



Department of Econometrics and Business Statistics

<http://business.monash.edu/econometrics-and-business-statistics/research/publications>

# **Macroeconomic forecasting for Australia using a large number of predictors**

Bin Jiang, George Athanasopoulos, Rob J Hyndman,  
Anastasios Panagiotelis, Farshid Vahid

December 2016

Working Paper 17/02

# Macroeconomic forecasting for Australia using a large number of predictors

**Bin Jiang**

Email: svenbinsven@gmail.com

**George Athanasopoulos\***

Email: george.athanasopoulos@monash.edu

**Rob J Hyndman**

Email: rob.hyndman@monash.edu

**Anastasios Panagiotelis**

Email: anastasios.panagiotelis@monash.edu

**Farshid Vahid**

Email: farshid.vahid@monash.edu

\* Correspondence to: George Athanasopoulos, Department of Econometrics and Business Statistics, Monash University, Caulfield East, Victoria 3145, Australia.

**Acknowledgment:** The authors acknowledge the support of Australian Research Council Discovery Project DP 1413220.

14 December 2016

**JEL classification:** C52, C53, C55

**Abstract**

A popular approach to forecasting macroeconomic variables is to utilize a large number of predictors. Several regularization and shrinkage methods can be used to exploit such high-dimensional datasets, and have been shown to improve forecast accuracy for the US economy. To assess whether similar results hold for economies with different characteristics, an Australian dataset containing observations on 151 aggregate and disaggregate economic series is introduced. An extensive empirical study is carried out investigating forecasts at different horizons, using a variety of methods and with information sets containing different numbers of predictors. The results share both differences and similarities with the conclusions from the literature on forecasting US macroeconomic variables. The major difference is that forecasts based on dynamic factor models perform relatively poorly compared to forecasts based on other methods which is the opposite of the conclusion made by Stock and Watson (2012) for the US. On the other hand, a conclusion that can be made for both the Australian and US data is that there is little to no improvement in forecast accuracy when the number of predictors is expanded beyond 20–40 variables.

**Keywords:** Australian economy, Bayesian VAR, bagging, dynamic factor model, ridge regression, least angular regression, shrinkage, regularization.

## 1 Introduction

Forecasts of macroeconomic variables, including key indicators such as GDP and inflation, are necessary inputs for government budget planning, central bank policy making and business decisions. Forming forecasts based on lags of other macroeconomic variables is an approach that dates back at least as far as efforts by Mitchell and Burns (1938) and Burns and Mitchell (1946) to find leading indicators of the business cycle. The use of time series approaches for macroeconomic forecasting gained impetus in the 1970s and 1980s as forecasts from univariate ARIMA models (Box & Jenkins, 1970) and vector autoregressions (VARs) (Sims, 1980) were shown to outperform structural macroeconomic models (for a discussion of this history see Diebold, 1997, and references therein). During this era, the information sets used to form forecasts typically contained only a small number of variables.

This situation changed in the early 2000s as researchers began to compose high-dimensional macroeconomic datasets. Two of the earliest and most widely used examples are the US dataset containing 149 variables measured at a monthly frequency featured in Stock and Watson (2002)

(hereafter referred to as the ‘Stock and Watson data’), and the Euro area dataset containing 447 variables measured at a monthly frequency featured in Forni, Hallin, Lippi and Reichlin (2003). In both studies, utilising a large number of predictors in a dynamic factor modelling (DFM) framework (see Geweke, 1977; Sargent & Sims, 1977) leads to improved forecasts of industrial production relative to standard benchmarks. A key factor in the popularity of this approach is its simplicity; principal components provide consistent estimates of the dynamic factors and can subsequently be used in auxiliary predictive regressions. There is an extensive literature that establishes how the DFM, when used with a large number of predictors, yields good forecasts for macroeconomic variables such as output and inflation for a number of different economies (see Eickmeier & Ziegler, 2008, for an overview).

Despite its success, the DFM is by no means the only framework for forecasting with a large number of predictors. Advances from the statistics and machine learning literature have also been exploited in the macroeconomic context. For example, De Mol, Giannone and Reichlin (2008) consider both ridge regression and the LASSO (see Tibshirani, 1996) for the Stock and Watson data and obtain forecasts that have a similar performance to those obtained from a DFM. Bai and Ng (2008) use least angular regression or LARS (Efron, Hastie, Johnstone & Tibshirani, 2004) to select a set of ‘targeted predictors’. Forecasts are then produced using either these targeted predictors on their own, or alternatively by using the principal components of the targeted predictors. Bai and Ng (2008) show that at least for some periods of the data, methods based on LARS produces better forecasts of CPI, personal income, retail sales, industrial production and total employment compared to the case where principal components are formed using the full information set. Methods that account for model uncertainty such as bootstrap aggregation or ‘bagging’ (see Breiman, 1996; Bühlmann & Yu, 2002; Lee & Yang, 2006) have been successful in forecasting inflation by Inoue and Kilian (2008). Finally, in the class of multivariate forecasting there has been a focus on “big” VARs estimated using Bayesian techniques. Examples include Kadiyala and Karlsson (1997), and more recently Bańbura, Giannone and Reichlin (2010), Carriero, Kapetanios and Marcellino (2011) and Koop (2013) who utilize shrinkage priors including the so-called Minnesota prior of Doan, Litterman and Sims (1984) and Litterman (1986).

Although the literature on macroeconomic forecasting with a large number of predictors is vast, it is possible to make some general conclusions. First, somewhat unsurprisingly, there is no single method that dominates all alternatives for every series at every horizon. Second, the benefit from expanding the information set beyond 20–40 variables is often small, a conclusion

particularly supported by Bai and Ng (2008), Bańbura et al. (2010) and Koop (2013). Third, forecasts based on principal components are highly competitive. In a thorough empirical study Stock and Watson (2012) conclude that *“it will be difficult to improve systematically upon DFM forecasts using time-invariant linear functions of the principal components of large macro datasets like the one considered here.”*

Since these conclusions have been formed on the basis of using only US data, it is worthwhile questioning whether the same results can be found for other economies with vastly different characteristics to the US. Therefore the first major contribution of this paper is to introduce an extensive Australian macroeconomic data set comparable in size to that of the US, comprising 151 quarterly Australian macroeconomic variables which naturally divide into 12 categories of macroeconomic activity. To the best of our knowledge, such a dataset has not been analyzed previously, and since Australia is a small open economy it provides an interesting point of contrast with the ‘default’ Stock and Watson US data.

The second contribution of the paper is to undertake an empirical comparison of the aforementioned approaches, including the dynamic factor model, in the context of forecasting Australian macroeconomic variables. We first focus our attention on forecasting three key variables, namely Gross Domestic Product (GDP), Consumer Price Index (CPI) and the overnight interbank rate (IBR). The IBR is closely related to the ‘cash rate’, the main monetary policy instrument targeted by the Australian central bank. We use the IBR here as the cash rate series only begins in August 1990. The two series are essentially identical over the period for which they are both available. To investigate the value of expanding the number of predictors, we consider information sets of increasing sizes similar to Bańbura et al. (2010) and Koop (2013).

The third major contribution is to compare different forecasting methods for each of the 151 available variables. This stands in contrast to many other papers in the literature which only consider a single variable such as inflation or industrial production. Since it is reasonable to expect that no single method will dominate all others, it is worthwhile investigating whether some methods tend to produce better forecasts for specific categories of macroeconomic variables. To facilitate this analysis, we complement tabulated results with a series of scatter plots which assist in effectively visualising a large amount of information.

The rest of the paper is organized as follows. In Section 2, we provide the details of the Australian macroeconomic data set. Section 3 describes the alternative forecasting approaches we implement in this paper. Section 4 introduces the measures of forecast accuracy we use

and Section 5 gives the main empirical results. We investigate the forecast performance across various forecast techniques considered for different categories of the target economic variables via scatter plots in Section 5.1. Section 6 concludes.

## 2 An Australian Macroeconomic Data Set

The Australian macroeconomic data set compiled for this study comprises 151 variables collected from the Australian Bureau of Statistics (ABS) and the Reserve Bank of Australia (RBA). The series IDs assigned by either the ABS or the RBA are recorded in Table 1. The variables naturally divide into 12 categories shown in Table 2. Each variable consists of 123 quarterly observations spanning the period Q4 1984 to Q2 2015. Variables which are observed at a monthly frequency are aggregated to quarterly by averaging over the 3 months in a quarter (as in Koop, 2013). Each variable is transformed to stationary (similar to Stock & Watson, 2012), with the transformations listed in Table 3.

The complete data set is available from the Australian Macro Database (AMD) at <http://ausmacrodata.org/research.php>. This link provides two files. The first contains the variables as used in this paper. Hence, the file contains all the variables observed over the above mentioned time span and after transformation to stationarity. The second file contains an up-to-date version for each of the raw (untransformed) variables. Each variable is updated automatically when new updates are released by the ABS or the RBA. For more details see Behlul, Panagiotelis, Athanasopoulos, Hyndman and Vahid (2017).

Let  $\mathcal{I}_K$  denote an *information set* containing  $K$  macroeconomic variables with  $|\mathcal{I}_K| = K$ . The full data set contains 151 variables and is denoted  $\mathcal{I}_{151}$ . In this paper, we evaluate the empirical performance of the competing forecasting methods with nested subsets of predictors with different values of  $K$ , specifically  $\mathcal{I}_3 \subset \mathcal{I}_{13} \subset \mathcal{I}_{23} \subset \mathcal{I}_{43} \subset \mathcal{I}_{151}$ . This allows us to investigate the impact of utilizing information sets of differing sizes.

The motivation for considering these nested sets is similar to that of Bańbura et al. (2010) and Koop (2013). Tables 4–8 provide a detailed description of each variable, along with the transformation applied to achieve stationarity, as well as the category to which each variable belongs.

To further illustrate, the smallest of the information subsets  $\mathcal{I}_3$  (see Table 4) contains GDP (the real Gross Domestic Product), CPI (the consumer price index) and IBR (the interbank overnight cash rate equivalent to the Federal funds rate in the US). These are widely considered as three of

the more important variables to forecast in macroeconomic forecasting and US counterparts of these variables have been used in many simple DSGE models (see for example An & Schorfheide, 2007; Bańbura et al., 2010; Christiano, Eichenbaum & Evans, 1999). The information subset  $\mathcal{I}_{13}$  (see Table 5) includes in addition to the three variables in  $\mathcal{I}_3$ , the Australian versions of those variables modelled in: the small VAR of Bańbura et al. (2010) (the total number of persons employed); the medium VAR of Koop (2013) (the industrial production index, private dwelling approvals and the S&P ASX AllOrds stock price index); and the monetary model of Christiano et al. (1999) (the commodity price index of Australia, M1 and total credit). Since Australia, in contrast to the US, is a small open economy we also include in this set, terms of trade, import and export volumes (Dungey & Pagan, 2009).

The set  $\mathcal{I}_{23}$  (see Table 6) contains the remaining 10 variables which are mostly aggregate information, e.g., consumption, labour, money, exchange rates. These variables account for other aspects of the economy not accounted for when  $K = 13$ . In general, the variables chosen in this scenario are analogous to those selected by Koop (2013). Bańbura et al. (2010) and Koop (2013) refer to these as “medium” scale models.

Combining the variables included in the preceding information subsets with an additional 20 variables leads to the information set  $\mathcal{I}_{43}$  (see Table 7). It consists of the majority of aggregate variables in the information set. Koop (2013) refers to these as “medium-large” scale models. Finally, we consider the largest information set  $\mathcal{I}_{151}$  (see Table 8) by adding the remaining 108 variables which are mostly disaggregate variables.

### 3 Forecasting Methods

We investigate a wide range of forecasting methods applied to the Australian macroeconomic data set described in the previous section. These techniques include univariate benchmarks such as a naïve forecast and the AR model, and some more advanced approaches for forecasting high-dimensional data sets such as the dynamic factor model, ridge regression and least angle regression as well as multivariate Bayesian VARs.

Let  $\mathbf{x}_t$  be a  $K$ -vector with elements  $x_{i,t} : i \in \mathcal{I}_K$ , where  $x_{i,t}$  is the value of macroeconomic variable  $i$  at time  $t$ , after it has (i) been transformed to stationarity, (ii) centered by subtracting the mean, and (iii) standardised by dividing by the standard deviation. It is worth noting that in the forecast comparisons that follow we consider the RMSE and the MASE, a scale independent measure due to Hyndman and Koehler (2006). For the RMSE, the comparisons are based on the

variables after transforming to stationary, and standardizing to mean zero and unit variance. For the MASE, the comparisons are performed on the original scale. More details follow in Section 4. Also define  $\mathbf{z}_t := (\mathbf{x}'_t, \mathbf{x}'_{t-1}, \dots, \mathbf{x}'_{t-p+1})'$  as a  $Kp$ -vector containing all the information (including lagged information) available at time  $t$ . Finally define  $y_t$  as the target variable that we are attempting to forecast. Although not necessary in general, in the context of this paper  $y_t$  will always be an element of the predictor set allowing us to write  $y_t = x_{j,t}$ . Since we are potentially interested in forecasting every macroeconomic variable, we will consider  $j = 1, 2, 3, \dots, K$  for all forecasting methods.

All methods that we consider provide forecasts which are linear combinations of the predictors, allowing us to use the following general form

$$\hat{y}_{t+1}^M = \mathbf{x}'_t \hat{\boldsymbol{\theta}}_1^M + \mathbf{x}'_{t-1} \hat{\boldsymbol{\theta}}_2^M + \dots + \mathbf{x}'_{t-p+1} \hat{\boldsymbol{\theta}}_p^M \quad (3.1)$$

where  $\hat{y}_{t+1}^M$  is an one-step-ahead forecast of the target variable using forecasting method  $M$  and  $\hat{\boldsymbol{\theta}}_\ell^M := (\hat{\theta}_{1,\ell}, \hat{\theta}_{2,\ell}, \dots, \hat{\theta}_{K,\ell})'$  where  $\hat{\theta}_{i,\ell}$  is the weight placed on the  $\ell$ th lag of the  $i$ th variable in information set  $\mathcal{I}_K$ . This can also be expressed in terms of the stacked form as

$$\hat{y}_{t+1}^M = \mathbf{z}'_t \hat{\boldsymbol{\theta}}^M \quad (3.2)$$

where  $\hat{\boldsymbol{\theta}}^M := (\hat{\boldsymbol{\theta}}_1^{M'}, \hat{\boldsymbol{\theta}}_2^{M'}, \dots, \hat{\boldsymbol{\theta}}_p^{M'})'$  is a  $Kp$ -vector. Note that Equations (3.1) and (3.2) specify a one-step-ahead forecast. For the univariate forecasting procedures that follow multiple-steps ahead forecasts are generated recursively by iterating the estimated model. The one exception is the dynamic factor model which implements direct forecasting. We discuss this in Section 3.2.

### 3.1 Benchmarks

Two benchmark models are used to facilitate the empirical evaluation undertaken in this paper. The first benchmark we consider is the sample mean which in this setting forms a natural naïve benchmark. Since the data are mean corrected,  $\hat{y}_{t+1}^{\text{naive}} = 0$  implying  $\hat{\boldsymbol{\theta}}^{\text{naive}} = 0$ . It is worth noting that for macroeconomic variables of interest such as GDP and the interbank overnight cash rate which usually require first-differencing for stationarity, the sample mean forecast is equivalent to assuming a random walk model with drift for the original variable.



The second benchmark we consider is the standard AR( $p$ ) model. Recalling that  $y_t = x_{j,t}$ , in the general framework of Equation (3.2),

$$\begin{aligned}\hat{\theta}_{i,\ell}^{AR} &\neq 0 & \text{if } i = j \\ \hat{\theta}_{i,\ell}^{AR} &= 0 & \text{if } i \neq j.\end{aligned}$$

The non-zero weights are found as estimates of an AR( $p$ ) model. It is worth noting that AR( $p$ ) forecasts only utilize the information in the target series and therefore form a natural benchmark against the univariate forecast procedures which extract information from a large number of predictors.

