

KRANNERT GRADUATE SCHOOL OF MANAGEMENT

Purdue University
West Lafayette, Indiana

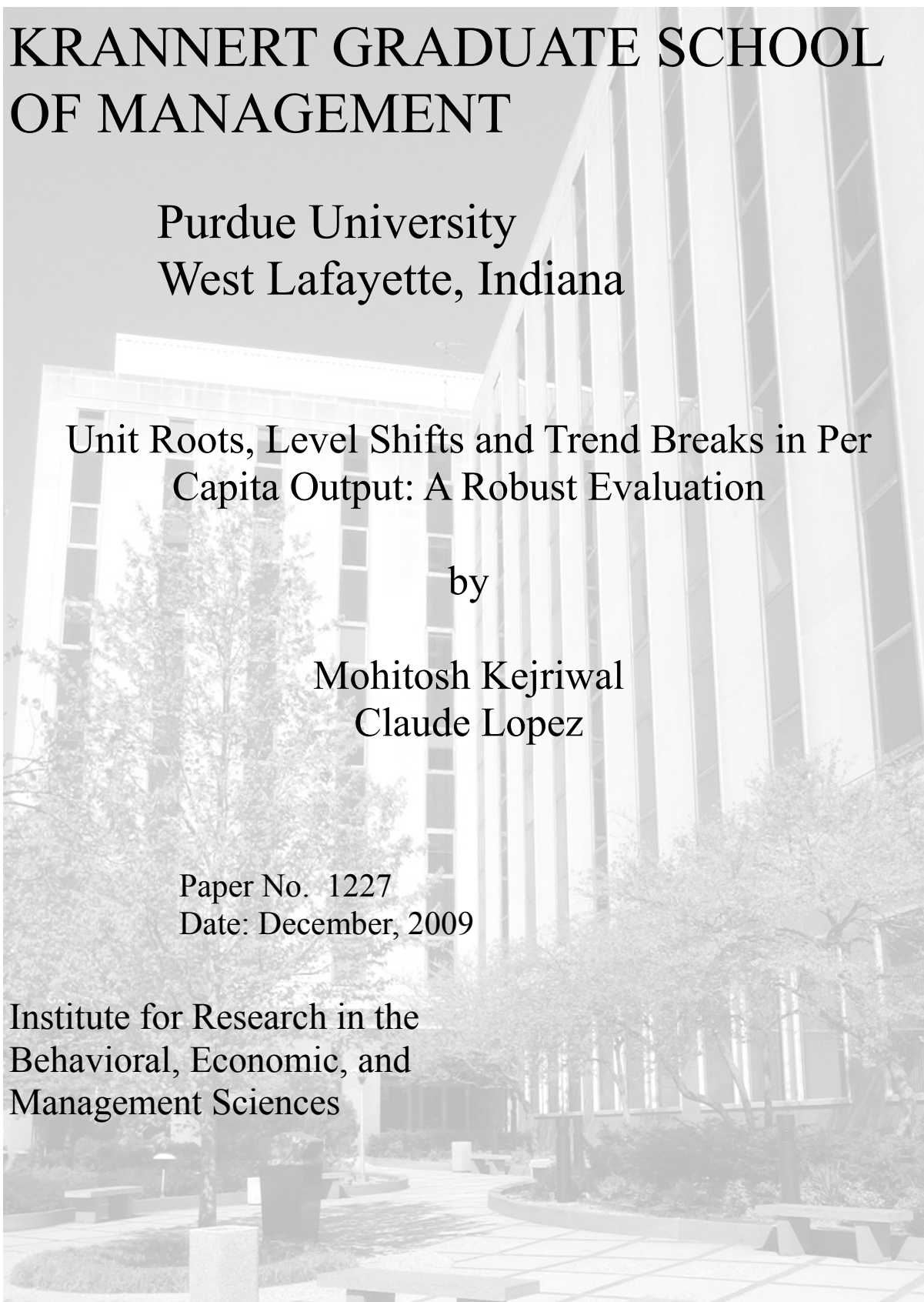
Unit Roots, Level Shifts and Trend Breaks in Per
Capita Output: A Robust Evaluation

by

Mohitosh Kejriwal
Claude Lopez

Paper No. 1227
Date: December, 2009

Institute for Research in the
Behavioral, Economic, and
Management Sciences



Unit Roots, Level Shifts and Trend Breaks in Per Capita Output: A Robust Evaluation*

Mohitosh Kejriwal[†]

Claude Lopez[‡]

Purdue University

University of Cincinnati

December 4, 2009

Abstract

Determining whether per capita output can be characterized by a stochastic trend is complicated by the fact that infrequent breaks in trend can bias standard unit root tests towards non-rejection of the unit root hypothesis. The bulk of the existing literature has focused on the application of unit root tests allowing for structural breaks in the trend function under the trend stationary alternative but not under the unit root null. These tests, however, provide little information regarding the existence and number of trend breaks. Moreover, these tests suffer from serious power and size distortions due to the asymmetric treatment of breaks under the null and alternative hypotheses. This paper estimates the number of breaks in trend employing procedures that are robust to the unit root/stationarity properties of the data. Our analysis of the per-capita GDP for OECD countries thereby permits a robust classification of countries according to the “growth shift”, “level shift” and “linear trend” hypotheses. In contrast to the extant literature, unit root tests conditional on the presence or absence of breaks do not provide evidence against the unit root hypothesis.

Keywords: growth shift, level shift, structural change, trend breaks, unit root

JEL Classification: C22

*We thank Arabinda Basistha, James Morley, David Papell, Pierre Perron, participants at the Econometrics seminar at University of Houston and the 2009 Midwest Econometric Group for useful comments, and Josep Lluís Carrion-i-Silvestre for sharing his Gauss code. We are solely responsible for any errors.

[†]Krannert School of Management, Purdue University, 403 West State Street, West Lafayette IN 47907 (mkejriwa@purdue.edu).

[‡]Department of Economics, University of Cincinnati, Cincinnati OH 45221-0371 (claude.lopez@uc.edu).

1 Introduction

Following the seminal work of Perron (1989), it is now well known that failure to account for structural changes in the trend can bias unit root tests in favor of the unit root model when the true process is subject to structural changes but is otherwise (trend) stationary within regimes specified by the break dates. Accordingly, it is now standard econometric practice to test for the presence of unit roots while allowing for structural changes in the trend function of the underlying time series. These testing procedures are typically based on the minimum t -statistic corresponding to the unit root parameter over the set of permissible break dates or alternatively computing this t -statistic at the break date that minimizes (or maximizes) the t -statistic associated with the break parameter (or maximizes its absolute value).

Recent developments in the econometrics literature highlight major drawbacks of commonly used unit root tests based on search procedures. When the break dates are unknown, it is useful to have information regarding the presence or absence of a change in order to investigate the potential presence a unit root. Indeed, unit root tests routinely employed in empirical analyses such as Zivot and Andrews (1992), Banerjee et al. (1992), Perron (1997) and Vogelsang and Perron (1998) are not invariant to the magnitude of trend breaks if the latter are present. Nunes et al. (1997), Lee and Strazicich (2001, 2003) and Kim and Perron (2009), among others, demonstrate that such tests suffer from serious power and size distortions due to the asymmetric treatment of breaks under the null and alternative hypotheses. For instance, the test of Zivot and Andrews (1992) assumes that if a break occurs, it only does so under the alternative hypothesis of trend stationarity. As a result, the test may reject the unit root null when the noise component is integrated but the trend is changing, leading to spurious evidence in favor of broken trend stationarity.

On the other hand, testing whether a time series can be characterized by a broken trend is complicated by the fact that the nature of persistence in the errors is usually unknown. Indeed, inference based on a structural change test on the level of the data depends on whether a unit root is present or not, given that asymptotic critical values are different in the two cases. Further, tests based on differenced data have very poor properties when the series contains a stationary component (Vogelsang, 1998). A circular testing problem therefore arises between tests on the parameters of the trend function and unit root tests.

To deal with this circular problem, various approaches have been suggested to test for the stability of the trend function that are robust to the nature of persistence in the noise component. Vogelsang (2001), building on prior work related to hypothesis testing on the

coefficients of a polynomial time trend reported in Vogelsang (1998), develops a Wald test statistic for structural change in the coefficients of a linear trend function with the same asymptotic critical values in both the stationary $I(0)$ and unit root $I(1)$ cases. More recently, Harvey et al. (2009a) [HLT henceforth] propose tests for a one-time break in the slope of the trend function based on a weighted average of the regression t -statistics appropriate for the case of $I(0)$ and $I(1)$ shocks. Perron and Yabu (2009a) [PY henceforth] suggest an alternative approach to assess the presence of changes in slope based on a Feasible Generalized Least Squares procedure that uses a super-efficient estimate of the sum of the autoregressive parameters α when $\alpha = 1$. Based on Monte Carlo experiments, HLT and PY show their respective procedures to be more powerful than that of Vogelsang (2001). Building on the work of Perron and Yabu (2009a), Kejriwal and Perron (2009) propose a sequential procedure that allows one to obtain a consistent estimate of the true number of breaks in the slope of the trend, irrespective of whether the errors are $I(1)$ or $I(0)$. Finally, Harvey et al. (2009b) propose robust tests for detecting multiple breaks in level conditional on a stable underlying slope.

Recent developments have also investigated issues related to the treatment of the breaks in the trend function when testing for the presence of a unit root. Harris et al. (2009) suggest the use of a GLS detrending procedure similar to that used by Elliott et al. (1996) and propose a unit root test that allows for a single change in the intercept and the slope of the trend function under both the null and alternative hypotheses. An alternative approach is advocated by Carrion et al. (2009), who propose extensions of the M class of tests analyzed in Ng and Perron (2001) and the feasible point optimal statistic of Elliott et al. (1996) that allow for multiple changes in the level and/or slope of the trend function. These tests have been shown to possess superior size and power properties relative to those that only allow for breaks under the alternative hypothesis.

A particularly important economic application where a broken trend model has received considerable attention, and consequently where the circular problem discussed above becomes relevant in practice, is the issue of determining whether output can be characterized by a stochastic trend.¹ Empirical evidence provided by commonly used unit root tests with trend breaks, however, varies considerably depending on the models used and the countries considered. Studies investigating US real GDP such as Perron (1989), Banerjee et al. (1990),

¹Kilian and Ohanian (2002) report an exhaustive list of the applications of unit root tests with breaks in the macroeconomic literature.

Balke and Fomby (1991), Christiano (1992), Zivot and Andrews (1992), and Papell and Prodan (2004), among others, show rather strong rejections of the unit root null hypothesis, regardless of how many breaks are considered, how they are selected and whether the data considered is aggregate or per capita. In contrast, studies focusing on international output offer a less clear conclusion. Allowing for one break in the trend while considering real GDP, Banerjee et al. (1992) and Bradley and Jansen (1995) can reject the unit root only for one and three out of the seven countries considered, respectively. Allowing for one break in both the level and the slope, Raj (1992) is able to reject the unit root null hypothesis for at least half of the per capita real GDP series considered. As an alternative, Ben-David et al. (2003) propose allowing for two breaks in the level and the slope and report rejections of the unit root for 12 out of the 16 countries for aggregate and per capita real GDP while Papell and Prodan (2009), allowing the first break to be in both the intercept and the slope while the second only in the slope, reject the unit root hypothesis for 14 out of the 18 OECD countries considered.

The aim of this paper is twofold. First, it proposes a formal econometric procedure that enables (i) robust detection of breaks in the level and/or the slope of the trend function, (ii) robust estimation of the number of breaks, (iii) reliable inference regarding the presence of a unit root conditional on the presence/absence of breaks, and (iv) reliable estimation of the break locations as well as the slope parameters in the regimes identified by the estimated break dates. Second, it applies this procedure to investigate the behavior of GDP per capita for nineteen OECD countries over the period 1870-2006. Our analysis permits a robust classification of countries according to the “growth shift” (shifts in the slope with possible shifts in the level), “level shift” (shifts in the level with no concurrent shifts in the slope) and “linear trend” (no shifts in the level or the slope) hypotheses. Moreover, in sharp contrast to the extant literature, results from unit root tests conditional on the presence or absence of breaks provide strong evidence in favor of the unit root hypothesis.

The rest of the paper is organized as follows. Section 2 proposes and describes the empirical methodology, including a discussion of the various limitations of the commonly employed procedures in the literature. Section 3 presents a set of Monte Carlo experiments to illustrate the merits of our procedure relative to those that have been routinely been applied in the literature. Section 4 reports the empirical results. Section 5 contains a discussion of our results and Section 6 offers some concluding remarks.

