GROWINPRO

Growth Welfare Innovation Productivity

Working Paper

Characterizing growth instability: new evidence on unit roots and structural breaks in long run time series

Emanuele Russo
Institute of Economics and EMbeDS Department, Scuola Superiore Sant'Anna, Pisa, Italy

Neil Foster-McGregor
United Nations University - Maastricht Economic and social Research institute on Innovation and Technology, Maastrich (The Netherlands)

Bart Verspagen
United Nations University - Maastricht Economic and social Research institute on Innovation and Technology, Maastrich (The Netherlands)

17/2019 **September**

This project has received funding from the European Union Horizon 2020 Research and Innovation action under grant agreement No 822781

Characterizing growth instability: new evidence on unit roots and structural breaks in long run time series

Emanuele Russo *[1], Neil Foster-McGregor †[2], and Bart Verspagen ‡ [2]

[1]Institute of Economics and EMbeDS, Scuola Superiore Sant'Anna, Pisa (Italy)
[2]United Nations University - Maastricht Economic and social Research institute on Innovation and Technology, Maastrich (The Netherlands)

Abstract

In this paper we investigate whether long run time series of income per capita are better described by a trend-stationary model with few structural changes or by unit root processes in which permanent stochastic shocks are responsible for the observed growth discontinuities. To this purpose, we develop a methodology to test the null of a generic I(1) process versus a set of stationary alternatives with structural breaks. Differently from other tests in the literature, the number of structural breaks under the alternative hypothesis is treated as an unknown (up to some ex ante determined maximum). Critical values are obtained via Monte Carlo simulations and finite sample size and power properties of the test are reported. An application is provided for a group of advanced and developing countries in the Maddison dataset, also using bootstrapped critical values. As compared to previous findings in the literature, less evidence is found against the unit root hypothesis. Failures to reject the I(1) null are particularly strong for a set of developing countries considered. Finally, even less rejections are found when relaxing the assumption of Gaussian shocks.

Keywords Long-run growth, structural breaks, unit roots

JEL classification O47, C22

1 Introduction

Among empirical growth economists a quite large consensus has emerged over the last years concerning the unstable nature of economic growth. It is now widely recognized that the vast majority of growth experiences, even when considering now-rich countries, do not comply with a simple steady-growth model (Ben-David and Papell, 1995; Papell and Prodan, 2014). In aggregate income data, it is common to observe several growth discontinuities of different kinds such as accelerations and collapses, sudden stops or level jumps (Easterly et al., 1993; Hausmann et al., 2005; Lamperti and

*emanuele.russo@santannapisa.it
†foster@merit.unu.edu
‡verspagen@merit.unu.edu

Mattei, 2018; Pritchett et al., 2000). Nevertheless, there is clearly less consensus when it comes to characterizing growth instability with econometric models.

A major issue is whether growth paths are better described by a trend-stationary model with relatively few structural breaks (either in growth or in levels) or by unit root processes in which permanent stochastic shocks are responsible for continuous changes. As pointed out by Perron et al. (2006), one should not restrict the analysis to these two limiting cases as there are several interesting instances in between. The key question is therefore: do the data reveal frequent and large growth discontinuities or are structural changes occurring at most occasionally? From this standpoint, testing for unit roots, rather than discerning definitively between stationary vis-à-vis integrated models, allows exactly to make inference on where do we stand between these two alternatives. Addressing this question has strong empirical and theoretical implications.

In terms of economic theory, a trend-stationary specification seems to be consistent with models from both the Neoclassical and the New Growth traditions. More precisely, the former typically suggests a log-linear trend with level shifts in response to changes in some policy variables while the latter also accounts for growth effects.[1] In new or endogenous growth models, exogenous shocks affecting the accumulation of physical and human capital (Lucas, 1988; Romer, 1986) or R&D expenditures (Aghion and Howitt, 1992; Grossman and Helpman, 1991) cause shifts in the equilibrium growth rate of the economy. On the contrary, evolutionary models emphasize out-of-equilibrium dynamics and can hardly be reconciled with a trend-stationary data generating process (Dosi et al., 2019; Nelson and Winter, 1982; Silverberg and Verspagen, 1994, 1995).[2] This is the case, for instance, in the so-called "K+S" family of models in which demand-driven fluctuations may affect the long run path of the economy via their interaction with firms' innovation patterns (Dosi et al., 2013, 2015, 2010, 2017; Lamperti et al., 2018).[3] In this framework, growth is inherently stochastic as it results from the aggregation of endogenous stochastic innovations at the microeconomic level and, therefore, models are probably better approximated by path-dependent processes akin to random walks. Policy shocks still play a crucial role but their effect is far from being deterministic, depending on the specific realization of events associated with the arrival of innovations and to their "disruptive" consequences on the economic system.[4]

From an empirical point of view, the relevance of discerning models with stochastic trends from stationary alternatives is twofold. First, it has been shown that the interpretation and the implications of convergence tests change greatly when output series follow I(1) processes (Lee et al., 1997). Most importantly, the standard practices adopted in order to identify growth episodes typically disregard prior unit roots testing, albeit with some exceptions (Ben-David and Papell, 1995; Papell and Prodan, 2014; Sobreira et al., 2014). The search for growth episodes is generally carried out either by formal tests for structural breaks (Berg et al., 2012; Jones and Olken, 2008; Kerekes, 2007) or by imposing filters based on subjective economic criteria (Aizenman and Spiegel, 2010; Bluhm et al.,

[1] The emergence of level effects is also a characteristic of semi-endogenous growth models (Jones, 1995, 2005). For an empirical classification of countries' growth paths according to the "constant trend", "level shifts" or "trend shifts" hypothesis see Papell and Prodan (2014) and Sobreira et al. (2014).

[2] For a comparative survey of evolutionary and endogenous growth theories see Castellacci (2007).

[3] A framework akin to the "K+S" model has been exploited in Dosi et al. (2019) to investigate long-run growth patterns among several inter-dependent economies.

[4] One may argue that the evolutionary view on the role of stochastic events and path dependence in growth trajectories is shared also by some economic historians (Abramovitz, 1986; David, 2001; Gerschenkron, 1962; Kuznets, 1971). On the contrary, theories pointing out different stages of growth (Rostow, 1960) may be more consistent with I(0) models featuring deterministic trend shifts.

2016; Hausmann et al., 2005, 2006).[5] Nevertheless, in both cases, the choice of a level vis-à-vis first difference specification is often not informed by prior evidence from unit root tests. Moreover, the economic filters adopted generally reflect time-invariant and deterministic characteristics which are not suited to capturing the stochastic nature of structural shifts observed in integrated models.

Stemming from Nelson and Plosser (1982), researchers have started to pay attention to the possible presence of stochastic trends in the data. This interest was originally motivated by the fact that in I(1) type processes the distinction between secular movements and business cycles becomes blurred as the trend component itself displays fluctuations. Nevertheless, when a time series exhibits a unit root, it is equally complicated to distinguish growth episodes occurring at medium run frequencies from the secular stochastic trend. The focus of this paper is on this second question, i.e. understanding whether the observed episodes are the result of segmented deterministic trends and level shifts or of stochastic forces affecting the secular component.

Following Perron (1989), it is now a widespread practice to incorporate structural breaks in unit root tests. It was shown that omitting dummies for structural change in Dickey-Fuller regressions would result in a failure to reject the unit root null hypothesis (Perron, 1989). Also, drawing on Zivot and Andrews (1992) and Christiano (1992) these tests now feature a data-dependent algorithm to determine the location of the structural shifts under the alternative hypothesis. However, a major drawback of such an approach concerns the assumption of a fixed number of breaks, typically determined ex ante.[6] This creates a gap with the empirical literature in which data-driven procedures are used, not only to identify break dates, but also to select the number of relevant structural changes. To deal with this issue, Kapetanios (2005) presents a test of the unit root hypothesis against I(0) alternatives with an unspecified number of breaks (up to some exogenously given maximum). The test, nevertheless, features a search algorithm based on the minimization of t-statistics which has been shown to perform poorly in identifying the correct number of shifts and their dates (Lee and Strazicich, 2001; Vogelsang and Perron, 1998).

Few empirical applications of unit root tests have been concerned with countries long-run growth paths. A range of contributions investigates the presence of unit roots in historical time series of real GDP per capita, generally for a few advanced countries.[7] Ben-David and Papell (1995) apply the test of Zivot and Andrews (1992) allowing for a break in both the trend and the constant in a sample of OECD countries and reject the unit root hypothesis for 7 out of 16 series. In a follow up paper, Ben-David et al. (2003) show that by incorporating an additional break it is possible to reject the null for 12 out 16 countries. Extending the previous works, Papell and Prodan (2014) consider various models with different break forms in a sample of 19 OECD countries and 7 Asian economies. Their results report respectively 15 rejections for the OECD group and 6 rejections for

[5] A compromise between formal testing and subjective criteria is found in Kar et al. (2013) and Pritchett et al. (2016). They propose a "fit and filter" methodology in which potential breaks are estimated looking at the best econometric fit and, secondly, relevant ones are selected according to economic filters.

[6] Another shortcoming is related to the fact that structural breaks are allowed only under the alternative hypothesis. Although we do not address directly this problem, we provide a discussion in Section 2. Recently, various tests have been put forward which rely on a GLS detrending procedure similar to that presented by Elliott et al. (1996). These new tests investigate unit roots in the noise function of a series and have the advantage of incorporating breaks under both the null and the alternative hypothesis (Carrion-i Silvestre et al., 2009; Harris et al., 2009; Harvey et al., 2013; Narayan and Popp, 2010, 2013).

[7] A skeptic point of view on this line of research is provided by Gaffeo et al. (2005). The authors run different unit roots tests and find substantial heterogeneity in the results depending on the type of procedure adopted. They interpret this evidence as questioning the possibility to characterize income per capita series with a sufficiently invariant statistical model.

the Asian one. An alternative framework is proposed by Kejriwal and Lopez (2013). They present an econometric procedure that uses in a sequential manner various tests allowing for up to two structural breaks under both the null and the alternative. With such methodological amendments, in contrast to the results of the previous literature, their approach indicates no evidence against the unit root hypothesis.

This paper contributes to this literature by introducing a methodology for testing the null of a generic I(1) process versus a set of stationary alternatives (with structural breaks) and by presenting an empirical application for long run time series of per capita GDP. Methodologically, we build upon Kapetanios (2005) and add novel features along three dimensions: (i) we treat the number of breaks (not only their location) as unknown; (ii) we exploit the sequential approach by Bai (1997) to extend the number of breaks to four and, consequently, we include in the analysis also a group of developing countries with more volatile series; (iii) we implement a robust search algorithm that resembles the practices for the identification of growth episodes adopted in the empirical literature.

Our results suggest less support in favour of trend stationarity than in previous contributions. In a sample of 34 countries we find 17 rejections. Interestingly, developing countries exhibit only four rejections, thus, showing a more complex and unstable dynamics. Moreover, even less evidence against the unit root hypothesis is found when we relax the assumption of Gaussian innovations by using bootstrapped critical values. More generally, the evidence presented in this paper complements the skeptical results of Kejriwal and Lopez (2013). This points to the general conclusion according to which the dismissal of the unit root hypothesis in GDP series may be premature. In particular, more attention should be devoted to investigating the role of the various search algorithms implemented in unit root tests, as well as of the assumptions on the functional form of the shocks, in driving the evidence against integrated models.

The reminder of this work proceeds as follows: Section 2 describes the methodology; Section 3 shows some Monte Carlo experiments to assess power and size properties of the test in finite samples; Section 4 presents the empirical strategy while in Section 5 we introduce and discuss the results; Section 6 concludes.

2 Methodology

Following Zivot and Andrews (1992) we consider the following null hypothesis:

$$y_t = \mu + y_{t-1} + \Psi^*(L)v_t, \tag{1}$$

where: $\Psi^*(L) = A^*(L)^{-1}B(L)$; $A^*(L)$ and $B(L)$ are lag polynomials respectively of order p and q with all the roots outside the unit circle and v is a zero-mean sequence of iid random variables.

The alternative model considered takes the form:

$$y_t = \mu + \beta t + \Psi(L)\big[\sum_{i=1}^{m}\theta_i DU(T_i)_t + \sum_{i=1}^{m}\gamma_i DT(T_i)_t + v_t\big] \tag{2}$$

where: $\Psi(L) = A(L)^{-1}B(L)$; $A(L) = (1-\alpha L)A^*(L)$. The intercept and trend break dummies are $DU(T_i)_t = \mathbb{1}(t > T_i)$ and $DT(T_i)_t = \mathbb{1}(t > T_i)(t - T_i)$ with $\mathbb{1}$ being the indicator function and T_i a generic break date. Notice that, according to the so-called *innovation outlier* specification, changes in the trend or in the constant evolve as any other shock. For instance, while the immediate impact of a generic variation in the constant is θ_i, the corresponding long-run effect will be $\Psi(1)\theta_i$.

