Journal of Econometrics 184 (2015) 262–279

Contents lists available at ScienceDirect

Journal of Econometrics

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

Confidence sets for the date of a break in level and trend when the order of integration is unknown☆

David I. Harvey *, Stephen J. Leybourne

School of Economics, University of Nottingham, UK

ARTICLE INFO

Article history:
Received 3 October 2013
Received in revised form
8 June 2014
Accepted 11 September 2014
Available online 20 September 2014

JEL classification:
C22

Keywords:
Level break
Trend break
Stationary
Unit root
Locally best invariant test
Confidence sets

ABSTRACT

We propose methods for constructing confidence sets for the timing of a break in level and/or trend that have asymptotically correct coverage for both I(0) and I(1) processes. These are based on inverting a sequence of tests for the break location, evaluated across all possible break dates. We separately derive locally best invariant tests for the I(0) and I(1) cases; under their respective assumptions, the resulting confidence sets provide correct asymptotic coverage regardless of the magnitude of the break. We suggest use of a pre-test procedure to select between the I(0)- and I(1)-based confidence sets, and Monte Carlo evidence demonstrates that our recommended procedure achieves good finite sample properties in terms of coverage and length across both I(0) and I(1) environments. An application using US macroeconomic data is provided which further evinces the value of these procedures.

1. Introduction

It has now been widely established that structural change in the time series properties of macroeconomic and financial time series is commonplace (see, *inter alia*, Stock and Watson, 1996), and much work has been devoted to this area of research in the literature. Focusing on the underlying trend function of a series, the primary issues to be resolved when considering the possibility of structural change are whether a break is present, and, if so, when the break occurred. The focus of this paper concerns the latter issue regarding the timing of the break, and is therefore complementary to procedures that focus on break detection. A proper understanding of the likely timing of a break in the trend function is crucial for modelling and forecasting efforts, and is also of clear importance when attempting to gain economic insight into the cause and impact of a break. While a number of procedures exist to determine a point estimate of a break in level and/or trend, this paper concentrates on ascertaining the degree of uncertainty surrounding break date estimation by developing procedures for calculating a confidence set for the break date, allowing practitioners to identify a valid set of possible break points with a specified degree of confidence.

The methodology of Bai (1994) allows construction of a confidence set for a break in level in a time series, extended in Bai (1997) to allow for a break in trend, with the confidence set comprised of a confidence interval surrounding an estimated break point, with the interval derived from the asymptotic distribution of the break date estimator. However, as Elliott and Müller (2007) [EM] argue, the asymptotic theory employed in this approach relies on the break magnitude being in some sense "large", in that the magnitude can be asymptotically shrinking only at a rate sufficiently slow to permit break detection procedures to have power close to one, so that although the magnitude is asymptotically vanishing, the break is still large enough to be readily detectable. EM argue that in many practical applications it is "small" breaks (for which detection is somewhat uncertain) that are typically encountered, and these authors go on to demonstrate that for smaller magnitude breaks, the Bai approach results in confidence sets that suffer from coverage rates substantially below the nominal level, with the true break date being excluded from the confidence set much too frequently. EM suggest an alternative approach to deriving confidence sets that achieve asymptotic validity, based on inverting a sequence of tests of the null that the break occurs at a maintained date, with the resulting confidence set comprised of all maintained dates for which the corresponding test did not reject. By deriving a locally

☆ The authors are grateful to two anonymous referees for very useful comments on an earlier version of this paper.

* Correspondence to: School of Economics, University of Nottingham, Nottingham, NG7 2RD, UK. Tel.: +44 0 115 9515481; fax: +44 0 115 9514159.
E-mail address: dave.harvey@nottingham.ac.uk (D.I. Harvey).

http://dx.doi.org/10.1016/j.jeconom.2014.09.004
0304-4076/

best invariant test that is invariant to the magnitude of the break under the null, the EM confidence sets have asymptotically correct coverage, regardless of the magnitude of the break (and therefore regardless of whether the magnitude is treated as fixed or asymptotically vanishing).

The EM model and assumptions pertain to a break in a linear time series regression, of which a break in level is a special case. They do not, however, consider the case of a break in linear trend, hence our first contribution is to develop an EM-type methodology for calculating asymptotically valid confidence sets for the date of a break in trend (and/or level). As in their approach, we derive a locally best invariant test of the null that the break occurs at a maintained date, and make an expedient choice for the probability measure used in deriving the test so as to render the resulting test statistic asymptotically invariant to the break timing.

When attempting to specify the deterministic component of an economic time series in practice, a critical consideration is the order of integration of the stochastic element of the process. Given the prevalence of integrated data, it is important to develop methods that are valid in the presence of I(1) shocks. Moreover, since there is very often a large degree of uncertainty regarding the order of integration in any given series, it is extremely useful to have available techniques that are robust to the order of integration, dealing with the potential for either stationary or unit root behaviour at the same time as specifying the deterministic component. A body of work has developed in recent years focusing on such concerns, developing order of integration-robust tests for a linear trend (e.g. Vogelsang, 1998, Bunzel and Vogelsang, 2005, Harvey et al., 2007, Perron and Yabu, 2009a), tests for a break in trend (e.g. Harvey et al., 2009, Perron and Yabu, 2009b, Saygınsoy and Vogelsang, 2011), and tests for multiple breaks in level (e.g. Harvey et al., 2010), *inter alia*. Most recently, Harvey and Leybourne (2014) have proposed methods for estimating the date of a break in level and trend that performs well for both I(0) and I(1) shocks.

