Hitotsubashi Journal of Economics 56 (2015), pp.117-134.

UNCERTAIN EFFECTS OF SHOCKS VS. UNCERTAIN UNIT ROOT: AN ALTERNATIVE VIEW OF U.S. REAL GDP*

Yu-Lieh Huang*

Department of Quantitative Finance, National Tsing-Hua University
Hsinchu 30013, Taiwan
ylihuang@mx.nthu.edu.tw

and

Chao-Hsi Huang

Department of Economics, National Tsing-Hua University
Hsinchu 30013, Taiwan
ghuang@mx.nthu.edu.tw

Received May 2014; Accepted October 2014

Abstract

Instead of exploring the uncertainty about the existence of a unit root in the long-span U.S. real GDP series as in previous studies, e.g., Rudebusch (1993), in this study we investigate the uncertainty about the state (permanence vs. transitoriness) of the output shock period by period by using the "innovation regime-switching" (IRS) model. In this model the effect of a shock may be permanent or transitory in different time periods. By applying the IRS model to the 1870-2008 annual U.S. real GDP data, we find that the output shocks in the periods of the 1893 depression, the 1907 financial panic, the two World Wars and the Great Depression are likely to have had a large but transitory effect, whereas the output shocks in the remaining periods are likely to have had a permanent effect. This result suggests that the long-span real GDP is neither a unit-root series nor a trend-stationary series.

Keywords: Innovation regime-switching, permanent innovation, trend-stationarity, uncertain unit root

JEL Classification Codes: C22, C51

* We thank Chung-Ming Kuan, Cheng Hsiao, Peter H. Calkins, Yi-Ting Chen and Shu-Ling Chen for valuable suggestions on early drafts of this paper. We are also indebted to the associate editor and one anonymous referee whose comments lead to a much improved version of the paper. All remaining errors are ours. Y.-L. Huang gratefully acknowledges research support of the National Science Council of the Republic of China (NSC-101-2410-H-007-008).

I. *Introduction*

Since the seminal work of Nelson and Plosser (1982), much effort has been devoted to exploring whether the U.S. real GDP series can be characterized as a difference-stationary (hereafter DS) or a trend-stationary (hereafter TS) process. In spite of numerous studies, the debate about the difference-stationarity versus trend-stationarity of the U.S. GDP remains open. For example, when exploring long-span U.S. real GDP data, Diebold and Senhadji (1996) found that conventional Dickey Fuller unit-root tests produce results favoring trend-stationarity, which casts doubt on the consensus opinion of the existence of unit root in post-war real GDP. Murray and Nelson (2000, 2002), on the other hand, have argued that the evidence against the unit root in the long-span real GDP series is mainly caused by the period of turmoil experienced from 1929 to 1946 due to the Great Depression and World War II. Once the heterogeneity in the data is taken into account, the long-span real GDP data may still contain a unit root. In contrast with Murray and Nelson (2002), Papell and Prodan (2004) provided evidence against unit roots in long-span real GDP by conducting Lamsdaine and Papell (1997) tests. Their result suggests that the real GDP can be viewed as a TS process with trend breaks occurring from 1929-1946.

The studies mentioned above bring relevant research full circle on the issue of the uncertainty about the existence of a unit root in U.S. real GDP. In the models of the pro difference-stationarity camp, e.g., Murray and Nelson (2002) and Kilian and Ohanian (2002), permanent shock is presented in each period while transitory shocks, usually with large short-term effects, occur only occasionally. In contrast, in the models of the pro trend-stationarity camp, e.g., Diebold and Senhadji (1996) and Papell and Prodan (2004), almost all shocks to output are transitory while permanent breaks (shocks) occur infrequently. Table 1 summarizes these empirical results in the literature.

TABLE 1. DYNAMIC BEHAVIOR OF U.S. REAL GDP

Papers	Suggestions	Notes
Nelson and Plosser (1982)	DS process	Conducting unit-root tests
Rudebusch (1993), Diebold and Senhadji (1996)	TS process	Using long-span data
Murray and Nelson (2000, 2002)	DS process	Considering heterogeneity in the data
Kilian and Ohanian (2002)	DS process	Using regimeswitching models
Ben-David and Papell (1995)	TS with breaks	Using Zivot and Andrews' (1992) tests
Papell and Prodan (2004,2007)	TS (with offsetting trend breaks)	Using Lumsdaine and Papell's (1997) tests

