Empirical Economics, 303-319, February 2003

UNIT ROOTS, POSTWAR SLOWDOWNS AND LONG-RUN GROWTH: EVIDENCE FROM TWO STRUCTURAL BREAKS

Dan Ben-David
Tel-Aviv University, NBER and CEPR

Robin L. Lumsdaine
Deutsche Bank and NBER

and

David H. Papell
University of Houston

ABSTRACT

This paper provides evidence on the unit root hypothesis and long-term growth by allowing for two structural breaks. We reject the unit root hypothesis for three-quarters of the countries – approximately 50% more rejections than in models that allow for only one break. While about half of the countries exhibit slowdowns following their postwar breaks, the others have grown along paths that have become steeper over the past 120 years. The majority of the countries, including most of the slowdown countries, exhibit faster growth after their second breaks than during the decades preceding their first breaks.

Correspondence:

Dan Ben-David
School of Government and Policy
Tel-Aviv University
Tel-Aviv, Israel 69978

Tel: 972 3 640-9912
Fax: 972 3 640-7382
E-Mail: danib@post.tau.ac.il

I. INTRODUCTION

The issue of unit roots in long-term output has been a matter of controversy ever since Nelson and Plosser (1982) – using annual data for the United States – could not reject the unit root hypothesis with an Augmented-Dickey-Fuller test for either aggregate or per capita GNP. A major challenge to the findings of Nelson and Plosser was mounted by Perron (1989). Noting that structural change becomes more likely with long spans of data, Perron demonstrates that observed "unit root behavior" could be due to the failure to account for structural change, resulting in misspecification of the deterministic trend. Perron argues that most macroeconomic series are not characterized by a unit root, but rather that persistence of shocks is limited to a few rare events while most shocks are transitory. Allowing a single change in the trend function after 1929, he finds that most U.S. macroeconomic variables, including aggregate and per capita real GNP, are trend stationary.[1]

Subsequent deterministic trend literature (e.g. Banerjee, Lumsdaine, and Stock (1992), Christiano (1992), and Zivot and Andrews (1992)) has treated the break date as unknown *a priori*. Using the same Nelson-Plosser data as Perron, Zivot and Andrews (1992) use a sequential Dickey-Fuller test to endogenously determine the break date and find evidence against the unit root hypothesis for fewer series. For aggregate and per capita real GNP, however, Zivot and Andrews, like Perron, reject the unit root null in favor of a trend stationary alternative with a single break.

Lumsdaine and Papell (1997) re-examine the unit root hypothesis for Nelson-Plosser data against a two-break trend stationary alternative. In general, they find evidence against unit roots

[1] An alternative characterization to the deterministic shifting trend model of Perron is the stochastic trend model, which characterizes infrequent shocks as occurring randomly. In the limit (as the number of shocks increases), this corresponds to the limit of the model considered in this paper, namely allowing for a structural change in each period. Balke and Fomby (1991) show that standard time series methods cannot distinguish between a data generating process driven by large, infrequent permanent shocks versus one in which permanent shocks arise frequently but are small in magnitude. Using outlier identification techniques, however, they find some evidence of infrequent, permanent shocks in US GNP.

for more series than Zivot and Andrews (1992), but fewer series than Perron (1989). In the case of output, however, they find results consistent with the one-break tests of both Perron and Zivot and Andrews: namely, rejection of the unit root hypothesis for aggregate and per capita real GNP. Thus, regardless of one's belief of how many breaks there were, or whether they should be endogenously determined, accounting for structural change suggests that both aggregate and per capita real GNP in the United States are trend stationary.

The evidence becomes less clearcut when the analysis is extended beyond the United States. To date, only one-break tests (allowing for a one-time change in both the intercept and the slope of the trend function) have been used to test the unit root hypothesis for the output of countries other than the U.S. – and the results from these tests have been mixed. Raj (1992), using per capita real GDP for 9 countries, Perron (1994), using aggregate real GDP for 11 countries, and Ben-David and Papell (1995), using both aggregate and per capita real GDP for 16 countries, all reject the unit root hypothesis for about half of the countries.

The international analysis of long-term GDP is broadened in this paper by testing for unit roots while allowing for multiple, specifically two, trend breaks. For each break date, we allow either a change in the intercept or a change in both the intercept and the slope of the trend function. Just as failure to allow one break can cause non-rejection of the unit root null by the Augmented-Dickey-Fuller test, failure to allow for two breaks, if they exist, can cause non-rejection of the unit root null by the tests which only incorporate one break.

By allowing for two breaks, we show that it is possible to reject the unit root hypothesis, using long-term Maddison (1991) aggregate and per capita GDP for 16 countries, in three-quarters of the cases. Put differently, we can reject the unit root null for half of the cases in which Ben-David and Papell (1995) are unable to reject. The rejections are much stronger when changes in the slope, as well as the intercept, are allowed. It is important to understand that these rejections are in favor of a regime-wise trend stationary alternative with two breaks in the intercept and slope of the trend function, not in favor of a trend stationary alternative without breaks. The "permanent" (in the context of the span of the data) changes in the slope of output correspond to "permanent" changes in output growth.

What are the implications of multiple structural breaks on the long-run growth paths of countries? Models that allow just one structural break face a tradeoff. Studies which use long-term data, such as Ben-David and Papell (1995), cannot determine if, and when, postwar slowdowns occurred because the breaks are dominated by the World Wars and the onset of the Great Depression. Studies which use postwar data, as in Ben-David and Papell (1998), are unable to gauge the magnitude of these slowdowns from a long-run perspective.