### 3.2 Dynamic Factor Model (DFM)

The dynamic factor model assumes that  $r_D$  unobserved dynamic factors can summarize the information set of the predictors  $\mathbf{x}_t$  where  $r_D \ll K$ . More precisely we assume that  $\mathbf{x}_t$  admits the approximate factor structure

$$\mathbf{x}_t = \mathbf{\Lambda} \mathbf{f}_t + \boldsymbol{\xi}_t$$

in which  $\mathbf{\Lambda}$  is a  $K \times r_D$  matrix of factor loadings,  $\mathbf{f}_t$  are the  $r_D$  unobserved factors and  $\boldsymbol{\xi}_t$  is a vector of idiosyncratic errors which can be weakly inter-correlated (e.g., Fan, Liao & Mincheva, 2013). Forni, Hallin, Lippi and Reichlin (2000) and Stock and Watson (2002) show that the first  $r_D$  principal components of the data can consistently estimate the  $r_D$  unobserved factors under the assumptions of the DFM. If  $\mathbf{W}$  is a  $K \times r_D$  matrix with columns given by the first  $r_D$  eigenvectors of  $S_X = (T - 1)^{-1} \sum_{t=1}^T \mathbf{x}_t \mathbf{x}_t'$ , then the factors at time  $t$  are estimated by

$$\hat{\mathbf{f}}_t = \mathbf{W}' \mathbf{x}_t \tag{3.3}$$

To operationalize this framework as a forecasting model for  $h$ -step-ahead forecasting,  $y_{t+h}$  is regressed upon  $\hat{\mathbf{f}}_t, \hat{\mathbf{f}}_{t-1}, \dots, \hat{\mathbf{f}}_{t-p+1}$ . The  $h$ -step-ahead forecasts from the dynamic factor model are derived directly as

$$\hat{y}_{t+h}^{\text{DFM}} = \hat{\mathbf{f}}_t' \hat{\boldsymbol{\delta}}_1 + \hat{\mathbf{f}}_{t-1}' \hat{\boldsymbol{\delta}}_2 + \dots + \hat{\mathbf{f}}_{t-p+1}' \hat{\boldsymbol{\delta}}_p, \tag{3.4}$$

where  $\hat{\boldsymbol{\delta}}_\ell$  is an  $r_D$ -vector of coefficients estimated by least squares, that corresponds to the  $\ell$ th lag of the factors in the auxiliary regression. Despite the use of a two-step procedure, the forecast

is still a linear function of the predictors. Substituting Equation (3.3) into Equation (3.4) and using the notation of Equation (3.1) yields

$$\hat{\theta}_\ell^{\text{DFM}} = \hat{W} \hat{\delta}_\ell.$$

### 3.3 Ridge Regression

Ridge regression is a method that shrinks the estimates of regression coefficients towards zero relative to the least squares solution. It can be motivated by adding a penalty on the  $L_2$  norm of the coefficients to the usual sum of squared errors objective function, or in a Bayesian context by assuming a conjugate Gaussian prior on the coefficients. In our context, ridge regression implies that the forecasts weights have the following closed form

$$\hat{\theta}^{\text{ridge}} = \left( \sum_{t=p}^T \mathbf{z}_{t-1} \mathbf{z}'_{t-1} + r_R \mathbf{I}_{Kp} \right)^{-1} \left( \sum_{t=p}^T \mathbf{z}_{t-1} y_t \right),$$

where  $r_R \geq 0$  is a parameter that controls the amount of shrinkage on the ridge coefficients and  $\mathbf{I}_{Kp}$  is a  $Kp \times Kp$  identity matrix.

It is worth noting some similarities between the dynamic factor modelling approach, which is effectively a principal components regression, and ridge regression. Principal components analysis considers a rotated version of the predictor space and shrinks the coefficients corresponding to smaller eigenvalues to be exactly zero. When ridge regression is considered in this rotated space, coefficients are shrunk by a factor of  $d_i^2 / (d_i^2 + r_R)$  where  $d_i$  is the  $i$ th largest eigenvalue of  $S_Z = (T - p - 1)^{-1} \sum_{t=p}^T \mathbf{z}_{t-1} \mathbf{z}'_{t-1}$ . Since the  $d_i^2 / (d_i^2 + r_R)$  is close to zero for small  $d_i$ , ridge regression shrinks the coefficients corresponding to smaller eigenvalues towards zero to a greater degree than coefficients corresponding to large eigenvalues, but in contrast to principal components regression, ridge regression does not shrink any of these coefficients exactly to zero.

### 3.4 Least Angle Regression (LARS)

Least angle regression is a shrinkage and variable selection method proposed by Efron et al. (2004) and is motivated by the forward selection and the forward stagewise methods for selecting regression models (e.g., Hastie, Tibshirani & Friedman, 2009). A comprehensive formulation of the technical details underlying the LARS algorithm can be found in Efron et al. (2004), but we summarise the main features here. The LARS algorithm is initialised at a null model that

includes no predictors. In our context, the dependent variable is  $y_t$  and the set of potential predictors is  $z_{t-1}$ , although in the following discussion, the time index  $t$  and  $t-1$  will be dropped for convenience. Let  $\mathcal{S}_k$  be the set of predictors included at step  $k$  of the LARS algorithm. At step  $k$  of the algorithm, the coefficients of the predictors are updated in an equiangular fashion towards the least squares solution of a regression of  $y$  on  $z_i : z_i \in \mathcal{S}_k$ , with the residual given by  $e_k$ . This proceeds until the correlations between  $e_k$  and the  $z_i : z_i \in \mathcal{S}_k$  are equal to the correlation between  $e_k$  and some  $z_{i'} \notin \mathcal{S}_k$ . The algorithm proceeds to step  $k+1$ , we define  $\mathcal{S}_{k+1} = \{\mathcal{S}_k, z_{i'}\}$  and the coefficients are now updated in the direction of the least squares solution of a regression of  $y$  on  $z_i \in \mathcal{S}_{k+1}$ .

It is worth noting that the selection process of the LARS algorithm actually delivers a ranking of the predictors. This is a great advantage when forecasting with a handful of economic variables as practitioners can get a feel for which variables are more important than the others. Moreover, at no further computational cost, the LARS algorithm can be used to implement forward stagewise selection and LASSO, the latter of which penalizes the  $L_1$  norm of the coefficient vector.

The total level of shrinkage is controlled by the number of variables selected which we denote as  $r_L$ . This implies that for LARS the forecasts in Equation (3.2) are given by a sparse linear combination of the predictors where,

$$\begin{aligned}\hat{\theta}_i^{\text{LARS}} &\neq 0 && \forall i : z_{i,t-1} \in \mathcal{S}_{r_L} \\ \hat{\theta}_i^{\text{LARS}} &= 0 && \forall i : z_{i,t-1} \notin \mathcal{S}_{r_L}.\end{aligned}$$

### 3.5 Bagging LARS

Bootstrap aggregation or bagging (e.g., Breiman, 1996) is a prevalent device for improving the accuracy and stability of statistical learning algorithms. Rather than training the parameters once, bootstrap samples are taken, the parameters are trained on each bootstrap sample and are then averaged.

In macroeconomic forecasting, bagging has been used in conjunction with pretest methods for variable selection whereby variables are included only if their corresponding t-statistics exceed some hard threshold (see Inoue & Kilian, 2008; Stock & Watson, 2012, for more details). Rather than follow Inoue and Kilian (2008) and Stock and Watson (2012) we use bagging in conjunction with LARS rather than the pretest method for three main reasons. First, for  $K = 151$ ,

the number of predictors is larger than the number of observations, a scenario that cannot be handled by the pretest methods but can be handled by LARS. Second, pretest methods can select predictors that are highly correlated with one another, which as Bai and Ng (2008) point out, can be avoided by LARS. Selecting highly correlated predictors reduces the benefit from bagging since model averaging is most effective when the predictors carry distinct information. Third, bagging with LARS has a computational cost of the same order as ordinary least squares. In the empirical evaluation that follows we refer to forecasts generated from bagging LARS as “BagL” or “Bag-LARS”. If we let  $\hat{\theta}_b^{\text{LARS}}$  be the coefficients obtained in the manner described in Section 3.4 but for bootstrap sample  $b$  then by the linearity of our forecasts

$$\hat{\theta}^{\text{BagL}} = \sum_{b=1}^B \hat{\theta}_b^{\text{LARS}},$$

where  $B$  is the total number of bootstrap samples.

### 3.6 A Bayesian VAR

The multivariate forecast method we focus on is a Bayesian VAR (see for example, Bańbura et al., 2010; Carriero et al., 2011; Koop, 2013). Unlike the conventional VARs that only include a small number of variables (typically less than ten), such Bayesian VARs can allow for quite a large number of predictors. For instance, the BVAR of Bańbura et al. (2010) contains 131 predictors. In the previous sections we have let  $\mathbf{x}_t$  be the  $K$ -dimensional vector the elements of which are the macroeconomic variables in the information set  $\mathcal{I}_K$ . Thus, a multivariate VAR( $p$ ) can be expressed as

$$\mathbf{x}_{t+1} = \mathbf{A}_1 \mathbf{x}_t + \mathbf{A}_2 \mathbf{x}_{t-1} + \cdots + \mathbf{A}_p \mathbf{x}_{t-p+1} + \mathbf{e}_{t+1}, \quad (3.5)$$

where  $\mathbf{e}_{t+1}$  is a vector of errors distributed independently  $N(0, \Sigma)$  across  $t$  and  $\{\mathbf{A}_i; i = 1, 2, \dots, p\}$  are  $K \times K$  autoregressive coefficient matrices. As in the previous sections trends and means are assumed to have been properly removed.

Recalling  $\mathbf{z}_t = (\mathbf{x}'_t, \mathbf{x}'_{t-1}, \dots, \mathbf{x}'_{t-p+1})'$  and denoting  $\mathbf{A} = (\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_p)'$ , the model (3.5) can be alternatively expressed as

$$\underset{(T^* \times K)}{\mathbf{X}} = \underset{(T^* \times Kp)}{\mathbf{Z}} \underset{(Kp \times K)}{\mathbf{A}} + \underset{(T^* \times K)}{\boldsymbol{\mathcal{E}}},$$

where  $\mathbf{X} = (x_{p+1}, \dots, x_T)'$ ,  $\mathbf{Z} = (z_p, \dots, z_{T-1})'$  and  $\boldsymbol{\mathcal{E}} = (e_{p+1}, \dots, e_T)'$  such that  $\text{vec}(\boldsymbol{\mathcal{E}}) \sim N(0, \boldsymbol{\Sigma} \otimes \mathbf{I}_{T^*})$ .

We employ a normal-inverted Wishart prior proposed by Kadiyala and Karlsson (1997)

$$\text{vec}(\mathbf{A}) | \boldsymbol{\Sigma} \sim N(\text{vec}(\mathbf{A}_0), \boldsymbol{\Sigma} \otimes \boldsymbol{\Omega}_0), \quad \boldsymbol{\Sigma} \sim iW(\nu_0, \mathbf{S}_0), \quad (3.6)$$

where the parameters  $\nu_0, \mathbf{S}_0, \mathbf{A}_0, \boldsymbol{\Omega}_0$  are hyper-parameters set as follows. The values of  $\mathbf{A}_0$  and  $\boldsymbol{\Omega}_0$  are set according to the Minnesota prior (Litterman, 1986):

$$\mathbb{E}[A_\ell^{(i,j)}] = \begin{cases} 0, & i = j, \ell = 1; \\ 0 & \text{otherwise;} \end{cases} \quad \text{Var}[A_\ell^{(i,j)}] = \begin{cases} \frac{r_V^2}{\ell^2}, & i = j; \\ \frac{r_V^2}{\ell^2} \frac{\sigma_i^2}{\sigma_j^2} & \text{otherwise;} \end{cases}$$

where  $A_\ell^{(i,j)}$  is element  $(i, j)$  of the coefficient matrix  $A_\ell$ . Our formulation here differs slightly from the original paper by Litterman (1986) who shrinks all non-stationary variables to a random walk by setting  $\mathbb{E}[A_\ell^{(i,j)}] = 1$  for all  $i$  corresponding to non-stationary variables. Since we transform all variables to be stationarity, this is not relevant in our context. The shrinkage on the coefficients is determined by the hyper-parameter  $r_V$ . The values of  $\nu_0$  and  $\mathbf{S}_0$  are chosen such that the prior expectation of  $\boldsymbol{\Sigma}$  is the fixed residual covariance matrix  $\text{diag}\{\sigma_1^2, \sigma_2^2, \dots, \sigma_K^2\}$ . This a priori expected value of  $\boldsymbol{\Sigma}$  is the same as for the Minnesota prior, however in contrast to the Minnesota prior, there is some prior support for non-zero correlations between error terms under the normal-inverted Wishart prior.

The prior given in (3.6) together with the likelihood given in 3.5 lead to the conditional posteriors which are also normal-inverted Wishart,

$$\text{vec}(\mathbf{A}) | \boldsymbol{\Sigma}, \mathbf{X}, \mathbf{Z} \sim N(\text{vec}(\bar{\mathbf{A}}), \boldsymbol{\Sigma} \otimes \bar{\boldsymbol{\Omega}}), \quad \boldsymbol{\Sigma} | \mathbf{X}, \mathbf{Z} \sim iW(\bar{\nu}, \bar{\mathbf{S}})$$

where  $\bar{\mathbf{A}} = (\boldsymbol{\Omega}_0^{-1} + \mathbf{Z}'\mathbf{Z})^{-1}(\boldsymbol{\Omega}_0^{-1}\mathbf{A}_0 + \mathbf{Z}'\mathbf{X})$ ,  $\bar{\boldsymbol{\Omega}} = (\boldsymbol{\Omega}_0^{-1} + \mathbf{Z}'\mathbf{Z})^{-1}$ . Due to conjugacy,  $\boldsymbol{\Sigma}$  can be integrated out analytically leading to a multivariate  $t$  distribution for  $\text{vec}(\mathbf{A}) | \mathbf{X}$  with posterior expectation  $\text{vec}(\bar{\mathbf{A}}) = \text{vec}\left((\boldsymbol{\Omega}_0^{-1} + \mathbf{Z}'\mathbf{Z})^{-1}(\boldsymbol{\Omega}_0^{-1}\mathbf{A}_0 + \mathbf{Z}'\mathbf{X})\right)$  (Zellner, 1971).

The Bayesian VAR also generates forecasts that are linear combinations of the predictors. To express this in the form of Equation 3.2 we isolate a single element of  $\mathbf{x}_t$ , namely  $y_t = x_{j,t}$ , and

the weights on the predictors for a one-step-ahead forecast are given by

$$\hat{\theta}^{\text{BVAR}} = \bar{a}_j,$$

where  $\bar{a}_j$  denotes the  $j$ th column of  $\bar{A}$

### 3.6.1 Setting Regularisation Parameters

The levels of regularization in the dynamic factor modelling approach, the ridge regression, LARS and Bayesian VAR depend on a parameter denoted  $r_D$ ,  $r_R$  and  $r_L$  and  $r_B$  respectively. In the case of the dynamic factor model we set  $r_D = 5$  a rather ad-hoc choice which nevertheless is widely used in the literature (e.g. Stock & Watson, 2012). As a robustness check, we also used  $r_D = 3$  and  $r_D = 8$  which had minimal impact on our final conclusions. For ridge and LARS we use cross-validation, a choice supported by the conclusion in Bergmeir, Hyndman and Koo (2015) that cross validation is valid when statistical learning methods are applied to purely autoregressive models. The default setting of 10-fold cross validation was used in the R-packages, `glmnet` (version 2.0.5) and `lars` (version 1.2) for estimating ridge and LARS respectively. As an alternative, for LARS we also implemented the  $C_p$ -type selection criterion proposed by Efron et al. (2004). However, its performance was inferior to that of cross validation and therefore we do not report the results. Finally, for selecting the shrinkage parameter of the Bayesian VAR we follow Bańbura et al. (2010) and Koop (2013). We note however that the three key macroeconomic variables we target namely the real Gross Domestic Product, consumer price index and interbank overnight cash rate in Australia are different to the ones used in the aforementioned studies.

