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