Time Series Tests of Constant Steady-State Growth

David H. Papell and Ruxandra Prodan

University of Houston

January 11, 2012

Abstract

We propose a new methodology to study the stability of steady-state growth. Long-run GDP per capita can be characterized by: (1) the *linear trend hypothesis*, where there are no long-run changes in GDP levels or growth rates, (2) the *level shift hypothesis*, where there are long-run level shifts, but not changes in growth rates, and (3) the *growth shift hypothesis*, where there are long-run changes in both GDP levels and growth rates. We formally test these hypotheses using time series techniques with over 135 years of data. The results are not favorable to the hypothesis of constant steady-state growth. While we find evidence supporting the *linear trend hypothesis* for the United States and Canada and the *level shift hypothesis* for three additional OECD countries, the *growth shift hypothesis* is supported for seven OECD and four Asian countries. The results are not driven by transition dynamics.

JEL Classification: O40, C22, C52

Keywords: Long-run economic growth, Time series tests, Structural change

We thank Chad Jones, Sebnem Kalemli-Ozcan and Dietz Vollrath for helpful comments and discussions.

Correspondence to:
David Papell, tel. (713) 743-3807, email: dpapell@mail.uh.edu, Department of Economics, University of Houston, Houston, TX 77204-5019
Ruxandra Prodan, tel. (713) 743-3836, email: rprodan@uh.edu, Department of Economics, University of Houston, Houston, TX 77204-5019

"Consider the following simple exercise. An economist living in the year 1929 (who has miraculous access to historical per capita GDP data) fits a simple linear trend to the natural log of per capita GDP for the United States from 1880 to 1929 in an attempt to forecast per capita GDP today, say in 1987. How far off would the prediction be?" Charles Jones (1995b)

1. Introduction

Is long-run economic growth best characterized by a constant growth rate of per capita GDP, which would be consistent with a simple linear trend of the natural log of per capita GDP? Alternatively, is long-run growth best characterized by permanent changes in the growth rate of per capita GDP, which would not in general be consistent with a simple linear trend of the natural log of per capita GDP? Finally, is long-run growth best characterized by transitory, but not permanent, changes in the growth rate of per capita GDP, which also would not in general be consistent with a simple linear trend of the natural log of per capita GDP?

The average U.S. growth rate appears to have been stable over the last century, being well described by a process with a constant mean and little persistence. This presents a puzzle from the perspective of both neoclassical and endogenous growth models. In the neoclassical growth framework, the documented significant increase in R&D intensity and time spent accumulating skills through formal education should have generated temporarily high growth rates and long run "level effects", but the evidence looks like an economy that is fluctuating around its balanced growth path.[1] In many endogenous growth models where ideas are nonrivalrous, such policy changes should lead to a permanent increase in the growth rate: in the models of Romer (1986), Lucas (1988) and Rebelo (1991) where investment in human and physical capital is the engine of growth or in the models of Romer (1990), Grossman and Helpman (1991) and Aghion and Howitt (1992) where the growth rate of the economy is an increasing function of the research effort.

In an attempt to reconcile growth models with these facts, Jones (1995b, 2002, 2005) proposes a "semi-endogenous" or "ideas-based" growth model. Once he allows for decreasing returns to scale for research, policy changes will have no effect on the long run growth rate, but only on the long-run level of per capita income.[2] Jones argues that these models exhibit "weak" scale effects versus the first-generation idea-based growth models which exhibit "strong" scale effects, that they are better representations for the trends in the data, and that they do not rely on the assumption of linearity.[3] He also gives several possible

[1] The neoclassical growth model (Solow, 1956) postulates stable equilibrium with a long run constant income growth rate. Following changes in variables affected by government policy, specifically savings and investment, the growth rate of the economy rises temporarily and then returns to its original value, but the level of income is permanently higher as a result.

[2] The long-run growth rate in these models is generally an increasing function of the growth rate of research effort, which depends on the population growth rate of the countries contributing to the world research. But this growth rate is taken to be exogenous in these models, producing the policy-invariance results.

[3] The law of motion for human and physical capital accumulation is assumed to be linear and the idea production function is assumed to be a linear differential equation.

explanations for the stable US growth rate: First, nothing in the US experience has had large persistent effects on the growth rate, second, persistent effects have been offsetting each other and, third, US growth is generated by a sequence of transition dynamics at a constant rate that is higher than the steady state rate.[4] An implication of this result is that growth rates could slow down at some point in the future, when the transition period ends.[5]

A large amount of empirical work, using cross-country and time series methodologies, has tried to make a clear choice between neoclassical and endogenous type growth models, but has failed to arrive at a consensus. Studies by Lee (1988), Kremer (1993) and others have argued in favor of endogenous growth models with "strong" scale effects. In contrast, Mankiw et al (1992), Karras (2001) and others find significant evidence in agreement with the neoclassical growth model. While there has not been as much research devoted to study the relevance of "weak" scale effects, cross-country studies that explicitly control for differences in international trade and institutional quality, such as Backus, Kehoe and Kehoe (1992), Frankel and Romer (1999), Jones (2001) and Alcala and Ciccone (2002) have found an important role for "weak" scale effects.

This paper proposes a new methodology to study an important set of "stylized facts" that aggregative models of economic growth should endeavor to explain, the pattern of output growth over 139 years for 19 OECD countries and 7 Asian countries.[6] While it is clear that many developing countries, especially Asian countries, experience regular changes in their growth rates over shorter time periods, it has not been established in the literature which empirical case applies over the long run.[7] We categorize patterns of long-run economic growth by three hypothesis: the *"linear trend hypothesis"* that output can be represented by a simple trend, the "*level shift hypothesis*" that there have been permanent changes in the level of income, but only temporary changes in growth rates, and the "*growth shift hypothesis"* that there have been permanent changes in long run growth rates. The linear trend and level shift hypotheses, but not the growth rate hypothesis, are characterized by constant steady-state growth. The purpose of this paper is to determine which countries are best characterized by which hypotheses. Our innovation consists in the fact that we formally test these hypotheses using time series techniques.

Tests for structural change in univariate time series provide a natural framework for investigating stylized facts of output growth. We test for multiple structural changes using a specification which allows

---

[4] Jones (2002) argues that 80% of the US growth in post-war period is due to the transition dynamics associated with increases in educational attainment or increase in the world R&D intensity. They seem to rise smoothly, generating an approximate stable growth path.

[5] Jones (1995b) also documents that growth rates of GDP per capita have remained roughly constant over the post-World War II era for several OECD countries

[6] We will henceforth use "output" or "income" to denote "the logarithm of per capita real GDP". Our data span choice is dictated by the cross-country data availability.

[7] Some papers that discuss changes in the growth rates in developing economies over shorter time scales are Hausmann, Pritchett and Rodrick (2005), Jones and Olken (2008), Berg, Ostry and Zettelmeyer (2007) and Jerzmanowski (2006).

for two types of structural breaks in trend output: one in both the intercept (level effect) and the slope (growth effect) of the trend function and the other(s) in just the slope. The inspiration for the model is threefold. First, a common characteristic of most of these countries is that they experience a sharp fall in output during the Great Depression, World War I, and/or (especially) World War II, followed by higher growth, which we capture by a change in the intercept and the slope for the first break.[8] Second, several countries experience extremely high growth rates immediately following World War II which level off in the late 1940s and early 1950s, which we characterize by a change in the slope, but not the intercept, for the second break. Third, a number of countries experience a slowdown in growth rates starting in the 1970s, which we also capture by a change in the slope, but not the intercept, for the second break.[9]

To formally test the above hypotheses, we develop tests for restricted structural change. The most restrictive specification, which corresponds to the linear trend hypothesis, constrains the trend following the last break to be a linear projection of the trend function preceding the first break. The second most restrictive specification, which corresponds to the level shift hypothesis, constrains the slope of the trend function following the last break to equal the slope of the trend function preceding the first break, but does not impose any constraints on the level of output. The least restrictive specification, which corresponds to the growth shift hypothesis, does not place any constraints on either the level or the growth rate of output following the last break.[10]

In order to test the null hypotheses of restricted breaks against the alternative hypothesis of unrestricted breaks, we must first establish that output can be characterized as regime-wise trend stationary with structural changes. Using a variety of unit root tests in the presence of structural change, we reject the unit root null against the regime-wise trend stationary alternative for 15 of 19 OECD countries and 6 out of 7 Asian countries. We next test for structural change. Among the "stationary" countries, we find evidence of three breaks for 3 out of 15 OECD countries, evidence of two breaks for 10 out of 15 OECD countries and 2 out of 6 Asian countries, and evidence of one break for all of the other countries.

Is long-run economic growth best characterized by the linear trend hypothesis? For the United States and Canada, the answer is positive. The model with the linear trend restriction cannot be rejected in favor of the model with unrestricted breaks at the 5 percent significance level. For the OECD and Asian countries as a whole, however, the answer is resoundingly negative. The model with the linear trend restriction is rejected in favor of the model with unrestricted breaks for every other OECD and Asian

[8] These sharp falls in output can be very substantial, the largest being an almost 50 percent drop in output for Japan in 1945.
[9] We also allow for a second break in the slope, but find evidence of three breaks for only two countries, Austria, Japan and Spain.
[10] In Papell and Prodan (2011), we use the models developed in this paper to investigate whether severe recessions associated with financial crises cause permanent reductions in potential GDP.

country for which structural change tests produced evidence of more than one break. In other words, there are significant changes in GDP levels, GDP growth rates, or both for most countries.

It is not surprising that the linear trend hypothesis is rejected for so many countries, for it is not a prediction of any class of growth models. A more compelling question is whether, after the end of a transition period, countries return to their pre-first-break rates of growth as predicted by the level shift hypothesis. The model that constrains the slope of the trend function following the last break to equal the slope of the trend function preceding the first break is rejected in favor of the model with unrestricted breaks at the 5 percent significance level for all countries except Austria, Denmark, Germany, Italy, and Sweden. In addition, for the countries that are characterized by one break, the null hypothesis of constant growth can be rejected in favor of the alternative hypothesis changing growth for 4 of the 5 countries, the exception being Switzerland, which experiences a positive intercept break, but no slope break, in 1944. Combining the two sets of results, both variants of the ideas-based growth model, the linear trend hypothesis and the level shift hypothesis are rejected in favor of the growth shift hypothesis for 7 out of 15 OECD countries and all 6 Asian countries.

While we use the latest available data, the endpoint is necessarily determined by when the paper was written. We cannot rule out *a priori* the possibility that the rejections of the level shift hypothesis in favor of the growth shift hypothesis occur because not enough time has passed since the last break to return to steady-state growth. In order to test the conjecture that our results may be caused by transition dynamics, we add 50 years of artificial data to each OECD series, assuming that the actual data generating process since 1973 will continue unchanged, and test the level shift null against the growth shift alternative. The results for the extended series are very similar to our previous findings, and do not support the conjecture that the rejections are caused by transition dynamics.

While the structural change methodology can provide statistical evidence to differentiate among the three hypotheses, it does not describe how large the changes are. In order to supplement the statistical evidence with economic evidence, we compare the growth rates before the first break with the growth rates after the last break. While, for most countries, the statistical and economic evidence coincides in the sense that the countries for which the linear trend and level shift hypotheses are rejected in favor of the growth shift hypothesis are also those countries with high ratios of post-last-break to pre-first-break growth rates, there are several exceptions. For three countries, Austria, Germany, and Italy, the constraint that the slope of the trend function following the second break to equal the slope of the trend function preceding the first break cannot be rejected even though the ratios are high, casting doubt on their classification as being in accord with the level shift hypothesis. Conversely, for the Philippines, the same constraint can be rejected even though the ratio is relatively low, casting doubt on its classification as being in accord with the growth shift hypothesis. The combination of the statistical and economic results is not supportive of the constant steady-state growth hypothesis. While we find evidence supporting either the *linear trend hypothesis* or the *level shift hypothesis* for 5 countries, the evidence points towards the *growth shift hypothesis* for 11 countries.