In the current context, it is clear that reliable specification of confidence sets for the date of a break in level/trend will be dependent on the order of integration of the data under consideration. Perron and Zhu (2005) extend the results of Bai (1994, 1997) to allow for I(1), as well as I(0), processes when estimating the timing of a break in trend or level and trend, and different distributional results are obtained under I(0) and I(1) assumptions. Similarly, and as would be expected, we show that the EM procedure for calculating confidence sets, which is appropriate for I(0) shocks, does not result in sets with asymptotically correct coverage when the driving shocks are actually I(1). However, extension to the I(1) case is possible via a modified approach applied to the first differences of the data, whereby the level break and trend break are transformed into an outlier and a level break, respectively. This development comprises the second main contribution of our paper. Since there is typically uncertainty surrounding the integration order in practice, we propose a unit root pre-test-based procedure for calculating confidence sets that are asymptotically valid regardless of the order of integration of the data. We find the new procedure allows construction of confidence sets with correct asymptotic coverage under both I(0) and I(1) shocks (irrespective of the magnitude of the break). We also examine the performance of our procedure under local-to-I(1) shocks, and find that it displays asymptotic over-coverage (i.e. coverage rates above the nominal level), hence the confidence sets are asymptotically conservative in such situations, including the true date in the confidence set at least as frequently as the nominal rate would suggest. Monte Carlo simulations demonstrate that our recommended procedure performs well in finite samples, in terms of both coverage and length (the number of dates included in the confidence set as a proportion of the sample size).

The paper is structured as follows. Section 2 sets out the level/trend break model. Section 3 derives the locally best invariant tests for a break at a maintained date in both the stationary and unit root environments. The large sample properties under the null of correct break placement are established when correct and incorrect orders of integration are assumed, with the implications discussed for the corresponding confidence sets based on these tests. The properties of feasible variants of these tests, and corresponding confidence sets, are subsequently investigated. In Section 4 we propose use of a unit root pre-test to select between I(0) and I(1) confidence sets when the order of integration is not known. The finite sample behaviour of the various procedures is examined in Section 5. Here we also consider trimming as a means of potentially shortening the confidence sets. Section 6 provides empirical illustrations of our proposed procedure using US macroeconomic data, while Section 7 concludes.

The following notation is also used: '$\lfloor\cdot\rfloor$' denotes the integer part, '$\Rightarrow$' denotes weak convergence, and '$1(\cdot)$' denotes the indicator function.

2. The model and confidence sets

We consider the following model which allows for a level and/or a trend break in either a stationary or unit root process. The DGP for an observed series y_t we assume is given by

$$y_t = \beta_1 + \beta_2 t + \delta_1 1(t > \lfloor\tau_0 T\rfloor) + \delta_2 (t - \lfloor\tau_0 T\rfloor) 1(t > \lfloor\tau_0 T\rfloor) + \varepsilon_t, \quad t = 1, \ldots, T \tag{1}$$

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t, \quad t = 2, \ldots, T, \qquad \varepsilon_1 = u_1 \tag{2}$$

with $\lfloor\tau_0 T\rfloor \in \{2, \ldots, T-2\} \equiv \Lambda_T$ the level and/or trend break point with associated break fraction τ_0. In (1), a level break occurs at time $\lfloor\tau_0 T\rfloor$ when $\delta_1 \neq 0$; likewise, a trend break occurs if $\delta_2 \neq 0$. The parameters β_1, β_2, δ_1 and δ_2 are unknown, as is the break point $\lfloor\tau_0 T\rfloor$, inference on which is the central focus of our analysis. Our generic specification for ε_t is given by (2) assuming that $-1 < \rho \le 1$ and that u_t is I(0).

For an assumed break point $\lfloor\tau T\rfloor \in \Lambda_T$, our interest centres on testing whether or not $\lfloor\tau_0 T\rfloor$ and $\lfloor\tau T\rfloor$ coincide, which we can write in hypothesis testing terms as a test of the null hypothesis $H_0 : \lfloor\tau_0 T\rfloor = \lfloor\tau T\rfloor$ against the alternative $H_1 : \lfloor\tau_0 T\rfloor \neq \lfloor\tau T\rfloor$. Then, following EM, a $(1-\alpha)$-level confidence set for τ_0 is constructed by inverting a sequence of α-level tests of $H_0 : \lfloor\tau_0 T\rfloor = \lfloor\tau T\rfloor$ for $\lfloor\tau T\rfloor \in \Lambda_T$, with the resulting confidence set comprised of all $\lfloor\tau T\rfloor$ for which H_0 is not rejected. Provided the test of $H_0 : \lfloor\tau_0 T\rfloor = \lfloor\tau T\rfloor$ has size α for all $\lfloor\tau T\rfloor$, the confidence set will have correct coverage, since the probability of excluding τ_0 from the confidence set (via a spurious rejection of H_0) is α. In terms of confidence set length, a shorter than $(1-\alpha)$-level confidence set arises whenever the tests of $H_0 : \lfloor\tau_0 T\rfloor = \lfloor\tau T\rfloor$ reject with probability greater than α under the alternative $H_1 : \lfloor\tau_0 T\rfloor \neq \lfloor\tau T\rfloor$ across $\lfloor\tau T\rfloor$. Other things equal, the more powerful a test is in distinguishing between H_0 and H_1, the shorter this confidence set should be. Note that this approach to constructing confidence sets does not guarantee that the set is comprised of contiguous sample dates, cf. EM (p. 1207).

In the next section, we consider construction of powerful tests of H_0 against H_1, deriving locally best invariant tests along the lines of EM when $\rho = 0$ and when $\rho = 1$, under a Gaussianity assumption for u_t. The large sample properties of these tests are subsequently established under weaker conditions for ρ and u_t.

3. Locally best invariant tests

For the purposes of constructing locally best invariant tests, we make the standard assumption that $u_t \sim NIID(0, \sigma_u^2)$, and we