International Journal of Forecasting 30 (2014) 996–1015

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Forecasting macroeconomic time series: LASSO-based approaches and their forecast combinations with dynamic factor models

Jiahan Li *, Weiye Chen

Department of Applied and Computational Mathematics and Statistics, University of Notre Dame, 153 Hurley Hall, Notre Dame, IN 46556, USA

ARTICLE INFO

Keywords:
High-dimensional time series
Model selection
Dynamic factor model
Combining forecasts

ABSTRACT

In a data-rich environment, forecasting economic variables amounts to extracting and organizing useful information from a large number of predictors. So far, the dynamic factor model and its variants have been the most successful models for such exercises. In this paper, we investigate a category of LASSO-based approaches and evaluate their predictive abilities for forecasting twenty important macroeconomic variables. These alternative models can handle hundreds of data series simultaneously, and extract useful information for forecasting. We also show, both analytically and empirically, that combing forecasts from LASSO-based models with those from dynamic factor models can reduce the mean square forecast error (MSFE) further. Our three main findings can be summarized as follows. First, for most of the variables under investigation, all of the LASSO-based models outperform dynamic factor models in the out-of-sample forecast evaluations. Second, by extracting information and formulating predictors at economically meaningful block levels, the new methods greatly enhance the interpretability of the models. Third, once forecasts from a LASSO-based approach are combined with those from a dynamic factor model by forecast combination techniques, the combined forecasts are significantly better than either dynamic factor model forecasts or the naïve random walk benchmark.

1. Introduction

The forecasting of macroeconomic variables plays a critical role in macroeconomic studies, financial economics and monetary policy analysis. Accurate forecasts lead to a better understanding of mechanisms of economic dynamics (Bai & Ng, 2008), better portfolio management and hedging strategies (Rapach, Strauss, & Zhou, 2010), and more effective monetary policies (Bernanke, Boivin, & Eliasz, 2005). In the data-rich environment that exists nowadays, large numbers of economic data series are tracked by economists and policy-makers. Low-dimensional models usually incorporate a few pre-specified economic predictors, and thus have difficulty in capturing the complex, dynamic patterns which underlie large panels of time series. Therefore, there is a daunting need to propose econometric models and analysis frameworks which aim to extend their low-dimensional counterparts in order to obtain better predictions.

Over the past decade, the Dynamic Factor Model (DFM; Stock & Watson, 2002a,b) and its variants have been used widely for extracting and organizing useful information from a large number of predictors. These methods summarize a large panel of time series using dynamic factors, make forecasts of dynamic factors, and then recover the dynamics of the original variable using its factor loadings. Motivated by dynamic factor models, Bernanke et al.

* Corresponding author. Tel.: +1 574 631 2741.
E-mail address: jli7@nd.edu (J. Li).

http://dx.doi.org/10.1016/j.ijforecast.2014.03.016
0169-2070/

(2005) proposed the Factor-Augmented Vector Autoregressive (FAVAR) approach for monetary policy analysis. Moench (2008) summarized the risk factors which drive the pricing kernel using dynamic factor models, based on which yield curves are predicted in a no-arbitrage asset pricing framework.

Despite the analytical tractability of dynamic factor models, however, a trade-off has to be made between the loss of information and the curse of dimensionality. In other words, if only the first few principal components are used to summarize the majority of the information in all time series, the remaining principal components could still explain a considerable proportion of the overall variation. However, if more factors are included in the model, the dimensionality of the resulting model increases and the degrees-of-freedom problem arises again. As a result, the number of factors should be limited in order to conserve the degrees-of-freedom, and the risk of losing useful information is hidden behind information compression and dimension reduction. Since such information can hardly be recovered in subsequent steps, an unsatisfactory predictive ability and biased structural inference may follow. Obviously, the larger the number of time series observed and the more heterogeneous these time series are, the more severe the information loss will be. In some recent empirical analyses, dynamic factor models have exhibited a lower predictive power in forecasting some economic indicators than Bayesian shrinkage approaches (see, e.g., Korobilis, 2013).

In this paper, we propose a category of alternative forecasting methods, where a large number of predictors are accommodated simultaneously and shrinkage estimation methods are employed. Within this framework, dimension reduction is *not* carried out before forecasting, but is guided by forecasting, thus avoiding the discard of potentially important information. Specifically, our methods depend on penalized least squares estimation, which is a generalization of ordinary least squares estimation, with an additional term that penalizes the size of regression coefficients. In doing so, it regularizes the model complexity, and avoids over-fitting that can cause the out-of-sample forecasting performance to deteriorate.

Common penalized least squares estimations include LASSO regression (Tibshirani, 1996) and ridge regression (Hoerl & Kennard, 1970), whose individual performances in forecasting economic variables were investigated by De Mol, Giannone, and Reichlin (2008) in a Bayesian framework. They concluded that the two methods produce highly correlated forecasts with similar predictive abilities. In particular, LASSO regressions tend to produce estimated regression coefficients that are exactly zeros, and thus can be used for variable selections, where only predictors with nonzero estimates are considered to be important. In macroeconomic forecasting, such a property has been explored by Bai and Ng (2008) for selecting a subset of predictors, from which factors in dynamic factor models are constructed.

In this paper, however, we will consider several LASSO-based approaches that generalize the classic LASSO regression. First, Zou and Hastie (2005) showed that the variable selection instability of LASSO is due to the parameter uncertainty in estimating a large covariance matrix. They showed that replacing the sample estimator of the covariance matrix with a shrinkage estimator made the resulting regression coefficients and variable selection process more stable. This is equivalent to imposing an additional L2 norm constraint in a LASSO regression problem. In the statistics literature, this method is known as the elastic net, since it is like a net that catches all "big fish" for better forecasts.

Second, since predictors in economic forecasting can be divided into different blocks (Hallin & Liška, 2011; Moench, Ng, & Potter, 2011), we impose sparsity constraints at the block level. This is done with a two-stage procedure. In the first stage, all predictors are grouped into different blocks. Then, in the second stage, group LASSO (Yuan & Lin, 2005) is employed so that predictors in the same block tend to be selected together. As can be seen from the empirical analysis, all LASSO-based approaches have very similar out-of-sample forecast performances, and in general outperform dynamic factor models, but elastic net regression and group LASSO regression give more consistent variable selection results over the whole out-of-sample evaluation period, leading to enhanced model interpretability.

Moreover, motivated by our results that, although LASSO-based approaches have better forecast accuracies in general, dynamic factor models could gain momentum from time to time, we propose to combine the forecasts of LASSO-based models with those of dynamic factor models using forecast combination techniques (Bates & Granger, 1969; Timmermann, 2006). We show analytically that the combined forecasts are associated with smaller mean square forecast errors (MSFE) at the population level in the presence of model uncertainty. Empirically, combined forecasts have significantly lower forecast errors than those from dynamic factor models for all of the economic variables we have predicted, and these forecasts are stabilized over time.

The advantages of these LASSO-based approaches are predictive accuracy and model interpretability. Other than model uncertainty, the forecasting gains can also be explained by the role of non-pervasive shocks. When the true data generating process is unknown, assuming common factors for all variables may ignore shocks that affect a group of variables (or non-pervasive shocks; see Luciani, 2014). As a result, LASSO-based regressions may capture the local correlation that was left behind by common factors in a factor model. As regards model interpretability, it is well known that variables selected by the classic LASSO are not stable over time, in the sense that, once another observation has been added into the estimation window, the estimated regression coefficients and the subset of important predictors may change dramatically. This phenomenon is observed in the statistical analysis of high-dimensional data (Fan & Li, 2001; Zou & Hastie, 2005), as well as in forecasting macroeconomic time series with many predictors (De Mol et al., 2008). Elastic net and group lasso regressions, on the other hand, could deliver relatively stable forecasts and enhanced model interpretation.

The rest of the paper is organized as follows. Section 2 introduces three versions of LASSO-based regressions and their estimation details. Section 3 presents forecast combination techniques that combine forecasts from LASSO-based methods and dynamic factor models. In Section 4 we evaluate the statistical predictabilities of the proposed methods. Section 5 presents the results of an empirical analysis of more than one hundred time series, where twenty macroeconomic variables are predicted by different approaches and the forecasting performances are evaluated and compared. Finally, in Section 6, we provide concluding remarks.

2. Forecasting with shrinkage estimation

2.1. LASSO regression

Assuming that J stationary time series of macroeconomic variables $\boldsymbol{x_t} = \left(x_{1,t}, \ldots, x_{J,t}\right)^T$, $t = 1, \ldots, T$, are observed, with J being very large, our goal is to forecast the kth variable $x_{k,T+h}$ using the forecasting model

$$x_{k,T+h} = \mu_k + \sum_{p=1}^{P} \sum_{j=1}^{J} \phi_{k,j}^{p} x_{j,T-p} + e_{k,T+h}, \tag{1}$$

or

$$x_{T+h} = \mu + \sum_{p=1}^{P} \sum_{j=1}^{J} \phi_{j}^{p} x_{j,T-p} + e_{T+h}, \tag{2}$$

after dropping the subscript for simplicity, where $\phi_{k,j}^{p}$, $p = 1, \ldots, P$ denotes the linear dependency of $x_{k,t}$ on $x_{j,t-p}$, for $j, k = 1, \ldots, J$, and the error term e_{T+h} is stationary and follows a normal distribution with mean zero and variance σ_e^2.

This high-dimensional time series model with a huge number of parameters has been studied before in the literature, with Bayesian shrinkage estimations and factor models being employed (Bai & Ng, 2008; De Mol et al., 2008; Korobilis, 2013). In this study we estimate ϕ_j^p, $j = 1, \ldots, J$, $p = 1, \ldots, P$, by minimizing the constrained least squares

$$\frac{1}{2} \sum_{t=1+P}^{T} \left(x_t - \mu - \sum_{p=1}^{P} \sum_{j=1}^{J} \phi_j^p x_{j,t-p} \right)^2$$

$$\text{subject to } \sum_{p=1}^{P} \sum_{j=1}^{J} |\phi_j^p| < s, \tag{3}$$

for some $s > 0$, which is equivalent to the following penalized least squares for some $\lambda > 0$:

$$\frac{1}{2} \sum_{t=1+P}^{T} \left(x_t - \mu - \sum_{p=1}^{P} \sum_{j=1}^{J} \phi_j^p x_{j,t-p} \right)^2 + \lambda \sum_{p=1}^{P} \sum_{j=1}^{J} |\phi_j^p|. \tag{4}$$

This is the classic LASSO regression (Tibshirani, 1996). When the total number of predictors is large, the ordinary least squares estimators tend to have larger variances, and thus the forecasts are highly volatile over time. This is especially a problem when we have low-frequency economic variables. With a L1 norm constraint on the regression coefficients, the LASSO regression avoids over-fitting and encourages sparse solutions, in the sense that a subset of zero coefficients can be identified and the corresponding predictors excluded from the final model.

From the perspective of Bayesian statistics, minimizing Eq. (4) is equivalent to maximizing the posterior distributions of regression coefficients in a linear regression model, where the prior distributions are independent Laplace distributions (Park & Casella, 2008). De Mol et al. (2008) applied LASSO regression in forecasting using large cross sections. However, our LASSO regression differs from that of De Mol et al. (2008) in two ways. First, we select the tuning parameter λ via cross-validation, which is a data-driven method that is designed to maximize the expected out-of-sample predictive accuracy. Moreover, we also extend the classic LASSO to a scenario with model uncertainty (see Section 3).

2.2. Elastic net regression

Zou and Hastie (2005) showed that, in matrix notation, the objective function in Eq. (4) is

$$\boldsymbol{\phi^T} \left(\boldsymbol{X}^T \boldsymbol{X}\right) \boldsymbol{\phi} - 2\boldsymbol{y^T X \phi} + \lambda |\boldsymbol{\phi}|_1, \tag{5}$$

where $\boldsymbol{y} = (x_{1+P}, \ldots, x_T)^{\boldsymbol{T}}$, $\boldsymbol{\phi} = \left(\phi_1^1, \phi_1^2, \ldots, \phi_J^p\right)^{\boldsymbol{T}}$ is a vector of regression coefficients, $\boldsymbol{X}$ is the corresponding design matrix, with the tth row given by $(x_{1,t-1}, x_{1,t-2}, \ldots, x_{J,t-P})$, $|\cdot|_1$ is the L1 norm of a vector, and λ is the same tuning parameter. Note that $\boldsymbol{X}^T\boldsymbol{X}$ is the sample covariance matrix of J time series. Since J is very large in economic forecasting, the number of unknown parameters in $\boldsymbol{\Sigma}$ is huge. By minimizing Eq. (5), we implicitly assume that $\hat{\boldsymbol{\Sigma}} = \boldsymbol{X}^T\boldsymbol{X}$ is a reasonable estimator of its population counterpart $\boldsymbol{\Sigma}$. However, it is well known that $\hat{\boldsymbol{\Sigma}}$ is far from being the optimal estimator when large numbers of asset prices or economic time series are considered (Kan & Zhou, 2007).