2 Methodology

The bulk of the current empirical literature investigating the persistence properties of macroeconomic time series has primarily focused on the application of unit root tests allowing for structural breaks in the trend function followed by the estimation of a level or first-differenced specification according to whether a unit root is present or not. These tests are generally obtained by minimizing the t -statistic on the unit root parameter over the set of permissible break dates or computing this t -statistic at the break date that minimizes (or maximizes) the t -statistic associated with the break parameter (or maximizes its absolute value). In order to provide the motivation for the econometric methodology advocated in this paper, it is useful to first discuss the potential drawbacks associated with the testing procedures that have typically been employed by existing studies.

First, the tests provide little information regarding the existence or number of trend breaks. At an intuitive level, it seems more natural to be first able to ascertain if breaks are at all present before proceeding to conduct unit root tests allowing for such breaks. In the absence of breaks, these tests suffer from low power due to the inclusion of extraneous break dummies thereby potentially leading the researcher to estimate a differenced specification when a level specification is in fact more appropriate. Indeed, as stressed by Campbell and Perron (1991), proper specification of the deterministic components is essential to obtaining unit root tests with reliable finite sample properties. Second, the unit root tests typically employed suffer from serious power and size distortions due to the asymmetric treatment of breaks under the null and alternative hypotheses. Moreover, if a break is present, this information is not exploited to improve the power of the testing procedure (a detailed discussion of this issue together with Monte Carlo evidence demonstrating the finite sample problems associated with this type of tests can be found in Kim and Perron, 2009, Nunes et al., 1997 and Lee and Strazicich, 2001 and 2003). Further, in most cases, the estimates of the break dates are obtained by minimizing/maximizing these unit root tests over all possible break dates which, in general, do not provide consistent estimates of the true break dates (Vogelsang and Perron, 1998).

Third, based on the prescription of unit root tests, the existing procedures often estimate a level specification and evaluate the joint significance of the intercept and slope dummies. However, a joint test is likely to conclude in favor of unstable growth rates even if the series has undergone a pure level shift, thereby making the interpretation of such tests quite difficult in practice (see section 3). Thus, if the objective is to distinguish between changes

in the level and the slope, it is essential to test for the stability of the slope parameter while allowing the intercept to vary across regimes and, conditional on the absence of slope shifts, test for level shifts.

Fourth, another common strategy is to start (before testing for a unit root) with a general level specification that incorporates both a changing slope as well as a changing intercept and then evaluate the significance of the individual t -statistics on the dummy variables. Depending on the outcome, the relevant model is estimated and used as the alternative model when testing for a unit root. There are two problems with such an approach. First, the limit distributions of the slope coefficient dummy estimates are different depending on whether a unit root is present so that prior information regarding the existence of a unit root is essential to validate significance based on t -statistics. Second, in the presence of a slope shift, the level shift parameters are not identified regardless of whether the noise component is stationary or not (Perron and Zhu, 2005).

Our econometric methodology is aimed at addressing each of the limitations discussed above and the proposed algorithm is summarized in Figure 1. The most general model considered can be described as:

$$y_t = \mu_0 + \beta_0 t + \sum_{i=1}^K \mu_i DU_{it} + \sum_{i=1}^K \beta_i DT_{it} + u_t, \quad t = 1, \dots, T \quad (1)$$

$$u_t = \alpha u_{t-1} + v_t, \quad t = 2, \dots, T, \quad u_1 = v_1 \quad (2)$$

where $DU_{it} = I(t > T_i)$, $DT_{it} = (t - T_i)I(t > T_i)$, $i = 1, \dots, K$. A break in the trend occurs at time $T_i = [T\lambda_i]$ when $\beta_i \neq 0$. The date of the breaks, T_i , and the number of breaks, K , are treated as unknown. The error u_t is allowed to be either $I(0)$ ($|\alpha| < 1$) or $I(1)$ ($\alpha = 1$). The stochastic process $\{v_t\}$ is assumed to be stationary (but not necessarily i.i.d. thereby permitting a general error structure for u_t). We are interested in the null hypothesis $H_0: \beta_i = 0$ against the alternative hypothesis $H_1: \beta_i \neq 0$.²

The first step tests for one structural break (that is $K = 1$ in (1)) in the slope of the trend function using procedures that are robust to the stationarity/non-stationarity properties of the data (HLT and PY). The tests employed are designed to detect a break in slope while allowing the intercept to shift. A rejection by these robust tests can therefore be interpreted

²Strictly speaking, the null hypothesis must be re-stated as $H_0: \mu_i = \beta_i = 0$ to obtain pivotal limiting distributions for the test statistics (see section 4.2 in HLT). This, however, does not mean that the tests are incorrectly sized in the presence of pure level shifts (see the simulation experiments in section 3).

as a change in the growth rate regardless of whether the level has changed.³ Given evidence in favor of a break by either of the single break tests, we then proceed to test for one versus two slope breaks (that is $K = 2$ in (1)) using the extension of PY proposed by Kejriwal and Perron (2009). Again, this latter test allows us to distinguish between one and two breaks while being agnostic to whether a unit root is present. Given the number of sample observations in our empirical analysis (137), we allow for a maximum of two breaks in our empirical analysis.⁴ While this may appear restrictive, allowing for a large number of breaks is not an appropriate strategy if one wants to determine if a unit root is present. The reason is that a unit root process can be viewed as a limiting case of a stationary process with multiple breaks, one that has a break (permanent shock) every period. Further, as discussed in Kejriwal and Perron (2009), the maximum number of breaks should be decided with regard to the available sample size. Otherwise, sequential procedures for detecting trend breaks will be based on successively smaller data subsamples (as more breaks are allowed) thereby leading to low power and/or size distortions. It is therefore important to allow for a sufficient number of observations in each segment and choose the maximum number of permissible breaks accordingly.⁵

A caveat associated with such a sequential procedure, as pointed out by Bai and Perron (2006) and Prodan (2008), is that single break tests may suffer from low power in finite samples in the presence of multiple breaks, especially if they are of opposite sign. To guard against such a possibility, we report the results of the one versus two breaks test regardless of whether a rejection is obtained from the single break tests.

Conditional on the presence of a stable slope at the initial step (that is $\beta_i = 0$ in (1) for $i = 1, \dots, K$), the focus becomes potential changes in the level of the trend and the hypotheses tested are $H_0: \mu_i = 0$ against the alternative hypothesis $H_1: \mu_i \neq 0$. Harvey et al. (2009b) propose a test for detecting multiple level breaks that is robust to the unit root/stationarity properties of the data.⁶ A rejection by this robust test can therefore be interpreted as changes in the level of the series. These authors also develop a sequential

³A potential strategy in this case to dissociate a level from a slope shift could be to use a t -statistic to test for the significance of the level shift parameter. Such a strategy is, however, flawed since, as shown in Perron and Zhu (2005), the level shift parameter is not identified in this case.

⁴This assumption is common to the majority of existing empirical studies.

⁵If a unit root is indeed present, the estimates of the break dates (obtained from the first-differenced specification) from an underspecified model are consistent for those break dates inserting which allow the greatest reduction in the sum of squared residuals and therefore correspond to the most dominant breaks in this sense (see Chong, 1995, Bai and Perron, 1998).

⁶The level breaks are modeled as local to zero in the $I(0)$ case and as increasing functions of sample size in the $I(1)$ case.

procedure which allows reliable estimation of the number of breaks. It is important to note here that the power issue associated with the sequential procedure for detecting slope shifts is not relevant in this case since the alternative hypothesis for the test in the latter case is that of at least one break in the level so that as the number of breaks increases, the test has an increasing number of opportunities to detect a level break (see the discussion in section 4.3 of Harvey et al., 2009b).

Given evidence in favor of instability in the slope (that is $\beta_i \neq 0$ in (1) for $i = 1, \dots, K$), we apply a new class of unit root tests which allows for breaks in the level and the slope under both the null and alternative hypotheses (Harris et al., 2009 and Carrion et al., 2009).⁷ Such a symmetric treatment of breaks alleviates these unit root tests from size and power problems that plague tests based on search procedures. Similarly, in the presence of at least one level shift, we apply unit root tests which allow for breaks in the level under both the null and alternative hypotheses (Carrion et al., 2009). If no evidence is found of instability either in the level or in the slope, we apply standard (no break) unit root tests developed by Elliott et al. (1996) and Ng and Perron (2001).

With breaks in the level and/or the slope, the trend coefficients are estimated from a first-differenced or level specification according to whether a unit root is present or not. Perron and Zhu (2005) show, in the presence of a break in the slope, that the estimates of the break dates as well as the parameters governing the slope of the trend function obtained from the level specification are consistent even in the presence of a unit root. However, more efficient estimates of the break dates (in terms of a faster rate of convergence to the true value) can be obtained from estimating a specification in first differences in this case (see section 3). The unit root tests thereby enable more precise estimation of the break dates. In models with pure level shifts, consistent estimates of the break dates are obtained using the procedure suggested by Harvey et al. (2009b) in the unit root case and by minimizing the sum of squared residuals from the level specification in the stationary case.

To obtain the trend parameter estimates in the stable linear trend case (that is $\beta_i = \mu_i = 0$ in (1) for $i = 1, \dots, K$), we apply the robust procedures proposed by Harvey et al. (2007) and Perron and Yabu (2009b). These procedures are simply the “no break” counterparts of the HLT and PY procedures respectively and are therefore not discussed in detail in the paper.

To ensure brevity of the main text as well as to enhance readability, we have relegated

⁷Note that Perron (1989, 1990) devised unit root testing procedures that are invariant to the magnitude of the shift in level and/or slope but his analysis was restricted to the known break date case.

the discussion of the various testing procedures including the notation for the different tests and estimates to the Appendix.

3 Monte Carlo Experiments

This section explores two aspects of the proposed procedure vis-a-vis currently existing procedures by means of Monte Carlo experiments: (1) the appropriateness of testing for shifts in the slope while allowing for shifts in the level as opposed to joint tests for shifts in both the level and the slope, when the objective is to detect shifts in the slope only and (2) the relative efficiency gains (in terms of mean squared error) obtained by estimating the break dates from a level specification when the noise component is stationary and a first-differenced specification when a unit root is present.

To investigate these issues, we consider the following data generating process with a single break:⁸

$$y_t = \mu_0^0 + \beta_0^0 t + \mu_1^0 I(t > T_1^0) + \beta_1^0 (t - T_1^0) I(t > T_1^0) + u_t \quad (3)$$

where the errors $\{u_t\}_{t=1}^T$ are generated as

$$u_t = \alpha^0 u_{t-1} + v_t, \quad t = 2, \dots, T, \quad u_1 = v_1 \quad (4)$$

where $\{v_t\}_{t=1}^T$ is a sequence of i.i.d. $N(0, \sigma^2)$ random variables. We use the superscript “0” to indicate the true value of a parameter. Regarding the choice of base parameter values, we closely follow the design employed by Perron and Zhu (2005). We thus set $\mu_0^0 = 1.72$, $\beta_0^0 = 0.03$, $\sigma = 0.01$ when $\alpha^0 = 1$ and $\sigma^2 = 0.1$ otherwise. We consider five values for the autoregressive parameter: $\alpha^0 = 0.5, 0.6, 0.7, 0.8, 0.9, 1$. Further, we set $\mu_1^0 = -0.04, -0.02, 0, 0.02, 0.04$ and the break fraction $\lambda_1^0 = 0.3, 0.5, 0.7$. Results are reported for two sample sizes: $T = 150, 200$.

3.1 Joint Tests versus Tests for Slope Shifts

Empirical research studying the stability of output growth has almost exclusively focused on joint tests for the presence of shifts in the level and the slope. However, as discussed in section 2, joint tests have power against processes which are characterized by shifts in the level only and are therefore likely to reject the null of stability even when there is no

⁸The experiments were also performed on a data generating process with two breaks. Results were qualitatively very similar and hence not reported. They are available upon request.

change in the slope of the trend function. In other words, a rejection by these tests does not provide useful information regarding whether a change in the slope is at all present. The suggested procedure has the correct asymptotic size regardless of the presence of level shifts. To illustrate the advantage of using the latter procedure, Table 1 and 2 report empirical rejection frequencies of the one break tests of HLT and PY when the data generating process (3) and (4) is characterized by a stable slope ($\beta_1^0 = 0$) but a shift in the level ($\mu_1^0 \neq 0$). For comparison, we also present the rejection rates for the joint test on both the intercept and the slope proposed in PY (denoted $ExpW_J$). Note that all three tests are robust to the nature of persistence in the noise component.

