Oxford Economic Papers 48 (1996), 134–162

DETERMINISTIC, STOCHASTIC, AND SEGMENTED TRENDS IN AGGREGATE OUTPUT: A CROSS-COUNTRY ANALYSIS

By YIN-WONG CHEUNG *and* MENZIE DAVID CHINN
Department of Economics, University of California, Santa Cruz, CA 95064, USA

This paper examines whether output *per capita* in 126 countries is better described as trend or difference stationary, using appropriate finite-sample critical values. Depending upon whether one uses solely a test with a trend stationary null, or solely one with a difference stationary null, very different conclusions are obtained. This outcome suggests that it is useful to consider the tests complementary, rather than competing. We find that when a definite characterization of GDP can be made, it is very likely to indicate a difference stationary process. However, the likelihood of making definite conclusions does vary positively with both income level and data quality.

1. Introduction

ONE of the most active debates in macroeconomics over the past decade has been whether output contains a unit root. The attention devoted to this topic is well deserved. Knowledge regarding the persistence of GDP is important for several reasons. First, whether GDP is trend or difference stationary has dramatic implications for the long run behavior of output dynamics. A shock to a trend stationary process has a transitory impact, whereas a shock to a difference stationary one permanently shifts the trend. Second, when GDP is used in regressions with other variables, the interpretation of the regression results can depend on whether the variables involved are trend or difference stationary. This phenomenon is related to the 'spurious regression' literature due to Granger and Newbold (1974). Third, the issue of output persistence is also relevant to the recent empirical studies examining the hypothesis of convergence across countries (e.g., Bernard and Durlauf, 1991). A necessary, though not sufficient, condition is that the output data have common long run persistence. Recently the finding of a unit root GDP process based on the conventional unit root tests, such as augmented Dickey–Fuller (ADF) and the Phillips–Perron tests, has been re-evaluated using different methodologies. Christiano and Eichenbaum (1990) and Rudebusch (1993), for example, argue that the standard unit root tests do not have much power to differentiate between the trend stationary and difference stationary properties of GDP. Thus these tests may generate spurious unit root results.

Another concern is the effect of structural instability on these unit root tests. Perron (1989b, 1990) shows that the conventional unit root tests tend to misinterpret a trend stationary time series with a structural break as a difference stationary series. For instance, Kormendi and Meguire (1990) find evidence of a unit root in both long historical and post Second World War GDP data. However, Zelhorst and De Haan (1994a) and De Haan and Zelhorst (1993) show that the unit root result may be spurious and induced by structural changes in the data generating process. The issue of structural breaks' effects on unit root tests becomes more prominent when one has to analyze GDP series over long time periods, during which changes in regime may occur.

This study attempts to provide a better insight on the persistence of GDP data in the following manner. First we use different types of tests to examine the unit root property. In addition to the standard ADF test which has the unit root process as the null hypothesis and trend stationarity as the alternative hypothesis, we also subject the data to: (i) a unit root test which has trend stationarity as the null and the unit root process as the alternative, developed by Kwiakowski, Phillips, Schmidt, and Shin (1992, hereafter KPSS), and; (ii) the unit root test which has trend stationarity and a structural break as the alternative, developed by Banerjee, Lumsdaine, and Stock (1992, hereafter BLS). Second, we assess the test results using the appropriate finite sample critical values. Finally, we examine GDP data across 126 countries to investigate if unit root persistence is a general phenomenon, or a characteristic that only applies to a specific group of countries.

Regarding the first point, consider the case of two econometricians, one using the ADF test, and the other, the KPSS test. Each independently might make quite different conclusions regarding the time series characteristics of GDP. In fact, in the data set used in this analysis, they would. Econometrician A, using the ADF, would conclude that most GDP series were difference stationary, while econometrician K, using the KPSS test, would conclude that most series were trend stationary.

Instead of viewing the tests as competing, the KPSS test results could be used to corroborate the information obtained from the ADF test, and vice versa. Suppose the ADF test fails to reject the unit root null because of low power. The KPSS test which has trend stationarity as the null should indicate the data have no unit roots. On the other hand, if the KPSS test rejects the trend stationarity null, then we have stronger evidence for unit root persistence. That is, consistent results from ADF and KPSS tests yield more persuasive evidence on data persistence while conflicting results indicate uncertainty associated with the interpretation of the individual test outcomes.

The BLS test is used to determine if the apparent unit root results are attributable to discrete trend breaks in the GDP data. One advantage of the BLS test is that this procedure is not conditional on a pre-assigned break point. By endogenizing the break point date in the procedure, the BLS test avoids the criticism of data-mining and provides an objective evaluation of the unit root and stationary-with-structural-break hypotheses.[1]

Regarding the second point, our study is careful to ensure that neither the use of asymptotic KPSS critical values, nor the use of the standard ADF critical values, results in inappropriate inferences. We expect in finite samples that the use of critical values will result in over-rejection. Hence, we will report results

1 Recently, Cheung (1994) has applied this test to the GDP data of the G-7 countries, for the period 1900–90.

using finite sample critical values; those results using asymptotic critical values are available from the authors.

Finally, this paper studies the unit root property of GDP data contained in the Summers–Heston data set, which provides the longest consistent output series for a large number of countries. This last attribute is important because we wish to identify patterns of acceptance and rejection that may be missed if we only test 'the usual suspects', such as the G-7 countries.[2] The patterns considered in this study include results one might obtain when examining differences between: (i) developed versus less-developed countries; (ii) different regions, and; (iii) low and high quality data. We are unaware of any other systematic investigation of GDP persistence using these techniques and over such a wide set of countries.[3]

There is one caveat. We do not attempt to fit our empirical findings to any specific economic theory. Rather, the interested reader is encouraged to decide for him- or herself the theoretical reasons underlying the results. In this sense, the current study is an explicitly empirical exercise.

This paper is organized as follows. In Section 2, we discuss the data, and then outline the econometric methodology used. The ADF and KPSS results are presented in Section 3, while in Section 4 we apply the test of BLS to those series which reject the trend stationary hypothesis. Section 5 concludes:

2. Data and methodology

2.1. *Data description*

The data used are the annual Gross Domestic Product (GDP) *per capita* in real terms using international prices, for 126 countries, obtained from the Summers and Heston (1991) Penn World Tables, Mark 5 (PWT5). In the following discussion, we refer to these log real *per capita* output as simply GDP for convenience.[4]

The use of GDP evaluated at international prices differentiates this study from nearly all previous ones. Typically, cross-country studies rely on domestically valued GDP, in constant prices. We choose to examine the characteristics of the Summers–Heston measure of output exactly because it is comparable across countries. Consequently, we abstract from idiosyncratic relative price

[2] Cogley (1990), Banerjee *et al.* (1992), for example. Another difference between the current work and previous studies is the sample period. Some previous studies examine long historical data that span up to 100 years or more. Because of the consistency of GDP data across a large number of countries, we choose the Heston–Summers data. The trade-off is the possible reduction in the test power.

[3] Although Riezman and Whiteman (1990) do apply Bayesian unit root tests to an almost equally large number of countries, they do not interpret the individual country results, nor do they analyze how the results vary over certain county groupings.

[4] The variable used is RGDPCH (the Summers–Heston mnemonic), which is real GDP *per capita* evaluated in 1985 international prices calculated using a Chain index. The use of international prices converts GDP data from different countries to a common denomination and facilitate cross-country comparison. The specifics of the data series are detailed in Summers and Heston (1991), to which the interested reader is referred.

changes in each country when characterizing the time series properties of output. The relevance of this study is reinforced by the widespread use of the Summers and Heston data set in a number of influential studies of the determinants of long-run growth.