Allowing for two breaks provides us with a model that is able to address both aspects of the tradeoff. While all of the countries have at least *one* break which is associated with a World War or the Depression, over half turn out to have a break which occurs in 1955 or later – an outcome that facilitates an analysis of postwar slowdowns within a long-run context. The results, however, do not support the view that postwar growth slowdowns were caused by the first oil shock. Only four of the breaks are associated with the 1973-75 period, and the others occur earlier. This is consistent with the recent findings of Bai, Lumsdaine, and Stock (1998). Considering quarterly US postwar consumption, output, and investment, they identify a common structural break in the first quarter of 1969 with a 90% confidence interval around the break date of (66:2,71:4).[2]

When the countries are viewed from such a long-run perspective, we find that the large majority of them exhibited faster growth in the latter years of the sample than during the early years. This is true even for countries that experienced slowdowns. The finding of increasing growth over the long run is consistent with the increasing growth predictions of Romer (1986) and many of the subsequent endogenous growth models. While we consider this to be a natural interpretation of our findings, they are also consistent with neoclassical or "semi-endogenous"

[2] Based on the rejections of unit roots in this (and other) papers, it would be possible to use the methods in Bai (1999) and Bai and Perron (1998,2000) to conduct a more extensive analysis of multiple structural changes in long-term GDP. Such an analysis, however, is beyond the scope of this paper.

growth models, such as Solow (1956) or Jones (1995), if the structural breaks are caused by exogenous rather than endogenous changes in technological progress. In addition, with a finite amount of data, univariate models of output cannot distinguish between changes in steady-state growth rates and arbitrarily long transitional dynamics.

The paper is organized as follows: the two-break sequential Dickey-Fuller tests are developed and used to investigate the unit root hypothesis for long-term aggregate and per capita GDP in Section II. An interpretation of the results, within a growth context, is provided in Section III. Conclusions are presented in Section IV.

II. TWO BREAK UNIT ROOT TESTS

This section describes sequential Dickey-Fuller tests for a unit root, allowing for two shifts in the deterministic trend at two distinct unknown dates. The distributional theory underlying these tests is developed and described in more detail in Lumsdaine and Papell (1997).[3]

Augmented-Dickey-Fuller tests which incorporate two breaks involve regressions of the following form,

(1) $$\Delta y_t = \mu + \beta t + \theta_1 DU1_t + \gamma_1 DT1_t + \theta_2 DU2_t + \gamma_2 DT2_t + \alpha y_{t-1} + C(L)\Delta y_{t-1} + \varepsilon_t \quad .$$

[3] Alternative tests for unit roots in the presence of multiple breaks have been developed by Park and Sung (1994) and Busetti and Harvey (2001). These tests, however, require that the break dates be known *a priori*.

A period at which the change in the parameters of the trend function occurs will be referred to as the time of break, or *TBi* ($i = 1, 2$). The break dummy variables have the following values: $DUi_t = 1$ and $DTi_t = t - TBi$ if $t > TBi$; 0 otherwise. Model (1) is estimated for all possible pairs (*TB1*, *TB2*), where $TBi = 2,...,T\text{-}1$, i = 1,2, and *T* is the number of observations after adjusting for those "lost" by first-differencing and lag length *k*. *C*(*L*) is a lag polynomial of known order *k*. The errors are assumed to satisfy the assumption that ε_t is a martingale difference sequence and satisfies $\mathrm{E}(\varepsilon_t^2 | \varepsilon_{t-1},...) = \sigma^2$, $\mathrm{E}(|\varepsilon_t|^i | \varepsilon_{t-1}, ...) = \kappa_i (i = 3,4)$, $\sup_t \mathrm{E}(|\varepsilon_t|^{4+\xi} | \varepsilon_{t-1}, ...) = \overline{\kappa} < \infty$ for some $\xi > 0$.

For each choice of *TBi*, the value of *k* is selected by the criteria advocated by Campbell and Perron (1991) and Ng and Perron (1995). Start with an upper bound on *k* chosen *a priori*. If the last included lag is significant, choose the upper bound. If not, reduce *k* by one until the last lag becomes significant. If no lags are significant, set $k = 0$. Following Perron (1989), we set the upper bound on *k* to equal 8 and the criterion for significance of the *t*-statistic on the last lag equal to 1.60.

The test is implemented by computing the *t*-statistic for α over all possible distinct pairs of break dates, *TB1* and *TB2*.[4] The null hypothesis, that the series $\{y_t\}$ is an integrated process without an exogenous structural break, is tested against the alternative hypothesis that $\{y_t\}$ is trend stationary with two breaks in the trend function which occur at two distinct unknown dates. Our specification is a two-break extension of Zivot and Andrews (1992). The estimated break dates are the values of *TB1* and *TB2* for which the *t*-statistic for α is minimized. If the minimum *t*-statistic is more negative than the associated critical value, then the null hypothesis is rejected.

Three types of models are estimated. Model AA allows for two breaks in the intercept, but not in the slope, of the trend function, and thus sets $\gamma_1 = \gamma_2 = 0$. Model CA allows for two breaks in the intercept and one break in the slope of the trend function, and thus sets $\gamma_2 = 0$.

