ARTICLE IN PRESS

Economics Letters ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at SciVerse ScienceDirect

Economics Letters

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

A note on estimating a structural change in persistence

Mohitosh Kejriwal[a], Pierre Perron[b,*]

[a] Krannert School of Management, Purdue University, 403 West State Street, West Lafayette IN 47907, United States

[b] Department of Economics, Boston University, 270 Bay State Rd., Boston MA 02215, United States

ARTICLE INFO

Article history:
Received 30 May 2012
Accepted 19 July 2012
Available online xxxx

JEL classification:
C22

Keywords:
Structural change
Break date
Persistence
Unit root
Autoregressive model

ABSTRACT

This paper studies issues related to the estimation of a structural change in the persistence of a univariate time series. The break is such that the process has a unit root [i.e., is $I(1)$] in the pre-break regime but reverts to a stationary [i.e., $I(0)$] process in the post-break regime or vice versa. Chong (2001) develops the limit theory for the estimation of such autoregressive processes and shows that the rate of convergence of the breakpoint estimator in the $I(1)$–$I(0)$ case is faster than that in the $I(0)$–$I(1)$ case, which enables the break date to be estimated much more precisely in the former case. In this paper, we show that the faster rate is an artifact of the assumed data generating process that is characterized by a spurious jump at the true breakpoint. Based on a reformulation that avoids this jump, the same rate of convergence prevails in both cases. An important implication of this result is that existing confidence intervals in the $I(1)$–$I(0)$ case have asymptotically zero coverage rates when the break magnitude is fixed. A small simulation study confirms the relevance of the asymptotic results in finite samples.

1. Introduction

The issues of detection and estimation of structural change constitute an important aspect of the analysis of economic and financial time series data. It is now well known that failure to account for parameter instability can result in biased parameter estimates and inaccurate forecasts. Accordingly, numerous procedures have been developed for detecting such instabilities as well as dating them in order to identify their potential sources. A recent strand of this literature has been concerned with developing tests and estimators for a change in the persistence of a univariate time series. Here, the change is characterized by a shift from a unit root process [$I(1)$] to a stationary process [$I(0)$] or vice versa at some unknown date over the sample under consideration. A wide variety of tests have been proposed to evaluate the null hypothesis of stable persistence against the alternative that there is a one-time shift in persistence at an unknown date. These tests are reviewed in Perron (2006) and we refer to this paper for appropriate references. Tests allowing for the possibility of multiple structural changes have been proposed by Leybourne et al. (2007) and Kejriwal et al. (forthcoming).

Conditional on the presence of a break, the goal is to estimate the breakpoint and form confidence intervals. Chong (2001, CH hereafter) studies the asymptotic properties of the estimated parameters including their limit distribution in the first order autoregressive model with a single break in persistence. One of the main results emanating from his analysis is that the rate of convergence of the estimated breakpoint is much faster in the $I(1)$–$I(0)$ scenario than that in the $I(0)$–$I(1)$ scenario, with the implication that the break date is estimated with much more precision in the former case. More formally, the difference between the estimated and the true break date is strictly bounded stochastically (i.e., is strictly $O_p(1)$) in the latter case while it converges to zero in the former case (i.e., is $o_p(1)$). This finding is also crucial for developing the corresponding limit distribution which forms the basis for the proposed confidence intervals.

This note reconsiders the analysis in CH for the $I(1)$–$I(0)$ case and shows that the faster convergence rate is an artifact of the assumed data generating process that is characterized by a spurious jump at the true breakpoint. Based on a reformulation that avoids this jump, the same rate of convergence prevails in both cases. An important implication of this result is that existing confidence intervals in the $I(1)$–$I(0)$ case have asymptotically zero coverage rates when the break magnitude is fixed. A small simulation study confirms the relevance of the asymptotic results in finite samples.

The rest of the paper is organized as follows. Section 2 presents the model and states the relevant assumptions. Section 3 details the asymptotic properties of the parameter estimates in our framework and includes a small simulation study. Section 4 concludes. All proofs are included in a separate mathematical appendix available online on the journal's website. As a matter

* Corresponding author. Tel.: +1 617 353 3026; fax: +1 617 353 4449.
E-mail addresses: mkejriwa@purdue.edu (M. Kejriwal), perron@bu.edu (P. Perron).

