The Uncertain Unit Root in U.S. Real GDP: Evidence with Restricted and Unrestricted Structural Change

David H. Papell and Ruxandra Prodan

David H. Papell *is a professor of Economics at the University of Houston. E-mail:* dpapell@uh.edu
Ruxandra Prodan *is a graduate student in Economics at the University of Houston. E-mail:*
rprodan@mail.uh.edu

1. Introduction

The trend-stationary model, where business cycles were modeled as stationary fluctuations around a linear deterministic trend, was the canonical representation of aggregate output until the publication of Nelson and Plosser (1982), who argued in favor of a difference-stationary model where current shocks have a permanent effect on the long-run level of macroeconomic and financial aggregates. Using Augmented-Dickey-Fuller (ADF) tests, they could not reject the unit root null hypothesis against the trend stationary alternative for 13 out of 14 long-term annual U.S. macro series, including real GNP.

Nelson and Plosser's findings of a unit root in U.S. real GNP have been called into question on a number of grounds. Much research has focused on studying the performance of conventional unit root tests: Rudebusch (1993) proposes a bootstrap approach to evaluate the tests. Using postwar data, he demonstrates that unit root tests have low power against economically relevant trend-stationary alternatives. Diebold and Senhadji (1996), using long span annual U.S. GNP data, argue that Rudebusch's procedure produces evidence that favors trend-stationarity.

Ben-David and Papell (1995) and Ben-David, Lumsdaine and Papell (2003), using tests for a unit root against the alternative of broken trend-stationarity allowing for one and respectively two endogenous break points, developed by Zivot and Andrews (1992) and Lumsdaine and Papell (1997), reject the unit root null in favor of broken trend-stationarity for long-term U.S. GDP. In all cases, the estimated breaks coincide with the Great Depression and/or World War II.

Murray and Nelson (2002) examine the performance of ADF tests in the context of models where the effects of the Great Depression are large, but transitory, based on a parametric bootstrap of a Markov switching model for real GDP from 1870-1994. They find that when a temporary component from 1930-1945 is added to an underlying unit root process, the unit root hypothesis is (incorrectly) rejected too often. Kilian and Ohanian (2002) examine the performance of Zivot and Andrews unit root tests when both the Great Depression and World War II produce transitory fluctuations. They also find that the unit root hypothesis is rejected too often.

While a single structural change is (by definition) permanent, multiple structural changes can, in principle, be consistent with transitory fluctuations. In this paper we use Lumsdaine and Papell (1997) tests for a unit root in the presence of unrestricted structural change (two endogenous breaks) and Papell and Prodan (2003) tests for a unit root in the presence of restricted structural change (two offsetting structural changes) to investigate long-term U.S. real GDP. The unit root null is rejected at the 1% level with both tests. Using the Murray and Nelson (2002) procedure to evaluate the rejections, we find that, in contrast with ADF tests, the unit root null is not rejected too often if the data generating process is a unit root with transitory fluctuations.

2.   Unit root tests for long-term real GDP

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:

$$\Delta y_t = \mu + \beta t + \gamma_1 DU1_t + \gamma_2 DU2_t + \alpha y_{t-1} + \sum_{i=1}^{k} c_i \Delta y_{t-i} + \varepsilon_t,$$

(1)

for t = 1,….,T, where $DU1_t$ and $DU2_t$ are indicator dummy variables for the intercept changes in the trend function, occurring at times TB_1 and TB_2. That is, $DU1 = 1$ ($t>TB_1$) and $DU2 = 1$ ($t>TB_2$).

We use the "general-to-specific" recursive t-statistic procedure suggested by 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 null hypothesis of a unit root is rejected in favor of trend stationarity if α is significantly different from zero. 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.

Three types of models are estimated. The standard ADF test sets $\gamma_1 = \gamma_2 = 0$ and tests the null hypothesis of a unit root in favor of the alternative of *trend-stationarity*. Model AA allows for two breaks

in the intercept of the trend function (Lumsdaine and Papell, 1997), and tests the null hypothesis of a unit root in favor of the alternative of *broken trend-stationarity*. Model AA restricted also allows for two breaks in the intercept of the trend function (Papell and Prodan, 2003), but restricts the coefficients on the dummy variables that depict the breaks to be equal and opposite in sign, $\gamma_1 + \gamma_2 = 0$, so that the trend following the second break is equal the trend prior to the first break. The alternative becomes *restricted trend-stationarity*.

Using the data of Maddison (2001) for the period 1870-1998, we reject the unit root null against the alternatives of trend-stationarity, broken trend-stationarity, and restricted trend-stationarity. Bootstrap critical values with the data generated under the null hypothesis, using 129 observations and 5000 replications, are calculated for each of the three 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. Unit Root Tests						
	α	Break1	Break 2	$\hat{\tau}$	p-values	Lag
ADF	-0.23	-	-	-3.86	0.027	6
Model AA:	-0.58	1929	1940	-7.68	0.000	7
Model AA restricted:	-0.38	1929	1940	-7.11	0.000	2

Critical values: ADF test: -4.22 (1%), -3.60 (5%), -3.29 (10%); Model AA: -6.68 (1%), -6.10 (5%), -5.85 (10%); Model AA restricted: -6.32 (1%), -5.80 (5%), -5.54 (10%).