When the errors are $I(0)$, the joint test is subject to mild size distortions regardless of the magnitude of the level shift. This is also true when the errors are $I(1)$ but the level shift component is small ($|\mu_1^0| \leq 0.02$). However, when the level shift is large, the joint test substantially over-rejects the null of stability. Importantly, these distortions are not mitigated as the sample size increases. In contrast, the tests which are designed to purely detect a break in the slope are much better sized for all values of the level shift although some size distortions are apparent in the unit root case, especially for the Harvey et al. (2009b) test.

3.2 Efficiency in Break Date Estimation

Perron and Zhu (2005) show that the estimate of the break date obtained from a level specification is consistent irrespective of whether the noise component is stationary or has a unit root. The break fraction estimate based on the level specification converges to the true value at rate T when the errors are $I(0)$ and at rate $T^{1/2}$ when the errors are $I(1)$. An improved rate of convergence in the $I(1)$ case (rate T) can, however, be obtained by estimating the break date from the specification in first differences. This follows from the results in Bai (1994), Bai (1997) and Bai and Perron (1998) who show that a shift in the mean of an $I(0)$ process can be estimated with a T rate of convergence. This improvement is likely to provide finite sample efficiency gains from estimating a specification in first differences compared to one in levels. In other words, information about the presence of a unit root can be exploited to facilitate more accurate estimation of the break dates.

In order to provide a quantitative assessment of the potential efficiency gains, we consider the single break DGP given by (3) and (4) with $\beta_1^0 = -0.02, 0.02$. Tables 3 and 4 reports the ratio of mean squared errors (MSE_D/MSE_L), where MSE_D and MSE_L are the mean squared errors of the estimated break dates from the specification in first differences and levels

respectively. The results confirm that knowledge regarding the presence/absence of a unit root can be used to obtain improved break date estimates. When the errors are $I(0)$, using a level model results in much lower mean squared errors while the first differenced model dominates when a unit root is present.

4 Empirical Results

This section presents an empirical analysis of the stability of the trend function as well as the notion of trend reverting behavior of long-run per capita output employing the proposed procedure. Specifically, it focuses on the commonly used Maddison dataset, considering the per capita GDP for nineteen OECD countries during the 1870-2006 period.⁹ An informal inspection of the plot in Figure 2 suggests the possibility of at least one level and/or slope shift in the trend function for most of the per-capita output series. For the sake of brevity, we refer to the logarithm of per capita GDP as output for the rest of the paper.

Following Papell and Prodan (2009)'s classification, we label a shifting slope with possible shifts in the level the "growth shift" hypothesis, shifts in the level with no concurrent shifts in the slope the "level shift" hypothesis and the absence of shifts in either the level or the slope the "linear trend hypothesis". Note that our statistical interpretation of the "growth shift" hypothesis differs from that of Papell and Prodan's in that we allow for the possibility of level shifts (rather than test for them) when testing for shifts in the slope. Indeed, as explained in the previous section, joint tests of significance on the slope and the level shift parameters are likely to generate misleading results regarding the classification of countries according to these three hypotheses. The proposed methodology allows us to reliably distinguish between the three hypotheses in addition to providing evidence regarding the potential presence of a stochastic trend in output.

The initial step of the analysis tests for the presence and the number of breaks in the trend function. The results are reported in Table 5. Evidence clearly favoring the growth shift hypothesis is obtained for thirteen countries. In particular, Finland, Norway, Portugal, Spain, Sweden and the UK show evidence of one break in slope while Austria, Belgium, Germany, Italy, Japan, Netherlands and New Zealand report two slope breaks. The level shift hypothesis is supported for USA and Denmark, with evidence of one and two level breaks, respectively; while the results for Australia and Canada support the linear trend hypothesis. Note that the results for France and Switzerland are not unambiguously in favor

⁹The dataset is obtained from Maddison (2009).

of a particular hypothesis given that single break tests cannot reject the null of stability while the test of one versus two breaks provides strong evidence of instability. Accordingly, for these two countries, we present unit root test results as well as trend parameter estimation results for both the stable trend case and the two slope breaks case.

Having categorized countries according to the above hypotheses, we next use this information to test for the presence of a unit root in output. In addition to providing important evidence about whether output can be characterized by a stochastic trend, the unit root tests will also allow us to choose the appropriate specification for estimating the model parameters. As shown in Table 6, none of the countries studied show any evidence of trend stationarity or regime-wise trend stationarity (None of the tests are significant at even the 10% level). This result is a clear departure from the commonly accepted notion that allowing for breaks strengthens the rejection of the unit root null for GDP data. This issue is further discussed in the next section.

Turning to the estimation results, we report estimates of the slope parameters and the associated 95% confidence intervals together with the estimates of the break dates. Estimates for the level parameters are not presented as these parameters are not identified in the presence of a unit root component (see Hatanaka and Yamada, 1999 and Perron and Zhu, 2005). Table 7a reports the parameter estimates for countries exhibiting growth shifts while Table 7b focuses on countries with pure level shifts and linear trends. In Table 7b, we denote the “no break” counterparts to the HLT and PY procedures as HLT^0 and PY^0 respectively. Both tables confirm that major historical/economic events have had a clear impact on these economies given that the break dates selected correspond to the two World Wars (WWI and WWII) for most of the European countries, WWII for USA and Japan, the first oil price shock for Japan and a change from colonial to independent dominion status as well as after-Depression effects for New Zealand. Interestingly, our results do not show WWI and WWII to have had any significant impact on growth rates in Portugal and Spain. For Portugal, the break date corresponds to a change in the political regime with an emphasis towards financial stability and therefore increased economic growth. Finally, the break for Spain is associated with the onset of the Spanish Civil War.

Table 7a also highlights several interesting patterns across countries. All countries experiencing a single change in the growth rate are subject to similar growth patterns: whether the break is around the first World War (Finland, Norway, Sweden and the UK) or the second (Portugal and Spain), the growth rate in the post-break period exceeds that in the pre-break period. This increase is the largest for the latter break with an average growth

ratio of 3.63 compared to 2.06 for countries with the earlier break.¹⁰ The two-break countries confirm such an outcome. The European countries corroborate the impact of the two World Wars and, with the exception of Belgium, they all report their most productive phase after the second break, that is after WWII. Japan also records its highest growth rate of 7.64% after WWII, followed by a steep decline in growth to a rate of 2.76% engineered by the first oil shock in 1973.

Growth ratios across regimes also provide some useful insights and a less homogenous description of the changes. Countries such as Belgium and Japan report a strong growth enhancement between the first and the second breaks with ratios of 4.91 and 4.42, respectively. In contrast, Austria, France and New Zealand report a slowdown, with ratios less than one, while Italy and the Netherlands observe a “meltdown” with a negative growth rate in the second segment. The latter five countries, however, all experience their largest growth improvement in the segment following the most recent break.

It is worth noting that the confidence intervals, however, render some different conclusions regarding the shorter term dynamics of output for several countries. Indeed, based on their 95% confidence intervals, the slope estimates of Austria, France, Germany, and the Netherlands are not significantly different from zero for the first two periods. The re-estimation of the models with only one break in the trend confirms the importance of WWII, that is, the initial second break.¹¹ Note that in all cases, the slope coefficient estimate for the last segment remains unchanged while the estimate in the first segment is still not significantly different from zero.

Overall, the short term variations in the output growth are country specific, yet the long term behavior is quite similar across the OECD countries observing growth shifts: they all report a strengthening of their growth when comparing the first and last segments, with an average ratio of 3.11. Finally, none of the countries whether they support the growth shifts, the level shifts or the linear trend hypothesis provide evidence of trend reversion in output.

5 Discussion

Our empirical results are generally not supportive of the neoclassical view that growth rates remain stable in the long run.¹² Rather, they are representative of the idea that the growth

¹⁰Ben-David and Papell (1995) reach a similar conclusion, although their ratios are smaller.

¹¹Unit root tests allowing for a single break were computed for these countries and again did not provide any evidence against the unit root null.

¹²According to the neoclassical growth model, changes in policy variables generate only temporary changes in the growth rate.

process is not continuous. Following Kuznets (1963) perspective on the need to provide a distinct demarcation between different periods of growth, our empirical analysis is directed towards identifying the time periods at which such discontinuities occur which allows us to delineate and study distinct growth regimes. Rosenstein-Rodan’s (1943) theory of the “big push” as well as Rostow’s (1961) theory of “takeoffs” provide examples of growth discontinuities. The results are also broadly consistent with the implication of endogenous growth models such as Romer (1986) that growth rates tend to increase over time. Relatively high postwar growth could be the result of a sustained movement towards the liberalization of trade and the creation of institutions such as Bretton Woods and GATT whose objective was to promote the flow of goods across international boundaries. An alternative explanation is advocated by Olson (1982), who suggests that major social upheavals can cause the elimination of old distributional coalitions resulting in a more efficient reallocation of resources and therefore increased economic growth. Our break dates corroborate commonly accepted conclusions among empirical studies on output over similar periods: major historical/economic events, such as World Wars and the first oil shock had an important effect on output growth rates of OECD countries.

The application of the proposed methodology leads to a clear departure from standard unit root test results when allowing for breaks in the trend function: none of the output series studied report evidence of trend stationarity or regime-wise trend stationarity.¹³ As shown in Tables 8a and 8b, several authors have studied the behavior of per capita GDP for several OECD countries over a similar sample period. Raj (1992), using both Perron’s (1989) and Zivot and Andrews’ (1992) tests, reports evidence of regime-wise trend stationarity for five out of nine countries when allowing for a break in the intercept and in the trend. Using the latter test and two different break date selection methods, Zelhorst and De Haan (1995) are able to reject the unit root null hypothesis for nine out of twelve countries. Using the same test, Ben-David and Papell (1995) investigate the behavior of both aggregate and per capita real GDP for 16 OECD countries. They consider a model that allows both the trend and the intercept to change, unless one of the shift dummies is not significant, in which case it is dropped and the model is re-estimated. Rejection of the unit root null hypothesis is obtained for twelve per capita real GDP series. Considering the same countries, Ben-David et al. (2003) employ an extension of Zivot and Andrews (1992) to two breaks and report

¹³In a related paper, Murray and Nelson (2000) challenge evidence favoring trend stationarity by arguing that false rejections of the unit root hypothesis can be triggered by size distortions associated with data-based lag selection and departures from the maintained hypothesis of temporal homogeneity.

results that depend on the models considered. If the model includes two breaks in both the intercept and the trend, the unit root null is rejected for twelve countries. Using a restricted version of the same test, Papell and Prodan (2004) show that the US reports evidence of trend stationarity (i.e., no change in the slope). Finally, Papell et Prodan (2009) show that fifteen out of the eighteen OECD countries considered report evidence of regime-wise trend stationarity. All these studies agree on the direct relation between evidence of long run output convergence and the inclusion of breaks in the model. Furthermore, Ben-David and Papell (1995) demonstrate how essential is the ability to reject the unit root null hypothesis when trying to link findings on output to the notion of stable growth and a steady state path.¹⁴

A careful analysis of these studies allows us to emphasize their technical similarities, to explain their corroborating conclusions as well as clarify ways in which ours differ. Clearly, these studies share three major technical concerns: (i) none test for the existence and the number of breaks but impose either one or two breaks under the (broken) trend stationary alternative, (ii) the unit root tests used do not allow for break(s) under the unit root null, and (iii) break date selection relies on maximizing the evidence against the unit root null. All three issues were discussed in section 3 and are addressed by our testing procedure. A comparison between our point estimates with the ones reported in Tables 8a and 8b illustrates the relevance of these issues when dealing with real data. Testing for the presence of breaks leads to differences in the model chosen. For example, in our analysis, Australia does not experience any breaks in the level or the trend. Its average growth rate is 1.4% compared to regime-specific growth rates of 1.31%, 0.65% and 1.87% reported by Ben-David et al. (2003) and 0.07% and 0.42% reported by Papell and Prodan (2009). The results for other countries show the importance of obtaining accurate estimates for the break dates. For instance, our study reports two breaks for France in 1917 and 1945, yet the slope in the first two periods being not significantly different from zero, the model can be reduced to a single trend break model without any significant change in the point estimates or the break date estimate. The average growth rate of 0.42% and 2.96% in the two regimes can be compared to 0.6% and 2.2% for Raj (1992), 0.53% and 1.68% for Ben-David et al. (1995) with break date estimates of 1940 and 1939, respectively. Imposing two breaks, Ben-David et al. (2003) report an average growth rate of 1.29% prior to 1939, 3.49% between 1940 and

¹⁴Ben-David and Papell (2000) investigate the stability of the growth process under the assumption that output is (broken) trend stationary. They find that although there is some evidence of individual periods of slowdowns, the overall tendency appears to be one of increasing steady state growth over the long run.

