ARTICLE IN PRESS

Computational Statistics and Data Analysis ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

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

Real-time factor model forecasting and the effects of instability

Michael P. Clements *,[1]

ICMA Centre, Henley Business School, University of Reading, Reading RG6 6BA, United Kingdom

ARTICLE INFO

Article history:
Received 24 June 2014
Received in revised form 19 January 2015
Accepted 19 January 2015
Available online xxxx

Keywords:
Factor models
Robust approaches
Financial crisis

ABSTRACT

Factor forecasting models are shown to deliver real-time gains over autoregressive models for US real activity variables during the recent period, but are less successful for nominal variables. The gains are largely due to the Financial Crisis period, and are primarily at the shortest (one quarter ahead) horizon. Excluding the pre-Great Moderation years from the factor forecasting model estimation period (but not from the data used to extract factors) results in a marked fillip in factor model forecast accuracy, but does the same for the AR model forecasts. The relative performance of the factor models compared to the AR models is largely unaffected by whether the exercise is in real time or is pseudo out-of-sample.

1. Introduction

Diffusion indices and factor models have become popular in the economic forecasting literature in recent years: see e.g., Stock and Watson (1989, 1999, 2002, 2009, 2011), Forni et al. (2000), Peña and Poncela (2004), Schumacher and Breitung (2008), Bai and Ng (2008) and Castle et al. (2013, 2015), *inter alia*. This can be explained by the greater availability of large databases, and the desire to draw on all relevant data when analysing the current state of the economy and scope for government intervention. For example, Bernanke and Boivin (2003, p. 526) refer to 'Central banker's reputations as data fiends'. Factor models can be seen as a formal way of allowing all the disparate data series available to the analyst to have some influence on the question of interest. Moreover, their popularity was also due to their good forecasting performance, at least initially. Stock and Watson (2011, p. 54) suggest that the evidence indicates that 'factor forecasts perform well to very well relative to competitors for many, but not all, macroeconomic series. For US real activity series, reductions in pseudo-out-of-sample mean squared forecast errors at the two- to four-quarter horizon are often in the range of 20% to 40%, although smaller or no improvements are seen for other series, such as US inflation after 1990'. However, other assessments of the forecast performance of factor models are more equivocal, as indicated by Eickmeier and Ziegler (2008) in their meta-analysis of factor forecasting applications, and by D'Agostino et al. (2006), who suggest the predictability of US macroeconomic series has greatly diminished in the Great Moderation period (mid 1980s to the 2007 Financial Crisis).

Although Eickmeier and Ziegler (2008) tease out a number of the determinants of factor model forecast performance from a careful meta-analysis of 52 studies, a number of potentially important determinants are not explicitly addressed. Chief among these are the potential effects of parameter non-constancy or structural breaks on the factor forecasting models (as well as on the rival or benchmark models). Structural breaks are sometimes viewed as the key culprit in causing forecast

* Tel.: +44 0 1183788239; fax: +44 0 1189314741.
E-mail address: m.p.clements@reading.ac.uk.

[1] Michael Clements is Professor of Econometrics at the ICMA Centre, Henley Business School, University of Reading, Whiteknights, Reading RG6 6BA, and is a member of the Institute for New Economic Thinking at the Oxford Martin School, University of Oxford.

http://dx.doi.org/10.1016/j.csda.2015.01.011
0167-9473/

failure in general (see, e.g., Clements and Hendry, 2006 for a recent review). We find that the relative performance of factor models is markedly better than that of AR benchmarks during the recent crisis period, with factor models exhibiting a degree of adaptability in times of change. We also consider split-sample estimation of the factors and factor forecasting model following the promising results found by Stock and Watson (2009) in their in-sample study. In addition we consider the adoption of rolling estimation windows, which is a common approach when it is felt that earlier observations may be less relevant.

We consider the usefulness of these ways of mitigating the effects of breaks in a real-time forecasting exercise, as opposed to the 'pseudo-out-of-sample' exercises that typify the majority of the literature. That is, the vast majority of the evidence that underpins the conclusions of studies such as Eickmeier and Ziegler (2008) is based on forecasting exercises that have used fully-revised data, rather than the vintages of data which would have been available at the time the forecasts were actually made. A focus of much of the recent literature has been on the implications of using fully-revised data versus real-time data for both historical analyses and forecasting. A number of papers have investigated whether the use of final-revised data (as in the pseudo out-of-sample) overstates the usefulness of various predictor variables relative to appraisals based solely on the data vintages available at the time the forecasts are constructed (see, for example, Diebold and Rudebusch, 1991, Faust et al., 2003, and the review by Croushore, 2006). There are fewer than a handful of papers which report proper real-time evidence on factor forecasting performance, so that one contribution of our paper is to compare real-time and pseudo out-of-sample performance for forecasting a relatively large number of variables over the recent period.