2. Methodology and Empirical Results

While structural change tests play an important role in studying long-run macroeconomic time series, there has not been much research on the presence of structural change when investigating output growth. Jones (1995b), analyzing long-run growth rates, finds evidence of significant mean shifts for the growth rates of 6 out of 14 OECD countries, almost all of them associated with World War II. He documents large movements in policy variables and lack of persistent changes for the US over the last century and for OECD countries for the post-war era, arguing against the predictions of endogenous growth models. Ben-David and Papell (1995), using tests for one structural change and 120 years of annual per capita GDP data for 16 OECD countries, find that most of the countries exhibit a major break around World War II, with either World War I or the Great Depression constituting the break for the remaining countries. In most cases, a shock to the economy is followed by sustained growth that exceeds the earlier steady state growth, which they interpret as being compatible with Romer-type endogenous growth models.

The innovation in this paper is that we propose formal tests to analyze the stylized facts of output growth. First, we implement a structural change model designed to capture the idea that the aggregates fell during the World Wars or The Great Depression (a change in the intercept and the slope for the first break), followed by a recovery and a smooth transition to a new growth path (a change in the slope, but not in the intercept, for the second and (possibly) third break). Next, we test two restrictions: (1) the trend following the last break is constrained to be a linear projection of the trend prior to the first break, and (2) the slope of the trend function following the last break is constrained to equal the slope of the trend function prior to the first break. The difference between the restrictions is that (1), but not (2), imposes a constraint on the level, as well as the growth rate, of output following the last break.

The objective of this paper is to use tests for restricted structural change to analyze stylized facts of output growth. These tests, however, are only valid if per-capita GDP is regime-wise trend stationary; that is, trend stationary around one or more changes in the intercept and/or the slope of the series. We therefore first test for unit roots, and then test for restricted structural change using only those series for which we can reject the unit root null.

2.1 Tests for Unit Roots

The issue of unit roots in long-term output has been a matter of controversy since Nelson and Plosser (1982) failed to reject the unit root hypothesis, using conventional unit root tests, for either aggregate or per capita GNP. Perron (1989) inaugurated testing for unit roots in the presence of structural change, but in a context where the break date was imposed exogenously. Ben-David and Papell (1995), using tests with endogenous break selection and one break point for 16 countries, rejected the unit root hypothesis for half of the countries using per-capita GDP. Ben-David et al (2003), allowing for two endogenous break points, rejected the unit root hypothesis for three quarters of the countries.

Augmented-Dickey-Fuller (ADF) tests for a unit root, both with and without allowing for shifts in the deterministic trend at unknown dates, can be described as follows:

$$\Delta y_t = \mu + \beta t + \gamma_1 DU_{1t} + \theta_1 DT_{1t} + \theta_2 DT_{2t} + \alpha y_{t-1+} \sum_{i=1}^{k} \rho_i \Delta y_{t-i} + u_t \,, \qquad (1)$$

for t = 1,….,T, where $DU1_t$ is an indicator dummy variables for the intercept change in the trend function, occurring at time TB_1 . That is DU1 = 1 ($t>TB_1$). $DT1_t$ and $DT2_t$ are the corresponding changes in the slope of the trend function. That is $DT1_t = (t - TB_1)1(t>TB_1)$ and $DT2_t = (t - TB_2)1(t>TB_2)$.

Three types of models are estimated. The standard ADF test sets $\gamma_1 = \theta_1 = \theta_2 = 0$ and tests the null hypothesis of a unit root in favor of the alternative of trend-stationarity. We also test the null hypothesis of a unit root in favor of the alternative of *broken trend-stationarity*: Model A($\theta_1 = \theta_2 = 0$) allows for one break in the intercept and Model C ($\theta_2 = 0$) allows for one break in the intercept and slope of the trend function (Zivot and Andrews, 1992). Model CB allows shifts in the intercept and the slope of the trend function corresponding to the first break and shifts in the slope corresponding to the second break (Papell and Prodan, 2007).

We use the "general-to-specific" recursive t-statistic procedure suggested by Hall (1994) and Ng and Perron (1995) to choose the number of lags. We set the maximum value of k equal to 8, and use a critical value of 1.645 from the asymptotic normal distribution to assess significance of the last lag. The unit root null hypothesis is rejected against the regime-wise trend stationarity alternative if $\alpha < 0$. Equation (1) is estimated sequentially for each break year Tbi = k+2, …,T – 2, where i =1,2 and T is the number of observations, $Tb_1 \neq Tb_2$ and $Tb_1 \neq Tb_2 \pm 1$. The chosen break is that for which the maximum evidence against the unit root null, in the form of the most negative t-statistic on α, is obtained. Bootstrap critical values with the data generated under the null hypothesis, using 139 observations and 5000 replications, are calculated for each one of the four tests. The critical values for the finite sample distributions are taken from the sorted vector of 5000 replicated statistics (t-statistic on α in equation 1).

Table 1 reports the unit root test results. We first test for a unit root without considering structural change. Using ADF tests on the data of Maddison (2009) for 19 OECD and 7 Asian countries for the period 1870-2008, we can reject the unit root null against the alternative of trend-stationarity at the 10% significance level for only the United States out of 26 countries.[11] We next test for a unit root in the presence of one structural change. With Model A (one-time change in the intercept), we reject the unit

[11] This is the same result as in Ben David and Papell (1995), who found evidence of trend stationarity only for the U.S. among 16 OECD countries (at the 5% significance level).

root null for three out of 26 countries (Switzerland, Philippines and South Korea) at the 5% level. With Model C (one-time change in the intercept and slope), we find evidence of broken-trend stationarity level for one out of 26 countries (the U.S.) at the 5% significance level and for 4 more at the 10% significance level.[12]

When testing GDP per capita in the presence of one break, the results are dominated by the World Wars and the Great Depression. In order to account for the possibility that countries experience more than one break, possibly related to the postwar slowdowns, we test a for unit root in the presence of two breaks using model CB (one change in the intercept and slope and another change in just the slope the slope). We find evidence of broken trend stationarity for 15 out of 26 countries at 5% significance level and for 4 countries at 10% significance level. Our results are comparable with Ben-David, Lumdaine and Papell (2003) who found evidence of broken trend stationarity for three quarters of the cases.[13] Combining the two tests we find evidence of broken trend stationarity for 21 out of 26 countries at the 10% significance level. Switzerland and Philippines are the only countries for which we report a rejection for a single-break model (A) but not for Model CB. We did not find any evidence of regime wise trend stationarity for Finland, India, New Zealand, Norway, and Portugal.

2.2 Tests for Structural Change

Having established which countries experience evidence of regime wise trend-stationarity, we focus next on tests for multiple structural changes. We utilize the Bai (1999) likelihood ratio test, which allows for lagged dependent variables and multiple changes in the intercept and the slope of the trend function. While the wide range of possible structural changes precludes tabulated critical values, straightforward methods can be used in order to calculate bootstrap critical values.[14]

In accord with the stylized fact that long-horizon real GDP for many countries has been adversely affected by one "big" event, possibly accompanied by faster growth, followed by one or more subsequent growth changes, we consider a model where we allow for three specific breaks. One break is in both the intercept and the slope, while the others are just in the slope, of the trend function. For example, much previous work suggests that the Great Depression had the largest negative effect on GDP for the U.S., followed by faster growth in the 1930s, a return to more normal growth in the 1940s, and (possibly) a growth slowdown in the 1970s. For other countries, the large adverse event was either World War I or (especially) World War II. The estimated model is presented below,

[12] We report considerably fewer rejections than Ben-David and Papell (1995) because, while they use Zivot and Andrews (1992) asymptotic critical values, we use correctly-sized bootstrapped critical values.
[13] Ben-David, Lumsdaine, and Papell (2003) also use bootstrapped critical values. They test several models, but find most of their rejections when estimating model CC (allowing for two breaks in the intercept and slope of the trend function). They do not estimate model CB.
[14] It has been documented by Diebold and Chen (1996) and Prodan (2005) that tests for one and respectively multiple structural changes suffer from serious size distortions when tabulated critical values calculated with iid errors are used. The authors recommend the use of bootstrap critical values in the case of highly persistent data.

$$y_t = \mu + \beta t + \gamma_1 DU_{1t} + \theta_1 DT_{1t} + \theta_2 DT_{2t} + \theta_3 DT_{3t} + \sum_{i=1}^{k} \rho_i y_{t-i} + u_t, \tag{2}$$

where $DU_t = 1$ if $t > Tb_i$ and 0 otherwise, and $DT_t = t - Tb_i = 1$ if $t > Tb_i$ and 0 otherwise. The lag length, k, is chosen by Schwarz Information Criterion (SIC), which involves minimizing the function of the residual sum of squares combined with a penalty for a large number of parameters in the previous regressions. While the motivation for the structure of the breaks comes from the World Wars, Great Depression, and growth slowdown, nothing in the estimation constrains the breaks to follow this or any other pattern. In particular, we do not require that the first break, rather than the second or the third, be the one characterized by changes in both the intercept and the slope.

The estimated break points are obtained by a global minimization of the sum of squared residuals. The test statistic is based on the difference between the minimized SSR associated with ℓ breaks and the minimized SSR associated with $\ell + 1$ breaks. The model is estimated optimally under both null and the alternative hypothesis, which means that ℓ breaks under the null and $\ell + 1$ breaks under the alternative are estimated simultaneously. When $\ell = 0$, the test reduces to the usual test of no change against a single change. When the test is performed repeatedly while augmenting the value of ℓ, the number of break points can be consistently estimated.

It has been previously documented that in certain configurations of changes, particularly breaks of opposite sign, the Bai sequential method is unable to reject the null hypothesis of 0 versus 1 break but it is not difficult to reject the null hypothesis of 0 versus of a higher number of breaks. In order to improve the power of the test, Prodan (2008) proposed the following procedure. First, look at the sequential method. If 0 against 1 break is rejected, continue with the sequential method until the first failure to reject. If 0 against 1 is not rejected, then test the hypothesis of no break versus a fixed number of breaks (in this case 2 breaks). If $\sup F_t(2)$ is significant, then the number of breaks can be decided upon examination of $F_t(2|1)$.

Critical values are calculated using parametric bootstrap methods. Generally we assume that the underlying process follows a stationary finite-order autoregression of the form:

$$A(L) y_t = e_t, \tag{3}$$

$e_t \sim iidN(0,1)$ with $E(e_t) = 0$ and $E(e^2{}_t) < \infty$. $Y = (y_{1,...,} y_T)'$ denotes the observed data. $A(L)$ is an invertible polynomial in the lag operator. The AR(p) model may be bootstrapped as follows: First determine the optimal AR(p) model, using the Schwartz criterion. Next, estimate the parameters $\hat{A}(L)$ for the optimal model. Following, to determine the finite-sample distribution of the statistics under the null hypothesis of no structural change, use the optimal AR model with $iidN(0, \sigma^2)$ innovations to construct a pseudo sample of size equal to the actual size of the data, where σ^2 is the estimated innovation variance of the AR model. Then calculate the bootstrap parameter estimates: $\hat{A}^*(L)$ and compute the statistics of interest. The critical values are taken from the sorted vector of 5000 replicated statistics.

The results are reported in Table 2. Evidence of some variant of structural change is found for all 21 countries. We are able to reject the null hypothesis of no breaks against the alternatives of three breaks for 3 countries, two breaks for 12 countries, and one break for the remaining 6 countries. For the 15 countries for which two or three breaks were found, the first break was always in both the intercept and slope, with the remaining breaks in just the slope.