1974, and 1.86% thereafter, thus leading to very different growth rate estimates relative to the one break case.

An alternative approach is taken by Balke and Fomby (1991), Bradley and Jansen (1995), and Darné and Diebolt (2004), who employ procedures designed for detecting outliers in order to identify and isolate permanent and temporary shocks. Tests for the presence of a unit root are then implemented on data corrected for outliers. The approach allows for the possibility of multiple breaks in trend occurring at unknown dates. The disadvantages of such procedures are that identification of the type of outlier can be quite sensitive to the original specification of the ARIMA model. Moreover, the presence of outliers, primarily level shifts, can cause the original ARIMA component to be misspecified. This can cause the procedure to incorrectly identify the types of outliers (see Balke and Fomby, 1991).

The econometric procedure advocated in this paper addresses several concerns commonly encountered in the empirical literature assessing the long-term behavior of GDP. It allows proper identification of the number of breaks, precise estimation of the break dates as well as an accurate assessment of the nature of the trend in per-capita real GDP. While our results confirm that major events have had an important impact on the level of GDP, they also provide strong evidence supporting the stochastic nature of its trend. From a macroeconomic perspective, the most important implication of the stochastic trend/unit root hypothesis is that random shocks have a permanent effect on the system. Contrary to the implications of business cycle theories, fluctuations are not merely transitory deviations around a stable deterministic trend but the secular component is itself subject to fluctuations. Furthermore, the shocks are frequent in that they occur every observation period with relatively small variance. In this context, trend breaks can be viewed as large, infrequent shocks or outliers. Level shifts in output correspond to temporary changes in the drift of the unit root process (the average growth rate) while slope shifts correspond to permanent changes in this drift.

6 Conclusion

This paper proposes and implements an econometric procedure that allows rigorous assessment of the stability of the trend function of a univariate time series as well as whether the series can be characterized by a stochastic trend. The break detection procedures used are robust to the persistence of the noise component and can therefore be applied when no a priori knowledge is available regarding whether the shocks are stationary or not. Contrary to the existing literature, it enables a clear dissociation between significant changes in the slope

(while allowing for potential shifts in the level) from those in the level (with no concurrent changes in the slope). Further, the unit root testing procedures employed are not subject to finite sample issues associated with empirical size and power that typically undermine the use of currently popular tests. The methodology advocated allows consistent estimation of the true number of changes, whether the changes occur in the slope or the level of the trend function, reliable inference regarding the presence/absence of a unit root as well as consistent and efficient estimation of the slope parameters and break dates.

The analysis of historical data for nineteen OECD countries over 1870-2006 provides evidence of an overall increase in the growth rate for fifteen of these countries while two report evidence of pure shifts in the level. The estimated break dates emphasize the importance of the two World Wars in determining the growth path of output as well as other world events such as the first oil shock or more local events (such as the change in status for New Zealand). Yet, the results for none of these countries indicate evidence of trend or regime-wise trend stationarity.

References

- Andrews, D.W.K. (1993), Tests for Parameter Instability and Structural Change with Unknown Change Point, *Econometrica* 61, 821-856.
- Andrews, D.W.K. and W. Ploberger (1994), Optimal Tests when a Nuisance Parameter is present only under the Alternative *Econometrica* 62, 1383-1414.
- Bai, J. (1994), Least Squares Estimation of a Shift in Linear Processes, *Journal of Time Series Analysis* 15, 453-472.
- Bai, J. (1997), Estimation of a Change Point in Multiple Regression Models, *The Review of Economics and Statistics* 79, 551-563.
- Bai, J. and P. Perron (1998), Estimating and Testing Linear Models with Multiple Structural Changes, *Econometrica* 66, 47-78.
- Balke, N.S. and T.B Fomby (1991), Shifting Trends, Segmented Trends, and Infrequent Permanent Shocks, *Journal of Monetary Economics* 28, 61-85.
- Banerjee, A., J. Dolado, and J.W. Galbraith (1990), Recursive Tests for Unit Roots and Structural Breaks in Long Annual GNP Series, Manuscript, University of Florida.
- Banerjee, A., R.L. Lumsdaine and J. H. Stock (1992), Recursive and Sequential Tests of the Unit Root and Trend-Break Hypotheses: Theory and International Evidence, *Journal of Business and Economic Statistics* 10, 271-287.
- Ben-David, D. and D.H. Papell (1995), The Great Wars, the Great Crash, and steady state growth: Some new Evidence about an Old Stylized Fact, *Journal of Monetary Economics* 36, 453-475.
- Ben-David, D. and D.H. Papell (2000), Some Evidence on the Continuity of the Growth Process among the G7 Countries, *Economic Inquiry* 38, 320-330.
- Ben-David, D., R.L. Lumsdaine, and D.H. Papell (2003), Unit Roots, Postwar Slowdowns and Long-run Growth: Evidence from two Structural Breaks, *Empirical Economics* 28, 303-319.
- Bradley, M.D. and D.W. Jansen (1995), Unit Roots and Infrequent Large Shocks: New International Evidence on Output Growth, *Journal of Money, Credit, and Banking* 27, 876-893.
- Carrion-i-Silvestre, J.L., D. Kim and P.Perron (2009), GLS-based Unit Root Tests with Multiple Structural Breaks both Under the Null and the Alternative Hypotheses, *Econometric Theory* forthcoming.
- Campbell, J.Y. and P. Perron (1991), Pitfalls and Opportunities: What Macroeconomists should know about Unit Roots, *NBER Macroeconomics Annual* 6, 141-201.
- Chong, T.T. (1995), Partial Parameter Consistency in a Misspecified Structural Change Model, *Economics Letters* 49, 351-357.

- Christiano, L.J. (1992), Searching for a Break in GNP, *Journal of Business and Economic Statistics* 10, 237–250.
- Darne, O. and C. Diebolt (2004), Unit Roots and Infrequent Large Shocks: New International Evidence on Output, *Journal of Monetary Economics* 51, 1449-1465.
- Dickey, D.A. and W.A. Fuller (1979), Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *Journal of the American Statistical Association* 74, 427-431.
- Elliott, G., Rothenberg, T., and J.H. Stock (1996), Efficient Tests for an Autoregressive Unit Root, *Econometrica* 64, 813-836.
- Harris, D., D.I. Harvey, S.J. Leybourne and A.M.R. Taylor (2009), Testing for a Unit Root in the Presence of a Possible Break in Trend, *Econometric Theory* forthcoming.
- Harvey, D. I., Leybourne, S. J. and A.M.R Taylor (2007), A Simple, Robust and Powerful Test of the Trend Hypothesis, *Journal of Econometrics* 141, 1302-1330
- Harvey, D.I., S.J. Leybourne, and A.M.R. Taylor (2009a), Simple, Robust and Powerful Tests of the Breaking Trend Hypothesis, *Econometric Theory* 25, 995-1029 .
- Harvey, D.I., S.J. Leybourne, and A.M.R. Taylor (2009b), Robust Methods for Detecting Multiple Level Breaks in Autocorrelated Time Series, Granger Centre Discussion Paper No. 09/01.
- Hatanaka, M. and K. Yamada (1999), A Unit Root Test in the Presence of Structural Changes in I(1) and I(0) models, in R.F. Engle and H. White (eds.), *Cointegration, Causality*
- Kejriwal, M. and P. Perron (2009), A Sequential Procedure to Determine the Number of Breaks in Trend with an Integrated or Stationary Noise Component, Krannert Working Paper No. 1217, Purdue University.
- Kilian, L. and L. Ohanian (2002), Unit Roots, Trend Breaks, and Transitory Dynamics: A Macroeconomic Perspective, *Macroeconomic Dynamics* 6, 614-631.
- Kim, D. and P. Perron (2009), Unit Root Tests Allowing for a Break in the Trend Function at an Unknown Time under Both the Null and Alternative Hypotheses, *Journal of Econometrics* 148, 1-13.
- Kuznets, S. (1963), Notes on the Take-off, in W.W. Rostow (ed.), *The Economics of Take-Off into Sustained Growth*, Macmillan.
- Lee, J., and M.C. Strazicich (2001), Break Point Estimation and Spurious Rejections with Endogenous Unit Roots, *Oxford Bulletin of Economics and Statistics* 63, 535-558.
- Lee, J., and M.C. Strazicich (2003), Minimum Lagrange Multiplier Unit Root Test with Two Structural Breaks, *The Review of Economics and Statistics* 85, 1082-1089.
- Lumsdaine, R.L. and D.H. Papell (1997), Multiple Trend Breaks and the Unit Root Hypothesis, *Review of Economics and Statistics* 79, 212–218.

Maddison, A. (2009), Statistics on World Population, GDP and Per Capita GDP, <http://www.ggdc.net/maddison/>.

Murray, C.J. and C.R. Nelson (2000), The Uncertain Trend in U.S. GDP, *Journal of Monetary Economics* 46, 79–95.

Ng, S. and P. Perron (2001), Lag Length Selection and the Construction of Unit Root Tests With Good Size and Power, *Econometrica* 69, 1519-1554.

Olson, M. (1982), *The Rise and Decline of Nations: Economic Growth, Stagflation and Social Rigidities*, Yale University Press, 1982.

Nunes, L.C., P. Newbold, and C.-M. Kuan (1997), Testing for Unit Roots with Breaks: Evidence on the Great Crash and the Unit Root Hypothesis Reconsidered, *Oxford Bulletin of Economics and Statistics* 59, 435–448.

Papell, D.H. and R. Prodan (2004), The Uncertain Unit Root in U.S. Real GDP: Evidence with Restricted and Unrestricted Structural Change, *Journal of Money, Credit and Banking* 36, 423-427.

Papell D.H. and R. Prodan (2009), Time Series Tests of Constant Steady-State Growth, Working Paper, University of Houston.

Perron, P. (1989), The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis, *Econometrica* 57, 1361–1401.

Perron, P. (1997), Further Evidence on Breaking Trend Functions in Macroeconomic Variables, *Journal of Econometrics* 80, 355–386.

Perron, P. and T. Yabu (2009a), Testing for Shifts in Trend with an Integrated or Stationary Noise Component, *Journal of Business and Economic Statistics* 27, 369-396.

Perron, P. and T. Yabu (2009b), Estimating Deterministic Trends with an Integrated or Stationary Noise Component, *Journal of Econometrics* 151, 56-69.

Perron, P. and X. Zhu (2005), Structural Breaks with Deterministic and Stochastic Trends, *Journal of Econometrics* 129, 65-119.

Prodan, R. (2008), Potential Pitfalls in Determining Multiple Structural Changes with an Application to Purchasing Power Parity, *Journal of Business and Economic Statistics* 26, 50-65.

Raj, B. (1992), International Evidence on Persistence in Output in the Presence of an Episodic Change, *Journal of Applied Econometrics* 7, 281–293.

Romer, P.M. (1986), Increasing Returns and Long-Run Growth, *Journal of Political Economy* 98, 71-102.

Rosenstein-Rodan, P.N. (1943), Problems of Industrialization of Eastern and South-Eastern Europe, *The Economic Journal* 53, 202-211.

Rostow, W.W. (1961), *The Stages of Economic Growth: A Non-Communist Manifesto*, Cambridge University Press.

Roy, A. and W.A. Fuller (2001), Estimation for Autoregressive Time Series With a Root Near 1, *Journal of Business and Economic Statistics* 19, 482-493.

Vogelsang, T.J. (1998), Trend Function Hypothesis Testing in the Presence of Serial Correlation, *Econometrica* 66, 123-148.

Vogelsang, T.J. (2001), Tests for a Shift in trend when Serial Correlation is of Unknown Form, Unpublished Manuscript, Department of Economics, Cornell University.

Vogelsang, T.J. and P. Perron (1998), Additional Tests for a Unit Root Allowing the Possibility of Breaks in the Trend Function, *International Economic Review* 39, 1073-1100.

Zelhorst, D. and J. De Haan (1995), Testing for a Break in Output: New International Evidence, *Oxford Economic Papers* 47, 357-362.

Zivot, E. and D.W.K. Andrews (1992), Further Evidence on the Great Crash, the Oil-price Shock, and the Unit-root Hypothesis, *Journal of Business and Economic Statistics* 10, 251-270.

