Journal of Econometrics 177 (2013) 134–152

Contents lists available at ScienceDirect

Journal of Econometrics

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

Optimal forecasts in the presence of structural breaks[☆]

M. Hashem Pesaran[a,b], Andreas Pick[c,d,*], Mikhail Pranovich[e]

[a] University of Southern California, USA
[b] Trinity College, Cambridge, United Kingdom
[c] Erasmus University Rotterdam, Netherlands
[d] De Nederlandsche Bank, Netherlands
[e] Joint Vienna Institute, Austria

ARTICLE INFO

Article history:
Available online 6 April 2013

JEL classification:
C22
C53

Keywords:
Forecasting
Structural breaks
Optimal weights
Robust optimal weights
Exponential smoothing

ABSTRACT

This paper considers the problem of forecasting under continuous and discrete structural breaks and proposes weighting observations to obtain optimal forecasts in the MSFE sense. We derive optimal weights for one step ahead forecasts. Under continuous breaks, our approach largely recovers exponential smoothing weights. Under discrete breaks, we provide analytical expressions for optimal weights in models with a single regressor, and asymptotically valid weights for models with more than one regressor. It is shown that in these cases the optimal weight is the same across observations within a given regime and differs only across regimes. In practice, where information on structural breaks is uncertain, a forecasting procedure based on robust optimal weights is proposed. The relative performance of our proposed approach is investigated using Monte Carlo experiments and an empirical application to forecasting real GDP using the yield curve across nine industrial economies.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

It is now widely recognized that parameter instability is an important source of forecast failure in macroeconomics and finance as documented by Pesaran and Timmermann (2002); Pesaran et al. (2006), Koop and Potter (2007); Giacomini and Rossi (2009); Inoue and Rossi (2011), among others. Clements and Hendry (1999, 2006); Rossi (2011) provide reviews. Broadly speaking, there are two basic approaches to modeling parameter instability: parameters are assumed to change either at discrete time intervals or continuously. Under the former, break dates are estimated and forecasts are typically constructed using the post-break observations.[1] Assuming that the break dates are accurately estimated, the forecasts based on observations after the last break are likely to be unbiased. However, as pointed out by Pesaran and Timmermann (2007), forecasts from the post-break window may not minimize the mean square forecast error (MSFE) as the estimation uncertainty may be large due to the relatively short post-break window. For this reason Pesaran and Timmermann (2007) suggest an optimal estimation window that may include pre-break observations. When the time and size of the break is uncertain, Pesaran and Timmermann (2007) consider averaging forecasts across estimation windows (AveW), which, as Pesaran and Pick (2011) show, improves forecasts without relying on estimates of break dates and sizes.

Under the continuously changing parameter model, the breaks are assumed to occur at every period, and observations are downweighted to take account of the slowly changing nature of the parameters. Within this framework, a prominent approach is exponential smoothing (ExpS), first proposed by Holt (1957); Brown (1959). Other approaches using Kalman filters have also been proposed as generalizations of ExpS. Hyndman et al. (2008) provide a comprehensive survey.

In this paper, we develop a unified approach to obtaining optimal forecasts under both types of structural breaks, focusing on one-step-ahead forecasts. We consider forecasts based on weighted observations as in the ExpS approach but derive weights that are optimal in the sense that the resulting forecasts minimize the MSFE. In the case of continuous breaks, the optimal weights approximate ExpS weights if T is large and the downweighting parameter of ExpS is not too close to unity. In contrast, when the breaks are assumed to occur at discrete time intervals the optimal weights can differ markedly from the ExpS weights. We show that, conditional on the break size and date, the optimal weights follow

☆ We would like to thank Don Harding and Ana Galvao for interesting discussions and two anonymous referees for their comments. Additionally, we are grateful for helpful comments from participants at the 31st Annual International Symposium on Forecasting in Prague, the 7th ECB workshop on forecasting techniques, the MMF/CEF workshop at Brunel University, the NESG meeting in Groningen, ESEM in Malaga, and a seminar at ESSEC.

* Correspondence to: P.O. Box 1738, 3000 DR Rotterdam, The Netherlands.
E-mail address: pick@ese.eur.nl (A. Pick).