In this study, we argue that what is important about output fluctuations is the nature of shock (permanence vs. transitoriness) in each period rather than the presence of an exact unit root. We avoid the usual dichotomy between difference-stationarity and trend-stationarity and consider the uncertainty of the state (permanence vs. transitoriness) of output shock period by period. There are a few reasons for doing so. First, from an econometric point of view, although a DS or TS view of output process simplifies the model structure, such views might greatly limit the dynamic patterns that can fully characterize the real output series. Essentially, there is no a priori reason that one type of output shock (permanent or transitory) should always prevail or dominate in a very long time span.[1] In addition, output shocks in periods of tranquility are highly likely to be fundamentally different in nature from those periods of turmoil. Second, on the theoretical side, as Sims (1988), Durlauf (1989), Christiano and Eichenbaum (1990) and many others have demonstrated, the existence of an exact unit root in the real output per se may not help in identifying the true economic structure. Instead, as has been convincingly argued by Christiano and Eichenbaum (1990), it is the relative importance of permanent and transitory shocks and economic agents' perception of these shocks that determine the dynamic properties of economic models. Accordingly, a model which specially accounts for the uncertainty of permanent and transitory shocks should have potential to shed light on the relative importance of these shocks and, as a result, the underlying structure of economic models. Third, from the point of view of output forecasting, the knowledge of the nature of recent output shocks is much more important than the knowledge of whether the output process contains a unit root. For example, in an event where real output has declined mainly due to transitory factors, e.g., monetary shocks, the output is expected to bounce back to its long-term mean or time trend. In contrast, in an event where output has declined mainly due to permanent factors, no such rebound of output is expected. That is, the different states of recent output shocks should have very different implications as to how the future output may change. Accordingly, the knowledge of the state of recent output shocks provides important information about the future output path.

To account for the uncertain effect of each shock, this study considers a more flexible model: the "innovation regime-switching" (henceforth IRS) model recently proposed by Kuan et al. (2005), and demonstrates how such a modeling strategy can be applied to analyzing the persistence of output fluctuation period by period. The IRS model is an unobserved-component model which treats a time series process as consisting of a unit root with a drift component and a TS component; whether a particular component is activated depends on an unobservable state variable whose law of motion is governed by certain probability laws. Thus, the effects of shocks in an IRS process are not fixed at all times but may be permanent or transitory in different time periods. If the components are state independent, the model is reduced to a conventional DS or TS model.

The IRS modeling approach has several merits worth mentioning. First, the IRS model accommodates both trend-reverting and trend-disturbing behaviors, and hence bridges the gap between TS and unit-root nonstationary models. As a result, the model provides us with a more flexible framework to explore the persistent nature of U.S. GDP. Second, to accommodate the heterogeneity in the data, the IRS model allows for potential asymmetry in volatility across different regimes by permitting switching variances in the random shocks. Third and more importantly, based on the data, the IRS model estimates whether or not the permanent or transitory state of shock activated is governed by the probability law, and the probabilities of the respective states in each period. Due to this particular feature of the model, no a priori assumption regarding the importance of permanent vis-à-vis transitory shocks in the GDP process is required. In other words, the model just lets the data speak for themselves. As a result, the estimation results of the IRS model can provide evidence on the relative importance of permanent and transitory shocks during the sample period. Consequently, the

[1] See, for example, Newbold et al. (2001) for the argument.

model can serve as a benchmark to evaluate the plausibility of various GDP models with differing assumptions regarding the state of the shocks, and this can help to shed light on the debate regarding the nature of real GDP fluctuations.

Our empirical study uses U.S. real GDP from 1870 to 2008. The study suggests that the proposed IRS modeling approach is able to reveal important features of output data. In particular, the estimation results show strong evidence that permanent and transitory output innovations prevail in different sub-sample periods. This suggests that over the very long time period analyzed, U.S. real GDP is neither a DS series nor a TS series (including pure TS or trend-break model a la Papell and Prodan, 2004) as indicated in previous studies. Our results reveal that the output shocks in the periods of the 1893 depression, the 1907 financial panic, the World War I, the Great Depression and the World War II are more likely to have had a transitory effect. The results indicating that the output shocks in the Great Depression and the two World Wars had large transitory effects are compatible with the theorizing of Friedman and Schwartz (1963), Lucas and Rapping (1969) and Barro (1981). Our results also show that output shocks for the post World War II period are more likely to have had a permanent effect, and this is consistent with the consensus opinion of the existence of a unit root in the post-war GDP.

To evaluate the performance of models with differing assumptions regarding the state of output innovations, we follow Rudebusch (1993) and conduct a Monte Carlo study of the rejection rates of several unit-root tests— namely, the efficient augmented Dickey-Fuller test of Elliott et al. (1996) and the sequential breakpoint selection tests of Lamsdaine and Papell (1997). In our simulation, we generate data from the best-fitting IRS model and other best-fitting benchmark models, including a pure unit-root ARIMA model, the regime-switching model of Murray and Nelson (2002), a TS ARMA model, and the trend-break model proposed by Papell and Prodan (2004). Our simulations show that the unit-root tests, while producing evidence unfavorable to the existence of a unit root, are not able to discriminate between TS/trend-break models and the empirical IRS model. In addition, the sample estimates of the unit-root test statistics often have values close to the center section of the finite-sample distribution of the statistics generated from the empirical IRS model. This, together with our estimation results, indicates that the proposed IRS model may serve as the best model of U.S. real GDP among many alternatives. Finally, in our Monte Carlo experiment of the sequential breakpoint selection tests of Lamsdaine and Papell (1997), we find that the tests tend to spuriously identify permanent trend breaks in periods where the data have actually been generated by transitory innovations according to the empirical IRS model, especially during the period from 1929-1946. This raises strong doubt about Papell and Prodan's (2004) contention that permanent trend breaks exist in the U.S. real GDP during the period of 1929-1946.

