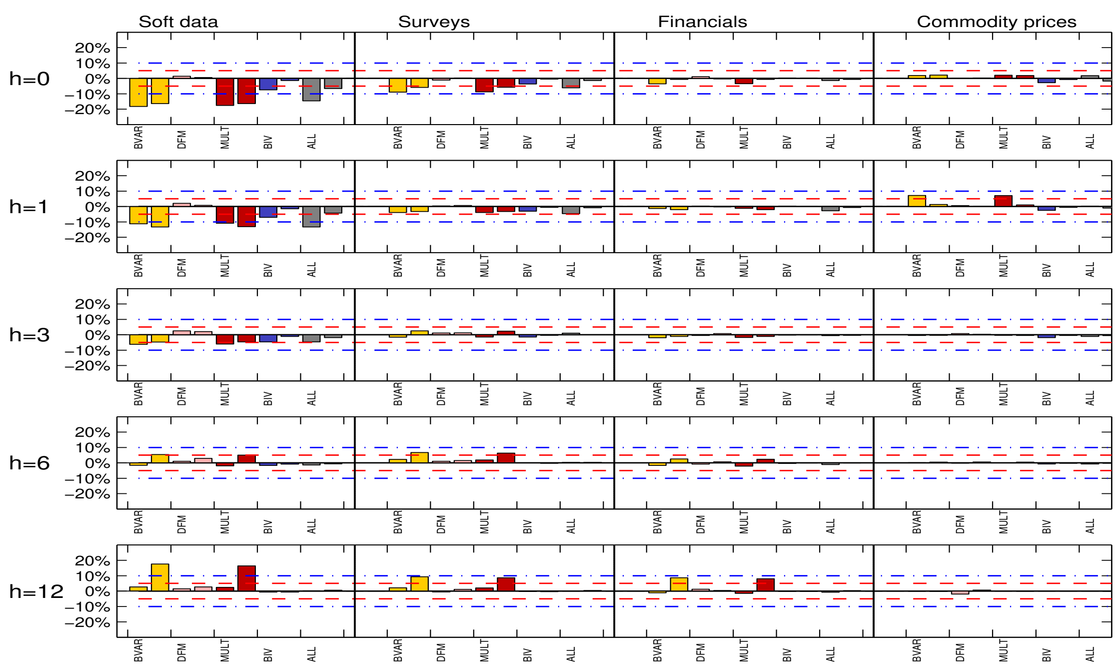
Notes: The figure shows for each model the percentage change in RMSFEs from including the different respective block(s) of soft data. In each sub-figure the two consecutive bars for a given horizon displays the results for the forecasts combination schemes used, i.e. av.10% and av.d-mse. The dashed and dashed dotted lines are drawn at the $+/-5\%$ and $+/-10\%$ threshold respectively.

Figure B.8: Marginal predictive ability of soft data for the pers. consumption expenditures price index

(a) Pre-crisis sample



(b) Crisis sample



Notes: The figure shows for each model the percentage change in RMSFEs from including the different respective block(s) of soft data. In each sub-figure the two consecutive bars for a given horizon displays the results for the forecasts combination schemes used, i.e. av.10% and av.d-mse. The dashed and dashed dotted lines are drawn at the $+/-5\%$ and $+/-10\%$ threshold respectively.