The dates of the breaks are consistent with the stylized facts; countries which were most affected by the war experienced a significant drop in the income levels around World War II, (European countries as Austria, Belgium, Denmark, France, Germany, Italy, Netherlands, Spain and Switzerland as well as Japan), followed by much faster rates of growth. Among these countries, several experienced a second change in the growth rate during the 1970s: Belgium, Denmark, France, Spain and Japan.[15] The remaining countries, Austria, Germany, Italy and Netherlands, experienced the second break during the 1950s.[16] An interesting case is Japan which experienced a third break in 1991 due to the economic downturn following the collapse in the stock and land market. The other countries were most affected by World War I, (Finland, Sweden and the UK), or the Great Depression, (Australia, Canada and the US).[17] The majority of the Asian countries experience one break around World War II.

3. Structural Change and Economic Growth Hypotheses

Using a combination of unit root and structural change tests, we have established that the process of economic growth is not smooth. Starting with 26 OECD and Asian countries, we were able to reject unit roots in per capita GDP for 21 countries. Among those 21 countries, all display evidence of structural change. We now turn to the central question of the paper and use restrictions on the form of structural change to differentiate among various growth hypotheses.

3.1. The linear trend hypothesis

Jones (1995, 2002) documents that "over the last 125 years the average growth rate of per capita GDP in the US economy has been a steady 1.8 percent per year". He argues that fitting a simple linear trend from 1880 to 1929 in order to forecast per capita GDP in 1987 would lead to a prediction which is very close to the actual growth rate (he concludes that the prediction overestimates US per capita GDP by

---

[15] Austria and Spain experienced a third change in the growth rate during the 1970's.

[16] For these countries, strongly affected by the WWII, that fact that we have not found a significant break in the seventies can be due to the low power of the structural change tests to capture this break as well as the lack of the available data.

[17] Ben-David and Papell (1995) tested for one structural change using Vogelsang's (1994) test for a structural change and found similar results.

only 5%). This result underlies the conventional view that the US economy is close to its long run steady state balanced growth path. Jones (2002) argues that this view is challenged by two significant changes that occurred in the last 50 years: a substantial increase in the human capital investment (time spent accumulating skills through formal education) and an intensified search for new ideas (increase in the number of scientists engaged in research and development). These changes should lead to either long run increases in the level of income per capita in the neoclassical models and semi-endogenous growth models ("weak" scale effects) or to long run increases in the growth rate of income per capita in the endogenous models ("strong" scale effects).

Trying to solve this puzzle, Jones (2002) proposes several theories that can explain the apparent stability of the long-run growth rate and income per capita level in the US: either (1) nothing has had large persistent effects on the growth rate, or (2) persistent effects have been offsetting each other, or (3) US growth is generated by a sequence of transition dynamics at a constant rate that is higher than the steady state rate. We will call the hypothesis of a constant long-run average growth rate *"the linear trend hypothesis"*.

The linear trend hypothesis can be satisfied in two distinct ways. First, there may be no structural change, which we have already ruled out by finding evidence of some variant of structural change for all 21 countries. Second, the trend following the last break may be a linear projection of the trend preceding the first break. In order to formally test this hypothesis, we propose a model in which we impose such a restriction. Note that, if there is one break, the linear trend hypothesis cannot hold, so we only consider the 15 countries for which we found evidence of two or three breaks.

We add two restrictions to equation (2):

$$\sum_{\ell=1}^{i} \theta_\ell = 0 \tag{4}$$

which imposes a constant trend following the last break, as well as prior to the first break, and

$$\gamma_1 + \theta_1 (Tb_i - Tb_1) = 0, \quad \text{for } i = 2, 3 \tag{5}$$

which restricts the trend following the last break to equal the trend prior to the first break.

Is the restricted model a better representation of the data than the unrestricted model? Using an F-test, we test the null hypothesis that the restrictions (4) and (5) hold against the alternative hypothesis that they do not hold. If we reject the restricted model described by equations (2), (4), and (5) against the unrestricted model in equation (2), we conclude that the linear trend hypothesis does not hold. If we cannot reject the restricted model against the unrestricted model in equation (2), we conclude that the evidence favors the linear trend hypothesis.

The results are presented in Table 3 for the 12 countries for which we found evidence of two breaks and the 3 countries for which we found evidence of three breaks. The United States and Canada are the only two countries where the linear trend hypothesis cannot be rejected, so that long-run GDP per capita from 1870 to 2008 can be represented by a linear trend. In both cases output per capita dropped significantly around the Great Depression followed by a quick recovery (around 1940's). For the other countries the linear trend hypothesis can be rejected. Most of these countries were severely affected by wars, the strong after-war recovery leaving them with a different level of income and/or different growth rate of income. The graphs for the US and Canada are depicted in Figure 1.

3.2 The level shift hypothesis

We have found that the United States and Canada are the only countries for which the restriction that the trend function following the second break is a linear projection of the trend function preceding the first break cannot be rejected. This should not come as a surprise. From the visual inspection of the data, a new long run growth path which is a projection of the old growth path seems to be an unrealistic assumption for most countries. A more common pattern is for countries to first experience a substantial drop in income levels due to the world wars, in particular World War II, followed by higher growth rates during the period following the war.[18].

According to the theory of transition dynamics, one would expect the growth effects to decline over time as recovery from a negative shock to the level of income took place. The question that arises is whether the new long run growth path is significantly different than the pre-break growth path. There are two important theories that predict long run increases in the level but not the growth rate of income per capita: the neoclassical model (Solow) and the semi-endogenous, or ideas-based, growth model with weak scale effects. Jones (1995, 2002, 2004) argues that there is significant evidence towards models that exhibit "weak" scale effects versus "strong" scale effects, which would lead to a change in the long-run growth rate of income.

In order to answer these questions, we relax the previous restriction, proposing a model which allows for permanent changes in the income level and transitory changes in the growth rate. We estimate the regression described by equation (2), both unrestricted and restricted only by equation (4), and use an F-test to determine statistical significance. If we reject the restricted model described by equations (2) and (4) against the unrestricted model in equation (2), we conclude that the level shift hypothesis does not hold. If we cannot reject the restricted model against the unrestricted model in equation (2), we conclude that the evidence favors the level shift hypothesis.

The results are presented in the Table 4 for 13 countries: the 15 countries for which we found evidence of two or three breaks minus the United States and Canada, for which we have already presented evidence in favor of the more restrictive linear trend hypothesis. Among the OECD countries, the restricted model can be rejected for Belgium, France, Japan, Netherlands, and Spain, and can also be rejected for the two Asian countries, Philippines and Sri Lanka. For these countries, the post-break growth

[18] Jones (1995) and Ben-David and Papell (1995) have previously documented that most countries which were severely affected by the world wars subsequently experienced a higher average growth rate.

rate is significantly different from the pre-break growth rate. The restricted model cannot be rejected for Austria, Denmark, Germany, Italy and Sweden. For these countries the post-break growth rate is not significantly different from the pre-break growth rate. Graphs that illustrate the level shift restriction for the OECD and Asian countries are presented in Figure 2.

The level shift hypothesis, unlike the linear trend hypothesis, can hold with one structural change. For the countries which experience only one break, Australia, Indonesia, Malaysia, South Korea, Switzerland and Taiwan, we estimate the regression described by equation (2) with a single break, so that $\theta_2 = \theta_3 = 0,$ by adding the following restriction:

$$\theta_1 = 0, \tag{6}$$

which imposes a constant trend prior to and following the break by restring the single break to be only in the intercept of the trend function. As reported in Table 5, the null hypothesis that there is no break in the slope is rejected for Australia, Indonesia, Malaysia, South Korea and Taiwan, so that Switzerland is the only country for which we find evidence of constant growth using this specification.

Summarizing the results, we reject the level shift hypothesis that per capita income growth is the same prior to the first break and following the last break for 7 of the 13 OECD countries and all 6 of the Asian countries. Since failure to reject the level shift hypothesis is consistent with the predictions of the ideas-based growth model, the evidence for the model is mixed for the OECD countries and non-existent for the Asian countries.

3.3. Transition Dynamics

While post-WWII growth in countries heavily affected by the war was initially driven primarily by the transition to a peacetime economy, the level shift hypothesis predicts that countries will eventually return to their previous long-run growth rates. Nevertheless, we find that for half of the OECD countries the new long run growth path is significantly different from the pre-break growth path. In order to test the conjecture that our results are driven by transition dynamics - not enough time has passed to observe a return to the steady-state growth path - we experiment by adding 50 years of data to each OECD series assuming that the data generating process from 1973 to 2008 will continue unchanged for the next 50 years.[19] For each OECD country, we determine the optimal AR(p) model, using the Schwartz criterion. We then use the optimal AR model with $iidN(0,\sigma^2)$ innovations to construct a 50 observation sample, where σ^2 is the estimated innovation variance of the model. Finally, we compute a sample of 189 observations, from 1870 to 2059.

[19] We pick 1973 as the year that marks the beginning of the slowdown, averaging the second break years that we find in our sample.

We consider the five OECD countries for which we have not found evidence of either the linear trend or the level shift hypothesis.[20] In order to differentiate between the level shift and growth shift hypotheses, we add restriction (4) to equation (2), which imposes a constant trend following 1973, as well as prior to the first break. Using an F-test, we test the null hypothesis that restriction (4) holds against the alternative hypothesis that it does not hold. The results are presented in Table 6. Using the extended data, we are able to add one more country, France, to our evidence of a constant trend, but find evidence in accord with our previous findings for the other four countries (Belgium, Japan, Netherlands and Spain). With the exception of France, our results are not driven by transition dynamics.

3.4 Changing Growth Rates

We have presented statistical evidence testing the hypothesis that the growth rate of long-run income per capita is constant, whether or not there are permanent long-run effects on income levels. While the hypothesis can be rejected for Asian countries, the evidence is mixed for OECD countries. We now investigate whether or not the economic evidence accords with the statistical evidence. In Table 7, we report the growth rate before the first break, the growth rate after the last break, and the ratio of the post-last-break and pre-first-break growth rates, as well as the break dates themselves. This is the same comparison of growth rates that we have previously made using formal statistical techniques, and represents a wide variety of break dates and number of breaks. For all 20 countries, the ratio of the post-last-break and pre-first-break growth rates is greater than one, with the ratio ranging from 1.08 (Switzerland) to 3.93 (Australia). This is consistent with the results of Ben-David, Lumsdaine, and Papell (2003), who impose two structural breaks *a priori*, and Chen and Zivot (2008), who use Bayesian methods and let the data determine the number and form of the breaks.

There is a very strong correspondence between the statistical and the economic evidence. First, almost all countries for which the ratio is between 0.80 and 1.20 - (Canada, Denmark, Sweden, Switzerland, the U.S) – we failed to reject the hypothesis of constant steady-state growth rates, the exception being Philippines. Second, all Asian countries have ratios above 1.90 and reject the constant steady-state growth rate hypothesis. Third, most OECD countries with ratios above 1.20 – (Australia, Belgium, France, Japan, Netherlands, Spain, and the U.K.) – also reject the constant steady-state growth rate hypothesis.

Japan experienced breaks in 1944, 1971, and 1991, and it is the only country for which there is a break after the 1970s. Growth increased sharply after 1944 (along with a very large fall in the level associated with World War II), decreased with the growth slowdown after 1971, and further decreased with the "lost decade" after 1991. In order to facilitate a comparison with other OECD countries, we report the growth rate of 2.42 percent and ratio of 1.32 after the second break in 1971. For the third break

[20] We consider only the OECD countries for which we found evidence of two or three breaks because, if there is one break, neither the linear trend nor the level shift hypothesis can hold.

in 1991, the growth rate after the last break is 1.02 percent and the ratio of the post-last-break and pre-first-break growth rates is 0.56. The conventional view of Japan's post-WWII growth is that it is driven primarily by transition dynamics and would be expected to grow at a similar rate as the U.S. in the long run. Since the post-second-break U.S. growth rate is 2.03 percent, lower than Japan's 2.42 percent post-second-break growth rate but higher than Japan's 1.02 percent post-third-break growth rate, we do not have enough data to resolve the issue.