0165-1765/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.econlet.2012.07.020

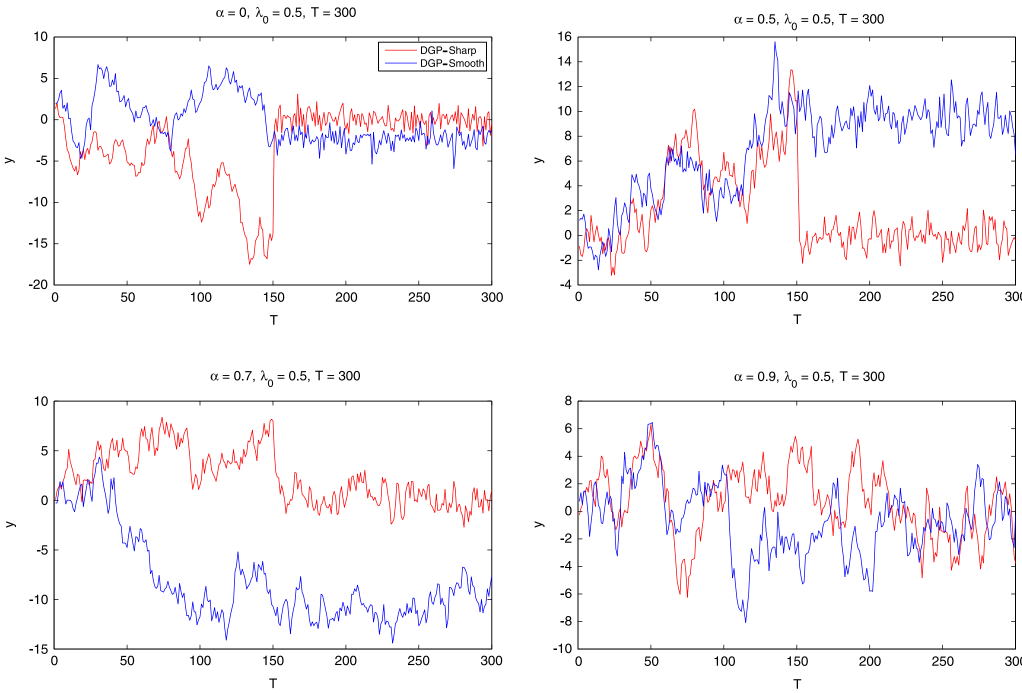

Fig. 1. Sharp versus smooth DGPs.

of notation, $\xrightarrow{p}$ and $\xrightarrow{d}$ denote convergence in probability and distribution respectively, and $\Rightarrow$ denotes weak convergence of the associated probability measures.

2. The model and assumptions

Consider a scalar random variable y_t generated as

$$\begin{aligned} y_t &= \mu_1 + u_t, \quad u_t = u_{t-1} + e_t \text{ for } t = 1, \ldots, [T\lambda_0] \\ y_t &= \mu_2 + u_t, \quad (1) \\ u_t &= u_{[T\lambda_0]} + \alpha(u_{t-1} - u_{[T\lambda_0]}) + e_t \text{ for } t = [T\lambda_0] + 1, \ldots, T \end{aligned}$$

where λ_0 denotes the true unknown break fraction while $k_0 = [T\lambda_0]$ denotes the true break date. Our analysis is based on the following assumptions.

Assumption 1. u_0 is a constant or an $O_p(1)$ random variable.

Assumption 2. $e_t \sim$ i.i.d.$(0, \sigma^2)$ for all t, $0 < \sigma^2 < \infty$, and $E(e_t^4) < \infty$.

Assumption 3. $\lambda_0 \in [\lambda_L, \lambda_H] \subset (0, 1)$.

Assumption 4. $|\alpha| < 1$.