Appendix: Description of Testing Procedures

A.1 Robust Tests for Breaks in Trend

A.1.1 The Harvey et al. (2009a) Test for a Break in Slope

Harvey et al. (2009a) propose test statistics that are constructed by taking a weighted average of the regression t -statistics from a regression in levels and a regression in differences. The weighting function is based on the KPSS stationarity statistics applied to the levels and differenced data. First differencing (1) [for $K = 1$] yields

$$\Delta y_t = \beta_0 + \mu_1 D_{1t} + \beta_1 DU_{1t} + \varepsilon_t, \quad t = 2, \dots, T \quad (\text{A.1})$$

where $\varepsilon_t = \Delta u_t$, $D_{1t} = I(t = T_1 + 1)$ and $DU_{1t} = I(t > T_1)$. Consider the t -statistics

$$t_0(\lambda_1) = \frac{\hat{\beta}_1(\lambda_1)}{\sqrt{\hat{\omega}_1^2(\lambda_1) [\{\sum_{t=1} x_{L1,t}(\lambda_1) x_{L1,t}(\lambda_1)'\}^{-1}]_{44}}} \quad (\text{A.2})$$

$$t_1(\lambda_1) = \frac{\tilde{\beta}_1(\lambda_1)}{\sqrt{\tilde{\omega}_1^2(\lambda_1) [\{\sum_{t=1} x_{D1,t}(\lambda_1) x_{D1,t}(\lambda_1)'\}^{-1}]_{33}}} \quad (\text{A.3})$$

In (A.2), $x_{L1,t}(\lambda_1) = \{1, t, DU_{1t}, DT_{1t}\}$, $DT_{1t} = (t - T_1)I(t > T_1)$, $\hat{\beta}_1(\lambda_1)$ is the OLS estimate of β_1 from (1) and $\hat{\omega}_1^2(\lambda_1)$ is an estimate of the long-run variance based on the OLS residuals $\hat{u}_t(\lambda_1) = y_t - \hat{\mu}_0(\lambda_1) - \hat{\beta}_0(\lambda_1)t - \hat{\mu}_1(\lambda_1)DU_{1t} - \hat{\beta}_1(\lambda_1)DT_{1t}$. In (A.3), $x_{D1,t}(\lambda_1) = \{1, D_{1t}, DU_{1t}\}$ and $\tilde{\beta}_1(\lambda_1)$ is the OLS estimate of β_1 from (A.1) and $\tilde{\omega}_1^2(\lambda_1)$ is an estimate of the long-run variance based on the residuals $\tilde{\varepsilon}_t(\lambda_1) = y_t - \tilde{\beta}_0(\lambda_1) - \tilde{\mu}_1(\lambda_1)D_{1t} - \tilde{\beta}_1(\lambda_1)DT_{1t}$. The following long-run variance estimators are used:

$$\begin{aligned} \hat{\omega}_1^2(\lambda_1) &= T^{-1} \sum_{t=1}^T \hat{u}_t^2(\lambda_1) + 2T^{-1} \sum_{j=1}^{T-1} (1 - j/(l+1)) \sum_{t=j+1}^T \hat{u}_t(\lambda_1) \hat{u}_{t-j}(\lambda_1) \\ \tilde{\omega}_1^2(\lambda_1) &= (T-1)^{-1} \sum_{t=1}^T \tilde{\varepsilon}_t^2(\lambda_1) + 2(T-1)^{-1} \sum_{j=1}^{T-2} (1 - j/(l+1)) \sum_{t=j+2}^T \tilde{\varepsilon}_t(\lambda_1) \tilde{\varepsilon}_{t-j}(\lambda_1) \end{aligned}$$

with $l = [4(T/100)^{1/4}]$. Next, consider stationarity statistics $S_0(\lambda_1)$ and $S_1(\lambda_1)$ calculated from the residuals $\{\hat{u}_t(\lambda_1)\}_{t=1}^T$ and $\{\tilde{\varepsilon}_t(\lambda_1)\}_{t=2}^T$ respectively:

$$\begin{aligned} S_0(\lambda_1) &= \frac{\sum_{t=1}^T (\sum_{i=1}^t \hat{u}_i(\lambda_1))^2}{T^2 \hat{\omega}_1^2(\lambda_1)} \\ S_1(\lambda_1) &= \frac{\sum_{t=2}^T (\sum_{i=2}^t \tilde{\varepsilon}_i(\lambda_1))^2}{(T-1)^2 \tilde{\omega}_1^2(\lambda_1)} \end{aligned}$$

The next step is to choose a weight function which converges to unity when u_t is $I(0)$ and to zero when u_t is $I(1)$. Based on the properties of the stationarity statistics, the weight function $\eta(S_0(\lambda_1), S_1(\lambda_1)) = \exp[-\{g_1 S_0(\lambda_1) S_1(\lambda_1)\}^{g_2}]$ is recommended. Finally, the proposed test statistic is

$$t_\eta = \left\{ \eta(S_0(\hat{\lambda}_1), S_1(\hat{\lambda}_1)) \right\} t_0(\hat{\lambda}_1) + m_\xi \left\{ \left[1 - \eta(S_0(\hat{\lambda}_1), S_1(\hat{\lambda}_1)) \right] \right\} t_1(\tilde{\lambda}_1) \quad (\text{A.4})$$

where $\hat{\lambda}_1 = \arg \sup_{s \in \Lambda_1} |t_0(\frac{s}{T})|$, $\tilde{\lambda}_1 = \arg \sup_{s \in \Lambda_1} |t_1(\frac{s}{T})|$ with $\Lambda_1 = [\epsilon T, (1 - \epsilon)T]$. The parameter ϵ determines the level of trimming used. The positive constant m_ξ is chosen such that, for a significance level ξ under H_0 , the asymptotic critical value in the $I(0)$ and $I(1)$ cases coincide. This ensures that the asymptotic null critical values of t_η are the same regardless of whether u_t is $I(0)$ or $I(1)$. Based on a range of Monte Carlo simulations on the finite sample size and power of the tests, they recommend choosing $g_1 = 500$ and $g_2 = 2$ for the construction of the weight function $\eta(\cdot)$. Note that both stationarity statistics are evaluated at the breakpoint estimator $\hat{\lambda}_1$, this being a consistent estimator of the true break fraction irrespective of whether u_t is stationary or not.

A.1.2 The Perron and Yabu (2009a) Test for a Break in Slope

Perron and Yabu (2009) propose an alternative approach to testing the stability of the trend function based on a Feasible Quasi Generalized Least Squares procedure. First, the OLS estimate of α is obtained from the autoregression

$$\hat{u}_t = \alpha \hat{u}_{t-1} + \sum_{i=1}^k \zeta_i \Delta \hat{u}_{t-i} + e_{tk} \quad (\text{A.5})$$

where k is chosen using the Bayesian Information Criterion (BIC) (k is allowed to be in the range $[0, [12(T/100)^{1/4}]]$). The corresponding estimate is denoted $\tilde{\alpha}$. To improve the finite sample properties of the tests, Perron and Yabu use a bias-corrected version of $\tilde{\alpha}$, denoted $\tilde{\alpha}_M$, proposed by Roy and Fuller (2001) (See Perron and Yabu, 2009a for details of the bias correction procedure). Next, Perron and Yabu propose the use of the following super-efficient estimate of α :

$$\tilde{\alpha}_{MS} = \begin{cases} \tilde{\alpha}_M & \text{if } |\tilde{\alpha}_M - 1| > T^{-1/2} \\ 1 & \text{if } |\tilde{\alpha}_M - 1| \leq T^{-1/2} \end{cases}$$

It is shown that using such a super-efficient estimate is crucial for obtaining procedures with nearly identical limit properties in the $I(0)$ and $I(1)$ cases. This estimate is then used to construct the quasi-differenced regression

$$\begin{aligned} (1 - \tilde{\alpha}_{MS}L)y_t &= (1 - \tilde{\alpha}_{MS}L)x'_{L1,t}\Psi + (1 - \tilde{\alpha}_{MS}L)u_t, \quad t = 2, \dots, T \\ y_1 &= x'_{L1,1}\Psi + u_1 \end{aligned} \quad (\text{A.6})$$

where $\Psi = (\mu_0, \beta_0, \mu_1, \beta_1)'$. Denote the resulting estimates by $\tilde{\Psi}^{FG} = (\tilde{\mu}_0^{FG}, \tilde{\beta}_0^{FG}, \tilde{\mu}_1^{FG}, \tilde{\beta}_1^{FG})'$. The Wald test $W_{QF}(\lambda_1)$ for a particular break fraction λ_1 , where the subscript QF stands for Quasi Feasible GLS, is given by

$$W_{QF}(\lambda_1) = \frac{\left(\tilde{\beta}_1^{FG}(\lambda_1)\right)^2}{\sqrt{\tilde{h}_v(\lambda_1) [(X^{\alpha'} X^{\alpha})^{-1}]_{44}}}$$

where $X^{\alpha} = \{x_{L1,1}, (1-\tilde{\alpha}_{MS}L)x_{L1,2}, \dots, (1-\tilde{\alpha}_{MS}L)x_{L1,T}\}'$. The quantity $\tilde{h}_v(\lambda_1)$ is an estimate of $(2\pi \text{ times})$ the spectral density function of $v_t = (1 - \alpha L)u_t$ at frequency zero. When $|\tilde{\alpha}_{MS}| < 1$, a kernel-based estimator

$$\tilde{h}_v(\lambda_1) = T^{-1} \sum_{t=1}^T \hat{v}_t^2(\lambda_1) + 2T^{-1} \sum_{j=1}^{T-1} k(j, \tilde{l}) \sum_{t=j+1}^T \hat{v}_t(\lambda_1) \hat{v}_{t-j}(\lambda_1)$$

is used where $\hat{v}_t(\lambda_1)$ are the OLS residuals from (A.6). The function $k(j, \tilde{l})$ is the quadratic spectral kernel and the bandwidth \tilde{l} is selected according to the plug-in method advocated by Andrews (1991) using an AR(1) approximation. When $|\tilde{\alpha}_{MS}| < 1$, the estimate suggested is a autoregressive spectral density estimate that can be obtained from the regression

$$\hat{v}_t = \sum_{i=1}^k \zeta_i \hat{v}_{t-i} + e_{tk} \quad (\text{A.7})$$

Denoting the estimate by $\hat{\zeta}(L) = 1 - \hat{\zeta}_1 L - \dots - \hat{\zeta}_k L^k$ and $\hat{\sigma}_{ek}^2 = (T - k)^{-1} \sum_{t=k+1}^T \hat{e}_{tk}^2$, $\tilde{h}_v = \hat{\sigma}_{ek}^2 / \hat{\zeta}(1)^2$. The order of the autoregression (A.7) is again selected using the BIC.

Following Andrews (1993) and Andrews and Ploberger (1994), Perron and Yabu considered the Mean, Exp and Sup functionals of the Wald test for different break dates. They found that with the Exp functional, the limit distributions in the $I(0)$ and $I(1)$ cases are nearly identical. They thus recommend the test statistic

$$ExpW = \log \left[T^{-1} \sum_{\lambda_1 \in \Lambda_1} \exp \left(\frac{1}{2} W_{QF}(\lambda_1) \right) \right]$$

A.1.3 The Harvey et al. (2009b) Test for Breaks in Level

Harvey et al. (2009b) propose a robust procedure for detecting multiple level breaks while accommodating a linear trend in the underlying data generating process. The model considered is

$$y_t = \mu_0 + \sum_{i=1}^n \mu_i I(t > T_i) + \beta_0 t + u_t$$

The null hypothesis is $H_0: \mu_i = 0$ for $i = 1, \dots, n$ while the alternative is that of at least one break in level; that is $H_1: \mu_i \neq 0$ for at least one $i \in \{1, \dots, n\}$. Let $\hat{\beta}_0$ denote the estimator of the trend coefficient, β_0 , from the OLS regression of y_t on $\{1, t\}$, $t = 1, \dots, T$. The proposed test statistic is based on the quantities

$$\begin{aligned} M &= \max_{t \in \Lambda_1} \left| M_{t, [mT]} - \hat{\beta}_0 \left[\frac{m}{2} T \right] \right| \\ S_0 &= (\hat{\omega}_v)^{-1} T^{-1/2} M \\ S_1 &= (\hat{\omega}_u)^{-1} T^{1/2} M \end{aligned}$$

where

$$M_{t, [mT]} = \frac{\sum_{i=1}^{\lfloor \frac{m}{2} T \rfloor} y_{t+i} - \sum_{i=1}^{\lfloor \frac{m}{2} T \rfloor} y_{t-i+1}}{\lfloor \frac{m}{2} T \rfloor}$$

and $\hat{\omega}_v, \hat{\omega}_u$ denoting long-run variance estimates appropriate for the case of $I(1)$ and $I(0)$ shocks, respectively (see Harvey et al., 2009b for details on the construction of these estimates). Based on the finite sample properties of the procedure, the choice $m = 0.10$ is recommended for practice. The proposed test is

$$U = \max \left\{ S_1, \left(\frac{cv_\xi^1}{cv_\xi^0} \right) S_0 \right\}$$

where cv_ξ^1 and cv_ξ^0 denote the ξ -level asymptotic critical values of S_1 under $I(1)$ errors and S_0 under $I(0)$ errors, respectively. The computed value of U is then compared with $\kappa_\xi cv_\xi^1$, where $\kappa_\xi = cv_\xi^{\max} / cv_\xi^1$, where cv_ξ^{\max} is the ξ -level critical value from the limit distribution of $\max \left\{ S_1, \left(\frac{cv_\xi^1}{cv_\xi^0} \right) S_0 \right\}$.