Both the null and the alternative model can be nested in a general DF-type of regression:

$$y_t = \mu + \beta t + \alpha y_{t-1} + \sum_{i=1}^{m} \theta_i DU(T_i)_t + \sum_{i=1}^{m} \gamma_i DT(T_i)_t + \sum_{j=1}^{k} c_j \Delta y_{t-j} + \epsilon_t \tag{3}$$

In our analysis, the number of breaks (m), the lag-truncation parameter (k) and the break dates ($T_1, ..., T_m$) are treated as unknown. Therefore, for a given number of breaks m, the null and the alternative hypothesis are defined as:

$$H_0 : \alpha = 1, \, \theta_i = \gamma_i = 0 \quad \forall \, i \in [1, m]$$

$$H_1 : \alpha < 1$$

Let us now focus on some methodological considerations. First, we are using the most general model that includes for each break both the intercept and the trend shift dummy. As discussed by Sen (2003), when the form of the breaks is unknown, the preferred strategy is to adopt a general specification allowing for changing intercept and trend in order to minimize power distortions.[8] Second, structural breaks are allowed only under the alternative hypothesis whereas the null model is described by an I(1) process without exogenous shifts in its deterministic components. Such asymmetric treatment of breaks characterizes several unit root tests proposed in the literature (Banerjee et al., 1992; Lumsdaine and Papell, 1997; Perron, 1997; Zivot and Andrews, 1992). However, Vogelsang and Perron (1998) and Lee and Strazicich (2001) show that size distortions arise when structural breaks are present under the null as a result of the nuisance parameter associated witho the trend function. Although it has been pointed out that serious distortions only emerge in the presence of large shifts and may not be particularly relevant in practice (Perron et al., 2006; Vogelsang and Perron, 1998), several works have directly addressed the issue (Carrion-i Silvestre et al., 2009; Harris et al., 2009; Harvey et al., 2013; Lee and Strazicich, 2003; Narayan and Popp, 2010, 2013). A common strategy is to change the specification of the Data Generating Process (DGP) to allow for a unit root only in the noise component, thus, avoiding the problem of nuisance parameters. The new DGP adopted is as follow:

$$y_t = \mu + \beta t + \sum_{i=1}^{m} \theta_i DU(T_i)_t + \sum_{i=1}^{m} \gamma_i DT(T_i)_t + e_t, \tag{4}$$

$$e_t = \alpha e_{t-1} + v_t \tag{5}$$

where v_t is an unobserved mean-zero process. A GLS detrending procedure is implemented and then the unit root test is performed on the estimated series of the noise term e.[9] Although such literature provides interesting results, in this paper we take the original approach of Zivot and Andrews (1992) and restrain from introducing exogenous shifts under the null model. The motivation underlying such a methodological choice is twofold. On the one hand, we want to investigate the presence of a unit root in y rather than in the error component, that implies models (cf. Equations 1 and 2) for which the structural break dummies have to be estimated simultaneously with the other parameters,

[8]Power simulations in Section 2 corroborate the results of Sen (2003). When adopting the mixed model (including both intercept and trend dummies) the test does not show dramatic power losses.

[9]The work of Narayan and Popp (2010) and Narayan and Popp (2013) start from the same DGP but follow a different procedure. Instead of detrending the series they fit a modified version of Equation 3 that includes lagged mean-shift dummies and impulse dummies. The break locations are then estimated by maximizing the Wald statistic for the joint significance of impulse dummies. The test, however, does not consider the possibility of pure trend breaks.

i.e. prior detrending is not feasible.[10] On the other hand, from an economic point of view, we want to test the radical hypothesis according to which growth episodes are generated by frequent stochastic events rather than by few exogenous structural changes. Hence, the results from the test proposed here have to be interpreted in a conservative way since rejections may occur when the data follows an integrated process with breaks. As will be reported, despite the evidence in favor of I(1) models tends to be negatively biased, our results still suggest fewer rejections than in previous works.

As for other tests in the literature, we implement a data-driven procedure to estimate the break locations. Nevertheless, in tune with the suggestion of Lumsdaine and Papell (1997), rather than narrowly consider a specific number of breaks we take an agnostic perspective and test for an unspecified number of structural changes, up to some maximum M. As stressed by Kejriwal and Lopez (2013), it is desirable to select the model with the appropriate number of breaks before proceeding with the unit root test as the imposition of extraneous dummy variables leads to considerable power losses. In this respect, the paper provides a first step in incorporating in the unit root test a methodology for the identification of structural shifts that is broadly in tune with the one actually used by practitioners in the field of growth empirics when looking for growth episodes (Berg et al., 2012; Jones and Olken, 2008; Kar et al., 2013; Kerekes, 2007). Along these lines, Kejriwal and Lopez (2013) adopt a sequential testing procedure to jointly identify the presence of structural breaks and unit roots in the error function, allowing for a maximum of two shifts of the *additive outlier* type. Their results for the historical series of GDP per capita are generally favorable to models with I(1) noise components. In this work we provide complementary evidence by testing the unit root hypothesis in the series itself, rather than in the noise function, and by introducing multiple breaks (possibly more than two, of the *innovation outlier* kind) under the null. Increasing the number of structural changes with respect to the existing literature is particularly desirable in order to include developing countries which are typically characterized by more unstable growth paths. Indeed, as shown by Perron (1989) for the case of one break and generalized by Ohara (1999), the omission of one or more structural changes will result in a failure to reject the null for models that are trend stationary in the segments between breaks.

The search algorithm used to choose m and $(T_1, ..., T_m)$ extends the approach by Kapetanios (2005) being grounded on the sequential (one-by-one) breaks estimation proposed by Bai (1997). As shown by simulations exercises (Lee and Strazicich, 2001; Vogelsang and Perron, 1998) the standard practice of locating breaks by minimizing the t-statistic for α generally leads to inconsistent estimates and, therefore, approaches based on minimization of the squared sum of residuals have to be preferred. Estimating breaks one at a time also has the advantage of being significantly computationally less expensive as compared to the grid search scheme by Bai and Perron (2003). However, the one-by-one procedure leads to limiting distributions of locations that tend to diverge from the ones obtained via simultaneous estimation. To guard from such a problem we implement in a second-step the repartition procedure suggested by Bai (1997).The algorithm can be described by the following steps:

- **Step 1. Sequential estimation:** For each $m \in [1, M]$ and holding $k = K$ fixed, where M and K refer to exogenously determined upper bounds respectively for the number of breaks and the truncation-lag parameter, obtain the break locations sequentially by minimizing the sum of squared residuals from Equation 3 conditional on past breaks estimation. Thus, a

[10] To the best of our knowledge, there are not yet statistical tests that could allow for a break under the null in the framework of Equation 1.

generic break date is estimated as:

$$\hat{T}_m = \text{argmin}_{T_m} S(\hat{T}_1, ..., \hat{T}_{m-1}, T_m), \tag{6}$$

where:

$$S(\hat{T}_1, ..., \hat{T}_{m-1}, T_m) = \sum_{t=k+2}^{T} \Bigg(y_t - \hat{\mu} - \hat{\beta} t - \hat{\alpha} y_{t-1} - \sum_{i=1}^{m-1} \hat{\theta}_i DU(\hat{T}_i)_t - \hat{\theta}_m DU(T_m) \\ - \sum_{i=1}^{m-1} \hat{\gamma}_i DT(\hat{T}_i)_t - \hat{\gamma}_m DT(T_m)_t - \sum_{j=1}^{K} \hat{c}_j \Delta y_{t-j} \Bigg)^2 \tag{7}$$

- **Step 2. Repartition procedure:** For each $m \in [2, M]$ and the associated partition $(\hat{T}_1, ..., \hat{T}_m)$, each break date is re-estimated by fitting a one-shift model in the data interval defined by $[\hat{T}_{i-1}; \hat{T}_{i+1}]$.[11] The new estimates $(T_1^*, ..., T_m^*)$ are consistent and share the same asymptotic distributions of those obtained by global maximization.[12] Notice that the whole search scheme is carried out imposing a trimming parameter h, expressed as a share of sample size, to ensure a minimum length for each segment between breaks.

- **Step3. Model selection:** As we are left with $M + 1$ possible partitions (including also the case with no breaks), the model with the appropriate number of breaks (m^*) is chosen using the BIC criteria. The truncation-lag parameter k^* is then selected using the general-to-specific approach advocated by Ng and Perron (1995), i.e. starting from the upper bound (K) we remove one lag at the time until the last lag in an autoregression of order k^* is significant while the last lag in an autoregression of order $k^* + 1$ is not significant.

Concerning model selection, different approaches have been proposed in the econometric literature. Kapetanios (2005) proposes to select the optimal partition by minimizing the t-statistic for α. As for selecting breaks locations, such an approach is unlikely to deliver satisfactory results since the imposition of more dummies will generally overestimate the true number of shifts.[13] The recommended strategy by Bai and Perron (2003) is to test for the presence of an additional shift in all the segments between break dates.[14] This $supF(l|l + 1)$ test allows one to discriminate between l and $l + 1$ breaks, and when used sequentially can be used to choose the model with the correct number of structural changes. Simulation evidence in Bai and Perron (2006) shows that both the sequential procedure and the BIC criteria perform better than other approaches. The former has the advantage of taking into account heterogeneity across segments and of being robust when serial correlation is present. Nevertheless, the sequential testing method presents serious power losses in small samples as it is typically carried out with ever less observations (Antoshin et al., 2008). Therefore, for this

[11] Notice that for $i = 1$, $\hat{T}_{i-1} = 1$ and for $i = m$, $\hat{T}_{i+1} = T$.

[12] Although asymptotic distributions are identical, they may diverge in finite samples. As a robustness check, we carried out simulations using also the simultaneous approach of Bai and Perron (2003) for a T equal to the average of our sample. Results are not considerably different and, therefore, we decided to opt for the repartition procedure. Simulation evidence on break location in finite samples is reported in Section 3.

[13] The reason is that I(1) can be seen as a limiting case of a I(0) process with several breaks, i.e. a I(0) process in which both the trend and the constant change permanently at any point in time. Hence imposing additional dummies leads to more evidence against the alternative and, accordingly, to a lower t-statistic. For a detailed discussion of the issue see Perron (1989). Simulation evidence in Section 3 corroborates such conclusion.

[14] The test is equivalent to the maximization of the Wald statistic $(F - test)$ over all the data points in a specific segment.

T	Statistic	$M=2$				$M=3$				$M=4$			
		1%	2.5%	5%	10%	1%	2.5%	5%	10%	1%	2.5%	5%	10%
							$\boldsymbol{h=0.05}$						
100	t_α	-7.15	-6.82	-6.51	-6.18	-8.18	-7.84	-7.50	-7.15	-8.99	-8.59	-8.27	-7.88
	F_T	14.13	12.19	11.02	10.00	13.35	11.96	11.01	10.06	12.83	11.81	10.96	10.15
150	t_α	-7.02	-6.65	-6.38	-6.09	-8.00	-7.67	-7.36	-7.03	-8.84	-8.46	-8.14	-7.81
	F_T	13.94	12.05	10.79	9.73	12.97	11.43	10.53	9.69	12.50	11.19	10.43	9.70
200	t_α	-6.91	-6.62	-6.37	-6.06	-7.91	-7.55	-7.29	-6.99	-8.74	-8.37	-8.11	-7.78
	F_T	13.73	11.86	10.68	9.76	12.64	11.23	10.37	9.64	12.14	11.00	10.24	9.59
							$\boldsymbol{h=0.1}$						
100	t_α	-7.27	-6.91	-6.59	-6.26	-8.09	-7.74	-7.41	-7.06	-8.67	-8.27	-7.98	-7.64
	F_T	14.53	12.44	11.32	10.30	13.54	12.23	11.30	10.36	13.43	12.19	11.29	10.38
150	t_α	-7.11	-6.80	-6.51	-6.19	-7.96	-7.66	-7.39	-7.05	-8.60	-8.26	-7.97	-7.61
	F_T	13.86	11.94	10.95	9.97	12.91	11.48	10.69	9.91	13.22	11.45	10.61	9.86
200	t_α	-7.06	-6.72	-6.48	-6.17	-7.91	-7.60	-7.34	-7.01	-8.59	-8.24	-7.88	-7.46
	F_T	13.60	12.08	10.96	9.93	13.52	11.61	10.66	9.82	12.72	11.39	10.57	9.76

Table 1: Finite sample critical values for t_α and F_T.

specific application, the BIC criteria appears to be more suited.[15] A general issue with the BIC criteria concerns its poor performance under the null (i.e. when breaks are not present) when serial correlation is not accounted for. In our case, however, such a problem is addressed by directly controlling for serial correlation via the inclusion of k lags in the regression.

Finally, having selected $(T_1^*, ..., T_m^*)$, m^* and k^*, we fit the corresponding regression and use as test statistics both the standard t-statistic (t_α) for the null of $\alpha = 1$ and the Wald statistic (F_T) for the joint null: $\alpha = 1$; $\theta_1 = ... = \theta_{m^*} = \gamma_1 = ... = \gamma_{m^*} = 0$.[16]

3 Finite sample size, power and break selection properties

In this section we present the critical values and explore the finite sample size and power properties of the test.[17] Table 1 reports finite sample critical values for different M, h and T.[18] Following Kapetanios (2005), critical values are obtained by approximating the distributions of t_α and F_T under the null via Monte Carlo simulations of standard random walks (10,000 replications).

We then present simulation results to investigate size and power properties of the test. The experimental design follows that of Vogelsang and Perron (1998) and Sen (2003). The simulated

[15]In this regard, we run some Monte Carlo exercises comparing the two approaches. Simulations results show the superiority of the BIC criteria, given the specificities of our application. We also found that the sequential procedure displays further power losses when, as in our case, the form of the breaks is not known a priori.

[16]The properties of F_T (in the case of one break) are largely explored in Sen (2003). Here we generalize to the case of multiple breaks. Thus, the statistic can be computed as: $F_T = \frac{(S_{UR}-S_R)/(1+2m^*)}{(1-S_{UR})/(T-3-2m^*-k^*)}$, where S_{UR} and S_R are for the sum of squared residuals respectively of the unrestricted and the restricted model.

[17]Considering the specific application of this paper, in which the average sample size of the GDP series is 164, we are only interested in the finite sample performance of the test. Accordingly, only finite sample critical values are derived.

[18]In deriving critical values the upper bound (K) for the lag truncation parameter is set to 7 for $h = 0.1$ and to 2 for $h = 0.5$. Results using other values are available upon request from the authors.

