Journal of Econometrics 177 (2013) 171–184

Contents lists available at ScienceDirect

Journal of Econometrics

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

Forecasting a long memory process subject to structural breaks

Cindy Shin-Huei Wang[a,b], Luc Bauwens[a], Cheng Hsiao[c,d,e,*]

[a] Université catholique de Louvain, CORE, B-1348 Louvain-La-Neuve, Belgium
[b] National Tsing Hua University, Department of Quantitative Finance, Taiwan
[c] University of Southern California, USA
[d] Hong Kong University of Science and Technology, Hong Kong
[e] WISE, Xiamen University, China

ARTICLE INFO

Article history:
Available online 17 April 2013

JEL classification:
C22
C53

Keywords:
Forecasting
Long memory process
Structural break
HAR model

ABSTRACT

We develop an easy-to-implement method for forecasting a stationary autoregressive fractionally integrated moving average (ARFIMA) process subject to structural breaks with unknown break dates. We show that an ARFIMA process subject to a mean shift and a change in the long memory parameter can be well approximated by an autoregressive (AR) model and suggest using an information criterion (AIC or Mallows' C_p) to choose the order of the approximate AR model. Our method avoids the issue of estimation inaccuracy of the long memory parameter and the issue of spurious breaks in finite sample. Insights from our theoretical analysis are confirmed by Monte Carlo experiments, through which we also find that our method provides a substantial improvement over existing prediction methods. An empirical application to the realized volatility of three exchange rates illustrates the usefulness of our forecasting procedure. The empirical success of the HAR-RV model can be explained, from an econometric perspective, by our theoretical and simulation results.

1. Introduction

Macroeconomic and financial time series are subject to occasional structural breaks (see Stock and Watson (1996) and Pesaran and Timmermann (2005)). It is often argued that ignoring the presence of breaks can lead to seriously biased estimates and forecasts (see Clements and Hendry (1998)). Accordingly, a conventional approach to forecast a time series with breaks is first to determine when the most recent break occurred and then to use the post-break data to estimate the forecasting model. Nevertheless, Pesaran and Timmermann (2005, 2007) showed that such an approach does not necessarily yield the optimal forecasting performance, especially when the time series are subject to multiple breaks, due to the difficulty in estimating the timing of breaks. Moreover, they illustrate that pre-break data can be useful for forecasting the after-break outcomes, provided that the break is not too large.

Many researchers use the autoregressive fractionally integrated moving average process of order p, d, q, denoted as ARFIMA (p, d, q), or $I(d)$ process, to model and forecast time series, where the differencing parameter d is a fractional number between -0.5 and 0.5. The main feature of the stationary $I(d)$ process is that its autocovariance function declines at a hyperbolic rate, slower than the geometric rate of stationary ARMA processes. For example, Ding et al. (1993) and Bollerslev and Mikkelsen (1996) showed that the persistence in stock market volatility could be well described by a long memory process. These findings further induced Hidalgo and Robinson (1996) to consider the issue of structural stability of a regression model with a long memory error term. However, Kuan and Hsu (1998) find that the Hidalgo and Robinson (1996) test could have large size distortions. Additionally, extending the results of Nunes et al. (1995) and Kuan and Hsu (1998) show that the conventional break tests for stationary long memory processes may misleadingly infer a structural break when there is none. Because of the possibility of misleading inference by the existing structural change tests for long memory processes, scant attention has been paid to suggesting optimal methods to forecast long memory processes in the presence of structural breaks.

The purpose of this paper is to propose an easy-to implement approach for forecasting a long memory process that may be subject to structural breaks. The conventional forecasting method based on post-break data can be suboptimal because the break detection approach may lead to spurious conclusions concerning the number of breaks even when there is none (e.g. Granger and Hyung (2004)). Moreover, Granger and Hyung (2004) and Choi et al. (2010) also showed that an increase in the number of mean breaks makes the memory of the process seemingly

* Correspondence to: Department of Economics, University of Southern California, Los Angeles, USA.
E-mail address: chsiao@usc.edu (C. Hsiao).

0304-4076/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jeconom.2013.04.006

more persistent. Choi et al. (2010) use a break-adjusted forecast method to reduce the forecast error for several realized volatility series, which are modelled as long memory processes with breaks, however, their method still depends on the knowledge of the accurate break dates. To avoid the risk of misleading inference on the break date or imprecise estimation of the fractional parameter d, we note that first, Granger (1980) has shown that when a long memory process has a break in the parameter d, the complete time series can be represented by another long memory process with memory parameter d^* that is a linear combination of the pre-and the post-break memory parameters d_1 and d_2. Second, Poskitt (2007) and Wang and Hsiao (2012) have shown that a stationary ARFIMA (p, d, q) process can be approximated by an autoregressive process with an order increasing at a certain rate of the sample size T. We therefore suggest using an autoregressive approximation to predict the outcome for long memory processes with breaks. We justify the use of an autoregression approximation when the time series follows an ARFIMA (p, d, q) process, for two types of structural breaks: a change in the long memory parameter and/or a shift in the mean when the shift size is within some magnitude of the standard deviation of the random noise. Furthermore, an AR approximation approach also has the advantage of avoiding the inaccurate parameter estimation and the break locations.