The sample period covers the post-War era for the developed countries, and somewhat shorter time lengths for some of the less-developed countries, for the period up to 1988. The data span at most 39 observations, and at minimum ten (St. Vincent). These sample sizes may appear small, and hence likely to worsen the power of the ADF test (for example, see Harris, 1992).[5] However, as pointed out by Shiller and Perron (1985) and Perron (1989a), the power of the tests for stationarity depend mainly on the length of the data time span and not on the number of observations. That is, the power of the test is essentially the same for both a sample containing 39 annual data points and a sample containing 39 × 12 monthly observations.

2.2. *Econometric methodology*

2.2.1. *Overview.* As mentioned in the introduction, GDP will be subjected to both the ADF and the KPSS tests. The results of these two tests are used to determine the nature of persistence in each GDP series. To investigate the possibility of misinterpreting a structural break as unit root persistence, we apply the BLS test to those series that both the ADF and KPSS tests indicate the presence of a unit root.

The null and alternative hypotheses of these tests can be summarized as follows:

Test	H_0:	H_A:
ADF	I(1)	I(0)
KPSS	I(0)	I(1)
BLS(A)	I(1)	I(0) w/shift in trend
BLS(B)	I(1)	I(0) w/shift in mean

2.2.2. *The ADF test.* Let $\{y_t\}$ be the GDP series. The ADF test for unit roots is based on the regression

$$\Delta y_t = c + \mu t + \pi y_{t-1} + \sum_{j=1}^{k} \gamma_j \Delta y_{t-j} + u_t \qquad (1)$$

The unit root null hypothesis is rejected if $\hat{\pi}$ is significantly less than zero according to the finite sample critical values calculated from simulated

[5] We report the results for all the series below. The overall pattern of rejections and non-rejections is virtually the same if we omit series with below 20 observations (there are only five such cases), or, surprisingly, even if we omit series with below 30 observations. However, in this latter case, the number of GDP series available is reduced from 126 to 75, or about 40%.

distributions which control for both sample size and lag structure (Cheung and Lai, 1995).

In applying this test, one must determine the lag length parameter, k. One possibility is to select an arbitrary, large k, to capture any possible series correlation in the data. Recently, Alastair Hall (1994) has shown that a general to specific modeling strategy that uses the data to determine k can improve both the size and power of the ADF test. Hence in this study, the Akaike Information Criterion and the Schwartz Bayesian Information Criterion (AIC and SBC respectively) are used to determine the lag parameter. In many cases the two criteria yield similar inferences and so in order to conserve space, we only report the results based on the AIC.

2.2.3. *Trend stationarity tests.* The KPSS test assumes that the time series is the sum of either a mean or a deterministic trend, a random walk, and a stationary error. It is a Lagrange Multiplier test for the null hypothesis that the error variance in the random walk component of the series is zero. In this study, we consider the case where the null is a trend stationary process.

To conduct the test, we first obtain the residual e_t from the regression of y_t on a constant and a trend. The KPSS statistics $\hat{\eta}_t$ is then given by

$$\hat{\eta}_t = T^{-2} \sum S_t^2 / s^2(l) \tag{2}$$

where S_t is the partial sum process defined by

$$S_t = \sum_{i=1}^{t} e_i \qquad t = 1, 2, \ldots, T \tag{3}$$

and $s^2(l)$ is the serial correlation and heteroskedasticity consistent variance estimator given by

$$s^2(l) = T^{-1} \sum_{t=1}^{T} e_t^2 + 2T^{-1} \sum_{s=1}^{l} w(s, l) \sum_{t=s+1}^{T} e_t e_{t-s} \tag{4}$$

$w(s, l)$ is an optimal weighting function corresponding to the choice of a spectral window.[6]

The null of trend stationarity is rejected in favor of the unit root alternative if the KPSS statistic is larger than the asymptotic critical values provided by KPSS. Based on simulation results, KPSS assert that their test has good size and power characteristics. As in the case of the ADF tests, to avoid spurious conclusions resulting from the use of asymptotic critical values, we also employ finite sample critical values (Cheung *et al.*, 1994). l in eq. (4) is a choice parameter to be determined. Following KPSS' suggestion we adopt the $l8$ rule, which sets $l = \text{INT}[8(T/100)^{1/4}]$.[7]

[6] We use a Bartlett window, $w(s, l) = 1 - s/(l + 1)$, as suggested by KPSS.

[7] Recently, Leybourne and McCabe (1994) developed a unit root test that also has stationarity as its null hypothesis. However, the alternative process in an I(1) process with an MA(1) component. This is in contrast with the KPSS test which has an I(0) process under the null and an I(1) process under the alternative. Further the use of finite-sample critical values that account for both sample-size and truncation-lag effects should minimize the possible size-distortion of the KPSS test in small samples.

2.2.4. *Structural breaks.* BLS propose several recursive and sequential tests that allow for structural breaks under the trend stationary alternative. Given the limited number of observations available, we use only the sequential tests.[8]

Consider the following model

$$\Delta y_t = \mu_0 + \mu_1 \tau_{1t}(p) + \mu_2 t + \alpha y_{t-1} + \beta(L)\Delta y_{t-1} + \varepsilon_t \quad (5)$$

where $\tau_{1t}(p)$ is a dummy variable. When a shift in the trend at time p is considered under the alternative hypothesis, then $\tau_{1t}(\text{p})$ is set equal to $(t - p)1(t > p)$, where $1(.)$ is an indicator function. Alternatively, when a shift in the mean at time p is considered, $\tau_{1t}(p)$ is set equal to $1(t > p)$. These two situations conform to BLS' Case A and Case B, respectively.

For case A, three sequential statistics are computed

$$\text{(i)} \quad F_{A,\max} \equiv \text{MAX}_{p_0 \leqslant p \leqslant T - p_0} F(p)$$

$$\text{(ii)} \quad t_A(\hat{p}), \text{ where } F(\hat{p}) = F_{A,\max}$$

$$\text{(iii)} \quad t_{A,\min} \equiv \text{MIN}_{p_0 \leqslant p \leqslant T - p_0} t_A(p)$$

where p_0 is the number of observations reserved by the trimming parameter δ_0, such that $p_0 = \delta_0 T$. We set $\delta_0 = 0.15$, as suggested by BLS.

$F(p)$ and $t_A(p)$ are, respectively, the F-statistic for the hypothesis of $\mu_1 = 0$ and the Dickey–Fuller t-statistic computed from the whole sample with $\tau_{1t}(p)$ defined as above. For case B, we compute similar statistics and label them as $F_{B,\max}$, $t_B(\hat{p})$ and $t_{B,\min}$.

The $\beta(L)$ polynomial serves to model the serial correlation in the data, a purpose analogous to the differenced terms in the ADF regression. To facilitate the comparison between the ADF and BLS tests, the lag orders chosen by the AIC used in the ADF regressions are also used to conduct the BLS tests.

The asymptotic distributions and the finite sample critical values of these sequential tests are given in BLS (1992). The smallest sample size considered by these authors is 100, which is relatively large compared to the sample sizes of our data set. To guard against inappropriate inferences caused by small sample effects, the tests conducted in this portion of the study are based on simulated critical values for the sample sizes 35 and 25.[9]

[8] In the literature there are other methods to detect the break-point. As Zelhorst and De Haan (1994b) and Cheung (1994) indicate various methods to determine the break data yield similar results. Hence the use of the BLS method should yield reasonable results.