Alternatively, shrinkage estimators of $\boldsymbol{\Sigma}$ have been proposed for replacing sample estimates. In portfolio management, Ledoit and Wolf (2003, 2004) proposed that the large covariance matrix be estimated using a shrinkage estimator

$$\hat{\boldsymbol{\Sigma}}_s = (1-\gamma)\, \hat{\boldsymbol{\Sigma}} + \gamma\, \hat{\boldsymbol{\Sigma}}_{\text{target}}, \tag{6}$$

where $\hat{\boldsymbol{\Sigma}}$ is the sample covariance matrix, $\hat{\boldsymbol{\Sigma}}_{\text{target}}$ is a shrinkage target, and $0 < \gamma < 1$ is the shrinkage intensity. $\hat{\boldsymbol{\Sigma}}_{\text{target}}$ could be either an identity matrix $\boldsymbol{I}$, or the covariance matrix implied by a factor model. In particular, we would like to investigate the benefit of the shrinkage estimator in Eq. (6) for economic forecasting via

$$\hat{\boldsymbol{\phi}}_{\boldsymbol{enet}} = \underset{\boldsymbol{\phi}}{\textbf{argmin}}\, \boldsymbol{\phi^T}((1-\gamma)\hat{\boldsymbol{\Sigma}} + \gamma\hat{\boldsymbol{\Sigma}}_{\text{target}})\boldsymbol{\phi} - 2\boldsymbol{y^T X\phi} + \lambda|\boldsymbol{\phi}|_1, \tag{7}$$

with $\hat{\boldsymbol{\Sigma}}_{\text{target}}$ being an identity matrix. This corresponds to the elastic net shrinkage estimation proposed by Zou and Hastie (2005).

It can be shown that elastic net regression can be regarded as a linear combination of LASSO regression ($\gamma = 0$) and ridge regression ($\gamma = 1$). Since many comparisons have been made between LASSO regression and ridge regression, with the conclusion that neither of them could uniformly dominate the other in the out-of-sample forecasting exercises (De Mol et al., 2008; Fu, 1998; Tibshirani, 1996), it seems reasonable to combine these two approaches adaptively. Ridge regression alone does not zero out any predictors, but shrinks all predictors. By combining it with LASSO regression, however, we could select a subset of variables using this flexible shrinkage estimation approach. Two shrinkage intensities here, γ and λ, are determined by cross-validation.

2.3. Group-LASSO regression

It has been recognized that large panels of economic data tend to have a block structure, where the variables within each block are similar economic measures. Examples of blocks include variables representing employment rates or price levels. However, since both LASSO regressions and elastic net regressions select variables individually, interpreting the final predictive model is not straightforward. Thus, it would be advantageous if such a block structure could be used, at least for better interpretations. Currently, block information has been incorporated into the construction of dynamic factors (Hallin & Liška, 2011; Moench et al., 2011). Song and Bickel (2011) proposed the application of a group LASSO penalty in the context of large vector autoregressions, where regression coefficients in the same group are shrunk to zero jointly. Their study proposed two different grouping strategies: universal grouping and segmented grouping. With universal grouping, all regression coefficients which use a given predictor to forecast different variables are defined as one group. In our problem of forecasting a single dependent variable, this strategy is equivalent to a simple LASSO penalty. Segmented grouping, on the other hand, uses a similar strategy, but estimates all regression coefficients segment by segment, where a segment has the same definition as our economically meaningful blocks. For example, in predicting a single dependent variable in one segment, this strategy is equivalent to a standard LASSO regression with two tuning parameters, one for coefficients in this segment and one for those outside.

For final model interpretability, we propose to apply the group LASSO penalty to groups of predictors, so that predictors in the same economic block could either enter or leave the final model together. This is done via a two-stage procedure. In the first stage, all predictors are grouped into different blocks, where either economic knowledge or statistical methods could be used. Then, in the second stage, sparsity constraints are imposed at the block level. In particular, we employ a group LASSO penalized regression (Yuan & Lin, 2005), so that predictors in the same block tend to be selected or excluded together in the final predictive regression. In this way, it is expected that both sparsity and model interpretability will be gained.

Suppose that all J predictors can be partitioned into L groups, with d_l being the number of predictors in group $l = 1, \dots, L$. Then, all regression coefficients ϕ_j^p, $j = 1, \dots, J$, $p = 1, \dots, P$, can be partitioned accordingly as $\{\boldsymbol{\phi}_1^1, \dots, \boldsymbol{\phi}_L^1, \dots, \boldsymbol{\phi}_L^P\}$, where the d_l-dimensional vector $\boldsymbol{\phi}_l^p$ contains all regression coefficients for variables in group l at lag p, $l = 1, \dots, L$, $p = 1, \dots, P$. We then estimate all regression coefficients by minimizing

$$\frac{1}{2}\sum_{t=1+P}^{T}\left(x_t - \mu - \sum_{p=1}^{P}\sum_{j=1}^{J}\phi_j^p x_{j,t-p}\right)^2 + \lambda\sum_{p=1}^{P}\sum_{l=1}^{L}\sqrt{d_l}\,\left\|\boldsymbol{\phi}_l^p\right\|_2, \tag{8}$$

where $\left\|\boldsymbol{\phi}_l^p\right\|_2$ is the L2 norm (not squared) of the vector $\boldsymbol{\phi}_l^p$. It can be shown that such a penalty encourages sparsity at the block level (see Yuan & Lin, 2005, for more discussions).

2.4. Computation and selection of tuning parameters

For all LASSO-based shrinkage methods, there is no direct solution. However, Yuan and Lin (2005) and Zou and Hastie (2005) have shown that the elastic net regression in Eq. (7) and the group LASSO regression in Eq. (8) can be transformed to LASSO regressions that take similar forms to that in Eq. (4). Several efficient algorithms have been proposed for LASSO regression, such as the least angle regression (Efron, Hastie, Johnstone, & Tibshirani, 2004) and the coordinate descent algorithm (Friedman, Hastie, & Tibshirani, 2010; Fu, 1998; Wu & Lange, 2008). Here, we will employ the coordinate descent algorithm, which is suitable for problems with large numbers of predictors.

Note that the strength of the shrinkage in Eqs. (4) and (8) depends on one tuning parameter λ, while that in Eq. (7) depends on two tuning parameters, γ and λ, with larger values of γ and λ corresponding to higher levels of shrinkage on the regression coefficients. In what follows, we describe how to determine these tuning parameters via cross-validations, which are expected to balance a model's in-sample fit and out-of-sample predictive ability.

This data-driven method has been used extensively in statistics (Arlot & Celisse, 2010) and finance (DeMiguel, Garlappi, Nogales, & Uppal, 2009). Given a combination of fixed values of γ and λ, say $(\gamma^{(1)}, \lambda^{(1)})$, we randomly partition the in-sample data into five pieces of roughly the same size. Then, for each piece, say piece k, we estimate the model with $\gamma = \gamma^{(1)}$ and $\lambda = \lambda^{(1)}$ using all data that are not included in piece k. Suppose that the estimated regression coefficients are $\hat{\boldsymbol{\phi}}^{(\boldsymbol{k})}$; we then go back to piece k, and make forecasts using the predictive regression, with the unknown parameters $\boldsymbol{\phi}$ being replaced by $\hat{\boldsymbol{\phi}}^{(\boldsymbol{k})}$. Since $\hat{\boldsymbol{\phi}}^{(\boldsymbol{k})}$ is estimated using data other than piece k, the forecasts of this part of data can be regarded as out-of-sample forecasts. This procedure is repeated for $k = 1, \dots, 5$, so that each observation in each of these five pieces receives its own out-of-sample forecast. We calculate the mean squared prediction error (MSPE) that is conditional on $(\gamma^{(1)}, \lambda^{(1)})$. Finally, the whole procedure is repeated for different values of γ and λ, and the combination that gives the lowest mean squared prediction error (MSPE) is used.[1] In order to forecast the variable of interest in the

[1] The sequence of each tuning parameter is specified by taking 100 tuning parameter values between 0 and an upper bound, where the upper bound is the smallest tuning parameter value whereby all regression coefficients are shrunk to zero, and is found by trial and error. Empirically, the impact of the upper bound on the estimates is negligible.

next month, we collect all of the in-sample data above, and use the selected tuning parameters.

3. Forecast combination of LASSO-based approaches and dynamic factor models

In general, all LASSO-based approaches and dynamic factor models are dimension reduction methods for forecasting using large numbers of time series. However, the method of dimension reduction varies. While LASSO-based approaches try to identify and aggregate individual predictors that are relevant to the variable to be predicted, the dynamic factors summarize all of the information using $\tilde{J}$ principal components before forecasting. As a result, if only a few predictors contain information for forecasting, LASSO-based methods could eliminate irrelevant predictors from the final predictive model, meaning that they are robust to cumulative noise in large data sets. However, if the variable of interest is driven by the co-movements of many time series, dynamic factor models are expected to find and characterize the underlying common factors, and outperform LASSO-based methods.

"Essentially, all models are wrong, but some are useful" (Box, 1976). This is true in economic forecasting as well. As our empirical analyses in Section 5 confirm, although LASSO-based approaches deliver better out-of-sample forecasts in general, they do not dominate dynamic factor models for all variables over the full evaluation period. We therefore propose to resolve this model uncertainty by forecast combination. Forecast combination was originally proposed by Bates and Granger (1969). Working with forecasts from individual models, it formulates a linear combination of all forecasts as a new forecast. Recently, forecast combination has received a lot of attention in the areas of economics and finance, due to its superior out-of-sample predictive ability and stabilized forecast errors (Della Corte & Tsiakas, 2012; Rapach et al., 2010; Stock & Watson, 2004; Timmermann, 2006).

3.1. Theoretical properties

Aiolfi, Capistrán, and Timmermann (2010) considered a model uncertainty specification where x_t is assumed to be generated by one of two factor models according to a Bernoulli random variable. In this section, we show analytically how a combination of the forecasts from a factor model and a sparse model can dominate two individual forecasts, where only a subset of predictors have nonzero regression coefficients in the sparse model, and LASSO-based approaches are used to identify these important variables.

Suppose that there is a shift in the data generating process of the variable to be predicted, x_t, between a factor model and a sparse model:

$$x_t = \mathrm{I}_{\{S_t=1\}} L^T F_t + \mathrm{I}_{\{S_t=2\}} \Phi^T Z_{t-1} + e_t, \tag{9}$$

where $\mathrm{I}_{\{A\}}$ is an indicator function that is equal to one if event A is true and zero otherwise, s_t is a Bernoulli random variable indicating two states with $\Pr(S_t = 1) = p$ and $\Pr(S_t = 2) = 1 - p,\ p \in (0, 1), F_t = (f_{1,t}, \ldots, f_{m_1,t})^T$ is an m_1-dimensional vector of factors, Z_{t-1} is a vector of lagged predictors, and e_t is a random error with mean zero and variance σ_e^2. According to these dynamics, x_t is generated by two different models in two different states. In state 1, x_t is governed by a vector of factors, where L^T is a vector of factor loadings; in state 2, x_t is determined by m_2 lagged predictors $Z_{t-1} = (x_{k_1,t-1}, \ldots, x_{k_{m_2},t-1})^T,\ k_j \in \{1, \ldots, J\}$, $j = 1, \ldots, m_2$, with Φ^T being an m_2-dimensional vector of regression coefficients. Since Z_{t-1} only contains a subset of all predictors in this true model, it is known as a sparse model. In practice, lagged predictors in Z_{t-1} are identified by LASSO-based methods.

We further assume the dynamics of F_t and Z_t to be

$$F_t = A_1 F_{t-1} + \epsilon_1, \tag{10}$$

and

$$Z_t = A_2 Z_{t-1} + \epsilon_2, \tag{11}$$

with $\epsilon_1 \sim N(0, \Sigma_1), \epsilon_2 \sim N(0, \Sigma_2)$ and $\mathrm{cov}(\epsilon_1, \epsilon_2) = \Sigma_{12}$. It can be shown (see Appendix A) that when a dynamic factor model is used for prediction, the associated mean squared forecast error (MSFE) is

$$\begin{aligned}\mathrm{MSFE}\left(\hat{x}_t^{(D)}\right) &= p\left(L^T - \beta_1^T\right)\Sigma_1(L - \beta_1) + (1-p) \\ &\quad \times (\Phi^T \Sigma_2 \Phi + \beta_1^T \Sigma_1 \beta_1 - 2\Phi^T \Sigma_{12} \beta_1) + \sigma_e^2,\end{aligned} \tag{12}$$

where

$$\beta_1 = pL + (1-p)\Sigma_1^{-1}\Sigma_{12}\Phi. \tag{13}$$

On the other hand, when a sparse model estimated by a LASSO-based approach is used for prediction, the associated MSFE is

$$\begin{aligned}\mathrm{MSFE}\left(\hat{x}_t^{(L)}\right) &= p(L^T \Sigma_1 L + \beta_2^T \Sigma_2 \beta_2 - 2L^T \Sigma_{12} \beta_2) \\ &\quad + (1-p)\left(\Phi^T - \beta_2^T\right)\Sigma_2(\Phi - \beta_2) + \sigma_e^2,\end{aligned} \tag{14}$$

where

$$\beta_2 = p\Sigma_2^{-1}\Sigma_{12}L + (1-p)\Phi. \tag{15}$$

Similarly, the MSFE of the equal-weighted, simple combined forecast $\hat{x}_t^{(C)} = \frac{1}{2}\left(\hat{x}_t^{(D)} + \hat{x}_t^{(L)}\right)$ is

$$\begin{aligned}\mathrm{MSFE}(\hat{x}_t^{(C)}) = {} & p[(L - \beta_1/2)^T \Sigma_1 (L - \beta_1/2) \\ & + (\beta_2/2)^T \Sigma_2 (\beta_2/2) - (L - \beta_1/2)^T \\ & \times \Sigma_{12}\beta_2] + (1-p)[(\Phi - \beta_2/2)^T \\ & \times \Sigma_2(\Phi - \beta_2/2) + (\beta_1/2)^T \Sigma_1 (\beta_1/2) \\ & - (\Phi - \beta_2/2)^T \Sigma_{12}\beta_1] + \sigma_e^2.\end{aligned} \tag{16}$$