[4] Note that the computational burden of the two-break model, relative to the one-break, is exponentially increasing in the sample size. With this approach, the number of computations required for estimation of an *n*-break model is T^n. Determining critical values for higher order models quickly becomes computationally infeasible.

Model CC allows the two breaks to be in both the intercept and in the slope of the trend function.[5] In accord with the literature on one break specifications, we do not attempt to select among the models. Finite sample critical values with the data generated under the null hypothesis, using 125 observations and 5000 replications, are given for each of the three tests in Tables 1-3.[6]

The tests are run on data compiled by Maddison (1991). He provides annual GDP data for 16 countries, starting in 1860 and ending in 1989. Indexes of annual aggregate real GDP (adjusted to exclude the impact of boundary changes) were converted into 1985 U.S. relative prices using OECD purchasing power parity units of national currency per U.S. dollar. Annual per capita GDPs were calculated by dividing the aggregate GDPs by the mid-year population levels. While the aggregate data begins in 1860, the per capita GDP is limited by the population data which begins in 1870.[7]

The results for Model AA, which do not show much evidence against the unit root hypothesis, are presented in Table 1. The unit root null can be rejected at the 5 percent (or higher) significance level for Canada, Germany, and the United States for both aggregate and per capita real GDP but, with the exception of aggregate GDP for the U.K., cannot be rejected for any other country.[8] The results for the United States are consistent with the findings in Lumsdaine and Papell (1997) which are based on a shorter time span of output data.

---

[5] For Models AA and CC, TB1 and TB2 are the years associated with the first and second breaks, respectively. For Model CA, TB1 is the year with both an intercept and a slope break and TB2 is the year with just an intercept break.

[6] The critical values differ slightly from those reported in Lumsdaine and Papell (1997) because, reflecting advances in available computational power, we perform more replications. Each set of critical values took 3-5 days, depending on the model, on a Pentium 400 computer. We could also compute critical values either by using the exact number of observations in each series or by estimating ARMA models for each series, (under the null hypothesis of no break), and using these as the underlying data generating processes. Evidence in Lumsdaine and Papell (1997) suggests that the use of such methods would not affect the results qualitatively, as the rejection levels reported here are quite strong.

[7] Not all countries have aggregate GDP data which goes back to 1860. The exceptions are Austria (1870), Canada (1870), Italy (1861), Japan (1885), Netherlands (1900), Norway (1865), Switzerland (1899), and the United States (1869). The per capita data begins in 1870 except for Japan, Netherlands, and Switzerland.

[8] We will use "rejected" as a shorthand for "rejected at the 5 percent or higher level of significance".

When the model is modified to allow for one break in the slope of the trend function (in addition to the two breaks in the intercept), the estimates of Model CA reported in Table 2 reveal more evidence against the unit root hypothesis, rejecting the null for 8 countries (aggregate) and 11 countries (per capita). The null is rejected for both aggregate and per capita real GDP for 7 countries: Belgium, Canada, Finland, France, Japan, the United Kingdom, and the United States; for aggregate (but not per capita) real GDP for Germany; and for per capita (but not aggregate) real GDP for Australia, Austria, Denmark, and Norway.

The results for Model CC, which allows for 2 breaks in both the slope and the intercept, are presented in Table 3. These show even more evidence against the unit root hypothesis, with rejection of the unit root null in 11 of the aggregate cases and 12 of the per capita cases. Specifically, the null is rejected for *both* aggregate and per capita real GDP for 10 countries: Austria, Belgium, Canada, Denmark, Finland, France, Japan, Norway, the United Kingdom, and the United States; for aggregate (but not per capita) real GDP for Germany; and for per capita (but not aggregate) real GDP for Sweden.

Overall, by allowing for two breaks, we can reject the unit root hypothesis in favor of the (broken) trend stationary alternative in three-quarters – 24 out of 32 – of the cases for at least one of the models. For aggregate real GDP, the unit root null is rejected for 11 countries while, for per capita real GDP, the null is rejected for 13 of the 16 countries.[9]

It is illustrative to compare these results with previous studies. The most direct comparison is with Ben-David and Papell (1995) who, performing endogenous one-break tests which allow a break (at the same date) in both the intercept and the slope of the trend function on

[9] The null is rejected for both aggregate and per capita real GDP for 11 countries: Austria, Belgium, Canada, Denmark, Finland, France, Germany, Japan, Norway, the United Kingdom, and the United States, and for per capita (but not aggregate) real GDP for Australia and Sweden. We cannot reject the unit root hypothesis in either case for Italy, Netherlands, or Switzerland.

identical data, reject the unit root null in just 16 out of 32 cases. Raj (1992) and Perron (1994), using similar data, also reject the unit root null for about half the countries.[10]

Our results illustrate the importance of allowing for breaks in the slope, as well as in the intercept, of the trend function. Model AA, which only incorporates intercept breaks, produces very few rejections of the unit root null. Model CA, which also allows for one slope break, produces additional rejections. Model CC, which also allows for two slope breaks, produces even more rejections. In fact, Model CC produces almost as many rejections of the null as are found in total in the estimation of all three models. Even though the critical values increase (in absolute value) from Model AA to Model CA to Model CC, the number of rejections of the unit root null increases. Unlike in Lumsdaine and Papell (1997), where estimation of Model CC versus Models AA and CA appeared to be associated with a loss of power, the long-term output data seem to exhibit patterns consistent with the possibility of two breaks in both intercept and slope.[11]

Of separate interest are the break dates themselves. In the single break tests of Raj (1992), Perron (1994), and Ben-David and Papell (1995) the breaks are mostly caused by wars. In Ben-David and Papell (1995) for example, World War II is the source of the breaks for Japan and all of the continental European countries. For several countries, Finland, Sweden, and the United Kingdom, the break is caused by World War I and, for those countries not as severely affected by the World Wars, Australia, Canada, and the United States, the breaks occur at the onset of the Great Depression.