These assumptions closely resemble those made in CH. Note that our framework is slightly more general in that we allow for a time varying mean. While the analysis could potentially be further generalized by allowing for more general error processes to account for higher order serial correlation, we have chosen to focus on the simple AR(1) model as well as retain CH's assumptions in order to facilitate a clearer comparison with his results. The difference between the data generating process (DGP) given by (1) and the one considered by CH lies in the way the post-break observations are generated. Specifically, CH assumes that the observations for $t = k_0 + 1, \ldots, T$ are generated according to

$$u_t = \alpha u_{t-1} + e_t. \quad (2)$$

To appreciate how different the two DGPs are, consider the special case $\mu_1 = \mu_2 = 0$ studied by CH. Based on (1), we can write

$$y_t = y_{[T\lambda_0]} + \sum_{j=0}^{t-[T\lambda_0]-1} \alpha^j e_{t-j} \quad \text{for } t = k_0 + 1, \ldots, T$$

so that the mean of the post-break data, conditional on the data in the pre-break regime, is given by

$$E(y_t | y_0, \ldots, y_{[T\lambda_0]}) = y_{[T\lambda_0]}.$$

This fact ensures a joining up of the $I(1)$ and $I(0)$ regimes. If, instead, the DGP for the post-break regime is formulated by (2), we obtain, for $t = k_0 + 1, \ldots, T$,

$$\begin{aligned} y_t &= \alpha^{t-[T\lambda_0]} y_{[T\lambda_0]} + \sum_{j=0}^{t-[T\lambda_0]-1} \alpha^j e_{t-j} \\ E(y_t | y_0, \ldots, y_{[T\lambda_0]}) &= \alpha^{t-[T\lambda_0]} y_{[T\lambda_0]}. \quad (3) \end{aligned}$$

In (3), the post-break conditional mean is a function of α. It is this fact that induces a spurious jump at the true break date k_0. In the extreme case, when $\alpha = 0$, the conditional mean is zero while for a positive α, the conditional mean exhibits a geometric decay with the speed of decay being governed by the value of α. To further illustrate this phenomenon, we simulate a sample of length $T = 300$ with a mid-sample break from each of the two DGPs assuming $\mu_1 = \mu_2 = 0$ and an i.i.d. standard normal distribution for $\{e_t\}$. The two samples are plotted in Fig. 1 for $\alpha = 0, 0.5, 0.7, 0.9$. For $\alpha < 0.9$, the spurious jump is evident for DGP (2). For $\alpha = 0.9$, the jump is less prominent, a feature that also follows from the conditional mean expression (3). The DGP given by (1) that avoids this jump was also adopted by Busetti and Taylor (2004) and Leybourne et al. (2007) in the context of developing procedures for detecting shifts in persistence (see also p. 278 of Banerjee et al., 1992).

The method of estimation is ordinary least squares (OLS). The regression model that forms the basis for estimation specifies that the dependent variable y_t is regressed on a constant and its lag y_{t-1} in each regime. For a generic break fraction λ (and associated break date $k = [T\lambda]$), the OLS estimates of the parameters are given by

Please cite this article in press as: Kejriwal, M., Perron, P., A note on estimating a structural change in persistence. Economics Letters (2012), doi:10.1016/j.econlet.2012.07.020

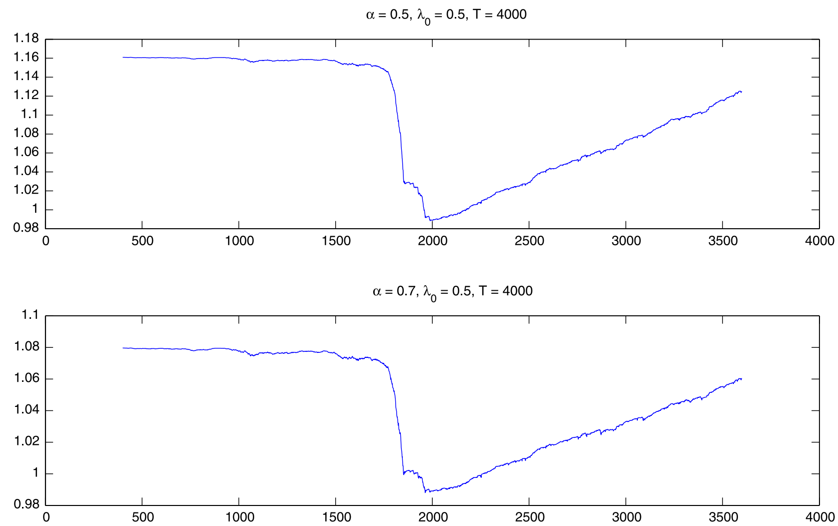

Fig. 2. $T^{-1}SSR(\lambda)$ for $T = 4000$.