[9] The simulated critical values are based on 10,000 replications. The computer code is available from the authors upon request.

3. Unit root and trend stationarity tests results

3.1. *Overview*

The specific country results are contained in Appendix 1. The overall results are summarized in the following format:

	KPSS accept	KPSS reject	
ADF accept	cell 1	cell 2	row total
ADF reject	cell 3	cell 4	row total
	column total	column total	total

Cell 1 reports the number and the percentage of cases that both the ADF and KPSS tests fail to reject (or more concisely and inaccurately, accept)[10] their respective null hypotheses. Cell 2 reports the number and the percentage of cases that the ADF accepts the unit root null while the KPSS test rejects the trend stationary null in favor of the unit root hypothesis. Cell 3 reports the number and the percentage of cases that the ADF rejects the unit root null in favor of the stationary alternative and the KPSS test accepts the trend stationary null. Cell 4 reports the number and the percentage of cases that both the ADF and KPSS tests reject their respective null hypotheses. The row total gives the total number and percentage from the ADF test while the column total gives those from the KPSS. The acceptance and rejection are based on the 5% significance level, and in general we will focus on the finite sample critical value results.

The interpretation of results from cells 2 and 3 is quite straightforward. Series that fall in cell 2 show a strong evidence of the presence of a unit root in the data while those in cell 3 show a strong evidence of trend stationarity. The cell 1 classification can be explained by the low power of both tests so that neither null is rejected: the data do not contain sufficient information to discriminate between the trend stationary and difference stationary hypotheses. On the other hand, the rejection of both the unit root and trend stationary null hypotheses, as in cell 4, cannot be attributed to the low power of one or both of the tests. One possible interpretation of such cases is that the data is governed by a more complex data generating process (DGP) than the standard trend and difference stationary classification.[11]

[10] For the sake of convenience, we will use the shorthand term 'accept' in place of the more cumbersome phrase 'fail to reject', while remaining cognizant of the difference.

[11] However, if the true series was trend stationary with a structural break, this would be insufficient to explain the results, as ADF tests are biased against rejection of the unit root null in such cases. Another possibility is that GDP is persistent, but not completely, as in fractionally integrated processes (e.g., Diebold and Rudebusch, 1989). The small samples we have precludes examining this possibility; for the Geweke and Porter-Hudak (1983) test, there would be only six effective observations. The presence of nonlinear trends is another possible explanation.

TABLE 1
Results for all countries

		KPSS test ACCEPT	KPSS test REJECT	
ADF test	ACCEPT	72 57%	39 31%	111 88%
	REJECT	14 11%	1 1%	15 12%
		86 68%	40 32%	126 100.0%

Notes: Results using the 5% significance level and finite sample critical values. 'Accept ADF and accept KPSS' indicates failure to reject both the unit root null and the trend stationary null hypotheses. 'Accept ADF and reject KPSS' indicates the failure to reject unit root null, but rejection of trend stationary null. 'Reject ADF and accept KPSS' indicates rejection of unit root null and failure to reject the trend stationary null. 'Reject ADF and Reject KPSS' indicates rejection of both the unit root and trend stationary null hypotheses. KPSS results refer to use of *l*8 rule.

3.2. *Results for all countries*

The results for the trend- and difference-stationarity tests for 126 countries are presented in Table 1.[12] For 111 (88%) of these series, the ADF test fails to reject the unit root null, and for 15 (12%), the test rejects (using a 5% marginal significant level). In contrast, 86 (68%) fail to reject the trend stationary null, and 40 (32%) reject, using the KPSS test. Hence, the econometrician relying solely on the ADF test might conclude that approximately 90% of all series contained unit roots, while the econometrician relying on the KPSS test would conclude that only one-third did.

When both test results are combined, the number of series that appear to be difference stationary is reduced considerably from that implied using solely the ADF. Admittedly, the most common outcome is Accept-Accept, which implies some degree of ambiguity (57%, or 72 series, occupy cell 1). However, the next most common outcome, 31% (39 series), is in cell 2, corresponding to the Accept ADF/Reject KPSS outcome. 11% (14 series) Reject ADF/Accept KPSS, which corresponds to a trend stationary view of GDP: Bahamas, Burundi, Canada, Congo, Cyprus, Hong Kong, Hungary, Iraq, Lesotho, Malta, Pakistan, the Philippines, South Korea, and Uganda. Only one series (Papua New Guinea) rejects both tests.

Two prominent overall results are illustrated in Table 1. For a large number of countries, no definite characterization can be made. However, for those countries that a definite characterization can be made, the difference stationary

[12] Results for samples restricted to series with 19+, and 29+, observations are reported in Appendix 2.

series outnumber the trend stationary series by about 3:1. This pattern holds true for most of the groupings examined below.

3.3. *Geographic differences*

One of the advantages of the wide selection availed by using the Summers and Heston data set is that we can stratify the sample by various criteria. One stratification we choose is by geographic location, or specifically, by continent. One motivation behind this choice is the recent trend of trading blocs in Europe, America, and Asia. Countries in a region linked by similar historical and cultural characteristics usually have similar economic conditions. This in turn makes it possible for these countries to cooperate and promote mutually beneficial economic goals. The results are presented in Table 2.

A consistent finding is that the proportion of failures to reject the ADF is always about 80%, rising to 100% for the case of Oceania (with, however, only five series). On the other hand, there is little variation in the tendency to reject the KPSS, except for the European region, where approximately one half of the series reject. Consequently, in all regions a large proportion fall into the ambiguous Accept ADF/Accept KPSS category, save in Europe, where 50% fall into the Accept ADF/Reject KPSS category. One common characteristic of the European countries is their relatively high income level. This suggests an alternative grouping.

3.4. *Level of development*

A plausible explanation for the previous results is that the GDP governing GDP is different, depending upon the level of development. In this view, less developed, primarily rural-agrarian societies are likely to have output series with very different time series properties than those from highly urban, industrialized economies. To examine this possibility, we segmented the results into three categories—low, medium, and high income countries, corresponding to the World Bank's (1989) definitions. Low income countries are those with GNP *per capita* of $480 or less in 1987; middle income countries are those with GNP *per capita* of more than $480 and less than $6,000; high income countries are those with GNP *per capita* in excess of $6,000.[13]

The results are tabulated by level of development in Table 3. The results for low- and middle-income countries are relatively similar: roughly 65% of the entries are Accept ADF/Accept KPSS, roughly 25% are Accept/Reject. The proportions change dramatically when one moves from these two categories to the high income category, primarily because then the KPSS test rejects trend

[13] It would have been preferable to use grouping based on incomes closer to the midpoint of the sample (e.g., 1970). However, earlier issues of the World Bank *World Development Report* grouped countries according to other development criteria besides income (see World Bank, 1989: x). Countries outside the World Bank classification scheme but in our sample are: (low) Gambia, Cape Verde, Guyana, Angola; (middle) Taiwan, Swaziland, St. Lucia, St. Vincent, Fiji, Suriname, Cyprus, Malta, Barbados; (high) Bahrain, Bahamas, Iceland, Luxembourg.

