Journal of Econometrics 177 (2013) 153–170

Contents lists available at ScienceDirect

Journal of Econometrics

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

Adaptive forecasting in the presence of recent and ongoing structural change

Liudas Giraitis[a], George Kapetanios[a,*], Simon Price[b,c]

[a] Queen Mary, University of London, United Kingdom
[b] Bank of England, London, United Kingdom
[c] City University, London, United Kingdom

ARTICLE INFO

Article history:
Available online 6 April 2013

JEL classification:
C100
C590

Keywords:
Recent and ongoing structural change
Forecast combination
Robust forecasts

ABSTRACT

We consider time series forecasting in the presence of ongoing structural change where both the time series dependence and the nature of the structural change are unknown. Methods that downweight older data, such as rolling regressions, forecast averaging over different windows and exponentially weighted moving averages, known to be robust to historical structural change, are found also to be useful in the presence of ongoing structural change in the forecast period. A crucial issue is how to select the degree of downweighting, usually defined by an arbitrary tuning parameter. We make this choice data-dependent by minimising the forecast mean square error, and provide a detailed theoretical analysis of our proposal. Monte Carlo results illustrate the methods. We examine their performance on 97 US macro series. Forecasts using data-based tuning of the data discount rate are shown to perform well.

© 2013 Bank of England. Published by Elsevier B.V. All rights reserved.

1. Introduction

It is widely accepted that structural change is a crucial issue in econometrics and forecasting. Clements and Hendry suggest forcefully (in e.g. Clements and Hendry, 1998a,b) that such change is the main source of the forecast error; Hendry (2000) argues that the dominant cause of forecast failures is the presence of deterministic shifts. Convincing evidence of structural change was offered by Stock and Watson (1996) who looked at many forecasting models of a large number of US time series, and found parameter instability in a substantial proportion. This issue remains relevant: in a survey of the literature on forecasting in the presence of instabilities for the *Handbook of Forecasting*, Rossi (2012) writes 'the widespread presence of forecast breakdowns suggests the need for improving ways to select good forecasting models in-sample'. Our work on robust and data driven forecasting is a contribution to precisely this end. As model parameters may change continuously, drift smoothly over time or change at discrete points in an unknown manner, and both within the sample and over the forecast period, we consider a general setting where the model structure and presence and type of structural change are all unknown.

There is a large literature on the identification of breaks, and forecasting methods robust to them (Rossi, 2012). However, the deeply practical need to forecast after a recent structural change, or during a period of such change, has received very little attention. As most forecast approaches are only effective in specific cases, the problem is compounded by the unknown and therefore unspecified nature of any structural change.

Detection of structural change has a long history, mainly in the context of structural breaks (although see Kapetanios (2007) for the case of smooth structural change). Seminal papers include Chow (1960), Andrews (1993) and Bai and Perron (1998). But the question of amendment of forecasting strategies then arises. While this has been tackled by many authors, a major contribution was made by Pesaran and Timmermann (2007). They concluded that, in the presence of breaks, forecast pooling using a variety of estimation windows provides a reasonably good and robust forecasting performance.

Nevertheless, most work on forecasting assumes that change has occurred when sufficient time has elapsed for post-break estimation.[1] In practice, the issue of change occurring in real time is a major consideration, which was partly addressed in Eklund et al. (2010). They considered a variety of forecasting strategies which can be divided into two distinct groups. In one case the forecaster monitors for change and adjusts methods once change has been detected. In the other the forecaster does not attempt to identify breaks, since that involves a substantial

* Corresponding author.
E-mail address: g.kapetanios@qmul.ac.uk (G. Kapetanios).

[1] Exceptions include the interesting work of Clements and Hendry (2006) and Castle et al. (2011).

0304-4076/$ – see front matter © 2013 Bank of England. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jeconom.2013.04.003

time lag. Instead break–robust forecasting strategies are used that essentially downweight data from older periods deemed to be irrelevant for the current conjuncture.

While moving in an interesting direction, Eklund et al. (2010) do not elaborate two issues: how much to downweight past data, and whether to do so monotonically. Clearly, any arbitrary discount factor is unlikely to be optimal. And neither may monotonicity: for example, if regimes (e.g., monetary policy) come and go then older data, from a period where the current regime previously held, might be more relevant than more recent data from other regimes.

In this paper, we suggest forecasting approaches that address these issues. Our main contribution is to introduce and analyse a cross-validation based method which selects a tuning parameter defining the downweighting rate of the older data. We show that the implied discount rate minimises the mean square error (MSE) of the forecast in the weighting schemes considered. Further, we consider a nonparametric method for determining a flexible weighting scheme. The latter does not assume any particular shape for the weight function, nor monotonicity. We explore the properties of the new forecasting methods for a variety of models in terms of theory, with a Monte Carlo exercise and empirically. It turns out that the method is valid under a wide range of forms of structural breaks and persistence, and can be generalised in a number of practically important dimensions, most notably allowing varying dynamic structures.

A byproduct of our results is a new way to accommodate trends of a generic nature in forecasting. Unlike many forecasting approaches that require the removal of stochastic or other trends before forecasting, our methods can be directly applied to the level of the forecast series.

The rest of the paper is organised as follows. Section 2 presents our approach for forecasting in the presence of recent structural breaks. We provide its theoretical justification and asymptotic MSE, and describe some robust forecasting strategies. Section 3 includes an extensive Monte Carlo study in which these strategies are evaluated. In Section 4 the methods are used to forecast a large number of US macroeconomic time series, where we find results broadly consistent with the Monte Carlo study. Section 5 concludes. Proofs are reported in an Appendix.

