International Journal of Forecasting 30 (2014) 20–29

Contents lists available at ScienceDirect

International Journal of Forecasting

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

Forecasting with approximate dynamic factor models: The role of *non-pervasive* shocks

Matteo Luciani *

F.R.S.-FNRS, 5 Rue d'Egmont, 1000, Brussels, Belgium
European Center for Advanced Research in Economics and Statistics (ECARES), Solvay Business School of Economics and Management, Université libre de Bruxelles, 50 av. F.D. Roosevelt, CP 114, B1050 Brussels, Belgium

A R T I C L E I N F O

Keywords:
Dynamic factor models
Penalized regressions
Local factors
Bayesian shrinkage
Forecasting

A B S T R A C T

This paper studies the role of *non-pervasive* shocks when forecasting with factor models. To this end, we first introduce a new model that incorporates the effects of *non-pervasive* shocks, an Approximate Dynamic Factor Model with a sparse model for the idiosyncratic component. Then, we test the forecasting performance of this model both in simulations, and on a large panel of US quarterly data. We find that, when the goal is to forecast a disaggregated variable, which is usually affected by regional or sectorial shocks, it is useful to capture the dynamics generated by *non-pervasive* shocks; however, when the goal is to forecast an aggregate variable, which responds primarily to macroeconomic, i.e. *pervasive*, shocks, accounting for *non-pervasive* shocks is not useful.

1. Introduction

In recent years, the literature has proposed two methods for coping with the curse of dimensionality problem, namely: factor models (Forni, Hallin, Lippi, & Reichlin, 2000; Stock & Watson, 2002a) and Bayesian shrinkage (De Mol, Giannone, & Reichlin, 2008). Roughly speaking, the main idea of factor models is to *summarize* the information content of a large number of predictors in a few factors, while the idea of Bayesian shrinkage is to limit the estimation uncertainty by *shrinking* the potentially complex model toward a simple *naïve* prior model.[1]

In factor models, each variable (x_{it}) can be decomposed into the sum of two mutually orthogonal components, one capturing the comovement among the data (χ_{it}), which is assumed to be driven by a small number of *pervasive* shocks ($\mathbf{u}_t$); and one capturing the idiosyncratic dynamics (ξ_{it}): $x_{it} = \chi_{it} + \xi_{it}$. Due to the strong comovement among macroeconomic time series, these models offer a realistic (and parsimonious) representation of the data. Moreover, these models can be estimated easily using the method of principal components under the assumption of "weakly" cross-sectionally-dynamically correlated idiosyncratic components (Bai, 2003; Bai & Ng, 2002; Forni et al., 2000; Forni, Hallin, Lippi, & Reichlin, 2005; Stock & Watson, 2002a), a likely feature in large macroeconomic databases where *non-pervasive* (sectorial or regional) shocks might affect groups of variables (local factors).[2]

Factor models have proved to be successful in predicting economic activity. A large body of literature has shown

* Correspondence to: European Center for Advanced Research in Economics and Statistics (ECARES), Solvay Business School of Economics and Management, Université libre de Bruxelles, 50 av. F.D. Roosevelt, CP 114, B1050 Brussels, Belgium. Tel.: +32 (0) 2 650 33 66.
E-mail address: matteo.luciani@ulb.ac.be.

[1] Other methods which are not used in this paper, but which are also able to forecast with large numbers of predictors, include partial least squares (Groen & Kapetanios, 2008), forecast combination (Bates & Granger, 1969), the Bayesian model average (Leamer, 1978), and bagging (Breiman, 1996). For a review of forecasting with many predictors, see Stock and Watson (2006).

[2] The literature refers to these models as *approximate* factor models, as distinct from *exact* factor models, which are characterized by cross-sectionally-dynamically uncorrelated idiosyncratic components, i.e., $\xi_{it} \sim iid(0, 1)$. For the sake of simplicity, throughout this paper we will refer to *approximate* factor models simply as "factor models".

0169-2070/$ – see front matter © 2013 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ijforecast.2013.05.001

how factor models can outperform common univariate benchmark forecasts by simply forecasting the common component (i.e., the one driven by *pervasive* shocks), while approximating with an autoregressive model the idiosyncratic component (i.e., the one driven by *non-pervasive* shocks). Specifically, this body of literature has shown that a forecast obtained as $x_{i,t+h|t} = \chi_{i,t+h|t} + \xi_{i,t+h|t}$, where $\chi_{i,t+h|t} = Proj\{x_{i,t+h}|\mathbf{u}_t, \mathbf{u}_{t-1}, \ldots\}$ and $\xi_{i,t+h|t} = Proj\{x_{i,t+h}|\xi_{it}, \xi_{it-1}, \ldots\}$, outperforms a univariate forecast such as $x_{i,t+h|t} = Proj\{x_{i,t+h}|x_{it}, , x_{it-1}, \ldots\}$. A (non-exhaustive) list of papers that use this approach is: Artis, Banerjee, and Marcellino (2005), Bańbura and Modugno (in press), Boivin and Ng (2005, 2006), Bulligan, Golinelli, and Parigi (2010), Camacho and Perez-Quiros (2010), D'Agostino and Giannone (2012), Forni, Hallin, Lippi, and Reichlin (2003), Forni et al. (2005), Marcellino, Stock, and Watson (2003), Schumacher (2007) and Stock and Watson (2002a,b). One limitation of this approach is that it uses factor models in forecasting as if the idiosyncratic components were mutually orthogonal. Consequently, the literature to date has not taken into account the correlation induced by *non-pervasive* shocks. This is a contradiction, though, since factor models are estimated under the hypothesis of "weakly" cross-sectionally-dynamically correlated errors. The questions then arise: does taking this correlation into account lead to any forecasting gains? Is it worth accounting for *non-pervasive* shocks when forecasting with factor models? This paper answers these questions.