Our study of the forecasting performance of factor models also acknowledges that factor models may work better for some macroeconomic variables than for others (as suggested by the quote above from Stock and Watson, 2011), and we are careful to distinguish between real activity variables, and price/nominal variables in the way we present results.

We also investigate why it might be that factor models offer little improvement on AR models for forecasting.

In focusing on structural breaks and real-time effects, there are a number of aspects that we do not cover which might potentially be important, so our paper is complementary to the large body of work on factor forecasting models summarized in Eickmeier and Ziegler (2008). Among the aspects we neglect are the following. We estimate the factors by principal components, and do not consider alternative methods such as that of Forni et al. (2005). We do not consider the role of the number of variables (N) used to estimate the factors, and because we adopt a fully real-time exercise, our N is necessarily at the lower end of the values used in pseudo out-of-sample forecasting exercises. We have $N = 51$. According to e.g., Bai and Ng (2002) and Boivin and Ng (2006), inter alia, this should not harm the forecasting performance of the factor models, although there is evidence to the contrary, e.g., Bernanke and Boivin (2003). Furthermore, we do not calculate factors from a set of 'targeted predictors', as in Bai and Ng (2008). Targeted predictors are variables selected to have predictive power for the variable of interest, based on hard or soft thresholding (such as LASSO, see e.g., Tibshirani, 1996). Nor do we consider block-factor approaches, such as Moench et al. (2009), where the data is divided up into a number of categories, and principal components are calculated for each category.

The plan of the remainder of the paper is as follows. Section 2 discusses the literature on factor forecasting using real-time data, and recent developments concerning factor models and instability. Section 3 describes the dataset and our implementation of factor forecasting. Section 4 presents the results of the empirical forecasting exercises, and Section 5 provides a focus on forecasting real GDP and inflation over the Financial Crisis period. Section 6 investigates why factor models are only marginally better than AR models in 'normal times'. Finally, Section 7 concludes.

2. Related literature

There are few studies of factor model forecasting using real-time data, presumably because the requirements of amassing datasets consisting of a large number of variables covering reasonably long historical periods, for each of a number of data vintages, constitute a demanding data collection exercise. One such study is Bernanke and Boivin (2003), who calculate the accuracy of recursively-generated forecasts of 3 variables, CPI inflation, industrial production, and unemployment, over the period 1970–1998, for both factor models based on real-time data, and on fully-revised data. For factors calculated from a dataset comprising 78 variables, for which they have real-time vintages and fully-revised data, they find forecast performance is broadly similar for the factor forecasting models estimated on fully-revised and real-time data, and the gains of the factor forecasts over an autoregressive model are modest: small for CPI, but larger for industrial production and unemployment. However, they find that if they use the 215 series fully-revised dataset of Stock and Watson (2002) to construct factors in a pseudo-out-of-sample exercise (that is, estimating the models using the fully-revised data), they obtain sizeable gains for all three variables, including CPI. They suggest that the number of variables used to construct the factors may matter more than whether fully-revised or real-time data is used.

Faust and Wright (2009) use the real-time datasets associated with the Greenbook forecasts for the FOMC meetings from 1980 to 2000 to analyse forecasts of GDP-deflator inflation and GDP growth. A key question is whether the superiority of the Greenbook forecasts is solely due to the Fed staff's knowledge of the current state of the economy and of recent developments which affect the short-term outlook. By providing the atheoretical models with the same information as the Greenbook, they find that the Fed's forecasts of the GDP-deflator measure of inflation remain superior, but their advantage forecasting GDP growth disappears. Their study covers factor models, as well as a raft of univariate and multivariate time-series models. They compare the relative performance of the Greenbook and model forecasts using *ex post* revised data and real-time data, and come to the same conclusion as Bernanke and Boivin (2003), that the relative performances are largely

Please cite this article in press as: Clements, M.P., Real-time factor model forecasting and the effects of instability. Computational Statistics and Data Analysis (2015), http://dx.doi.org/10.1016/j.csda.2015.01.011