The critical values for all the above mentioned tests are based on the maintained hypothesis of homoskedasticity. Murray and Nelson (2002) point out that the variance was much higher around the period of the Great Depression, and generally was lower after World War II. Consequently, the heterogeneity present in the data can lead to incorrect rejection of the unit root null using standard unit root tests. We proceed to investigate whether this criticism extends to the tests for unit roots in the presence of restricted and unrestricted multiple structural changes.

3. Unit root tests and transitory fluctuations

We perform the simulation experiment proposed by Murray and Nelson (2002) for the three tests. First, we estimate a simple reduced-form model that captures the view that there were large transitory

fluctuations during the Great Depression. This model generates time series as a sum of the latent random walk with drift and occasional large transitory movements, driven by a regime switching process.

The model is the following:

$$y_t = \tau_t + S_t z_t, \quad \text{where} \quad S_t = 1 \quad \text{from 1930-1945} \tag{2}$$

$$\tau_t = g + \tau_{t-1} + v_t \tag{3}$$

$$z_t = \varphi_1 z_{t-1} + \varphi_2 z_{t-2} + u_t \tag{4}$$

$$\begin{pmatrix} v_t \\ u_t \end{pmatrix} \approx N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_v & 0 \\ 0 & \sigma^2_u \end{bmatrix} \right) \tag{5}$$

The observed time series (y_t) consists of a trend component (τ_t) and a second component (z_t), which will be present only if an indicator variable (S_t), governed by a Markov process, is unity. This second component, if present, adds to the volatility of output and may also have different dynamics, so we allow it to have an AR(2) structure. Estimation is done by the approximate maximum likelihood method of Kim (1994). The parameters are reported in Table 2. Next, we use this model as a data generating process in our bootstrap experiments and generate 5000 series. We perform the previously discussed unit root tests on these constructed series and compute the test-statistic (denoted by $\hat{\tau}$) for each realization. Then we calculate the exact probability of obtaining the sample value of the test statistic using each test.

Table 2. Parameter Estimates of Markov-Switching Model for Annual U.S. Real GDP	
Parameter	Estimate
ϕ_1	1.2357
ϕ_2	-0.3817
g	0.0337
σ_v	0.0263
σ_u	0.0591

Using the above data generating process and the coefficient estimates we find p-values for $\hat{\tau}$:

ADF: $\qquad$ *Probability* $[\hat{\tau} < -3.86] = 0.094$

Model AA: $\qquad$ *Probability* $[\hat{\tau} < -7.68] = 0.054$

Model AA restricted: *Probability* $[\hat{\tau} < -7.11] = 0.050$

Our findings are threefold: First, we confirm Murray and Nelson's (2002) result for the ADF test that the unit root null is incorrectly rejected too often.[i] With a data generating process that contains a unit root, 9.4% of the realized test statistics are greater than (in absolute value) the t-statistic for the actual data. Second, the marginal significance level for rejection of the null hypothesis for the AA model is 0.054, which demonstrates that is even less likely that this test would incorrectly reject the null. Third, the marginal significance level for rejection of the null hypothesis for the restricted AA model is 0.05, making it even more unlikely that this test would incorrectly reject the null. We conclude that the rejections of the unit root null in favor of broken trend-stationarity and restricted trend-stationarity for long-span U.S. real GDP are not subject to the Murray and Nelson critique.

References

Ben-David, Dan, and David H. Papell (1995). "The Great Wars, The Great Crash and The Unit Root Hypothesis," *Journal of Monetary Economics*, December, 453-475.

Ben-David, Dan, Robin Lumsdaine, and David H. Papell (2003). "Unit Root, Postwar Slowdowns and Long-Run Growth: Evidence from Two Structural Breaks." *Empirical Economics*, February, 303-319.

Diebold, Francis X., and Abdelhak S. Senhadji (1996). "The Uncertain Root in Real GNP: Comment." *American Economic Review*, 86, 1291-1298.

Kilian, Lutz, and Lee Ohanian (2002). "Unit Roots, Trend Breaks and Transitory Dynamics: A Macroeconomic Perspective." *Macroeconomic Dynamics*, 6, 614-631.

Kim, Chang-Jin (1994). "Dynamic Linear Models with Markov Switching." *Journal of Econometrics,* 60, 1-22.

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

Maddison, Angus (2001). "The World Economy - A Millennial Perspective." *Development Centre Studies*, OECD.

Murray, Christian, and Charles R. Nelson (2002). "The Great Depression and Output Persistence." *Journal of Money, Credit and Banking*, 34, 1090-1098.

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

Papell, David H., and Ruxandra Prodan (2003). "Restricted Structural Change and the Unit Root Hypothesis." working paper, University of Houston.

Rudebusch, Glenn D. (1993). "The Uncertain Unit Root in Real GNP." *The American Economic Review*, 83, 264-272.

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, 251-270.

---

[i] Murray and Nelson (2002), using coefficient estimates from U.S. real GDP 1870-1994 data, found a probability of rejection of 0.108. The difference comes from the four additional years and data revisions by Maddison.