The outline of this paper is as follows. In section II, we describe and explore the long-span U.S. real GDP data employed in study. In particular, we conduct two unit-root tests for the data: the efficient augmented Dickey-Fuller test from Elliott et al. (1996) and the sequential breakpoint selection tests proposed by Lamsdaine and Papell (1997). In section III, we apply the IRS model to the long-span real GDP data and discuss the estimation results. In section IV, we employ the Rudebusch's (1993) bootstrap procedure to explore the finite distribution properties of the unit-root test statistics of the IRS model as well as some benchmark DS and TS models. This procedure is also used to evaluate the plausibility of the real GDP data being generated from the processes. The concluding remarks are given in section V.

II. *Unit-Root Testing Results*

We now examine the state of shocks on annual U.S. real GDP from 1870 to 2008 which is extracted from Angus Maddison's homepage: www.ggdc.net/maddison/. This is the long-span real GDP data widely discussed and employed in recent studies, e.g., Murray and Nelson (2000, 2002) and Papell and Prodan (2004). We first conduct unit-root tests on the log of the GDP series, y_t. Instead of using the standard augmented Dickey-Fuller (ADF) test, we apply the efficient ADF-GLS $^\tau$ test of Elliott et al. (1996) which is based on the following auxiliary regression:

$$(1-B)y_t^\tau=\alpha_0+\alpha_1 t+\psi_0 y_{t-1}^\tau+\sum_{i=1}^{m}\psi_i(1-B)y_{t-i}^\tau+\varepsilon_t, \qquad (1)$$

where B is the lag operator and y_t^τ is given by $y_t^\tau=y_t-\tilde{\beta}'z_t$ with $z_t=(1,t)'$ and with $\tilde{\beta}$ the OLS estimate of β obtained from regressing

$$(\tilde{y}_1, \tilde{y}_2, ..., \tilde{y}_T)=(y_1, (1-aB)y_2, ..., (1-aB)y_T)$$

on

$$(\tilde{z}_1, \tilde{z}_2, ..., \tilde{z}_T)=(z_1, (1-aB)z_2, ..., (1-aB)z_T)$$

at $a=1-13.5/T$. We reject the null hypothesis of a unit root if ψ_0 in (1) is significantly large.

Because conventional unit-root tests tend to misinterpret a trend-break series as DS series, we also employ tests that take into account the possibility of structural breaks in the trend. Here, we follow Papell and Prodan (2004) and adopt the procedure set up by Lumsdaine and Papell (1997) to test the unit-root null against a TS alternative with two breaks:

$$(1-B)y_t=\alpha_0+\alpha_1 t+\delta_1 DU_{1,t}+\delta_2 DU_{2,t}+\psi_0 y_{t-1}+\sum_{i=1}^{m}\psi_i(1-B)y_{t-i}+\varepsilon_t, \qquad (2)$$

where $DU_{i,t}=\mathbf{1}_{\{t>T_{b_i}\}}$ for $i=1,2$, $\mathbf{1}_{\{\cdot\}}$ is the indicator function and T_{b_i} is the time at which the change in the trend function occurs. We reject the null hypothesis of a unit root in favor of TS with breaks if ψ_0 in (2) is significantly different from zero. The model (2) is estimated sequentially for each break date $T_{b_i}=m+i,...,T-i$ where $i=1,2$, $T_{b_1}\neq T_{b_2}$, and $T_{b_1}\neq T_{b_2}+1$. We select breaks for which the maximum evidence against the unit-root null.

To choose the number of augmented lags m, the Schwarz information criterion (SIC) and the modified information criterion of Ng and Perron (2001) are used for the DF-GLS$^\tau$ test. In addition, the "general-to-specific" recursive t statistic procedure of Ng and Perron (1995) is conducted to select the lag length for the Lumsdaine-Papell test. For all tests considered, the maximum value of m is set to 12. The asymptotic critical values for the ADF-GLS$^\tau$ test statistic are provided by Elliott et al. (1996). The finite-sample critical values for the Lumsdaine-Papell test statistic are calculated based on the Monte Carlo method with 5,000 replications. Details of these bootstrap critical values are omitted to save space but can be found in Lumsdaine and Papell (1997).

The testing results are summarized in Table 2. From the table it can be seen that, for all the tests considered, the null hypothesis of the existence of a unit root in the real GDP process is rejected. For example, the ADF-GLS$^{\tau}$ test rejects the unit-root null at the 1% level, while the Lumsdaine-Papell test also rejects the unit-root null in favor of broken trend-stationarity at the 1% level. These results are consistent with the findings of Diebold and Senhadji (1996) in that, with longer span data, one tends to obtain evidence in favor of trend stationarity in the real GDP. The latter results are also compatible with the conclusions of Ben-David and Papell (1995) and Papell and Prodan (2004) in that the real GDP series may exhibit significant trend breaks during the early part of the last century.[2] The overall test results here appear to suggest that, almost all shocks to real GDP are transitory while permanent breaks (shocks) occur very infrequently.