A.2 Procedures for Selecting the Number of Breaks

A.2.1 The Kejriwal and Perron (2009) Sequential Procedure for Slope Breaks

Buliding on the work of Perron and Yabu (2009a), Kejriwal and Perron (2009) propose a sequential procedure that allows one to obtain a consistent estimate of the true number of breaks irrespective of whether the errors are $I(1)$ or $I(0)$. The first step is to conduct a test for no break versus one break. Conditional on a rejection, the estimated break date is obtained by a global minimization of the sum of squared residuals. The strategy proceeds by testing each of the two segments (obtained using the estimated partition) for the presence of an additional break and assessing whether the maximum of the tests is significant. Formally, the test of one versus two breaks is expressed as

$$ExpW(2|1) = \max_{1 \leq i \leq 2} \{ ExpW^{(i)} \}$$

where $ExpW^{(i)}$ is the one break test in segment i . We conclude in favor of a model with two breaks if $ExpW(2|1)$ is sufficiently large.¹

A.2.2 The Harvey et al. (2009b) Sequential Procedure for Level Breaks

Harvey et al. (2009b) also propose the following sequential procedure for selecting the number of level breaks in addition to the U test discussed above. First, if $S_1 > \kappa_\xi cv_\xi^1$, denote $\tilde{t}_1 = \operatorname{argmax}_{t \in \Lambda_1} (\hat{\omega}_v)^{-1} T^{-1/2} \left| M_{t, [mT]} - \hat{\beta}_0 \left[\frac{m}{2} T \right] \right|$. Then, denoting $\Lambda_2 = [\tilde{t}_1 - [mT] + 1, \tilde{t}_1 + [mT] - 1]$, if $\max_{t \in \Lambda_1 - \Lambda_2} (\hat{\omega}_v)^{-1} T^{-1/2} \left| M_{t, [mT]} - \hat{\beta}_0 \left[\frac{m}{2} T \right] \right| \leq cv_\xi^1$, we conclude that the procedure based on S_1 selects one break; otherwise, two breaks are selected. The number of breaks is denoted n'_1 . A similar procedure based on S_0 gives n'_0 breaks. The number of breaks selected by the sequential procedure based on U is then $n_U = \max(n'_1, n'_0)$. For a given number of breaks, consistent estimates of the break dates in the presence of $I(1)$ errors are also suggested (See Harvey et al., 2009b for details).

A.3 Unit Root Tests

A.3.1 The Harris et al. (2009) Test

Harris et al. (2009) propose a test for a unit root in the presence of a possible trend break based on a GLS detrending procedure similar to that used by Elliott et al. (1996) in the stable trend case. Consider the model given by (1) and (2). The first step is to obtain an estimate of the break fraction by minimizing the sum of squared residuals from OLS estimation of the first differenced regression (A.1). This is denoted $\tilde{\lambda}_1$. Applying a quasi-differenced transformation to (1) yields

$$(1 - \alpha(\tilde{\lambda}_1)L)y_t = (1 - \alpha(\tilde{\lambda}_1)L)x'_{L1,t}(\tilde{\lambda}_1)\Psi + (1 - \alpha(\tilde{\lambda}_1)L)u_t, \quad \alpha(\tilde{\lambda}_1) = 1 - \frac{c(\tilde{\lambda}_1)}{T} \quad (\text{A.8})$$

where $c(\tilde{\lambda}_1)$ denotes the value at which the asymptotic Gaussian local power envelope for a break fraction $\tilde{\lambda}_1$ at a given significance level has power equal to .50. Letting $\tilde{\Psi}_{c(\tilde{\lambda}_1)}$ and $\tilde{u}_{t,c(\tilde{\lambda}_1)}$ denote the OLS estimate and residuals from (A.8), the next step is to estimate the Augmented Dickey -Fuller type regression

$$\Delta \tilde{u}_{t,c(\tilde{\lambda}_1)} = \phi \tilde{u}_{t-1,c(\tilde{\lambda}_1)} + \sum_{j=1}^{k_1} \delta_j \Delta \tilde{u}_{t-j,c(\tilde{\lambda}_1)} + e_{k_1,t}, \quad t = k_1 + 2, \dots, T \quad (\text{A.9})$$

¹For the general model with k breaks, the estimated break points are obtained by a global minimization of the sum of squared residuals. The strategy proceeds by testing each $k + 1$ segment (obtained using the estimated partition) for the presence of an additional break. The test thus amounts to the application of $k + 1$ tests of the null hypothesis of no change versus the alternative hypothesis of a single change and assessing whether the maximum is significant. See Kejriwal and Perron (2009) for more details.

The unit root statistic, denoted H , is then the t -statistic for $\phi = 0$ in (A.9). The lag length k_1 is selected using the modified Akaike Information Criterion (MAIC) proposed in Ng and Perron (2001).

A.3.2 The Carrion et al. (2009) Tests

Carrion et. al (2009) propose an alternative testing procedure which allows for multiple structural breaks in the level and/or slope of the trend function under both the null and alternative hypotheses. The tests are extensions of the M class of tests analyzed in Ng and Perron (2001) and the feasible point optimal statistic of Elliott et al. (1996). We will provide a brief description of the tests for the two breaks model. The model is

$$y_t = \mu_0 + \beta_0 t + \mu_1 DU_{1t} + \beta_1 DT_{1t} + \mu_2 DU_{2t} + \beta_2 DT_{2t} + u_t$$

where $DU_{it} = I(t > T_i)$, $DT_{it} = (t - T_i)I(t > T_i)$ ($i = 1, 2$) and the errors u_t are generated as in (2). First, the estimates of the break fractions $\lambda = (\lambda_1, \lambda_2)$ and the regression parameters are obtained by minimizing the sum of squared residuals from the quasi-differenced regression analogous to (A.8). The sum of squared residuals evaluated at these estimates is denoted $S(\alpha(\hat{\lambda}), \hat{\lambda})$ with $\alpha(\hat{\lambda}) = 1 - \frac{c(\hat{\lambda})}{T}$. The feasible point optimal statistic is then

$$P_T^{gls}(\hat{\lambda}) = \frac{S(\alpha(\hat{\lambda}), \hat{\lambda}) - \alpha(\hat{\lambda})S(1, \hat{\lambda})}{s^2(\hat{\lambda})}$$

where $s^2(\hat{\lambda})$ is an autoregressive estimate of the spectral density of v_t at frequency zero:

$$s^2(\hat{\lambda}) = s_{ek}^2 / (1 - \hat{b}(1))^2 \quad (\text{A.10})$$

where $s_{ek}^2 = (T - k)^{-1} \sum_{t=k+1}^T \hat{e}_{tk}^2$, $\hat{b}(1) = \sum_{j=1}^k \hat{b}_j$, with \hat{b}_j and \hat{e}_{tk} obtained from the OLS estimation of

$$\Delta \tilde{y}_t = b_0 \tilde{y}_{t-1} + \sum_{j=1}^k b_j \Delta \tilde{y}_{t-j} + e_{tk}$$

with

$$\tilde{y}_t = y_t - \hat{\Psi}'_2 x_{L2,t}(\hat{\lambda}), \quad x_{L2,t}(\hat{\lambda}) = \left\{ 1, t, DU_{1t}(\hat{\lambda}), DU_{2t}(\hat{\lambda}), DT_{1t}(\hat{\lambda}), DT_{2t}(\hat{\lambda}) \right\} \quad (\text{A.11})$$

and $\hat{\Psi}_2$ being the OLS estimate obtained from the quasi-differenced regression.

Carrion et al. (2009) also consider extensions of the M -class of tests analyzed in Ng and

Perron (2001). These are given by

$$\begin{aligned}
MZ_{\alpha}^{gls}(\hat{\lambda}) &= (T^{-1}\tilde{y}_T^2 - s^2(\hat{\lambda}))(2T^{-2} \sum_{t=2}^T \tilde{y}_{t-1}^2)^{-1} \\
MSB^{gls}(\hat{\lambda}) &= (T^{-2} \sum_{t=2}^T \tilde{y}_{t-1}^2)^{1/2} / s^2(\hat{\lambda}) \\
MZ_t^{gls}(\hat{\lambda}) &= (T^{-1}\tilde{y}_T^2 - s^2(\hat{\lambda}))(4s^2(\hat{\lambda})T^{-2} \sum_{t=2}^T \tilde{y}_{t-1}^2)^{-1/2} \\
MP_T^{gls}(\hat{\lambda}) &= [c^2(\hat{\lambda})T^{-2} \sum_{t=2}^T \tilde{y}_{t-1}^2 + (1 - c(\hat{\lambda}))T^{-1}\tilde{y}_T^2] / s^2(\hat{\lambda}) \quad (\text{A.12})
\end{aligned}$$

where $s^2(\hat{\lambda})$ and \tilde{y}_t are as defined in (A.10) and (A.11). These test statistics (with obvious modifications) are also used to detect pure level breaks with a stable slope parameter. See Carrion et al. (2009) for details.