TABLE 2
Results by geographical region

African countries: $n = 43$

		KPSS test: ACCEPT	KPSS test: REJECT	
ADF test	ACCEPT	25 58%	14 33%	39 91%
	REJECT	4 9%	0 0%	4 9%
		29 67%	14 33%	43 100.0%

Asian countries: $n = 26$

		KPSS test: ACCEPT	KPSS test: REJECT	
ADF test	ACCEPT	17 65%	4 15%	21 81%
	REJECT	5 19%	0 0%	5 19%
		22 85%	4 15%	26 100.0%

European countries: $n = 23$

		KPSS test: ACCEPT	KPSS test: REJECT	
ADF test	ACCEPT	8 35%	12 52%	20 87%
	REJECT	3 13%	0 0%	3 13%
		11 48%	12 52%	23 100.0%

Continued

TABLE 2—*continued*

North American countries: $n = 17$

		KPSS test: ACCEPT	KPSS test: REJECT	
ADF test	ACCEPT	12 71%	3 18%	15 88%
	REJECT	2 12%	0 0%	2 12%
		14 82%	3 18%	17 100.0%

Oceania: $n = 5$

		KPSS test: ACCEPT	KPSS test: REJECT	
ADF test	ACCEPT	2 40%	2 40%	4 80%
	REJECT	0 0%	1 20%	1 20%
		2 40%	3 60%	5 100.0%

South American countries: $n = 12$

		KPSS test: ACCEPT	KPSS test: REJECT	
ADF test	ACCEPT	8 67%	4 33%	12 100%
	REJECT	0 0%	0 0%	0 0%
		8 67%	4 33%	12 100.0%

Notes: Results using 5% significance level. See notes to Table 1.

TABLE 3
Results by income level

Low income countries: $n = 38$

		KPSS test ACCEPT	*KPSS test* REJECT	
ADF test	ACCEPT	25 66%	9 24%	34 89%
	REJECT	4 11%	0 0%	4 11%
		29 76%	9 24%	38 100.0%

Middle income countries: $n = 59$

		KPSS test ACCEPT	*KPSS test* REJECT	
ADF test	ACCEPT	37 63%	14 24%	51 86%
	REJECT	7 12%	1 2%	8 14%
		44 75%	14 25%	59 100.0%

High income countries: $n = 29$

		KPSS test ACCEPT	*KPSS test* REJECT	
ADF test	ACCEPT	10 34%	16 55%	26 90%
	REJECT	3 10%	0 0%	3 10%
		13 45%	16 55%	29 100.0%

Notes: Results using 5% significance level. See notes to Table 1.

stationarity more often. Then the majority of cases fall into the Accept ADF/Reject KPSS cell (16 out of 29 cases, or 55%), while a smaller proportion fall into the Accept/Accept cell.

This particular pattern illustrates the importance of casting the net wider. The relatively unambiguous support in favor of unit roots (against the alternative of a simple linear trend) in GDP appears to be more a rich-country phenomenon, than a low or middle income one. In fact, among the OECD countries, only one country—Canada—rejects the unit root null.

One is tempted to judge this pattern of results as support for the view that high income (and especially developed)[14] economies behave differently than low and medium income ones. Caution in interpretation is necessary because high income countries also tend to be countries with high quality data, and the tendency to find more apparently trend stationary, or at least ambiguous, cases in the lower income countries may be due to the manner in which GDP series are calculated by these countries' statistical agencies. Alternatively, the valuation of GDP at international prices may be more precisely calculated for higher income countries.[15] In order to evaluate this possibility we stratify the GDP series by the quality of national income data.

3.5. *Data quality*

Summers and Heston (1991) conveniently provide a grade, ranging from D (worst) to A (best) for each country's data. While there is an apparent correlation between level of development and data quality, the correlation is not exact. All low income countries have grade D and C quality data, but middle income Greece and Portugal have grade A quality data. In contrast, high income countries such as the Bahamas and Singapore have D and C quality data. Moreover, the Middle East oil exporting nations are also classified as high income countries, but have D quality data. Overall, the correlation between income and data quality is 0.71.[16]

The results stratified by data quality are presented in Table 4. For the low quality data the results are similar across grades C and D: roughly 23% are in the Accept ADF/Reject KPSS category, with 65% in the ambiguous Accept/Accept cell. When one moves to the grade A and B quality data, one finds that a majority of the results indicate unit root processes, since 16 out of 28 series Accept ADF and Reject KPSS.

A priori, there is no particular reason to believe that poorly constructed data will appear more trend stationary than well constructed data. However, it is conceivable that LDCs with limited statistics-gathering resources may use

[14] The distinction is important, since some high income countries are oil exporters.

[15] The PWT5 data set covers a large number of countries for which detailed price data are not available. For these, usually less-developed, countries, the valuation calculations are based on a survey of capital city prices.

[16] The correlation coefficient was calculated by assigning numerical scores to the data quality indices, such as, 3.3 for B +, and correlating this with the income category (1 for low, 3 for high).

TABLE 4
Results by data quality

D and C quality data: $n = 98$

		KPSS test: ACCEPT	KPSS test: REJECT	
ADF test	ACCEPT	64 65%	23 23%	87 89%
	REJECT	10 10%	1 1%	11 11%
		74 76%	24 25%	98 100.0%

B and A quality data: $n = 28$

		KPSS test: ACCEPT	KPSS test: REJECT	
ADF test	ACCEPT	8 29%	16 57%	24 86%
	REJECT	4 14%	0 0%	5 14%
		12 43%	16 57%	28 100.0%

Notes: Results using the 5% significance level. See notes to Table 1.

methods of calculating GDP involving linear interpolation and smoothing which generate data that look like trend stationary processes. For instance, the output of the traditional, or subsistence, sector is often assumed to grow with population, which is in turn interpolated between infrequent population censuses (Blades and Marczewski, 1974; UNRISD, 1977). Since the traditional sector often constitutes a large proportion of total output, this can be a significant source of measurement error. At the limit, the reported *per capita* GDP figures would be either a constant or a deterministic trend reflecting assumed productivity growth.[17]

[17] An alternative data-based explanation is that higher quality data also tends to span longer periods. As the sample size grows the power of the KPSS test increases so that (given that the true processes are difference stationary) more entries fall into the 'Accept/Reject' cell. In fact, all grade A quality series span the entire 39 years. Grade C series range from 26 to 39 years, while grade D series are as short as ten years.

4. Structural break test results

It is possible that the apparent finding of unit roots is due to the presence of structural breaks in otherwise trend stationary processes. We apply the BLS sequential test to GDP series for which both the ADS and KPSS tests indicate the presence of a unit root (detailed individual country results are contained in Appendix 3). Series for four countries appear to have significant breaks in the trend *per capita* GDP (break years in brackets): Gambia (1978), Jamaica (1973), Spain (1975), and Yugoslavia (1980). The Japanese (1974) series also rejects the null hypothesis of a unit root at the 10% significance level, along with trend breaks essentially the same as those obtained by BLS. No series reject the null of no mean shift.[18]

In sum, for the 39 series for which we found evidence of a unit root, only four can be explained by the presence of structural breaks. Hence, discrete shifts do not explain the tendency to find a relatively large number of series that fail to reject the ADF and reject the KPSS.

5. Conclusions

We have found that for a large proportion of GDP series, it is not possible to make definitive conclusions regarding their time series properties. However, for those series that a conclusion can be made (either Accept/Reject or Reject/Accept), we find evidence in support of unit roots as opposed to simple linear trends, in the ratio of approximately 3:1. This conclusion is based on the application of the ADF test, which has the unit root process as a null hypothesis, in conjuction with the KPSS test, which has a null of trend stationarity. Moreover we assess whether the apparent acceptance of the unit root view is due to the presence of structural breaks, of the nature examined by Banerjee *et al.* (1992), and find that the cases of structural break account for a definite minority of the apparent unit root processes. In all these tests, we have controlled for possible biases due to the use of inappropriate asymptotical critical values.

Finally, we have also found that there is substantial variation in how strongly the accept ADF/reject KPSS result holds. In particular, low incomes and poor data quality appear to be associated with greater ambiguity.