TABLE 2. UNIT-ROOT TESTS FOR U.S. REAL GDP FROM 1870 TO 2008

Test	ψ_0	Test Statistic	Lag (m)	Break 1	Break 2
ADF-GLS$^{\tau}$	−0.1498	−3.8265**	1	—	—
Lumsdaine-Papell	−0.3813	−7.2032**	2	1929	1939

Note: Critical values for ADF-GLS$^{\tau}$ are −3.544 (1%), −3.000 (5%) and −2.710 (10%); those for Lumsdaine-Papell are −6.69 (1%), −6.13 (5%) and −5.89 (10%). Test statistics with two asterisks are significant at 1% level.

III. *The IRS Model*

Although the testing results in the previous section suggest the rejection of the null hypothesis of a unit root, it should be noted that the rejection of a unit root does not necessarily imply that the GDP series must be stationary. The dynamic properties of real GDP may be more complex than those of a unit root model or a TS model. Moreover, Sims (1988), Durlauf (1989), Christiano and Eichenbaum (1990) and many others point out that what is important for the dynamics of output is not the presence of an exact unit root per se. Rather, it is the persistence of each shock and the relative importance of temporary and permanent shocks. It is thus rather premature to draw any conclusion about the state of shocks based only on these tests. These notions underpin the models that allow for the fractional difference parameter, e.g., Diebold and Rudebusch (1989), and stochastic unit roots, e.g., Granger and Swanson (1997). Such notions also motivate the "current depth of recession" model proposed by Beaudry and Koop (1993), which differentiates the persistent nature of the real GDP shocks in expansions and recessions.

1. The Proposed Model

Instead of adopting a simplified dichotomy between DS and TS specifications, we consider a more flexible model, a variant of the IRS model of Kuan et al. (2005), to examine the output dynamics. We assume that the log of U.S. real GDP consists of two components— namely, $y_t = y_{1,t} + y_{0,t}$:

[2] Note that, as shown in the table, the estimated break dates of the real GDP coincide with the outset of the Great Depression as well as that of the World War II.

$$\begin{aligned}(1-B)y_{1,t}&=\alpha_0+s_t v_t,\\ \Psi(B)y_{0,t}&=\Phi(B)(1-s_t)v_t,\end{aligned} \tag{3}$$

where $\Psi(B)=1-\phi_1 B-\cdots-\phi_m B^m$ and $\Phi(B)=1-\varphi_1 B-\cdots-\varphi_n B^n$ are finite-order polynomials of the lag operator such that they have no common factors and their roots are all outside the unit circle. $s_t=\{0,1\}$ denotes an unobserved first-order Markov chain with the transition matrix

$$P=\begin{bmatrix}P(s_t=0|s_{t-1}=0) & P(s_t=1|s_{t-1}=0)\\ P(s_t=0|s_{t-1}=1) & P(s_t=1|s_{t-1}=1)\end{bmatrix}=\begin{bmatrix}p_{00} & p_{01}\\ p_{10} & p_{11}\end{bmatrix},$$

and $\upsilon_t=\alpha_1+\varepsilon_t$ is a non-zero mean innovation. The term $\{\varepsilon_t\}$ is a white noise with mean zero and variance depending on s_t, i.e., $\sigma^2_{s_t}$. This model is referred to hereafter as an IRS(1; m, n) model, signifying that one component ($y_{1,t}$) has a random walk structure and the other ($y_{0,t}$) has a stationary ARMA(m, n) structure. Compared with the IRS model originally considered by Kuan et al. (2005), the model (3) accommodates potential asymmetry in volatility across different regimes by permitting switching variances in the random shocks. Such a specification is capable of describing the heterogeneity in long-span U.S. real GDP.

A novel feature in model (3) is that only one component is activated in a given time period, depending on the realization of s_t. When $s_t=1$, the first component $y_{1,t}$ is excited by the random shock, while $y_{0,t}$ keeps evolving according to ARMA dynamics without the new shock. As long as $s_t=1$, the corresponding random shock has a permanent effect on future y_{t+j} ($j>0$) and generates unit-root type dynamics. When $s_t=0$, the random shock activates $y_{0,t}$ while leaving $y_{1,t}$ intact. The random shock thus has a transitory effect on future y_{t+j} and induces stationary ARMA dynamics. This model specification allows the effect of a random shock to alternate from time to time and thus is able to capture both nonstationary and stationary behaviors. It is worth noting that in a special case where $s_t=1$ ($s_t=0$) with probability one for all t, the model (3) is simply reduced to a conventional DS random walk model or a TS ARMA model, as discussed in the previous section.

The proposed model is also able to capture potential trend-breaks in the data. To see this, note that the model (3) can be expressed as

$$y_t=\alpha_0 t+\sum_{i=1}^{t} s_i v_i+\psi(B)^{-1}(1-s_t)v_t, \tag{4}$$

with $y_0=0$ and $v_i=0$ for $i\leq 0$, where the last component of (4) is a weakly stationary process generated by transitory innovations and gives rise to short-run fluctuations. If $s_t=1$ at $t=T_{b_1}+1$ and $s_t=0$ otherwise, then (4) becomes

$$y_t=\alpha_0 t+v_t DU_{1,t}+\omega_t,$$

where ω_t is a weakly stationary process. As such, y_t is a TS process with one endogenous break, in the sense that the break is due to the presence of permanent shocks v_t. In this case, the expected magnitude of the trend break is $E(v_t)=\alpha_1$. Similarly, the proposed model is able to approximate a TS process with two breaks if $s_t=1$ at $t=T_{b_1}+1$ and at $t=T_{b_2}+1$, and $s_t=0$ otherwise.