There are three exceptions to the commonality between the statistical and economic evidence among OECD countries - Austria, Germany and Italy. While we have previously found evidence that the post-break growth rate was not significantly different from the pre-break growth rate with our restricted model, the ratio of the post-break and pre-break growth rates is above 2.0. These three countries were Axis powers during World War II and experienced both a sharp decline in GDP at the end of the war and a very fast recovery (especially Austria and Germany). As shown in Figure 2, their subsequent growth paths display a considerable amount of curvature, which is not well-suited for a model with shifting linear trends. Since the statistical and economic evidence does not coincide, we cannot draw strong conclusions about these countries.

The other country for which the statistical and economic evidence differs is the Philippines. Although the ratio between the post-break and pre-break growth rates is 1.10 and no obvious changes in trend can be seen in Figure 2, the no-change null is decisively rejected. The Philippines was extremely adversely affected by World War II and experienced a very sharp and fast recovery. As with Austria and Germany, the trend break model does not appear to be well-suited to characterize this type of behavior.

4. Conclusions

We propose a new methodology in order to study the stability of output growth over 139 years for 19 OECD and 7 Asian countries. Our innovation consists in the fact that, by specifying a model with one break in both the intercept and the slope, followed by one or more breaks in the slope, we can impose restrictions and formally test hypotheses using time series methods. In order to complement the statistical results, we also investigate economic significance by comparing the pre-first-break and post-last-break growth rates.

We find evidence supporting the *linear trend hypothesis*, that the level of per capita GDP can be represented by a simple trend, for the U.S. and Canada, the two countries that were the least affected by the World Wars. For a second group of three OECD countries, Denmark, Sweden, and Switzerland, we find a constant growth rate but a permanent change in the income level after a shock, showing evidence for the *level shift hypothesis*. There is a third group of seven OECD and four Asian countries, Australia, Belgium, France, Indonesia, Japan, Netherlands, South Korea, Spain, Sri Lanka, Taiwan, and the U.K. where we find a permanent change in the growth rate and the income level after a shock, pointing towards the *growth shift hypothesis*. Finally, there are four countries, Austria, Germany, Italy, and the Philippines, for which the statistical and economic evidence differs.

The results are not supportive of the constant steady-state growth hypothesis. Among the 16 countries for which the statistical and economic evidence coincides, 11 are characterized by permanent changes in their long-run growth rates. Out of the 5 OECD countries for which the evidence favors the growth shift hypothesis over the level shift hypothesis, the result appears to be caused by transition dynamics for only one country, France. The results also highlight the differential effects of World War II on the countries' growth experiences. Within the group of 11 countries for which the evidence points towards the growth shift hypothesis, 7 have a break in the intercept and slope during World War II. For the 5 countries where the evidence is more consistent with the linear trend or level shift hypotheses, only Denmark, where the break is relatively small, and Switzerland, where the break in the intercept is positive, have a break in the intercept and slope during World War II.

References

Aghion, Philippe and Peter Howitt (1992), "A Model of Growth through Creative Destruction," *Econometrica*, *60* (2), 323-351.

Alcala, Francisco and Antonio Ciccone (2004), "Trade and Productivity," *Quarterly Journal of Economics,* CXIX (2), 613-646.

Backus, David K., Patrick J. Kehoe, and Timothy J. Kehoe (1992), "In Search of Scale Effects in Trade and Growth," *Journal of Economic Theory*, *58*, 377-409.

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*, December, 453-475.

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

Chen, Yi-Chi and Eric Zivot (2008), "Postwar Slowdowns and Long-Run Growth: A Bayesian Analysis of Structural-Break Models," manuscript, University of Washington.

Frankel, Jeffrey A. and David Romer (1999), "Does Trade Cause Growth?" *American Economic Review*, *89* (3), 379-399.

Grossman, Richard and Elhanan Helpman (1991), "Quality ladders in the theory of growth", *Review of Economic Studies*, 58, 43-61.

Hall, A. (1994), "Testing for a Unit Root in Time Series with Pretest Data-Based Model Selection", *Journal of Business & Economic Statistics*, 12 (2), 461–470.

Jones, Charles I. (1995a), "R&D-Based Models of Economic Growth," *Journal of Political Economy*, *103* (4), 759-784.

Jones, Charles I. (1995b), "Time Series Tests of Endogenous Growth Models," *Quarterly Journal of Economics*, *110* (441), 495-525.

Jones, Charles I. (1999), "Growth: With or Without Scale Effects?" *American Economic Association Papers and Proceedings*, *89*, 139-144.

Jones, Charles I. (2001), "Was an Industrial Revolution Inevitable? Economic Growth Over the Very Long Run," *Advances in Macroeconomics*, August 2001, Vol. 1, No. 2.

Jones, Charles I. (2002), "Sources of U.S. Economic Growth in a World of Ideas," *American Economic Review*, *92* (1), 220-239.

Jones, Charles I. (2003), "Population and Ideas: A Theory of Endogenous Growth," in P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford, eds., *Knowledge, Information, an Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, Princeton University Press.

Jones, Charles I. (2005), "Growth and Ideas," in P. Aghion and S. Durlauf, eds., *Handbook of Economic Growth*, Elsevier, Volume 1B, 1063-1101.

Karras (2001), "Long-Run Economic Growth in Europe: Is It Endogenous or Neoclassical?" *International Economic Journal*, 15(2), 63-76.

Kremer, Michael (1993), “Population Growth and Technological Change: One Million B.C. to 1990,” *Quarterly Journal of Economics*, *108* (4), 681-716.

Lee, Ronald D.(1988), “Induced Population Growth and Induced Technological Progress: Their Interaction in the Accelerating Stage,” *Mathematical Population Studies*, *1* (3), 265-288.

Lucas, Robert E.(1988), “On the Mechanics of Economic Development,” *Journal of Monetary Economics*, *22* (1), 3-42.

Maddison, Angus, *The World Economy: Historical Statistics*, Paris: Organization for Economic Cooperation and Development, 2005.

Mankiw, N. Gregory, David Romer, and David Weil (1992), “A Contribution to the Empirics of Economic Growth,” *Quarterly Journal of Economics*, *107*(2), 407-438.

Nelson, C. R., and C. I. Plosser (1982), “Trends and Random Walks in Macro-Economic Time Series: Some Evidence and Implications.” *Journal of Monetary Economics*, 10, 139-162.

Ng, S., and P. Perron (1995), “Unit Root Test in ARMA Models with Data dependent Methods for the Selection of the Truncation Lag.” *Journal of the American Statistical Association*, 90, 268-81.

Papell, David and Ruxandra Prodan (2007), “Restricted Structural Change and Unit Root Hypothesis,” *Economic Inquiry*, *45*, 834-853.

Papell, David and Ruxandra Prodan (2011), “The Statistical Behavior of GDP after Financial Crises and Severe Recessions,” unpublished, University of Houston

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

Rebelo, Sergio (1991), “Long-Run Policy Analysis and Long-Run Growth,” *Journal of Political Economy*, *99*, 500-521.

Romer, Paul M. (1986), “Increasing Returns and Long-Run Growth,”*Journal of Political Economy*, *94*, 1002-1037.

Romer, Paul (1990), “Endogenous Technological Change,” *Journal of Political Economy*, *98* (5), 71-102.

Romer Paul (1992), “Two Strategies for Economic Development: Using Ideas and Producing Ideas,” *Proceedings of the World Bank Annual Conference on Development Economics,* 63-115.

Segerstrom, Paul (1998), “Endogenous Growth Without Scale Effects,” *American Economic Review*, *88* (5), 1290-1310.

Solow, Robert M.(1956), “A Contribution to the Theory of Economic Growth,” *Quarterly Journal of Economics*, *70* (1), 65-94.

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*, 3, 251-70

Table 1: Unit root tests with and without structural change

$$\Delta y_t = c + (\beta t) + \gamma_1 DU_{1t} + \theta_1 DT_{1t} + \theta_2 DT_{2t} + \alpha y_{t-1} + \sum_{i=1}^{k} \rho_i \Delta y_{t-i} + u_t$$

Country	ADF test $\theta_1 = \theta_2 = 0$ $\gamma_1 = 0$	Model A $\theta_1 = \theta_2 = 0$	Model C $\theta_2 = 0$	Model CB
	t-stat	t-stat	t-stat	t-stat
OECD				
Australia	-1.84	-3.87	-5.12	-6.24*
Austria	-1.78	-3.43	-4.41	-8.77***
Belgium	-1.20	-2.83	-4.64	-7.37***
Canada	-2.17	-4.48	-4.80	-6.73**
Denmark	-2.22	-4.21	-3.91	-9.11***
Finland	-2.27	-3.48	-4.36	-5.26
France	-1.72	-4.04	-4.63	-10.1***
Germany	-3.03	-4.63	-5.17*	-9.01***
Italy	-2.20	-3.69	-3.45	-6.50**
Japan	-1.68	-4.21	-3.36	-13.67***
Netherlands	-2.01	-3.55	-4.68	-7.13***
New Zealand	-2.01	-3.58	-3.81	-5.02
Norway	-2.58	-3.52	-3.90	-4.46
Portugal	-1.64	-4.14	-2.98	-4.43
Spain	-1.24	-3.51	-5.29*	-6.92**
Sweden	-2.39	-3.94	-4.08	-8.77**
Switzerland	-2.44	-5.14**	-4.29	-5.47
UK	-0.49	-2.61	-4.97	-6.85**
US	-3.27*	-4.49	-6.64***	-8.27***
ASIA				
India	2.68	-3.25	-2.59	-3.90
Indonesia	-0.26	-1.94	-4.49	-6.38*
Malaysia	-0.68	-4.35	-4.02	-6.32*
Philippines	-2.31	-5.15**	-4.52	-5.48
Taiwan	-1.34	-3.48	-5.94**	-8.19***
South Korea	-1.35	-8.39***	-9.86***	-9.96***
Sri Lanka	1.57	-2.66	-2.67	-6.07*

Critical values: ADF test: -3.15 (10%), -3.45 (5%), 4.04 (1%);
Model A: -4.79 (10%), -5.04 (5%), -5.69(1%);
Model C: -5.15 (10%), -5.42 (5%), -5.96(1%);
Model CB: -6.05 (10%), -6.40 (5%), -6.96 (1%);

Table 2. Structural change tests allowing for a maximum of 3 breaks

$$y_t = \mu + \beta t + \gamma_1 DU_{1t} + \theta_1 DT_{1t} + \theta_2 DT_{2t} + \theta_3 DT_{3t} + \sum_{i=1}^{k} \rho_i y_{t-i} + u_t$$

