International Journal of Forecasting 30 (2014) 144–160

Contents lists available at ScienceDirect

International Journal of Forecasting

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

A new structural break model, with an application to Canadian inflation forecasting

John M. Maheu [a,b,*], Yong Song [c,b]

[a] DeGroote School of Business, McMaster University, 1280 Main Street West, Hamilton, ON, Canada, L8S 4M4
[b] RCEA, Italy
[c] Economics Discipline Group, University of Technology, Sydney, Australia

ARTICLE INFO

Keywords:
Multiple change-points
Regime duration
Inflation targeting
Predictive density
MCMC

ABSTRACT

This paper develops an efficient approach to modelling and forecasting time series data with an unknown number of change-points. Using a conjugate prior and conditioning on time-invariant parameters, the predictive density and the posterior distribution of the change-points have closed forms. Furthermore, the conjugate prior is modeled as hierarchical in order to exploit the information across regimes. This framework allows breaks in the variance, the regression coefficients, or both. The regime duration can be modelled as a Poisson distribution. A new, efficient Markov chain Monte Carlo sampler draws the parameters from the posterior distribution as one block. An application to a Canadian inflation series shows the gains in forecasting precision that our model provides.

1. Introduction

This paper develops an efficient Bayesian approach to modelling and forecasting time series data with an unknown number of change-points. The approach simplifies structural break analysis and reduces the computational burden relative to existing approaches in the literature. A conjugate prior is modeled as hierarchical in order to exploit information across regimes. The regime duration can be inferred from a fixed structural change probability, or modelled as a Poisson distribution. Compared to existing time series models of Canadian inflation, including alternative structural break models, our specification produces superior density forecasts and point forecasts.

Accounting for structural instability in macroeconomic and financial time series models is important. Empirical applications by Clark and McCracken (2010), Geweke and Jiang (2011), Giordani, Kohn, and van Dijk (2007), Liu and Maheu (2008), Stock and Watson (1996), and Wang and Zivot (2000), among others, demonstrate a significant degree of instability.

The problem of forecasting in the presence of structural breaks has recently been addressed by Koop and Potter (2007), Maheu and Gordon (2008), Maheu and McCurdy (2009) and Pesaran, Pettenuzzo, and Timmermann (2006) using Bayesian methods. These approaches do provide feasible solutions, but they are all computationally intensive.

The purpose of this paper is to provide a change-point model which is suitable for out-of-sample forecasting and has the attractive features of the previous approaches, but which is computationally less demanding. The parameters in each regime are drawn independently from a hierarchical prior. This allows for learning about the structural change process and its effect on model parameters, and is convenient for computation. We introduce a new Markov chain Monte Carlo (MCMC) sampler for drawing all of the parameters, including the hierarchical prior, the parameters of the durations, the change-points, and the parameters characterizing each regime, from their posterior

* Corresponding author at: DeGroote School of Business, McMaster University, 1280 Main Street West, Hamilton, ON, Canada, L8S 4M4. Tel.: +1 905 525 9140x26198.
E-mail addresses: maheujm@mcmaster.ca (J.M. Maheu), ysong1@gmail.com (Y. Song).

0169-2070/$ – see front matter
http://dx.doi.org/10.1016/j.ijforecast.2013.06.004

distributions jointly. As a result, the mixing of the chain is better than that of a regular Gibbs sampling scheme as per Chib (1998). Lastly, different types of break dynamics, including having breaks in the variance, the regression coefficients or both, are nested in this framework.

We extend the work of Maheu and Gordon (2008) and Maheu and McCurdy (2009) in four directions. First, a conjugate prior for the parameters that characterize each regime is adopted. Conditional on this prior and the time-invariant parameters, the predictive density has a closed form, which reduces the computational burden compared to that of Maheu and Gordon (2008).[1] Second, a hierarchical structure for the conjugate prior is introduced to allow the pooling of information across regimes, as did Pesaran et al. (2006). Third, we show how the regime duration can be modeled as a Poisson distribution, which implies duration-dependent break probabilities. Lastly, we show how to produce the smoothed distribution of the change-points.