These expressions are quite general. To gain more insights, we simplify the model to the case where $m_1 = m_2 = 1, A_1 = A_2 = 0$, and both L and Φ are identity matrices. This gives $F_t \sim N\left(0, \sigma_1^2\right), Z_t \sim N\left(0, \sigma_2^2\right)$, and $\mathrm{cov}(F_t, Z_{t-1}) = \sigma_{12}$. As a result, the MSFE associated with factor model forecasts is

$$\begin{aligned}\mathrm{MSFE}\left(\hat{x}_t^{(D)}\right) &= \left[p(1-\beta_1)^2 + (1-p)\beta_1^2\right]\sigma_1^2 \\ &\quad + (1-p)\sigma_2^2 - 2(1-p)\beta_1\sigma_{12} + \sigma_e^2.\end{aligned} \tag{17}$$

The MSFE of forecasts from the sparse model is

$$\mathrm{MSFE}\left(\hat{x}_t^{(L)}\right) = p\sigma_1^2 + \left[p\beta_2^2 + (1-p)(1-\beta_2)^2\right]\sigma_2^2 - 2p\beta_2\sigma_{12} + \sigma_e^2, \tag{18}$$

and the MSFE of the simple combined forecast $\hat{x}_t^{(C)}$ is

$$\begin{aligned}\mathrm{MSFE}\left(\hat{x}_t^{(C)}\right) = &\left[p(1-\beta_1/2)^2 + (1-p)\beta_1^2/4\right]\sigma_1^2 \\ &+ \left[p\beta_2^2/4 + (1-p)(1-\beta_2/2)^2\right]\sigma_2^2 \\ &- \left[p(1-\beta_1/2)\beta_2 + (1-p)\beta_1 \right.\\ &\left.\times (1-\beta_2/2)\right]\sigma_{12} + \sigma_e^2.\end{aligned} \tag{19}$$

For simplicity, suppose that $\sigma_{12} = 0$. Then, by plugging in the definitions of β_1 and β_2, the population MSFE of the equal-weighted combined forecast, $\mathrm{MSFE}\left(\hat{x}_t^{(C)}\right)$, will be lower than those of the individual models if the following condition holds:

$$\frac{1}{3}\left(\frac{p}{1-p}\right)^2 < \frac{\sigma_2^2}{\sigma_1^2} < 3\left(\frac{p}{1-p}\right)^2, \tag{20}$$

where p is the probability of being in state 1. Fig. 1 shows how the MSFEs of three models change with σ_2^2/σ_1^2 and p. Clearly, as long as two models have comparable predictive accuracies (all three scenarios), and the probability of being in one state is not extreme, the forecast combination has the lowest MSFE.

3.2. Empirical implementation

In empirical applications, we estimate the combined forecast $\hat{x}_{T+h}^{(C)}$ as the weighted average of the forecast from a LASSO-based model, $\hat{x}_{T+h}^{(L)}$, and that from a dynamic factor model, $\hat{x}_{T+h}^{(D)}$:

$$\hat{x}_{T+h}^{(C)} = w_{T+h}\hat{x}_{T+h}^{(L)} + (1-w_{T+h})\hat{x}_{T+h}^{(D)}, \tag{21}$$

where $0 \le w_{T+h} \le 1$ is the weight of the LASSO-based forecast. For a simple combined forecast, $w_{T+h} = 0.5$. Alternatively, w_{T+h} can also be determined by the historical relative performance of the LASSO-based method. Following Stock and Watson (2004), among others, we measure the past performance of each model by the inverse of its Discounted MSE (DMSE), or

$$\mathrm{DMSE}_{T+h}^{(L)} = \sum_{q=T-\kappa}^{T-1} \theta^{T-1-q}\left(x_q - \hat{x}_q^{(L)}\right)^2, \tag{22}$$

where κ is the window size, so that DMSE is evaluated based on the κ most recent forecasts, θ is the discount factor, and $x_q - \hat{x}_q^{(L)}$ is the forecasting error of the LASSO-based method within that window. Similarly,

$$\mathrm{DMSE}_{T+h}^{(D)} = \sum_{q=T-\kappa}^{T-1} \theta^{T-1-q}\left(x_q - \hat{x}_q^{(D)}\right)^2. \tag{23}$$

Since the lower the DMSE_{T+h}, the better the predictive ability of a model before time T, the weight of the LASSO-based method in the combined forecast in Eq. (21) is proportional to the inverse of $\mathrm{DMSE}_{T+h}^{(L)}$:

$$w_{T+h} = \frac{\left(\mathrm{DMSE}_{T+h}^{(L)}\right)^{-1}}{\left(\mathrm{DMSE}_{T+h}^{(L)}\right)^{-1} + \left(\mathrm{DMSE}_{T+h}^{(D)}\right)^{-1}}. \tag{24}$$

When the discount factor $\theta = 1$ (no discounting), the forecast combination method is the one investigated by Bates and Granger (1969). Following Della Corte and Tsiakas (2012), in addition to simple forecast combinations, we also consider discount factors $\theta = \{1, 0.95, 0.9\}$, and evaluate past out-of-sample performances over the last κ months, where $\kappa = \{12, 36, 60\}$.

4. Evaluation of statistical predictability

Since the performance of the dynamic factor model has been superior for forecasting economic variables, outperforming univariate autoregressions, small vector autoregressions, and leading indicator models (Stock & Watson, 2002a), we will use the dynamic factor model as a benchmark against which to compare the proposed strategies.

In what follows, we describe various statistical criteria for comparing the out-of-sample predictive abilities of alternative approaches. First, we compute the root mean squared error (RMSE) difference statistic, ΔRMSE, which was used by Della Corte and Tsiakas (2012) and Welch and Goyal (2008) for evaluating the predictability of asset pricing models. We define

$$\Delta\mathrm{RMSE} = \sqrt{\frac{\sum_{t=T+1}^{M}(x_t - \bar{x}_t)^2}{M-T}} - \sqrt{\frac{\sum_{t=T+1}^{M}(x_t - \hat{x}_t)^2}{M-T}}, \tag{25}$$

where $M-T$ is the number of out-of-sample forecasts, x_t is the observed value of the variable being predicted at time t, $t = T+1, \ldots, M$, $\bar{x}_t$ is the forecast from a benchmark model, and $\hat{x}_t$ is the forecast from a proposed model. A positive ΔRMSE suggests that the proposed model outperforms the benchmark (the dynamic factor model).

In addition to comparing the RMSEs directly, statistical tests are also employed for testing the significance of the differences between forecasts. Since the dynamic factor model and LASSO-based approaches are non-nested, we employ the Diebold–Mariano test (Diebold & Mariano, 1995) for comparing the forecast accuracies of two methods. Specifically, if one defines the difference in squared forecast errors as

$$d_t = e_{1,t}^2 - e_{2,t}^2, \quad t = T+1, \ldots, M, \tag{26}$$

then the Diebold–Mariano test statistic is

$$\mathrm{DM} = \left[\frac{1}{M-T}\left(\hat{\gamma}_0 + 2\sum_{k=1}^{h-1}\hat{\gamma}_k\right)\right]^{-1/2} \times \frac{1}{M-T}\sum_{t=T+1}^{M} d_t, \tag{27}$$

where $\hat{\gamma}_k$, $k = 0, 1, \ldots, k$ is the estimated kth autocovariance of the time series d_t, and h is the forecast horizon.

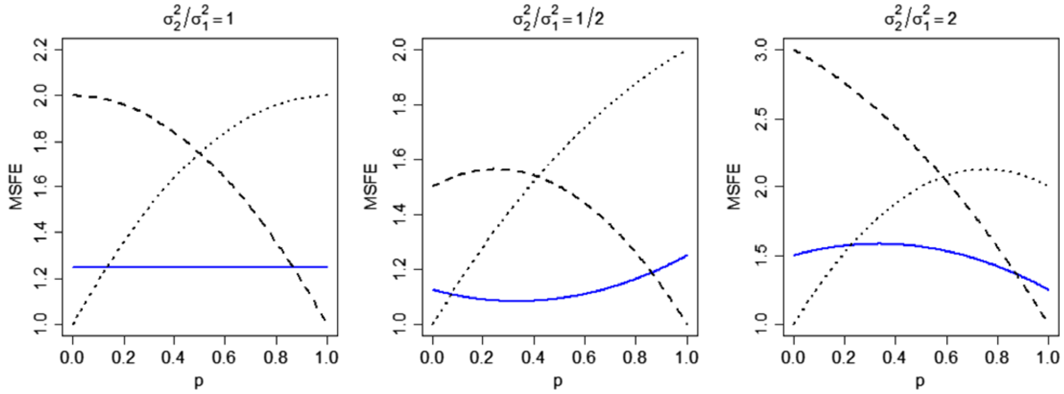

Fig. 1. Mean squared forecast errors (MSFE) from a factor model (dashed line – – –), a sparse model (dotted line · · ·), and the forecast combination (blue solid line –) as a function of p.

Under the null hypothesis of an equal expected forecast performance, this statistic has an asymptotic standard normal distribution. Such a standard asymptotic distribution of the test statistic is advantageous here. Since the cross-validation used for selecting tuning parameters is a resampling scheme, hypothesis tests that are free of bootstrapping are desired, in order to maintain affordable computational costs in the out-of-sample evaluation.

When testing the null hypothesis that the forecast combination and the dynamic factor model have equal predictive abilities, however, the problem involves testing two nested models. We apply the testing procedure developed recently by Clark and West (2006, 2007), which acknowledges that, even if two models have the same forecasting accuracy at the population level, the RMSE from a larger model is expected to be higher. This test statistic has been used in evaluating the forecasting accuracies of alternative foreign exchange rate models (Della Corte & Tsiakas, 2012), as well as of alternative asset pricing models for stock returns (Rapach, Strauss, & Zhou, 2013). Specifically, given the time series of forecasting errors of two models, $e_{1,t}$ and $e_{2,t}$, we define

$$\hat{f}_t = e_{2,t}^2 - e_{1,t}^2 + \left(e_{1,t} - e_{2,t}\right)^2. \tag{28}$$

Then, the test statistic MSE-t is defined as the t-statistic of the intercept in the following linear regression without independent variables

$$\hat{f}_t = \alpha + e_t. \tag{29}$$

Clark and West (2006, 2007) showed that the standard normal distribution could approximate closely the distribution of the test statistic MSE-t, based on which an associated p-value can be obtained for hypothesis testing. Again, the null hypothesis is that the two nested models have equal predictive abilities.

5. Empirical results

5.1. Data and estimation

The data set for empirical analysis contains 107 macroeconomic indicators, including industrial production, price levels, credit conditions, employment and inflation. This monthly data set was used by Stock and Watson (2012), and can be downloaded from Mark Watson's website:[2] http://www.princeton.edu/~mwatson/publi.html. According to the common practice in the macroeconomic literature, these data series have been transformed to be stationary. The descriptions of all data series, together with their corresponding transformations, are provided in Appendix B.

We examine the out-of-sample forecasting performances of the proposed forecasting strategies in a pseudo out-of-sample forecasting exercise as follows. Following Bai and Ng (2008) and Stock and Watson (2002b), the first out-of-sample forecast is based on a 10-year (120-month) estimation window. Although the original data series are from January 1959 to December 2008, the variable transformations and the inclusion of up to 4 lags in the predictors mean that the estimation window starts at July 1959, and the first set of variables to be predicted is in July 1969. Then, for any month t between July 1969 and December 2008 (474 months in total), we define an in-sample period which includes all data series in the previous 10 years. Based on the in-sample data, we estimate a predictive regression with the maximum lag order $P = 4$. Next, we produce one-step-ahead forecasts with the estimated model. Finally, we drop the first observation in the rolling estimation window and include a new one at the end. This whole process is repeated until the end of the out-of-sample evaluation period.

LASSO-based predictive regressions are estimated according to the discussions in the previous sections. For the group LASSO regression, in particular, we partition the 107 variables according to their economic meanings (see Appendix B). Unreported statistical tests indicate that the variables within each block are highly correlated, which is consistent with the findings of Moench et al. (2011). In terms of dynamic factor models, we follow the specification of Bernanke et al. (2005) by considering three dynamic factors and lag order $P = 4$.

We examine the out-of-sample forecasting performance in detail for twenty representative macroeconomic

[2] We are grateful to Mark Watson for making this data set available.

Table 1
Out-of-sample root mean square error (RMSE) differentials (ΔRMSE) of forecasts from LASSO-assisted regressions and dynamic factor models (DFM).