ACKNOWLEDGEMENTS

Helpful comments were received from Tim Cogley, David Papell, Doug Steigerwald, Shang-Jin Wei, seminar participants at UC Santa Barbara, Harvard University, and our colleagues at UCSC. We thank Tuan Tran for research assistance, and the Pacific Rim Project of the University of California for financial support. Any remaining errors remain solely ours.

[18] It is interesting that the indicated break points coincide with significant political events. For instance, Spain's and Yugoslavia's implied breaks match with the deaths of Franco and Tito, respectively, while the break in Jamaica coincides with Manley's attemped restructuring of the economy.

REFERENCES

BANNERJEE, 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 and Economic Studies*, **10**, 271–87.

BERNARD, A. B. and DURLAUF, S. N. (1991). 'Convergence of International Output Movements', NBER Working Paper No. 3717.

BLADES, D. W. and MARCZEWSKI, W. (1974). 'National Accounts Priorities for Low Income Countries from a Statistician's Point of View', in OECD Development Centre, *National Accounts and Development Planning in Low Income Countries*, Proceedings of a Study Session, Paris 1972, OECD, Paris.

CHEUNG, Y.-W. (1994). 'Aggregate Output Dynamics in the 20th Century', *Economics Letters*, **45**, 15–22.

CHEUNG, Y.-W. and LAI, K. S. (1995). 'Lag Order and Critical Values of the Augmented Dickey–Fuller Test', *Journal of Business and Economics Statistics*, **13**, 277–80.

CHEUNG, Y.-W., LAI, K. S., and TRAN, T. A. (1994). 'Finite-Sample Critical Values of the KPSS Test: A Response Surface Approach', unpublished typescript, Department of Economics, University of California, Santa Cruz.

CHRISTIANO, L. J. and EICHENBAUM, M. (1990). 'Unit Roots in Real GNP: Do We Know, and Do We Care?' in K. Brunner and A. Meltzer (eds), *Carnegie-Rochester Conference Series on Public Policy*, **32**, 7–62.

COGLEY, T. W. (1990). 'International Evidence on the Size of the Random Walk in Output', *Journal of Political Economy*, **98**, 501–18.

DE HAAN, J. and ZELHORST, D. (1993). 'Does Output have a Unit Root? New International Evidence', *Applied Economics*, **25**, 953–60.

DIEBOLD, F. X. and RUDEBUSCH, G. D. (1989). 'Long Memory and Persistence in Aggregate Output', *Journal of Monetary Economics*, **24**, 189–209.

GEWEKE, J. and PORTER-HUDAK, S. (1983). 'The Estimation and Application of Long Memory Time Series Models', *Journal of Time Series Analysis*, **4**, 221–38.

GRANGER, C. W. J. and NEWBOLD, P. (1974). 'Spurious Regressions in Econometrics', *Journal of Econometrics*, **2**, 111–20.

HALL, A. R. (1994). 'Testing for a Unit Root in Time Series with Pretest Data-Based Model Selection', *Journal of Business and Economics Statistics*, **12**, 461–70.

HARRIS, R. I. D. (1992). 'Small Sample Testing for Unit Roots', *Oxford Bulletin of Economics and Statistics*, **54**, 615–25.

KORMENDI, R. C. and MEGUIRE, P. (1990). 'A Multicountry Characterization of the Nonstationarity of Aggregate Output', *Journal of Money, Credit, and Banking*, **22**, 77–93.

KWIATKOWSKI, D., PHILLIPS, P. C. B., SCHMIDT, P. J., and SHIN, Y. (1992). 'Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root?', *Journal of Econometrics*, **54**, 159–78.

LEYBOURNE, S. J. and MCCABE, B. P. M. (1994). 'A Consistent Test for a Unit Root', *Journal of Business & Economic Statistics*, **12**, 157–66.

PERRON, P. (1990). 'Testing for a Unit Root in a Time Series with a Changing Mean', *Journal of Business and Economic Statistics*, **8**, 153–62.

PERRON, P. (1989a). 'Testing for a Random Walk: A Simulation Experiment When the Sampling Interval Is Varied', in B. Raj (ed.), *Advances in Econometrics and Modeling*, Kluwer Academic Publishers, Dordrecht and Boston.

PERRON, P. (1989b). 'The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis', *Econometrica*, **57**, 1361–401.

RIEZMAN, R. G. and WHITEMAN, C. H. (1990). 'Worldwide Persistence, Business Cycles, and Economic Growth', unpublished typescript, California Institute of Technology, Pasadena, CA.

RUDEBUSCH, G. D. (1993). 'The Uncertain Unit Root in Real GNP', *American Economic Review*, **83**, 264–72.

SHILLER, R. J. and PERRON, P. (1985). 'Testing the Random Walk Hypothesis: Power versus Frequency of Observation', *Economics Letters*, **18**, 381–6.

SUMMERS, R. and HESTON, A. (1991). 'The Penn World Table (Mark 5): An Expanded Set of International Comparisons', *Quarterly Journal of Economics*, **106**, 327–68.

UN RESEARCH INSTITUTE FOR SOCIAL DEVELOPMENT (1977). *Research Data Bank of Development Indicators: Volume IV Notes on the Indicators*, Report 77.2, United Nations, Geneva.

WORLD BANK (1989). *World Development Report, 1989*, Oxford University Press, Oxford.

ZELHORST, D. and DE HAAN, J. (1994a). 'The Nonstationarity of Aggregate output: Some Additional Evidence', *Journal of Money, Credit and Banking*, **26**, 9–22.

ZELHORST, D. and DE HAAN, J. (1994b). 'Testing for a Break in Output: New International Evidence', *Oxford Economic Papers*, **47**, 357–62.

APPENDIX 1

1. Detailed unit root test results

TABLE A 1

Series	*NOB*	*QLT*	*ALAG*	*ASTAT*	*AA95*	*AF95*	*L8*	*KA895*	*KF895*	*AKA895*	*AKF895*
Afghanistan	26	D	2	−2.836	A	A	0.086	A	A	AA	AA
Algeria	29	D	6	1.025	A	A	0.118	A	A	AA	AA
Angola	26	D	8	−1.933	A	A	0.123	A	A	AA	AA
Argentina	39	C	8	−3.187	A	A	0.154	R	R	AR	AR
Australia	39	A−	6	−1.884	A	A	0.141	A	R	AA	AR
Austria	39	A−	4	−1.259	A	A	0.178	R	R	AR	AR
Bahamas	11	D	4	−4.747	R	R	0.227	R	A	RR	RA
Bahrain	14	D	5	0.494	A	A	0.146	R	A	AR	AA
Banglades	27	C−	1	−2.267	A	A	0.105	A	A	AA	AA
Barbado	26	C	7	0.103	A	A	0.145	A	A	AA	AA
Belgium	39	A	1	−0.683	A	A	0.129	A	A	AA	AA
Benin	30	D+	2	−2.795	A	A	0.073	A	A	AA	AA
Bolivia	39	C	5	−2.151	A	A	0.107	A	A	AA	AA
Botswana	27	C	8	−2.208	A	A	0.091	A	A	AA	AA
Brazil	38	C−	1	−1.162	A	A	0.128	A	A	AA	AA
Burkina Fasco	24	D	8	−2.498	A	A	0.113	A	A	AA	AA
Burma	36	D	5	−2.726	A	A	0.088	A	A	AA	AA
Burundi	29	D	5	−4.227	R	R	0.105	A	A	RA	RA
Cameroon	29	C−	6	−3.120	A	A	0.090	A	A	AA	AA
Canada	39	A−	8	−3.594	R	R	0.097	A	A	RA	RA
Cape Verde Is	26	D	3	−2.662	A	A	0.094	A	A	AA	AA
Centra Africa Rep	29	D	1	−1.744	A	A	0.135	A	A	AA	AA
Chad	26	D	1	−2.163	A	A	0.129	A	A	AA	AA
Chile	39	C	2	−2.700	A	A	0.154	R	R	AR	AR