Country	Test	Statistic	Critical values	Nr. of breaks	Break dates	β	γ	θ_1	$\theta_2 ; \theta_3$
OECD									
Australia	0 to 1 1 to 2 2 to 3	22.79* 13.27 18.95	21 15.2 16.59	1	1931	0.0054	-0.0460	0.0158	
Austria	0 to 1 1 to 2 2 to 3	23.55 99.28 15.46*	22.18 14.52 15.01	3	1944 1950 1976	0.0084	-1.0063	0.1646	-0.1280 -0.0254
Belgium	0 to 1 1 to 2 2 to 3	18.37 31.86* 5.52	23.05 15.35 17.2	2	1939 1976	0.0086	-0.3205	0.0251	-0.0143
Canada	0 to 1 1 to 2 2 to 3	13.38 15.22* 3.24	18.56 11.67 13.34	2	1930 1940	0.0197	-0.5589	0.0498	-0.0472
Denmark	0 to 1 1 to 2 2 to 3	15.01 48.08* 10.02	20.69 11.64 13.2	2	1939 1969	0.0163	-0.2519	0.0179	-0.0154
France	0 to 1 1 to 2 2 to 3	23.22 53.41* 3.35	20.06 13.76 15.56	2	1939 1973	0.0127	-0.4711	0.0326	-0.0298
Germany	0 to 1 1 to 2 2 to 3	19.15 66.71* 5.67	19.07 12.23 13.98	2	1944 1950	0.0131	-0.8467	0.1485	-0.1353
Italy	0 to 1 1 to 2 2 to 3	26.39 13.83* 9.05	21.73 14.25 17.34	2	1942 1948	0.0137	-0.6290	0.1168	-0.0996
Japan	0 to1 1 to 2 2 to 3	16.45 161.92 25.34*	23.01 15.13 18.09	3	1944 1971 1991	0.0183	-0.7899	0.0599	-0.0456 -0.0225
Netherlands	0 to1 1 to 2 2 to 3	29.81 26.09* 1.97	20.7 13.44 15.09	2	1945 1951	0.0076	-0.1783	0.0748	-0.0589
Spain	0 to 1 1 to 2 2 to 3	28.97 24.23 25.83*	24.24 16.13 19.65	3	1935 1959 1971	0.0093	-0.3737	0.0132	0.0489 -0.0444
Sweden	0 to 1 1 to 2 2 to 3	14.19 23.27* 11.54	22.51 14.69 16.11	2	1915 1970	0.0140	-0.2575	0.0157	-0.0124

Table 2. Structural change tests allowing for a maximum of 3 breaks - continued

Switzerland	0 to 1 1 to 2 2 to 3	25.28* 8.69 5.95	20.46 13.24 15.89	1	1944	0.0153	0.3156	0.0013	
UK	0 to 1 1 to 2 2 to 3	32.39 19.64* 15.89	20.99 13.61 16.41	2	1939 1945	0.0086	0.2938	-0.0444	0.0563
US	0 to 1 1 to 2 2 to 3	15.23 25.26* 12.34	17.53 11.41 13.18	2	1929 1942	0.0169	-0.4010	0.0323	-0.0289
ASIA									
Indonesia	0 to 1 1 to 2 2 to 3	28.02* 6.54 6.00	23.4 16.37 19.35	1	1941	0.0107	-0.5429	0.0160	
Philippines	0 to 1 1 to 2 2 to 3	16.28 22.22* 6.31	20.8 13.53 16.82	2	1946 1952	0.0114	-0.5844	0.0843	-0.0832
South Korea	0 to 1 1 to 2 2 to 3	85.43* 5.86 21.46	15.59 24.32 27.22	1	1944	0.0166	-1.0686	0.0413	
Sri Lanka	0 to 1 1 to 2 2 to 3	14.39 23.29* 5.47	24.53 15.69 17.38	2	1900 1966	0.0156	-0.0057	-0.0144	0.0285
Malaysia	0 to 1 1 to 2 2 to 3	19.65* 7.98 6.02		1	1944	0.0178	-0.5592	0.0192	
Taiwan	0 to 1 1 to 2 2 to 3	40.27* 9.26 19.41	23.61 16 19.21	1	1942	0.0192	-0.9518	0.0400	

Note:
For the three countries where we find three breaks, Austria ,Spain and Japan we report the slope coefficient on the third break (θ_3) instead of the slope coefficient on the second break (θ_2).

Table 3. F- tests – restricted structural change – constant trend and same income

$$y_t = c + (\beta t) + \gamma_1 DU_{1t} + \theta_1 DT_{1t} + \theta_2 DT_{2t} + \theta_3 DT_{3t} + \sum_{i=1}^{k} \rho_i y_{t-i} + u_t$$

$\sum_{\ell=1}^{3} \theta_\ell = 0$ and $\gamma_1 + \theta_1 (Tb_i - Tb_1) = 0$, for i = 2, 3

Country	Break dates	F-statistic	p-value	R/FTR
OECD				
Austria	1944, 1950, 1976	26.871	0.000	R
Belgium	1939, 1976	23.792	0.000	R
Canada	1930, 1940	0.172	0.842	FTR
Denmark	1939, 1969	22.445	0.000	R
France	1939, 1973	31.074	0.000	R
Germany	1944, 1950	4.724	0.010	R
Italy	1942, 1948	5.839	0.003	R
Japan	1944,1971,1991	11.601	0.000	R
Netherlands	1945,1951	22.25	0.000	R
Spain	1935, 1959, 1971	19.29	0.000	R
Sweden	1915,1970	17.06	0.000	R
UK	1939,1945	24.43	0.000	R
US	1929, 1942	2.336	0.100	FTR
ASIA				
Philippines	1946,1952	13.87	0.000	R
Sri Lanka	1900,1966	17.61	0.000	R

Table 4. F- tests – restricted structural change – constant trend

$$y_t = c + (\beta t) + \gamma_1 DU_{1t} + \theta_1 DT_{1t} + \theta_2 DT_{2t} + \theta_3 DT_{3t} + \sum_{i=1}^{k} \rho_i y_{t-i} + u_t \text{, where} \sum_{\ell=1}^{3} \theta_\ell = 0$$

Country	Break dates	F-Statistic	p-value	R/FTR
OECD				
Austria	1944, 1950, 1976	3.68	0.057	FTR
Belgium	1939, 1976	19.15	0.000	R
Denmark	1939, 1969	2.289	0.132	FTR
France	1939, 1973	4.360	0.038	R
Germany	1944, 1950	2.395	0.124	FTR
Italy	1942, 1948	1.428	0.234	FTR
Japan	1944,1971,1991	12.568	0.000	R
Netherlands	1945,1951	31.53	0.000	R
Spain	1935, 1959, 1971	19.988	0.000	R
Sweden	1915,1970	2.495	0.116	FTR
UK	1939,1945	48.470	0.000	R
ASIA				
Philippines	1946,1952	12.93	0.000	R
Sri Lanka	1900,1966	7.23	0.008	R

Table 5. F- tests – restricted structural change – constant trend for countries with 1 break

$y_t = c + (\beta t) + \gamma_1 DU_{1t} + \theta_1 DT_{1t} + \sum_{i=1}^{k} \rho_i y_{t-i} + u_t$, **where** $\theta_1 = 0$

Country	Break date	F-statistic	p-value	R/FTR
Australia	1931	18.48	0.000	R
Switzerland	1944	0.292	0.589	FTR
Indonesia	1941	23.53	0.000	R
South Korea	1944	70.64	0.000	R
Malaysia	1944	14.90	0.000	R
Taiwan	1942	37.89	0.000	R

Table 6. F- tests – restricted structural change – constant trend (1870-2058)

$y_t = c + (\beta t) + \gamma_1 DU_{1t} + \theta_1 DT_{1t} + \theta_2 DT_{2t} + \theta_3 DT_{3t} + \sum_{i=1}^{k} \rho_i y_{t-i} + u_t$ **, where** $\sum_{\ell=1}^{3} \theta_\ell = 0$

Country	Break dates	F-Statistic	p-value	R/FTR
OECD				
Belgium	1939, 1973	15.27	0.00	R
France	1939, 1973	1.35	0.25	FTR
Japan	1944,1973	25.37	0.00	R
Netherlands	1945,1973	15.99	0.00	R
Spain	1935,1973	50.27	0.00	R

Table 7. Pre - (first) break and post - (last) break growth rates

Country	Break dates	Pre - (first) break growth rate	Post - (last) break growth rate	Ratio Post - (last) break/Pre - (first) break
Australia	1931	0.54%	2.12%	3.93
Austria	1944, 1950, 1976	0.84%	1.96%	2.33
Belgium	1939, 1976	0.86%	1.94%	2.26
Canada	1930, 1940	1.97%	2.23%	1.13
Denmark	1939, 1969	1.63%	1.88%	1.15
France	1939, 1973	1.27%	1.55%	1.22
Germany	1944, 1950	1.31%	2.63%	2.01
Italy	1942, 1948	1.37%	3.09%	2.26
Japan	1944, 1971	1.83%	2.42%	1.32
Netherlands	1945, 1951	0.76%	2.35%	3.09
Spain	1935, 1959, 1971	0.93%	2.70%	2.90
Sweden	1915,1970	1.40%	1.73%	1.24
Switzerland	1944	1.53%	1.66%	1.08
UK	1939, 1945	0.86%	2.05%	2.38
US	1929, 1942	1.69%	2.03%	1.20
Indonesia	1941	1.07%	3.85%	2.50
Philippines	1946,1952	1.14%	1.25%	1.10
South Korea	1944	1.66%	5.79%	3.49
Sri Lanka	1900,1966	1.56%	2.97%	1.90
Malaysia	1944	1.78%	3.70%	2.08
Taiwan	1942	1.92%	5.92%	3.08