	LASSO vs. DFM		ENET vs. DFM		G-LASSO vs. DFM	
	ΔRMSE	*p*-value	ΔRMSE	*p*-value	ΔRMSE	*p*-value
Federal funds rate	0.017	0.339	0.023	0.330	−0.012	0.597
Industrial production	0.054	**0.024**	0.061	**0.006**	0.031	0.161
CPI	0.114	**0.001**	0.126	**0.000**	0.129	**0.000**
3m Treasury bills	−0.058	0.762	−0.034	0.693	−0.056	0.778
5y Treasury bonds	−0.009	0.582	0.009	0.411	−0.012	0.607
M1	0.123	**0.022**	0.050	0.288	0.073	0.206
M2	0.099	**0.024**	0.096	**0.019**	0.069	0.073
Exchange rate: Japan	0.024	0.153	0.017	0.245	0.034	0.085
Commodity price index	0.558	**0.000**	0.551	**0.000**	0.413	**0.000**
Capacity utilization	0.381	**0.000**	0.380	**0.000**	0.091	**0.000**
Personal consumption	0.112	**0.009**	0.102	**0.007**	0.120	**0.005**
Durable consumption	0.068	**0.002**	0.063	**0.004**	0.068	**0.001**
Nondurable consumption	0.038	0.082	0.051	**0.045**	0.118	**0.000**
Unemployment rate	0.019	0.192	0.045	**0.016**	−0.002	0.533
Employment rate	0.006	0.365	0.009	0.305	−0.024	0.867
Avg hourly earnings	0.296	**0.000**	0.285	**0.000**	0.275	**0.000**
Housing starts	0.355	**0.000**	0.350	**0.000**	0.233	**0.000**
New orders index	0.228	**0.000**	0.223	**0.000**	0.215	**0.000**
Price/dividend ratio	0.086	**0.002**	0.084	**0.002**	0.069	**0.009**
Consumer expectations	0.028	0.112	0.033	0.059	0.022	0.166

The LASSO-assisted regressions include the LASSO regression, the elastic net regression (ENET) and the group lasso regression (G-LASSO), where the tuning parameters are selected through cross validations for each month over the out-of-sample evaluation period. The forecasting accuracies are tested formally using the Diebold–Mariano test, where the *p*-values associated with the test statistics are reported. *p*-values which indicate significant differences in predictive abilities are given in bold (the significance level is 0.05).

indicators that were investigated carefully by Bernanke et al. (2005), which include the federal funds rate (FFR), industrial production (IP), the consumer price index for all urban consumers (CPI), and other measures of price levels, real activities and consumption. These variables include important indicators for the economy and monetary policies.

5.2. Predictive accuracy

In Table 1 we report the out-of-sample root mean square error (RMSE) differentials (ΔRMSE) of forecasts from LASSO-based approaches and dynamic factor models (DFM), where the former category includes the LASSO regression (columns 1–2), the elastic net regression (columns 3–4) and the group LASSO regression (columns 5–6), and the tuning parameters are selected by cross validations for each month throughout the out-of-sample evaluation period. For each comparison of each variable, the Diebold–Mariano test is conducted and the associated *p*-value is reported, with *p*-values less than 0.05 indicating a rejection of the null of equal predictive abilities at the significance level of 0.05.

By comparing the forecasts from LASSO-based approaches with those from dynamic factor models, we conclude that, in general, LASSO-based approaches perform better out-of-sample. For example, when using LASSO regressions, 18 variables out of 20 have positive ΔRMSE values, of which 12 are statistically significant. The other two methods give 19 and 15 positive ΔRMSE values, respectively, of which more than 65% are statistically significant. Interestingly, the variables whose forecasts from two models are not significantly different include interest rate related series (Federal Funds Rate, 3M Treasury Bills, 5Y Treasury Bonds), employment/unemployment rates and the index of consumer expectations. Since the dynamic factors measured by the first few principal components could capture the most obvious data variations over time, we conclude that such variations provide satisfactory summaries of the information that is useful for forecasting these variables.

Moreover, all models have comparable performances in forecasting foreign exchange rates. It has been documented that foreign exchange rate movements are not connected with economic fundamentals, a fact which is known as the "exchange rate disconnect puzzle" (Meese & Rogoff, 1983; Obstfeld & Rogoff, 2001) in the international economics literature. Thus, the best foreign exchange rate model at short horizons is the random walk model which predicts the future foreign exchange rate by its historical mean. On the other hand, in the economic forecasting literature, many works have focused on forecasting industrial production and the consumer price index (CPI). It can be seen that our new methods perform well for forecasting these two economic aggregates.

Table 2 presents similar forecast performance measures that compare forecasts combinations and dynamic factor models. Forecast combinations combine forecasts from the dynamic factor model with forecasts from one of the LASSO-based approaches, including the LASSO regression, elastic net regression and group LASSO regression, where a discount factor of $\theta = 0.9$ and the most recent $\kappa = 60$ forecasts are used. Since the forecast combination and the dynamic factor model are nested in this scenario, the improvements in forecast accuracies are tested formally via MSE-*t* statistics (Clark & West, 2006, 2007), and the *p*-values associated with the test statistics are reported. It can be seen that once the forecasts are combined, the ΔRMSE values are positive for all variables, no matter which LASSO-based approach is considered. Moreover, the improvements in forecast accuracies are statistically

Table 2
Out-of-sample root mean square error (RMSE) differentials (ΔRMSE) of forecasts from forecast combinations and dynamic factor models (DFM).

	Combination 1 vs. DFM		Combination 2 vs. DFM		Combination 3 vs. DFM	
	ΔRMSE	*p*-value	ΔRMSE	*p*-value	ΔRMSE	*p*-value
Federal funds rate	0.041	**0.009**	0.038	**0.016**	0.030	**0.011**
Industrial production	0.068	**0.000**	0.063	**0.000**	0.062	**0.000**
CPI	0.102	**0.000**	0.107	**0.000**	0.111	**0.000**
3m Treasury bills	0.012	0.077	0.017	0.068	0.022	**0.023**
5y Treasury bonds	0.028	**0.001**	0.033	**0.000**	0.030	**0.001**
M1	0.113	**0.004**	0.069	**0.012**	0.102	**0.016**
M2	0.084	**0.003**	0.077	**0.002**	0.067	**0.008**
Exchange rate: Japan	0.034	**0.000**	0.027	**0.001**	0.039	**0.000**
Commodity price index	0.530	**0.000**	0.523	**0.000**	0.378	**0.000**
Capacity utilization	0.368	**0.000**	0.367	**0.000**	0.185	**0.000**
Personal consumption	0.085	**0.001**	0.077	**0.000**	0.095	**0.000**
Durable consumption	0.056	**0.000**	0.053	**0.000**	0.056	**0.000**
Nondurable consumption	0.042	**0.000**	0.052	**0.000**	0.091	**0.000**
Unemployment rate	0.038	**0.000**	0.051	**0.000**	0.033	**0.000**
Employment rate	0.026	**0.000**	0.024	**0.000**	0.019	**0.000**
Avg hourly earnings	0.245	**0.000**	0.235	**0.000**	0.228	**0.000**
Housing starts	0.339	**0.000**	0.335	**0.000**	0.253	**0.000**
New orders index	0.206	**0.000**	0.199	**0.000**	0.202	**0.000**
Price/dividend ratio	0.079	**0.000**	0.077	**0.000**	0.073	**0.000**
Consumer expectations	0.033	**0.000**	0.034	**0.000**	0.026	**0.002**

The forecast combinations combine forecasts from LASSO-assisted regressions and dynamic factor models, where the former category includes the LASSO regression (combination 1), the elastic net regression (combination 2) and the group LASSO regression (combination 3), with the tuning parameters being selected by cross validation, and $\theta = 0.9$ and $\kappa = 60$. The forecasting accuracies are tested by MSE-t statistics (Clark & West, 2006, 2007), with the *p*-values associated with the test statistics being reported. *p*-values which indicate significant differences in predictive abilities are given in bold (the significance level is 0.05).

significant for at least 19 variables, as the MSE-t statistics show. For this data set, the simple combined forecast and other specifications of θ and κ present very similar results, and thus are not reported. We conclude that forecast combination is an effective strategy for forecasting macroeconomic time series when large numbers of predictors are available. Due to the statistical properties of principal components, the dynamic factor model is capable of capturing the large variations underlying multiple time series, and predicting a variable based on its connections with the temporal variation. LASSO-based approaches, on the other hand, are capable of identifying important predictors for recovering a sparse predictive model. Once the forecasts from two strategies are combined, model uncertainty is resolved and predictive power is gained. This is consistent with our analytical results in Section 3.1.

Furthermore, we show the extent to which the forecasts produced by different methods differ by reporting the correlations among these forecasts (see Table 3). First of all, the forecasts from LASSO regressions and elastic net regressions are highly correlated for almost all variables. However, their forecasts are less correlated with the forecasts from group LASSO regressions (columns 2–3). As a result, the forecasts from LASSO regressions and elastic net regressions are correlated with those produced by dynamic factor models to a similar extent (columns 4–5), but the forecasts from group LASSO regressions and dynamic factor models are slightly less correlated (column 6). This may indicate that group LASSO regressions and dynamic factor models provide complementary information, which explains the superior out-of-sample performance of their combined forecasts.

It can sometimes be interesting to see how predictive regressions perform over time. In Fig. 2 we plot the evolution of ΔRMSE for each variable over the full sample period, where the forecasts are produced by two models: LASSO regressions and dynamic factor models. At a given time point, a positive ΔRMSE implies that the LASSO regression has better forecasts so far. Meanwhile, an increasing trend of ΔRMSE implies that, at this point in time, the forecast from the LASSO regression has a smaller RMSE, and thus is better. Since these patterns are similar for all LASSO-based approaches, those for the elastic net regression and group LASSO regression are not reported. In Fig. 3 we further visualize the dynamics of ΔRMSE when the alternative model is the forecast combination of the LASSO regression and the dynamic factor model. It can be seen that variables with negative ΔRMSEs over the full evaluation period in Table 1 usually exhibit sharp decreases in Fig. 2 at some time points. However, once forecasts from two models are combined, the severity of such decreases is ameliorated, and ΔRMSE is smoothed over time.

Usually, forecasters are concerned about how consistently forecasts are close to the true values. We answer this question using two measures: the maximum RMSE and the variance of RMSE. In Table 4, we report the ratio of the maximum RMSE from various alternative models to that from dynamic factor models, where the alternative models include the LASSO regression, elastic net regression, group LASSO regression, and the forecast combinations of LASSO-based regressions and dynamic factor models. Ratios less than one are in bold, and suggest that the alternative model produces less extreme forecasting errors. Table 5 further reports the ratios of RMSE variances from the same pairs of models, where ratios less than one imply that the predictive ability of the alternative model is more stable over time.

Tables 4 and 5 both confirm that all six of the newly proposed strategies generally produce less extreme and more stable forecast errors. Moreover, although the combined

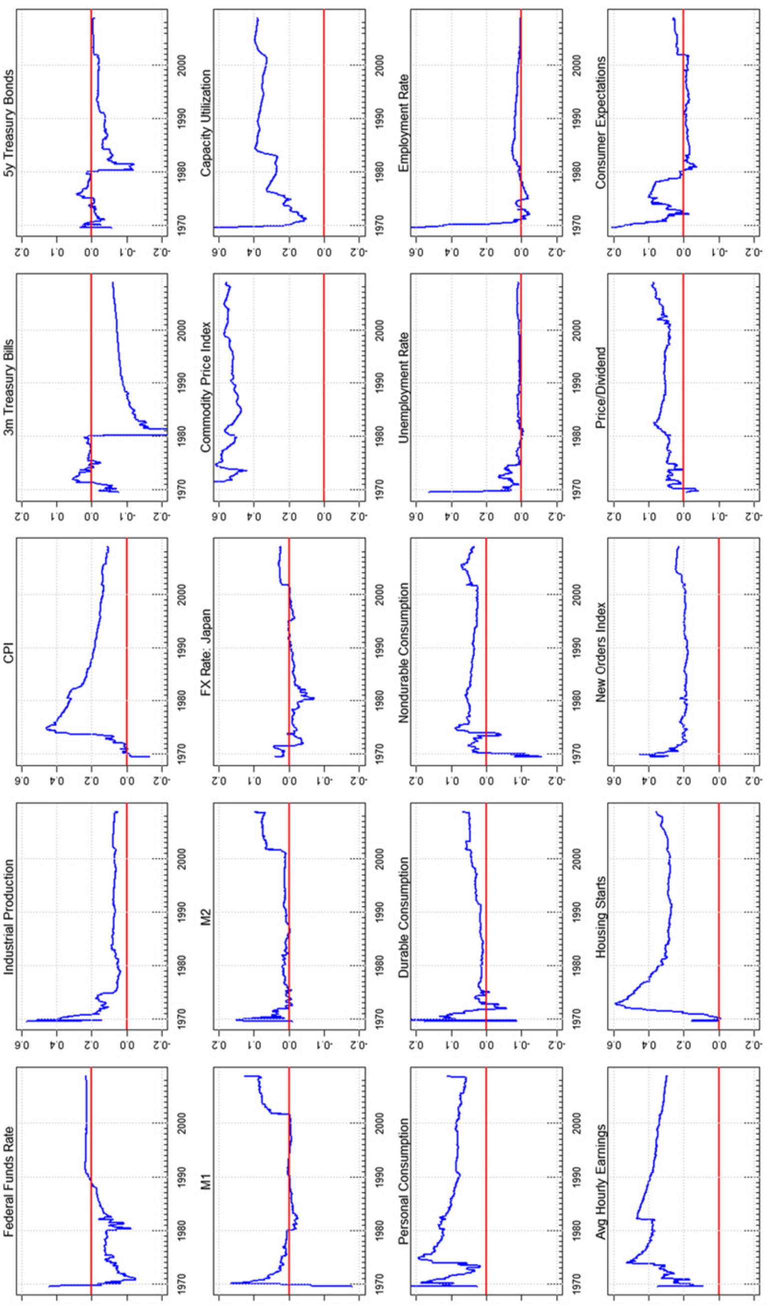

Fig. 2. Out-of-sample root mean square error (RMSE) differentials (ΔRMSE) over time, where the forecasts are from two models: LASSO regressions and dynamic factor models (DFM).