continued

TABLE A 1—*continued*

Series	*NOB*	*QLT*	*ALAG*	*ASTAT*	*AA95*	*AF95*	*L8*	*KA895*	*KF895*	*AKA895*	*AKF895*
China	28	D	2	−3.107	A	A	0.142	A	A	AA	AA
Columbia	39	C	6	−2.758	A	A	0.104	A	A	AA	AA
Congo	27	D+	8	−4.647	R	R	0.064	A	A	RA	RA
Costa Rica	39	C	8	−1.575	A	A	0.137	A	A	AA	AA
Cyprus	39	C	8	−3.781	R	R	0.057	A	A	RA	RA
Denmark	39	A−	8	−2.269	A	A	0.152	R	R	AR	AR
Dominican Rep	39	C	1	−1.773	A	A	0.097	A	A	AA	AA
Ecuador	39	C	3	−2.214	A	A	0.085	A	A	AA	AA
Egypt	39	D+	1	−2.393	A	A	0.135	A	A	AA	AA
El Salvador	39	C	2	−1.844	A	A	0.153	R	R	AR	AR
Ethiopia	37	D+	4	−2.722	A	A	0.076	A	A	AA	AA
Finland	38	A−	7	−0.983	A	A	0.141	A	R	AA	AR
France	39	A	4	−0.765	A	A	0.164	R	R	AR	AR
Fuji	28	D	1	−0.276	A	A	0.120	A	A	AA	AA
Gabon	26	D	8	−2.204	A	A	0.126	A	A	AA	AA
Gambia	26	D	5	0.007	A	A	0.152	R	R	AR	AR
Germany	39	A	7	−2.473	A	A	0.181	R	R	AR	AR
Ghana	34	D	6	−0.776	A	A	0.157	R	R	AR	AR
Greece	39	A−	1	0.404	A	A	0.157	R	R	AR	ZR
Guatemala	39	C	8	−2.581	A	A	0.114	A	A	AA	AA
Guinea	27	D	1	−2.418	A	A	0.142	A	A	AA	AA
Guinea Bissau	29	D	7	−3.081	A	A	0.133	A	A	AA	AA
Guyana	39	D	1	−1.376	A	A	0.121	A	A	AA	AA
Haiti	29	D	1	−2.149	A	A	0.090	A	A	AA	AA
Honduras	39	C	5	−2.856	A	A	0.090	A	A	AA	AA
Hong Kong	29	B−	2	−4.406	R	R	0.091	A	A	RA	RA
Hungary	19	B	8	−12.736	R	R	0.136	A	A	RA	RA
Iceland	39	B+	4	−2.217	A	A	0.096	A	A	AA	AA
India	39	C	1	−2.648	A	A	0.111	A	A	AA	AA

Indonesia	27	C	2	−1.460	A	A	0.102	A	A	AA	AA
Iran	31	C−	2	−2.213	A	A	0.121	A	A	AA	AA
Iraq	33	D	7	−4.060	R	R	0.092	A	A	RA	RA
Ireland	39	A−	2	−0.917	A	A	0.139	A	A	AA	AA
Israel	36	B	1	−1.308	A	A	0.165	R	R	AR	AR
Italy	39	A	1	−1.240	A	A	0.182	R	R	AR	AR
Ivory Coast	29	C−	1	0.360	A	A	0.160	R	R	AR	AR
Jamaica	35	C	8	−3.250	A	A	0.162	R	R	AR	AR
Japan	39	A	2	−0.294	A	A	0.170	R	R	AR	AR
Jordan	35	D	6	−3.355	A	A	0.083	A	A	AA	AA
Kenya	39	C	2	−1.737	A	A	0.092	A	A	AA	AA
Korea Rep. of	36	B−	6	−4.140	R	R	0.105	A	A	RA	RA
Kuwait	27	D	1	−1.754	A	A	0.151	R	R	AR	AR
Lesotho	26	D	8	−4.359	R	R	0.077	A	A	RA	RA
Liberia	27	D	5	−0.153	A	A	0.144	A	A	AA	AA
Luxembourg	39	A−	2	−3.153	A	A	0.080	A	A	AA	AA
Madagascar	29	D+	1	−0.942	A	A	0.157	R	R	AR	AR
Malawi	35	D+	1	−1.315	A	A	0.151	R	R	AR	AR
Malaysia	34	C	2	−3.047	A	A	0.089	A	A	AA	AA
Malta	35	C	7	−3.528	R	R	0.110	A	A	RA	RA
Mauritania	29	D	1	−1.560	A	A	0.153	R	R	AR	AR
Mauritius	39	D+	1	−1.430	A	A	0.160	R	R	AR	AR
Mexico	39	C	5	0.120	A	A	0.128	A	A	AA	AA
Morocco	39	C−	3	−3.475	R	A	0.084	A	A	RA	AA
Mozambique	29	D	2	−2.337	A	A	0.149	R	R	AR	AR
Nepal	26	D+	1	−2.526	A	A	0.116	A	A	AA	AA
Netherland	39	A	3	0.235	A	A	0.162	R	R	AR	AR
New Zealand	39	A−	2	−1.411	A	A	0.166	R	R	AR	AR
Nicaragua	37	D	1	−1.541	A	A	0.166	R	R	AR	AR
Niger	29	D	1	−2.163	A	A	0.134	A	A	AA	AA
Nigeria	39	D+	2	−2.163	A	A	0.094	A	A	AA	AA

continued

TABLE A 1—*continued*

Series	*NOB*	*QLT*	*ALAG*	*ASTAT*	*AA95*	*AF95*	*L8*	*KA895*	*KF895*	*AKA895*	*AKF895*
Norway	39	A−	4	−3.065	A	A	0.135	A	A	AA	AA
Pakistan	39	C−	6	−3.841	R	R	0.082	A	A	RA	RA
Panama	37	C	2	−0.801	A	A	0.128	A	A	AA	AA
Papua New Guinea	29	D	8	−3.833	R	R	0.155	R	R	RR	RR
Paraguay	39	C	1	−2.226	A	A	0.142	A	R	AA	AR
Peru	39	C	4	0.352	A	A	0.170	R	R	AR	AR
Philippines	39	C	2	−3.601	R	R	0.137	A	A	RA	RA
Portugal	39	A−	1	−0.885	A	A	0.130	A	A	AA	AA
Rwanda	29	D+	6	−2.636	A	A	0.084	A	A	AA	AA
Saudi Arabia	26	D	8	2.092	A	A	0.157	R	R	AR	AR
Senegal	29	C−	1	−3.414	R	A	0.068	A	A	RA	AA
Sierra Leone	28	D+	1	−1.529	A	A	0.172	R	R	AR	AR
Singapore	26	C	8	1.372	A	A	0.079	A	A	AA	AA
Somalia	28	D	1	−2.617	A	A	0.093	A	A	AA	AA
South Africa	39	C−	1	−1.129	A	A	0.146	R	R	AR	AR
Spain	38	A−	8	−1.764	A	A	0.163	R	R	AR	AR
Sri Lanka	38	C−	2	−0.765	A	A	0.140	A	A	AA	AA
St. Lucia	11	D	4	−2.238	A	A	0.160	R	A	AR	AA
St. Vincent	10	D	1	−3.495	R	A	0.241	R	A	RR	AA
Sudan	34	D	1	−2.407	A	A	0.129	A	A	AA	AA
Suriname	26	D	2	−1.401	A	A	0.128	A	A	AA	AA
Swaziland	26	D+	8	−1.196	A	A	0.164	R	R	AR	AR
Sweden	39	A−	6	−2.360	A	A	0.162	R	R	AR	AR
Switzerland	39	B+	3	−1.911	A	A	0.170	R	R	AR	AR
Syrian Arab Emir	29	C−	8	−2.071	A	A	0.113	A	A	AA	AA
Taiwan	39	D−	6	−3.309	A	A	0.119	A	A	AA	AA
Tanzania	29	C−	4	−2.041	A	A	0.158	R	R	AR	AR
Thailand	39	C−	6	−3.131	A	A	0.097	A	A	AA	AA
Trinidad	39	C	2	−0.241	A	A	0.134	A	A	AA	AA
Tunisia	29	C−	1	0.096	A	A	0.110	A	A	AA	AA
Turkey	39	C	8	−0.856	A	A	0.104	A	A	AA	AA
U. Arab Emirates	16	D	1	−2.547	A	A	0.179	R	A	AR	AA