Figure 1. The *"linear trend"* restriction

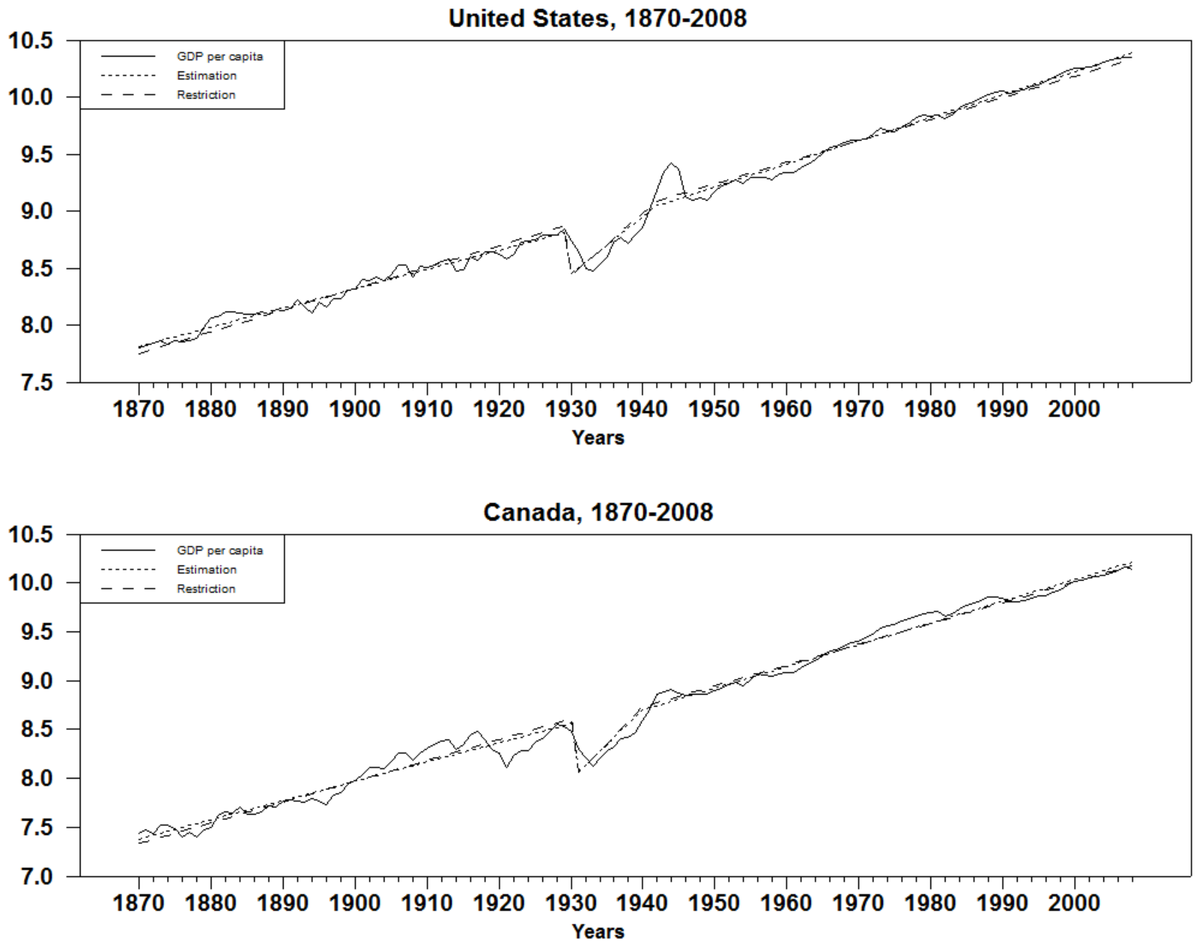
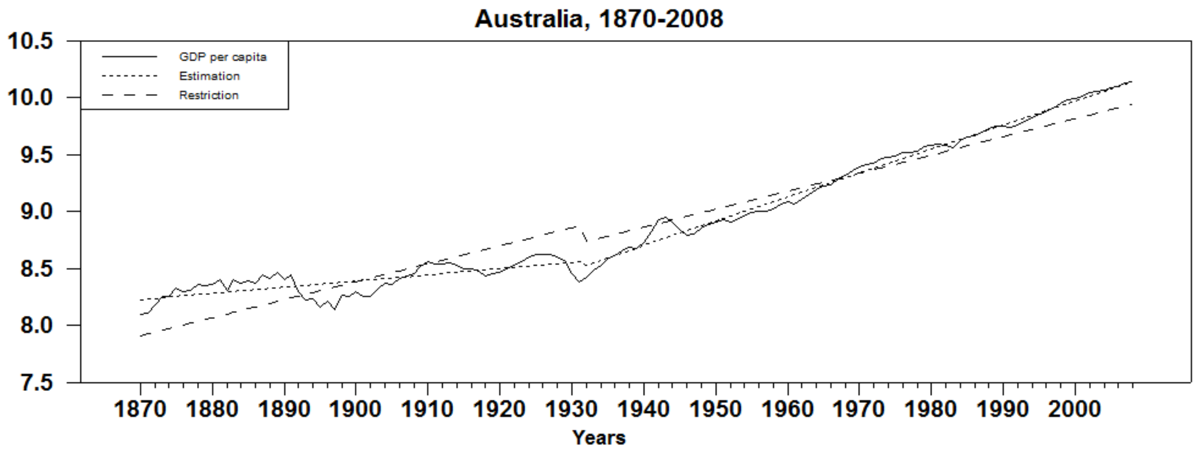
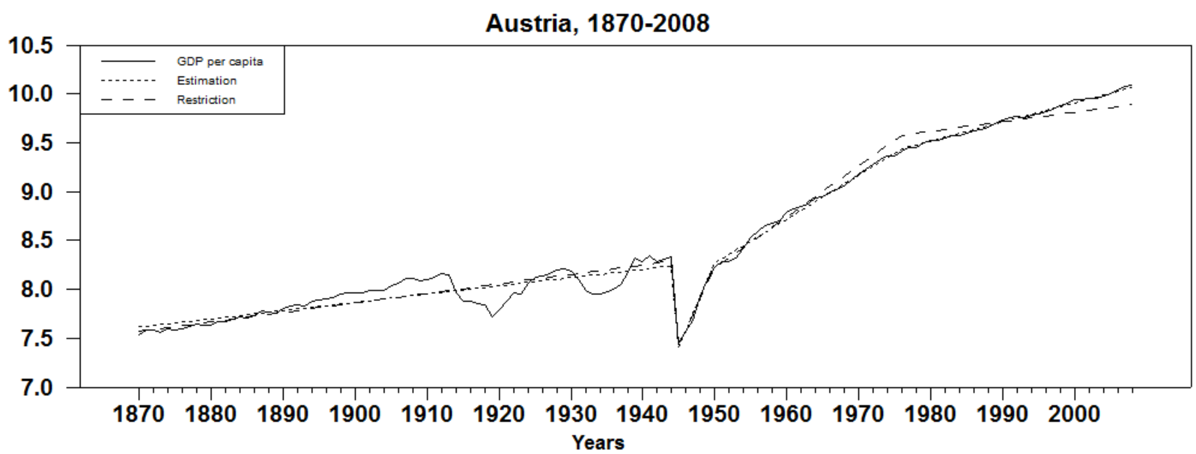
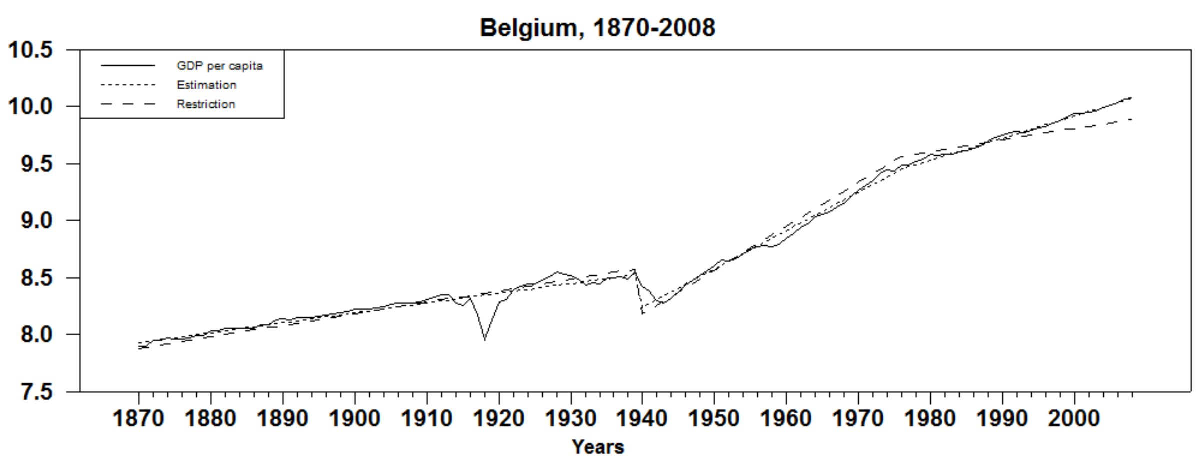