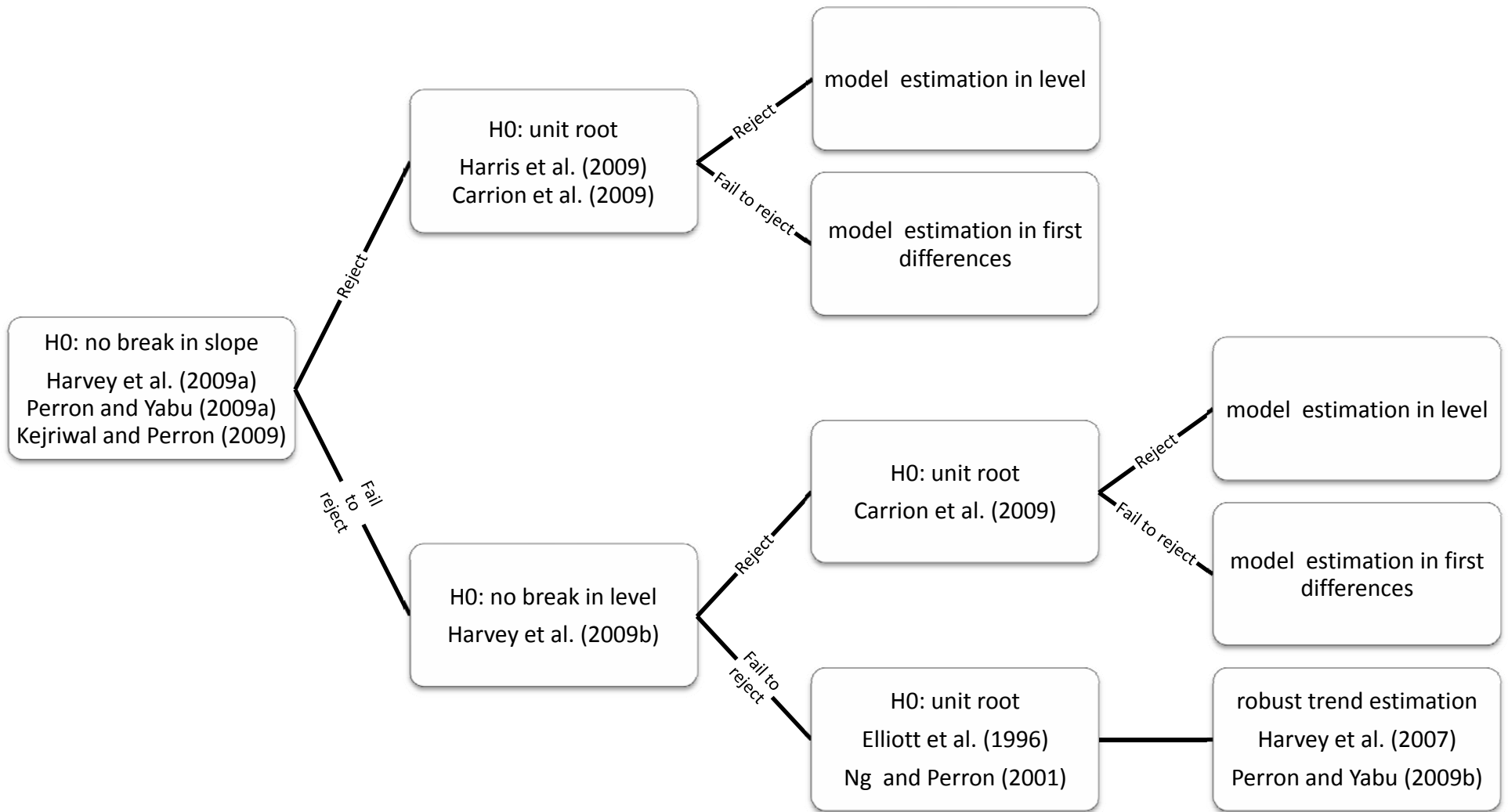


Figure 1: The Proposed Algorithm

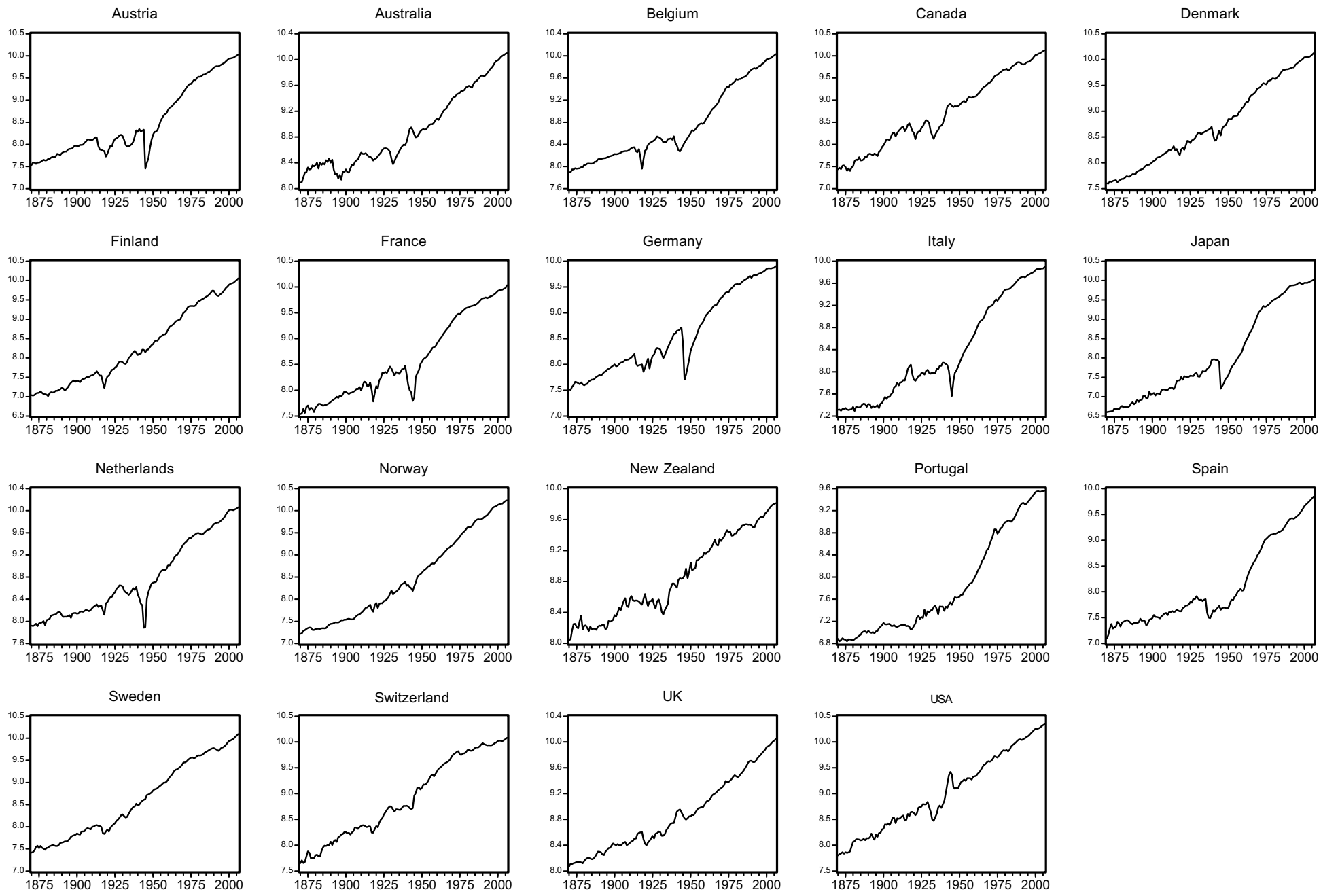


Figure 2: Per-Capita GDP of OECD Countries

Table 1: Empirical Size of Joint and Individual Tests, $T = 150$

μ_b	$\rho = .5$			$\rho = .6$			$\rho = .7$			$\rho = .8$			$\rho = .9$			$\rho = 1$		
	$ExpW_J$	$ExpW$	t_η	$ExpW_J$	$ExpW$	t_η	$ExpW_J$	$ExpW$	t_η	$ExpW_J$	$ExpW$	t_η	$ExpW_J$	$ExpW$	t_η	$ExpW_J$	$ExpW$	t_η
(A) $\lambda = 0.3$																		
-0.04	.10	.09	.06	.10	.07	.08	.10	.07	.09	.07	.04	.11	.05	.02	.09	.60	.11	.19
-0.02	.09	.09	.06	.09	.07	.08	.10	.07	.09	.08	.04	.12	.05	.02	.09	.16	.11	.19
.02	.09	.09	.07	.09	.07	.08	.09	.07	.09	.07	.05	.12	.05	.02	.09	.16	.11	.20
.04	.09	.09	.06	.09	.07	.07	.09	.07	.09	.07	.05	.12	.05	.02	.09	.61	.12	.20
(B) $\lambda = 0.5$																		
-0.04	.10	.08	.07	.09	.07	.08	.09	.07	.10	.07	.04	.11	.05	.02	.09	.64	.11	.16
-0.02	.10	.09	.06	.09	.06	.08	.09	.07	.10	.07	.04	.12	.05	.02	.09	.14	.11	.19
.02	.09	.08	.06	.09	.07	.08	.10	.07	.09	.07	.04	.12	.05	.02	.09	.15	.11	.18
.04	.09	.08	.06	.09	.07	.08	.09	.07	.09	.07	.04	.12	.05	.02	.09	.60	.11	.20
(C) $\lambda = 0.7$																		
-0.04	.09	.09	.07	.09	.07	.07	.09	.07	.09	.07	.05	.12	.05	.02	.10	.61	.09	.18
-0.02	.09	.09	.07	.09	.07	.08	.09	.07	.10	.07	.05	.12	.05	.02	.09	.14	.10	.18
.02	.09	.08	.06	.10	.07	.08	.10	.07	.10	.08	.04	.12	.05	.02	.09	.14	.12	.20
.04	.09	.08	.06	.10	.07	.08	.10	.07	.09	.08	.04	.12	.05	.02	.09	.14	.11	.20

Table 2: Empirical Size of Joint and Individual Tests, $T = 200$

μ_b	$\rho = .5$			$\rho = .6$			$\rho = .7$			$\rho = .8$			$\rho = .9$			$\rho = 1$		
	$ExpW_J$	$ExpW$	t_η	$ExpW_J$	$ExpW$	t_η	$ExpW_J$	$ExpW$	t_η	$ExpW_J$	$ExpW$	t_η	$ExpW_J$	$ExpW$	t_η	$ExpW_J$	$ExpW$	t_η
(A) $\lambda = 0.3$																		
-0.04	.09	.08	.09	.07	.08	.11	.09	.07	.11	.06	.04	.15	.05	.03	.09	.57	.11	.17
-0.02	.09	.08	.09	.07	.07	.11	.09	.07	.11	.06	.04	.15	.05	.03	.09	.12	.12	.16
.02	.09	.08	.09	.07	.07	.10	.08	.07	.12	.06	.05	.15	.05	.03	.09	.14	.11	.16
.04	.09	.08	.09	.08	.07	.10	.09	.07	.11	.06	.05	.15	.05	.03	.09	.62	.10	.15
(B) $\lambda = 0.5$																		
-0.04	.09	.08	.09	.07	.07	.10	.09	.07	.12	.07	.04	.15	.05	.03	.09	.59	.10	.14
-0.02	.09	.08	.10	.07	.07	.10	.09	.07	.12	.06	.04	.15	.05	.03	.09	.13	.11	.15
.02	.09	.08	.09	.08	.07	.07	.09	.06	.09	.06	.04	.06	.05	.03	.05	.13	.12	.13
.04	.09	.08	.09	.07	.08	.10	.09	.07	.11	.06	.05	.15	.05	.03	.09	.58	.12	.15
(C) $\lambda = 0.7$																		
-0.04	.09	.08	.09	.08	.07	.10	.09	.07	.11	.06	.04	.15	.05	.03	.09	.60	.11	.17
-0.02	.09	.08	.10	.07	.07	.10	.08	.07	.11	.06	.04	.15	.05	.03	.09	.14	.12	.18
.02	.09	.08	.09	.07	.08	.10	.09	.07	.12	.06	.04	.15	.05	.03	.09	.14	.11	.17
.04	.10	.08	.09	.07	.07	.10	.08	.07	.11	.06	.07	.15	.05	.04	.09	.57	.10	.16

Table 3: MSE-Ratio (MSE_D/MSE_L), $T = 150$

β_1^0	μ_1^0	$\alpha^0 = .5$	$\alpha^0 = .6$	$\alpha^0 = .7$	$\alpha^0 = .8$	$\alpha^0 = .9$	$\alpha^0 = 1$
(A) $\lambda_1^0 = 0.3$							
-0.02	-0.04	2.219	1.764	1.409	1.204	1.189	0.0105
	-0.02	2.148	1.599	1.288	1.107	1.043	0.2045
	0.00	2.040	1.695	1.378	1.165	1.063	0.3775
	0.02	2.150	1.674	1.342	1.161	1.107	0.2982
	0.04	2.092	1.672	1.339	1.198	1.074	0.0979
0.02	-0.04	2.187	1.679	1.393	1.234	1.202	0.0754
	-0.02	2.083	1.647	1.342	1.147	1.049	0.3696
	0.00	2.172	1.608	1.304	1.086	1.034	0.4002
	0.02	2.129	1.682	1.400	1.171	1.119	0.1636
	0.04	2.121	1.638	1.385	1.191	1.075	0.0282
(B) $\lambda_1^0 = 0.5$							
-0.02	-0.04	1.652	1.370	1.198	1.097	1.178	0.0096
	-0.02	1.672	1.436	1.210	1.118	1.157	0.1291
	0.00	1.681	1.362	1.179	1.099	1.169	0.3860
	0.02	1.695	1.382	1.182	1.044	1.118	0.3219
	0.04	1.645	1.372	1.181	1.083	1.145	0.1122
0.02	-0.04	1.683	1.378	1.198	1.091	1.161	0.1262
	-0.02	1.681	1.433	1.231	1.164	1.183	0.3394
	0.00	1.738	1.399	1.196	1.084	1.126	0.3945
	0.02	1.703	1.426	1.162	1.083	1.161	0.1129
	0.04	1.587	1.369	1.188	1.073	1.159	0.0239
(C) $\lambda_1^0 = 0.7$							
-0.02	-0.04	2.197	1.680	1.365	1.124	1.047	0.0036
	-0.02	2.314	1.861	1.496	1.282	1.207	0.0915
	0.00	2.195	1.706	1.356	1.201	1.167	0.3262
	0.02	2.248	1.744	1.362	1.183	1.163	0.3645
	0.04	2.282	1.806	1.432	1.186	1.183	0.1821
0.02	-0.04	2.076	1.578	1.328	1.109	1.030	0.1984
	-0.02	2.105	1.655	1.392	1.242	1.206	0.3674
	0.00	2.313	1.756	1.499	1.269	1.153	0.3455
	0.02	2.334	1.811	1.491	1.195	1.186	0.0888
	0.04	2.102	1.687	1.402	1.229	1.180	0.0085