---

[10] It should be noted that allowing for additional breaks does not necessarily produce more rejections of the unit root null. The critical values rise (in absolute value), causing a loss of power if too many breaks are included. We therefore view the additional rejections as evidence in favor of the two-break model. It is possible that a three (or more) break model would produce even more rejections, but only if the t-statistic oṇ rises by enough to compensate for the power loss.

[11] The models which we estimate do not allow for breaks in just the slope, but not the intercept, of the trend function, and it is possible that estimating such models would provide more power. The results, in any event, would not be less strong than those reported here.

Allowing for two breaks produces a richer set of results. For Model CC, which accounts for breaks in both the intercept and the trend, one of the two breaks occurs in the same year as the single break found by Ben-David and Papell (1995) in 15 out of the 16 cases where they reject the unit root null.[12] We also find a number of post World War II breaks. Again focusing on Model CC, we find that over half of the countries have one of their breaks in 1955 or later, and a number occur in the 1970s.[13]

For a number of countries, the break dates change considerably across models. For example, with aggregate real GDP, one of the breaks for the United States is always 1929. The other break, however, ranges from 1978 (AA) to 1942 (CA) to 1955 (CC). Since these differences reflect allowing for additional changes in long-term growth, there is no reason *a priori* to expect the break dates to either change or to remain the same.[14]

In order to check the robustness of our empirical findings of post World War II breaks, we perform Zivot and Andrews (1992) tests for a unit root in the presence of a single break on the 1946 - 1989 sub-sample. Using Model C, which allows for a change in both the intercept and the slope, we reject the unit root null (at the 5% level) in 8 cases with aggregate data and 9 cases with per capita data (out of 16 countries). This is about the same as found by Ben-David and Papell (1995) with single break tests on the full sample. While some of the break dates were the same as the post World War II breaks found with the multiple break tests, others were considerably different.[15]

---

[12] The exception is aggregate GDP for Finland, where the break date changes from 1913 to 1916.

[13] We find less evidence of post World War II breaks for Models AA and CA.

[14] While it would be desirable to have a method for choosing between models AA, CA, and CC, we are not aware that any such method exists in the literature.

[15] We thank an anonymous referee for this suggestion. The critical values were calculated for the exact number of observations (54) in the sub-sample.

III. GROWTH IMPLICATIONS

Having identified the potential break dates in the previous section, the emphasis now shifts to an examination of the corresponding changes in per capita output levels and growth rates. Since there are no generally accepted statistical techniques to choose among the models, the choice is necessarily subjective. In general, the strength of the unit root rejections increases between models AA and CA, and either increases or stays the same between models CA and CC. We therefore use model CC but, as described below, explore the implications of other choices.

The timing of the trend breaks can be used to delineate the 16 countries into three groups (see Table 4). The largest of these groups (A) includes the 9 countries that experienced their second trend breaks during the years following the Second World War. All of these countries experienced a post-WWII slowdown. For the majority of these countries (Germany, Austria, Italy, Japan, the Netherlands, Denmark and France), the timing of their first trend breaks coincided with WWII. World War I served as the first trend break for Sweden while the onset of the Great Depression coincided with the first trend break for the United States. Group B includes the 5 countries (Belgium, Norway, Finland, Switzerland, and the United Kingdom) whose two breaks occurred during, or in very close proximity with, the two World Wars. Group C consists of the last two countries, Australia and Canada, who experienced their first breaks prior to World War I and their second breaks between the two World Wars, during the late twenties.

For each country in group A, the first break corresponds with a substantial drop in income levels. The group's average $\hat{\theta}_1$, the coefficient for the first intercept dummy variable, equals -0.252. These sharp drops in income levels were usually followed by much higher growth rates, with a 3.44 ratio of second period to first period average growth rates (the first trend break marks the division between the two periods).

In contrast, the second trend break signals the end of the period of fast growth and marks a return to what might be considered the new long-run growth path. Though the 1970s are usually thought of as the turning point for most countries, this is not as obvious when viewed within the long-run context. Only four of the nine countries experiencing a post-WWII slowdown had their second trend break during the seventies. The five remaining countries experienced their second breaks during the fifties (3 countries) and sixties (2 countries).

The last column in Table 4 provides an indication of the extent of the postwar slowdowns from the long-run perspective. After the post-WWII slowdown, three of the countries returned to the average growth rates that they had exhibited prior to their first breaks. However, the six other countries continued along faster growth paths – with average growth rates that were twice as high as the growth rates prior to their first breaks. In the case of the countries in group B, average third period growth rates were 261 percent of first period rates, while third period growth rates in group C were 131 percent of first period rates.

Another interesting feature of the estimation results is the apparent link between the extent of the shocks to output levels and the extent of the subsequent changes in growth rates during the following period. Define DS_i as the difference between $\hat{\theta}_{1i}$ and $\hat{\theta}_{2i}$ (the coefficients for the first and second intercept dummies for country *i*). This provides a measure of the relative magnitude of the shocks for each country. Define DG_i as the difference between second and third period average growth rates for country *i* (where the second period spans the years between *TB1* and *TB2*, while the third period includes the years following *TB2*).