Figure 2. The *"level shift"* restriction

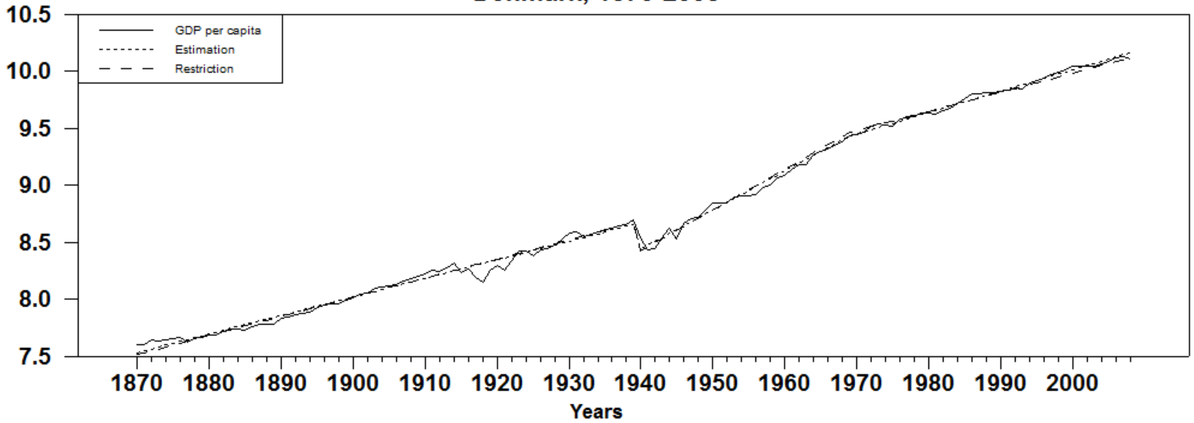
OECD Countries



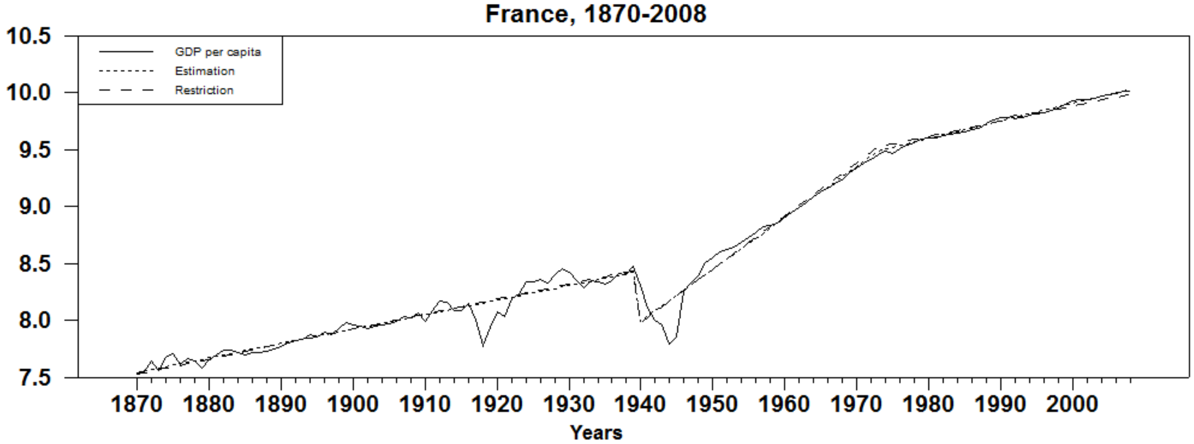




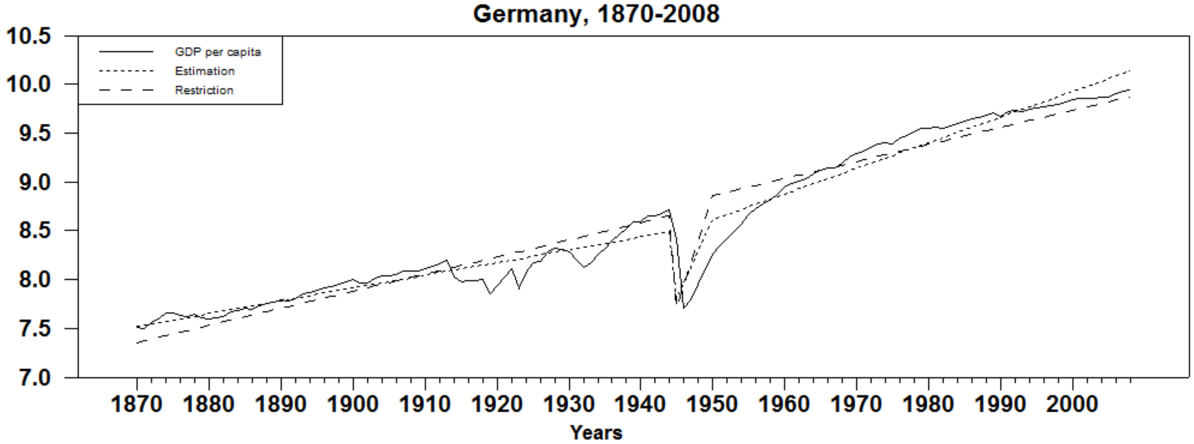
Denmark, 1870-2008



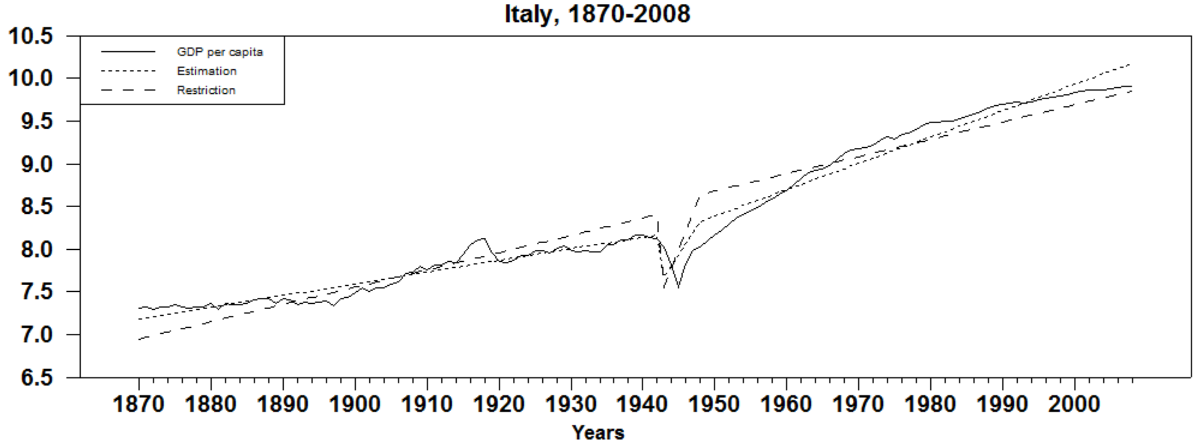
France, 1870-2008



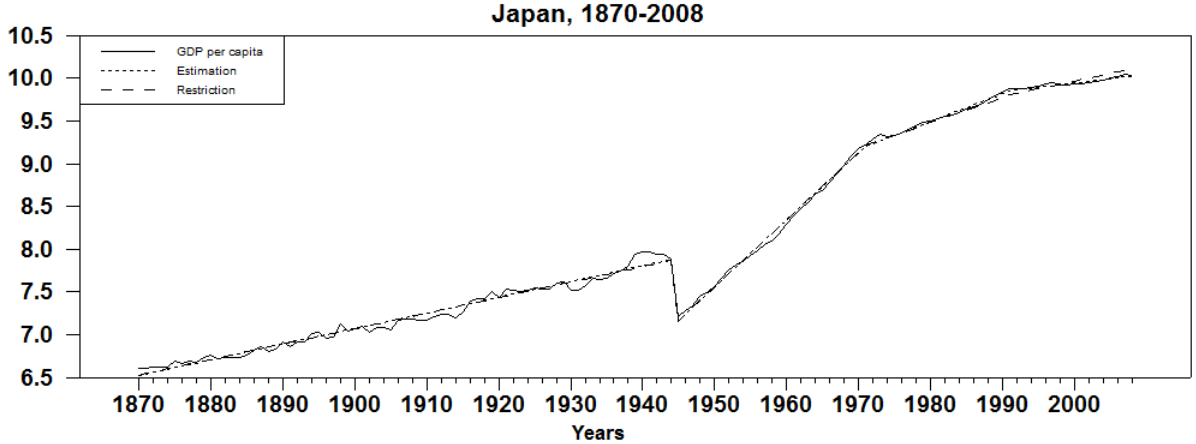
Germany, 1870-2008



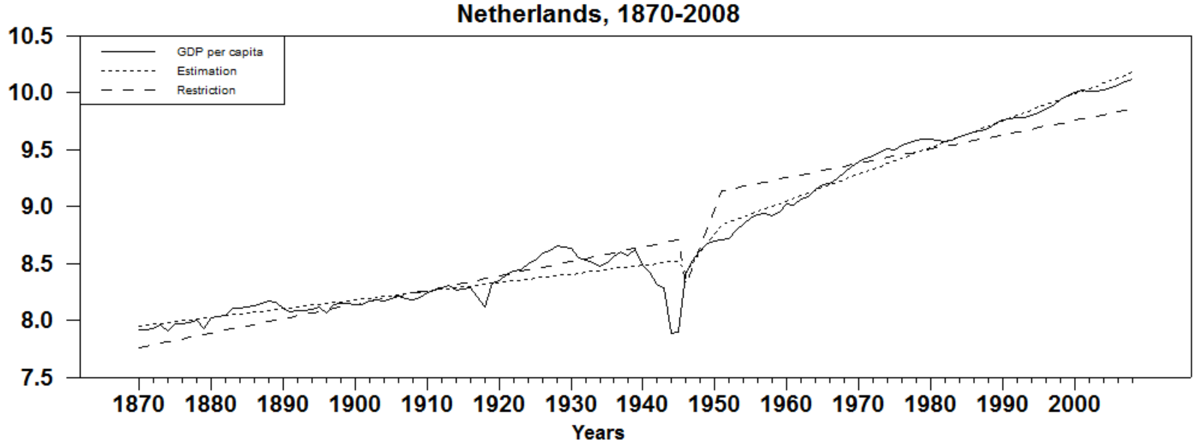
Italy, 1870-2008

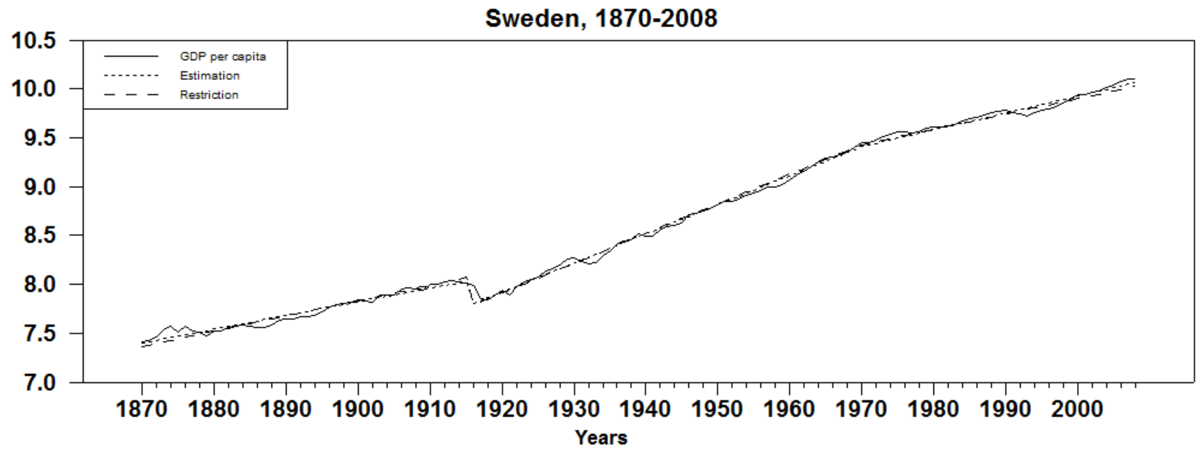


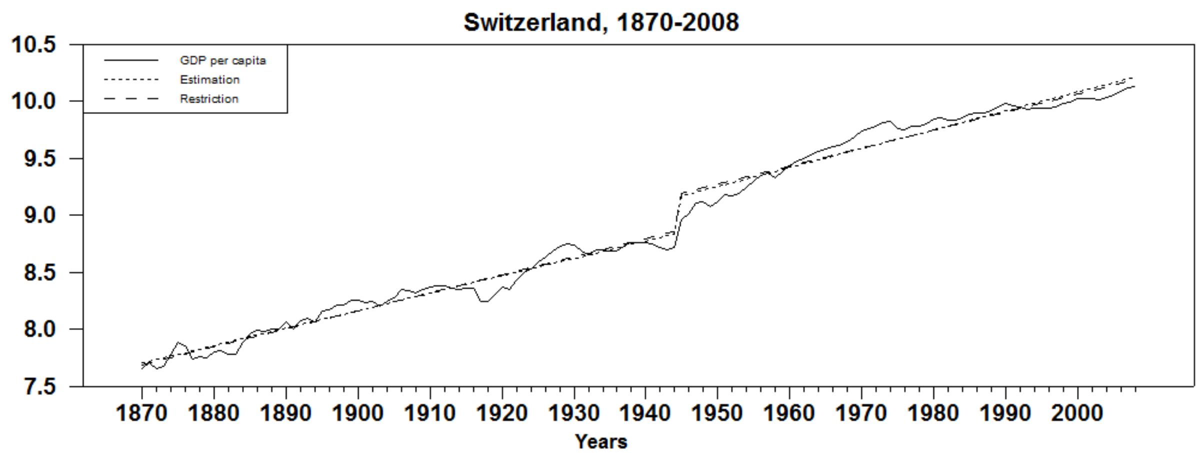
Japan, 1870-2008

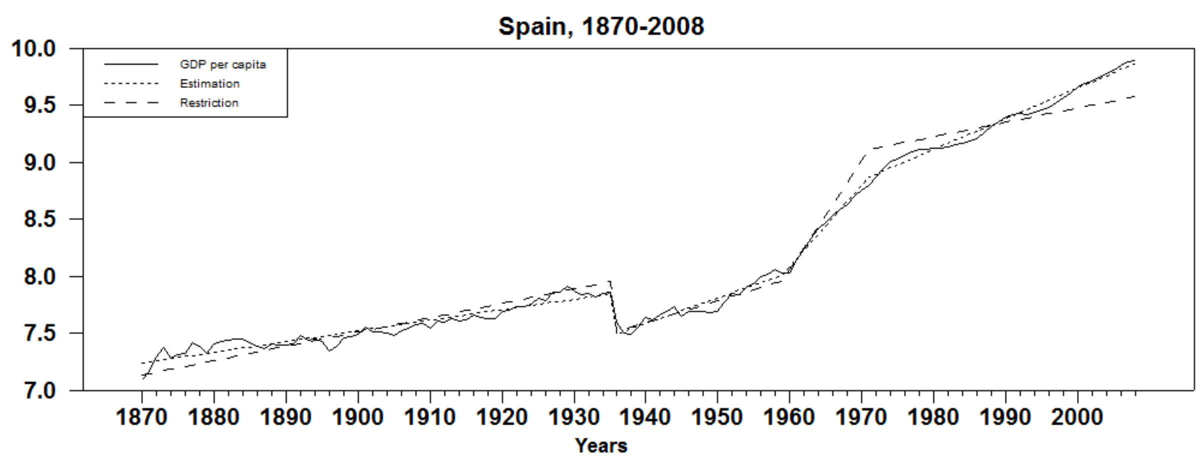


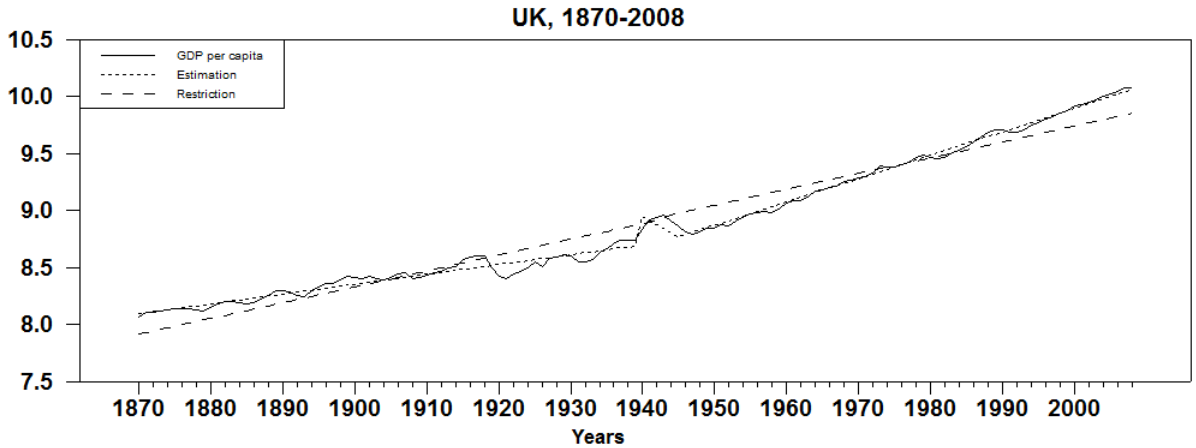
Netherlands, 1870-2008



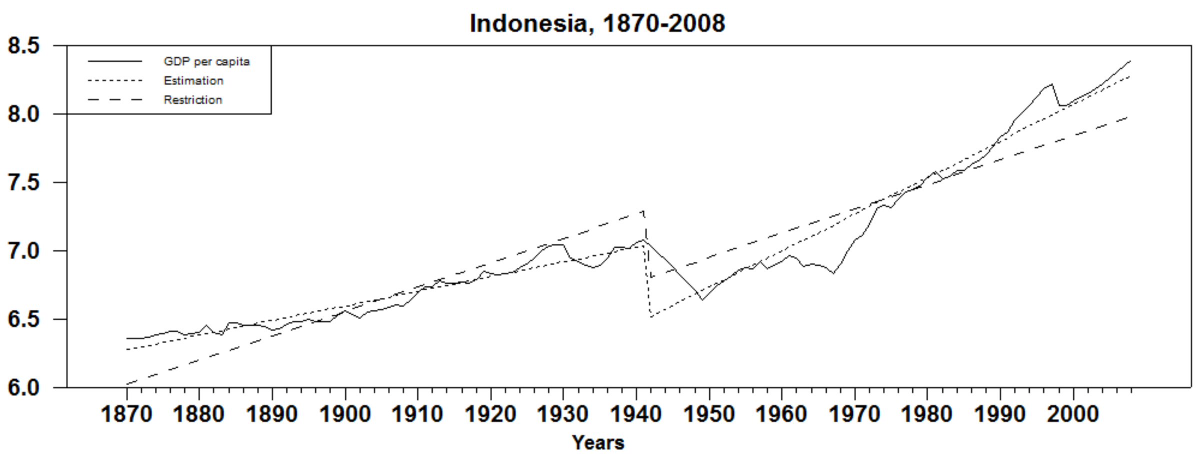

Sweden, 1870-2008
GDP per capita
Estimation
Restriction
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000
Years

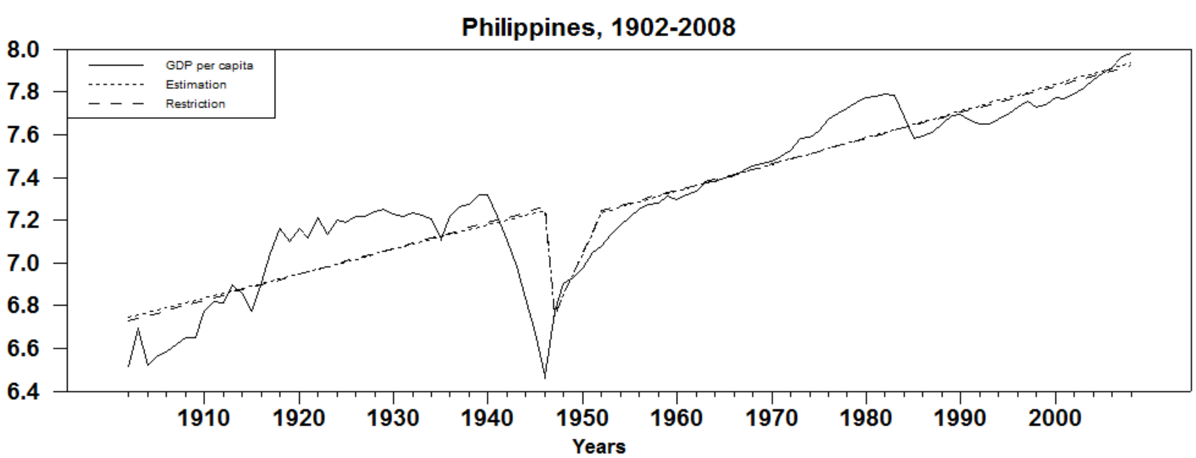

Switzerland, 1870-2008
GDP per capita
Estimation
Restriction
10.5
10.0
9.5
9.0
8.5
8.0
7.5
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000
Years