Koop and Potter (2007) also model regime durations, but they assume a heterogeneous distribution for the duration in each regime. Their approach augments the state space with regime durations, meaning that there are $O(T^2)$ states, which implies a large transition matrix. In contrast, we assume that the regime durations are drawn from the same distribution. This simplification results in the number of states in our model being $O(T)$. Koop and Potter (2007) assume that, after a structural change, the parameters in the new regime are related to those in the previous regime through a random walk. This path dependence in parameters increases the computation time further.

Different versions of our model are applied to a Canadian inflation series in order to investigate its dynamic stability. Canadian inflation is challenging to forecast, as inflation targeting was introduced in 1991. This raises the question of the usefulness of the data prior to this date for forecasting after 1991. We also show that incorporating exogenous subjective information from policy changes into our model can improve the forecasts further.

The log-predictive likelihood is used as the criterion for model comparison. The best model is the hierarchical model, which allows breaks in the regression coefficients and the variance simultaneously. This model provides large improvements compared to both linear no-break models and autoregressive benchmarks with a GARCH parametrization. A sub-sample analysis is consistent with the results from the full sample. We also show how exogenous information or variables can be incorporated in our framework for out-of-sample forecasting. A posterior analysis based on the optimal model identifies four major change-points in the Canadian inflation dynamics. The duration-dependent break probability is not a significant feature of the data.

[1] Maheu and Gordon (2008) assume a conditional conjugate prior and use Gibbs sampling to compute the predictive density. The computational benefits of our approach are that it requires a conjugate prior and a structural break process which are simplified relative to other models in the literature.

The paper is organized as follows. Section 2 introduces the model, and a MCMC method is proposed for efficient sampling from the posterior distribution. Section 3 extends the non-hierarchical prior to a hierarchical one in order to exploit the information across regimes. Different extensions of the hierarchical model are introduced in Section 4, including models with breaks only in the variance, breaks only in the regression coefficients, or independent breaks in both. A duration-dependent break probability is also modeled by assuming a Poisson distribution for the regime durations. Section 5 applies the model to a Canadian inflation time series. Finally, Section 6 concludes.

2. Structural break model with a conjugate prior

In what follows, we assume that two consecutive structural breaks define a regime. A regime consists of a set of contiguous data drawn from a data density using a fixed model parameter θ. Different regimes will have different values of θ, which is assumed to be drawn from a specified distribution. The number of observations in a regime denotes the duration of a regime. We discuss how to compute the posterior density of θ for each regime, as well as the predictive density. Section 2.1 then gives specifics and shows how all possible structural break points (regimes) can be integrated out in order to form predictions.

If time i is the starting point of the most recent regime, it is assumed that the data before time i are not informative for the posterior of the parameter θ which governs the current regime.

If the most recent break is at time i ($i \leq t$), then the duration of the current regime at time t is defined as $d_t = t - i + 1$. The duration is used as a state variable in what follows, for two reasons. First, we wish to study not only the forecasting problem but also the ex-post analysis of multiple change-points in-sample.[2] Second, working with d_t facilitates the modeling of regime durations directly, which we discuss later.

Formally, define d_t as the duration of the most recent regime up to time t, and $d_t \in \{1, \ldots, t\}$ by construction. If a break happens at time t, then $d_t = 1$. If $d_t = t$, then there is no break throughout the whole sample. Define $Y_{i,t} = (y_i, \ldots, y_t)$ for $1 \leq i \leq t$. If $i > t$, $Y_{i,t}$ is an empty set.

In order to form the predictive density for y_{t+1} conditional on the duration d_{t+1}, we require the posterior density based on data $Y_{1,t}$. Let the data density of y_{t+1}, given the model parameter θ and information set $Y_{1,t}$, be denoted as $p(y_{t+1} \mid \theta, Y_{1,t})$. There are two cases to consider. The first case is that the regime continues for one more period, while the second case is the occurrence of a structural change, with a new draw of the parameter θ occurring between t and $t + 1$. If $p(\theta)$ is the prior for θ, then, conditional on the duration d_{t+1}, the posterior is

$$p(\theta|d_{t+1}, Y_{1,t}) \equiv p(\theta|Y_{t-d_{t+1}+2,t})$$

$$\propto \begin{cases} p(y_{t-d_{t+1}+2}, \ldots, y_t|\theta)p(\theta) & d_{t+1} > 1 \\ p(\theta) & d_{t+1} = 1. \end{cases} \quad (1)$$

[2] Maheu and Gordon did not consider the smoothed distribution of breaks, but focused on the filtered distribution of change points.