## 4 Forecast Evaluation

We consider  $h = 1$  to 4-steps ahead forecasts for each of the five information sets described in Section 2. We should reemphasise that all the estimation and calculations that follow are based on the variables after transformation to stationarity. The forecast evaluation begins using a training window of 10 years, i.e., 40 observations. Each model is estimated within this window from which  $h = 1$  to 4-steps-ahead forecasts are generated. The window is then rolled forward one quarter at a time until the end of the sample (similar to Stock & Watson, 2012). Sample means, sample variances and models are re-estimated and forecasts are generated with each step. This results in  $75 - h$  out-of-sample forecasts for each forecast horizon  $h$ , which are used to

evaluate the forecast performances of the competing models. All forecast approaches make use of 4 lags of the predictors which includes 4 lags of the dependent variable as well.

We consider two measures of forecasting accuracy. First the RMSE (root mean squared error). As mentioned in Section 3, each variable is standardised using its sample mean and variance before the methods are applied. The RMSE is a scale dependant measure. Therefore we calculate the RMSE for the forecasts generated from each method based on the standardised variables.

As an alternative, we also consider the MASE (mean absolute scaled error) proposed by Hyndman and Koehler (2006). For an  $h$ -step-ahead out-of-sample forecast, MASE is defined as

$$\text{MASE} = \frac{1}{J} \sum_{j=1}^J |q_j|,$$

where  $J$  is the number of  $h$ -step-ahead forecasts available for evaluation and  $q_j$  is the following scaled error

$$q_j = \frac{e_j}{\frac{1}{T-1} \sum_{t=2}^T |y_t - y_{t-1}|} \quad (4.1)$$

where the numerator is simply the out-of-sample forecast error based on a sample of  $T$  observations and the denominator is the in-sample mean absolute error from the one-step naïve random walk forecast. Equation (4.1) implies that MASE is a scale independent measure and therefore comparing forecast accuracy across macroeconomic time series measured in distinct units is valid. Hence, we calculate the MASE across the forecasts generated for the transformed to stationary non-standardised series, i.e., after returning to the forecasts the in-sample mean and standard deviation.

## 5 Main Results

Table 9 presents the forecast accuracy for  $h = 1$  to 4-steps ahead of the alternative approaches across the three key macroeconomic variables of interest: GDP, CPI and IBR. Each entry shows the RMSE of the forecast approach relative to the naïve benchmark.  $K$  denotes the number of variables included in the information set. The entries in bold show the minimum RMSE achieved by each alternative approach using information sets of varying sizes. Hence, our primary interest here was to first and foremost assess the value added to forecasting key Australian macroeconomic variables by increasing the size of the information set. The results

based on the MASE are presented in Table 10. The results are qualitatively similar for both measures and therefore the discussion that follows refers to both sets of results.

Most entries in Table 9 are close to 1. This indicates that in general the competing forecasting approaches do not forecast the three target variables more accurately than the naïve approach. In most cases at best the competing methods match the naïve benchmark or only slightly outperform it. Hence it seems that in particular with Australian GDP using a random walk with drift is sufficient.

There are some other observations worth noting. With the exception of the DFM5 all competing methods perform considerably better than the naïve benchmark in forecasting the IBR for at least  $h = 1$  and for some cases 2-steps ahead. Interestingly the AR(4) benchmark performs remarkably well for  $h = 1$ -step ahead for both CPI and IBR. For both these variables at least for very short forecasting, information from the additional predictors does not seem to assist. Using only their own past information suffices.

In contrast to the general perception from forecasting the US economy, the dynamic factor model performs very poorly with the Australian dataset independent of the size of the information set. The DFM5 is considerably less accurate than the naïve benchmark and the AR(4). In most cases it seems that the best performance for the dynamic factor model comes from using the smallest possible information set of  $K = 3$ , hence using the DFM3. As we have previously mentioned, we experimented with also using (for all information sets) the DFM3 and the DFM8. The results were qualitatively similar to the DFM5. We should acknowledge that these results can possibly be improved by estimating the unobserved factors using principal components of targeted predictors, as suggested by Bai and Ng (2008), which are often selected by either hard or soft thresholding algorithms such as pretest and LARS.

In an effort to investigate the effect of the information set sizes, we discover some interesting findings. The results for Ridge show that there are considerable improvements in accuracy going from  $K = 3$  to  $K = 13$ , with minor improvements after that. For LARS there does not seem to be any differences stemming from the information set used. As expected, bagging LARS shows some improvements (although minor) in many cases. There is a substantial improvement in the accuracy of the BVAR by increasing the information set from  $K = 3$  to  $K = 13$ , and some marginal improvements after that.

Table 11 shows the average error measures relative to the naïve benchmark from forecasting all  $K$  variables within the different information sets. Hence it is only appropriate to compare



results within each panel where the number of variables forecast is  $K$ . The results send some very clear messages. LARS and bagging LARS are the best performing methods when using an information set of only  $K = 3$  predictors, the variables we consider as the ones of main interest. As the information set increases beyond these three variables, Ridge and BVAR catch up and provide qualitatively very similar forecasting performance to the LARS results. It is only for very few cases, for  $K \geq 13$  and only for  $h = 1$ , that these four methods show any improvement over the naïve benchmark; however this improvement is not substantial. The dynamic factor model is, in all cases, the least accurate.

In summary, the main results we obtain with Australian macroeconomic data are as follows. First, the utilization of additional information can improve the forecast accuracy of small models. Second, the results provide evidence from an Australian perspective, supporting the finding of Bańbura et al. (2010) and Koop (2013), that conjunctural information (the majority of which are disaggregate variables) beyond the 23 important aggregate macroeconomic variables only leads to either minor or no improvements in forecasting key macroeconomic variables. For Australia, using  $K = 13$  predictors seems to be enough. This result is not just for the Bayesian VAR but also observed with the other approaches. Third, the forecast performance of the naïve benchmark becomes extremely challenging to outperform, especially for  $h \geq 2$ , both for overall and for specific key macroeconomic variables. This is particularly the case for forecasting Australian GDP, where using a random walk with drift model can yield high quality forecasts at all quarters. Fourth, simple implementation of the dynamic factor model does not perform well for forecasting the Australian economy. Fifth, employing either multivariate or univariate models seems not to be practically important for forecasting with many predictors for Australia. Finally, for one-step-ahead forecasts of CPI, it is hard to beat the AR(4) benchmark.

### 5.1 A Visualization Assisted Evaluation

It is conventional in the literature to report the forecast accuracy for only a few key macroeconomic variables, or at best summary measures over a group of variables, in order to evaluate the forecasting performance of alternative approaches. This is what we do with Tables 9–11. However, there is a plethora of information that hides behind the individual variables used in the information sets. This is what we attempt to explore in this section. In particular, we focus on forecasting the variables within the information sets of  $K = 43$  (i.e., the majority of aggregate variables) and  $K = 151$  (i.e., all variables we have collected for the Australian economy). To achieve this by tabulating results would be impossible, as this would require tables of very large

dimensions. We suggest comparing the forecast results via scatter plots, which can make these comparisons feasible in a very simple way. What we attempt to introduce is the idea of “data visualization” (see Lindquist, 2011 for a recent survey) to macroeconomic forecasting.

Figures 1 and 2 compare RMSE and MASE values for each variable, for  $h = 1$  to 4-steps ahead, for the various forecasting approaches and the naïve benchmark, when the information set is  $K = 43$ . The colors represent the 12 categories to which the variables belong, as shown in Table 2. The shapes identify the competing forecast approaches implemented; i.e., DFM5, Ridge, LARS, Bag-LARS, BVAR and AR(4). Figures 3 to 10 compare these for all  $K = 151$  variables with each row of plots showing the results within each category of variables. For each point, the value on the vertical axis denotes the percentage deviation of the error measure of the competing approach relative to the error measure of the naïve benchmark for that particular variable. The corresponding value on the horizontal axis denotes the error measure for the naïve benchmark. Hence a point below the horizontal axis shows the percentage reduction in the error measure achieved by the forecasting approach relative to the naïve benchmark. The further to the right the point, the larger the error measure for the naïve benchmark; i.e., the harder the variable is to forecast.

The first clear observation from Figures 1 and 2 is that most points related to the DFM5 and the AR(4) benchmark (naughts and crosses) lie above the horizontal axis, with some points showing more than 25% increase in RMSE compared to the naïve benchmark. Hence these two are clearly outperformed by the naïve benchmark in most cases. Second, for  $h = 1$  there is a clustering of data points below the horizontal axis mainly for the categories of interest rates (IR), housing, inventories and money, indicating that the alternative approaches forecast more accurately than naïve for these categories of variables. Third, as the forecast horizon  $h$  increases, the points below the horizontal axis move towards it and are scattered around this line. Hence any gains in forecast accuracy achieved by using Ridge, LARS, Bag-LARS or BVAR, seem to disappear as  $h$  increases. This signals that the predictive power of the large information set is not persistent for Australia. In general, these observations are consistent with the conclusions reached from the tabulated results. Obviously the advantage here is that we are able to identify the categories of variables beyond the three variables of interest for which there are some considerable gains in forecasting accuracy compared to the naïve benchmark, from using the information set of  $K = 43$ , especially for  $h = 1$  and in some cases  $h = 2$ . We next delve into assessing the results for all  $K = 151$  variables through visualisation.

Some very interesting observations can be made from analysing the scatter plots for each of the  $K = 151$  variables within each of the twelve categories across Figures 3 to 10. Again, in general the DFM5 is clearly the least accurate, and not a suitable method for forecasting Australian macroeconomic variables. In contrast to the previous results, there are some very noticeable exceptions for the case of the AR(4) benchmark where it now performs much better in some categories. We comment on these below.

The first row in both scatter plots in Figures 3 and 7 confirm what we have concluded so far — that it is very challenging to beat the naïve benchmark for forecasting GDP — but also generalise this results across all the variables belonging to this category, with very few exceptions for some variables and mostly for  $h = 1$ . In contrast to the GDP category, the second row of the scatter plots shows some sizeable improvements in forecast accuracy over the naïve benchmark for the industrial production (IP) category, especially for  $h = 1$  and 2 quarters ahead.

The final rows of Figures 3 and 7 and the first rows of Figures 4 and 8 show the results for the employment and unemployment categories respectively. With the exception of the DFM5 and the AR(4), which are once again clearly outperformed by the naïve benchmark, the other competing approaches show some gains for some of the variables over the naïve benchmark, especially for  $h = 1$ .

The results of the housing category, which are shown in the middle rows of Figures 4 and 8, are very interesting. In general, the results show gains over the naïve benchmark, especially for  $h = 1$ , and the only exception is for the DFM5. In fact the AR(4) benchmark shows gains for some variables in the RMSE, up to  $h = 4$ , highlighting the dynamics in the housing variables. The other two categories of variables where gains are observed over the naïve benchmark (including the AR(4) in this case but excluding the DFM5) are prices and interest rates for  $h = 1$ .

In summarising this section we can conclude that visualising the results for both  $K = 43$  and  $K = 151$  has assisted in identifying specific categories of variables where some gains in forecast accuracy over the naïve benchmark can be achieved. Also overall we can see that at the very worst case, the alternative approaches considered converge to the forecasting accuracy of the naïve benchmark due to the shrinkage and selection properties of the approaches.

## 6 Conclusion

The dataset comprising a large number of Australian macroeconomic variables provides an alternative empirical platform for research on macroeconomic forecasting to the ‘default’ US

data of Stock and Watson (2002). Despite significant differences between the US and Australian economies, we find some similar empirical results.

From an Australian perspective, adding extra predictors can improve the forecast accuracy over small models, especially for the shorter forecast horizons of one or two quarters ahead. In particular, we find substantial gains in the accuracy of forecasts of the key macroeconomic variables of interest after increasing the information set from 3 to 13 predictors. However, the benefits from using conjunctural information (mostly disaggregate variables) beyond the main aggregate variables are found to be minor. The most accurate forecasts for the key macroeconomic variables are largely obtained with no more than 23 predictors, which are aggregates that characterize important aspects of the Australian economy. This finding is in line with Bańbura et al. (2010) and Koop (2013) who find little gain in expanding the information set beyond 20 important aggregates for the US data when forecasting with Bayesian VARs. We find that this is a general result, and robust to the different approaches to forecasting that we consider in the paper.

A major point of difference with the conclusion made by Stock and Watson (2012) is that, in contrast to the US data, the dynamic factor model does not retain its good forecast performance for the Australian economy. This result holds for different target variables, different sizes of information sets, and for different forecast horizons. Furthermore, this general conclusion is also robust to including either fewer or more dynamic factors. It will be interesting to investigate whether shrinkage methods outperform the dynamic factor model for other small open economies.

An important caveat to all results is that it is difficult to outperform the naïve sample mean benchmark. For example, using the sample mean as a forecast for GDP growth is qualitatively similar to forecasting with more sophisticated methods. This indicates that it is difficult to outperform a random walk with drift for GDP in levels. This result is exacerbated at longer forecast horizons, implying that Australian GDP data has very low persistence after transformation via differencing (even more so than the US data). This is further highlighted by the fact that the AR(4), a benchmark usually considered for the US economy (see for example Stock & Watson, 2002), becomes mostly irrelevant for the Australian case.

Our results suggest a few ways in which forecasts of Australian macroeconomic variables could be further improved. As is the case for the US data, using models with non linearities and/or time varying parameters in this context is an interesting avenue of future research. However,

even for forecasts that are linear combinations of predictors, it may be worthwhile including other types of macroeconomic variables in the information set. In particular, for a small open economy such as Australia that is reliant on commodity exports, macroeconomic data from major trading partners may be more useful predictors of Australian GDP than domestic data. Although pursuing such an approach leads to even larger information sets, the statistical methods used in this paper highlight a few of the tools that can help to tackle the problem of macroeconomic forecasting.

## A Data Appendix

The complete data set as used in the paper can be downloaded the Australian Macro Database from <http://ausmacrodata.org/research.php>. An up-to-date complete data set can also be downloaded from this page.