A	B	Break coefficients	Autoregressive coefficient (α)											
			$=1$		$=0.9$		$=0.8$		$=0.7$		$=0.6$		$=0.5$	
			t_α	F_T	t_α	F_T	t_α	F_T	t_α	F_T	t_α	F_T	t_α	F_T
		$\rho = 0; \lambda = 0$												
0	0	-	0.054	0.050	0.108	0.188	0.173	0.419	0.247	0.609	0.329	0.709	0.480	0.778
2	0	θ=(2,2)	-	-	0.026	0.436	0.149	0.420	0.359	0.565	0.569	0.675	0.665	0.731
2	0	θ=(4,4)	-	-	0.429	0.999	0.847	0.999	0.926	0.999	0.962	1.000	0.971	1.000
4	0	θ=(2,2,2,2)	-	-	0.027	0.261	0.235	0.239	0.475	0.282	0.636	0.398	0.709	0.524
4	0	θ=(4,4,4,4)	-	-	0.277	0.998	0.946	0.998	0.992	0.998	0.997	0.998	0.997	0.999
2	2	θ=(2,2); γ=(0.025, 0.025)	-	-	0.025	0.787	0.133	0.768	0.335	0.850	0.544	0.899	0.667	0.918
2	2	θ=(4,4); γ=(0.05, 0.05)	-	-	0.235	1.000	0.796	1.000	0.923	1.000	0.962	1.000	0.975	1.000
4	4	θ=(2,2,2,2); γ=(0.025, 0.025, 0.025, 0.025)	-	-	0.015	0.471	0.158	0.425	0.414	0.510	0.610	0.639	0.701	0.729
4	4	θ=(4,4,4,4); γ=(0.05, 0.05, 0.05, 0.05)	-	-	0.115	1.000	0.892	1.000	0.978	1.000	0.995	1.000	0.996	1.000
		$\rho = 0.5; \lambda = 0$												
0	0	-	0.044	0.057	0.147	0.398	0.272	0.683	0.463	0.802	0.584	0.843	0.641	0.890
2	0	θ=(2,2)	-	-	0.179	0.455	0.549	0.676	0.713	0.773	0.778	0.818	0.795	0.816
2	0	θ=(4,4)	-	-	0.897	0.999	0.973	0.999	0.982	0.999	0.987	1.000	0.986	1.000
4	0	θ=(2,2,2,2)	-	-	0.295	0.294	0.652	0.452	0.739	0.603	0.780	0.668	0.815	0.682
4	0	θ=(4,4,4,4)	-	-	0.957	0.998	0.998	1.000	1.000	1.000	1.000	1.000	0.999	1.000
2	2	θ=(2,2); γ=(0.025, 0.025)	-	-	0.151	0.802	0.524	0.899	0.696	0.929	0.763	0.946	0.785	0.956
2	2	θ=(4,4); γ=(0.05, 0.05)	-	-	0.864	1.000	0.971	1.000	0.978	1.000	0.983	1.000	0.985	1.000
4	4	θ=(2,2,2,2); γ=(0.025, 0.025, 0.025, 0.025)	-	-	0.221	0.484	0.614	0.677	0.740	0.780	0.792	0.803	0.823	0.823
4	4	θ=(4,4,4,4); γ=(0.05, 0.05, 0.05, 0.05)	-	-	0.924	1.000	0.992	1.000	1.000	1.000	0.999	1.000	0.999	1.000
		$\rho = -0.5; \lambda = 0$												
0	0	-	0.045	0.051	0.085	0.119	0.125	0.265	0.162	0.424	0.198	0.572	0.238	0.665
2	0	θ=(2,2)	-	-	0.007	0.530	0.047	0.397	0.141	0.414	0.237	0.459	0.339	0.540
2	0	θ=(4,4)	-	-	0.093	0.998	0.613	1.000	0.828	1.000	0.899	1.000	0.936	1.000
4	0	θ=(2,2,2,2)	-	-	0.004	0.253	0.050	0.209	0.176	0.214	0.336	0.245	0.481	0.262
4	0	θ=(4,4,4,4)	-	-	0.024	0.990	0.636	0.998	0.947	0.999	0.985	0.999	0.993	0.999
2	2	θ=(2,2); γ=(0.025, 0.025)	-	-	0.018	0.838	0.026	0.749	0.104	0.746	0.217	0.803	0.331	0.860
2	2	θ=(4,4); γ=(0.05, 0.05)	-	-	0.017	1.000	0.448	1.000	0.776	1.000	0.869	1.000	0.923	1.000
4	4	θ=(2,2,2,2); γ=(0.025, 0.025, 0.025, 0.025)	-	-	0.009	0.458	0.033	0.432	0.129	0.418	0.272	0.429	0.421	0.500
4	4	θ=(4,4,4,4); γ=(0.05, 0.05, 0.05, 0.05)	-	-	0.005	0.999	0.436	1.000	0.889	1.000	0.961	1.000	0.988	1.000
		$\rho = 0; \lambda = 0.5$												
0	0	-	0.057	0.074	0.100	0.177	0.156	0.371	0.199	0.556	0.245	0.645	0.295	0.716
2	0	θ=(2,2)	-	-	0.026	0.402	0.121	0.412	0.261	0.490	0.375	0.576	0.483	0.632
2	0	θ=(4,4)	-	-	0.413	0.998	0.838	0.999	0.914	0.999	0.946	0.999	0.955	0.999
4	0	θ=(2,2,2,2)	-	-	0.026	0.224	0.200	0.260	0.387	0.272	0.505	0.334	0.587	0.401
4	0	θ=(4,4,4,4)	-	-	0.232	0.993	0.935	0.997	0.984	0.999	0.997	1.000	0.999	1.000
2	2	θ=(2,2); γ=(0.025, 0.025)	-	-	0.025	0.768	0.104	0.770	0.242	0.807	0.366	0.852	0.455	0.884
2	2	θ=(4,4); γ=(0.05, 0.05)	-	-	0.208	1.000	0.775	1.000	0.894	1.000	0.933	1.000	0.951	1.000
4	4	θ=(2,2,2,2); γ=(0.025, 0.025, 0.025, 0.025)	-	-	0.018	0.384	0.141	0.427	0.315	0.478	0.443	0.547	0.537	0.602
4	4	θ=(4,4,4,4); γ=(0.05, 0.05, 0.05, 0.05)	-	-	0.100	0.998	0.856	1.000	0.970	1.000	0.987	1.000	0.996	1.000
		$\rho = 0; \lambda = -0.5$												
0	0	-	0.204	0.205	0.276	0.408	0.396	0.648	0.497	0.757	0.596	0.841	0.656	0.888
2	0	θ=(2,2)	-	-	0.085	0.398	0.374	0.573	0.623	0.722	0.748	0.792	0.797	0.825
2	0	θ=(4,4)	-	-	0.415	0.999	0.883	0.999	0.959	0.999	0.972	0.999	0.983	0.999
4	0	θ=(2,2,2,2)	-	-	0.052	0.186	0.403	0.330	0.649	0.555	0.764	0.654	0.828	0.690
4	0	θ=(4,4,4,4)	-	-	0.364	0.998	0.982	0.998	0.997	0.998	0.997	0.999	0.999	0.999
2	2	θ=(2,2); γ=(0.025, 0.025)	-	-	0.089	0.771	0.341	0.867	0.644	0.921	0.748	0.939	0.803	0.965
2	2	θ=(4,4); γ=(0.05, 0.05)	-	-	0.244	1.000	0.855	1.000	0.958	1.000	0.982	1.000	0.985	1.000
4	4	θ=(2,2,2,2); γ=(0.025, 0.025, 0.025, 0.025)	-	-	0.058	0.379	0.376	0.531	0.649	0.710	0.763	0.797	0.827	0.836
4	4	θ=(4,4,4,4); γ=(0.05, 0.05, 0.05, 0.05)	-	-	0.204	0.999	0.947	1.000	0.994	1.000	0.997	1.000	0.997	1.000

Table 2: Size and power results under different parametrizations

model takes the general form:

$$[1-(\alpha+\rho)L+\rho L^2]y_t = (1+\lambda L)[\sum_{i=1}^{A}\theta_i DU(T_i)_t + \sum_{i=1}^{B}\gamma_i DT(T_i)_t + e_t], \tag{8}$$

where $e_t \sim \mathcal{N}(0,1)$. For each experiment we run 1000 replications of length $T = 200$ and report the rejection rate at the 5% level using the appropriate critical values for $M = 4$ and $h = 0.1$. The following combinations of ρ and λ are tested: $\{(0,0);(0.5,0);(-0.5,0);(0,0.5);(0,-0.5)\}$. In the size simulations we impose $\alpha = 1$ and $A = B = 0$, while for the power simulations we experiment for $\alpha \in \{0.9;0.8;0.7;0.6;0.5\}$ introducing different number of breaks of different forms and magnitudes.[19] Results are reported in Table 2. Let us now emphasize some key features emerging from simulations:

1. The size of t_α and F_T is reasonably close to the nominal value. A well-known exception is the case with a negative moving average component in which both the test statistics are slightly over-sized.

2. In the absence of breaks, F_T displays uniformly higher power than t_α across all the experiments.

3. When the number of structural changes increases, some loss in power has to be expected, *ceteris paribus*, as a result of the introduction of additional dummies (see Kapetanios, 2005, for a discussion of this issue).

4. Convergence to 100% power occurs fast as the magnitude of the breaks increases. As documented by Sen (2003), F_T converges faster than the standard t-statistic since it incorporates information on the presence of breaks.

5. The power generally increases monotonically as we move away from the null (i.e. as α decreases). Nevertheless, in the presence of a negative autoregressive term, the power of F_T may slightly decrease between $\alpha = 0.9$ and $\alpha = 0.6$.

6. For $\alpha = 0.9$, F_T has a higher power than t_α in almost all the experiments, i.e. it is better suited to investigate cases with the autoregressive parameter close to unity.

Let us now compare the power performance of t_α and F_T with the Kapetanios test. The latter provides a natural benchmark for comparison as it generalizes the framework of Zivot and Andrews (1992) to the case of an unknown number of breaks. Results are reported in Figure 1 for $M \in \{2;3;4\}$ and different parameter values of the simulated model.[20] Some important aspects stand out from the simulations. First, as pointed out by Ohara (1999), the power of all the statistics falls dramatically when M is lower than the true number of breaks. Second, for t_α and the Kapetanios test statistic a less pronounced reduction in power also appears when increasing M, given the number of true breaks. The performance of F_T, on the contrary, remains largely unaffected by this second effect. Most importantly, the Kapetanios test exhibits higher power than both t_α and F_T in only some marginal instances when no breaks are present under the null. Generally, F_T tend to outperform the other statistics, especially when the the upper bound M increases (cf. the panels with $M = 4$ in Figure 1).

In Figures 2 and 3 we also report a comparison with the Kapetanios test concerning the ability to correctly identify the number of breaks and their locations.[21] Figure 2 assumes the number of breaks

[19] In all the experiments we assume break locations to be symmetrically distributed across the time span.
[20] In Figure 1, with "small breaks" we refer to $(\theta,\gamma) = (2, 0.025)$, while "large breaks" stands for $(\theta,\gamma) = (4, 0.05)$.
[21] In both Figures 2 and 3 we assume the size of each break to be $(\theta,\gamma) = (2, 0.025)$.

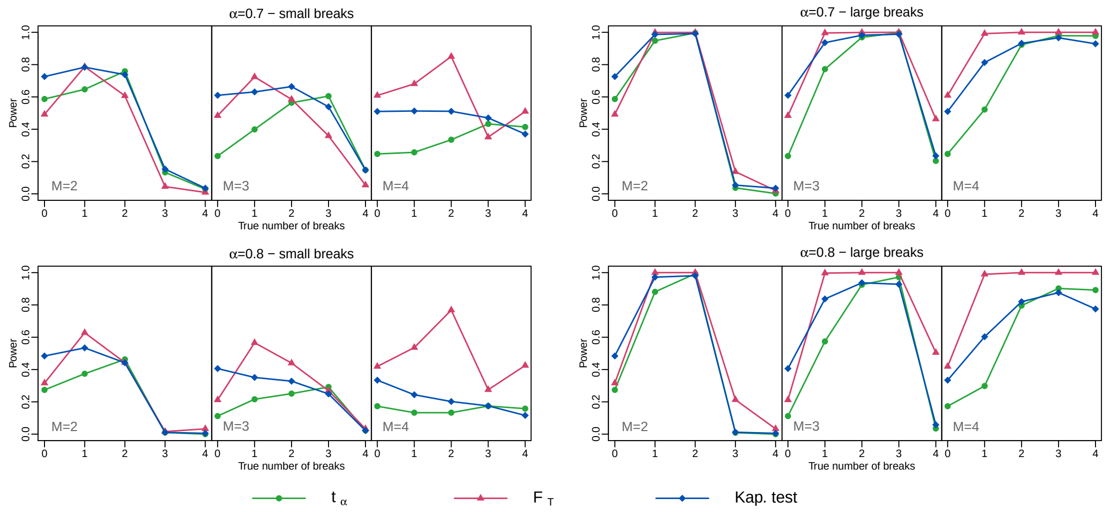

Figure 1: General power comparison - F_T, t_α, the Kapetanios statistic (min-t_α)

to be known (equal to 4) and plots the distributions of the estimated break dates under different degrees of serial correlation. It contrasts the performance of our approach based on the minimization of the sum of squared residuals in two steps (sequential and repartition) with the standard one based on the minimization of the t-statistic. While in the former case the distribution of break locations are symmetric and centered around the true dates, in the latter the distributions tend to display long left tails, in particular for the first two breaks. Figure 3, instead, compares the outcomes of the criteria used for selecting the number of structural shifts. Once again, the minimization of the t-statistic performs poorly as it tend to always select a number of breaks equal to the upper bound M, resulting in a general overestimation. In this respect, the minimization of the BIC criteria provides substantial improvements and the probability to choose the appropriate number of breaks increases significantly.

Overall, according to our simulation exercises, the testing procedure proposed here turns out to yield gains in terms of both power performance and the precision of breaks estimation. In particular, one should expect a generally higher power for F_T than for t_α.[22] However, since the Wald statistic may exhibit non-monotonic power in the few specific instances described above, in the empirical application we will also report results using t_α.

4 The empirical strategy

We investigate the presence of unit roots and structural breaks in income per capita series. Table 3 summarizes the results from previous studies. Data are taken from the last release of the Maddison database (Bolt et al., 2018).[23] To preserve the robustness of our analysis we focus exclusively on time series with at least 100 consecutive observations, that is, we are left with a sample of 34 countries (20 OECD and 14 developing).

Concerning the choice of M, as documented in Section 2, a parsimonious specification of M

[22] This is in tune with the evidence reported by Sen (2003) for the case of a single break suggesting generally higher power of the Wald statistic.

[23] More precisely, we use the variable *RGDPNApc* based on a single price benchmark (1990 US dollars).