To what extent is the magnitude of the shocks in levels related to the magnitude of the post-shock growth rates? The correlation coefficient of -0.81 provides an indication of a strong negative relationship between DG_i and DS_i. Figure 1, which plots the two variables, corroborates this with visual evidence that the sharper the drop in levels, the higher the subsequent growth rate. This attribute of the results corresponds to the standard neoclassical growth model prediction that shocks to an economy will be followed by initially faster (than steady state) growth. As the impact of these shocks dissipates over time, output growth should slow down and the economy should return to its original steady state path.

But does output actually return to its original steady state path? Consider the 12 countries for which the unit root null was significantly rejected. Using the coefficients estimated earlier, it is possible to extrapolate the growth paths of the first two periods into the subsequent periods. To the extent that the actual third period growth path of a country is below the extrapolated second period path, a country can be characterized as a slowdown country. Other than the United States, there are 5 countries – France, Japan, Denmark, Sweden, and Austria – that fit this description of a slowdown (we will return to the case of the U.S. in a moment). One of these, France, is depicted in Panel A of Figure 2. Extrapolations of its pre-break paths are plotted (using the coefficients derived in the estimation of Equation 1) for the periods following the estimated break points to better facilitate the growth comparisons.

The two trend breaks found in the estimation of Equation 1 for France are 1939 and 1974. Note that while the onset of World War I coincided with a sharp drop in real per capita output levels, the country rebounded after the war and returned to its old path – as the neoclassical growth model predicts. Thus, in the case of France, none of the years associated with WWI came out as significant break years in the estimation.

The rebound that occurred after the Second World War however, was of a different nature. By the 1950s, the new growth path had already exceeded the old path and the fast growth continued until 1974 – by which time the old path had been left far below the new one. At that point, the economy began to slow down. But as the results in Table 4 indicate, despite the slowdowns, average growth rates in the last period were still 145 percent of those in the first period. This pattern of increasing growth is also found in the cases of Austria, Japan and Sweden.

While slowdown behavior characterizes a number of countries in the sample, it is not the only kind of result borne out by the two-break tests. Remaining with the 12 countries for which the unit root null was rejected, there is another group of countries, this time consisting of 4 countries (Norway, Belgium, the United Kingdom, and Finland) that managed to exhibit faster growth in each of the subsequent periods. Take the case of Norway, for example, a country that is depicted in Panel B of Figure 2. While it too suffered strong negative level shocks, it managed to grow along ever-higher growth paths. This group of countries provides even stronger evidence that growth rates for the majority of countries have been increasing over the long run.

Of the three remaining countries for which the unit root null can be rejected, the United States (depicted in Panel C of Figure 2) and Canada comprise the third category of countries.[16] These two countries, for the most part, appear to have grown along their original growth paths for most of the past 120 years. They did not experience first-hand the cataclysmic shocks of the two World Wars (to the extent that Japan and the European countries did), nor were they subject to major domestic and external institutional changes as were the majority of the other countries. Nonetheless, both countries still ended the postwar period slightly above their original (first period) growth paths.

The United States and Canada also comprise two (out of three) countries for which the unit root null can be rejected with Model AA, which includes intercept, but not slope, changes. While 1929 is still a break year for the United States, the other break changes from 1955 to 1940.[17] Since the unit root null can be rejected in favor of an alternative without slope changes, these results are most naturally interpreted in the context of exogenous or semi-endogenous growth models which feature level, but not growth, effects.[18]

Germany is the other country for which the unit root null can be rejected with Model AA. In contrast with the United States and Canada, the null cannot be rejected once slope changes are allowed in Models CA and CC. With Model AA, the breaks are for 1944 and 1946, reflecting

[16] The last country, Australia, experienced a very sharp drop in income levels in 1891 and, in contrast with the cases described above, did not rebound back to its original path during its second period. Only after the additional sharp fall in levels in 1927 did the country move to a new, steeper, growth path that eventually surpassed the earlier paths.

[17] For Canada, while the first break year changes from 1908 to 1917, the second break, 1928, stays the same, making less of a difference to the results.

[18] In addition, as reported in Ben-David and Papell (1995), the United States is the only country for which the unit root null can be rejected by ADF tests without any form of structural change.

both the destruction of the German economy at the end of the war and its very fast (relative to other Continental European countries) recovery.

IV. CONCLUSIONS

This paper examines the output growth paths of 16 industrialized countries during a span of 120 years. The twin focal points of the paper – unit roots in long-term international output and long-run growth behavior – center around a test of the unit root hypothesis which allows for two endogenously determined structural breaks.

Previous tests of the unit root hypothesis for long-term international output allow for no more than one structural break in the trend process. Accounting for the existence of two breaks, we investigate the unit root hypothesis for aggregate and per capita real GDP. We reject the unit root null for more countries (12 out of 16) than models that allow for only one break.

Inclusion of the two structural breaks also enables us to investigate the long-run growth patterns in the 16 countries. While most of the countries exhibit one break that is associated with a war, over half have a break which occurs in 1955 or later. The estimated breaks are associated with substantial changes in output levels and growth rates. There is a clear negative relationship between the level response to shocks and the subsequent change in growth rates (with sharp drops in income levels followed by high growth rates).