**Table 1:** *Abbreviations used for each series and their Australian Bureau of Statistics or Reserve Bank of Australia Series ID. Data on the AllOrds were obtained from Yahoo Finance*

Name	Series ID	Name	Series ID	Name	Series ID
RGDP	A2304402X	IP_NMMP	A2716605L	HCE-Post	A3604720V
Cons_Food	A2303246R	IP_MP	A2716603J	HCE-TelEqu	A3604721W
Cons_CigTob	A2303248V	IP_ME	A2716604K	HCE-AVCE	A3604738T
Cons_Alc	A2303250F	RetSal	h03hist	HCE-DomTra	A3604739V
Cons_CloFoo	A2303252K	JobVac	A590698F	HCE-IntTra	A3604740C
Cons_RenDwe	A2303254R	Emp_LFPer	h05hist	HCE-Pet	A3604719K
Cons_EGF	A2303256V	Emp_HoursFtPer	A3346482L	PEXP	f15hist
Cons_FurEqu	A2303258X	Emp_HoursPtPer	A3346483R	PIMP	f15hist
Cons_Hea	A2303260K	Emp_HoursPer	A3346481K	COMMMP	i02hist
Cons_PurVeh	A2303262R	Emp_FtPer	A181513R	RHDI	h02hist
Cons_OpeVeh	A2303264V	Emp_PtPer	A181514T	TotalComp	A2302607T
Cons_TraSer	A2303266X	Emp_TotalPer	A181515V	AveComp	A2302608V
Cons_Com	A2303268C	URate_OutofLF	h05hist	IBR	f01hist
Cons_RecCul	A2303270R	URate_LookforFtPer	A181526A	90 days Bk bills	f01dhist
Cons_Edu	A2303272V	URate_Person	A181525X	5 yr T-bond	f02histhist
Cons_HCR	A2303274X	UNum_LookforFtPer	A181516W	10 yr T-bond	f02histhist
Cons_InsFin	A2303276C	UNum_LookforPtPer	A181517X	10 yr NSW-T-bond	f02hist
Cons_Other	A2303278J	UNum_Person	A181518A	Currency	d03hist
Cons_Final	A2303280V	HSTotalexfinNum	A2412508V	M1	d03hist
FixInv	A2304051J	HSTotalexfinVal	A2412509W	M3	d03hist
ResInv	A2304054R	HStarts: NSW	A2412632C	BM	d03hist
NResInv-Equip	A2304041C	HStarts: VIC	A2412636L	Mbase	d03hist
NResInv-CBR	A2716097C	HStarts: QLD	A2412644L	Credit_Person	d02hist
NResInv-IPP	A2716583K	HStarts: SA	A2412648W	Credit_Bus	d02hist
Exports	A2304114F	HStarts: WA	A2412656W	Credit_Total	d02hist
Imports	A2304115J	HStarts: TAS	A2412662T	Credit_Nar	d02hist
Gov	A2304080V	HStarts: NT	A2412668F	LoaAdv	d02hist
Gov Nat	A2304078J	HStarts: ACT	A2412672W	Ex rate: JPY	f11hist
Gov Nat/Loc	A2304079K	HStarts: Total	A2412624C	Ex rate: USD	f11hist
DomSales	A2302599C	Hstarts_PDA	h03hist	EX rate: KRW	f11hist
IP_Mining	A2716600A	InvLevel	A2302598A	Ex rate: NZD	f11hist
IP_Man	A2716601C	InvtoSales	A2302601C	Ex rate: CNY	f11hist
IP_EGWWS	A2716610F	CPI-ALL	A2325846C	Ex rate: GBP	f11hist
IP_TotalInd	A2298671X	HCE-UTF	A3604727K	Ex rate: SGD	f11hist
IP_FBT	A2716602F	HCE-MV	A3604722X	EX rate: IDR	f11hist
IP_TCO	A2716608V	HCE-ParAcc	A3604725F	Ex rate: HKD	f11hist
IP_WPP	A2716609W	HCE-Fuel	A3604723A	Ex rate: MYR	f11hist
IP_PRM	A2716607T	HCE-MaiRep	A3604726J	Ex rate: avg	f11hist
IP_PCCRP	A2716606R	HCE-OthMot	A3604724C	SP ASX AllOrds	Yahoo Finance
EPI_LA	A2295549R	EPI_TFW	A2295588F	EPI_PPA	A2295645R
EPI_Meat	A2295570J	EPI_CFCM	A2295591V	EPI_IS	A2295654T
EPI_DPBE	A2295573R	EPI_MOMS	A2295594A	EPI_NFM	A2295651K
EPI_FCMAI	A2295552C	EPI_CCB	A2295606X	EPI_PGME	A2295612V
EPI_Cer	A2295555K	EPI_PPRM	A2295600K	EPI_TESEM	A2295615A
EPI_VF	A2295558T	EPI_GNM	A2295603T	EPI_RV	A2295618J
EPI_SH	A2295561F	EPI_AOF	A2295609F	EPI_OTE	A2295621W
EPI_CTCSM	A2295576W	EPI_MPP	A2295663V	EPI_PSCIA	A2295627K
EPI_FSA	A2295564L	EPI_PNPF	A2295669J	EPI_PAESO	A2295630X
EPI_HSFR	A2295582T	EPI_CMP	A2295675C	EPI_MMA	A2295633F
EPI_CW	A2295585X	EPI_LLMDF	A2295639V	EPI_Gold	A2295636L
TermsTrade	A2304200A				

**Table 2:** Categories of variables in the Australian macroeconomic data set

Category (C)	Description	Number of Variables
1	GDP components	30
2	Industrial Production	13
3	Employment	9
4	Unemployment rate	5
5	Housing	12
6	Inventories	2
7	Prices	50
8	Wages	3
9	Interest rates	5
10	Money	10
11	Exchange rates	11
12	ASX S&P All Ordinaries index	1

**Table 3:**  $w_t$  denotes an observed variable in levels. Transformation (T) denotes the transformation implemented to achieve stationarity: 1 = no transformation; 2 = first difference; 3 = second difference; 4 = log; 5 = first difference of logged variables; and 6 = second difference of logged variables.

Transformation (T)	
1	$w_t$
2	$w_t - w_{t-1}$
3	$(w_t - w_{t-1}) - (w_{t-1} - w_{t-2})$
4	$\ln(w_t)$
5	$\ln(w_t/w_{t-1})$
6	$\ln(w_t/w_{t-1}) - \ln(w_{t-1}/w_{t-2})$

**Table 4:** The three key macroeconomic variables in  $\mathcal{I}_3$ . These are used for forecasting in all settings.

Name	T	C	Description
RGDP	5	1	Gross domestic product: Chain volume measures
CPI-ALL	6	7	Index Numbers; All groups CPI
IBR	2	9	Interbank overnight cash rate in Australia



**Table 5:** Variables in  $\mathcal{I}_{13} \setminus \mathcal{I}_3$  where  $A \setminus B$  denotes the elements of  $A$  that are not also elements of  $B$ .

Name	T	C	Description
IP_TotalInd	5	2	Total industrial industries; Index
Emp_TotalPer	5	3	Employed – total; Persons
Hstarts_PDA	5	5	Private dwelling approvals
COMMP	6	7	Index of commodity prices; All items; AUD, Index, 2013/14=100
Exports	5	1	Exports of goods and services
Imports	5	1	Imports of goods and services
TermsTrade	5	7	Terms of Trade; Index
M1	6	10	M1
Credit_Total	6	10	Credit; Total
SP ASX AllOrds	5	12	S& P ASX AllOrds adjusted closing prices

**Table 6:** Variables in  $\mathcal{I}_{23} \setminus \mathcal{I}_{13}$ .

Name	T	C	Description
Cons_Final	5	1	Household Final Consumption Expenditure: Chain volume measures
Urate_Per	3	4	Unemployment rate; Persons
AveComp	5	8	Average non-farm compensation per employee: Current prices
10 yr T-bond	2	9	10 yrs Australian Government
M3	6	10	M3
BM	6	10	Broad money
Mbase	6	10	Money base
Credit_Per	6	10	Credit; Other personal
Credit_Bus	6	10	Credit; Business
Ex rate: avg	5	11	Australian Dollar Trade-weighted Index

**Table 7:** Variables in  $\mathcal{I}_{43} \setminus \mathcal{I}_{23}$ .

Name	T	C	Description
Gov	5	1	General government; Final consumption expenditure
DomSales	5	1	Domestic sales; Current prices
RetSal	2	2	Retail sales; All industries; Current price
Emp_LFPer	5	3	Labour Force; Persons
Emp_HoursPer	5	3	Aggregate Monthly Hours Worked; Persons
HSTotalexfinNum	5	5	Total excluding refinancing of established dwellings – No.
HSTotalexfinVal	5	5	Total excluding refinancing of established dwellings – Value
HStarts: Total	5	5	Total (State); Number of Commitments
InvLevel	5	6	Private non-farm inventory levels; book values; Current prices
InvtoSales	2	6	Inventories to total sales; Ratio
PEXP	5	7	Real export-weighted index; Original
PIMP	5	7	Real import-weighted index; Original
RHDI	5	8	Real household disposable income
90 days Bk bills	2	9	90 days Bank accepted bills
5 yr T-bond	2	9	5 yrs Australian Government
10 yr NSW T-bond	2	9	10 yrs NSW Treasury Corporation
Currency	6	10	Currency: Seasonally adjusted
Credit_Nar	6	10	Narrow credit including loans and advances by AFIs plus Bills on issue
LoaAdv	6	10	Loans and advances; Banks

**Table 8:** Variables in  $\mathcal{I}_{151} \setminus \mathcal{I}_{43}$ .

Name	T	C	Description
Cons_Food	5	1	CVM (Chain Volume Measure) Household Cons Food
Cons_CigTob	5	1	CVM Household Cons Cigarettes and tobacco
Cons_Alc	5	1	CVM Household Cons Alcoholic beverages
Cons_CloFoo	5	1	CVM Household Cons Clothing and footwear
Cons_RenDwe	5	1	CVM Household Cons Rent and other dwelling services
Cons_EGF	5	1	CVM Household Cons Electricity, gas and other fuel
Cons_FurEqu	5	1	CVM Household Cons Furnishings and household equipment
Cons_Hea	5	1	CVM Household Cons Health
Cons_PurVeh	5	1	CVM Household Cons Purchase of vehicles
Cons_OpeVeh	5	1	CVM Household Cons Operation of vehicles
Cons_TraSer	5	1	CVM Household Cons Transport services
Cons_Com	5	1	CVM Household Cons Communications
Cons_RecCul	5	1	CVM Household Cons Recreation and culture
Cons_Edu	5	1	CVM Household Cons Education services
Cons_HCR	5	1	CVM Household Cons Hotels, cafes and restaurants
Cons_InsFin	5	1	CVM Household Cons Insurance and other financial services
Cons_Other	5	1	CVM Household Cons Other goods and services
ResInv	5	1	Gross fixed capital formation-Dwellings
NResInv-Equip	5	1	Gross fixed capital formation-Machinery and equipment
NResInv-CBR	5	1	Gross fixed capital formation-Cultivated biological resources
NResInv-IPP	5	1	Gross fixed capital formation-Intellectual property products
Gov Nat	5	1	Government-National; Final Cons. expenditure
Gov Nat/Loc	5	1	Government-State and local; Final Cons. expenditure
IP_Mining	5	2	Mining excluding exploration and mining support services
IP_Man	5	2	Manufacturing
IP_EGWWS	5	2	Electricity, gas, water and waste services
IP_FBT	5	2	Food, beverage and tobacco products
IP_TCO	5	2	Textile, clothing and other manufacturing
IP_WPP	5	2	Wood and paper products
IP_PRM	5	2	Printing and recorded media
IP_PCCRP	5	2	Petroleum, coal, chemical and rubber products
IP_NMMP	5	2	Non-metallic mineral products
IP_MP	5	2	Metal products
IP_ME	5	2	Machinery and equipment
JobVac	5	3	Job Vacancies; Australia
Emp_HoursFtPer	5	3	Aggregate Monthly Hours Worked (Emp f/t); Persons
Emp_HoursPtPer	5	3	Aggregate Monthly Hours Worked (Emp p/t); Persons
Emp_FtPer	5	3	Employed – full-time; Persons
Emp_PtPer	5	3	Employed – part-time; Persons
URate_OutofLF	3	4	Unemployed persons as percentage of labour force
URate_LookforFtPer	3	4	Unemployment rate – looking for full-time work; Persons

*Continued on next page*

Table 8: *continued from previous page*

Name	T	C	Description
UNum_LookforFtPer	5	4	Unemployed – looking for full-time work; Persons
UNum_LookforPtPer	5	4	Unemployed – looking for part-time work; Persons
UNum_Person	5	4	Unemployed – total; Persons
HStarts: NSW	5	5	New South Wales; Number of Commitments
HStarts: VIC	5	5	Victoria; Number of Commitments
HStarts: QLD	5	5	Queensland; Number of Commitments
HStarts: SA	5	5	South Australia; Number of Commitments
HStarts: WA	5	5	Western Australia; Number of Commitments
HStarts: TAS	5	5	Tasmania; Number of Commitments
HStarts: NT	5	5	Northern Territory; Number of Commitments
HStarts: ACT	5	5	Australian Capital Territory; Number of Commitments
HCE-UTF	6	7	Index; Urban transport fares
HCE-MV	6	7	Index; Motor vehicles
HCE-ParAcc	6	7	Index; Spare parts and accessories for motor vehicles
HCE-Fuel	6	7	Index; Automotive fuel
HCE-MaiRep	6	7	Index; Maintenance and repair of motor vehicles
HCE-OthMot	6	7	Index; Other services in respect of motor vehicles
HCE-Post	6	7	Index; Postal services
HCE-TelEqu	6	7	Index; Telecommunication equipment and services
HCE-AVCE	6	7	Index; Audio, visual and computing equipment
HCE-DomTra	6	7	Index; Domestic holiday travel and accommodation
HCE-IntTra	6	7	Index; International holiday travel and accommodation
HCE-Pet	6	7	Index; Pets and related products
TotalComp	5	8	Non-farm; Total compensation of employees: Current prices
Ex rate: JPY	5	11	AUD/JPY Exchange Rate
Ex rate: USD	5	11	AUD/USD Exchange Rate
EX rate: KRW	5	11	AUD/KRW Exchange Rate
Ex rate: NZD	5	11	AUD/NZD Exchange Rate
Ex rate: CNY	5	11	AUD/CNY Exchange Rate
Ex rate: GBP	5	11	AUD/GBP Exchange Rate
Ex rate: SGD	5	11	AUD/SGD Exchange Rate
EX rate: IDR	5	11	AUD/IDR Exchange Rate
Ex rate: HKD	5	11	AUD/HKD Exchange Rate
Ex rate: MYR	5	11	AUD/MYR Exchange Rate
EPI_LA	5	7	Index Numbers; 00 Live animals other than division 03
EPI_Meat	5	7	Index Numbers; 01 Meat and meat preparations
EPI_DPBE	5	7	Index Numbers; 02 Dairy products and birds' eggs
EPI_FCMAI	5	7	Index Numbers; 03 Fish, crustaceans, molluscs and aquatic invertebrates and preparations thereof
EPI_Cer	5	7	Index Numbers; 04 Cereals and cereal preparations
EPI_VF	5	7	Index Numbers; 05 Vegetables and fruit
EPI_SH	5	7	Index Numbers; 06 Sugars, sugar preparations and honey

*Continued on next page*

**Table 8:** *continued from previous page*

Name	T	C	Description
EPI_CTCSM	5	7	Index Numbers; 07 Coffee, tea, cocoa, spices and manuf. thereof
EPI_FSA	5	7	Index Numbers; 08 Feeding stuff for animals (not including unmilled cereals)
EPI_HSFR	5	7	Index Numbers; 21 Hides, skins and furskins, raw
EPI_CW	5	7	Index Numbers; 24 Cork and wood
EPI_TFW	5	7	Index Numbers; 26 Textile fibres and their wastes
EPI_CFCM	5	7	Index Numbers; 27 Crude fertilizers, other than those of division 56, and crude minerals (excluding coal, petroleum)
EPI_MOMS	5	7	Index Numbers; 28 Metalliferous ores and metal scrap
EPI_CCB	5	7	Index Numbers; 32 Coal, coke and briquettes
EPI_PPRM	5	7	Index Numbers; 33 Petroleum, petroleum products and related materials
EPI_GNM	5	7	Index Numbers; 34 Gas, natural and manufactured
EPI_AOF	5	7	Index Numbers; 41 Animal oils and fats
EPI_MPP	5	7	Index Numbers; 54 Medical and pharmaceutical products
EPI_PNPF	5	7	Index Numbers; 58 Plastics in non-primary forms
EPI_CMP	5	7	Index Numbers; 59 Chemical materials and products, n.e.s.
EPI_LLMDF	5	7	Index Numbers; 61 Leather, leather manufactures, n.e.s., and dressed furskins
EPI_PPA	5	7	Index Numbers; 64 Paper, paperboard, and articles of paper pulp, of paper or of paperboard
EPI_IS	5	7	Index Numbers; 67 Iron and steel
EPI_NFM	5	7	Index Numbers; 68 Non-ferrous metals
EPI_PGME	5	7	Index Numbers; 71–75 Power generating, general industrial and other specialised machinery and equipment
EPI_TESEM	5	7	Index Numbers; 76–77 Telecommunications equipment, sound recording apparatus and electrical machinery, n.e.s.
EPI_RV	5	7	Index Numbers; 78 Road vehicles (incl. air-cushion vehicles)
EPI_OTE	5	7	Index Numbers; 79 Other transport equipment
EPI_PSCIA	5	7	Index Numbers; 87 Professional, scientific and controlling instruments and apparatus, n.e.s.
EPI_PAESO	5	7	Index Numbers; 88 Photographic apparatus, equipment and supplies and optical goods, n.e.s.
EPI_MMA	5	7	Index Numbers; 89 Miscellaneous manufactured articles, n.e.s.
EPI_Gold	5	7	Index Numbers; 97 Gold, non-monetary (excluding gold ores and concentrates)

**Table 9:** Forecast accuracy for  $h = 1$  to 4. Each entry shows RMSE relative to the naïve benchmark.  $K$  denotes the number of predictors in the information set used. Note that DFM5 in fact denotes a DFM with 3 factors for  $K = 3$ . BagL denotes Bagging-LARS. Bold entries indicate the lowest error measure achieved by the competing approaches for the variable of interest across each row, i.e., using an increasing number of predictors.