2. Adaptive forecasting: econometric framework

2.1. Forecasting strategies

In this section we work with a simple location forecasting framework that is as general as possible while consistent with clear theoretical results. It may be summarised as

$$y_t = \beta_t + u_t, \quad t = 1, \ldots, T, \tag{2.1}$$

where β_t is an unobserved persistent process, and u_t is a stationary dependent noise that is independent of β_t. Unlike most previous work we wish to place as little structure as possible on the process β_t. We do not specify whether β_t is stochastic or deterministic, or whether it is discontinuous or smooth. The noise process u_t is a stationary linear process with mean zero and finite variance σ_u^2. The persistent component $\beta_t \equiv \beta_{T,t}$ is allowed to be a triangular array, and can be a stochastic (unit root) or deterministic (bounded) trend. This set-up provides sufficient flexibility to our theoretical analysis of forecasting y_t, allowing for $\beta_{T,t}$ such as those used in locally stationary models (e.g. Dahlhaus (1996)), or in persistent stochastic unit root trend models. For simplicity of notation, we write $y_{T,t}$ as y_t and $\beta_{T,t}$ as β_t. It should be stressed that in robust forecasting, which is our focus, the structure of β_t is neither known nor estimated. Concerning our simple location conditional mean modelling, we note that our analysis can allow both the use of a generic model of the conditional mean of the process and robust forecasting around that model. We discuss details related to this extension in Section 2.8.

Eklund et al. (2010) find that simple forecasting of y_t, based on weighting schemes that discount past data, works well in practice. Examples include exponential weighting and forecast combinations based on different estimation windows. By varying a tuning parameter, such methods impose different shapes on the weight functions that downweight past data. Their weakness is that it is not clear how to select the tuning parameters. So data-dependent tuning methods for choosing these parameters are of great interest.

One way to calibrate parameters is by optimising on in-sample forecasting performance. This idea is not new. For example, Kapetanios et al. (2006) suggest forecasts where different models are averaged with weights that depend on the forecasting performance of each model in the recent past. In what follows we formalise the above ideas, presenting a data-driven weighting strategy and developing its theoretical analysis.

We consider a linear forecast of y_t, based on (local) averaging of past values $y_{t-1}, \ldots, y_1$:

$$\hat{y}_{t|t-1,H} = \sum_{j=1}^{t-1} w_{tj,H} y_{t-j} = w_{t1,H} y_{t-1} + \cdots + w_{t,t-1,H} y_1, \tag{2.2}$$

with weights $w_{tj,H} \geq 0$ such that $w_{t1,H} + \cdots + w_{t,t-1,H} = 1$, parametrised by a single tuning parameter H. The latter defines the rate of downweighting the past observations (e.g., the width of the rolling window). The structure of weights $w_{tj,H}$ is described in Assumption 1. We assume that H takes values in the interval $I_T = [\alpha, H_{\max}]$, where $\alpha > 0$.

Assumption 1. The function $K(x) \geq 0$, $x \geq 0$ is continuous and differentiable on its support, such that $\int_0^\infty K(u)du = 1$, $K(0) > 0$, and for some $C > 0$, $c > 0$

$$K(x) \leq C\exp(-c|x|), \qquad |\dot{K}(x)| \leq C/(1+x^2), \quad x > 0, \tag{2.3}$$

where $\dot{K}$ is the first derivative of K. For $t \geq 1$, $H \in I_T$, set $k_{j,H} = K(j/H)$ and define

$$w_{tj,H} = \frac{k_{j,H}}{\sum_{s=1}^{t-1} k_{s,H}}, \quad j = 1, \ldots, t-1. \tag{2.4}$$

Example 1. The main classes of commonly employed weights satisfy this assumption.

(i) *Rolling window* weights, with $K(u) = I(0 \leq u \leq 1)$.
(ii) *Exponential weighted moving average* (EWMA) weights, with $K(u) = e^{-u}$, $u \in [0, \infty)$. Then, with $\rho = \exp(-1/H)$, $k_{j,H} = \rho^j$ and $w_{tj,H} = \rho^j / \sum_{k=1}^{t-1} \rho^k$, $1 \leq j \leq t-1$.
(iii) *Triangular window* weights, with $K(u) = 2(1-u)I(0 \leq u \leq 1)$.

While the rolling window simply averages the H previous observations, the EWMA forecast uses all observations $y_1, \ldots, y_{t-1}$, increasingly downweighting the more distant past. In practice, forecasting of a unit root or trending process y_t is often conducted by averaging over the last few observations. When persistence in y_t falls, wider windows may be expected to yield smaller forecast MSE. It is also plausible that for a stationary process $\{y_t\}$ when dependence is sufficiently strong a forecast discounting past data will outperform the sample mean forecast $(y_t + \cdots + y_1)/t$. These observations, supported by the theory below, indicate that the 'optimal' selection of H depends on the unknown type of persistence in y_t. Thus, contrary to the usual practice of using a preselected value of H, a data based selection method for H is indicated.

2.2. Selection of the tuning parameter H

Given a sample $y_1, \ldots, y_T$, computation of the forecast $y_{T+1|T,H}$ requires selection of the parameter H. We use a cross-validation

Download English Version:

https://daneshyari.com/en/article/5096265

Download Persian Version:

https://daneshyari.com/article/5096265

Daneshyari.com