The majority of countries – both those for which the unit root can and cannot be rejected – exhibit substantial increases in their rates of growth over the past 120 years, with third period growth rates roughly double (on average for all 16 countries) the first period growth rates. While we believe that such a pattern of increasing growth is most naturally interpreted as consistent with the predictions of recent endogenous growth models, it can also be interpreted, as described in the introduction, in the context of neoclassical or semi-endogenous growth models.

The United States appears to be an exception to this pattern, with growth rates after 1955 that are only 105 percent of those before 1929. The fact that the U.S. growth path does not appear to be representative of output growth for the other countries illustrates the risks inherent in studies that tend to generalize too much from the experience of the United States in their evaluations of economic growth theories.

This paper also provides evidence which reconciles postwar growth slowdowns with the endogenous growth models' prediction that increasing steady state growth rates are possible. Only about half of the countries exhibit slowdowns following their postwar breaks, and even the majority of slowdown countries exhibit faster growth during their slowdown periods than during the decades preceding their first breaks.

REFERENCES

Bai, Jushan and Pierre Perron (1998), "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica*, *66*, 47-78.

Bai, Jushan and Pierre Perron (2000), "Computation and Analysis of Multiple Structural Changes," manuscript, Boston University, February.

Bai, Jushan (1999), "Likelihood Ratio Tests for Multiple Structural Changes," *Journal of Econometrics*, *91*, 299-323.

Bai, Jushan, Robin L. Lumsdaine, and James H. Stock (1998), "Testing for and Dating Common Breaks in Multivariate Time Series." *Review of Economic Studies*, *65*, 395-422.

Balke, Nathan, and Thomas Fomby (1991), "Shifting Trends, Segmented trends, and Infrequent Permanent Shocks," *Journal of Monetary Economics*, *28*, 61-85.

Banerjee, Anindya, Robin L. Lumsdaine, and James H. Stock (1992), "Recursive and Sequential Tests of the Unit Root and Trend Break Hypotheses: Theory and International Evidence," *Journal of Business and Economic Statistics*, *10*, 271-287.

Ben-David, Dan, and David H. Papell (1995), "The Great Wars, the Great Crash, and Steady State Growth: Some New Evidence About an Old Stylized Fact," *Journal of Monetary Economics*, *36*, 453-475.

Ben-David, Dan, and David H. Papell (1998), "Slowdowns and Meltdowns: Post-War Growth Evidence from 74 Countries," *Review of Economics and Statistics*, *LXXX*, 561-571.

Busetti, Fabio and Andrew Harvey (2001), "Testing for the Presence of a Random Walk in Series with Structural Breaks, *Journal of Time Series Analysis*, *22*, 127-150.

Campbell, John and Pierre Perron (1991), "Pitfalls and Opportunities: What Macroeconomists Should Know About Unit Roots," *NBER Macroeconomics Annual*, 141-201.

Christiano, Lawrence, (1992), "Searching for a Break in GNP," *Journal of Business and Economic Statistics*, *10*, 237-250.

Jones, Charles I. (1995), "R % D-Based Models of Economic Growth," *Journal of Political Economy*, *103*, 759-784.

Lumsdaine, Robin L. and David H. Papell (1997), "Multiple Trend Breaks and the Unit Root Hypothesis," *Review of Economics and Statistics*, *LXXIX*, 212-218.

Maddison, Angus (1991), *Dynamic Forces in Capitalist Development: A Long-Run Comparative View*, Oxford: Oxford University Press.

Nelson, Charles, and Charles Plosser, (1982), "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics*, *10*, 139-162.

Ng, Serena, and Pierre Perron, (1995), “Unit Root Tests in ARMA Models With Data Dependent Methods for Selection of the Truncation Lag,” *Journal of the American Statistical Association*, *90*, 268-281.

Park, Joon and Jaewhan Sung (1994), Testing for Unit Roots in Models with Structural Change, *Econometric Theory*, *10*, 917-936.

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

Perron, Pierre (1994), "Trend, Unit Root and Structural Change in Macroeconomic Time Series," in B.B. Rao, ed., *Cointegration for the Applied Economist*, London: Macmillan.

Raj, Baldev (1992), "International Evidence on Persistence in Output in the Presence of an Episodic Change," *Journal of Applied Econometrics*, *7*, 281-93.

Romer, Paul M. (1986), "Increasing Returns and Long Run Growth," *Journal of Political Economy*, *94*, 1002-38.

Solow, Robert M. (1956), "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics, 70*, Reprinted in Joseph E. Stiglitz and Hirofumi Uzawa (eds.), *Readings in the Modern Theory of Economic Growth*, Cambridge: MIT Press, 58-87.

Zivot, Eric and Donald Andrews (1992), "Further Evidence on the Great Crash, the Oil Price Shock, and the Unit Root Hypothesis," *Journal of Business and Economic Statistics*, *10*, 251-270.