Model (3) can be written as a special case of a general dynamic model with state-dependent coefficients; see Kuan et al. (2005) for more details. Once the model is written in the state-space form with switching coefficients, the estimation algorithm (also the algorithms for calculating the filtering and smoothing probabilities) developed in Kim (1994) can be applied. In this study, we follow Kuan et al. (2005) by writing (3) in a Markov-switching state-space form, as proposed by Kim (1994), and compute the approximate quasi-maximum likelihood estimates (QMLE):

$$\theta=(\psi_1,\ldots,\psi_m,\varphi_1,\ldots,\varphi_n,\alpha_0,\sigma_0^2,\sigma_1^2,p_{00},p_{11})'.$$

By applying Kim's (1994) estimation algorithm, we obtain the filtering probabilities $P(s_t=0|Y^t;\theta)$, the smoothing probabilities $P(s_t=0|Y^T;\theta)$ and quasi-log-likelihood function as byproducts, where $Y^t=\{y_1,\ldots,y_t\}$ is the collection of all the observed variables up to time t. We shall use these probabilities to examine the effect of output shock in each period.

2. Estimation Results

To assess the empirical relevance of the proposed IRS model, we estimate (3) based on the annual data of U.S. real GDP from 1870 through 2008, with a total of 139 observations. We estimate an array of IRS(1; *m, n*) models for the U.S. real GDP with m and n no greater than 4. The parameter θ is estimated using the algorithm described in Kim (1994) and Kuan et al. (2005). This algorithm is initialized by a broad range of random initial values. The covariance matrix of θ is $-H(\hat{\theta})^{-1}$, the Hessian matrix of the log-likelihood function evaluated at the QMLE $\hat{\theta}$. Among all the models considered, the IRS(1; 2, 0) model is selected based on the SIC. The estimation results are summarized in Table 3. As the table shows, all parameter estimates (except α_0) are statistically significant at the 5% level.

TABLE 3. QUASI-MAXIMUM LIKELIHOOD ESTIMATES OF THE PROPOSED MODEL

Parameter	Estimate	Standard Error	t-statistic
α_0	−0.0013	0.0204	−0.063
α_1	0.0276	0.0133	2.075*
ψ_1	0.8222	0.0924	8.984*
ψ_2	−0.3348	0.0593	−5.645*
σ_0	0.1516	0.0523	2.898*
σ_1	0.0319	0.0091	3.505*
p_{00}	0.6300	0.1642	
p_{11}	0.9473	0.0231	
Log-Likelihood = −218.910		SIC = −398.519	
$Q(12)=15.123$		$Q(24)=31.782$	
ARCH(2) = 1.788		ARCH(4) = 4.290	

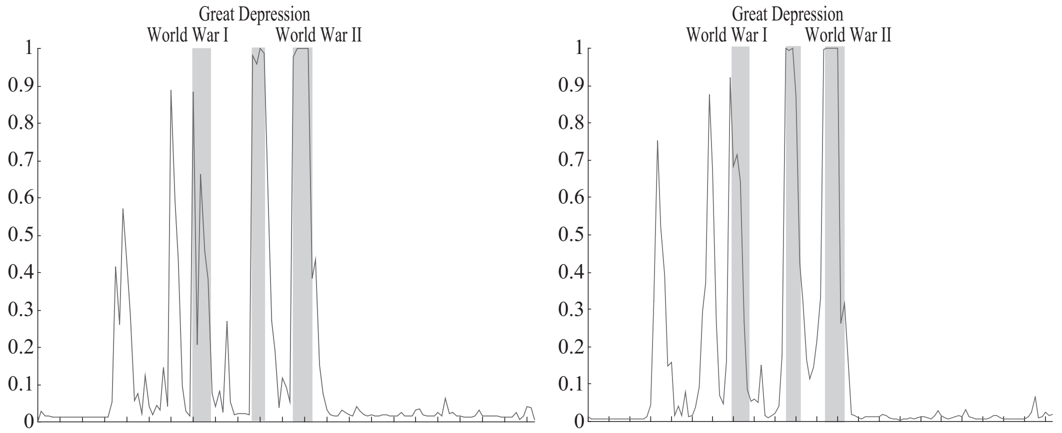
Note: t-statistics with anasterisk are significant at 5% level. The term $Q(\cdot)$ is the Q statistic in Ljung-Box (1978) and ARCH($\cdot$) denotes the LM statistic in Engel (1982).