In this paper, we study the role of *non-pervasive* shocks when forecasting with factor models. To this end, we first introduce a new model that incorporates the effects of *non-pervasive* shocks, then test its forecasting performance both in simulations, and on a large panel of US quarterly data.

Our model augments the factor model with a sparse model for the idiosyncratic component, thus taking into account, and exploiting in forecasting, the fact that, in approximate dynamic factor models, the idiosyncratic component is "weakly" cross-sectionally-dynamically correlated. Our model produces a forecast as $x_{i,t+h|t} = \chi_{i,t+h|t} + \xi_{i,t+h|t}$, where $\chi_{i,t+h|t} = Proj\{x_{i,t+h}|\mathbf{u}_t, \mathbf{u}_{t-1}, \ldots\}$ and $\xi_{i,t+h|t} = Proj\{x_{i,t+h}|\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t-1}, \ldots\}$, with $\boldsymbol{\xi}_t = [\xi_{1,t}, \xi_{2,t}, \ldots, \xi_{N,t}]'$. This forecast is obtained by mixing factor models and L_1 penalized regressions, which are equivalent to Bayesian shrinkage with double exponential priors, or boosting. We choose L_1 penalized regressions and boosting because, by performing both shrinkage and variable selection, they impose a sparse structure on the idiosyncratic component. This sparse structure is particularly appropriate for our purpose, since we are interested in capturing *non-pervasive* shocks that, by definition, affect only a limited number of variables.

The literature recently suggested a different forecasting strategy which also involves L_1 penalized regressions. This method was used by Bai and Ng (2008a) and De Mol et al. (2008). The former suggest extracting the factors only from those variables that are really informative for forecasting the target variable. The latter suggest selecting the predictors and estimating the model using only the selected predictors.

Although our approach uses the same method as those of Bai and Ng (2008a) and De Mol et al. (2008), it is theoretically different: while they impose a sparse structure on the whole dataset, we impose a sparse structure only on the idiosyncratic component. That is, we begin by extracting what is common, then impose sparsity on what is left.

An alternative method of accounting for *non-pervasive* shocks, used in forecasting by Bańbura, Giannone, and Reichlin (2011), involves estimating a factor model with both global and local factors using either maximum likelihood techniques (Doz, Giannone, & Reichlin, 2012) or Bayesian methods (Kose, Otrok, & Whiteman, 2008; Moench, Ng, & Potter, in press).[3] However, we do not consider this approach here, since it requires *a priori* information on the structure of the economy in order to identify *non-pervasive* shocks. In contrast, our method identifies *non-pervasive* shocks automatically by performing variable selection.

The rest of the paper is organized as follows. We illustrate our model in Section 2. In Section 3, by means of a simulation exercise, we study whether and when it is useful to account for *non-pervasive* shocks when forecasting with factor models. We test our model in Section 4 by means of a pseudo real time forecasting exercise on US quarterly data against the factor model of Forni et al. (2005), and against the methods of Bai and Ng (2008a) and De Mol et al. (2008). In Section 5, we verify the robustness of our model to the composition of the database. Finally, we present our conclusions in Section 6.

2. Methodology

Let $\mathbf{x}_t$ be an $N \times 1$ vector of stationary variables. Suppose that we are interested in forecasting the ith variable h steps ahead, $x_{i,t+h|t}$, by using all N potential predictors. In this case, the best linear prediction, defined as

$$x_{i,t+h|t} = Proj\{x_{i,t+h}|\Omega_t\}, \tag{1}$$

where $\Omega_t = span\{\mathbf{x}_{t-p}, p = 0, 1, \ldots\}$, might be extremely inefficient, or even impossible, due to the lack of degrees of freedom. This is the well-known curse of dimensionality problem. In recent years, the literature has suggested two solutions: factor models (Forni et al., 2000; Stock & Watson, 2002a) and Bayesian shrinkage (De Mol et al., 2008).

If the comovement of the N variables in $\mathbf{x}_t$ can be approximated well by a small number $q \ll N$ of *pervasive* (or common) shocks $\mathbf{u_t}$, while the variable specific dynamics ($\boldsymbol{\xi}_t$) are only mildly correlated, then the information set can be split into two orthogonal spaces: the space spanned by the common shocks and the space spanned by the idiosyncratic components ($\Omega_t = \Omega_t^u \oplus \Omega_t^\xi$, where $\Omega_t^u = span\{\mathbf{u}_{t-p}, p = 0, 1, \ldots\}$, and $\Omega_t^\xi = span\{\boldsymbol{\xi}_{t-p}, p = 0, 1, \ldots\}$, with $\Omega_t^F \cap \Omega_t^\xi = \{0\}$). The idea of factor models is to approximate the linear projection on the whole information set by the sum of the linear projection on the space

[3] Hallin and Liška (2011) suggest a method involving dynamic principal components, which is able to account for *non-pervasive* shocks but cannot be used in forecasting.