An important issue in finite sample is to select the appropriate order of the AR model. We suggest using Mallow's C_p criterion to select the order of an AR(k) model fitted to a long memory process with structural change. Our Monte Carlo experiments show that the lag length based on Mallow's criterion to approximate the ARFIMA process with structural breaks is usually small in a sample of two hundred observations and the residual variance estimate is very close to the true error variance. Our simulation experiments confirm the theoretical analysis clearly by demonstrating that an AR-approximation forecast method for forecasting a long memory process with structural breaks outperforms conventional methods, namely the two naive ARFIMA-based methods, the post-break and Tiao and Tsay (1994) adaptive forecasting methods, even in cases where the structure of an ARFIMA model, including its parameters and lag orders, changes dramatically after breaks. Furthermore, for the special case in which the structural breaks take place immediately prior to the forecast period, our AR-approximation also performs better. An empirical forecasting exercise of realized volatility for the DEM/USD, USD/YEN and GBP/USD spot exchange rates shows that our AR-approximation dominates the existing methods.

We present the basic model and theoretical results in Section 2. Section 3 provides the mean-squared prediction errors of forecasts generated by an AR approximation, a post-break model, and two naive ARFIMA-based forecast models. Section 4 provides the finite sample simulation comparison. Section 5 provides the comparison of different methods for predicting volatilities. Concluding remarks are in Section 6. Proofs are in the Appendix.

2. The model and theoretical results

2.1. The basic model

Let

$$(1-L)^{d_1}(\eta_t^{(1)}-\mu_1)=e_t \tag{1}$$

$$(1-L)^{d_2}(\eta_t^{(2)}-\mu_2)=e_t \tag{2}$$

be two ARFIMA $(0, d, 0)$ processes where (i) $d_1, d_2 \in (-0.5, 0.5)$, $d_1 \neq 0$ and $d_2 \neq 0$ are differencing parameters; (ii) e_t is an independently and identically distributed process, with $E(e_t)=0$, $E(e_t^2)=\sigma^2$, and $E\left(e_t^4\right)<\infty$. We suppose the T time series observations, η_t take the form

$$\eta_t=\eta_t^{(1)} \quad \text{for } t=1,2,\ldots,T_1,$$

and

$$\eta_t=\eta_t^{(2)} \quad \text{for } t=T_1+1,\ldots,T,$$

where $T_1=\kappa T$, $\kappa \in (0,1)$. We consider two scenarios:

case I: changes in the differencing parameter only, i.e., $d_1 \neq d_2, \mu_1=\mu_2=\mu$.

case II: changes in both differencing parameter and mean, i.e., $d_1 \neq d_2, \mu_1 \neq \mu_2$.

We focus on breaks in the mean and long memory parameters in the DGP. We could, of course, examine ARFIMA(p, d, q) models with breaks in the coefficients of the AR and MA terms. However, this would substantially complicate the notations and derivations without gaining insight.

Lemma 1. [1] *When the DGP satisfies the basic model, with $T_1=\kappa T$, $\kappa \in (0,1)$, the observed data can be represented by a long memory process with long memory parameter d^* that is a linear combination of the pre-and post-break parameters d_1 and d_2, and with a mean μ^* that is a linear combination of the pre-and post-break means μ_1 and μ_2, that is*

$$(1-L)^{d^*}(\eta_t-\mu^*)=e_t, \tag{3}$$

where $\mu^*=\kappa\mu_1+(1-\kappa)\mu_2$, $d^*=\lambda d_1+(1-\lambda)d_2$, $0\le\lambda\le 1$. When T is finite, $\kappa\to 0, \lambda\to 0$, and $\kappa\to 1, \lambda\to 1$. When $T\to\infty$, $d^*=max(d_1,d_2)$ and $\mu^*=\kappa\mu_1+(1-\kappa)\mu_2$.

Brockwell and Davis (1991) have shown that a long memory process can be represented by an infinite order autoregressive process,

$$\eta_t=\mu+\sum_{j=1}^{\infty}\beta_j\eta_{t-j}+e_t, \tag{4}$$

where $\beta_j=\Gamma(j-d)/[\Gamma(j+1)\Gamma(d)]$, $d\in(-0.5,0.5)$, $d\neq 0$, and e_t is a zero mean white noise process with constant variance σ^2. Poskitt (2007) and Wang and Hsiao (2012) have shown that η_t can be approximated by an ever increasing order autoregressive model, AR(k), as T increases,

$$\eta_t=e_{t,k}+\sum_{j=1}^{k}\beta_{jk}\eta_{t-j}+\beta_{0k}, \tag{5}$$

where $e_{t,k}$ is the prediction error and $\beta_{jk}, j=0,1,2,\ldots,k$ are the coefficients of the minimum mean squared predictor of η_t based only on a constant term and the past observations $\eta_{t-1},\eta_{t-2},\ldots,\eta_{t-k}$.

Theorem 1. [2] *When the DGP satisfies the basic model, with $T_1=\kappa T$, $\kappa\in(0,1)$, when $d^*\in(-0.5,0.5)$ there exists an AR(k) approximation of η_t, as $T\to\infty$, $k=O(T^r)$, $r>2d^*/(1+2d^*)$, such that*

1. $\|\widehat{\beta}(k)-\beta(k)\|=O_p((k\log T/T)^{0.5-d^*})$,

2. $\widehat{\sigma}_{t,k}^2=\frac{1}{T-k}\sum_{t=k+1}^{T}\widehat{e}_{t,k}^2=\sigma^2+O_p(k^{-2d^*-1}T^{2d^*})=\sigma^2+o_p(1)$, where $\widehat{\beta}(k)$ is the OLS estimator of $\beta(k)$ and $\widehat{e}_{t,k}$ the OLS residual.

Given Lemma 1 and Theorem 1, we suggest using an AR(k) model to approximate the data generating process of η_t and using it to generate post-sample predictions.

[1] Lemma 1 is a generalization to that of Granger (1980) with less restrictive conditions on the distribution of the DGP and also allows breaks in the mean.

[2] The rate of convergence of $\|\widehat{\beta}(k)-\beta(k)\|$ is shown faster than that of Poskitt (2007) or Wang and Hsiao (2012) because we use a different methodology.