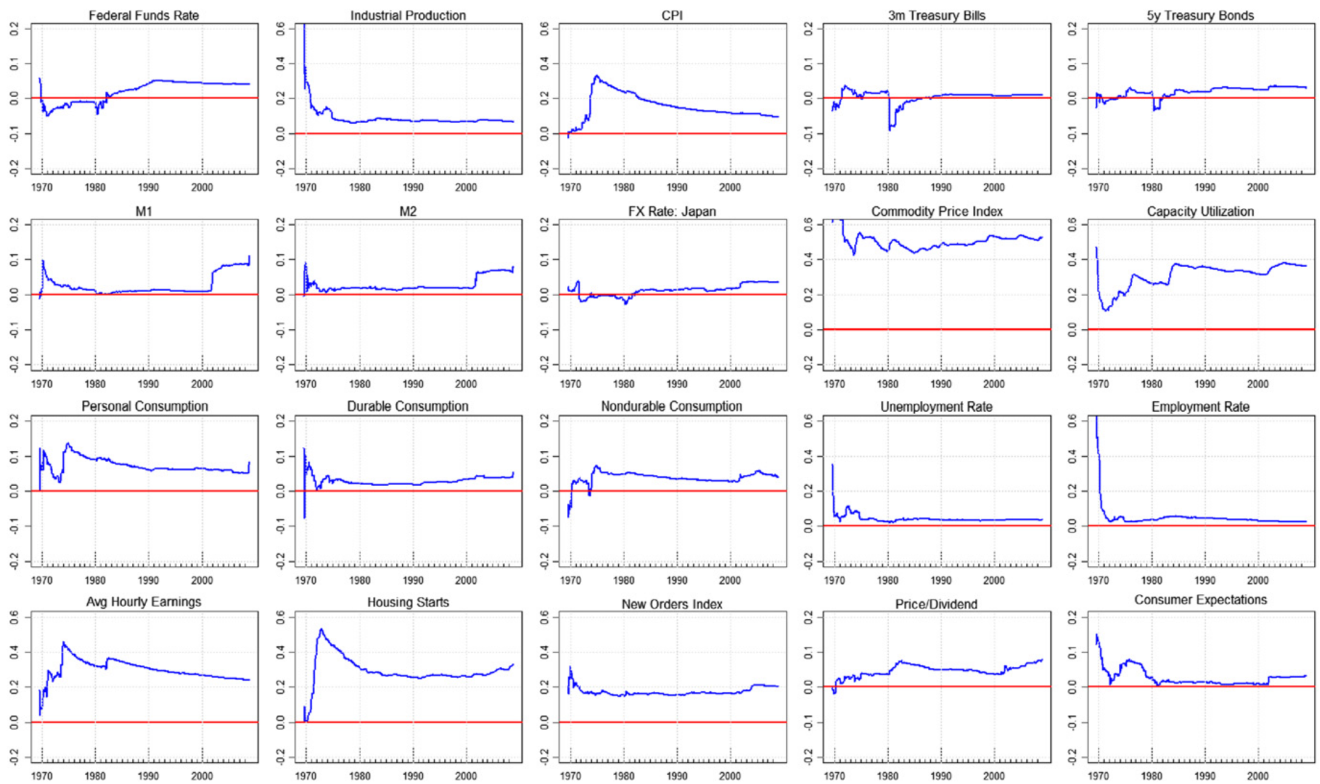

Fig. 3. Out-of-sample root mean square error (RMSE) differentials (ΔRMSE) over time, where the forecasts are from two models: (a) the forecast combination of LASSO regressions and dynamic factor models (DFM), and (b) dynamic factor models (DFM).

Table 3
Correlations of out-of-sample forecasts from two different models, including LASSO regression, elastic net regression (ENET), group LASSO regression (G-LASSO), dynamic factor model (DFM), and forecast combinations that combine the forecasts from LASSO-assisted regressions and dynamic factor models.

	LASSO, ENET	LASSO, GLASSO	GLASSO, ENET	LASSO, DFM	ENET, DFM	GLASSO, DFM	Comb. 1, Comb. 2	Comb. 1, Comb. 3	Comb. 2, Comb. 3
Federal funds rate	0.936	0.799	0.762	0.726	0.758	0.612	0.978	0.953	0.943
Industrial production	0.914	0.688	0.681	0.687	0.784	0.649	0.982	0.937	0.939
CPI	0.918	0.680	0.691	0.418	0.440	0.319	0.968	0.869	0.877
3m Treasury bills	0.898	0.614	0.640	0.602	0.715	0.479	0.978	0.926	0.933
5y Treasury bonds	0.800	0.486	0.537	0.540	0.565	0.411	0.963	0.906	0.908
M1	0.561	0.454	0.846	0.415	0.616	0.407	0.875	0.819	0.950
M2	0.896	0.720	0.670	0.477	0.591	0.418	0.966	0.906	0.897
Exchange rate: Japan	0.763	0.481	0.473	0.338	0.465	0.270	0.949	0.880	0.883
Commodity price index	0.998	0.944	0.943	0.555	0.556	0.547	0.998	0.947	0.946
Capacity utilization	1.000	0.883	0.884	0.826	0.824	0.713	1.000	0.961	0.961
Personal consumption	0.937	0.700	0.720	0.508	0.575	0.472	0.982	0.907	0.907
Durable consumption	0.942	0.634	0.643	0.347	0.327	0.266	0.975	0.839	0.844
Nondurable consumption	0.901	0.629	0.681	0.547	0.524	0.441	0.963	0.848	0.867
Unemployment rate	0.901	0.688	0.706	0.658	0.686	0.570	0.971	0.925	0.927
Employment rate	0.921	0.596	0.594	0.638	0.688	0.421	0.983	0.925	0.925
Avg hourly earnings	0.975	0.863	0.858	0.214	0.240	0.178	0.981	0.898	0.897
Housing starts	0.998	0.924	0.924	0.801	0.798	0.761	0.999	0.961	0.961
New orders index	0.993	0.955	0.949	0.805	0.818	0.773	0.996	0.976	0.973
Dividend yield	0.999	0.996	0.995	0.963	0.964	0.961	1.000	0.999	0.999
Consumer expectations	0.752	0.546	0.512	0.151	0.333	0.221	0.936	0.888	0.872
Average	0.900	0.714	0.735	0.561	0.613	0.494	0.972	0.914	0.920

Table 4
The ratio of the maximum root mean square error (RMSE) from an alternative model to the maximum root mean square error (RMSE) from dynamic factor models (DFM).

	LASSO vs. DFM	ENET vs. DFM	G-LASSO vs. DFM	Comb. 1 vs. DFM	Comb. 2 vs. DFM	Comb. 3 vs. DFM
Federal funds rate	1.123	1.241	1.163	1.057	1.125	1.080
Industrial production	1.146	1.145	1.138	1.069	1.070	1.063
CPI	**0.893**	**0.892**	**0.974**	**0.942**	**0.942**	**0.985**
3m Treasury bills	1.872	1.741	1.752	1.422	1.360	1.390
5y Treasury bonds	**0.996**	**0.974**	1.062	**0.998**	**0.976**	**0.988**
M1	**0.919**	**0.900**	**0.843**	**0.956**	**0.944**	**0.914**
M2	**0.755**	**0.721**	**0.818**	**0.865**	**0.845**	**0.900**
Exchange rate: Japan	**0.950**	**0.962**	**0.963**	**0.933**	**0.938**	**0.931**
Commodity price index	**0.496**	**0.507**	**0.589**	**0.474**	**0.469**	**0.566**
Capacity utilization	**0.326**	**0.298**	**0.800**	**0.329**	**0.330**	**0.676**
Personal consumption	**0.662**	**0.671**	**0.618**	**0.720**	**0.767**	**0.726**
Durable consumption	**0.795**	**0.779**	**0.881**	**0.887**	**0.885**	**0.942**
Nondurable consumption	**0.966**	**0.918**	**0.933**	**0.983**	**0.959**	**0.910**
Unemployment rate	**0.849**	**0.774**	**0.940**	**0.908**	**0.892**	**0.879**
Employment rate	1.093	1.091	1.045	1.049	1.047	1.024
Avg hourly earnings	**0.586**	**0.591**	**0.633**	**0.736**	**0.740**	**0.770**
Housing starts	**0.579**	**0.612**	**0.713**	**0.565**	**0.572**	**0.653**
New orders index	**0.799**	**0.796**	**0.679**	**0.809**	**0.808**	**0.704**
Price/dividend ratio	**0.970**	**0.970**	**0.924**	**0.967**	**0.970**	**0.951**
Consumer expectations	1.001	1.044	1.007	1.000	1.025	1.004

The alternative models include LASSO regression, elastic net regression and group LASSO regression, as well as their forecast combinations with dynamic factor models. Ratios less than one are in bold, and suggest that the alternative model produces less extreme forecasting errors.

forecasts are derived from a linear function of two forecasts, the forecast error instability is not increased empirically. In fact, all of the forecast combination methods have lower forecast error variances than the dynamic factor models. This is consistent with previous applications of forecast combination to the prediction of the equity premium (Rapach et al., 2010), where combined forecasts were found to reduce the forecast volatility relative to individual forecasts. The reduction in forecast error variances is most obvious for four data series: the commodity price index, average hourly earnings, housing starts and the new orders index, suggesting that the large variations captured by dynamic factors have trouble explaining these variables.

As referees pointed out, it would be interesting to include the ridge regression and a variant of the LASSO regression that considers both individual predictors and dynamic factors simultaneously. These two shrinkage methods can be implemented when one is forecasting with a large number of predictors, but does not assume sparse structures or model uncertainties in the data generating processes. Table 6 compares the out-of-sample root mean square error (RMSE) differentials (ΔRMSE) of forecasts from alternative methods with those from dynamic factor models (DFM), where the alternative methods include: (1) ridge regressions with the tuning parameters determined by cross validation (Ridge); (2) a combination of

Table 5
The ratio of the root mean square error (RMSE) variances from an alternative model to that from the dynamic factor model (DFM).

	LASSO vs. DFM	ENET vs. DFM	G-LASSO vs. DFM	Comb. 1 vs. DFM	Comb. 2 vs. DFM	Comb. 3 vs. DFM
Federal funds rate	**0.966**	**0.954**	1.023	**0.921**	**0.926**	**0.942**
Industrial production	**0.875**	**0.859**	**0.930**	**0.846**	**0.857**	**0.862**
CPI	**0.805**	**0.785**	**0.781**	**0.824**	**0.816**	**0.810**
3m Treasury bills	1.118	1.069	1.111	**0.978**	**0.969**	**0.960**
5y Treasury bonds	1.017	**0.984**	1.023	**0.951**	**0.940**	**0.944**
M1	**0.798**	**0.915**	**0.877**	**0.814**	**0.884**	**0.831**
M2	**0.833**	**0.837**	**0.882**	**0.857**	**0.869**	**0.885**
Exchange rate: Japan	**0.959**	**0.972**	**0.942**	**0.942**	**0.953**	**0.933**
Commodity price index	**0.143**	**0.149**	**0.279**	**0.166**	**0.173**	**0.318**
Capacity utilization	**0.076**	**0.076**	**0.688**	**0.091**	**0.091**	**0.440**
Personal consumption	**0.812**	**0.827**	**0.799**	**0.855**	**0.869**	**0.839**
Durable consumption	**0.863**	**0.874**	**0.863**	**0.886**	**0.893**	**0.887**
Nondurable consumption	**0.933**	**0.912**	**0.800**	**0.926**	**0.909**	**0.843**
Unemployment rate	**0.957**	**0.904**	1.002	**0.918**	**0.891**	**0.928**
Employment rate	**0.984**	**0.978**	1.060	**0.938**	**0.943**	**0.955**
Avg hourly earnings	**0.493**	**0.508**	**0.522**	**0.568**	**0.582**	**0.593**
Housing starts	**0.240**	**0.247**	**0.452**	**0.263**	**0.269**	**0.410**
New orders index	**0.449**	**0.458**	**0.475**	**0.493**	**0.507**	**0.501**
Price/dividend ratio	**0.857**	**0.860**	**0.887**	**0.871**	**0.873**	**0.879**
Consumer expectations	**0.952**	**0.943**	**0.962**	**0.943**	**0.942**	**0.954**

The alternative models include the LASSO regression, elastic net regression, and group LASSO regression, as well as their forecast combinations with dynamic factor models. Ratios less than one are in bold, and suggest that the predictive ability of the alternative model is more stable over time.

Table 6
Out-of-sample root mean square error (RMSE) differentials (ΔRMSE) of forecasts from alternative predictive regressions and dynamic factor models (DFM).

	Ridge vs. DFM		Combination 4 vs. DFM		LASSO4All vs. DFM	
	ΔRMSE	*p*-value	ΔRMSE	*p*-value	ΔRMSE	*p*-value
Federal funds rate	0.020	0.364	0.043	0.052	0.045	0.118
Industrial production	0.071	**0.001**	0.064	**0.000**	0.055	**0.017**
CPI	0.100	**0.001**	0.080	**0.000**	0.110	**0.001**
3m Treasury bills	−0.030	0.662	0.018	0.081	−0.059	0.797
5y Treasury bonds	0.014	0.373	0.037	**0.001**	−0.007	0.567
M1	0.075	0.100	0.085	**0.007**	0.101	**0.040**
M2	0.080	**0.027**	0.067	**0.004**	0.090	**0.031**
Exchange rate: Japan	0.028	0.108	0.031	**0.001**	0.010	0.329
Commodity price index	0.220	**0.000**	0.205	**0.000**	0.561	**0.000**
Capacity utilization	0.062	**0.000**	0.083	**0.000**	0.381	**0.000**
Personal consumption	0.076	**0.047**	0.062	**0.003**	0.088	**0.014**
Durable consumption	0.027	0.065	0.028	**0.004**	0.059	**0.007**
Nondurable consumption	0.056	**0.011**	0.045	**0.000**	0.020	0.227
Unemployment rate	0.047	**0.004**	0.042	**0.000**	0.031	0.060
Employment rate	0.020	0.096	0.026	**0.000**	0.014	0.177
Avg hourly earnings	0.151	**0.000**	0.108	**0.000**	0.290	**0.000**
Housing starts	0.103	**0.000**	0.101	**0.000**	0.351	**0.000**
New orders index	0.116	**0.000**	0.096	**0.000**	0.229	**0.000**
Price/dividend ratio	0.085	**0.002**	0.075	**0.000**	0.087	**0.002**
Consumer expectations	0.023	0.125	0.025	**0.002**	0.037	0.055

The alternative predictive regressions include ridge regressions (Ridge), the forecast combinations of ridge regressions and dynamic factor models (combination 4), and LASSO regressions on both individual predictors and the dynamic factors simultaneously (LASSO4All). For ridge regressions, tuning parameters are selected by cross validations for each month throughout the out-of-sample evaluation period. The forecasting accuracies are tested formally using the Diebold–Mariano test, with the *p*-values associated with the test statistics being reported. *p*-values which indicate significant differences in predictive abilities are given in bold (the significance level is 0.05).

ridge regressions and dynamic factor models (combination 4); and (3) LASSO regressions on both individual predictors and the dynamic factors simultaneously (LASSO4All). Clearly, although ridge regressions are significantly better than dynamic factor models for forecasting many variables, their statistical gains are not as large as those from LASSO-based approaches (see, e.g., commodity price index, capacity utilization, earnings and housing starts). As a result, the combined forecasts from ridge regressions and dynamic factor models are less advantageous. The LASSO4All method, on the other hand, has a forecasting performance similar to those of LASSO regressions. However, its performance is less satisfactory for forecasting exchange rates and consumption, and the interpretation of the regression coefficients is not straightforward.

