International Journal of Forecasting 32 (2016) 1284–1305

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Forecasting and nowcasting economic growth in the euro area using factor models

Irma Hindrayanto [a], Siem Jan Koopman [b,c], Jasper de Winter [a,*]

[a] Economics and Research Division, De Nederlandsche Bank, The Netherlands
[b] Department of Econometrics, Vrije Universiteit Amsterdam and Tinbergen Institute, The Netherlands
[c] CREATES Aarhus University, Denmark

ARTICLE INFO

Keywords:
Factor models
Dynamic analysis
State space method
Kalman filter
Forecasting competition
Real-time data
Mixed frequency

ABSTRACT

Many empirical studies have provided evidence that the use of factor models, which use large data sets of economic variables, can contribute to the computation of more accurate forecasts. In this study, we examine the performances of four different factor models in a pseudo real-time forecasting competition for the euro area and five of its largest countries. Our aim is to identify empirically the best factor model approach for the forecasting and nowcasting of the quarterly gross domestic product growth rate. We also propose some modifications of existing factor model specifications, with the aim of improving their forecast performances empirically. We conclude that factor models consistently outperform the benchmark autoregressive model, both before and during the crisis. Moreover, we find that the best forecast accuracy is generally produced by the collapsed dynamic factor model.

1. Introduction

It is acknowledged widely that the forecasting of macroeconomic time series is of critical importance, both for economic policy makers and for the general public. Reliable short-term forecasts are in particularly high demand when the economic environment is uncertain. Many different methodologies for producing such forecasts exist, ranging from basic time series models to sophisticated structural dynamic models. Over the last decade, dynamic factor models have become a popular tool for short-term forecasting amongst both practitioners and econometricians, due to their good forecast performances in many studies; see for example Giannone, Reichlin, and Small (2008) and Stock and Watson (2002) for the United States, Angelini, Camba-Mendez, Giannone, Reichlin, and Rünstler (2011) and Rünstler et al. (2009) for the euro area, and Schumacher and Breitung (2008) for Germany. In all empirical studies concerning dynamic factor models, there are various decisions that need to be made before forecasting can start. We provide three examples. First, the optimal number of factors in the model needs to be determined by following procedures such as those of Ahn and Horenstein (2013), Alessi, Barigozzi, and Capasso (2010), Bai and Ng (2002), Hallin and Liška (2007) and Onatski (2010). Second, the selection of the database for extracting the factors, and its size, are important determinants of a successful forecasting procedure; see for example the discussions by Boivin and Ng (2005), Caggiano, Kapetanios, and Labhard (2011) and den Reijer (2013). Third, the number of lagged terms of the target variable in the forecasting model needs to be set. The gain in forecast accuracy from including one or more lags of the target variable in the forecast

* Corresponding author.
E-mail addresses: a.i.w.hindrayanto@dnb.nl (I. Hindrayanto), s.j.koopman@vu.nl (S.J. Koopman), j.m.de.winter@dnb.nl (J. de Winter).

http://dx.doi.org/10.1016/j.ijforecast.2016.05.003
0169-2070/

equation has not been documented well. However, recent studies indicate that including more autoregressive terms may increase the forecast accuracy; see for example Clements and Galvão (2008), Jansen, Jin, and de Winter (2016) and Kuzin, Marcellino, and Schumacher (2011). It is an empirical question as to whether this finding holds for all factor model specifications. Such matters have also been discussed in related empirical studies; see for example Bańbura, Giannone, Modugno, and Reichlin (2011), Jungbacker and Koopman (2015), Lahiri and Monokroussos (2013), Liebermann (2014), and Matheson (2013).