Uganda	36	D	5	−4.290	R	R	0.096	A	A	RA	RA
United Kingdom	39	A	3	−2.767	A	A	0.130	A	A	AA	AA
United States	39	A	1	−2.378	A	A	0.079	A	A	AA	AA
Uruguay	39	C−	4	−2.821	A	A	0.089	A	A	AA	AA
Venezuela	38	C	1	−0.703	A	A	0.096	A	A	AA	AA
Yemen	20	D	1	−0.643	A	A	0.140	A	A	AA	AA
Yugoslavia	28	B	1	−0.289	A	A	0.155	R	R	AR	AR
Zaire	39	D	1	−1.971	A	A	0.157	R	R	AR	AR
Zambia	34	D+	4	−1.493	A	A	0.147	R	R	AR	AR
Zimbabwe	35	C−	1	−1.409	A	A	0.092	A	A	AA	AA

KA895 / AA95	A	R
A	69 0.55%	39 0.31%
R	15 0.11%	3 0.02%

KA895 / AA95	A	R
A	72 0.57%	39 0.31%
R	14 0.11%	1 0.01%

Notes:
NOB is the number of observations. QLT is the data quality assigned by Summers and Heston (1991). ALAG is the lag length chosen by the AIC. ASTAT is the ADF statistic. AA95 and AF95 refer to whether the ADF test rejects at the 5% significance level, using either the standard critical values, controlling for only sample size (AA95), or the critical values controlling for both sample size and lag structure (AF95). An A entry denotes failure to reject, and R entry denotes rejection of the null. *l*8 denotes the KPSS statistic using the *l*8 rule. The next two columns denote failure to reject, or rejection, based on the KPSS *l*8 rule and asymptotic (KA895) or finite sample (KF895) critical values.

The next two columns present summaries of combination of results for the ADF and KPSS tests; for instance AR denotes failure to reject the ADF, and rejection of the KPSS, tests. AKA895 refers to standard ADF/asymptotic KPSS critical values and KPSS *l*8 rule; AKF895 refers to ADF critical values controlling for sample size and lag structure, finite sample KPSS critical values and KPSS *l*8 rule.

More detailed results, reporting the lag length, data quality, etc., are available from the authors