5.3. Model sparsity and interpretability

So far we have seen three LASSO-based approaches having similar out-of-sample forecast performances. Since these approaches are all built upon the selection of important variables, with forecasting being based on these

Table 7
Final model sparsity of LASSO regressions, elastic net regressions and group LASSO regressions, where the sparsity in each month is defined as the proportion of independent variables used to explain the dependent variable.

	LASSO		ENET		G-LASSO	
	Sparsity: mean	Sparsity: stdev	Sparsity: mean	Sparsity: stdev	Sparsity: mean	Sparsity: stdev
Federal funds rate	0.152	0.112	0.609	0.357	0.459	0.235
Industrial production	0.126	0.074	0.480	0.364	0.388	0.239
CPI	0.145	0.114	0.323	0.282	0.426	0.257
3m Treasury bills	0.159	0.076	0.563	0.373	0.434	0.237
5y Treasury bonds	0.093	0.084	0.429	0.339	0.373	0.256
M1	0.125	0.094	0.293	0.266	0.451	0.296
M2	0.133	0.099	0.402	0.310	0.392	0.306
Exchange rate: Japan	0.058	0.077	0.440	0.398	0.187	0.180
Commodity price index	0.191	0.081	0.220	0.091	0.597	0.157
Capacity utilization	0.191	0.083	0.207	0.092	0.346	0.143
Personal consumption	0.099	0.082	0.185	0.180	0.328	0.260
Durable consumption	0.069	0.053	0.104	0.114	0.185	0.171
Nondurable consumption	0.119	0.097	0.304	0.275	0.390	0.282
Unemployment rate	0.168	0.113	0.559	0.350	0.580	0.307
Employment rate	0.135	0.097	0.501	0.360	0.468	0.337
Avg hourly earnings	0.175	0.106	0.243	0.149	0.446	0.287
Housing starts	0.150	0.068	0.176	0.083	0.330	0.215
New orders index	0.208	0.073	0.279	0.118	0.495	0.228
Price/dividend ratio	0.023	0.031	0.490	0.444	0.172	0.182
Consumer expectations	0.071	0.075	0.459	0.402	0.175	0.237

selected variables, it is interesting to check how they differ in terms of variable selections. In Table 7, we summarize the final model sparsity of LASSO regressions, elastic net regressions and group LASSO regressions, where sparsity is summarized monthly, and is defined as the proportion of the independent variables that are used to explain the dependent variable. Therefore, the average sparsity indicates how many variables are used to predict the economic indicator of interest, on average, and the standard deviation of model sparsity shows whether it is consistent over time. It can be seen that, in general, the LASSO regression gives the most parsimonious model, and the sparsity is consistently small over the full out-of-sample period. On the other hand, the elastic net regression and the group LASSO regression have comparable sizes for the final predictive models. Moreover, it has been confirmed that predicting interest rate related variables and employment/unemployment rates requires a relatively large set of independent variables.

Although the *numbers* of variables selected by the elastic net and group LASSO are relatively large, the selected variables are quite stable over time, which facilitates the model interpretation. In Fig. 4, we show the variable selection results of the LASSO regression, elastic net regression and group LASSO regression in forecasting the consumer price index (CPI). Compared with the LASSO regression, the elastic net regression generally selects more predictors, and these predictors can be included in a predictive regression for an extended period of time. By imposing a sparsity constraint at the block level, the group LASSO delivers the most interpretable results, in that the selected independent variables are within the same economically meaningful block in a given month. For example, in addition to consumer price indices (block K in Fig. 4), personal consumption indices (block J) are also strong predictors of the particular CPI time series under investigation. However, the CPI is not a predictor of personal consumption (see Table 8), which suggests a possible lead–lag relationship between these two variables. Similar patterns are observed in forecasting other variables. To present the results in a concise manner, in Table 8 we summarize the average number of important blocks selected, together with the three most frequently selected blocks for forecasting each variable. In forecasting industrial production, for example, variables that show consistent predictive power include interest rate differences (block H), housing starts (block F) and inventories and orders (block N). These findings are consistent with the economic interpretations of Bernanke (1990), Harvey (1989) and Koenig (2002), for example. In this regard, the group LASSO regression is the most advantageous as far as producing highly interpretable results for economists and policy makers. Based on its variable selection results, a clear, dynamic network of numerous economic indicators is revealed.

5.4. Robustness

Lastly, we investigate the robustness of these approaches by looking at predictive performances over various sub-samples and making comparisons with a different benchmark. Specifically, Table 9 compares the combined forecasts from group LASSO regressions and dynamic factor models with forecasts from dynamic factor models directly over two sub-samples, where the root mean square error (RMSE) differentials (ΔRMSE) and MSE-t statistics are reported. This forecast combination is the best one in Table 2. Following D'Agostino, Gambetti, and Giannone (2011), among others, we separate the Great Moderation period from the full out-of-sample evaluation period, and report the forecasting gains before and after January 1985. During the Great Moderation period, two models produce similar forecasts for dividend yield and consumer expectation. Moreover, by comparing these results with the full-sample results in Table 2, we find that four variables (M1, M2, employment rate and unemployment rate) are slightly more difficult to forecast using the proposed method after the mid-1980s.

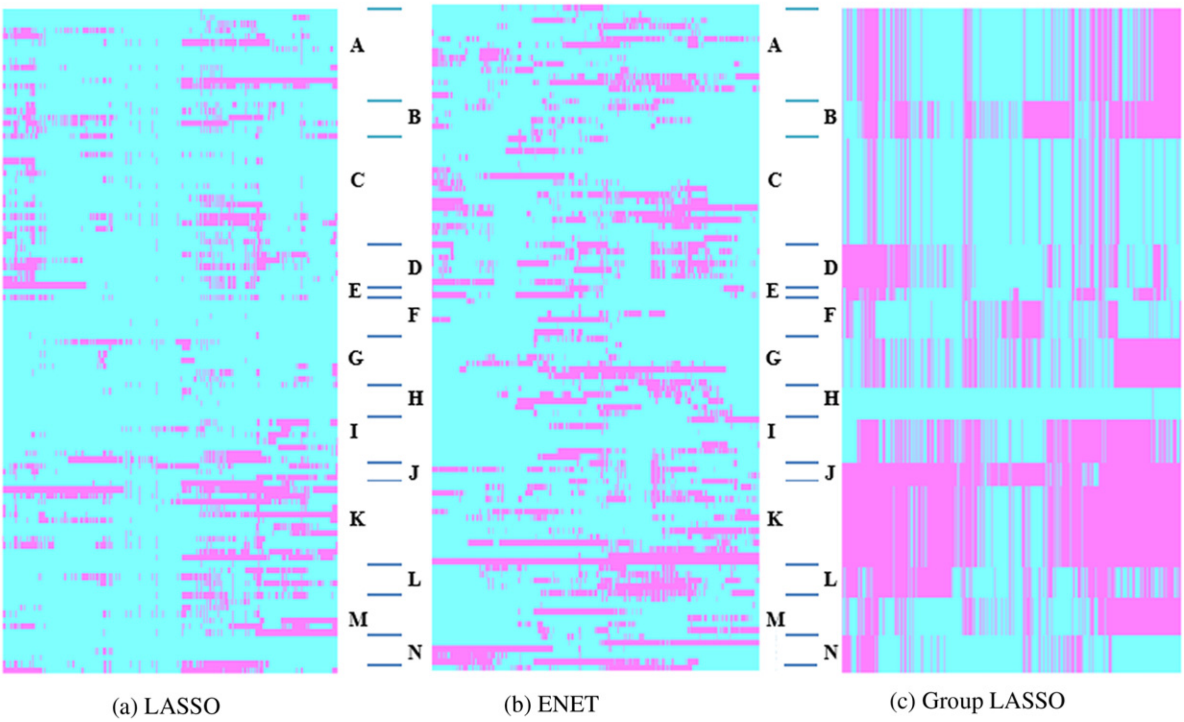

Fig. 4. Variable selected for forecasting the consumer price index (CPI) using (a) the LASSO regression, (b) the ENET regression, and (c) the group LASSO regression, where the *x*-axis shows the month within the out-of-sample evaluation period, and the *y*-axis is a total of 107 potential independent variables. Magenta color indicates that a variable is selected and included in the predictive regression in a given month. (For the interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 8
Variable selection results of group LASSO regressions.

	Avg. no. of active blocks	Most frequently selected blocks		
		1st	2nd	3rd
Federal funds rate	6.9	Interest rates	Inventories and orders	Stock prices
Industrial production	**6.1**	**Inventories and orders**	**Interest rates differences**	**Housing starts and sales**
CPI	**6.1**	**Consumptions**	**Price indexes**	**Money and credit**
3m Treasury bills	6.7	Interest rates	Hours	Interest rate differences
5y Treasury bonds	6.0	Interest rates	Stock prices	Interest rate differences
M1	6.4	Money and credit	Interest rates	Unemployment
M2	5.7	Money and credit	Unemployment	Interest rates
Exchange rate: Japan	3.3	Exchange rates	Interest rate differences	Stock prices
Commodity price index	8.1	Price indexes	Inventories and orders	Interest rate differences
Capacity utilization	5.5	Interest rate differences	Hours	Housing starts and sales
Personal consumption	**5.1**	**Consumptions**	**Hourly earnings**	**Money and credit**
Durable consumption	3.4	Consumptions	Unemployment	Hourly earnings
Nondurable consumption	5.6	Consumptions	Hourly earnings	Money and credit
Unemployment rate	8.4	Inventories and orders	Interest rate differences	Hours
Employment rate	6.9	Housing starts and sales	Inventories and orders	Unemployment
Avg hourly earnings	6.8	Hourly earnings	Hours	Consumptions
Housing starts	5.6	Housing starts and sales	Interest rate differences	Inventories and orders
New orders index	7.8	Interest rate differences	Inventories and orders	Housing starts and sales
Price/dividend ratio	3.0	Interest rate differences	Hours	Stock prices
Consumer expectations	2.8	Stock prices	Consumptions	Hours

For each dependent variable, some blocks of independent variables (active blocks) are selected out of the 14 blocks in each month. We report the average number of active blocks over the full out-of-sample evaluation period, as well as the three most frequently selected blocks. The three variables discussed in Section 5.3 are in bold.

Naïve forecasting models that do not include any regressors could be hard to beat. We compare three of the forecast combination methods in Table 2 with naïve models, and report the results in Table 10. All of the results are qualitatively similar, but the forecasting gains over the naïve benchmarks decrease for three variables: the federal funds rate, 5-year interest rate, and Japanese yen exchange rate. This suggests that, on average, the naïve models deliver better forecasts than the dynamic factor models for these three financial variables.

Table 9
Out-of-sample root mean square error (RMSE) differentials (ΔRMSE) of forecasts from forecast combinations and dynamic factor models (DFM) over two sub-samples.

	1969–1984		1985–2008	
	ΔRMSE	*p*-value	ΔRMSE	*p*-value
Federal funds rate	0.011	**0.012**	**0.096**	**0.000**
Industrial production	0.076	**0.000**	**0.055**	**0.000**
CPI	0.191	**0.017**	**0.047**	**0.001**
3m Treasury bills	−0.003	**0.000**	**0.081**	**0.000**
5y Treasury bonds	0.017	0.194	**0.053**	**0.000**
M1	0.016	**0.004**	**0.134**	**0.042**
M2	0.001	**0.037**	**0.099**	**0.018**
Exchange rate: Japan	0.028	**0.027**	**0.047**	**0.002**
Commodity price index	0.210	**0.000**	**0.490**	**0.000**
Capacity utilization	0.233	**0.000**	**0.190**	**0.000**
Personal consumption	0.091	**0.020**	**0.100**	**0.000**
Durable consumption	0.020	**0.016**	**0.080**	**0.000**
Nondurable consumption	0.087	**0.019**	**0.096**	**0.000**
Unemployment rate	0.054	**0.000**	**0.018**	**0.012**
Employment rate	0.034	**0.000**	**0.004**	**0.012**
Avg hourly earnings	0.327	**0.000**	**0.126**	**0.000**
Housing starts	0.193	**0.000**	**0.279**	**0.000**
New orders index	0.177	**0.000**	**0.217**	**0.000**
Price/dividend ratio	0.064	**0.003**	**0.081**	0.564
Consumer expectations	0.009	**0.030**	**0.034**	0.503

The forecast combinations combine forecasts from group LASSO regressions (combination 3) and dynamic factor models, with the tuning parameters being selected via cross validation, and $\theta = 0.9$ and $\kappa = 60$. The forecasting accuracies are tested by MSE-*t* statistics (Clark & West, 2006, 2007), with the *p*-values associated with the test statistics being reported. *p*-values which indicate significant differences in predictive abilities are given in bold (the significance level is 0.05).