Table 1

Model AA

Aggregate Real GDP

Country	TB1	TB2	t-statistic for α	k	Significance Level (in %)
Australia	1891	1925	-4.97	0	
Austria	1944	1947	-5.69	2	
Belgium	1913	1961	-4.01	3	
Canada	1884	1928	-6.29	7	5
Denmark	1939	1958	-5.29	3	
Finland	1876	1913	-4.53	4	
France	1939	1945	-4.80	7	
Germany	1944	1948	-6.68	2	2.5
Italy	1943	1945	-4.80	3	
Japan	1944	1950	-5.23	5	
Netherlands	1939	1947	-4.99	2	
Norway	1944	1970	-3.49	2	
Sweden	1941	1959	-3.91	5	
Switzerland	1944	1958	-5.52	1	
U.K.	1918	1962	-7.27	7	1
U.S.A.	1929	1978	-7.29	8	1

Per Capita Real GDP

Country	TB1	TB2	t-statistic for α	k	Significance Level (in %)
Australia	1891	1925	-5.22	2	
Austria	1944	1947	-6.06	2	10
Belgium	1958	1968	-3.71	3	
Canada	1917	1928	-7.00	7	1
Denmark	1939	1958	-4.61	2	
Finland	1898	1913	-4.23	4	
France	1939	1945	-5.18	7	
Germany	1944	1946	-7.31	2	1
Italy	1941	1945	-4.55	0	
Japan	1944	1950	-4.60	5	
Netherlands	1939	1947	-4.80	2	
Norway	1944	1970	-3.33	2	
Sweden	1897	1913	-4.04	5	
Switzerland	1944	1958	-5.42	1	
U.K.	1900	1918	-5.48	5	
U.S.A.	1929	1940	-8.87	8	1

Note: The critical values are -6.74 (1 percent), -6.43 (2.5 percent), -6.16 (5 percent), and -5.89 (10 percent).

Table 2

Model CA

Aggregate Real GDP

Country	TB1	TB2	*t*-statistic for *α*	*k*	Significance Level (in %)
Australia	1891	1925	-5.46	2	
Austria	1944	1913	-6.53	2	10
Belgium	1939	1916	-7.60	6	1
Canada	1928	1908	-7.76	7	1
Denmark	1939	1979	-6.31	3	
Finland	1916	1939	-7.51	3	1
France	1939	1953	-8.29	4	1
Germany	1944	1952	-6.85	1	2.5
Italy	1942	1906	-5.11	1	
Japan	1944	1980	-9.64	4	1
Netherlands	1939	1980	-5.72	3	
Norway	1916	1939	-5.95	3	
Sweden	1916	1980	-5.58	5	
Switzerland	1958	1944	-5.42	1	
U.K.	1918	1944	-7.80	5	1
U.S.A.	1942	1929	-7.93	8	1

Per Capita Real GDP

Country	TB1	TB2	*t*-statistic for *α*	*k*	Significance Level (in %)
Australia	1927	1891	-6.72	8	5
Austria	1944	1913	-7.75	8	1
Belgium	1939	1913	-6.90	6	2.5
Canada	1917	1928	-7.53	7	1
Denmark	1939	1979	-7.43	4	1
Finland	1916	1938	-7.93	3	1
France	1939	1979	-8.34	4	1
Germany	1913	1953	-6.44	7	10
Italy	1942	1980	-5.24	1	
Japan	1944	1980	-9.70	4	1
Netherlands	1939	1980	-5.93	3	
Norway	1920	1939	-6.62	3	5
Sweden	1915	1959	-6.47	4	10
Switzerland	1944	1921	-6.16	0	
U.K.	1918	1944	-7.92	5	1
U.S.A.	1929	1940	-9.13	8	1

Note: The critical values are -7.19 (1 percent), -6.85 (2.5 percent), -6.62 (5 percent), and -6.37 (10 percent).

Table 3

Model CC

Aggregate Real GDP

Country	**TB1**	**TB2**	***t*-statistic for *α***	***k***	**Significance Level (in %)**
Australia	1891	1928	-6.00	2	
Austria	1913	1944	-7.53	8	1
Belgium	1916	1939	-7.56	6	1
Canada	1908	1928	-7.71	7	1
Denmark	1939	1974	-7.16	3	2.5
Finland	1916	1939	-7.45	3	1
France	1939	1975	-9.32	4	1
Germany	1944	1963	-8.15	1	1
Italy	1942	1966	-5.72	1	
Japan	1944	1974	-13.79	0	1
Netherlands	1939	1975	-6.17	3	
Norway	1903	1939	-7.57	4	1
Sweden	1916	1963	-6.20	5	
Switzerland	1940	1969	-5.14	8	
U.K.	1918	1959	-7.18	5	2.5
U.S.A.	1929	1955	-7.81	8	1

Per Capita Real GDP

Country	**TB1**	**TB2**	***t*-statistic for *α***	***k***	**Significance Level (in %)**
Australia	1891	1927	-6.70	8	10
Austria	1944	1959	-8.27	2	1
Belgium	1916	1939	-7.17	6	2.5
Canada	1908	1928	-7.70	7	1
Denmark	1939	1975	-7.73	4	1
Finland	1916	1943	-8.11	3	1
France	1939	1974	-8.87	4	1
Germany	1944	1958	-6.13	1	
Italy	1942	1966	-5.92	1	
Japan	1944	1973	-14.34	0	1
Netherlands	1939	1975	-6.18	3	
Norway	1917	1939	-7.20	4	1
Sweden	1916	1963	-7.28	5	1
Switzerland	1921	1944	-5.95	0	
U.K.	1918	1945	-7.02	5	2.5
U.S.A.	1929	1955	-8.88	8	1

Note: The critical values are -7.19 (1 percent), -6.95 (2.5 percent), -6.75 (5 percent), and -6.48 (10 percent).