To confirm the adequacy of the model, we conduct some diagnostic checks on the estimated model, including the Ljung-Box (1978) Q test and the LM test of Engle (1982) on the ARCH effect. As shown in Table 3, the resulting statistics for the residuals are $Q(12)=15.123$, $Q(24)=31.782$, ARCH(2)= 1.788 and ARCH(4) = 4.290. These statistics are all insignificant even at a 10 percent level under the $\chi^2(12)$, $\chi^2(24)$, $\chi^2(2)$ and $\chi^2(4)$ distributions, respectively. Hence, it appears that there is no serial correlation and conditional heteroskedasticity in the residuals. Following Engel and Hamilton (1990), we also test whether the state variables are independent over time, i.e., $p_{00}+p_{11}=1$. The resulting Wald statistic is 13.270, rejecting the null at a 1 percent level under the $\chi^2(1)$ distribution. This result is consistent with the Markovian specification.

We now turn to some interesting results from our IRS model estimation. In Figure 1 we plot the estimated filtering and smoothing probabilities of $s_t=0$ (i.e., transitory shock), where the shaded areas denote the periods of World War I (1914-1918), the Great Depression (1929-1933) and World War II (1941-1946), respectively. We find that during the sample period there are 19 years (about 14 percent of the sample) where the estimated smoothing probability $P(s_t=0|Y^T;\theta)$ is greater than 0.5. This result reveals that permanent innovations are more likely to prevail in about 86 percent of the sample period while transitory innovations dominate in the remaining period. The finding that both permanent and transitory shocks occur frequently is quite different from the assertions of the traditional TS model, the trend-break model and the unit-root model as discussed in Diebold and Senhadju (1996), Papell and Prodan (2004) and Murray and Nelson (2000, 2002). In contrast, our findings accord well with Newbold et al. (2001), in that neither simple TS nor DS specifications are found to adequately characterize U.S. long-span real GDP data. In addition, our estimation results show that, with $\hat{\sigma}_0$ about 4 times of $\hat{\sigma}_1$ (as Table 3 shows), the U.S. real GDP tends to be much more volatile in the periods where transitory shocks are present.

Figure 1 also reveals distinctly different dynamic patterns of real GDP for the pre- and post-1947 periods. For the post-1947 period (with 60 years), the smoothing probabilities of transitory shocks are all less than 0.5, indicating that permanent shocks are the dominant driving force behind the GDP fluctuations and that unit-root nonstationarity is the prevailing dynamic pattern. The results are consistent with the consensus findings of Campbell and

FIGURE 1. ESTIMATED FILTERING (LEFT) AND SMOOTHING (RIGHT) PROBABILITIES OF $s_t=0$ FOR U.S. ANNUAL REAL GDP FROM 1870 TO 2008



Mankiw (1987), Murray and Nelson (2000) and Newbold et al. (2001), among many others. In contrast, Figure 1 also shows that transitory shocks occur frequently during the pre-1947 period. For example, the estimated filtering and smoothing probabilities of $s_t=0$ are greater than 0.5 for the periods of the 1893 financial panic, the 1907 banking crisis, the Great Depression and the two World Wars. This, together with the estimation results of σ_0 and σ_1, suggests that the aforementioned events had large but transitory effects on output. This result is similar to the finding of Murray and Nelson (2000, 2002) which suggests that the Great Depression and World War II are two major episodes in which output shocks had large effects. A major difference is that, in the empirical IRS(1; 2, 0) model, output innovations in the periods of the 1893 depression, the 1907 financial panic, the Great Depression and the two World Wars are more likely to have had a transitory effect. The model of Murray and Nelson (2002), on the other hand, postulates a priori that the permanent shocks are presented in each of the sample periods, even in the period of the pre-1947 turmoil. As shown in the following simulation experiments, such a difference may influence the rejection frequencies of unit-root tests dramatically.

IV. *Simulation Results*

We now proceed to evaluate the likelihood of various benchmark models in generating the U.S. output data by employing Rudebusch's (1993) bootstrap procedure. We first consider the empirical IRS(1; 2, 0) model as well as models representing the two opposing approaches of aggregate output modeling: TS vs. DS models. In the TS camp, we explore the pure TS ARMA(m, n) model and the TS model with two breaks, as in Papell and Prodan (2004). In the DS camp, we study the pure ARIMA(m, 1, n) model and the model of Murray and Nelson (2002):

$$\begin{aligned} y_t &= \tilde{y}_{1,t} + \tilde{s}_t \tilde{y}_{0,t}, \\ \tilde{y}_{1,t} &= \alpha_0 + \tilde{y}_{1,t-1} + v_t, \\ \tilde{y}_{0,t} &= \phi_1 \tilde{y}_{0,t-1} + \phi_2 \tilde{y}_{0,t-2} + u_t, \end{aligned} \tag{5}$$

where $\tilde{s}_t=\{0,1\}$ denotes an unobserved state variable whose law of motion is governed by a first-order Markov chain with the transition matrix P, and the error terms v_t and u_t are Gaussian white noises with $cov(v_t,u_t)=0$ for all t. We estimate an array of TS ARMA(m, n) and ARIMA (m, 1, n) models with m and n no greater than 8 and choose an appropriate specification based on the SIC.