6. Concluding remarks

In economics and finance, forecasting multivariate time series using large numbers of predictors is an important and challenging problem. Various attempts have been made in the literature to improve the model tractability and predictive ability, but there has been less consideration of model interpretation. This paper proposes and tests a collection of LASSO-assisted predictive regressions, where the LASSO components are capable of eliminating irrelevant predictors from the predictive model based on data-driven techniques. In this way, dimension reduction is guided by the out-of-sample performance, and both the predictive accuracy and the model interpretability are enhanced. Although the extents to which our three LASSO-based approaches improve the dynamic factor model are similar, the group LASSO that shrinks variables at the block level has the most easily interpretable results.[3] Moreover, when combining the forecasts from two categories of methods, the resulting forecasts have significantly higher predictive accuracies than those from dynamic factor models. In this regard, LASSO-based approaches could serve as an additional information source that complements dynamic factor-based approaches.[4]

Using LASSO-based approaches, Bai and Ng (2008) refined dynamic factor models by selecting the most important information for constructing dynamic factors. In our study, forecast combination can be regarded as another way to enhance dynamic factor models using shrinkage estimation. These highly integrated forecasting techniques provide valuable tools for economists in the presence of complex, dynamic and high-dimensional economic variables.

Acknowledgments

The authors are grateful to Rob Hyndman (the Editor-in-Chief), three anonymous referees, Mark Watson and Thomas Cosimano for helpful comments.

Appendix A. MSFE of combined forecasts

Suppose that the model uncertainty takes a shift in the data generating process (DGP) of the variable to be predicted, x_t, between a factor model and a sparse model:

$$x_t = \mathrm{I}_{\{S_t=1\}} L^T F_t + \mathrm{I}_{\{S_t=2\}} \Phi^T Z_{t-1} + e_t,$$

where s_t is a Bernoulli random variable indicating two states, with $\Pr(S_t = 1) = p$ and $\Pr(S_t = 2) = 1 - p$, $p \in (0, 1)$; $F_t = (f_{1,t}, \ldots, f_{m_1,t})^T$ is an m_1-dimensional vector of factors; Z_{t-1} is a vector of lagged predictors; and e_t is a random error with mean zero and variance σ_e^2. When a forecast is produced by the factor model and the true DGP is $S_t = 1$, the population value of the projection coefficient of x_t on $\hat{x}_t$, $\beta_{1|S_t=1}$, is given by

$$\beta_{1|S_t=1} = \mathrm{E}\left[F_t F_t^T\right]^{-1} \mathrm{E}\left[F_t \left(L^T F_t + e_t\right)^T\right] = \Sigma_1^{-1} \Sigma_1 L = L.$$

When a forecast is produced by the same factor model but the true DGP is $S_t = 0$,

$$\beta_{1|S_t=0} = \mathrm{E}\left[F_t F_t^T\right]^{-1} \mathrm{E}\left[F_t \left(\Phi^T Z_{t-1} + e_t\right)^T\right] = \Sigma_1^{-1} \Sigma_{12} \Phi.$$

Therefore, the unconditional value of the projection coefficient is

$$\beta_1 = pL + (1-p)\Sigma_1^{-1}\Sigma_{12}\Phi,$$

and the associated mean squared forecast error (MSFE) is

$$\begin{aligned}
\mathrm{MSFE}\left(\hat{x}_t^{(D)}\right) &= p\,\mathrm{Var}\left(x_t - \hat{x}_t^{(D,S_t=1)}\right) + (1-p) \\
&\quad \times \mathrm{Var}\left(x_t - \hat{x}_t^{(D,S_t=2)}\right) \\
&= p\,\mathrm{Var}\left(L^T F_t + e_t - \beta_1^T F_t\right) + (1-p) \\
&\quad \times \mathrm{Var}(\Phi^T Z_{t-1} + e_t - \beta_1^T F_t) \\
&= p\left(L^T - \beta_1^T\right)\Sigma_1 (L - \beta_1) + (1-p) \\
&\quad \times \left(\Phi^T \Sigma_2 \Phi + \beta_1^T \Sigma_1 \beta_1 - 2\Phi^T \Sigma_{12}\beta_1\right) + \sigma_e^2.
\end{aligned}$$

[3] Note, however, that this method requires a knowledge of the group membership. Determining the group membership and its effect on the stability of the results could be interesting topics for future research. We thank a referee for pointing this out.

[4] Luciani (2014) showed that, in factor models, a variable can be decomposed into two mutually orthogonal components: a common component, which is driven by a small number of pervasive shocks, and an idiosyncratic component, which is driven by non-pervasive shocks. In particular, the idiosyncratic component can be described by a sparse model. Therefore, LASSO-based approaches could help in capturing the cross-correlation which is left over once the principal components have been extracted. We thank a referee for pointing this out.

Table 10
Out-of-sample root mean square error (RMSE) differentials (ΔRMSE) of forecasts from forecast combinations and Naïve models without regressors (Naïve).

	Combination 1 vs. Naïve		Combination 2 vs. Naïve		Combination 3 vs. Naïve	
	ΔRMSE	*p*-value	ΔRMSE	*p*-value	ΔRMSE	*p*-value
Federal funds rate	0.042	**0.001**	0.039	**0.000**	0.031	**0.004**
Industrial production	0.066	**0.000**	0.063	**0.000**	0.064	**0.000**
CPI	0.101	**0.000**	0.106	**0.000**	0.110	**0.000**
3m Treasury bills	0.012	**0.000**	0.018	**0.000**	0.023	**0.000**
5y Treasury bonds	0.030	**0.001**	0.037	**0.000**	0.033	**0.000**
M1	0.110	**0.003**	0.067	**0.044**	0.099	**0.031**
M2	0.082	**0.001**	0.075	**0.000**	0.066	**0.004**
Exchange rate: Japan	0.034	**0.015**	0.027	**0.002**	0.040	**0.000**
Commodity price index	0.525	**0.000**	0.518	**0.000**	0.374	**0.000**
Capacity utilization	0.365	**0.000**	0.365	**0.000**	0.207	**0.000**
Personal consumption	0.084	**0.000**	0.076	**0.000**	0.097	**0.000**
Durable consumption	0.054	**0.000**	0.052	**0.000**	0.056	**0.000**
Nondurable consumption	0.041	**0.002**	0.051	**0.000**	0.092	**0.000**
Unemployment rate	0.037	**0.000**	0.049	**0.000**	0.034	**0.000**
Employment rate	0.026	**0.000**	0.023	**0.000**	0.018	**0.000**
Avg hourly earnings	0.239	**0.000**	0.229	**0.000**	0.224	**0.000**
Housing starts	0.328	**0.000**	0.325	**0.000**	0.244	**0.000**
New orders index	0.201	**0.000**	0.194	**0.000**	0.200	**0.000**
Price/dividend ratio	0.077	**0.002**	0.076	**0.012**	0.074	0.260
Consumer expectations	0.032	**0.005**	0.033	**0.000**	0.026	**0.047**

Forecast combinations combine forecasts from LASSO-assisted regressions and dynamic factor models, where the former category includes the LASSO regression (combination 1), the elastic net regression (combination 2) and the group LASSO regression (combination 3), with the tuning parameters being selected via cross validation, and $\theta = 0.9$ and $\kappa = 60$. The forecasting accuracies are tested using MSE-*t* statistics (Clark & West, 2006, 2007), with the *p*-values associated with the test statistics being reported. *p*-values which indicate significant differences in predictive abilities are given in bold (the significance level is 0.05).

Table B.1
Data transformations and descriptions.

Real output and income (block A)			
1*	**IPS10**	**5**	**INDUSTRIAL PRODUCTION INDEX—TOTAL INDEX**
2	IPS11	5	INDUSTRIAL PRODUCTION INDEX—PRODUCTS, TOTAL
3	IPS299	5	INDUSTRIAL PRODUCTION INDEX—FINAL PRODUCTS
4	IPS12	5	INDUSTRIAL PRODUCTION INDEX—CONSUMER GOODS
5	IPS13	5	INDUSTRIAL PRODUCTION INDEX—DURABLE CONSUMER GOODS
6	IPS18	5	INDUSTRIAL PRODUCTION INDEX—NONDURABLE CONSUMER GOODS
7	IPS25	5	INDUSTRIAL PRODUCTION INDEX—BUSINESS EQUIPMENT
8	IPS32	5	INDUSTRIAL PRODUCTION INDEX—MATERIALS
9	IPS34	5	INDUSTRIAL PRODUCTION INDEX—DURABLE GOODS MATERIALS
10	IPS38	5	INDUSTRIAL PRODUCTION INDEX—NONDURABLE GOODS MATERIALS
11	IPS43	5	INDUSTRIAL PRODUCTION INDEX—MANUFACTURING (SIC)
12	IPS307	5	INDUSTRIAL PRODUCTION INDEX—RESIDENTIAL UTILITIES
13	IPS306	5	INDUSTRIAL PRODUCTION INDEX—FUELS
14	PMP	1	NAPM PRODUCTION INDEX (PERCENT)
15*	**UTL11**	**1**	**CAPACITY UTILIZATION—MANUFACTURING (SIC)**
Hourly earnings (block B)			
16*	**CES275**	**6**	**AVG HRLY EARNINGS, PROD WRKRS, NONFARM—GOODS-PRODUCING**
17	CES277	6	AVG HRLY EARNINGS, PROD WRKRS, NONFARM—CONSTRUCTION
18	CES278	6	AVG HRLY EARNINGS, PROD WRKRS, NONFARM—MFG
19	CES275R	5	REAL AVG HRLY EARNINGS, PROD WRKRS, NONFARM—GOODS-PRODUCING
20	CES277R	5	REAL AVG HRLY EARNINGS, PROD WRKRS, NONFARM—CONSTRUCTION
21	CES278 R	5	REAL AVG HRLY EARNINGS, PROD WRKRS, NONFARM—MFG
Employment (block C)			
22	CES002	5	EMPLOYEES, NONFARM—TOTAL PRIVATE
23	CES003	5	EMPLOYEES, NONFARM—GOODS-PRODUCING
24	CES006	5	EMPLOYEES, NONFARM—MINING
25	CES011	5	EMPLOYEES, NONFARM—CONSTRUCTION
26	CES015	5	EMPLOYEES, NONFARM—MFG
27	CES017	5	EMPLOYEES, NONFARM—DURABLE GOODS
28	CES033	5	EMPLOYEES, NONFARM—NONDURABLE GOODS
29	CES046	5	EMPLOYEES, NONFARM—SERVICE-PROVIDING
30	CES048	5	EMPLOYEES, NONFARM—TRADE, TRANSPORT, UTILITIES
31	CES049	5	EMPLOYEES, NONFARM—WHOLESALE TRADE
32	CES053	5	EMPLOYEES, NONFARM—RETAIL TRADE

(*continued on next page*)

Table B.1 (*continued*)