	$K = 3$	13	23	43	151	3	13	23	43	151	3	13	23	43	151
	GDP					CPI					IBR				
	$h = 1$														
DFM5	<b>1.06</b>	1.07	1.07	1.09	1.09	1.16	<b>1.12</b>	1.15	1.15	1.23	1.32	1.29	<b>1.28</b>	1.33	1.31
Ridge	1.03	1.03	<b>1.00</b>	1.04	1.02	1.03	0.99	1.00	0.99	<b>0.96</b>	0.88	0.86	0.85	<b>0.78</b>	0.80
LARS	<b>0.99</b>	1.01	1.06	1.01	1.00	1.03	1.00	0.99	1.00	<b>0.97</b>	0.92	0.93	0.94	<b>0.87</b>	<b>0.87</b>
BLAR	1.04	1.02	1.04	1.02	<b>1.01</b>	<b>0.92</b>	0.99	0.96	0.96	0.97	0.82	0.81	0.83	<b>0.73</b>	0.78
BVAR	1.34	1.02	<b>1.01</b>	<b>1.01</b>	<b>1.01</b>	<b>0.91</b>	0.98	1.01	0.99	0.98	0.98	<b>0.83</b>	0.87	0.85	0.89
AR(4)		1.07					0.86					0.75			
	$h = 2$														
DFM5	<b>0.96</b>	1.04	1.01	1.06	1.10	<b>1.03</b>	1.09	1.05	1.08	1.10	<b>1.10</b>	1.22	1.22	1.23	1.28
Ridge	1.09	1.03	<b>1.02</b>	<b>1.02</b>	1.03	1.14	1.01	1.00	<b>0.99</b>	1.02	0.97	0.95	0.97	<b>0.94</b>	<b>0.94</b>
LARS	1.04	<b>1.00</b>	<b>1.00</b>	1.01	1.03	1.04	1.01	1.04	<b>0.99</b>	<b>0.99</b>	0.99	<b>0.97</b>	1.00	0.99	0.99
BLAR	1.04	1.03	1.03	<b>1.02</b>	<b>1.02</b>	1.02	1.02	1.01	<b>1.00</b>	1.01	0.95	<b>0.94</b>	0.98	0.95	0.96
BVAR	1.35	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	1.12	1.01	1.01	<b>1.00</b>	<b>1.00</b>	1.11	<b>0.96</b>	0.98	<b>0.96</b>	0.97
AR(4)		1.04					1.06					0.95			
	$h = 3$														
DFM5	<b>1.02</b>	1.08	1.14	1.12	1.17	<b>1.02</b>	1.10	1.06	1.09	1.08	<b>1.05</b>	1.21	1.22	1.26	1.30
Ridge	1.08	1.03	1.02	<b>1.01</b>	1.03	1.09	0.99	<b>0.98</b>	<b>0.98</b>	1.00	1.03	0.99	1.01	<b>0.97</b>	0.99
LARS	1.01	1.01	1.01	1.03	<b>1.00</b>	1.00	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	1.05	<b>1.00</b>	1.02	1.03	1.03	<b>1.00</b>
BLAR	1.02	1.02	1.02	1.03	<b>1.01</b>	1.02	1.00	<b>0.99</b>	<b>0.99</b>	1.01	0.99	0.99	1.01	<b>0.98</b>	0.99
BVAR	1.34	1.01	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	1.09	1.00	1.00	<b>0.99</b>	<b>0.99</b>	1.12	<b>1.00</b>	1.01	<b>1.00</b>	<b>1.00</b>
AR(4)		1.05					1.04					1.00			
	$h = 4$														
DFM5	<b>1.01</b>	1.06	1.14	1.19	1.22	1.05	1.13	1.13	1.12	<b>1.04</b>	<b>1.12</b>	1.24	1.23	1.28	1.37
Ridge	1.03	1.02	<b>1.01</b>	<b>1.01</b>	1.02	1.04	0.99	<b>0.98</b>	<b>0.98</b>	0.99	1.05	0.99	1.00	<b>0.98</b>	1.00
LARS	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	1.01	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	1.02	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	1.01	<b>1.00</b>
BLAR	<b>1.00</b>	1.02	1.02	1.01	<b>1.00</b>	1.00	<b>0.99</b>	<b>0.99</b>	1.00	1.00	<b>0.99</b>	1.00	1.01	1.00	1.00
BVAR	1.24	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	1.20	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	1.19	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>
AR(4)		1.02					1.04					1.03			

**Table 10:** Forecast accuracy for  $h = 1$  to 4. Each entry shows MASE relative to the naïve benchmark.  $K$  denotes the number of predictors in the information set used. DFM5 is not applicable for  $K = 3$ . We use it in this case to refer to DFM3 to simplify the table presentation. BagL denotes Bagging-LARS. Bold entries indicate the lowest error measure achieved by the competing approaches for the variable of interest across each row, i.e., using an increasing number of predictors.

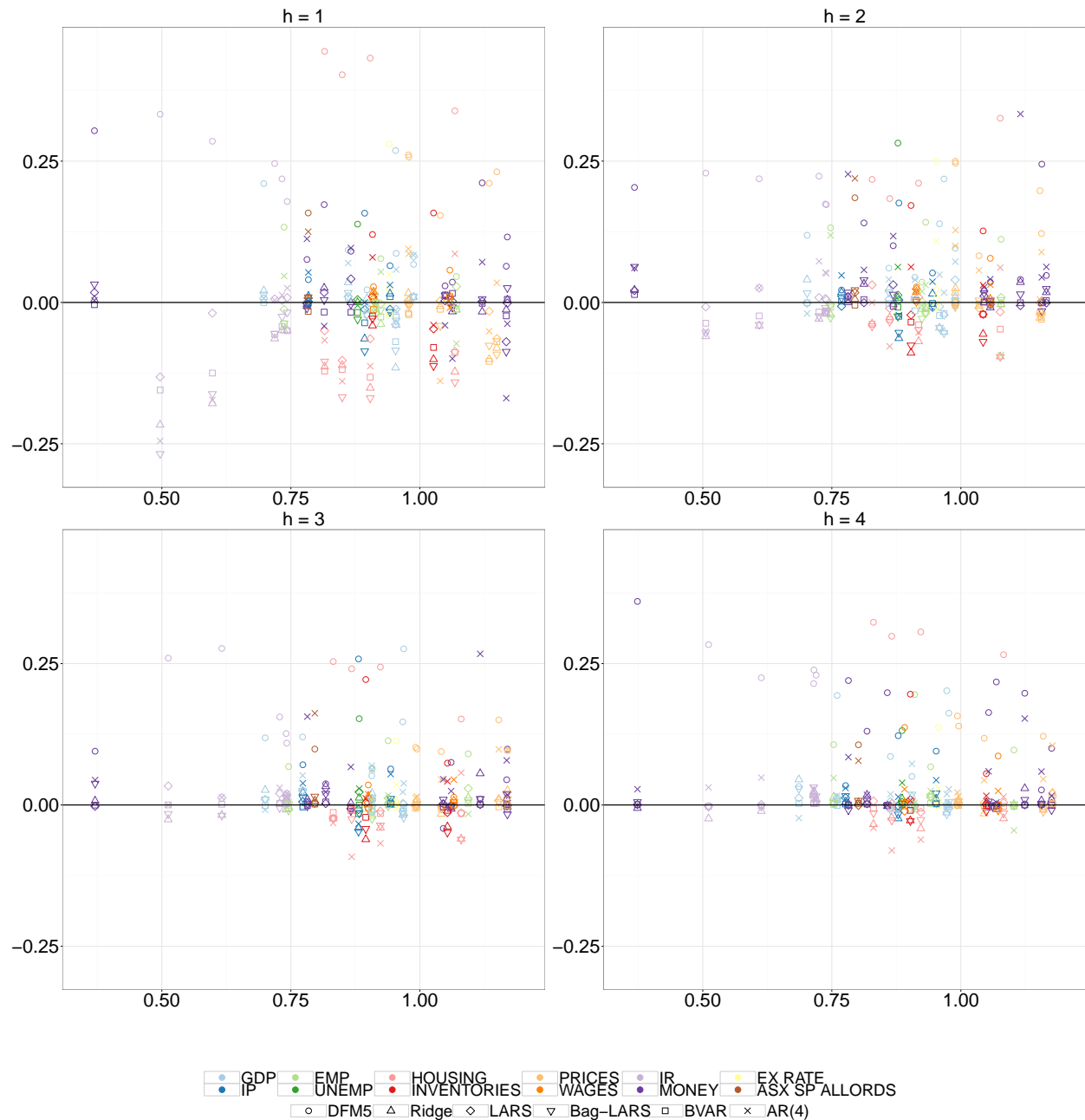
	$K = 3$	13	23	43	151	3	13	23	43	151	3	13	23	43	151
	GDP					CPI					IBR				
	$h = 1$														
DFM5	1.05	1.06	1.09	1.08	<b>1.04</b>	1.21	<b>1.17</b>	1.21	1.25	1.28	<b>1.28</b>	1.29	1.29	1.38	1.37
Ridge	1.03	<b>1.02</b>	<b>1.02</b>	1.05	1.03	0.98	1.01	0.98	<b>0.97</b>	0.98	0.91	0.87	0.84	<b>0.75</b>	0.80
LARS	0.99	1.01	1.05	1.02	<b>0.97</b>	1.03	1.00	1.03	<b>0.97</b>	1.01	0.87	0.91	0.92	0.80	<b>0.78</b>
BLAR	<b>1.02</b>	1.05	1.05	1.05	<b>1.02</b>	0.94	0.96	0.96	<b>0.93</b>	0.95	0.78	0.83	0.80	<b>0.70</b>	0.75
BVAR	1.33	1.03	<b>1.01</b>	<b>1.01</b>	<b>1.01</b>	<b>0.95</b>	0.96	0.99	0.98	0.97	1.07	0.83	0.84	<b>0.80</b>	0.85
AR(4)			1.08					0.87					0.76		
	$h = 2$														
DFM5	<b>0.98</b>	1.05	1.03	1.10	1.14	<b>1.05</b>	1.08	1.08	1.11	1.16	<b>1.14</b>	1.29	1.30	1.32	1.44
Ridge	1.11	1.06	1.04	1.03	<b>1.01</b>	1.18	1.03	1.00	<b>0.98</b>	1.02	1.04	<b>0.95</b>	0.99	0.96	1.01
LARS	<b>0.99</b>	<b>0.99</b>	1.00	1.02	1.05	1.05	1.02	1.01	1.01	<b>0.97</b>	1.00	1.00	1.00	0.99	<b>0.98</b>
BLAR	1.05	1.05	<b>1.02</b>	1.03	<b>1.02</b>	1.03	1.02	0.99	0.99	<b>0.98</b>	<b>0.92</b>	0.97	0.97	0.95	0.96
BVAR	1.40	<b>1.01</b>	1.02	1.02	1.02	1.21	1.01	1.00	<b>0.99</b>	1.00	1.27	0.97	0.97	<b>0.96</b>	0.97
AR(4)			1.07					1.07					0.98		
	$h = 3$														
DFM5	<b>0.99</b>	1.06	1.11	1.14	1.14	<b>1.08</b>	1.17	1.11	1.13	1.16	<b>1.07</b>	1.36	1.32	1.33	1.44
Ridge	1.09	1.04	1.04	1.04	<b>1.01</b>	1.11	1.01	<b>0.97</b>	<b>0.97</b>	1.02	1.12	<b>0.99</b>	1.04	1.01	1.01
LARS	1.00	<b>0.99</b>	<b>0.99</b>	1.05	1.05	1.01	1.00	1.01	0.98	<b>0.97</b>	<b>0.98</b>	1.02	1.02	1.01	<b>0.98</b>
BLAR	1.04	1.03	1.03	1.04	<b>1.02</b>	1.01	1.01	0.99	<b>0.98</b>	<b>0.98</b>	<b>0.96</b>	0.98	1.01	0.97	<b>0.96</b>
BVAR	1.32	1.01	<b>0.97</b>	<b>0.97</b>	1.02	1.20	1.00	<b>0.99</b>	<b>0.99</b>	1.00	1.19	1.00	1.01	1.00	<b>0.97</b>
AR(4)			1.06					1.05					1.05		
	$h = 4$														
DFM5	<b>1.00</b>	1.08	1.14	1.18	1.23	<b>1.09</b>	1.16	1.19	1.21	1.10	<b>1.17</b>	1.34	1.32	1.31	1.43
Ridge	1.04	1.04	1.04	1.04	<b>1.03</b>	1.09	1.00	<b>0.99</b>	<b>0.99</b>	1.01	1.13	0.98	0.99	<b>0.97</b>	1.03
LARS	<b>1.00</b>	<b>1.00</b>	1.03	1.01	1.01	1.00	0.99	1.00	1.00	<b>0.98</b>	0.99	1.00	1.02	1.00	<b>0.98</b>
BLAR	<b>1.00</b>	1.03	1.04	1.02	<b>1.00</b>	1.01	0.99	0.99	<b>0.97</b>	0.99	<b>0.95</b>	0.97	0.99	0.99	0.98
BVAR	1.29	<b>1.00</b>	1.03	1.03	1.03	1.29	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	1.25	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>
AR(4)			1.04					1.06					1.10		

**Table 11:** Each entry shows the average error measure relative to the naïve benchmark from forecasting all  $K$  variables in the information set. DFM5 is not applicable for  $K = 3$ . We use it in this case to refer to DFM3 to simplify the table presentation. BagL denotes Bagging-LARS. Bold entries indicate the lowest error measure achieved across the competing approaches for each forecast horizon.