Before presenting the empirical results of these benchmark models, it would be of interest to discuss the differences between the IRS model and the regime switching model of Murray and Nelson (2002). By comparing the specifications in (5) and (3), it is apparent that the Murray and Nelson's (2002) model actually shares a similar setup with the IRS model used in this study. In particular, both models assume that the real output process is driven by two types of shock: permanent and transitory shocks. However, importantly, the two models differ in their treatment of permanent shock in the real output process, y_t. In the IRS model, the state of the output shock at each sample point is either permanent or transitory and there is no assurance that the permanent shock should always occur. In contrast, the Murray and Nelson's

(2002) model imposes, a priori, that the permanent shock is certain at each sample point; that is, the real output follows a unit-root process. The setup of Murray and Nelson (2002) may not always be appropriate, however. For example, Beaudry and Koop (1993) and Bradley and Jansen (1997) show that positive shocks to U.S. GDP are more persistent than negative shocks; see also Kuan et al. (2005). Thus, by allowing the innovation to excite only one component, the proposed model allows for distinct (unit-root and stationary) dynamics in different periods.

We now turn to discuss the empirical results of these models. The estimation results of these benchmark models are summarized in Table 4. As the table shows, the best fitting benchmark models are: the TS ARMA (2, 0) model, the trend-break model in (2) with $T_{b_1}=1929$, $T_{b_2}=1939$ and $m=2$, the ARIMA(2, 1, 1) model and the model (5) with transition probabilities $p_{00}=0.9887$ and $p_{11}=0.9740$. Note that the parameter estimates of model (5) are similar to those reached in Papell and Prodan (2004). It is also of interest to know that, based on the estimated temporary component $\tilde{y}_{0,t}$ in model (5), the output shocks during the 1929-1945 period have much greater volatility than those during the rest of the sample period, confirming Murray and Nelson's (2000, 2002) argument for the heterogeneity in the output data.[3] In the subsequent simulations a la Rudebusch (1993), the data are generated from these best-fitting benchmark models and the empirical IRS(1; 2, 0) model.

TABLE 4. PARAMETER ESTIMATES OF BENCHMARK MODELS FOR U.S. REAL GDP

Parameter	Murray-Nelson		ARIMA(2, 1, 1)		TS ARMA(2, 0)		Trend-Break	
	Estimate	Std.	Estimate	Std.	Estimate	Std.	Estimate	Std.
α_0	0.0335*	0.001	0.0322*	0.001	2.1798*	0.497	4.4586*	0.614
α_1	–	–	–	–	0.0060*	0.001	0.0129*	0.001
δ_1	–	–	–	–	–	–	−0.1205*	0.022
δ_2	–	–	–	–	–	–	0.0839*	0.019
ψ_0	–	–	–	–	–	–	−0.3813*	0.062
ψ_1	1.2227*	0.019	1.1557*	0.082	1.1532*	0.082	0.3167*	0.076
ψ_2	−0.3963*	0.020	−0.3365*	0.082	−0.3391*	0.083	0.1608	0.081
ϕ_1	–	–	−0.9974*	0.013	–	–	–	–
σ_ε	–	–	0.0496*	0.010	0.0494*	0.031	0.0441*	0.023
σ_v	0.0199*	0.001	–	–	–	–	–	–
σ_u	0.0624*	0.001	–	–	–	–	–	–
p_{00}	0.9887	1.217	–	–	–	–	–	–
p_{11}	0.9740	0.772	–	–	–	–	–	–

Note: The term Std. denotes the standard error. Estimates with an asterisk are significant at 5% level.

In our bootstrap experiments, we simulate data according to the estimation results reported in Tables 3 and 4. Moreover, we generate the state variables of the IRS model by setting $s_t=0$ for the periods where the smoothing probability for the state of transitory shock to occur is greater than 0.5 and $s_t=1$ for the rest of the sample period. For the Murray and Nelson's model of equation (5), we set $s_t=1$ for the 1929-1945 period and $s_t=0$ for the rest. Based on 100,000 replications of bootstrap simulations with a sample size of 139, we obtain the finite-sample distributions of the ADF-GLS$^\tau$ and the Lumsdaine-Papell test statistics for the various best-fitting models. To make exposition easier, we label the finite-sample distributions of the

[3] The estimated temporary component of model (5) is available from the authors upon request.

IRS(1; 2, 0) model, the TS ARMA(2, 0) model, the TS model with two breaks, the ARIMA(2, 1, 1) model, and the model of Murray and Nelson (2002) as $f_{IRS}(\hat{\tau})$, $f_{TS}(\hat{\tau})$, $f_{TB}(\hat{\tau})$, $f_{DS}(\hat{\tau})$ and $f_{MN}(\hat{\tau})$, respectively, where $\hat{\tau}=\{\hat{\tau}_1,\hat{\tau}_2\}$ is the ADF-GLS$^\tau$ test statistic in (1) or the Lumsdaine-Papell test statistic in (2).

