ARTICLE IN PRESS

Journal of Econometrics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at SciVerse ScienceDirect

Journal of Econometrics

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

Forecasting by factors, by variables, by both or neither?

Jennifer L. Castle[a], Michael P. Clements[b], David F. Hendry[c,*]

[a] Magdalen College and Institute for New Economic Thinking, Oxford Martin School, University of Oxford, UK

[b] Economics Department, Warwick University, UK

[c] Economics Department and Institute for New Economic Thinking, Oxford Martin School, University of Oxford, UK

ARTICLE INFO

Article history:
Available online xxxx

JEL classification:
C51
C22

Keywords:
Model selection
Factor models
Forecasting
Impulse-indicator saturation
Autometrics

ABSTRACT

We consider forecasting with factors, variables and both, modeling in-sample using *Autometrics* so all principal components and variables can be included jointly, while tackling multiple breaks by impulse-indicator saturation. A forecast-error taxonomy for factor models highlights the impacts of location shifts on forecast-error biases. Forecasting US GDP over 1-, 4- and 8-step horizons using the dataset from Stock and Watson (2009) updated to 2011:2 shows factor models are more useful for nowcasting or short-term forecasting, but their relative performance declines as the forecast horizon increases. Forecasts for GDP levels highlight the need for robust strategies, such as intercept corrections or differencing, when location shifts occur as in the recent financial crisis.

1. Introduction and historical background

There are three venerable traditions in economic forecasting based respectively on economic-theory derived empirical econometric models, 'indicator' or 'factor' approaches combining many sources of information, and mechanistic approaches, which are typically univariate.

Members of the first group are exemplified by early models like Smith (1927, 1929) and Tinbergen (1930), smaller systems in the immediate post-war period (such as Klein, 1950; Tinbergen, 1951; Klein et al., 1961), leading onto large macroeconometric models (Duesenberry et al., 1969, and Fair, 1970, with a survey in Wallis, 1989), and now including both dynamic stochastic general equilibrium (DSGE) models widely used at Central Banks (see e.g., Smets and Wouters, 2003), and global models, first developed by project Link (see e.g., Waelbroeck, 1976) and more recently, global vector autoregressions (GVARs: see Dees et al., 2007; Pesaran et al., 2009; Ericsson, 2010).

The second approach commenced with the ABC curves of Persons (1924), followed by leading indicators as in Zarnowitz and Boschan (1977) with critiques in Diebold and Rudebusch (1991) and Emerson and Hendry (1996). Factor analytic and principal component methods have a long history in statistics and psychology (see e.g., Spearman, 1927; Cattell, 1952; Anderson, 1958; Lawley and Maxwell, 1963; Joreskog, 1967; Bartholomew, 1987) and have seen some distinguished applications in economics (e.g., Stone, 1947, for an early macroeconomic application; and Gorman, 1956, for a microeconomic one). Diffusion indices and factor models are now quite widely used for economic forecasting; see e.g., Stock and Watson (1989, 1999, 2009), Forni et al. (2000), Peña and Poncela (2004) and Schumacher and Breitung (2008).

The third set includes methods like exponentially weighted moving averages, the closely related Holt–Winters approach (see Holt, 1957; Winters, 1960), damped trend (see e.g., Fildes, 1992) and autoregressions, including the general time-series approach in Box and Jenkins (1970). Some members of this class were often found to dominate in forecasting competitions: see Makridakis et al. (1982) and Makridakis and Hibon (2000), and are the 'neither' in the title.

Until recently, while the first two approaches often compared their forecasts with various 'naive' methods selected from the third group, there were few direct comparisons between them, and almost no studies included both. Here we consider models selected from very general initial specifications, which might be motivated as approximating the reduced forms of the models of the first group, and compare these directly with the factor models of the second group. Our automatic model selection algorithm permits the inclusion of variables and factors on an equal footing, allowing in-sample selection over both of these based on their explanatory power for the target variable. This remedies the dearth of direct comparisons of the two approaches in the literature. But, as emphasized in Section 2, a good in-sample fit does not guarantee

* Corresponding author.
E-mail addresses: jennifer.castle@magd.ox.ac.uk (J.L. Castle), M.P.Clements@warwick.ac.uk (M.P. Clements), david.hendry@nuffield.ox.ac.uk (D.F. Hendry).

0304-4076/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jeconom.2013.04.015

a good out-of-sample forecast performance (see e.g., Clements and Hendry, 2005a), so that a detailed analysis of forecasting performance is undertaken.

The structure of the paper is as follows. Section 2 describes some of the issues that arise in any analysis of forecasting models or methods. Section 3 outlines the statistical framework used to analyze forecasting with factors or variables. Section 4 develops the analysis of forecasting from factor models when there are location shifts. Section 5 discusses model selection when including both factors and variables. Section 6 briefly reviews alternatives to principal components proposed in the statistics literature, including those which form part of the empirical analysis. Section 7 illustrates the analysis using US GDP forecasts. Section 8 concludes.

2. Setting the scene

Many interacting issues need to be addressed when analyzing forecasting, the complexity of which entails that the answer to the title's question is likely to be context specific. Although general guidelines are rare, it is fruitful to consider eight aspects: (i) the pooling of both variables and factors in forecasting models; (ii) the role of in-sample model selection in that setting; (iii) whether or not breaks over the forecast horizon are unanticipated; (iv) more versus less information in forecasting; (v) the type of forecasting model in use, specifically whether it is an equilibrium-correction mechanism (EqCM); (vi) measurement errors in the data, especially near the forecast origin; (vii) how to evaluate the 'success or failure' of forecasts; (viii) the nature of the data-generating process (DGP). We briefly consider these in turn, and indicate whether in principle our approach of selecting over variables and factors should be advantageous. An advantage of our approach is that it allows us to be agnostic as to the nature of the data generating process (DGP), especially whether the DGP can be usefully represented as having a factor structure.