Table 4

Trend Breaks and Per Capita Growth Rates*

	Trend Breaks		Level Changes		Average Growth Rates (by period)			Ratios of Growth Rates		
	TB1	**TB2**	**TB1**	**TB2**	**1**	**2**	**3**	**2/1**	**3/2**	**3/1**
Overall Average			-0.155	-0.046	1.32%	3.39%	2.39%	2.63	1.04	1.93
Group A										
U.S.A.	1929	1955	-0.193	-0.084	1.77%	2.13%	1.85%	1.20	0.87	1.05
Germany	1944	1958	-0.181	-0.039	1.83%	6.81%	2.76%	3.72	0.41	1.51
Austria	1944	1959	-0.404	-0.038	1.09%	8.73%	3.11%	8.01	0.36	2.86
Sweden	1916	1963	-0.118	0.072	1.27%	2.93%	2.11%	2.31	0.72	1.67
Italy	1942	1966	-0.150	-0.003	1.03%	4.15%	2.84%	4.04	0.68	2.76
Japan	1944	1973	-0.622	-0.086	1.69%	7.68%	3.29%	4.54	0.43	1.95
Neth.	1939	1975	-0.224	-0.013	1.09%	2.87%	1.09%	2.64	0.38	1.01
Denmark	1939	1975	-0.141	-0.013	1.53%	2.70%	1.53%	1.77	0.57	1.01
France	1939	1974	-0.234	-0.016	1.29%	3.49%	1.86%	2.71	0.53	1.45
Group A Average			-0.252	-0.025	1.40%	4.61%	2.27%	3.44	0.55	1.69
Group B										
Belgium	1916	1939	-0.063	-0.163	0.90%	1.63%	2.62%	1.81	1.61	2.90
Norway	1917	1939	-0.001	-0.106	1.10%	3.14%	3.21%	2.85	1.02	2.91
Finland	1916	1943	-0.118	-0.089	1.11%	3.20%	3.38%	2.88	1.06	3.05
Switz.	1921	1944	0.097	0.129	0.98%	1.21%	2.27%	1.23	1.88	2.31
U.K.	1918	1945	-0.103	-0.043	1.12%	1.13%	2.13%	1.00	1.90	1.90
Group B Average			-0.038	-0.054	1.04%	2.06%	2.72%	1.95	1.49	2.61
Group C										
Australia	1891	1927	-0.110	-0.113	1.31%	0.65%	1.87%	0.49	2.88	1.43
Canada	1908	1928	0.092	-0.127	1.98%	1.76%	2.35%	0.89	1.33	1.19
Group C Average			-0.009	-0.120	1.65%	1.20%	2.11%	0.69	2.11	1.31

* Omission of the four countries with non-significant breaks does not appreciably affect the group averages reported in this table.

Figure 1

Relationship Between Changes in Per Capita Output Levels and Subsequent Changes in Per Capita Growth Rates

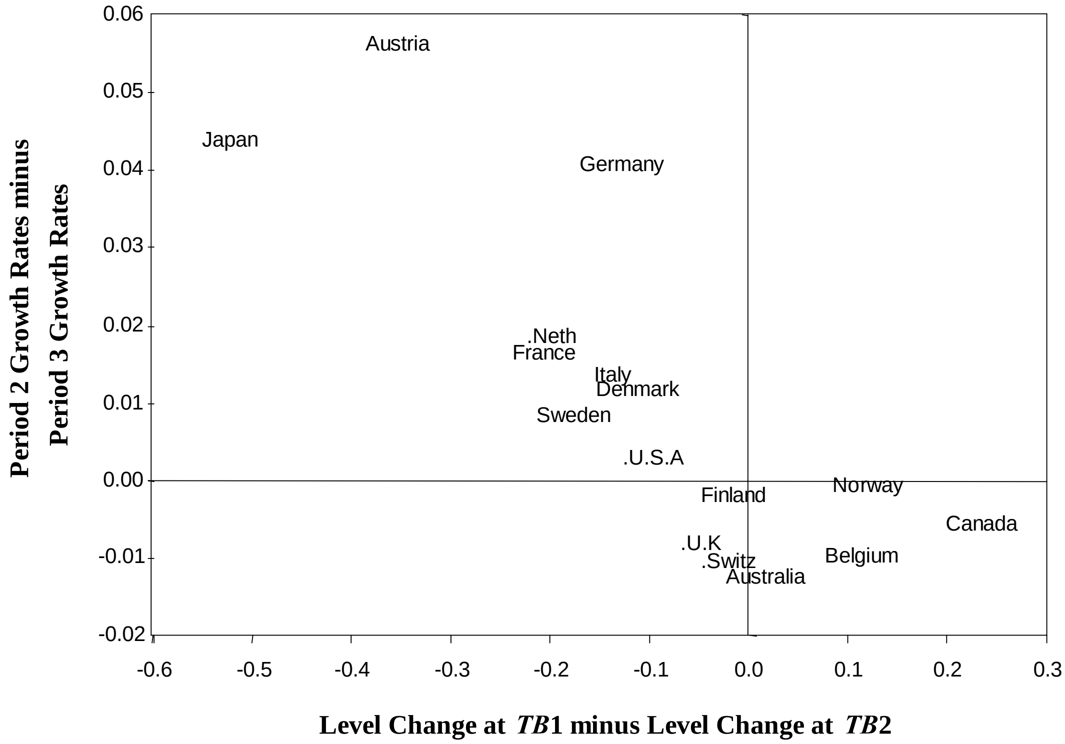

Figure 2

Comparisons of Growth Paths