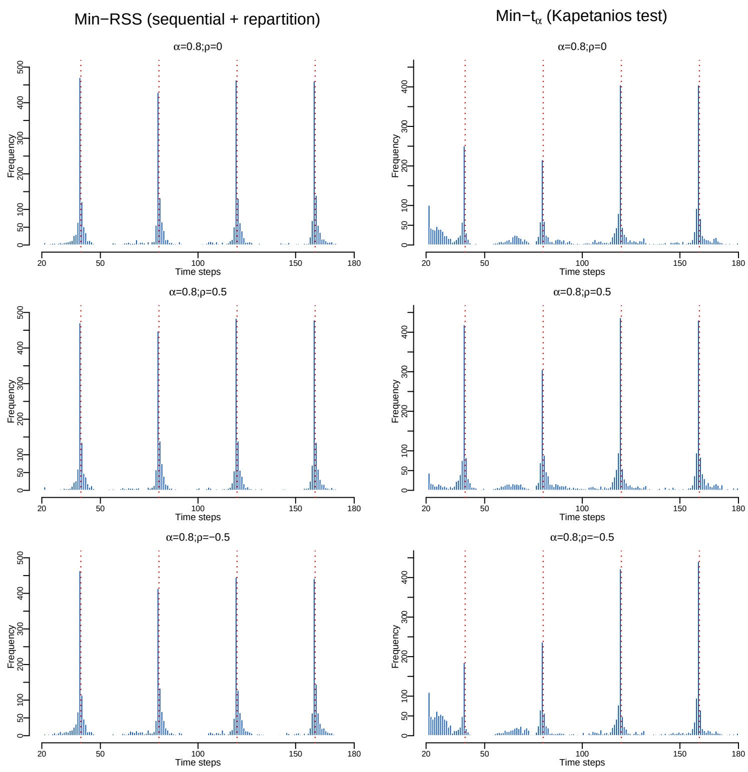

Figure 2: Distribution of break dates - Two step minimization of SSR (left panels) vis-à-vis sequential minimization of t_α (right panels)

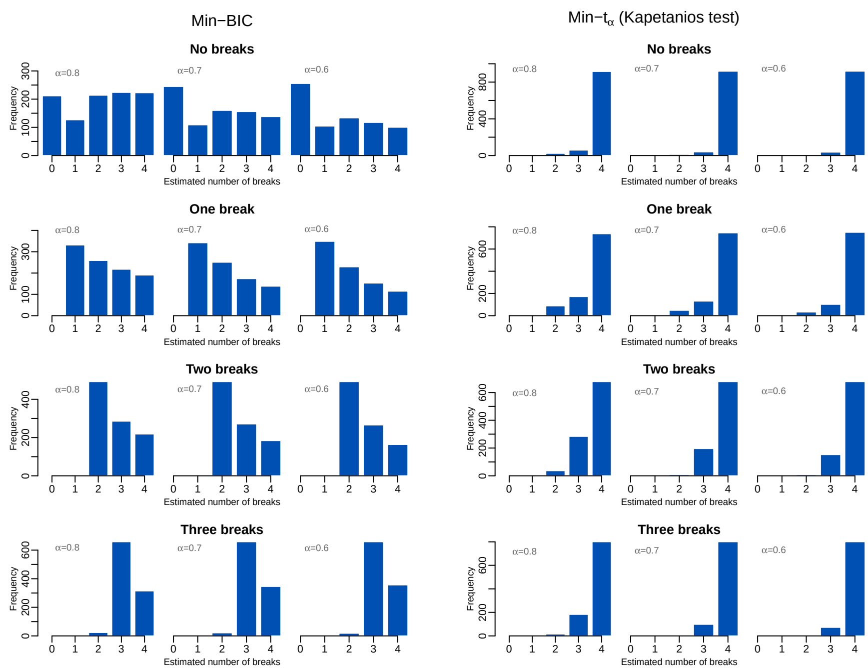

Figure 3: Frequency of selected number of breaks - BIC criteria (left panels) vis-à-vis sequential minimiziation of t_α (right panels)

may improve the power of the test when the true number of shifts is less than or equal to the specified maximum. However, large power losses exist when the number of breaks is greater than the maximum allowed. Therefore, we report results for both $M = 3$ and $M = 4$ and find no evidence of power losses.[24] Indeed, it is reassuring that the number of rejections does not fall when allowing for an additional break. In all the tests the trimming parameter is set to $h = 0.1$.

Differently from previous studies which rely on asymptotic critical values (Kejriwal and Lopez, 2013; Lumsdaine and Papell, 1997; Papell and Prodan, 2014), for each country in the sample (and each M) we derive finite-sample critical values to take into account the specific characteristic of different time series. The key intuition is that under the null the first differences of the series can be described by a stationary $ARMA$ process. Following Christiano (1992) and Zivot and Andrews (1992), for each series we take first differences and estimate a battery of $ARMA(p, q)$ models. To determine the appropriate number of lags p and q we use the BIC criteria. The distribution of both t_α and F_T as well as the associated critical values are then approximated via Monte Carlo simulations (with 5000 replications) of the selected model. Consistently with the power simulations reported in in Section 3, we find a higher number of rejections when using the Wald statistic. In the simulations we allow for two alternative assumptions regarding the nature of the stochastic disturbances: (i) Normal shocks with zero mean and standard deviation estimated from the residuals; (ii) Randomly drawn shocks (with replacement) from the distribution of residuals. Hence, critical values are computed both assuming the Gaussianity of the shocks and via bootstrapping (cf. Table B.1 in the Appendix B). The latter technique has the advantage of restraining from parametric assumptions but may lead to spurious results in small samples, in particular when the criteria used for model selection fail to identify serial correlation in the error term. As a consequence, results are reported for both approaches in Table 4. Figure 4 shows the dynamics of GDP per capita and the estimated break dates. We also performed some robustness checks. First, we ran the test assuming a fixed number of structural changes in order to identify possible power losses arising in the selection of the appropriate number of breaks. Results are reported in Appendix A (cf. Table A.1). Although showing general consistency with the baseline case, they indicate even less rejections, thus, excluding the possibility for our results to be driven by power losses due to the selection procedure adopted. As a second robustness check, we run the test imposing a smaller trimming parameter ($h = 0.5$, cf. Table A.2) in order to allow for more consecutive break dates. This results in three extra rejections for OECD countries while the coefficient of New Zealand loses its significance. Hence, allowing for shorter growth segments only provides little additional evidence against the unit root hypothesis.

5 Discussion of results

For OECD countries, our methodology rejects the null of the unit root in only 13 of 20 instances under the assumption of Gaussian shocks. Although our results do not contrast strongly with the previous literature (cf. Table 3), we find additional failures to reject the null (i.e. Canada, Denmark, Sweden, Switzerland). These differences reflect the different break search methodology adopted and, possibly, the use of series-specific critical values vis-à-vis asymptotic ones. Somewhat consistently with Kejriwal and Lopez (2013), relying on the minimization of the SSR rather than the t-statistics

[24] In tune with the discussion in Kejriwal and Lopez (2013), allowing for a greater number of breaks is not desirable given the available sample sizes (which range from 111 to 197 observations).

produces less evidence against the unit root hypothesis.[25]

This paper also presents new evidence for developing countries. In particular, we find only 4 rejections in a sample of 14 economies. Intuitively, those countries tend to experience more erratic growth processes with persistent and frequent (possibly more than four) shifts in both level and trend. This is in line with several contributions emphasizing the ubiquitous presence of growth discontinuities in poor- and middle-income countries (Hausmann et al., 2005; Lamperti and Mattei, 2018; Pritchett et al., 2000).

Another relevant contribution of our work regards the possibility of departing from the assumption of Gaussian shocks by deriving bootstrapped critical values. In other terms, a key question addressed here concerns how results from standard unit root tests may change when being agnostic about the functional form of the innovations. Rejection levels using bootstrapped critical values are reported in brackets in Table 4. Interestingly, this leads to considerably less evidence against the unit root hypothesis. In Figure 5, the empirical distribution of the residuals under the null is contrasted with the best Normal fit. Departures from Normality appear to exist in some countries in terms of skewness and, most importantly, excess kurtosis.[26] This seems to suggest that the assumption of Gaussianity, typically adopted in standard testing procedures, may bias the results in favour of trend-stationary models. One may conjecture, instead, that GDP time series may be well described by I(1) models with fat-tailed innovations. Such a characterization is consistent with empirical findings which identify Laplacian distributions of aggregate growth shocks (Fagiolo et al., 2008). Fat-tailed distributions of shocks entail a growth process driven by large and lumpy events. They typically emerge when some of the assumptions of the central limit theorem are violated. In particular, it has been pointed out that the presence of dynamic increasing returns and strong correlating mechanisms (e.g. competition, network externalities) at the firm level may lead to non-trivial aggregation of microeconomic shocks, thus, being responsible for the emergence of fat tails in macroeconomic data (Bottazzi and Secchi, 2006; Dosi et al., 2007; Fagiolo et al., 2008). An I(1) characterization of the GDP per capita series with non-Gaussian innovations is common to many evolutionary growth models.[27] These models generally describe the growth process as a result of complex interactions across individuals and organizations which, in turn, lead to path dependency and irreversibility of shocks as well as to the emergence of fat-tailed distributions at all the levels of aggregation. The lack of evidence against I(1) processes may be interpreted as pointing towards strong degrees of "complexity" and inter-relatedness across economic units, thus, providing support to evolutionary models. For instance, Dosi et al. (2019) present a multi-country agent-based model in which firms interact both domestically and in international markets following idiosyncratic learning trajectories. Simulation results show that countries endogenously differentiate and cluster into two groups of winners and losers exhibiting extremely erratic paths with fat-tailed distributions of growth rates.

[25] Zerbo and Darné (2018) apply the methodology in Kejriwal and Lopez (2013) to the GDP per capita series of 28 sub-Saharan African countries for the period 1960-2014. Although their results may suffer from small sample bias, they also find no evidence against the unit root hypothesis.

[26] Investigating the precise functional form of the growth residuals goes beyond the scope of this work and would probably require longer time series. Here it suffices to mention that visual inspection of the density seems to suggest the presence of fat-tails in several countries. In Section 6 we sketch some future lines of research including the use of quantile regression to dig into the extreme quantiles of the distribution.

[27] See for instance early evolutionary growth models Dosi et al. (1994); Nelson and Plosser (1982); Silverberg and Verspagen (1995); Verspagen (1992). For some agent-based evolutionary models see Ciarli et al. (2010), Dosi et al. (2010), Dawid et al. (2014), Caiani et al. (2016), Lorentz et al. (2016), Ciarli et al. (2017), Caiani et al. (2018), Dawid et al. (2018), Dosi et al. (2019).

Break dates are estimated under the I(0) alternative and, therefore, they have a meaningful interpretation when the unit root null is rejected. Nevertheless, it should be kept in mind that break locations for all countries tend to capture major historical events such as wars, booms and crisis. In this respect, the endogenous identification of relevant episodes provides a further validation of the search algorithm proposed here. Moreover, consistent with previous contributions, there is no evidence of a single steady state model as each country displays at least one structural break. In Table 5 we report estimates of break dummy coefficients for the series which appear to be stationary. Most countries with I(0) time series tend to exhibit significant changes in both their intercepts and trends. As an illustrative example consider the case of France whose experience is representative of those of many OECD countries. Our break selection procedure suggests two major crashes associated with the two world wars which are both accompanied by subsequent periods of growth acceleration. The phase of strong catching up in the aftermath of World War two is then followed by a period of relative stagnation (i.e. a negative trend shift) at the end of the 1970s. The presence of (relatively few) changes in growth rates within-country, possibly associated also to level shifts, is a feature of endogenous growth models exhibiting "strong" scale effects. Less evidence is found supporting pure Neoclassical and semi-endogenous models which predict only level effects. This is broadly consistent with the results of Papell and Prodan (2014), who find growth effects in the majority of the time series considered.[28]

The evidence presented here has some relevant implications for applied work in the field of growth empirics. First, the presence of unit roots in many GDP series affects significantly the identification of specific kinds of growth episodes. Several empirical papers disregard prior unit root testing when looking for structural changes in the data. The choice of a level vis-à-vis first-difference specification is however crucial for the appropriate implementation of structural breaks search procedures. Our results indicate that for most GDP time series, especially in developing countries, the first-difference variant has to be preferred. Moreover, they call into question the widespread practice of using simple economic filters, based on invariant criteria (e.g. a jump in growth rates of a given amount lasting for some years), to identify growth shifts. In fact, the evidence in favour of I(1) models hints at extremely frequent growth discontinuities which hardly obey deterministic and recurrent characteristics.

6 Conclusion

In this paper we developed a methodology to test for the unit root hypothesis in long-run income time series. Our approach extends the test in Kapetanios (2005) by introducing a more robust search procedure which provides substantial improvements in terms of power and breaks identification (cf. the evidence in Section 3).

As argued in Section 1 discerning I(1) models from stationary alternatives has relevant theoretical and empirical implications in the field of economic growth. The tension between integrated and trend stationary models (with breaks) can be summarized by the following question: how frequently do countries experience structural breaks in their GDP per capita series? In the limit, unit root models are stationary processes in which both the intercept and the trend change permanently at any point in time. Hence, if structural breaks occur particularly often, the distinction between I(1) and I(0)

[28] More mixed evidence is presented by Sobreira et al. (2014). Using structural breaks tests robust to the presence of unit roots they find that countries distribute quite uniformly across the "constant trend", "level shifts" or "trend shifts" hypothesis.