$$\hat{\alpha}_1(\lambda) = \frac{\sum_{t=2}^{[T\lambda]} (y_t - \bar{y}_1)(y_{t-1} - \bar{y}_{1,-1})}{\sum_{t=2}^{[T\lambda]} (y_{t-1} - \bar{y}_{1,-1})^2},$$

$$\hat{\alpha}_2(\lambda) = \frac{\sum_{t=[T\lambda]+1}^{T} (y_t - \bar{y}_2)(y_{t-1} - \bar{y}_{2,-1})}{\sum_{t=[T\lambda]+1}^{T} (y_{t-1} - \bar{y}_{2,-1})^2}$$

where $\bar{y}_1 = (k-1)^{-1} \sum_{t=2}^{k} y_t$, $\bar{y}_{1,-1} = (k-1)^{-1} \sum_{t=2}^{k} y_{t-1}$, $\bar{y}_2 = (T-k)^{-1} \sum_{t=k+1}^{T} y_t$, $\bar{y}_{2,-1} = (T-k)^{-1} \sum_{t=k+1}^{T} y_{t-1}$. The breakpoint estimator is then obtained as $\hat{\lambda} = \text{argmin}_{\lambda \in (0,1)} SSR(\lambda)$, where

$$SSR(\lambda) = \sum_{t=2}^{[T\lambda]} [y_t - \bar{y}_1 - \hat{\alpha}_1(\lambda)(y_{t-1} - \bar{y}_{1,-1})]^2 + \sum_{t=[T\lambda]+1}^{T} [y_t - \bar{y}_2 - \hat{\alpha}_2(\lambda)(y_{t-1} - \bar{y}_{2,-1})]^2.$$

The final estimates of the autoregressive parameters are obtained as $\hat{\alpha}_1 = \hat{\alpha}_1(\hat{\lambda})$ and $\hat{\alpha}_2 = \hat{\alpha}_2(\hat{\lambda})$.

3. Asymptotic results

We now present asymptotic results pertaining to the consistency, rate of convergence and limit distribution of the parameter estimates when the data are generated by (1). Section 3.1 considers consistency and rate of convergence while Section 3.2 deals with confidence intervals for the breakpoint.

3.1. Consistency and rate of convergence

To study the consistency of $\hat{\lambda}$, define the set

$$A_{-}^{T} = [\lambda_L, \lambda_0) \cap \left\{ \lambda = \lambda_0 - \frac{j}{T}; \; j = 1, 2, \ldots, [\lambda_0 T], \; \lim_{T\to\infty} \frac{j}{\sqrt{T}} = \infty \right\}.$$

As in CH, the set A_{-}^{T} excludes sequences of λ that converge at a rate faster than $T^{1/2}$ to λ_0. We work with this set since, as shown below, the probability limit of $T^{-1}SSR(\lambda)$ is discontinuous at the left neighborhood of λ_0. An important difference from CH's analysis is that he shows, under his assumed DGP, that the discontinuity occurs both at the left and right of λ_0. We first state the following lemma that characterizes the asymptotic behavior of $T^{-1}SSR(\lambda)$ and thereby establishes the consistency of $\hat{\lambda}$.

Lemma 1. *Under Assumptions 1–4 with $\{y_t\}$ generated by (1),*

(a) $T^{-1}SSR(\lambda) \xrightarrow{p} \sigma^2 + \frac{(1-\alpha)(1-\lambda_0)\sigma^2}{1+\alpha}$ *for* $\lambda \in A_{-}^{T}$.

(b) $\sup_{\lambda_0 \le \lambda \le \lambda_H} \left| T^{-1}SSR(\lambda) - \sigma^2 - \frac{(1-\alpha)(\lambda-\lambda^0)\sigma^2}{1+\alpha} \right| = o_p(1)$.