We compare the short-term forecast performances of different factor models for quarterly gross domestic product (GDP) growth in the euro area and its five largest countries, before and during the financial crisis. The one- to three-month-ahead forecasts for the current quarter are referred to as nowcasts. The short-term forecasting of key economic variables using dynamic factor analysis has been reviewed by Bai and Ng (2008), Breitung and Eickmeier (2006) and Stock and Watson (2011), with Luciani (2014) discussing more recent contributions. We consider four estimation procedures for the dynamic factor model: the basic principal components method of Stock and Watson (2002), who initiated the current literature on factor models; the widely used two-step approach of Doz, Giannone, and Reichlin (2011); the more elaborate quasi-maximum likelihood method of Doz, Giannone, and Reichlin (2012); and the more recently proposed maximum likelihood method of Bräuning and Koopman (2014), based on a collapsed dynamic factor model. All of these estimation approaches rely to some extent on principal components that summarize the information in a large set of monthly indicators. The estimation methods proposed by Bräuning and Koopman (2014) and Doz et al. (2011) use the principal components as approximations of the dynamic factors. Doz et al. (2012) use the principal components as an initialisation of quasi-maximum likelihood estimation. The Kalman filter plays a key role in all three of these approaches.

For all dynamic factor approaches, we analyze the target variable and the common factors simultaneously in a multivariate unobserved component time series model. All modeling frameworks allow for panels with mixed-frequencies and with the monthly time series having different publication delays and starting dates. This leads to a data matrix of monthly time series with so-called "jagged" or "ragged" edges at the beginning and end of the sample. The two-step approach developed by Doz et al. (2011) was applied to the euro area by Angelini et al. (2011) and Bańbura and Rünstler (2011). The first step involves the computation of the principal components and the estimation of their dynamic properties by means of a vector autoregressive model. The second step involves obtaining the factor estimates and forecasts from the Kalman filter and smoother. Doz et al. (2011) provide the asymptotic properties of the factor estimates and use the model to forecast the quarterly GDP growth using monthly variables that contain jagged edges at the beginning and end of the sample. Bańbura and Rünstler (2011) developed this approach further by including the quarterly GDP growth as a latent variable in the state vector, so that the contributions of different variables to the forecasts can be quantified using the algorithms of Koopman and Harvey (2003). The quasi-maximum likelihood approach of Doz et al. (2012) was applied to the euro area by Bańbura et al. (2011) and Bańbura and Modugno (2014). It is shown that this approach obtains consistent estimates of the factors as the size of the cross-section goes to infinity. Bańbura and Modugno (2014) extend the framework of Doz et al. (2012) by introducing modifications in relation to missing entries (at random) and the dynamic treatment of idiosyncratic effects; see also Luciani (2014).

The collapsed dynamic factor model of Bräuning and Koopman (2014) effectively adopts a low-dimensional unobserved components time series model for both the target variable and a set of principal components. This multivariate model is then used to forecast the target variable based on its past realizations and the principal components. The idiosyncratic part of the target variable is modeled explicitly and dealt with jointly with the dynamic factors. It mitigates the challenge of estimating the factors and forecasting the target variable in a joint analysis based on a large macroeconomic panel. The unknown parameters in this parsimonious model are estimated by maximum likelihood: the loglikelihood function is evaluated by the Kalman filter and maximized numerically with respect to the unknown parameters. The score function can be evaluated using a corresponding smoothing algorithm. The forecasts of the target variable are generated by the Kalman filter.

The main contributions of this paper are twofold. First, we propose small modifications for the different estimation approaches, with the aim of placing them on a somewhat more equal footing. For example, we extend the model of Doz et al. (2011) by including more autoregressive terms, as in Bräuning and Koopman (2014) and Stock and Watson (2002). We also develop a more effective way of handling the jagged edges for the collapsed dynamic factor method of Bräuning and Koopman (2014). We propose a three-step method: (i) analyze each univariate time series by using an unobserved components model to extract the main signal for imputing the jagged edges; (ii) extract the principal components; (iii) estimate the parameters simultaneously. This handling of the jagged edges improves the forecast accuracy. Second, our main contribution is our empirical study, in which we verify the forecast accuracies of the four dynamic factor approaches. We present a systematic comparison of the different modeling treatments for the euro area, and conclude that the factor modeling approaches systematically produce more accurate forecasts than those of the benchmark autoregressive model. This good performance is not limited to the pre-financial crisis period: the factor models also outperform the benchmark model during the financial crisis by up to 77%, in terms of mean squared errors, depending on the factor model, country and forecast horizon. Overall, the collapsed dynamic factor approach is the most successful for forecasting and nowcasting in our empirical study of the euro area.

The remainder of the paper is organized as follows. Section 2 gives an overview of the four different dynamic factor model approaches considered in our study, and