Country	Ben-David et al. (2003)		Kejriwal and Lopez (2013)		Papell and Prodan (2014)	
	Break dates	Rej. lev.	Break dates	Rej. lev.	Break dates	Rej. lev
OECD						
Australia	1891, 1927	10%	1891, 1929	-	1931	10%
Austria	1944, 1959	1%	1913, 1944	-	1944, 1950, 1976	1%
Belgium	1916,1939	5%	1917, 1939	-	1939, 1976	1%
Canada	1908, 1928	1%	-	-	1930, 1940	5%
Denmark	1939, 1975	1%	1914, 1939	-	1939, 1969	1%
Finland	1916, 1943	1%	1917	-	-	-
France	1939, 1974	1%	1917, 1945	-	1939, 1973	1%
Germany	-	-	1922, 1945	-	1944, 1950	1%
Italy	-	-	1918, 1944	-	1942, 1948	5%
Japan	1944, 1973	1%	1944, 1973	-	1944, 1971, 1991	1%
Netherlands	-	-	1918, 1945	-	1945, 1951	1%
New Zealand			1907, 1935	-	-	-
Norway	1917, 1939	1%	1921	-	-	-
Portugal			1936	-	-	-
Spain			1937	-	1935, 1959, 1971	5%
Sweden	1916, 1963	1%	1917	-	1915, 1970	5%
Switzerland	-	-	1916, 1944	-	1944	5%
UK	1918, 1945	5%	1919	-	1939, 1945	5%
USA	1929, 1945	1%	1931, 1945	-	1929, 1942	1%
Asia						
India					-	-
Indonesia					1941	10%
Malaysia					1944	10%
Philippines					1946, 1952	5%
Taiwan					1942	1%
South Korea					1944	1%
Sri Lanka					1900, 1966	10%

Notes: Blank spaces denote countries not included in the study while ‘ - ’ indicates the failure to reject at the 10% confidence level. Papell and Prodan (2014) only report break dates obtained from structural break tests for stationary series. The dates may not coincide with those emerging from unit root tests.

Table 3: Rejection rates and break dates from other studies using Maddison data

Country	T	M=3			M=4		
		Breaks	t_α	F_T	Breaks	t_α	F_T
OECD							
Australia	197	1851, 1891, 1928	$-8.647^{***}_{(***)}$	$12.876^{**}_{(**)}$	1851, 1891, 1928	$-8.647^{***}_{(**)}$	$12.876^{**}_{(**)}$
Austria	147	1913, 1944, 1959	-10.495^{***}	19.158^{***}	1913, 1944, 1959, 1979	-14.118^{***}	29.757^{***}
Belgium	171	1918, 1943	1.135	$16.188^{***}_{(*)}$	1918, 1943	1.135	$16.188^{***}_{(*)}$
Canada	147	1908, 1933, 1988	-6.501	8.829	1908, 1933, 1988	-6.501	8.829
Denmark	197	1914, 1939, 1963	-5.174	7.624	1850, 1916, 1939, 1968	-7.269	7.712
Finland	157	1916, 1968	-5.586	8.697	1916, 1932, 1953, 1978	-7.103	7.676
France	197	1916, 1939, 1975	$-11.503^{***}_{(***)}$	$20.189^{***}_{(***)}$	1916, 1939, 1975	$-11.503^{***}_{(***)}$	$20.189^{***}_{(***)}$
Germany	167	1913, 1945, 1970	$-11.852^{***}_{(*)}$	21.63^{***}	1913, 1945, 1963, 1981	$-13.513^{***}_{(*)}$	22.993^{***}
Greece	184	1912, 1939, 1978	$-9.583^{***}_{(***)}$	$14.622^{***}_{(**)}$	1912, 1939, 1978	$-9.583^{***}_{(**)}$	$14.622^{***}_{(**)}$
Italy	197	1945	-2.964	$20.474^{***}_{(**)}$	1945	-2.964	$20.474^{***}_{(**)}$
Japan	147	1944, 1971	-10.444^{***}	35.444^{***}	1944, 1971	-10.444^{***}	35.444^{***}
Netherlands	197	1922, 1943, 1963	-8.019^{***}	11.552^{**}	1922, 1943, 1963	-8.019^{*}	11.552^{**}
New Zealand	147	1885, 1907, 1935	-4.6	7.865	1895, 1910, 1935, 1976	-8.003^{*}	10.068
Norway	187	1944, 1995	-3.938	9.697	1944, 1995	-3.938	9.697
Portugal	152	1934, 1973	-0.589	8.213	1934, 1973	-0.589	8.213
Spain	167	1935, 1960, 1999	-7.137^{*}	10.773^{*}	1935, 1960, 1999	-7.137	10.773^{**}
Sweden	197	1894, 1939, 1970	-5.236	5.685	1894, 1939, 1970	-5.236	5.685
Switzerland	166	1913, 1940, 1968	-7.31	9.186	1913, 1940, 1968	-7.31	9.186
UK	197	1918, 1944, 1996	$-9.184^{***}_{(***)}$	$13.664^{**}_{(**)}$	1875, 1918, 1944, 1996	$-8.787^{***}_{(**)}$	$11.911^{**}_{(**)}$
USA	197	1878, 1929, 1949	$-7.82^{**}_{(**)}$	$10.393^{*}_{(*)}$	1878, 1929, 1949, 1996	$-8.872^{***}_{(**)}$	$10.315^{*}_{(*)}$
Asia							
India	133	1944, 1964	-0.867	7.75	1944, 1964	-0.867	7.75
Taiwan	116	1945	4.029	19.866^{***}	1945	4.029	19.866^{***}
Sri Lanka	147	1904, 1965, 1992	-6.442	7.762	1900, 1943, 1975, 2000	-7.244	8.285
Latin America							
Argentina	142	1891, 1929, 1984	-7.025	8.2	1891, 1929, 1984	-7.025	8.2
Bolivia	127	1955, 1971, 1998	-6.248	6.945	1930, 1952, 1978, 2000	-6.069	7.21
Brazil	167	1891, 1928, 1970	-6.612	9.35ì	1891, 1929, 1962, 1980	-6.788	8.509
Chile	197	1918, 1981	-7.579^{**}	$11.971^{**}_{(*)}$	1918, 1981	-7.579	$11.971^{**}_{(*)}$
Colombia	147	1906	-3.182	8.93	1906	-3.182	8.93
Ecuador	117	1945, 1972, 1998	-6.505	7.792	1945, 1972, 1998	-6.505	7.792
Mexico	122	1930, 1981	-4.94	8.273	1915, 1930, 1942, 1981	-4.265	6.539
Panama	111	1929, 1945, 1987	-6.417	7.45	1929, 1945, 1981, 2003	-6.472	7.484
Peru	197	1876, 1987	-7.636^{**}	13.435^{**}	1876, 1987	-7.636^{*}	$13.435^{***}_{(**)}$
Uruguay	147	1913	-4.851	8.229	1897, 1913, 1948, 1981	-6.922	6.396
Venezuela	187	1895, 1944	-5.095	$10.627^{*}_{(*)}$	1895, 1944	-5.095	$10.627^{**}_{(*)}$

Notes: Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Significance levels in brackets refer to rejections using bootstrapped critical values. The trimming parameter is set to $h = 0.1$ and the lag-truncation parameter to $K = 7$.

Table 4: Results from the unit root tests and estimated break dates for $M = 3$ and $M = 4$

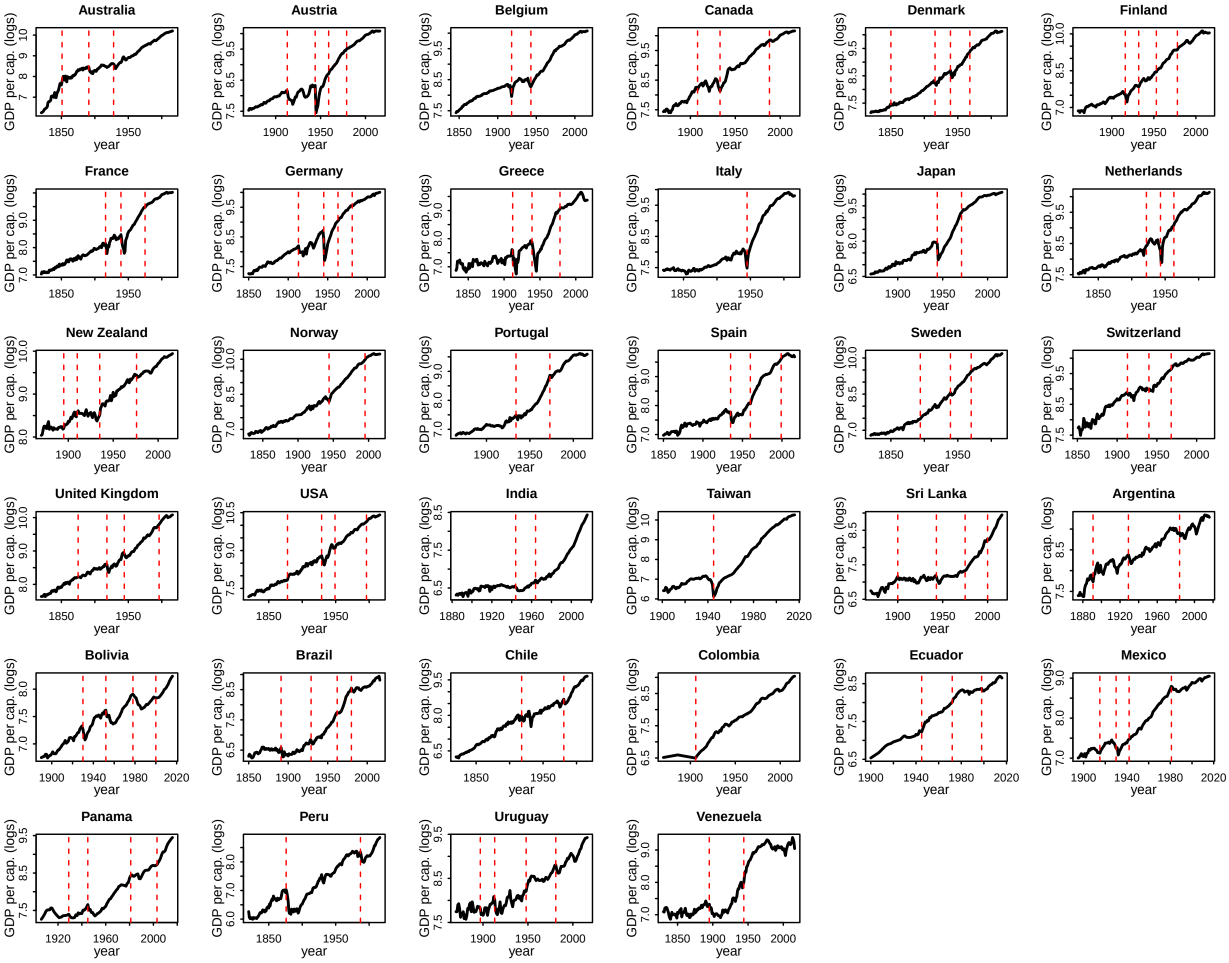

Figure 4: Time series of income per capita and estimated break dates

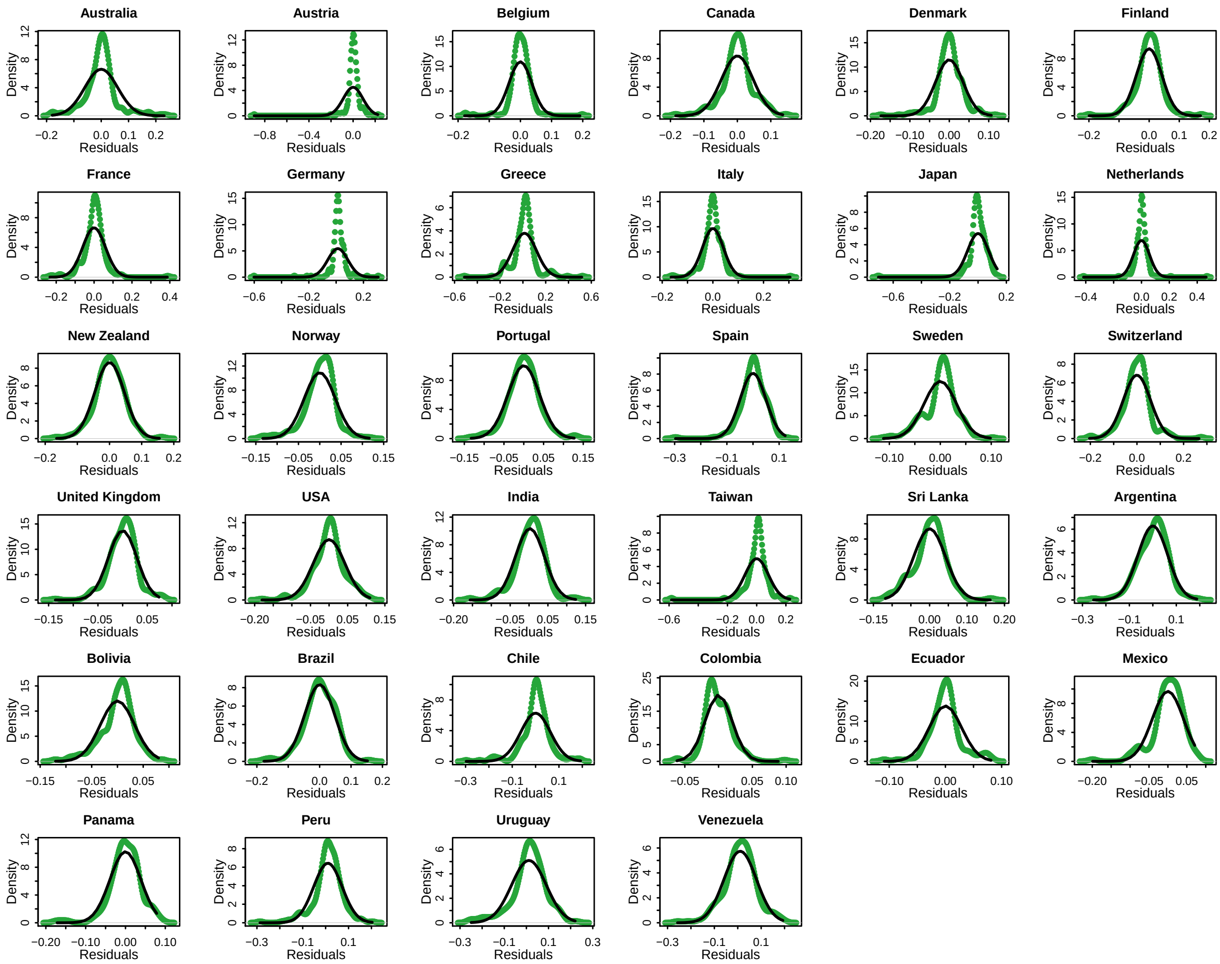

Figure 5: Residuals from the selected I(1) model - empirical density (in green) vs. Normal fit