2.1. Pooling of information

Factor models are a way of forecasting using a large number of predictors, as opposed to pooling over the forecasts of a large number of simple, often single-predictor, models. When there are many variables in the set from which factors are formed (the 'external' variables), including both the set of factors and the original variables will often result in the number of candidate variables, N, being larger than the sample size, T. Model selection when $N > T$ may have seemed insurmountable in the past, but is not now. Let $\mathbf{z}_t$ denote the set of n 'external' variables' from which the factors $\mathbf{f}_t = \mathbf{H}\mathbf{z}_t$ (say) are formed, then $\mathbf{f}_t, \ldots \mathbf{f}_{t-s}, \mathbf{z}_t, \ldots \mathbf{z}_{t-s}$ comprise the initial set of candidate variables. Automatic model selection can use multi-path searches to eliminate irrelevant variables with mixtures of expanding and contracting block searches, so can handle settings with both perfect collinearity and $N > T$; see Hendry and Krolzig (2005) and Doornik (2009b). The simulations in Castle et al. (2011) show the feasibility of such an approach when $N > T$ in linear dynamic models. Investigators are, therefore, not forced to allow for only a small number of factors, or just the factors and a few lags of the variable being forecast, as candidates. Since model selection is unavoidable when $N > T$, we consider that next.

2.2. Model selection

The search algorithm in *Autometrics* within *PcGive* (see Doornik, 2009a; Doornik and Hendry, 2009) seeks the local DGP (denoted LDGP), namely the DGP for the set of variables under consideration (see e.g., Hendry, 2009) by formulating a general unrestricted model (GUM) that nests the LDGP, checking its congruence when feasible (estimable once $N \ll T$ and perfect collinearities are removed). Search thereafter ensures congruence, so all selected models are valid restrictions of the GUM, and should parsimoniously encompass the feasible GUM. Location shifts are removed in-sample by impulse-indicator saturation (IIS, see Hendry et al., 2008, Johansen and Nielsen, 2009, and the simulation studies in Castle et al., 2012b), which also addresses possible outliers. Thus, if $\left\{1_{\{j=t\}}, t = 1, \ldots, T\right\}$ denotes the complete set of T impulse indicators, we allow for $\mathbf{f}_t, \ldots \mathbf{f}_{t-s}, \mathbf{z}_t, \ldots \mathbf{z}_{t-s}$ and $\left\{1_{\{j=t\}}, t = 1, \ldots, T\right\}$ all being included in the initial set of candidate variables to which multi-path search is applied. Hence $N > T$ will always occur when IIS is used, but the in-sample feasibility of this approach is shown in Castle et al. (2012a). Here we are concerned with the application of models selected in this way to a forecasting context when the DGP is non-stationary due to location shifts. Since there are few analyses of how well a factor forecasting approach would then perform (see however, Stock and Watson, 2009; Corradi and Swanson, 2011), we explore its behavior when faced with location shifts at the forecast origin. Section 5 discusses automatic model selection.

2.3. Unanticipated location shifts

Third, *ex ante* forecasting is fundamentally different from *ex post* modeling when unanticipated location shifts occur. Breaks can always be modeled after the event (at worst by indicator variables), but will cause forecast failure when not anticipated. Clements and Hendry (1998) proposed a general theory of economic forecasting using mis-specified models in a world of structural breaks, and emphasized that it had radically different implications from a forecasting theory based on stationarity and well-specified models (as in Klein, 1971, say). Moreover, those authors also show that breaks other than location shifts are less pernicious for forecasting (though not for policy analyses). Pesaran and Timmermann (2005) and Pesaran et al. (2006) consider forecasting time series subject to multiple structural breaks, and Pesaran and Timmermann (2007) examine the use of moving windows in that context. Castle et al. (2011) investigate how breaks themselves might be forecast, and if not, how to forecast during breaks, but draw somewhat pessimistic conclusions due to the limited information that will be available at the time any location shift occurs. Thus, we focus the analysis on the impacts of unanticipated location shifts in factor-based forecasting models.

2.4. Role of information in forecasting

Factor models can be interpreted as a particular form of 'pooling of information', in contrast to the 'pooling of forecasts' literature discussed in (e.g.) Hendry and Clements (2004). Pooling information ought to dominate pooling forecasts based on limited information, except when all variables are orthogonal (see e.g., Granger, 1989). However, the taxonomy of forecast errors in Clements and Hendry (2005b) suggests that incomplete information by itself is unlikely to play a key role in forecast failure, so using large datasets may not correct one of the main problems confronting forecasters, namely location shifts, unless that additional information is pertinent to forecasting breaks. Moreover, although we use model selection from a very general initial candidate set, combined with congruence as a basis for econometric modeling, it cannot be proved that congruent modeling helps for forecasting when facing location shifts (see e.g., Allen and Fildes, 2001). While Makridakis and Hibon (2000) conclude that parsimonious models do best in forecasting competitions, Clements and Hendry (2001) argue that such findings may conflate parsimony and robustness to location shifts: most of the parsimonious models were relatively