(c) *The asymptotic transition from λ_0 to A_{-}^{T} is monotonically increasing.*

According to Lemma 1, $T^{-1}SSR(\lambda)$ converges to a constant greater than σ^2 for $\lambda \in A_{-}^{T}$ while it converges uniformly to an upward sloping linear function for $\lambda_0 \le \lambda \le \lambda_H$. Note that in CH, not only is the function discontinuous on both sides of λ_0, the limits are also random due to the presence of a Brownian motion term (see Eqs. (16) and (17) in CH). Fig. 2 presents a simulation to illustrate the behavior of $T^{-1}SSR(\lambda)$ with i.i.d. normal errors for $\alpha = 0.5, 0.7$ and $T = 4000$. The plot shows that the function is relatively flat to the left of λ_0, takes a plunge near λ_0, and is linear and upward sloping for $\lambda > \lambda_0$, thus providing support for our theory. In fact, the plot is simply a mirror image of the plot for the $I(0)$–$I(1)$ case (see Fig. 2 in CH).

Next, to show that the rate of convergence of $\hat{\lambda}$ is strictly T, we first prove, using a contradiction argument, that $T(\hat{\lambda} - \lambda_0)$ is stochastically bounded, i.e., $O_p(1)$ in a weak sense. Subsequently, we establish that $\lim_{T\to\infty} P(\hat{k} = k_0) \neq 1$, where $\hat{k} = [T\hat{\lambda}]$. The rate of convergence is therefore the same as that for the $I(0)$–$I(1)$ case. This rate is still fast enough to ensure that the limit distributions of the autoregressive parameter estimates are the same as those that apply in the known break date case. We collect these results in the following theorem.

Theorem 1. *Under Assumptions 1–4, we have*

(a) $\hat{\lambda} \xrightarrow{p} \lambda_0$

(b) $T(\hat{\lambda} - \lambda_0) = O_p(1)$ *(in a strict sense, i.e., not $o_p(1)$)*

(c) $T(\hat{\alpha}_1 - 1) \Rightarrow \left(\int_0^{\lambda_0} \widetilde{W}(r)^2 \right)^{-1} \int_0^{\lambda_0} \widetilde{W}(r) dW(r)$, *where $\widetilde{W}(\cdot)$ is a standard Brownian motion demeaned over $[0, \lambda_0]$.*

(d) $\sqrt{T}(\hat{\alpha}_2 - \alpha) \xrightarrow{d} N\left(0, (1-\alpha^2)(1-\lambda_0)^{-1}\right)$.

Please cite this article in press as: Kejriwal, M., Perron, P., A note on estimating a structural change in persistence. Economics Letters (2012), doi:10.1016/j.econlet.2012.07.020

3.2. Confidence intervals for the breakpoint

To develop the limit distribution of the breakpoint estimator, CH adopts a shrinking break framework whereby the magnitude of the break shrinks to zero at a rate sufficiently slow to still retain the consistency of the estimator. The assumption of shrinking break facilitates the application of a functional central limit theorem that makes the limit independent of the distribution of the errors. Specifically, CH assumes that $\alpha_T = 1 - [\sqrt{Tg(T)}]^{-1}$, where $g(T) > 0$, $g(T) \to \infty$, $g(T)/\sqrt{T} \to 0$ as $T \to \infty$.

In what follows, we show that the confidence interval proposed in CH have a zero asymptotic coverage rate under a fixed break magnitude. To see this, observe that his confidence interval is constructed based on the result that $(1-\alpha_T)^2 T^2(\hat{\lambda} - \lambda_0) = O_p(1)$ (see Theorem 4 in CH). Given that his limit distribution is symmetric, a particular breakpoint, say, λ_C, will be included in a $100(1-\beta)\%$ confidence interval if

$$\left|(1-\alpha_T)^2 T^2(\hat{\lambda} - \lambda_C)\right| \leq cv(1-\beta/2) \quad (4)$$

where $cv(1-\beta/2)$ is the $(1-\beta/2)$ quantile of the limit distribution of $(1-\alpha_T)^2 T^2(\hat{\lambda} - \lambda_0)$. From Theorem 1, $\left|\hat{\lambda} - \lambda_0\right| = O_p(T^{-1})$ so that with a fixed break, $\left|(1-\alpha_T)^2 T^2(\hat{\lambda} - \lambda_0)\right| = O_p(T)$ which diverges for all $\lambda_0 \in (0, 1)$. Hence, the corresponding confidence interval has a zero asymptotic coverage rate. This is stated in the following corollary.