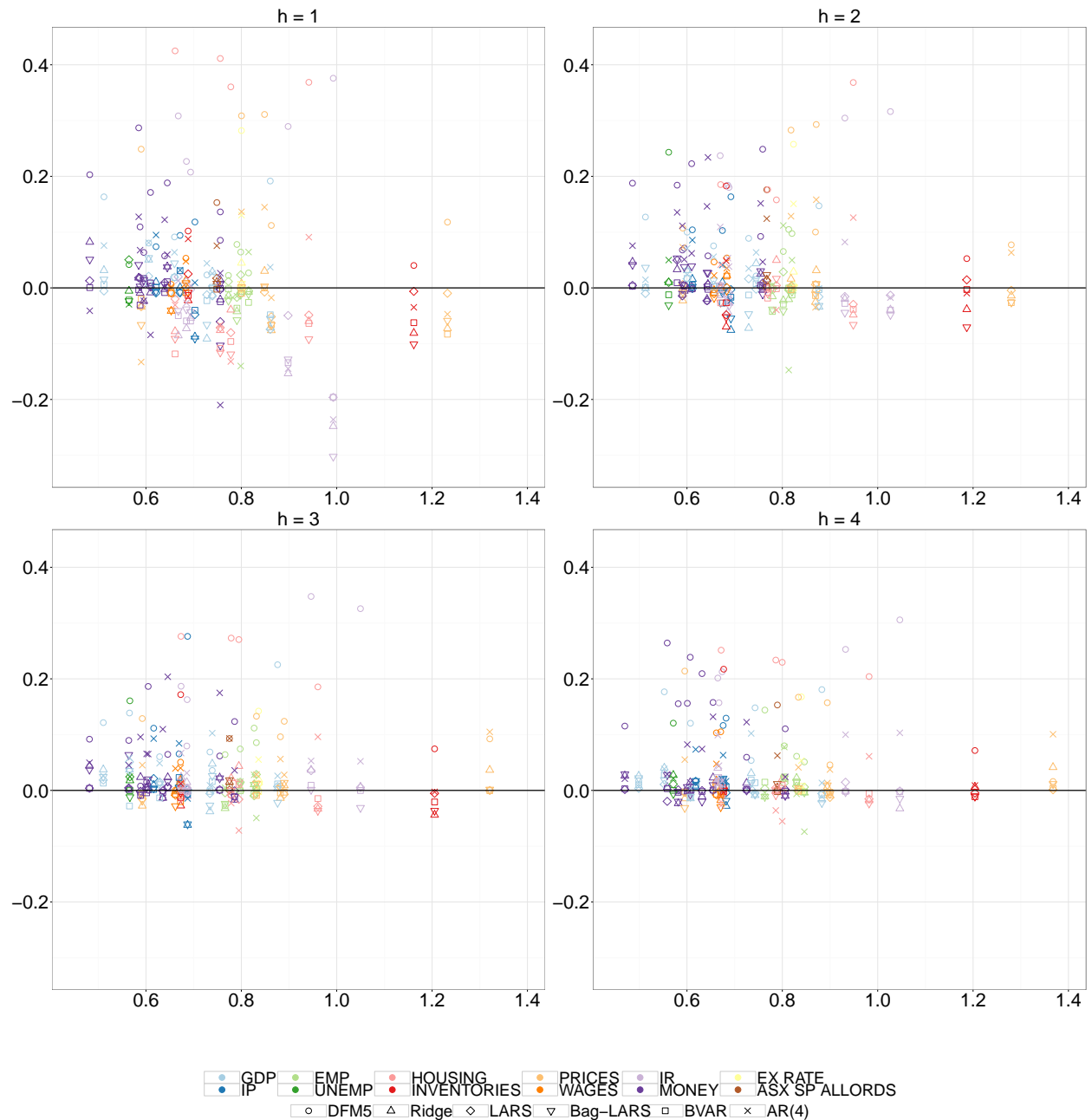
Average RMSE from forecasting all $K$ variables												
	$h = 1$	2	3	4	$h = 1$	2	3	4	$h = 1$	2	3	4
	$K = 3$				$K = 13$				$K = 23$			
DFM5	1.16	1.02	1.03	1.05	1.15	1.11	1.11	1.11	1.15	1.12	1.08	1.15
Ridge	1.03	1.09	1.08	1.05	0.97	<b>0.99</b>	<b>1.00</b>	<b>1.00</b>	<b>0.98</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>
LARS	0.99	<b>1.01</b>	<b>1.02</b>	<b>1.00</b>	1.00	1.00	<b>1.00</b>	<b>1.00</b>	1.00	<b>1.00</b>	1.01	<b>1.00</b>
BagL	0.94	<b>1.01</b>	<b>1.02</b>	<b>1.00</b>	0.98	1.00	<b>1.00</b>	<b>1.00</b>	<b>0.98</b>	<b>1.00</b>	1.01	1.01
BVAR	1.08	1.20	1.18	1.21	<b>0.96</b>	<b>0.99</b>	<b>1.00</b>	<b>1.00</b>	<b>0.98</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>
AR(4)	<b>0.91</b>	1.03	1.04	1.03	1.00	1.06	1.06	1.03	1.01	1.08	1.07	1.05
	$K = 43$				$K = 151$							
DFM5	1.16	1.14	1.11	1.15	1.15	1.12	1.10	1.13				
Ridge	0.97	<b>0.99</b>	1.00	<b>1.00</b>	0.99	<b>1.00</b>	1.01	1.01				
LARS	0.99	1.00	1.00	<b>1.00</b>	0.99	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>				
BagL	<b>0.96</b>	<b>0.99</b>	<b>0.99</b>	<b>1.00</b>	<b>0.98</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>				
BVAR	0.97	<b>0.99</b>	1.00	<b>1.00</b>	<b>0.98</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>				
AR(4)	1.00	1.05	1.04	1.03	1.02	1.08	1.08	1.09				
Average MASE from forecasting all $K$ variables												
	$K = 3$				$K = 13$				$K = 23$			
DFM5	1.20	1.07	1.05	1.10	1.14	1.13	1.13	1.14	1.15	1.14	1.11	1.15
Ridge	0.97	1.11	1.12	1.08	0.97	<b>0.99</b>	<b>1.00</b>	<b>1.00</b>	<b>0.98</b>	<b>1.00</b>	1.01	<b>1.00</b>
LARS	0.95	1.02	1.02	1.01	0.98	1.00	<b>1.00</b>	<b>1.00</b>	0.99	1.01	1.01	<b>1.00</b>
BagL	0.89	<b>0.99</b>	<b>1.00</b>	<b>0.98</b>	0.97	1.00	<b>1.00</b>	<b>1.00</b>	<b>0.98</b>	<b>1.00</b>	1.01	<b>1.00</b>
BVAR	1.11	1.29	1.23	1.27	<b>0.96</b>	<b>0.99</b>	<b>1.00</b>	<b>1.00</b>	<b>0.98</b>	<b>1.00</b>	<b>1.00</b>	<b>1.00</b>
AR(4)	<b>0.88</b>	1.03	1.05	1.07	0.99	1.05	1.06	1.05	1.01	1.07	1.07	1.06
	$K = 43$				$K = 151$							
DFM5	1.17	1.15	1.13	1.15	1.15	1.13	1.11	1.13				
Ridge	0.97	<b>0.99</b>	1.01	<b>1.00</b>	1.00	1.01	1.01	1.01				
LARS	0.98	1.00	<b>1.00</b>	<b>1.00</b>	0.99	1.00	<b>1.00</b>	<b>1.00</b>				
BagL	<b>0.96</b>	<b>0.99</b>	<b>1.00</b>	<b>1.00</b>	<b>0.97</b>	<b>0.99</b>	<b>1.00</b>	<b>1.00</b>				
BVAR	<b>0.96</b>	<b>0.99</b>	<b>1.00</b>	<b>1.00</b>	0.98	1.00	<b>1.00</b>	<b>1.00</b>				
AR(4)	1.00	1.06	1.05	1.04	1.01	1.06	1.05	1.05				



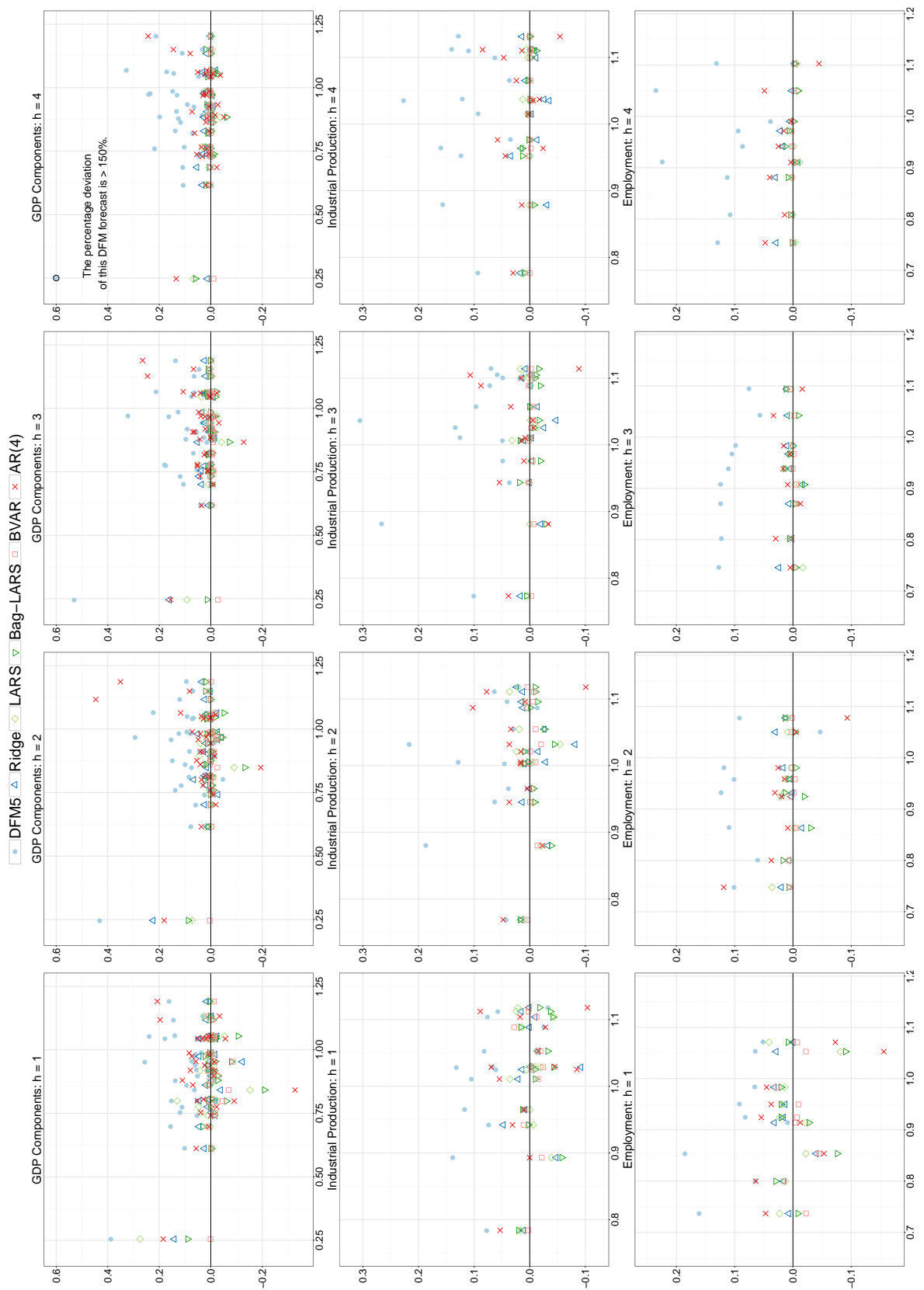
**Figure 1:** Comparing the  $h = 1$  to 4-steps ahead RMSE between the alternative forecasting approaches and the naïve benchmark for each variable in the information set when  $K = 43$ . The colors indicate the 12 categories the variables belong to and the shapes identify the competing forecast approaches implemented. The vertical axis denotes the percentage deviation in RMSE for the competing approaches relative to the naïve benchmark. The horizontal axis denotes the RMSE of the naïve benchmark.



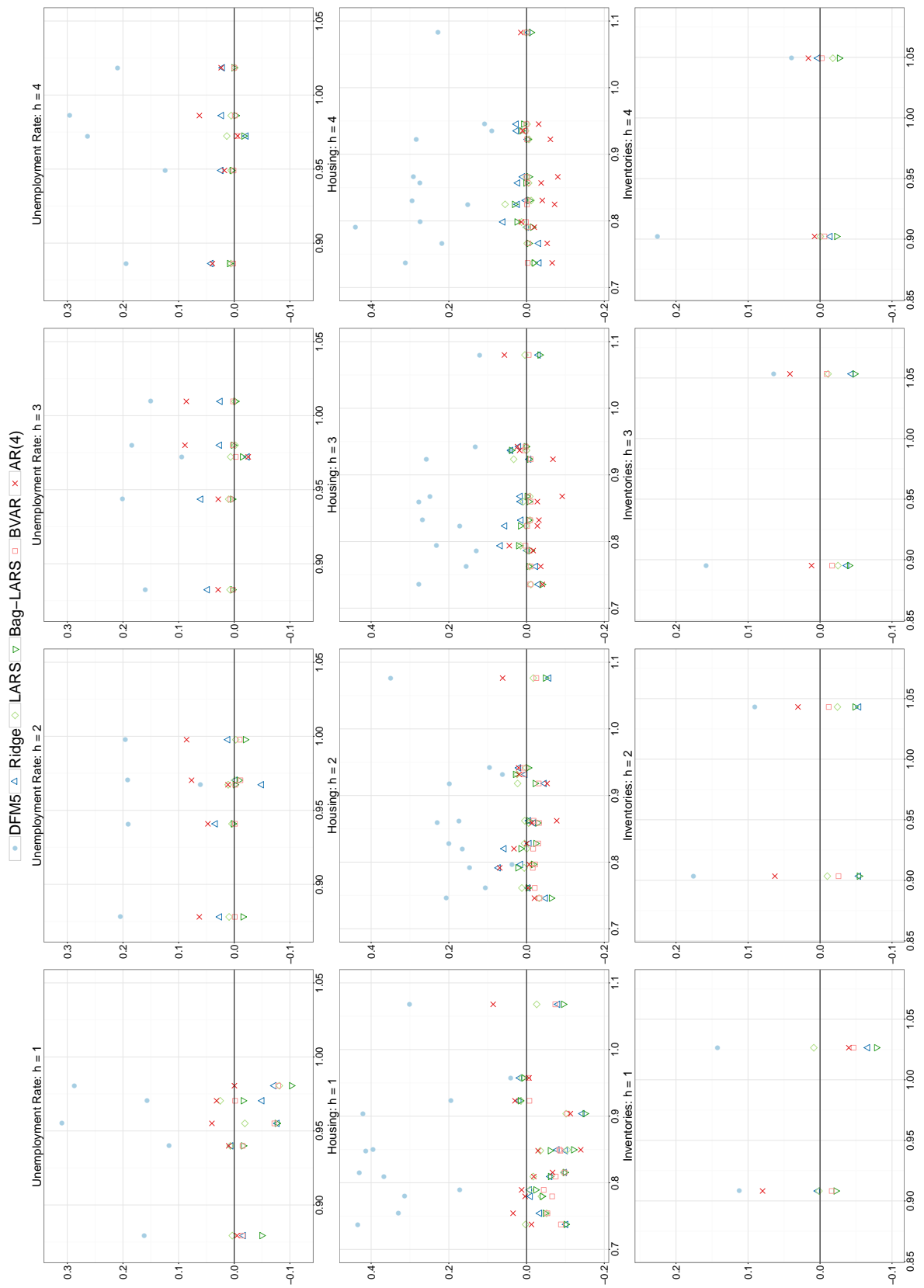
**Figure 2:** Comparing the  $h = 1$  to 4-steps ahead MASE between the alternative forecasting approaches and the naïve benchmark for each variable in the information set when  $K = 43$ . The colors indicate the 12 categories the variables belong to and the shapes identify the competing forecast approaches implemented. The vertical axis denotes the percentage deviation in MASE for the competing approaches relative to the naïve benchmark. The horizontal axis denotes the MASE of the naïve benchmark.



