Forthcoming, *Journal of Money, Credit and Banking*

**The Great Depression and Output Persistence:**
A Reply to Papell and Prodan

**Christian J. Murray**
*University of Houston*

**Charles R. Nelson**
*University of Washington*

November 2003

Corresponding author:
Charles R. Nelson
Box 353330
University of Washington
Seattle, WA 981195-3330
206 685 1382
cnelson@u.washington.edu
In a letter to this journal, Papell and Prodan (2003a) (hereafter PP), conduct unit root tests [from Lumsdaine and Papell (1997) and Papell and Prodan (2003b)] that allow for two level shifts in the trend function for annual U.S. real GDP, 1870-1998. The level shifts in the Lumsdaine and Papell (1997) test are unrestricted, whereas the level shifts in Papell and Prodan (2003b) are restricted to offset one another. PP report p-values for tests of the unit root null of about 0.01 from a parametric bootstrap of their regression equation under homoskedasticity, but 0.05 if the bootstrap is instead based on the unobserved components representation of Murray and Nelson (2002) (hereafter MN) which is designed to incorporate into the null the heteroskedastic turmoil of the Great Depression and WWII. PP reach two conclusions. First, contrary to MN, the evidence favors the alternative hypothesis of a trend stationarity for long run U.S. GDP, with a temporary structural break during the Depression and WWII period. Since they restrict the structural change to be transitory, an implication of their model is that there were no permanent shocks to GDP from 1870 to 1998. Second, rejection of the unit root is not subject to the MN critique that heteroskedasticity matters.

We consider the second point first. It seems to us that there is a very great difference between a p-value of 0.01 (under homoskedasticity) and 0.05 (under heteroskedasticity) as measures of the strength of the evidence against the unit root hypothesis, since one may care a lot whether the probability of a Type 1 error is 1 in 100 or 1 in 20. The difference could be irrelevant if one approaches the data with a particular p-value in mind, say 0.05 or above, as the decision criterion to accept or reject. But if one’s decision required a p-value of 0.04, then inference is reversed by the MN critique.

Turning to PP’s rejection of the unit root, we note that it is based the empirical distribution of their test statistic in MN’s parametric model of real GDP that incorporates the unit root as well as other assumptions: trend is a random walk, the Great Depression–WWII turmoil is a stationary AR(2) that switches on or off according to a Markov process, there is no other business cycle component, and shocks to the trend and Great Depression–WWII components are Normal, homoskedastic, and orthogonal. The p-value of 0.05 reported by PP may reflect actual trend stationarity in the economy, or it may be the result of the violation of one or more of these other maintained assumptions. Since the MN data generating process is driven by the higher volatility of the period through WWII, and noting that the PP break dummies switch in 1929 and 1940, we think it is relevant to consider the evidence from the
more homogeneous post-war period. Figure 1 plots the p-value for the ADF test statistic\(^1\) as a function of the beginning year of the sample. Since the last break in PP occurs in 1940, we compute unit root statistics with data beginning in 1941. For starting dates from 1941 through 1944, p-values are essentially zero, but jump abruptly as the data from WWII is deleted, settling in around 0.50 by the early 1950s. Thus, post-war data contains no evidence at all against the unit root null. This could be because stationarity is so gradual that it does not become evident even over half a century, or it could be that the lower p-values for the longer time series reflect failure to model adequately the extreme heterogeneity of the data over the entire 1870-1998 period.

Next we consider the economic implications of the trend stationary hypothesis. Fitting the PP trend function with restricted structural change, we obtain the stationary deviations from trend plotted in Figure 2. This implied cycle contains a number of features which call into question its credibility as a measure of the cyclical component of economic activity. Consider the following points:

1. The cycle is heterogeneous in spite of break dummies, with WWII an outlier and higher volatility prior to introduction of national income accounting in 1929. Measurement issues and their relation to unit root tests are discussed in Murray and Nelson (2000).

2. The implied cycle is dominated by very low frequency variation. If modeled as an AR(2), it contains a spectral peak corresponding to a period of 34 years. If true, this would help explain the lack of evidence against the unit root hypothesis in the post-war data, since there has been only one complete cycle. Clearly, this is not the ‘business cycle’ of contemporary macroeconomics, which is defined as having a period of not more than 8 years; see Hodrick and Prescott (1997) and Baxter and King (1999). This long wave phenomenon has been overlooked by economists, and explanations would shift the focus of macroeconomics from monetary policy shocks to longer run processes such as demographics or capital replacement.

3. The sign of the PP cycle is wrong. While the 1920’s do roar, the post-war pattern is perverse. The period 1946-1964 is entirely below trend, while 1965-80 is above trend, except for 1975-76. The ‘New Economy’ decade 1990’s is entirely below trend.

Finally, we consider the implications of having, hypothetically, taken the trend-stationary view on the basis of the 1870-1994 series that we used in MN. Refitting the trend function with restricted structural breaks in 1929 and 1940 to this series, we have extrapolated the trend through 2003 and compare it with actual data in Figure 3. If the insight that the U.S. economy was then more than 5% below trend had led us to predict faster growth

\(^{1}\) We use general-to-specific lag selection with a maximum lag of 8.
to return the economy to trend, we would have been disappointed. The implication going forward is for growth in excess of the trend rate of 3.2% as we close the gap, now about 7%, in coming years.

To sum up, accepting the trend-stationary view of long term output dynamics presents several challenges to macroeconomists. The low frequency cycle it implies and the perverse correlation of this cycle with conventional perception of economic conditions beg for explanations not to be found in contemporary economic theory.

References


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

Figure 1. ADF P-Values by Beginning Year
Figure 2. Papell and Prodan Cycle

Figure 3. Real GDP and Extrapolated Trend: 1994-2003