Corollary 1. *Suppose Assumptions 1–4 hold, $\{y_t\}$ generated by (1) and the magnitude of the break is fixed. Then the confidence interval based on Theorem 4 in CH has a zero asymptotic coverage rate.*

In order to assess the finite sample adequacy of the confidence interval proposed in CH when data are generated by (1), we conduct a small Monte Carlo study. We set $\mu_1 = \mu_2 = 0$ and consider three values of the autoregressive parameter in the $I(0)$ regime: $\alpha = 0.5, 0.7, 0.9$. Three sample sizes are considered: $T = 200, 400, 600$. The errors $\{e_t\}_{t=1}^T$ are i.i.d. $N(0, 1)$ random variables. Table 1 presents the empirical coverage rates of the 90% confidence interval given by (4) for three values of the true break location: $\lambda_0 = 0.3, 0.5, 0.7$. As predicted by the theory, the coverage rates are very low in all cases with the maximum coverage attained across all configurations being only 32%. Moreover, given a value of α, the coverage rates decline as the sample size increases, a feature that is again in accordance with the theory. These results clearly illustrate that a spurious jump associated with (2) lead to confidence intervals that are too optimistic in that they are unlikely to cover the true value of the break location in practice.

Table 1
Empirical coverage rates of 90% asymptotic confidence intervals.

α	T								
	200	400	600	200	400	600	200	400	600
	$\lambda_0 = 0.3$			$\lambda_0 = 0.5$			$\lambda_0 = 0.7$		
0.5	0.08	0.02	0.02	0.09	0.03	0.03	0.05	0.03	0.03
0.7	0.08	0.02	0.02	0.07	0.02	0.01	0.07	0.02	0.01
0.9	0.32	0.11	0.05	0.24	0.07	0.03	0.22	0.08	0.04

4. Conclusion

This paper considers issues related to the estimation of a structural change in the persistence of a univariate time series. Existing results seem to suggest that it is much easier to estimate the break date when the pre-break regime is $I(1)$ as opposed to $I(0)$ owing to a faster rate of convergence of the estimated breakpoint. We show that this finding is an artifact of the assumed data generating process that is affected by a spurious jump at the true breakpoint. A reformulation that avoids this jump delivers the same rate of convergence in both the $I(1)$–$I(0)$ and $I(0)$–$I(1)$ cases. Moreover, existing confidence intervals are shown to have zero coverage rates asymptotically for a fixed magnitude of the break. A comprehensive analysis of the problem of detecting structural changes in persistence including the limit distributions of the parameter estimates in general time series models is currently under investigation by the authors.

Appendix. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.econlet.2012.07.020.

References

Banerjee, A., Lumsdaine, R., Stock, J., 1992. Recursive and sequential tests of the unit root and trend break hypotheses: theory and international evidence. Journal of Business and Economic Statistics 10, 271–288.

Busetti, F., Taylor, A.M.R., 2004. Tests of stationarity against a change in persistence. Journal of Econometrics 123, 33–66.

Chong, T.T.L., 2001. Structural change in AR(1) models. Econometric Theory 17, 87–155.

Kejriwal, M., Perron, P., Zhou, J., 2012. Wald tests for detecting multiple structural changes in persistence. Econometric Theory (forthcoming), Available on CJO http://dx.doi.org/10.1017/S0266466612000357.

Leybourne, S.J., Kim, T., Taylor, A.M.R., 2007. Detecting multiple changes in persistence. Studies in Nonlinear Dynamics & Econometrics 11 (3), Article 2.

Perron, P., 2006. Dealing with structural breaks. In: Patterson, K., Mills, T.C. (Eds.), Palgrave Handbook of Econometrics. Palgrave Macmillan, pp. 278–352.

Please cite this article in press as: Kejriwal, M., Perron, P., A note on estimating a structural change in persistence. Economics Letters (2012), doi:10.1016/j.econlet.2012.07.020