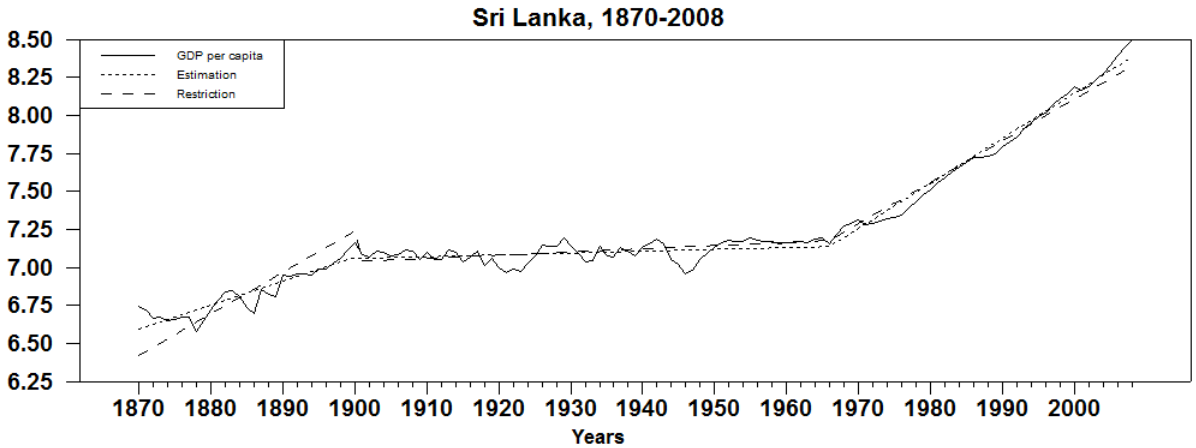

Spain, 1870-2008
GDP per capita
Estimation
Restriction
10.0
9.5
9.0
8.5
8.0
7.5
7.0
1870 1880 1890 1900 1910 1920 1930 1940 1950 1960 1970 1980 1990 2000
Years

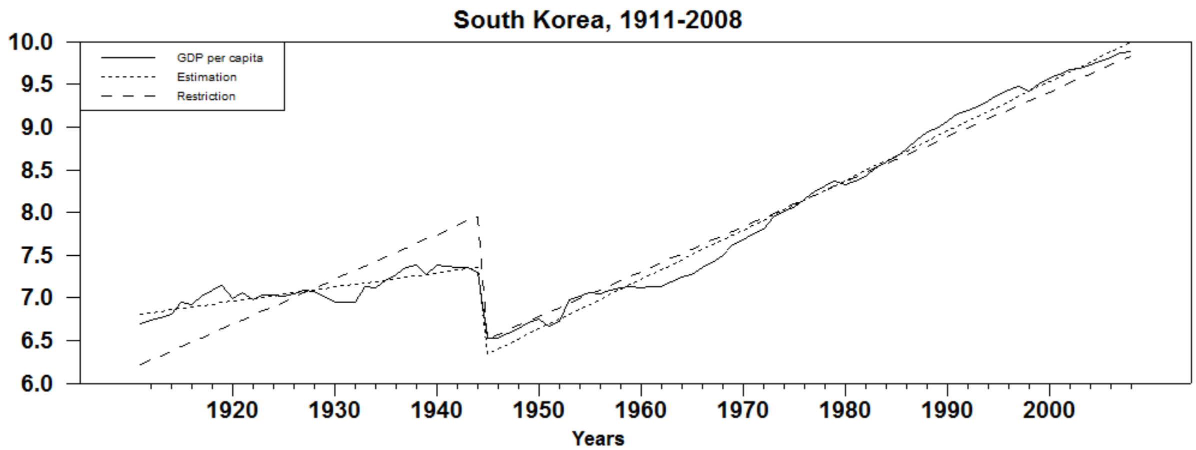


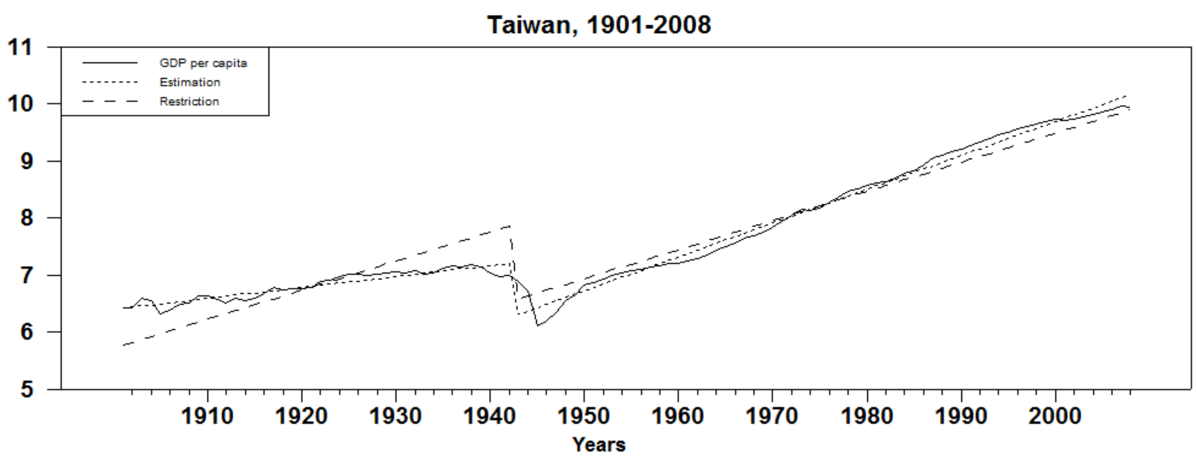
Asian countries

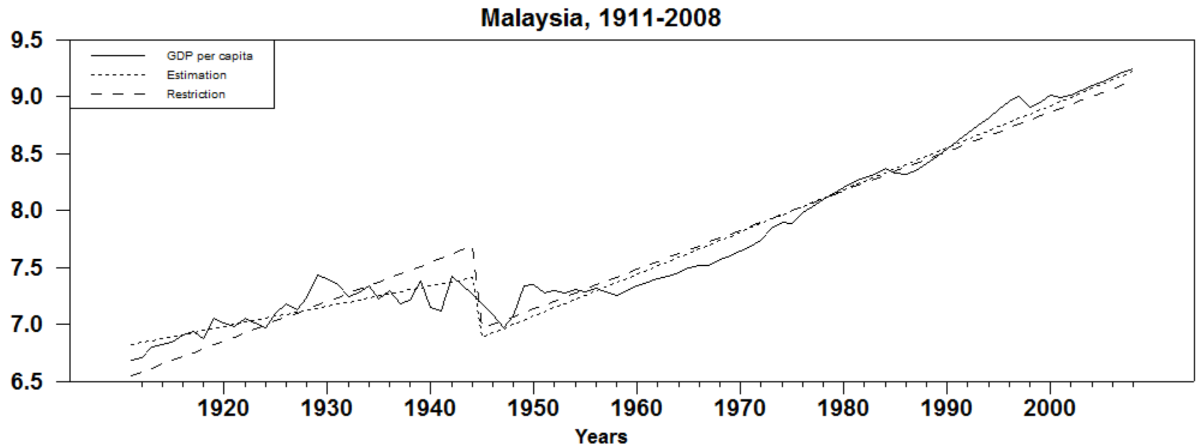





Sri Lanka, 1870-2008
GDP per capita
Estimation
Restriction
8.50
8.25
8.00
7.75
7.50
7.25
7.00
6.75
6.50
6.25
1870
1880
1890
1900
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000
Years


South Korea, 1911-2008
GDP per capita
Estimation
Restriction
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
1920
1930
1940
1950
1960
1970
1980
1990
2000
Years


Taiwan, 1901-2008
GDP per capita
Estimation
Restriction
11
10
9
8
7
6
5
1910
1920
1930
1940
1950
1960
1970
1980
1990
2000
Years



Malaysia, 1911-2008
9.5
9.0
8.5
8.0
7.5
7.0
6.5
GDP per capita
Estimation
Restriction
1920
1930
1940
1950
1960
1970
1980
1990
2000
Years