Country	DU_1 .	DT_1	DU_2	DT_2	DU_3	DT_3	DU_4	DT_4
OECD								
Australia	0.0965^{***}	-0.0177^{***}	-0.1761^{***}	-0.0046^{***}	-0.1039^{***}	0.0059^{***}		
	(0.0301)	(0.0024)	(0.0286)	(0.0013)	(0.0223)	(0.0011)		
Austria	-0.2826^{***}	-0.0027^{*}	-0.5855^{***}	0.0566^{***}	-0.0166	-0.0292^{***}	0.006	-0.0218^{***}
	(0.0302)	(0.0014)	(0.0462)	(0.0045)	(0.0391)	(0.0044)	(0.0292)	(0.0027)
Belgium	0.1154^{***}	-0.0052^{***}	0.1237^{***}	0.0049^{***}				
	(0.0163)	(0.0000)	(0.0173)	(0.001)				
France	-0.0551^{**}	0.006^{***}	-0.3241^{***}	0.0131^{***}	0.0078	-0.0175^{***}		
	(0.022)	(0.0015)	(0.0349)	(0.0019)	(0.0215)	(0.0016)		
Germany	-0.2329^{***}	0.0071^{***}	-0.6129^{***}	0.038^{***}	-0.0359	-0.0318^{***}	-0.0067	-0.0148^{***}
	(0.0271)	(0.0012)	(0.0521)	(0.0035)	(0.0354)	(0.0039)	(0.0284)	(0.0027)
Greece	-0.1709^{***}	0.0147^{***}	-0.4532^{***}	0.0069^{***}	-0.0052	-0.0179^{***}		
	(0.0403)	(0.0025)	(0.0577)	(0.0025)	(0.0397)	(0.0023)		
Italy	0.0975^{***}	0.0000						
	(0.0129)	(0.0000)						
Japan	-0.4978^{***}	0.0376^{***}	0.0198	-0.0394^{***}				
	(0.0435)	(0.0029)	(0.0343)	(0.0031)				
Netherlands	-0.148^{***}	0.0309^{***}	0.0303	-0.0168^{***}	0.1533^{***}	-0.0075^{***}		
	(0.0438)	(0.0043)	(0.0257)	(0.0029)	(0.0289)	(0.0019)		
Spain	-0.1398^{***}	0.0071^{***}	0.0932^{***}	0.0000	0.0299	-0.0121^{***}		
	(0.0239)	(0.0013)	(0.0236)	(0.0015)	(0.0246)	(0.0025)		
United Kingdom	-0.0194^{*}	-0.0012^{***}	-0.118^{***}	0.0058^{***}	-0.0757^{***}	-0.0000	0.036^{***}	-0.0047^{***}
	(0.01)	(0.0000)	(0.0153)	(0.0000)	(0.0128)	(0.0000)	(0.0131)	(0.001)
USA	0.0333^{**}	0.003^{***}	-0.1517^{***}	0.0099^{***}	-0.0042	-0.0067^{***}	0.0252	-0.0056^{***}
	(0.0143)	(0.0000)	(0.0236)	(0.0018)	(0.0217)	(0.0017)	(0.0194)	(0.0015)
Asia								
Taiwan	0.4096^{***}	-0.0071^{**}						
	(0.065)	(0.0028)						
Latin America								
Chile	-0.0937^{***}	-0.0024^{***}	-0.066^{**}	0.0103^{***}				
	(0.0209)	(0.0000)	(0.0255)	(0.0017)				
Peru	-0.1917^{***}	0.0000	-0.1664^{***}	0.0036^{***}				
	(0.0285)	(0.0000)	(0.0279)	(0.0012)				
Venezuela	-0.1348^{***}	0.0039^{***}	0.1279^{***}	-0.0057^{***}				
	(0.0288)	(0.0000)	(0.0352)	(0.0000)				

Notes: Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Only countries for which it is possible to reject the unit root hypothesis (10% significance or lower) are included.

Table 5: Estimates of structural break dummies

specifications becomes extremely blurred. In this perspective, testing for unit roots amounts to testing for the frequency of structural changes. The procedure introduced in this paper has the aim of distinguishing between models with several permanent changes in mean and trend and alternatives with relatively few variations. Our results are more favorable to the first alternative.

Even in advanced countries we find less evidence against I(1) processes in comparison to previous studies that tend to find a relatively large number of rejections (Ben-David et al., 2003; Papell and Prodan, 2014), with our results being more in line with new results pointing at a resurgence of the unit root hypothesis in GDP data (Kejriwal and Lopez, 2013; Zerbo and Darné, 2018). Another contribution of this paper is the inclusion of developing countries in the analysis. However, even by allowing for up to four breaks, we fail to reject the null of a unit root in most of the countries considered. Such results suggest the presence of strong growth discontinuities in backward economies which make their growth paths hardly distinguishable from a random walk. Finally, the number of rejections fall when using bootstrapped critical values instead of Gaussian shocks, possibly hinting at the presence of I(1) models with fat-tailed innovations.

In Section 5, such results have been interpreted as providing support to evolutionary growth models which stress path dependency, nonlinearities and the non-trivial aggregation of microeconomic shocks. At the macroeconomic level, these characteristics typically lead to the emergence of series exhibiting several growth shifts, similar to I(1) models.

From the point of view of growth empirics, we emphasize the importance of unit root testing prior to (or jointly with) structural break identification. Indeed, if countries exhibit growth trajectories similar to random walks, the practice of fitting structural change models on the series in levels may lead to spurious and inconsistent results.

Our results also suggest some future lines of investigation. First, it becomes crucial to move towards testing methodologies that are robust to the presence of fat-tailed shocks. Quantile autoregressions (QAR) are a natural candidate in this respect, as they allow for the investigation of persistence properties of a time series at different quantiles of the conditional distribution (Koenker and Xiao, 2004, 2006). Recently, structural break tests have been developed in the framework of QAR (Oka and Qu, 2011; Qu, 2008). Incorporating unit root tests in this setting would clearly be a key achievement. Second, there is a lot to learn from the growth dynamics of developing countries. The unstable and complex patterns shown by this group of economies call for further research efforts. As a matter of fact, most empirical papers investigating growth episodes in less developed countries tend to adopt a deterministic characterization of growth discontinuities, relying on constant and recurrent criteria (e.g. 2% acceleration in growth rates for a minimum number of years) to define episodes. The evidence presented here partially challenges this approach since we have shown that for developing countries, growth shifts are extremely frequent and exhibit random characteristics in terms of form and magnitude. Unfortunately, long run time series are available only for a limited sample of economies while both unit root and structural break tests suffer from finite sample biases. As a first attempt to address the issue, Antoshin et al. (2008) present a methodology for structural break testing suited for short time series. More generally, improving the small sample performance of unit root tests would allow one to perform a similar investigation using post-war data for a larger set of economies.

Acknowledgments

We thank the participants to the internal LEM seminar in Pisa (18/12/2018, Scuola Superiore Sant'Anna); the PEGDECH seminar in Groningen (10/04/2019, University of Groningen); the WEHIA 2019 conference in London. All usual disclaimers apply. Emanuele Russo acknowledges the support by the European Union's Horizon 2020 research and innovation program under grant agreement No. 822781 - GROWINPRO.

References

Abramovitz, M. (1986). Catching up, forging ahead, and falling behind. *The Journal of Economic History*, 46(02):385–406.

Aghion, P. and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica: Journal of the Econometric Society*, pages 323–351.

Aizenman, J. and Spiegel, M. M. (2010). Takeoffs. *Review of Development Economics*, 14(2):177–196.

Antoshin, S., Souto, M., and Berg, A. (2008). *Testing for structural breaks in small samples*. Number 8-75. International Monetary Fund.

Bai, J. (1997). Estimating multiple breaks one at a time. *Econometric theory*, 13(03):315–352.

Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of applied econometrics*, 18(1):1–22.

Bai, J. and Perron, P. (2006). Multiple structural change models: a simulation analysis. *Econometric theory and practice: Frontiers of analysis and applied research*, pages 212–237.

Banerjee, A., Lumsdaine, R. L., and Stock, J. H. (1992). Recursive and sequential tests of the unit-root and trend-break hypotheses: theory and international evidence. *Journal of Business & Economic Statistics*, 10(3):271–287.

Ben-David, D., Lumsdaine, R. L., and Papell, D. H. (2003). Unit roots, postwar slowdowns and long-run growth: evidence from two structural breaks. *Empirical Economics*, 28(2):303–319.

Ben-David, D. and Papell, D. H. (1995). The great wars, the great crash, and steady state growth: Some new evidence about an old stylized fact. *Journal of Monetary Economics*, 36(3):453–475.

Berg, A., Ostry, J. D., and Zettelmeyer, J. (2012). What makes growth sustained? *Journal of Development Economics*, 98(2):149–166.

Bluhm, R., de Crombrugghe, D., and Szirmai, A. (2016). The dynamics of stagnation: A panel analysis of the onset and continuation of stagnation. *Macroeconomic Dynamics*, 20(8):2010–2045.

Bolt, J., Inklaar, R., de Jong, H., and van Zanden, J. L. (2018). Rebasing 'maddison': new income comparisons and the shape of long-run economic development. *GGDC Research Memorandum*, 174.

Bottazzi, G. and Secchi, A. (2006). Explaining the distribution of firm growth rates. *The RAND Journal of Economics*, 37(2):235–256.

Caiani, A., Catullo, E., and Gallegati, M. (2018). The effects of fiscal targets in a monetary union: a multi-country agent-based stock flow consistent model. *Industrial and Corporate Change*, 27(6):1123–1154.

Caiani, A., Godin, A., Caverzasi, E., Gallegati, M., Kinsella, S., and Stiglitz, J. E. (2016). Agent based-stock flow consistent macroeconomics: Towards a benchmark model. *Journal of Economic Dynamics and Control*, 69:375–408.

Carrion-i Silvestre, J. L., Kim, D., and Perron, P. (2009). Gls-based unit root tests with multiple structural breaks under both the null and the alternative hypotheses. *Econometric theory*, 25(6):1754–1792.

Castellacci, F. (2007). Evolutionary and new growth theories. are they converging? *Journal of Economic Surveys*, 21(3):585–627.

Christiano, L. J. (1992). Searching for a break in gnp. *Journal of Business & Economic Statistics*, 10(3):237–250.

Ciarli, T., Lorentz, A., Savona, M., and Valente, M. (2010). The effect of consumption and production structure on growth and distribution. a micro to macro model. *Metroeconomica*, 61(1):180–218.

Ciarli, T., Lorentz, A., Valente, M., and Savona, M. (2017). Structural changes and growth regimes. *Journal of Evolutionary Economics*, pages 1–58.

David, P. A. (2001). Path dependence, its critics and the quest for 'historical economics'. In *Evolution and path dependence in economic ideas: Past and present*, pages 15–40. Cheltenham.

Dawid, H., Harting, P., and Neugart, M. (2014). Economic convergence: Policy implications from a heterogeneous agent model. *Journal of Economic Dynamics and Control*, 44:54–80.

Dawid, H., Harting, P., and Neugart, M. (2018). Cohesion policy and inequality dynamics: Insights from a heterogeneous agents macroeconomic model. *Journal of Economic Behavior & Organization*, 150:220–255.

Dosi, G. et al. (2007). Statistical regularities in the evolution of industries. a guide through some evidence and challenges for the theory. *Perspectives on innovation*, pages 1110–1121.

Dosi, G., Fabiani, S., Aversi, R., and Meacci, M. (1994). The dynamics of international differentiation: a multi-country evolutionary model. *Industrial and corporate change*, 3(1):225–242.

Dosi, G., Fagiolo, G., Napoletano, M., and Roventini, A. (2013). Income distribution, credit and fiscal policies in an agent-based keynesian model. *Journal of Economic Dynamics and Control*, 37(8):1598–1625.

Dosi, G., Fagiolo, G., Napoletano, M., Roventini, A., and Treibich, T. (2015). Fiscal and monetary policies in complex evolving economies. *Journal of Economic Dynamics and Control*, 52:166–189.

Dosi, G., Fagiolo, G., and Roventini, A. (2010). Schumpeter meeting keynes: A policy-friendly model of endogenous growth and business cycles. *Journal of Economic Dynamics and Control*, 34(9):1748–1767.

Dosi, G., Pereira, M. C., Roventini, A., and Virgillito, M. E. (2017). When more flexibility yields more fragility: The microfoundations of keynesian aggregate unemployment. *Journal of Economic Dynamics and Control*.

Dosi, G., Roventini, A., and Russo, E. (2019). Endogenous growth and global divergence in a multi-country agent-based model. *Journal of Economic Dynamics and Control*.

Easterly, W., Kremer, M., Pritchett, L., and Summers, L. H. (1993). Good policy or good luck? *Journal of Monetary Economics*, 32(3):459–483.

Elliott, G., Rothenberg, T. J., Stock, J. H., et al. (1996). Efficient tests for an autoregressive unit root. *Econometrica*, 64(4):813–836.

Fagiolo, G., Napoletano, M., and Roventini, A. (2008). Are output growth-rate distributions fat-tailed? some evidence from oecd countries. *Journal of Applied Econometrics*, 23(5):639–669.

Gaffeo, E., Gallegati, M., and Gallegati, M. (2005). Requiem for the unit root in per capita real gdp? additional evidence from historical data. *Empirical Economics*, 30(1):37–63.

Gerschenkron, A. (1962). Economic backwardness in historical perspective: a book of essays. Technical report, Belknap Press of Harvard University Press Cambridge, MA.

Grossman, G. M. and Helpman, E. (1991). Quality ladders in the theory of growth. *The Review of economic studies*, 58(1):43–61.

Harris, D., Harvey, D. I., Leybourne, S. J., and Taylor, A. R. (2009). Testing for a unit root in the presence of a possible break in trend. *Econometric Theory*, 25(6):1545–1588.

Harvey, D. I., Leybourne, S. J., and Taylor, A. R. (2013). Testing for unit roots in the possible presence of multiple trend breaks using minimum dickey–fuller statistics. *Journal of Econometrics*, 177(2):265–284.

Hausmann, R., Pritchett, L., and Rodrik, D. (2005). Growth accelerations. *Journal of economic growth*, 10(4):303–329.

Hausmann, R., Rodriguez, F. R., and Wagner, R. A. (2006). Growth collapses.