MSE_D and MSE_L are the mean squared errors of the estimated break dates from the specification in first differences and levels

Table 4: MSE-Ratio (MSE_D/MSE_L), $T = 200$

β_1^0	μ_1^0	$\alpha^0 = .5$	$\alpha^0 = .6$	$\alpha^0 = .7$	$\alpha^0 = .8$	$\alpha^0 = .9$	$\alpha^0 = 1$
(A) $\lambda_1^0 = 0.3$							
-0.02	-0.04	4.854	3.095	2.012	1.358	1.164	0.0139
	-0.02	4.456	3.030	1.958	1.416	1.131	0.1051
	0.00	4.817	2.943	1.973	1.382	1.098	0.2686
	0.02	4.889	3.164	1.996	1.401	1.101	0.2393
	0.04	4.702	3.073	2.037	1.416	1.147	0.0782
0.02	-0.04	4.716	3.132	2.050	1.444	1.207	0.0637
	-0.02	4.869	3.106	1.986	1.411	1.201	0.2520
	0.00	4.537	3.032	2.018	1.428	1.174	0.2670
	0.02	4.695	2.990	1.936	1.434	1.096	0.1062
	0.04	4.774	3.236	2.090	1.447	1.212	0.0122
(B) $\lambda_1^0 = 0.5$							
-0.02	-0.04	3.170	2.359	1.678	1.273	1.164	0.0019
	-0.02	3.055	2.146	1.569	1.206	1.121	0.0901
	0.00	3.122	2.293	1.673	1.278	1.148	0.2784
	0.02	2.980	2.185	1.620	1.249	1.151	0.2473
	0.04	2.999	2.195	1.553	1.189	1.116	0.0953
0.02	-0.04	3.162	2.234	1.648	1.255	1.105	0.0823
	-0.02	3.002	2.213	1.585	1.188	1.118	0.2661
	0.00	3.083	2.290	1.668	1.276	1.150	0.2689
	0.02	3.061	2.238	1.605	1.270	1.147	0.0848
	0.04	3.011	2.170	1.569	1.238	1.128	0.0067
(C) $\lambda_1^0 = 0.7$							
-0.02	-0.04	4.491	2.963	1.985	1.443	1.098	0.0064
	-0.02	4.427	3.072	1.979	1.352	1.127	0.0722
	0.00	4.397	2.925	2.052	1.479	1.204	0.2545
	0.02	4.157	3.008	1.988	1.428	1.205	0.3096
	0.04	4.302	2.894	1.957	1.463	1.156	0.1549
0.02	-0.04	4.808	3.246	2.180	1.529	1.138	0.1159
	-0.02	4.239	3.023	2.033	1.410	1.168	0.2961
	0.00	4.615	3.175	2.017	1.451	1.142	0.2211
	0.02	4.419	3.029	2.092	1.489	1.242	0.0802
	0.04	4.413	3.136	2.125	1.486	1.166	0.0039

MSE_D and MSE_L are the mean squared errors of the estimated break dates from the specification in first differences and levels

Table 5: Robust Tests for Breaks in Trend

Country\Test	Slope Breaks				Level Breaks	
	<i>ExpW</i>	<i>ExpW</i> (2 1)	t_{η}	#Breaks	<i>U</i>	#Breaks
Australia	0.87	0.41	2.67	0	0.37	0
Austria	2.82*	235.78*	5.06*	2	-	-
Belgium	2.36*	13.18*	5.89*	2	-	-
Canada	-0.21	-0.03	1.50	0	0.50	0
Denmark	0.38	1.90	2.63	0	0.56	2
Finland	3.41*	0.05	4.30*	1	-	-
France	1.56	14.34*	2.87	0/2	0.52	0
Germany	0.40	8.14*	3.18*	2	-	-
Italy	0.99	12.70*	3.36*	2	-	-
Japan	2.50*	26.79*	3.82*	2	-	-
Netherlands	1.36	7.64*	3.08*	2	-	-
Norway	2.35*	0.07	5.96*	1	-	-
New Zealand	-0.01	5.84*	3.76*	2	-	-
Portugal	4.85*	1.02	4.10*	1	-	-
Spain	4.99*	1.03	3.82*	1	-	-
Sweden	1.80*	1.33	3.08*	1	-	-
Switzerland	-0.04	29.52*	1.47	0/2	0.44	0
UK	12.67*	0.12	2.23	1	-	-
USA	-0.05	0.68	1.38	0	0.57	1

Here '*' denotes significance at the 5% level.

Table 6: Unit Root Tests

Country\Test	Slope Breaks						Level Breaks					No Breaks				
	MZ_a^{gls}	MSB^{gls}	MZ_t^{gls}	P_T^{gls}	MP_T^{gls}	H	MZ_a^{gls}	MSB^{gls}	MZ_t^{gls}	P_T^{gls}	MP_T^{gls}	MZ_a^{gls}	MSB^{gls}	MZ_t^{gls}	P_T^{gls}	MP_T^{gls}
Australia	-	-	-	-	-	-	-	-	-	-	-	-5.09	0.29	-1.47	18.12	17.37
Austria	-21.34	0.15	-3.21	11.76	10.08	-	-	-	-	-	-	-	-	-	-	-
Belgium	-14.81	0.18	-2.66	14.49	14.03	-	-	-	-	-	-	-	-	-	-	-
Canada	-	-	-	-	-	-	-	-	-	-	-	-6.94	0.27	-1.85	13.64	13.15
Denmark	-	-	-	-	-	-	-2.02	0.48	-0.98	52.78	43.54	-	-	-	-	-
Finland	-14.68	0.18	-2.68	7.19	7.14	-2.81	-	-	-	-	-	-	-	-	-	-
France	-14.29	0.18	-2.62	16.06	14.57	-	-	-	-	-	-	-4.07	0.34	-1.38	23.00	21.79
Germany	-17.21	0.17	-2.86	12.05	12.02	-	-	-	-	-	-	-	-	-	-	-
Italy	-13.32	0.19	-2.49	19.00	15.62	-	-	-	-	-	-	-	-	-	-	-
Japan	-19.38	0.16	-3.10	10.53	10.06	-	-	-	-	-	-	-	-	-	-	-
Netherlands	-13.57	0.19	-2.59	16.16	15.16	-	-	-	-	-	-	-	-	-	-	-
Norway	-15.21	0.17	-2.72	7.07	6.92	-2.83	-	-	-	-	-	-	-	-	-	-
New Zealand	-11.54	0.21	-2.39	18.63	17.62	-	-	-	-	-	-	-	-	-	-	-
Portugal	-7.70	0.24	-1.89	13.15	12.90	-1.92	-	-	-	-	-	-	-	-	-	-
Spain	-6.94	0.26	-1.86	15.22	14.21	-2.10	-	-	-	-	-	-	-	-	-	-
Sweden	-7.13	0.25	-1.82	14.33	14.40	-1.61	-	-	-	-	-	-	-	-	-	-
Switzerland	-12.07	0.20	-2.39	17.98	17.19	-	-	-	-	-	-	-7.84	0.25	-1.94	11.80	11.72
UK	-14.57	0.18	-2.62	7.71	7.47	-2.76	-	-	-	-	-	-	-	-	-	-
USA	-	-	-	-	-	-	-7.84	0.25	-1.97	12.81	11.66	-	-	-	-	-

Table 7b: Parameter Estimates - Countries with a Stable Slope (in Percentage)

Country\Estimate	One Level Break		Two Level Breaks			No Level Breaks	
	$\hat{\beta}_0$	Date	$\hat{\beta}_0$	Date 1	Date 2	$\hat{\beta}_0$ (HLT ⁰)	$\hat{\beta}_0$ (PY ⁰)
Australia	-	-	-	-	-	1.47 [0.85,2.10]	1.48 [0.53,2.42]
Canada	-	-	-	-	-	1.99 [1.31,2.67]	1.97 [1.85,2.09]
Denmark	-	-	2.06 [1.54,2.58]	1914	1939	-	-
France	-	-	-	-	-	1.84 [0.76,2.92]	1.84 [0.50,3.17]
Switzerland	-	-	-	-	-	1.81 [1.30,2.32]	1.79 [1.08,2.49]
USA	2.06 [1.07,3.05]	1945	-	-	-	-	-

Table 8a: Studies of the Maddison Dataset allowing for One Endogenous Break

Country	Raj (1992)			Zelhorst et al. (1995)	Ben-David et al. (1995)		
	Date	$\hat{\beta}_0$	$\sum_{i=0}^1 \hat{\beta}_i$	Date	Date	$\hat{\beta}_0$	$\sum_{i=0}^1 \hat{\beta}_i$
Australia	1928	0.1	0.7	1927	1927*	0.16	0.65
Austria	-	-	-	-	1944*	0.4	2.23
Belgium	-	-	-	-	1939*	0.29	1.08
Canada	1930*	0.9	1.7	1928*	1928*	1.00	1.26
Denmark	1940*	0.8	1.6	1939*	1939*	0.66	1.19
Finland	-	-	-	1913*	1913*	0.40	0.75
France	1940*	0.6	2.2	1939*	1939*	0.53	1.68
Germany	-	-	-	1946/1953*	1955*	1.67	2.46
Italy	1943	0.4	0.7	1945/1942*	1939	0.27	0.81
Japan	-	-	-	-	1944*	1.02	3.17
Netherlands	-	-	-	1945/1939	1939	0.71	1.49
Norway	-	-	-	1944/1939	1939	0.33	0.68
Sweden	1917	0.6	0.8	1915*	1916*	0.47	0.91
Switzerland	-	-	-	-	1944	0.26	-
UK	1919*	0.3	0.5	1918*	1918*	0.29	0.52
USA	1930*	1.1	1.3	1929*	1929*	0.9	1.13

Note: Here '**' denotes rejection of the unit root hypothesis.

Table 8b: Studies of the Maddison Dataset allowing for Two Endogenous Breaks

Country	Ben-David et al. (2003)				Papell and Prodan (2009)			
	(Date 1,Date 2)	$\hat{\beta}_0$	$\sum_{i=0}^1 \hat{\beta}_i$	$\sum_{i=0}^2 \hat{\beta}_i$	(Date 1,Date 2)	$\hat{\beta}_0$	$\sum_{i=0}^1 \hat{\beta}_i$	$\sum_{i=0}^2 \hat{\beta}_i$
Australia	(1891,1927)*	1.31	0.65	1.87	1931*	0.07	0.42	-
Austria	(1944,1959)*	1.09	8.73	3.11	(1944,1976)*	0.29	12.86	1.00
Belgium	(1916,1939)*	0.90	1.63	2.62	(1939,1976)*	0.29	1.12	0.61
Canada	(1908,1928)*	1.98	1.76	2.35	(1930,1940)*	0.46	2.25	0.49
Denmark	(1939,1975)*	1.53	2.70	1.53	(1939,1969)*	0.77	1.65	0.82
Finland	(1916,1943)*	1.11	3.20	3.38	-	-	-	-
France	(1939,1974)*	1.29	3.49	1.86	(1939,1972)*	0.62	2.24	0.81
Germany	(1944,1958)	1.83	6.81	2.76	(1944,1950)*	0.2	9.74	0.33
Italy	(1942,1966)	1.03	4.15	2.84	(1942,1948)*	0.19	5.27	0.32
Japan	(1944,1973)*	1.69	7.68	3.29	(1944,1972)*	1.23	5.38	1.60
Netherlands	(1939,1975)	1.09	2.87	1.09	(1943,1946)*	0.24	27.71	0.65
Norway	(1917,1939)*	1.10	3.14	3.21	-	-	-	-
Spain	-	-	-	-	(1935,1972)*	0.18	1.01	0.54
Sweden	(1916,1963)*	1.27	2.93	2.11	(1916,1968)*	0.48	1.07	0.55
Switzerland	(1921,1944)	0.98	1.21	2.27	1944*	0.25	0.23	-
UK	(1918,1945)*	1.12	1.13	2.13	1918*	0.26	0.45	-
USA	(1929,1955)*	1.77	2.13	1.85	(1929,1942)*	0.61	2.00	0.71

Note: Here '**' denotes rejection of the unit root hypothesis.