33	CES088	5	EMPLOYEES, NONFARM—FINANCIAL ACTIVITIES
34	CES140	5	EMPLOYEES, NONFARM—GOVERNMENT
35	LHEL	2	INDEX OF HELP-WANTED ADVERTISING IN NEWSPAPERS (1967 = 100;SA)
36	LHELX	2	EMPLOYMENT: RATIO; HELP-WANTED ADS: NO. UNEMPLOYED CLF
37*	**LHEM**	**5**	**CIVILIAN LABOR FORCE: EMPLOYED, TOTAL (THOUS., SA)**
38	LHNAG	5	CIVILIAN LABOR FORCE: EMPLOYED, NONAGRIC. INDUSTRIES (THOUS., SA)
Unemployment (block D)			
39*	**LHUR**	**2**	**UNEMPLOYMENT RATE: ALL WORKERS, 16 YEARS and OVER (%, SA)**
40	LHU680	2	UNEMPLOY. BY DURATION: AVERAGE (MEAN) DURATION IN WEEKS (SA)
41	LHU5	5	UNEMPLOY. BY DURATION: PERSONS UNEMPL. LESS THAN 5 WKS (THOUS., SA)
42	LHU14	5	UNEMPLOY. BY DURATION: PERSONS UNEMPL. 5 TO 14 WKS (THOUS., SA)
43	LHU15	5	UNEMPLOY. BY DURATION: PERSONS UNEMPL. 15 WKS + (THOUS., SA)
44	LHU26	5	UNEMPLOY. BY DURATION: PERSONS UNEMPL. 15 TO 26 WKS (THOUS., SA)
45	LHU27	5	UNEMPLOY. BY DURATION: PERSONS UNEMPL. 27 WKS + (THOUS., SA)
Hours (block E)			
46	CES151	1	AVG WKLY HOURS, PROD WRKRS, NONFARM—GOODS-PRODUCING
47	CES155	2	AVG WKLY OVERTIME HOURS, PROD WRKRS, NONFARM—MFG
Housing starts and sales (block F)			
48	HSBR	4	HOUSING AUTHORIZED: TOTAL NEW PRIV HOUSING UNITS (THOUS., SAAR)
49*	**HSFR**	**4**	**HOUSING STARTS: TOTAL FARM and NONFARM (THOUS., USA)**
50	HSNE	4	HOUSING STARTS: NORTHEAST (THOUS., USA)
51	HSMW	4	HOUSING STARTS: MIDWEST (THOUS., USA)
52	HSSOU	4	HOUSING STARTS: SOUTH (THOUS., USA)
53	HSWST	4	HOUSING STARTS: WEST (THOUS., USA)
Interest rates (block G)			
54*	**FYFF**	**2**	**INTEREST RATE: FEDERAL FUNDS (EFFECTIVE) (% PER ANNUM, NSA)**
55*	**FYGM3**	**2**	**INTEREST RATE: US TREASURY BILLS, SEC MKT, 3-MO. (% PER ANN, NSA)**
56	FYGM6	2	INTEREST RATE: US TREASURY BILLS, SEC MKT, 6-MO. (% PER ANN, NSA)
57	FYGT1	2	INTEREST RATE: US TREASURY CONST MATURITIES, 1-YR. (% PER ANN, NSA)
58*	**FYGT5**	**2**	**INTEREST RATE: US TREASURY CONST MATURITIES, 5-YR. (% PER ANN, NSA)**
59	FYGT10	2	INTEREST RATE: US TREASURY CONST MATURITIES, 10-YR. (% PER ANN, NSA)
60	FYAAAC	2	BOND YIELD: MOODY'S AAA CORPORATE (% PER ANNUM)
61	FYBAAC	2	BOND YIELD: MOODY'S BAA CORPORATE (% PER ANNUM)
Interest rate differences (block H)			
62	SFYGM6	1	FYGM6-FYGM3
63	SFYGT1	1	FYGT1-FYGM3
64	SFYGT10	1	FYGT10-FYGM3
65	SFYAAAC	1	FYAAAC-FYGT10
66	SFYBAAC	1	FYBAAC-FYGT10
Money and credit quantity aggregates (block I)			
67*	**FM1**	**6**	**M1(CURR, TRAV.CKS, DEM DEP, OTHER CK'ABLE DEP, BIL$, SA)**
68	MZMSL	6	MZM (SA) FRB ST. LOUIS
69*	**FM2**	**6**	**M2(M1 + O'NITE RPS, EURO$, G/P&B/D MMMFS&SAV&SM TIME DEP, BIL$, SA)**
70	FMFBA	6	MONETARY BASE, ADJ FOR RESERVE REQUIREMENT CHANGES (MIL$, SA)
71	FMRRA	6	DEPOSITORY INST RESERVES: TOTAL, ADJ FOR RESERVE REQ CHGS (MIL$, SA)
72	BUSLOANS	6	COMMERCIAL AND INDUSTRIAL LOANS AT ALL COMMERCIAL BANKS (BI$, SA)
73	CCINRV	6	CONSUMER CREDIT OUTSTANDING—NONREVOLVING (G19)
Consumption (block J)			
74*	**PI071**	**6**	**PERSONAL CONSUMPTION, PRICE INDEX (2000 = 100)**
75*	**PI072**	**6**	**PERSONAL CONSUMPTION—DURABLE GOODS, PRICE INDEX (2000 = 100)**
76*	**PI073**	**6**	**PERSONAL CONSUMPTION—NONDURABLE GOODS, PRICE INDEX (2000 = 100)**
77	PI074	6	PERSONAL CONSUMPTION—SERVICES, PRICE INDEX (2000 = 100)

(*continued on next page*)

The MSFE associated with a forecast produced by a sparse model and that associated with a combined forecast can be derived similarly.

Appendix B. Data description

All of the macroeconomic data series considered in the empirical study are taken directly from the DRI/McGraw Hill Basic Economic Database. The four columns in the tables correspond to the series number, series mnemonic, transformation code and series description. The transformation codes are 1: no transformation; 2: first difference; 3: second difference; 4: logarithm; 5: first difference of logarithm; and 6: second difference of logarithm. The twenty variables investigated in detail are given in bold and marked with asterisk ($*$) (see Table B.1).

Table B.1 (*continued*)

Price indices (block K)			
78*	**CPIAUCSL**	**6**	**CPI ALL ITEMS (SA) FRED**
79	CPILFESL	6	CPI LESS FOOD AND ENERGY (SA) FRED
80	PCEPILFE	6	PCE PRICE INDEX LESS FOOD AND ENERGY (SA) FRED
81	PWFSA	6	PRODUCER PRICE INDEX: FINISHED GOODS (82 = 100, SA)
82	PWFCSA	6	PRODUCER PRICE INDEX: FINISHED CONSUMER GOODS (82 = 100, SA)
83	PWIMSA	6	PRODUCER PRICE INDEX: INTERMED MAT.SUPPLIES & COMPONENTS (82 = 100, SA)
84	PWCMSA	6	PRODUCER PRICE INDEX: CRUDE MATERIALS (82 = 100, SA)
85	PWCMSAR	5	REAL PRODUCER PRICE INDEX: CRUDE MATERIALS (82 = 100, SA)
86	PSCCOM	6	SPOT MARKET PRICE INDEX: BLS & CRB: ALL COMMODITIES (1967 = 100)
87	PSCCOMR	5	REAL SPOT MARKET PRICE INDEX: BLS & CRB: ALL COMMODITIES (1967 = 100)
88	PW561	6	PRODUCER PRICE INDEX: CRUDE PETROLEUM (82 = 100, NSA)
89	PW561R	5	PPI CRUDE (RELATIVE TO CORE PCE) (PW561/PCEPILFE)
90*	**PMCP**	**1**	**NAPM COMMODITY PRICES INDEX (PERCENT)**
Exchange rates (block L)			
91	EXRUS	5	UNITED STATES; EFFECTIVE EXCHANGE RATE (MERM) (INDEX NO.)
92	EXRSW	5	FOREIGN EXCHANGE RATE: SWITZERLAND (SWISS FRANC PER US$)
93*	**EXRJAN**	**5**	**FOREIGN EXCHANGE RATE: JAPAN (YEN PER US$)**
94	EXRUK	5	FOREIGN EXCHANGE RATE: UNITED KINGDOM (CENTS PER POUND)
95	EXRCAN	5	FOREIGN EXCHANGE RATE: CANADA (CANADIAN $ PER US$)
Stock prices and miscellaneous (block M)			
96	FSPCOM	5	S&P'S COMMON STOCK PRICE INDEX: COMPOSITE (1941–43 = 10)
97	FSPIN	5	S&P'S COMMON STOCK PRICE INDEX: INDUSTRIALS (1941–43 = 10)
98*	**FSDXP**	**2**	**S&P'S COMPOSITE COMMON STOCK: PRICE/DIVIDEND RATIO (% PER ANNUM)**
99	FSPXE	2	S&P'S COMPOSITE COMMON STOCK: PRICE-EARNINGS RATIO (%, NSA)
100	FSDJ	5	COMMON STOCK PRICES: DOW JONES INDUSTRIAL AVERAGE
101*	**HHSNTN**	**2**	**U. OF MICH. INDEX OF CONSUMER EXPECTATIONS (BCD-83)**
Inventories and orders (block N)			
102	PMI	1	PURCHASING MANAGERS' INDEX (SA)
103*	**PMNO**	**1**	**NAPM NEW ORDERS INDEX (PERCENT)**
104	PMDEL	1	NAPM VENDOR DELIVERIES INDEX (PERCENT)
105	PMNV	1	NAPM INVENTORIES INDEX (PERCENT)
106	MOCMQ	5	NEW ORDERS (NET)—CONSUMER GOODS & MATERIALS, 1996 DOLLARS (BCI)
107	MSONDQ	5	NEW ORDERS, NONDEFENSE CAPITAL GOODS, IN 1996 DOLLARS (BCI)

References

Aiolfi, M., Capistrán, C., & Timmermann, A. (2010). *Forecast combinations*. CREATES research paper 21.

Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Surveys*, *4*, 40–79.

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

Bates, J. M., & Granger, C. W. (1969). The combination of forecasts. *Operations Research Quarterly*, *20*, 451–468.

Bernanke, B. (1990). *On the predictive power of interest rates and interest rate spreads.* National Bureau of Economic Research. No. W3486.

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

Box, G. E. (1976). Science and statistics. *Journal of the American Statistical Association*, *71*(356), 791–799.

Clark, T. E., & West, K. D. (2006). Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis. *Journal of Econometrics*, *135*(1), 155–186.

Clark, T. E., & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, *138*(1), 291–311.

D'Agostino, A., Gambetti, L., & Giannone, D. (2011). Macroeconomic forecasting and structural change. *Journal of Applied Econometrics*, *28*(1), 82–101.

Della Corte, P., & Tsiakas, I. (2012). Statistical and economic methods for evaluating exchange rate predictability. In J. James, L. Sarno, & I. W. Marsh (Eds.), *Handbook of exchange rates* (pp. 221–264). London: Wiley.

DeMiguel, V., Garlappi, L., Nogales, F. J., & Uppal, R. (2009). A generalized approach to portfolio optimization: improving performance by constraining portfolio norms. *Management Science*, *55*(5), 798–812.

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, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, *13*(3), 253–263.

Efron, B., Hastie, T., Johnstone, I., & Tibshirani, R. (2004). Least angle regression. *The Annals of Statistics*, *32*, 407–499.

Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, *96*(456), 1348–1360.

Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, *33*, 1–22.

Fu, W. (1998). Penalized regressions: the bridge versus the LASSO. *Journal of Computational and Graphical Statistics*, *7*, 397–416.

Hallin, M., & Liška, R. (2011). Dynamic factors in the presence of blocks. *Journal of Econometrics*, *163*(1), 29–41.

Harvey, C. R. (1989). Forecasts of economic growth from the bond and stock markets. *Financial Analysts Journal*, *45*, 38–45.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*, *12*(1), 55–67.

Kan, R., & Zhou, G. (2007). Optimal portfolio choice with parameter uncertainty. *Journal of Financial and Quantitative Analysis*, *42*(3), 621.

Koenig, E. F. (2002). Using the purchasing managers' index to assess the economy's strength and the likely direction of monetary policy. *Federal Reserve Bank of Dallas Economic and Financial Policy Review*, *1*(6), 1–14.

Korobilis, D. (2013). Hierarchical shrinkage priors for dynamic regressions with many predictors. *International Journal of Forecasting*, *29*(1), 43–59.

Ledoit, O., & Wolf, M. (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance*, *10*(5), 603–621.

Ledoit, O., & Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, *88*(2), 365–411.

Luciani, M. (2014). Forecasting with approximate dynamic factor models: the role of non-pervasive shocks. *International Journal of Forecasting*, *30*(1), 20–29.
Meese, R. A., & Rogoff, K. (1983). Empirical exchange rate models of the seventies: do they fit out of sample? *Journal of International Economics*, *14*(1), 3–24.
Moench, E. (2008). Forecasting the yield curve in a data-rich environment: a no-arbitrage factor-augmented VAR approach. *Journal of Econometrics*, *146*(1), 26–43.
Moench, E., Ng, S., & Potter, S. (2011). *Dynamic hierarchical factor models*. FRB of New York staff report, 412.
Obstfeld, M., & Rogoff, K. (2001). The six major puzzles in international macroeconomics: is there a common cause? In *NBER macroeconomics annual 2000, Vol. 15* (pp. 339–412). MIT Press.
Park, T., & Casella, G. (2008). The Bayesian Lasso. *Journal of the American Statistical Association*, *103*(482), 681–686.
Rapach, D. E., Strauss, J. K., & Zhou, G. (2010). Out-of-sample equity premium prediction: combination forecasts and links to the real economy. *Review of Financial Studies*, *23*(2), 821–862.
Rapach, D. E., Strauss, J. K., & Zhou, G. (2013). International stock return predictability: What is the role of the United States? *Journal of Finance*, *68*(4), 1633–1662.
Song, S., & Bickel, P. J. (2011). *Large vector auto regressions*. ArXiv preprint arXiv:1106.3915.
Stock, J. H., & Watson, M. W. (2002a). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics*, *20*, 147–162.
Stock, J. H., & Watson, M. W. (2002b). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, *97*(460), 1167–1179.
Stock, J. H., & Watson, M. W. (2004). Combination forecasts of output growth in a seven-country data set. *Journal of Forecasting*, *23*(6), 405–430.
Stock, J. H., & Watson, M. W. (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, Series B*, *58*, 267–288.
Timmermann, A. (2006). Forecast combinations. In *Handbook of economic forecasting, Vol. 1* (pp. 135–196).
Welch, I., & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, *21*(4), 1455–1508.
Wu, T. T., & Lange, K. (2008). Coordinate descent algorithms for Lasso penalized regression. *Annals of Applied Statistics*, *2*, 224–244.
Yuan, M., & Lin, Y. (2005). Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, *68*(1), 49–67.
Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)*, *67*(2), 301–320.

Jiahan Li is an assistant professor of statistics at the Department of Applied and Computational Mathematics and Statistics (ACMS), the University of Notre Dame. He received his M.S. degree from the University of Florida in 2008 and a Ph.D. in statistics from the Pennsylvania State University in 2011. His research interests include large data analytics, financial econometrics and forecasting methods for bioinformatics and business. His work won the Best Paper Award at the Fourth Annual Applied Finance with R Conference, 2012, and The Finalist of Best Paper Competition Award of the Service Science Section, INFORMS, 2009.

Weiye Chen is a Ph.D. candidate at the Department of Applied and Computational Mathematics and Statistics (ACMS), the University of Notre Dame. She received her B.S. degree in mathematics from the University of Science and Technology of China. Her research interests include statistical modeling and forecasting methods. She is the recipient of a travel grant for the Women in Mathematical Sciences conference, Institute for Mathematics and Its Applications (IMA), 2013.