Jones, B. F. and Olken, B. A. (2008). The anatomy of start-stop growth. *The Review of Economics and Statistics*, 90(3):582–587.

Jones, C. I. (1995). R&d-based models of economic growth. *Journal of political Economy*, 103(4):759–784.

Jones, C. I. (2005). Growth and ideas. In *Handbook of economic growth*, volume 1, pages 1063–1111. Elsevier.

Kapetanios, G. (2005). Unit-root testing against the alternative hypothesis of up to m structural breaks. *Journal of Time Series Analysis*, 26(1):123–133.

Kar, S., Pritchett, L., Raihan, S., and Sen, K. (2013). Looking for a break: Identifying transitions in growth regimes. *Journal of Macroeconomics*, 38:151–166.

Kejriwal, M. and Lopez, C. (2013). Unit roots, level shifts, and trend breaks in per capita output: A robust evaluation. *Econometric Reviews*, 32(8):892–927.

Kerekes, M. (2007). *Analyzing patterns of economic growth: a production frontier approach*. Number 2007/15. Diskussionsbeiträge des Fachbereichs Wirtschaftswissenschaft der Freien Universität Berlin.

Koenker, R. and Xiao, Z. (2004). Unit root quantile autoregression inference. *Journal of the American Statistical Association*, 99(467):775–787.

Koenker, R. and Xiao, Z. (2006). Quantile autoregression. *Journal of the American Statistical Association*, 101(475):980–990.

Kuznets, S. S. (1971). *Economic growth of nations*. Mass., Belknap Press of Harvard University Press.

Lamperti, F., Dosi, G., Napoletano, M., Roventini, A., and Sapio, A. (2018). Faraway, so close: coupled climate and economic dynamics in an agent-based integrated assessment model. *Ecological Economics*, 150:315–339.

Lamperti, F. and Mattei, C. E. (2018). Going up and down: Rethinking the empirics of growth in the developing and newly industrialized world. *Journal of Evolutionary Economics*, 28(4):749–784.

Lee, J. and Strazicich, M. C. (2001). Break point estimation and spurious rejections with endogenous unit root tests. *Oxford Bulletin of Economics and statistics*, 63(5):535–558.

Lee, J. and Strazicich, M. C. (2003). Minimum lagrange multiplier unit root test with two structural breaks. *Review of Economics and Statistics*, 85(4):1082–1089.

Lee, K., Pesaran, M. H., and Smith, R. (1997). Growth and convergence in a multi-country empirical stochastic solow model. *Journal of applied Econometrics*, 12(4):357–392.

Lorentz, A., Ciarli, T., Savona, M., and Valente, M. (2016). The effect of demand-driven structural transformations on growth and technological change. *Journal of Evolutionary Economics*, 26(1):219–246.

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of monetary economics*, 22(1):3–42.

Lumsdaine, R. L. and Papell, D. H. (1997). Multiple trend breaks and the unit-root hypothesis. *Review of economics and Statistics*, 79(2):212–218.

Narayan, P. K. and Popp, S. (2010). A new unit root test with two structural breaks in level and slope at unknown time. *Journal of Applied Statistics*, 37(9):1425–1438.

Narayan, P. K. and Popp, S. (2013). Size and power properties of structural break unit root tests. *Applied Economics*, 45(6):721–728.

Nelson, C. R. and Plosser, C. R. (1982). Trends and random walks in macroeconmic time series: some evidence and implications. *Journal of monetary economics*, 10(2):139–162.

Nelson, R. and Winter, S. (1982). *An evolutionary theory of economic change.* Cambridge [etc.]: Harvard University Press.

Ng, S. and Perron, P. (1995). Unit root tests in arma models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association*, 90(429):268–281.

Ohara, H. I. (1999). A unit root test with multiple trend breaks: A theory and an application to us and japanese macroeconomic time-series. *The Japanese Economic Review*, 50(3):266–290.

Oka, T. and Qu, Z. (2011). Estimating structural changes in regression quantiles. *Journal of Econometrics*, 162(2):248–267.

Papell, D. H. and Prodan, R. (2014). Long run time series tests of constant steady-state growth. *Economic Modelling*, 42:464–474.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica: Journal of the Econometric Society*, pages 1361–1401.

Perron, P. (1997). Further evidence on breaking trend functions in macroeconomic variables. *Journal of econometrics*, 80(2):355–385.

Perron, P. et al. (2006). Dealing with structural breaks. *Palgrave handbook of econometrics*, 1(2):278–352.

Pritchett, L. et al. (2000). Understanding patterns of economic growth: searching for hills among plateaus, mountains, and plains. *World Bank Economic Review*, 14(2):221–250.

Pritchett, L., Sen, K., Kar, S., and Raihan, S. (2016). Trillions gained and lost: Estimating the magnitude of growth episodes. *Economic Modelling*, 55:279–291.

Qu, Z. (2008). Testing for structural change in regression quantiles. *Journal of Econometrics*, 146(1):170–184.

Romer, P. M. (1986). Increasing returns and long-run growth. *Journal of political economy*, 94(5):1002–1037.

Rostow, W. W. (1960). *The stages of growth: A non-communist manifesto.* Cambridge University Press Cambridge.

Sen, A. (2003). On unit-root tests when the alternative is a trend-break stationary process. *Journal of Business & Economic Statistics*, 21(1):174–184.

Silverberg, G. and Verspagen, B. (1994). Learning, innovation and economic growth: a long-run model of industrial dynamics. *Industrial and Corporate Change*, 3(1):199–223.

Silverberg, G. and Verspagen, B. (1995). An evolutionary model of long term cyclical variations of catching up and falling behind. *Journal of Evolutionary Economics*, 5(3):209–227.

Sobreira, N., Nunes, L. C., and Rodrigues, P. M. (2014). Characterizing economic growth paths based on new structural change tests. *Economic Inquiry*, 52(2):845–861.

Verspagen, B. (1992). *Uneven growth between interdependent economies: an evolutionary view on technology gaps, trade and growth.* PhD thesis, Universiteit Maastricht.

Vogelsang, T. J. and Perron, P. (1998). Additional tests for a unit root allowing for a break in the trend function at an unknown time. *International Economic Review*, pages 1073–1100.

Zerbo, E. and Darné, O. (2018). Unit root and trend breaks in per capita output: evidence from sub-saharan african countries. *Applied Economics*, 50(6):634–658.

Zivot, E. and Andrews, D. W. (1992). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business & Economic Statistics*, 10(3):251–270.

Appendix A Robustness checks

Country	T	M=3			M=4		
		Breaks	t_α	F_T	Breaks	t_α	F_T
OECD							
Australia	197	1851, 1891, 1928	$-8.647^{***}_{(***)}$	$12.876^{***}_{(***)}$	1851, 1891, 1928, 1961	$-8.764^{***}_{(**)}$	$10.466^{***}_{(***)}$
Austria	147	1944, 1959, 1974	-9.18^{***}	14.595^{***}	1913, 1944, 1959, 1974	-13.742^{***}	28.369^{***}
Belgium	171	1916, 1939, 1958	−4.364	5.886	1916, 1939, 1958, 1976	−7.094	7.777
Canada	147	1908, 1933, 1970	−6.111	8.115	1894, 1917, 1933, 1970	−6.252	6.713
Denmark	197	1914, 1939, 1963	−5.174	7.624	1850, 1916, 1939, 1963	−6.289	7.615
Finland	157	1916, 1932, 1968	−6.446	6.919	1916, 1932, 1953, 1978	−7.103	7.676
France	197	1916, 1939, 1968	$-10.85^{***}_{(***)}$	$18.711^{***}_{(***)}$	1869, 1916, 1939, 1968	$-11.127^{***}_{(***)}$	$15.265^{***}_{(**)}$
Germany	167	1913, 1945, 1972	$-11.823^{***}_{(*)}$	21.289^{***}	1913, 1945, 1963, 1980	$-13.535^{***}_{(**)}$	23.066^{***}
Greece	184	1912, 1939, 1962	−6.712	9.787^{**}	1862, 1910, 1939, 1962	−5.832	6.319
Italy	197	1913, 1938, 1968	−4.89	5.299	1852, 1905, 1938, 1968	−5.951	5.435
Japan	147	1944, 1959, 1990	−5.464	20.585^{***}	1944, 1959, 1977, 1995	-12.407^{***}	24.009^{***}
Netherlands	197	1943, 1963, 1996	−6.433	7.034	1922, 1943, 1963, 1990	-8.186^{**}	9.357^{**}
New Zealand	147	1907, 1935, 1975	−4.189	4.946	1895, 1910, 1935, 1975	−6.715	8.599
Norway	187	1935, 1954, 1997	−3.284	4.594	1905, 1935, 1954, 1987	−4.957	4.263
Portugal	152	1914, 1934, 1973	−0.992	6.686	1914, 1934, 1973, 1992	−1.73	6.756
Spain	167	1935, 1960, 1999	-7.137^{*}	10.773^{**}	1935, 1960, 1980, 1999	−7.616	9.666^{**}
Sweden	197	1894, 1939, 1968	−5.302	5.496	1869, 1916, 1939, 1968	−5.9	5.845
Switzerland	166	1913, 1940, 1974	−7.021	9.038	1883, 1921, 1940, 1974	−4.094	6.753
United Kingdom	197	1918, 1944, 1996	$-9.184^{***}_{(***)}$	$13.664^{***}_{(***)}$	1875, 1918, 1944, 1996	$-8.787^{***}_{(**)}$	$11.911^{***}_{(***)}$
USA	197	1929, 1949, 1996	−6.549	7.519	1878, 1929, 1949, 1996	$-8.872^{***}_{(**)}$	$10.315^{***}_{(**)}$
Asia							
India	133	1930, 1945, 1964	−0.952	5.98	1916, 1935, 1950, 1964	−0.621	4.788
Taiwan	116	1943, 1955, 1985	−4.179	10.327^{**}	1943, 1955, 1985, 1997	−4.399	8.287
Sri Lanka	147	1904, 1965, 2000	−3.822	6.485	1900, 1943, 1975, 2000	−7.244	8.285
Latin America							
Argentina	142	1891, 1929, 1980	−6.548	7.792	1896, 1917, 1963, 1980	−5.291	5.101
Bolivia	127	1891, 1928, 1971	−6.248	6.945	1930, 1952, 1978, 1998	−5.874	7.197
Brazil	167	1891, 1928, 1971	−6.457	9.116^{*}	1891, 1929, 1962, 1986	−5.58	5.74
Chile	197	1918, 1957, 1981	-7.539^{**}	8.572	1888, 1913, 1936, 1981	−6.901	7.096
Colombia	147	1934, 1967, 1997	−2.027	2.13	1917, 1934, 1967, 1998	−2.636	3.257
Ecuador	117	1945, 1972, 1998	−6.505	7.792	1930, 1947, 1972, 1998	−7.065	7.732
Mexico	122	1930, 1942, 1981	−5.491	7.176	1930, 1942, 1981, 2002	−5.372	5.661
Panama	111	1929, 1945, 1985	−6.444	7.382	1929, 1945, 1981, 2005	−6.317	7.18
Peru	197	1877, 1929, 1982	$-8.379^{***}_{(**)}$	$14.592^{***}_{(***)}$	1877, 1929, 1959, 1982	$-8.622^{***}_{(**)}$	$12.347^{***}_{(**)}$
Uruguay	147	1913, 1943, 2000	−5.806	5.695	1913, 1948, 1981, 2001	−6.843	6.013
Venezuela	187	1895, 1947, 1979	−3.969	8.364	1861, 1895, 1947, 1979	−4.274	7.624

Notes: Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Significance levels in brackets refer to rejections using bootstrapped critical values. The trimming parameter is set to $h = 0.1$ and the lag-truncation parameter to $K = 7$.

Table A.1: Results from unit roots tests imposing a fixed number of breaks for $M = 3$ and $M = 4$