[1] There are many statistical procedures that can be used for detection of break dates, such as Brown et al. (1975); Andrews (1993); Andrews et al. (1996); Bai and Perron (1998, 2003); Altissimo and Corradi (2003).

0304-4076/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jeconom.2013.04.002

a step function that allocates constant weights within regimes but different weights between regimes. A striking result emerges under multiple breaks: observations of the last regime that continues into the forecast period may not receive the highest weight. The intuition for this result is that the bias component of the MSFE can be reduced by giving the largest weights to observations in an earlier regime to counterbalance biases of the opposite sign in another regime.

In practice, dates and sizes of the breaks are unknown and must be estimated. As such estimates tend to be quite imprecise and their use in practice leads to a deterioration of forecasts, which can be quite substantial. In order to address this problem, we develop weights that are robust to the uncertainty that surrounds the dates and the sizes of the breaks. Robust optimal weights are derived by integrating the optimal weights with respect to uniformly distributed break dates. An interesting insight from these derivations is that the effect of uncertainty of the break size on the weights is of order T^{-2} if the break is in the slope coefficient, and of order T^{-3} if the break is in the error variances, where T is the full sample size that includes the pre-break observations. In contrast, the uncertainty around the break date is of order T^{-1}, which suggests that dating a break correctly is generally more important than knowing the precise size of the break.

We conduct Monte Carlo experiments that compare the forecasts from optimal and robust optimal weights to a range of alternative forecasting methods. It emerges that the key factor for the relative performance of different forecasting methods under a discrete break is the size of the break. A larger break leads to more precise estimates of the break date and improves forecasts that are conditional on these estimates, which include the optimal weights forecast, post-break forecasts, and optimal window forecasts. In contrast, when the break is small relative to the noise in the DGP, the robust optimal weights produce the best forecasts as they do not make use of the often imprecisely estimated break dates and sizes. When the break process is continuous, ExpS forecasts that estimate the down-weighting parameter perform well. However, even under the continuous break process the forecasts from the robust optimal weights perform well and in some settings provide the best forecasts.

We use the different methods considered in this paper to forecast real GDP using the slope of the yield curve across nine industrial economies over the period 1994Q1–2009Q4. The general finding is that breaks are difficult to estimate with sufficient accuracy and, similar to the Monte Carlo results, forecasts based on estimates of break dates perform poorly. Forecasts based on robust optimal weights deliver the largest improvements over forecasts based on equal weights, and these improvements are statistically significant.

The rest of the paper is set out as follows. Using a linear regression model, derivations of optimal weights under different break processes are set out in Section 2, and the MSFE outcomes are compared across different forecasting methods. Optimal weights that are robust to the uncertainty of the break process are motivated and derived in Section 3. Monte Carlo evidence on the comparative performance of the different forecasting methods is discussed in Section 4. Empirical results are presented in Section 5. The paper ends with some concluding remarks in Section 6. A few of the less essential derivations are collected in a mathematical appendix. Additional material can be found in a web supplement.

2. Optimal weights under different break processes

Consider the linear regression model

$$y_t = \boldsymbol{\beta}_t'\mathbf{x}_t + \sigma_t\varepsilon_t, \qquad \varepsilon_t \sim iid(0,1), \\ t = 1, 2, \ldots, T, T+1 \tag{1}$$

where $\mathbf{x}_t$ is a $k \times 1$ vector of stationary regressors, and the $k \times 1$ coefficient vector, $\boldsymbol{\beta}_t$, and the scalar error variance, σ_t^2, are subject to breaks. We consider two possible types of break processes. A continuous break process whereby $\boldsymbol{\beta}_t$ changes in every period by a relatively small amount. A prominent example is the random walk model

$$\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{S}_\beta \mathbf{v}_t, \quad \text{where } \mathbf{v}_t \sim iid(\mathbf{0}, \boldsymbol{I}_k),$$

where $\boldsymbol{I}_k$ is the identity matrix of order k, and the break variance, $\boldsymbol{\Sigma}_\beta = \boldsymbol{S}_\beta \boldsymbol{S}_\beta'$, is assumed to be small relative to σ_t^2.[2] Additionally, σ_t may be subject to a similar break process. Alternatively, the breaks could be discrete where the parameters change at a small number of distinct points in time, $T_{b,i}$, $i = 1, 2, \ldots, n$,[3]

$$\boldsymbol{\beta}_t = \begin{cases} \boldsymbol{\beta}_{(1)} & \text{for } 1 < t \le T_{b,1} \\ \boldsymbol{\beta}_{(2)} & \text{for } T_{b,1} < t \le T_{b,2} \\ \vdots & \\ \boldsymbol{\beta}_{(n+1)} & \text{for } T_{b,n} < t \le T. \end{cases}$$