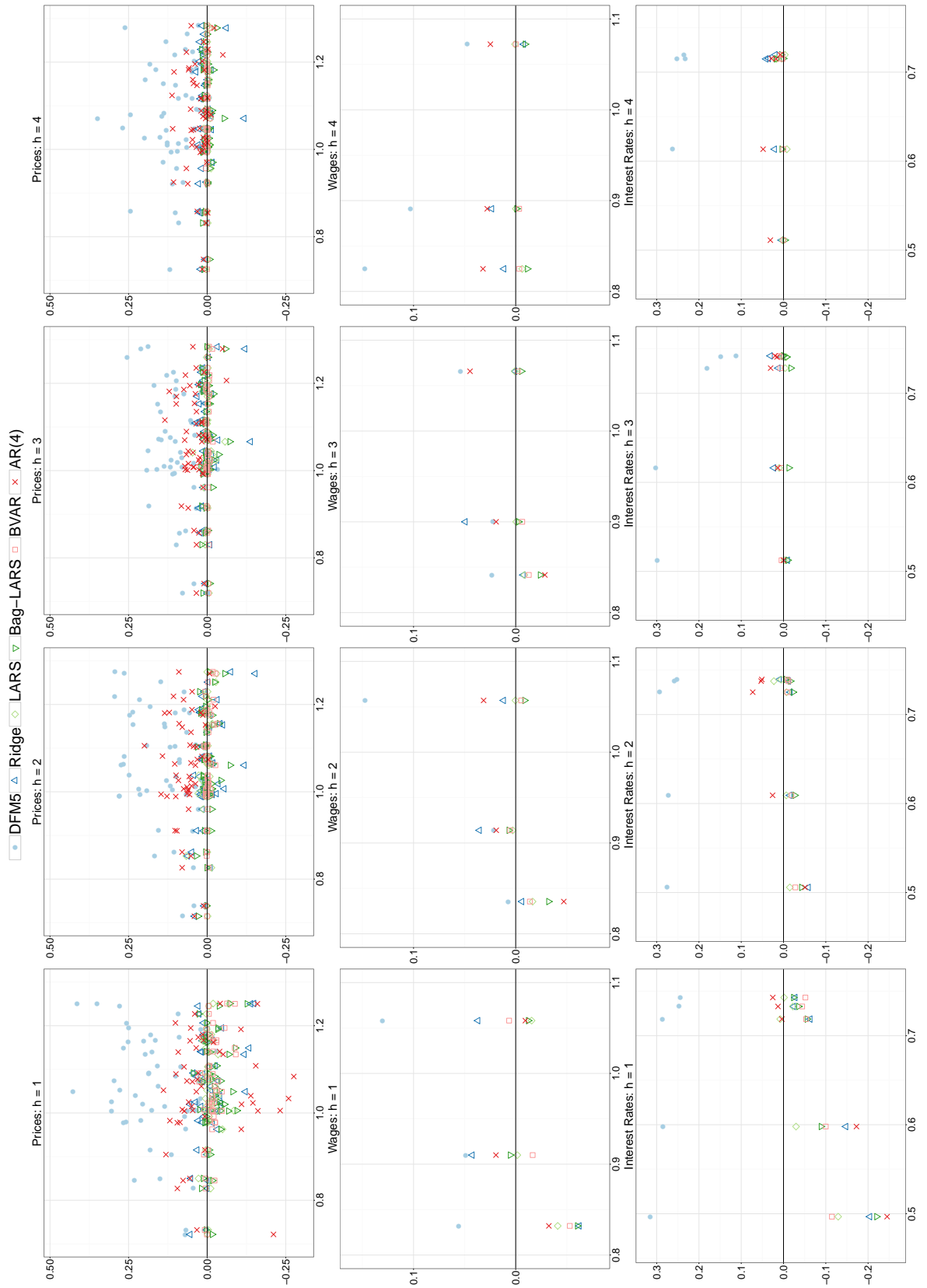
**Figure 3:** Comparing the  $h = 1$  to 4-steps ahead RMSE between the alternative forecasting approaches and the naive benchmark for each variable in the information set when  $K = 151$ . The colors indicate the 12 categories the variables belong to and the shapes identify the competing forecast approaches implemented. The vertical axis denotes the percentage deviation in RMSE for the competing approaches relative to the naive benchmark. The horizontal axis denotes the RMSE of the naive benchmark.



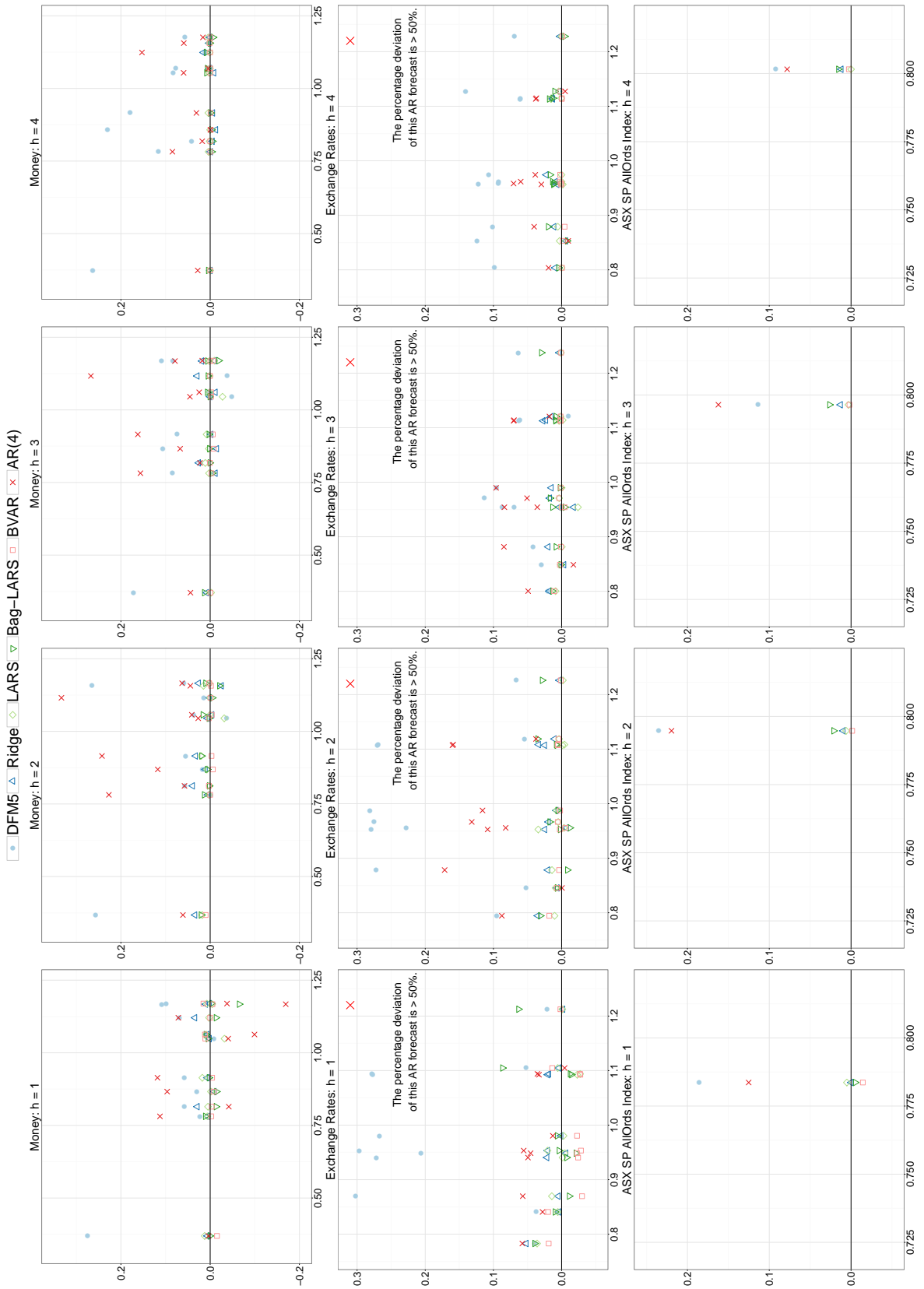
**Figure 4:** Comparing the  $h = 1$  to 4-steps ahead RMSE between the alternative forecasting approaches and the naive benchmark in the information set when  $K = 151$ . The colors indicate the 12 categories the variables belong to and the shapes identify the competing forecast approaches implemented. The vertical axis denotes the percentage deviation in RMSE for the competing approaches relative to the naive benchmark. The horizontal axis denotes the RMSE of the naive benchmark.



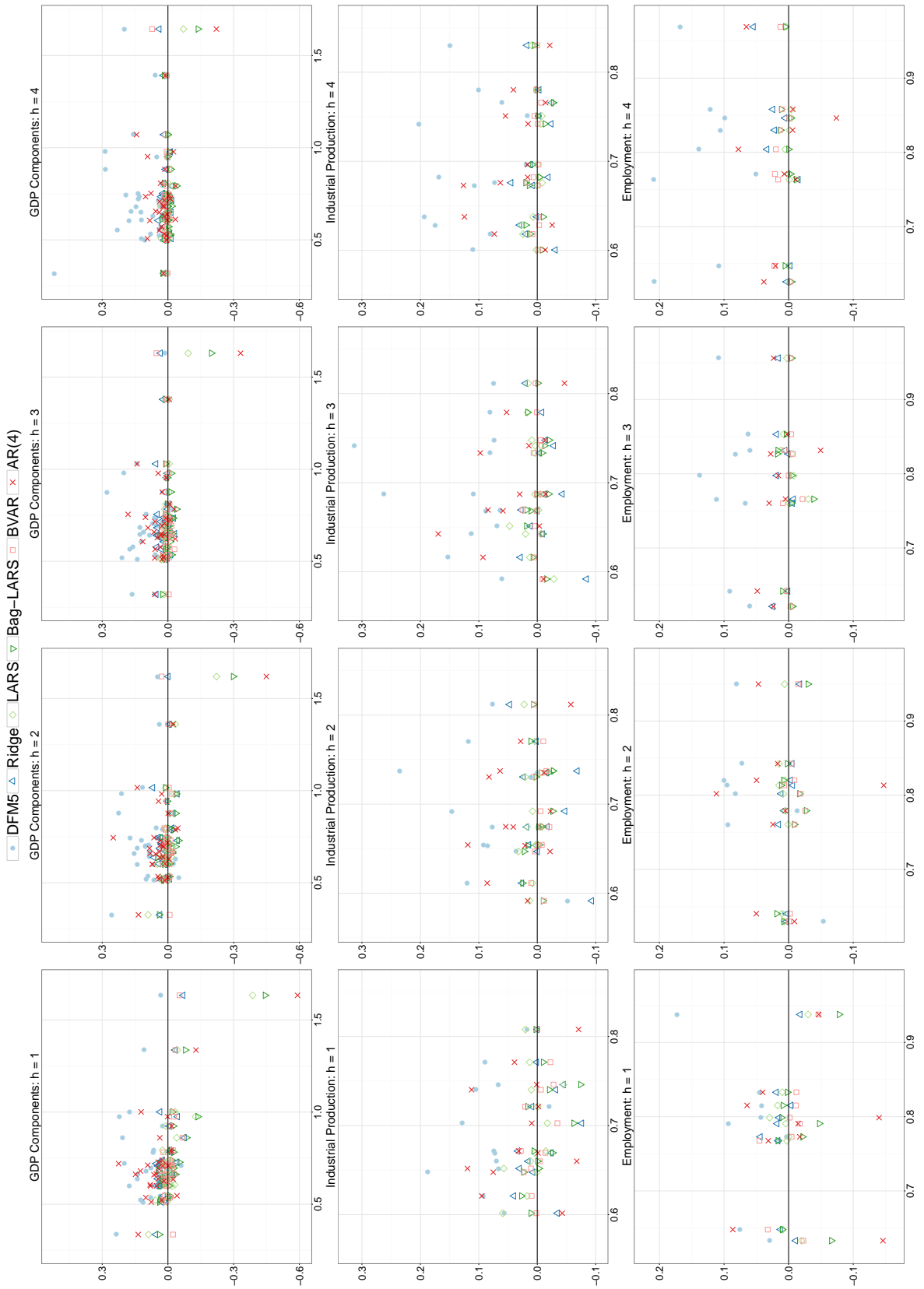
**Figure 5:** Comparing the  $h = 1$  to 4-steps ahead RMSE between the alternative forecasting approaches and the naive benchmark for each variable in the information set when  $K = 151$ . The colors indicate the 12 categories the variables belong to and the shapes identify the competing forecast approaches implemented. The vertical axis denotes the percentage deviation in RMSE for the competing approaches relative to the naive benchmark. The horizontal axis denotes the RMSE of the naive benchmark.



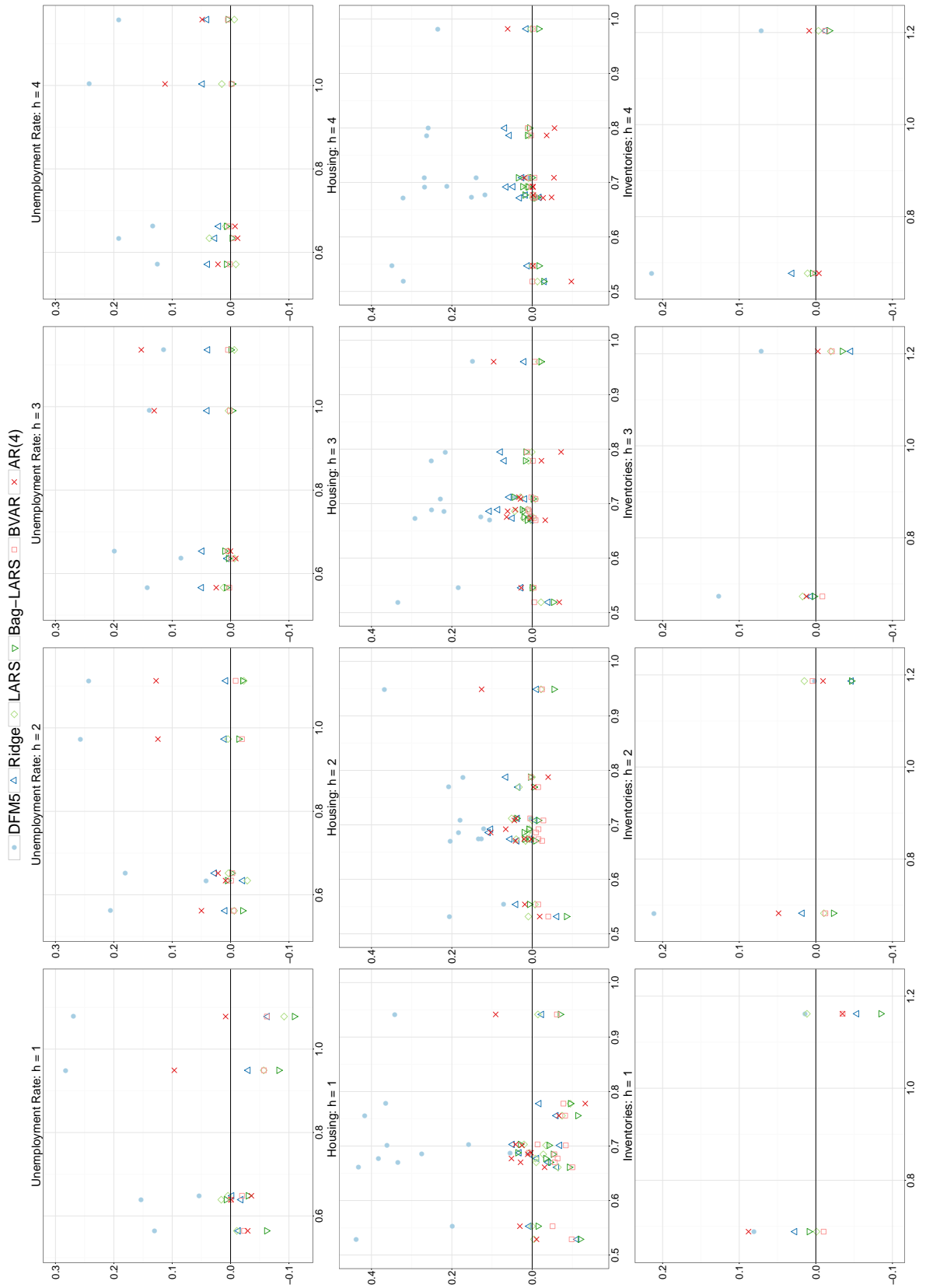
**Figure 6:** Comparing the  $h = 1$  to 4-steps ahead RMSE between the alternative forecasting approaches and the naive benchmark for each variable in the information set when  $K = 151$ . The colors indicate the 12 categories the variables belong to and the shapes identify the competing forecast approaches implemented. The vertical axis denotes the percentage deviation in RMSE for the competing approaches relative to the naive benchmark. The horizontal axis denotes the RMSE of the naive benchmark.