Country	T	M=3			M=4		
		Breaks	t_α	F_T	Breaks	t_α	F_T
OECD							
Australia	197	1851, 1891, 1928	$-8.647^{***}_{(***)}$	$12.876^{**}_{(**)}$	1840, 1853, 1891, 1928	$-8.778^{**}_{(***)}$	$13.42^{***}_{(**)}$
Austria	147	1913, 1943, 1973	-9.044^{***}	13.07^{***}	1913, 1943, 1973	-9.044^{***}	13.07^{***}
Belgium	171	1918, 1943	0.837	$16.135^{***}_{(*)}$	1918, 1943	0.837	16.135^{***}
Canada	147	1908, 1939, 1970	$-7.87^{**}_{(**)}$	9.331	1917, 1930, 1942, 1970	-6.779	$11.323^{**}_{(*)}$
Denmark	197	1939, 1950	-1.055	$13.947^{***}_{(**)}$	1939, 1950	-1.055	$13.947^{***}_{(**)}$
Finland	157	1916, 1943, 1978	-7.331	9.081	1916, 1943, 1978	-7.331	9.081
France	197	1916, 1939, 1975	$-11.503^{***}_{(***)}$	$20.189^{***}_{(***)}$	1916, 1929, 1939, 1974	$-12.198^{***}_{(***)}$	$19.432^{***}_{(***)}$
Germany	167	1913, 1944, 1953	-5.497	26.788^{***}	1913, 1944, 1953	-5.497	26.788^{***}
Greece	184	1912, 1939, 1978	$-9.583^{***}_{(***)}$	$14.622^{***}_{(**)}$	1912, 1939, 1978	$-9.583^{***}_{(**)}$	$14.622^{***}_{(**)}$
Italy	197	1945	-2.964	$20.474^{***}_{(**)}$	1945	-2.964	$20.474^{***}_{(**)}$
Japan	147	1943, 1972	-8.522^{***}	20.321^{***}	1943, 1972	-8.522^{**}	20.321^{***}
Netherlands	197	1918, 1943, 1954	-3.715	11.958^{**}	1918, 1929, 1943, 1953	-3.736	14.913^{***}
New Zealand	147	1878, 1907, 1934	-4.523	8.785	1878, 1907, 1934	-4.523	8.785
Norway	187	1935, 1944, 1995	-3.29	$10.734^{**}_{(*)}$	1935, 1944, 1995	-3.29	10.734^{**}
Portugal	152	1934, 1973, 1986	-1.459	8.022	1934, 1973, 1986	-1.459	8.022
Spain	167	1935, 1960, 2002	-7.19^{*}	10.835^{**}	1935, 1960, 2002	-7.19	10.835^{**}
Sweden	197	1894, 1939, 1970	-5.236	5.685	1894, 1939, 1970	-5.236	5.685
Switzerland	166	1867, 1945	-1.702	7.895	1867, 1912, 1927, 1945	-2.472	7.161
United Kingdom	197	1918, 1943, 2006	-6.379	$13.458^{***}_{(**)}$	1918, 1943, 2006	-6.379	$13.458^{***}_{(**)}$
USA	197	1929, 1944	-3.266	$15.03^{***}_{(***)}$	1878, 1931, 1944, 2006	-6.777	$12.038^{**}_{(**)}$
Asia							
India	133	1944, 1964	-1.176	10.369	1944, 1964	-1.176	10.369
Taiwan	116	1923, 1943, 1998	-8.39^{***}	11.257^{**}	1923, 1943, 1998	-8.39^{**}	11.257^{**}
Sri Lanka	147	1907, 1972	-5.291	7.848	1907, 1972	-5.291	7.848
Latin America							
Argentina	142	1882	-4.27	7.745	1899, 1913, 1929, 1980	-7.638	7.612
Bolivia	127	1929, 1952, 1981	-7.051	8.943	1929, 1952, 1981	-7.051	8.943
Brazil	167	1891, 1928, 1970	-6.612	9.35	1891, 1929, 1962, 1980	-6.788	8.509
Chile	197	1918, 1932, 1981	-5.985	8.522	1918, 1930, 1974, 1991	$-9.567^{***}_{(**)}$	$11.773^{**}_{(*)}$
Colombia	147	1906, 1939, 1997	-5.766	9.444	1906, 1939, 1997	-5.766	9.444
Ecuador	117	1945, 1972, 2003	-5.544	7.8	1945, 1972, 2003	-5.544	7.8
Mexico	122	1925, 1932, 1981	-4.899	12.139^{**}	1925, 1932, 1981	-4.899	12.139^{**}
Panama	111	1929, 1945, 1987	-7.039^{*}	8.491	1929, 1945, 1987	-7.039	8.491
Peru	197	1876, 1987	$-7.636^{**}_{(*)}$	$13.435^{***}_{(**)}$	1876, 1987	-7.636	$13.435^{***}_{(**)}$
Uruguay	147	1919, 1930, 1957	-6.459	7.647	1897, 1913, 1930, 1957	-7.093	7.286
Venezuela	187	1895, 1944, 2003	-3.979	$10.436^{*}_{(*)}$	1895, 1944, 2003	-3.979	$10.436^{**}_{(*)}$

Notes: Significance levels: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$. Significance levels in brackets refer to rejections using bootstrapped critical values. The trimming parameter is set to $h = 0.05$ and the lag-truncation parameter to $K = 7$.

Table A.2: Results from unit roots tests imposing $h = 0.05$ for $M = 3$ and $M = 4$

Appendix B Series-specific critical values

Country	ARMA order	Statistic	Gaussian shocks (M=3)			Gaussian shocks (M=4)			Bootstrapping (M=3)			Bootstrapping (M=4)		
			1%	5%	10%	1%	5%	10%	1%	5%	10%	1%	5%	10%
Australia	(0, 0)	t_α	−7.98	−7.38	−7.04	−8.63	−8.00	−7.65	−8.48	−7.84	−7.47	−9.34	−8.59	−8.21
		F_T	13.91	10.75	9.85	13.41	10.61	9.81	13.89	11.60	10.72	13.67	11.62	10.74
Austria	(0, 0)	t_α	−7.91	−7.38	−7.00	−8.65	−7.99	−7.62	−23.72	−19.65	−17.76	−26.36	−21.76	−19.61
		F_T	12.83	10.79	9.92	12.38	10.72	9.92	95.65	70.27	55.72	101.90	73.63	58.58
Belgium	(0, 1)	t_α	−8.00	−7.38	−7.06	−8.76	−8.04	−7.67	−10.56	−9.44	−8.85	−11.75	−10.55	−9.92
		F_T	13.34	10.84	9.94	12.99	10.69	9.92	20.10	16.52	14.82	20.96	17.10	15.46
Canada	(1, 0)	t_α	−7.97	−7.36	−6.99	−8.64	−7.93	−7.56	−8.42	−7.67	−7.29	−9.13	−8.38	−7.95
		F_T	12.67	10.82	9.96	12.40	10.72	9.96	13.90	11.47	10.50	13.80	11.51	10.60
Denmark	(0, 0)	t_α	−7.98	−7.38	−7.04	−8.63	−8.00	−7.65	−8.94	−8.05	−7.65	−9.82	−8.94	−8.49
		F_T	13.91	10.75	9.85	13.41	10.61	9.81	14.96	12.39	11.17	15.00	12.46	11.37
Finland	(1, 1)	t_α	−8.10	−7.41	−7.03	−8.68	−8.00	−7.58	−8.96	−8.13	−7.64	−9.81	−8.89	−8.35
		F_T	13.23	10.89	10.02	12.93	10.74	10.00	15.25	12.75	11.63	15.31	12.97	11.84
France	(0, 0)	t_α	−7.98	−7.38	−7.04	−8.63	−8.00	−7.65	−9.91	−8.84	−8.29	−11.14	−9.84	−9.20
		F_T	13.91	10.75	9.85	13.41	10.61	9.81	17.66	14.61	13.12	18.12	14.75	13.38
Germany	(0, 1)	t_α	−8.04	−7.44	−7.08	−8.64	−8.03	−7.67	−15.53	−13.11	−11.80	−17.68	−15.00	−13.39
		F_T	13.84	11.02	10.08	13.21	10.86	10.01	43.20	31.17	25.78	46.30	33.58	27.67
Greece	(0, 0)	t_α	−8.00	−7.36	−7.03	−8.62	−8.01	−7.68	−9.24	−8.38	−7.92	−10.28	−9.37	−8.84
		F_T	13.35	10.67	9.86	12.86	10.55	9.81	15.82	13.13	12.05	15.88	13.32	12.17
Italy	(0, 1)	t_α	−8.09	−7.39	−7.07	−8.72	−8.05	−7.68	−10.92	−9.55	−8.94	−12.19	−10.62	−9.84
		F_T	14.12	10.95	10.01	13.53	10.85	9.95	21.21	17.15	15.06	21.78	17.30	15.45
Japan	(0, 0)	t_α	−7.91	−7.38	−7.00	−8.65	−7.99	−7.62	−18.54	−15.50	−14.14	−20.44	−17.02	−15.29
		F_T	12.83	10.79	9.92	12.38	10.72	9.92	59.94	43.71	36.29	61.58	44.85	37.04
Netherlands	(0, 1)	t_α	−8.00	−7.38	−7.06	−8.63	−8.02	−7.66	−14.94	−12.50	−11.55	−16.87	−14.27	−13.18
		F_T	14.21	10.83	9.93	13.46	10.71	9.86	38.53	28.14	24.08	41.09	29.49	25.35
New Zealand	(2, 0)	t_α	−8.61	−7.73	−7.27	−9.33	−8.36	−7.89	−9.10	−8.03	−7.51	−9.94	−8.76	−8.20
		F_T	14.17	11.32	10.27	13.84	11.13	10.25	14.89	12.00	10.82	14.64	11.94	10.91
Norway	(0, 0)	t_α	−7.94	−7.35	−7.02	−8.57	−7.98	−7.64	−8.39	−7.77	−7.41	−9.31	−8.50	−8.10
		F_T	13.47	10.72	9.83	13.11	10.58	9.76	14.07	11.54	10.55	13.81	11.56	10.60
Portugal	(0, 0)	t_α	−8.06	−7.41	−7.04	−8.66	−8.03	−7.64	−8.34	−7.65	−7.27	−9.15	−8.33	−7.93
		F_T	13.20	10.79	9.95	12.84	10.77	9.93	13.51	11.30	10.43	13.34	11.34	10.50
Spain	(0, 0)	t_α	−7.99	−7.41	−7.06	−8.61	−8.00	−7.64	−9.84	−8.60	−8.02	−10.67	−9.48	−8.87
		F_T	13.49	10.83	9.94	13.15	10.68	9.87	17.22	14.11	12.64	17.21	14.27	12.83
Sweden	(0, 0)	t_α	−7.98	−7.38	−7.04	−8.63	−8.00	−7.65	−8.19	−7.51	−7.18	−8.94	−8.25	−7.88
		F_T	13.91	10.75	9.85	13.41	10.61	9.81	13.72	11.05	10.14	13.01	10.92	10.13
Switzerland	(0, 2)	t_α	−9.92	−8.64	−7.99	−10.27	−9.11	−8.53	−10.33	−9.16	−8.49	−11.10	−9.87	−9.21
		F_T	19.11	13.74	11.94	18.68	13.25	11.61	18.13	14.27	12.70	17.60	14.02	12.56
United Kingdom	(0, 1)	t_α	−8.00	−7.37	−7.05	−8.63	−8.01	−7.66	−8.39	−7.72	−7.34	−9.22	−8.45	−8.07
		F_T	14.07	10.78	9.88	13.25	10.70	9.83	14.14	11.44	10.51	13.81	11.50	10.54
USA	(0, 0)	t_α	−7.98	−7.38	−7.04	−8.63	−8.00	−7.65	−8.30	−7.59	−7.24	−9.07	−8.33	−7.98
		F_T	13.91	10.75	9.85	13.41	10.61	9.81	13.99	11.23	10.36	13.77	11.14	10.30
India	(1, 2)	t_α	−9.65	−7.78	−6.82	−10.64	−9.07	−7.89	−10.06	−8.15	−7.04	−11.06	−9.60	−8.40
		F_T	17.79	13.89	12.34	17.56	13.82	12.35	17.99	14.69	13.15	17.66	14.72	13.23
Taiwan	(1, 0)	t_α	−8.11	−7.39	−7.00	−8.65	−7.89	−7.47	−13.92	−11.94	−10.82	−15.32	−13.02	−11.82
		F_T	13.66	11.16	10.16	13.32	11.06	10.16	38.33	28.80	24.41	39.52	29.75	25.46
Sri Lanka	(0, 3)	t_α	−8.00	−7.38	−7.00	−8.70	−7.99	−7.59	−8.20	−7.55	−7.20	−8.95	−8.20	−7.77
		F_T	13.23	11.01	10.15	12.85	10.88	10.14	14.04	11.51	10.57	13.95	11.54	10.62
Argentina	(0, 0)	t_α	−7.89	−7.40	−7.05	−8.60	−7.98	−7.64	−8.39	−7.69	−7.33	−9.09	−8.36	−7.93
		F_T	13.52	10.77	9.91	12.95	10.73	9.90	13.72	11.56	10.59	13.57	11.55	10.59
Bolivia	(0, 1)	t_α	−8.10	−7.47	−7.12	−8.76	−8.03	−7.60	−8.50	−7.72	−7.34	−9.19	−8.36	−7.91
		F_T	13.24	11.12	10.18	13.06	11.11	10.19	14.39	11.98	10.95	14.39	12.01	10.98
Brazil	(0, 0)	t_α	−7.99	−7.41	−7.06	−8.61	−8.00	−7.64	−8.30	−7.62	−7.25	−9.10	−8.36	−7.96
		F_T	13.49	10.83	9.94	13.15	10.68	9.87	14.38	11.34	10.35	14.01	11.35	10.44
Chile	(0, 0)	t_α	−7.98	−7.38	−7.04	−8.63	−8.00	−7.65	−8.74	−7.95	−7.59	−9.60	−8.80	−8.39
		F_T	13.91	10.75	9.85	13.41	10.61	9.81	15.16	12.19	11.18	14.86	12.24	11.25
Colombia	(1, 0)	t_α	−7.98	−7.37	−7.00	−8.65	−7.93	−7.55	−8.78	−7.94	−7.50	−9.59	−8.73	−8.21
		F_T	13.06	10.84	10.02	12.51	10.77	10.02	15.02	12.53	11.38	15.08	12.75	11.64
Ecuador	(0, 0)	t_α	−8.03	−7.39	−7.01	−8.65	−7.98	−7.56	−8.78	−7.92	−7.52	−9.58	−8.67	−8.18
		F_T	13.13	10.99	10.17	12.87	10.97	10.21	14.49	12.24	11.23	14.69	12.50	11.45
Mexico	(0, 0)	t_α	−8.11	−7.41	−7.04	−8.67	−7.94	−7.57	−9.37	−8.31	−7.79	−10.16	−9.02	−8.47
		F_T	12.97	11.05	10.12	12.59	10.90	10.10	17.02	13.76	12.24	17.33	13.95	12.45
Panama	(2, 1)	t_α	−8.07	−7.35	−6.97	−8.55	−7.82	−7.44	−9.24	−8.24	−7.67	−10.00	−8.83	−8.25
		F_T	13.58	11.24	10.35	13.58	11.29	10.41	17.76	14.34	12.83	17.98	14.55	13.20
Peru	(1, 0)	t_α	−7.97	−7.33	−7.01	−8.63	−7.99	−7.62	−8.64	−7.83	−7.46	−9.51	−8.60	−8.19
		F_T	13.83	10.81	9.89	13.12	10.72	9.82	14.60	11.99	10.95	14.15	12.05	10.94
Uruguay	(0, 0)	t_α	−7.91	−7.38	−7.00	−8.65	−7.99	−7.62	−8.46	−7.68	−7.31	−9.19	−8.39	−7.98
		F_T	12.83	10.79	9.92	12.38	10.72	9.92	14.28	11.68	10.62	14.21	11.78	10.74
Venezuela	(0, 0)	t_α	−7.94	−7.35	−7.02	−8.57	−7.98	−7.64	−8.13	−7.54	−7.19	−8.94	−8.20	−7.84
		F_T	13.47	10.72	9.83	13.11	10.58	9.76	13.63	11.02	10.08	13.46	10.95	10.07

Notes: Monte Carlo simulations for each time series assume $h = 0.1$ and $K = 7$.

Table B.1: Series-specific critical values for M=3 and M=4 – derived using the Monte Carlo method under both Gaussian shocks and bootstrapping