APPENDIX 2

1. Patterns in test results using restricted samples

TABLE A 2.1

Results for all countries

OBS >= 20, N = 120

KPSS

		A	R	
ADF	A	68 0.57	39 0.33	107 0.89
	R	12 0.10	1 0.01	13 0.11
		80 0.67	40 0.33	120 1.00

OBS >= 30, N = 75

KPSS

		A	R	
ADF	A	39 0.52	28 0.37	67 0.89
	R	8 0.11	0 0.00	8 0.11
		47 0.63	28 0.37	75 1.00

TABLE A 2.2

Results by geographical region

African countries

OBS >= 20, N = 43

KPSS

		A	R	
ADF	A	25 0.58	14 0.33	39 0.91
	R	4 0.09	0 0.00	4 0.09
		29 0.67	14 0.33	43 1.00

OBS >= 30, N = 15

KPSS

		A	R	
ADF	A	8 0.53	6 0.40	14 0.93
	R	1 0.07	0 0.00	1 0.07
		9 0.60	6 0.40	15 1.00

Continued

TABLE A 2.2—*continued*

Asian countries

OBS >= 20, N = 24
KPSS

		A	R	
ADF	A	15 0.63	4 0.17	19 0.79
	R	5 0.21	0 0.00	5 0.21
		20 0.83	4 0.17	24 1.00

OBS >= 30, N = 14
KPSS

		A	R	
ADF	A	8 0.57	2 0.14	10 0.71
	R	4 0.29	0 0.00	4 0.29
		12 0.86	2 0.14	14 1.00

European countries

OBS >= 20, N = 22
KPSS

		A	R	
ADF	A	8 0.36	12 0.55	20 0.91
	R	2 0.09	0 0.00	2 0.09
		10 0.45	12 0.55	22 1.00

OBS >= 30, N = 21
KPSS

		A	R	
ADF	A	8 0.38	11 0.52	19 0.90
	R	2 0.10	0 0.00	2 0.10
		10 0.48	11 0.52	21 1.00

North American countries

OBS >= 20, N = 12
KPSS

		A	R	
ADF	A	10 0.71	3 0.21	13 0.93
	R	1 0.07	0 0.00	1 0.07
		11 0.79	3 0.21	14 1.00

OBS >= 30, N = 12
KPSS

		A	R	
ADF	A	8 0.67	3 0.25	11 0.92
	R	1 0.08	0 0.00	1 0.08
		9 0.75	3 0.25	12 1.00

TABLE A 2.2—*continued*

Oceania countries

OBS > = 20, N = 5

		KPSS A	*KPSS* R	
ADF	A	2 0.40	2 0.40	4 0.80
	R	0 0.00	1 0.20	1 0.20
		2 0.40	3 0.60	5 1.00

OBS > = 30, N = 2

		KPSS A	*KPSS* R	
ADF	A	0 0.00	2 1.00	2 1.00
	R	0 0.00	0 0.00	0 0.00
		0 0.00	2 1.00	2 1.00

South American countries

OBS > = 20, N = 12

		KPSS A	*KPSS* R	
ADF	A	8 0.67	4 0.33	12 1.00
	R	0 0.00	0 0.00	0 0.00
		8 0.67	4 0.33	12 1.00

OBS > = 30, N = 11

		KPSS A	*KPSS* R	
ADF	A	7 0.64	4 0.36	11 1.00
	R	0 0.00	0 0.00	0 0.00
		7 0.64	4 0.36	11 1.00

TABLE A 2.3
Results by income level

Low income countries

OBS >= 20, N = 38
KPSS

		A	R	
ADF	A	25 0.66	9 0.24	34 0.89
	R	4 0.11	0 0.00	4 0.11
		29 0.76	9 0.24	38 1.00

OBS >= 30, N = 14
KPSS

		A	R	
ADF	A	8 0.57	4 0.29	12 0.86
	R	2 0.14	0 0.00	2 0.14
		10 0.71	4 0.29	14 1.00

Middle income countries

OBS >= 20, N = 56
KPSS

		A	R	
ADF	A	35 0.63	14 0.25	49 0.88
	R	6 0.11	1 0.02	7 0.13
		41 0.73	15 0.27	56 1.00

OBS >= 30, N = 39
KPSS

		A	R	
ADF	A	24 0.62	10 0.26	34 0.87
	R	5 0.13	0 0.00	5 0.13
		29 0.74	10 0.26	39 1.00

High income countries

OBS >= 20, N = 26
KPSS

		A	R	
ADF	A	8 0.31	16 0.62	24 0.92
	R	2 0.08	0 0.00	2 0.08
		10 0.38	16 0.62	26 1.00

OBS >= 30, N = 22
KPSS

		A	R	
ADF	A	7 0.32	14 0.64	21 0.95
	R	1 0.05	0 0.00	1 0.05
		8 0.36	14 0.64	22 1.00

TABLE A 2.5
Results by data quality

D and C quality data

OBS >= 20, N = 93
KPSS

		A	R	
ADF	A	60 0.65	23 0.25	83 0.89
	R	9 0.10	1 0.01	10 0.11
		69 0.74	24 0.26	93 1.00

OBS >= 30, N = 50
KPSS

		A	R	
ADF	A	31 0.62	13 0.26	44 0.88
	R	6 0.12	0 0.00	6 0.12
		37 0.74	13 0.26	50 1.00

B and A quality data

OBS >= 20, N = 27
KPSS

		A	R	
ADF	A	8 0.30	16 0.59	24 0.89
	R	3 0.11	0 0.00	3 0.11
		11 0.41	16 0.59	27 1.00

OBS >= 30, N = 25
KPSS

		A	R	
ADF	A	8 0.32	15 0.60	23 0.92
	R	2 0.08	0 0.00	2 0.08
		10 0.40	15 0.60	25 1.00

Notes:
These tables provide information analogous to that provided in Table 1–4, for restricted samples. On the left-hand side are tables for samples restricted to series with 20 and above observations; on the right-hand side, those with 30 and above.

APPENDIX 3

1. Detailed structural break test results

TABLE A 3

Unit root sequential test—*GDP*

Country	*AICLAG*	*GSMXF*	*GSYR*	*GSFDF*	*GSMID*	*GBMXF*	*GBYR*	*GBFDF*	*GBMID*
Argentina	8	12.9095	1980	−4.1417	−4.1417	13.2939	1964	−2.9518	−2.9518
Australia	6	8.3415	1971	−3.539	−3.539	8.0787	1962	−2.5759	−3.0971
Austria	4	14.6348	1975	−4.1185	−4.1185	6.6749	1981	−2.8466	−2.8466
Chile	2	8.1827	1971	−4.1211	−4.1211	8.708	1975	−4.1887	−4.1887
Denmark	8	12.3923	1973	−4.0243	−4.0372	3.4876	1964	−3.0284	−3.0284
El Salvador	2	16.3598	1978	−4.6185	−4.6185	22.9572	1979	−4.4379	−4.6527
Finland	7	20.3863	1977	−4.664	−4.664	8.2733	1969	−2.5122	−2.7671
France	4	15.2312	1973	−3.8609	−3.9399	8.1038	1974	0.6071	−2.1892
Gambia	5	*26.9081	1978	*−5.0643	*−5.0643	12.5173	1971	−2.5658	−2.5658
Germany	7	6.0328	1974	−3.2902	−3.2902	8.5615	1968	−3.9284	−4.0451
Ghana	6	18.6205	1975	−4.4217	−4.4217	10.3851	1975	−0.0541	−2.4832
Greece	1	22.3283	1973	−3.5859	−3.8189	20.5443	1974	1.2889	−1.3324
Israel	1	17.3268	1974	−4.3249	−4.3249	15.3684	1968	−2.906	−2.906
Italy	1	9.5269	1970	−3.3647	−3.3647	7.3011	1959	−2.7897	−2.8382
Ivory Coast	1	22.8796	1979	−4.4411	−4.4411	6.0854	1970	−1.1941	−1.1941
Jamaica	8	*25.3984	1973	*−4.8829	*−4.9152	5.4956	1973	0.9635	−1.2601
Japan	1	24.5732	1971	−4.7703	−4.7703	14.8099	1974	−1.4192	−2.2932
Kuwait	1	10.2102	1973	−3.7166	−3.7166	16.4108	1980	−4.3057	−4.3057
Madagascar	1	13.9728	1972	−3.9372	−3.9372	8.56	1968	−3.0672	−3.0672
Malawi	1	14.7578	1976	−4.1866	−4.1866	15.9011	1965	−3.3702	−3.3702
Mauritania	1	14.3673	1977	−4.158	−4.158	7.0268	1978	−1.8199	−2.8463
Mauritius	1	4.5058	1969	−2.7395	−2.7395	13.1197	1974	−3.8868	−3.8868

TABLE A 3—*continued*

				Unit root sequential test—*GDP*					
Country	*AICLAG*	*GSMXF*	*GSYR*	*GSFDF*	*GSMID*	*GBMXF*	*GBYR*	*GBFDF*	*GBMID*
Mozambique	2	11.72	1972	−4.2974	−4.2974	7.6824	1974	−2.5777	−3.3124
Netherland	3	4.755	1975	−3.3822	−3.3935	7.2489	1981	−3.8786	−3.8786
New Zealand	2	11.3243	1974	−3.7241	−3.7241	11.1965	1960	−3.4771	−3.4771
Nicaragua	1	15.7618	1975	−4.1839	−4.1839	17.4653	1978	−3.9939	−4.0262
Paraguay	1	3.3689	1981	−1.5282	−2.3797	14.5553	1982	0.5417	−3.4124
Peru	4	11.9879	1975	−3.112	−3.1194	8.1689	1962	−1.4674	−1.4674
Saudi Arab	8	9.1905	1975	3.1327	−1.2994	7.0408	1976	3.7385	0.5138
Sierra Leone	1	9.5262	1974	−3.6098	−3.6098	9.699	1969	−3.5276	−3.5276
South Africa	1	11.756	1975	−3.6505	−3.6505	10.2498	1982	−2.9245	−2.9245
Spain	8	*32.0704	1975	*−6.2887	*−6.2887	5.5051	1963	−2.77	−2.77
Swaziland	8	10.0747	1980	−3.5056	−3.5056	8.7379	1973	−3.3246	−3.3246
Sweden	6	9.2121	1973	−3.6068	−3.6251	8.7492	1961	−3.9471	−3.9471
Switzerland	3	11.3837	1971	−3.9428	−3.9428	14.1738	1975	−4.0615	−4.0615
Tanzania	4	6.0524	1981	−3.3596	−3.3596	10.9759	1983	−4.0629	−4.0629
Yugoslavia	1	*30.0643	1980	*−5.3081	*−5.3081	6.0396	1981	−1.7207	−1.9027
Zaire	1	5.1789	1973	−2.8488	−2.8488	18.5238	1977	−4.6799	−4.6799
Zambia	4	9.0666	1974	−3.4302	−3.431	11.0658	1976	−2.0529	−3.2381

Statistics with * are significant at the 0.05 level.

Notes:
AICLAG is the lag length chosen by the AIC.
GSMXF is the maximum *F* statistic for a shift in the trend (change in slope)
GSYR is the indicated break year.
GSFDF is the ADF statistic assuming a break at GSYR.
GSMID is the minimum ADF statistic.

GBMXF is the maximum *F* statistic for a shift in the mean of the trend.
GBYR is the indicated year break.
GBFDF is the ADF statistic assuming a break at GBYR.
GBMID is the minimum ADF statistic.