**Figure 7:** Comparing the  $h = 1$  to 4-steps ahead MASE between the alternative forecasting approaches and the naive benchmark for each variable in the information set when  $K = 151$ . The colors indicate the 12 categories the variables belong to and the shapes identify the competing forecast approaches implemented. The vertical axis denotes the percentage deviation in MASE for the competing approaches relative to the naive benchmark. The horizontal axis denotes the MASE of the naive benchmark.

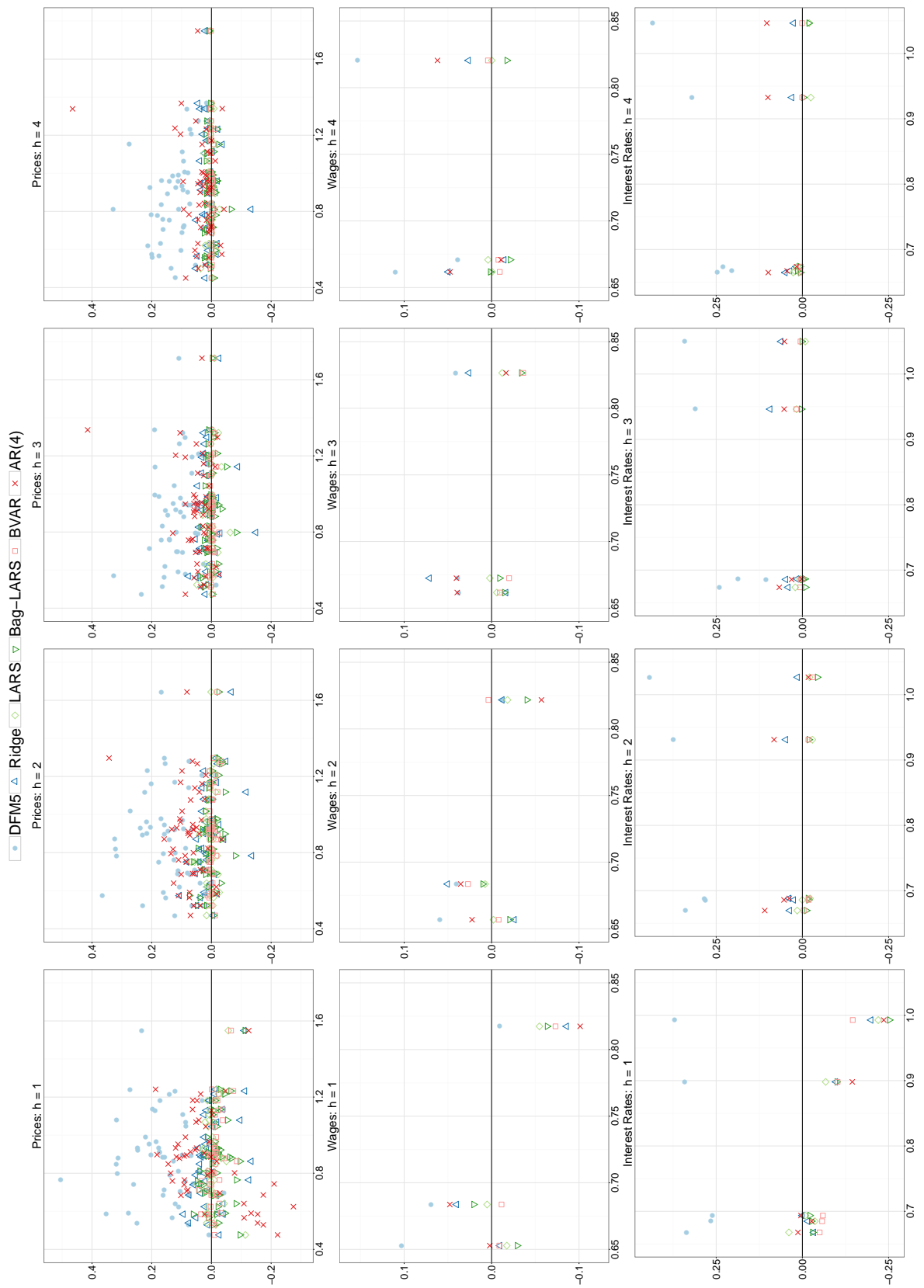


**Figure 8:** Comparing the  $h = 1$  to 4-steps ahead MASE between the alternative forecasting approaches and the naïve benchmark for each variable in the information set when  $K = 151$ . The colors indicate the 12 categories the variables belong to and the shapes identify the competing forecast approaches implemented. The vertical axis denotes the percentage deviation in MASE for the competing approaches relative to the naïve benchmark. The horizontal axis denotes the MASE of the naïve benchmark.

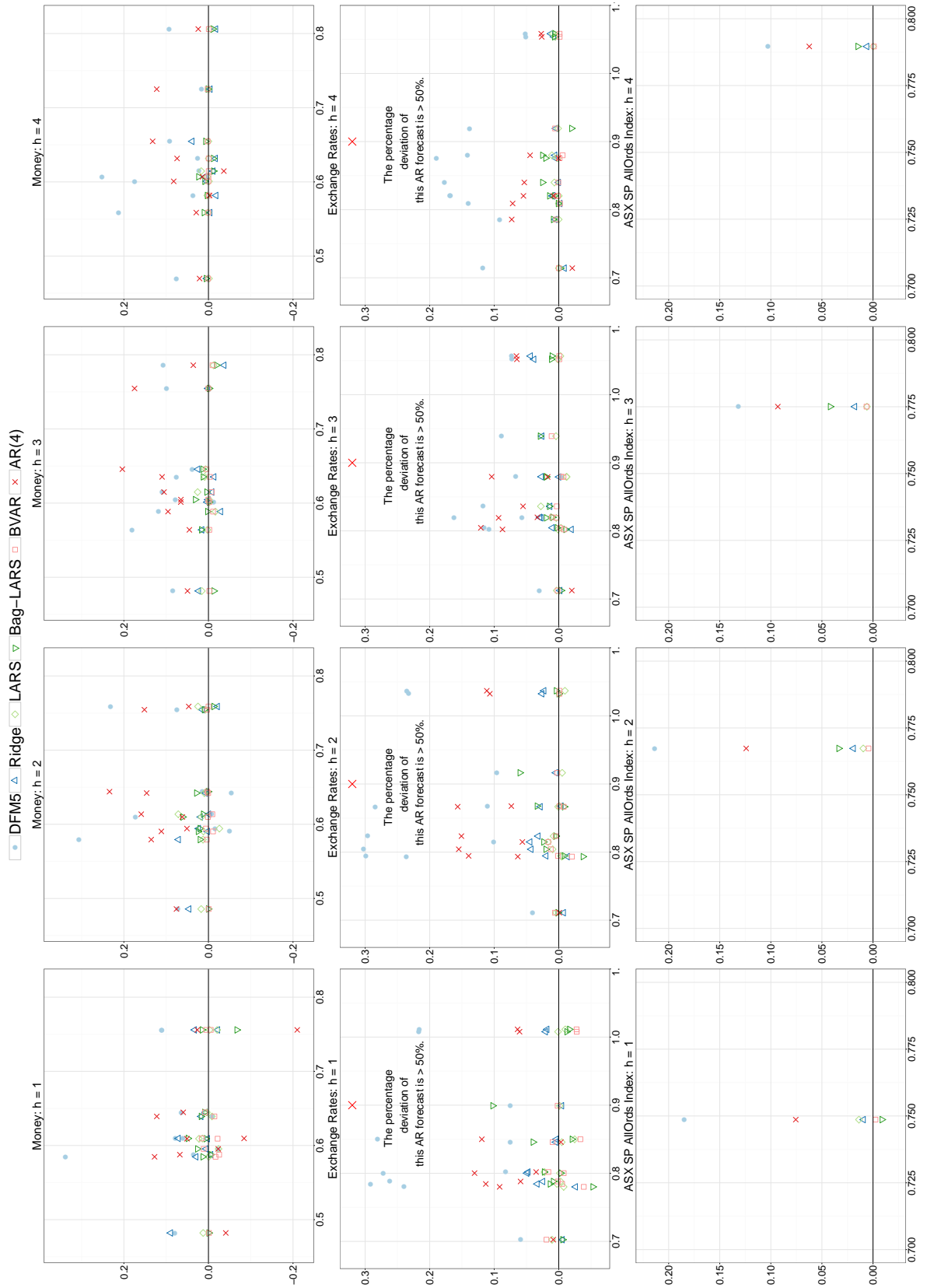




**Figure 9:** Comparing the  $h = 1$  to 4-steps ahead MASE between the alternative forecasting approaches and the naive benchmark for each variable in the information set when  $K = 151$ . The colors indicate the 12 categories the variables belong to and the shapes identify the competing forecast approaches implemented. The vertical axis denotes the percentage deviation in MASE for the competing approaches relative to the naive benchmark. The horizontal axis denotes the MASE of the naive benchmark.



**Figure 10:** Comparing the  $h = 1$  to 4-steps ahead MASE between the alternative forecasting approaches and the naïve benchmark for each variable in the information set when  $K = 151$ . The colors indicate the 12 categories the variables belong to and the shapes identify the competing forecast approaches implemented. The vertical axis denotes the percentage deviation in MASE for the competing approaches relative to the naïve benchmark. The horizontal axis denotes the MASE of the naïve benchmark.



## References

- An, S & Schorfheide, F. (2007). Bayesian Analysis of DSGE Models. *Econometric Reviews*, **26**(2-4), 113–172
- Bai, J & Ng, S. (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics*, **146**(2), 304–317
- Bañbura, M, Giannone, D & Reichlin, L. (2010). Large Bayesian Vector Auto Regressions. *Journal of Applied Econometrics*, **25**, 71–92.
- Behlul, T, Panagiotelis, A, Athanasopoulos, G, Hyndman, R & Vahid, F. (2017). *Australian Macro Database: An online resource for macroeconomic research in Australia* (Working paper No. 17/01). Monash University.
- Bergmeir, C, Hyndman, RJ & Koo, B. (2015). A Note on the Validity of Cross-Validation for Evaluating Time Series Prediction. *Monash University, Working Paper 10/15*.
- Box, GEP & Jenkins, GM. (1970). *Time Series Analysis: Forecasting and Control*. San Francisco, California, Holden Day.
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, **24**(2), 123–140
- Bühlmann, P & Yu, B. (2002). Analyzing bagging. *Annals of Statistics*, **30**(4), 927–961.
- Burns, AF & Mitchell, WC. (1946). Measuring business cycles. *NBER Books*.
- Carriero, A, Kapetanios, G & Marcellino, M. (2011). Forecasting large datasets with Bayesian reduced rank multivariate models. *Journal of Applied Econometrics*, **26**, 735–761.
- Christiano, LJ, Eichenbaum, M & Evans, CL. (1999). Monetary policy shocks: What have we learned and to what end? In J Taylor & M Woodford (Eds.), *Handbook of macroeconomics* (Vol. 1, pp.65–148). Elsevier, Amsterdam.
- De Mol, C, Giannone, D & Reichlin, L. (2008). Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components? *Journal of Econometrics*, **146**(2), 318–328
- Diebold, FX. (1997). *The past, present, and future of macroeconomic forecasting*. National Bureau of Economic Research.
- Doan, T, Litterman, RB & Sims, CA. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews*, **3**, 1–10.
- Dungey, M & Pagan, A. (2009). Extending a SVAR model of the Australian economy. *Economic Record*, **85**(268), 1–20.
- Efron, B, Hastie, T, Johnstone, I & Tibshirani, R. (2004). Least angle regression. *The Annals of Statistics*, **32**(2), 407–499.

- Eickmeier, S & Ziegler, C. (2008). How successful are dynamic factor models at forecasting output and inflation? a meta-analytic approach. *Journal of Forecasting*, **27**(3), 237–265.
- Fan, J, Liao, Y & Mincheva, M. (2013). Large Covariance Estimation by Thresholding Principal Orthogonal Complements. *Journal of the Royal Statistical Society B*, **75**(4), 603–680.
- Forni, M, Hallin, M, Lippi, M & Reichlin, L. (2000). The generalized dynamic factor model: Identification and estimation. *The Review of Economics and Statistics*, **82**(4), 540–554
- Forni, M, Hallin, M, Lippi, M & Reichlin, L. (2003). Do financial variables help forecasting inflation and real activity in the euro area? *Journal of Monetary Economics*, **50**(6), 1243–1255.
- Geweke, J. (1977). The dynamic factor analysis of economic time series. In D Aigner & A Goldberger (Eds.), *Latent variables in socio economic models*. Amsterdam: North Holland.
- Hastie, T, Tibshirani, R & Friedman, J. (2009). *The Elements of Statistical Learning* (2nd). Springer.
- Hyndman, RJ & Koehler, AB. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, **22**, 679–688.
- Inoue, A & Kilian, L. (2008). How Useful Is Bagging in Forecasting Economic Time Series? A Case Study of U.S. Consumer Price Inflation. *Journal of the American Statistical Association*, **103**(482), 511–522
- Kadiyala, KR & Karlsson, S. (1997). Numerical Methods for Estimation and Inference in Bayesian VAR-models. *Journal of Applied Econometrics*, **12**, 99–132.
- Koop, GM. (2013). Forecasting with medium and large Bayesian VARs. *Journal of Applied Econometrics*, **28**, 117–203.
- Lee, TH & Yang, Y. (2006). Bagging binary and quantile predictors for time series. *Journal of Econometrics*, **135**, 465–497.
- Lindquist, E. (2011). Surveying the World of Visualization. *Working paper, Australian National University*. Retrieved from
- Litterman, RB. (1986). Forecasting with Bayesian Vector Autoregressions – Five years of experience. *Journal of Business and Economic Statistics*, **4**, 25–38.
- Mitchell, WC & Burns, AF. (1938). Statistical indicators of cyclical revivals. In *Statistical indicators of cyclical revivals* (pp.1–12). NBER.
- Sargent, TJ & Sims, CA. (1977). Business cycle modelling without pretending to have too much of prior economic theory. In CA Sims (Ed.), *New methods in business cycle research* (pp.45–100). Federal Reserve Bank of Minneapolis.
- Sims, CA. (1980). Macroeconomics and Reality. *Econometrica*, **48**, 1–48.

Stock, JH & Watson, MW. (2002). Forecasting Using Principal Components from a Large Number of Predictors. *Journal of the American Statistical Association*, **97**(460), 1167–1179

Stock, JH & Watson, MW. (2012). Generalized shrinkage methods for forecasting using many predictors. *Journal of Business and Economic Statistics*, **30**(4), 481–493

Tibshirani, R. (1996). Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society B*, **58**, 267–288.

Zellner, A. (1971). *An introduction to Bayesian inference in econometrics*. Wiley, New York.