In contrast to the continuously changing parameter model, the number of discrete breaks, n, is assumed to be small, although the break sizes, measured by $\left\|\boldsymbol{\beta}_{(i)} - \boldsymbol{\beta}_{(i-1)}\right\|$ could be large relative to σ_t. There are merits in both specifications, and a choice between them would depend on the particular data at hand.

We propose a general approach to achieve a minimum mean square forecast error (MSFE) under both break processes. We weight past observations by weights w_t in the estimation

$$\hat{\boldsymbol{\beta}}_T(\mathbf{w}) = \left(\sum_{t=1}^{T} w_t \mathbf{x}_t \mathbf{x}_t'\right)^{-1} \sum_{t=1}^{T} w_t \mathbf{x}_t \mathbf{y}_t,$$

subject to the restriction $\sum_{t=1}^{T} w_t = 1$. The weights $\mathbf{w} = (w_1, w_2, \ldots, w_T)'$ are chosen such that the resulting MSFE of the one-step ahead forecast, $\hat{y}_{T+1} = \hat{\boldsymbol{\beta}}_T'\mathbf{x}_{T+1}$, is minimized.

Closed form solutions under the continuous break process are only available when we simplify the model to one without time-varying regressors. In this setting the optimal weights recover the exponential smoothing forecast. For the discrete break process we derive new results for the same simple model and also for models with one or more regressors.

2.1. Optimal weights in a model with continuous breaks

Consider the following model

$$y_t = \beta_t + \sigma_\varepsilon \varepsilon_t, \tag{2}$$

where $\beta_t = \beta_{t-1} + \sigma_v v_t$, and ε_t and v_t are $iid(0, 1)$. The optimal weights for a one-step ahead forecast can be found by minimizing $\mathrm{E}(y_{T+1} - \Sigma_{t=1}^{T} w_t y_t)^2$ with respect to w_t, $t = 1, 2, \ldots, T$, subject to $\sum_{t=1}^{T} w_t = 1$. For a solution to this problem we first note that the forecast error is given by

$$e_{T+1} = y_{T+1} - \hat{\beta}_{T+1}(\mathbf{w}) = \beta_{T+1} - \mathbf{w}'\boldsymbol{\beta} + \sigma_\varepsilon(\varepsilon_{T+1} - \mathbf{w}'\boldsymbol{\varepsilon}),$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_T)'$. But using the random walk formulation of $\boldsymbol{\beta}$ we have $\boldsymbol{\beta} = \beta_0 \boldsymbol{\iota}_T + \sigma_v \mathbf{H}\mathbf{v}$, where $\mathbf{v} = (v_1, v_2, \ldots, v_T)'$,

$$\mathbf{H} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 1 & \cdots & 1 & 0 \\ 1 & 1 & \cdots & 1 & 1 \end{pmatrix}, \quad \text{and} \quad \boldsymbol{\iota}_T = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix}.$$

[2] The covariance matrix $\boldsymbol{\Sigma}_\beta$ is said to be small relative to σ_t if $\left\|\boldsymbol{\Sigma}_\beta\right\| / \sigma_t$ is small, where $\|\boldsymbol{A}\|^2 = \mathrm{tr}(\mathbf{A}\mathbf{A}')$ denotes the Euclidean norm of matrix $\mathbf{A}$.

[3] Note that parentheses around subscripts denote subsamples between breaks, such that β_t is the parameter at period t but $\beta_{(i+1)}$ the parameter after break i.

Download English Version:

https://daneshyari.com/en/article/5096264

Download Persian Version:

https://daneshyari.com/article/5096264

Daneshyari.com