2. Finite-Sample Distributions of the ADF-GLS$^\tau$ Statistic

In Figure 2, the finite-sample distributions $f_{TS}(\hat{\tau}_1)$, $f_{DS}(\hat{\tau}_1)$, and $f_{MN}(\hat{\tau}_1)$ of the ADF-GLS$^\tau$ statistic are plotted, where $\hat{\tau}_{1,\text{sample}}=-3.8265$ denotes the actual sample value of ADF-GLS$^\tau$ test statistic obtained in Section 2. In this figure the shaded area under $f_{DS}(\hat{\tau}_1)$ and to the left of $\hat{\tau}_{1,\text{sample}}$ represents the probability of obtaining a value of the ADF-GLS$^\tau$ test equal to or smaller than -3.8265, conditional on the best fitting ARIMA(2, 1, 1) model. This p-value, denoted as

$$P(\hat{\tau}_1 \leq \hat{\tau}_{1,\text{sample}} | f_{DS}(\hat{\tau}_1)),$$

is equal to 4.29%. Therefore, given the sample test statistic, the ARIMA model is rejected at the five percent level. The other hatched area in the figure is the one under $f_{TS}(\hat{\tau}_1)$ and to the right of $\hat{\tau}_{1,\text{sample}}$. The area represents the probability of obtaining a value of the t test equal to or greater than -3.8265, conditional on the best-fitting TS ARMA(2, 0) model. This p-value, denoted as

$$P(\hat{\tau}_1 \geq \hat{\tau}_{1,\text{sample}} | f_{TS}(\hat{\tau}_1)),$$

is equal to 36.95%. Thus, it is highly unlikely to obtain $\hat{\tau}_{1,\text{sample}}$ when the true data generating process is ARIMA(2, 1, 1) but it is very likely when the data generating process is TS ARMA (2, 0). The result here confirms previous findings that when U.S. real output data of longer span are employed, the test results often point toward the rejection of a unit-root while in favor of a deterministic trend, e.g., Ben-David and Papell (1995), Cheung and Chinn (1997), and Diebold and Senhadji (1996). Nevertheless, the p-value corresponding to the area under $f_{MN}(\hat{\tau}_1)$

FIGURE 2. FINITE-SAMPLE DISTRIBUTIONS OF THE ADF-GLS$^\tau$ STATISTIC FOR THE TS ARMA, THE MURRAY-NELSON AND THE ARIMA MODELS

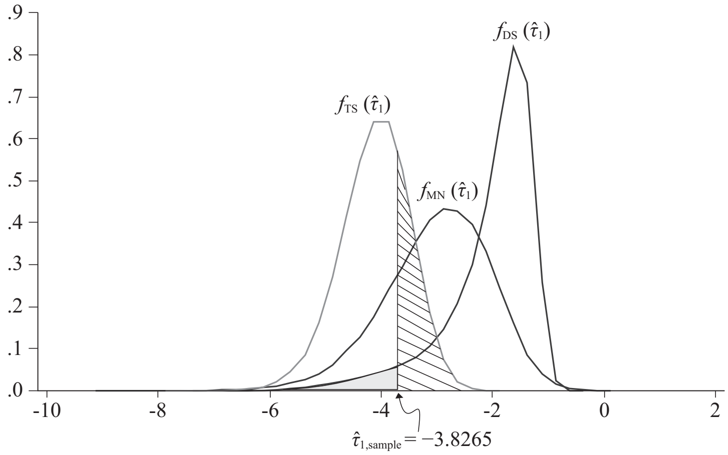
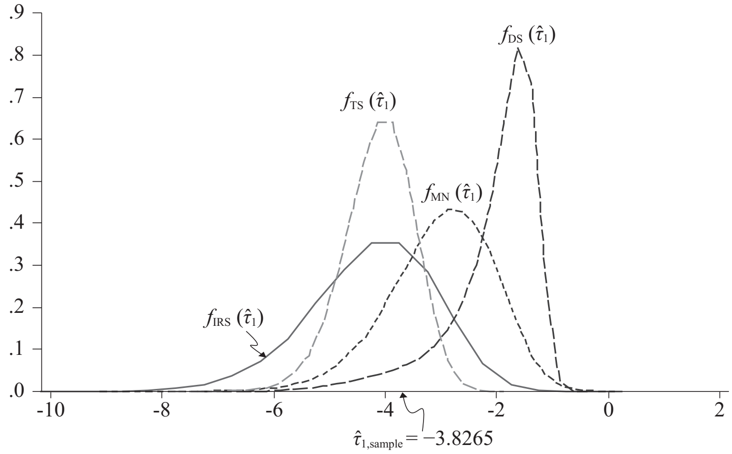


FIGURE 3. FINITE-SAMPLE DISTRIBUTIONS OF THE ADF-GLS$^{\tau}$ STATISTIC FOR THE IRS MODEL (SOLID LINE) AND OTHER MODELS (DASHED LINES)



and to the left of $\hat{\tau}_{1,\text{sample}}$, i.e.,

$$P(\hat{\tau}_1 \leq \hat{\tau}_{1,\text{sample}} | f_{MN}(\hat{\tau}_1)),$$

is 18.08%, indicating that the unit-root hypothesis cannot be rejected at the five percent level. This result confirms the finding of Murray and Nelson (2002) which suggests that the unit-root hypothesis is incorrectly rejected too often when the underlying model is a unit root process augmented with a transitory component to account for the heterogeneity in the historical GDP series.

In order to evaluate the empirical IRS(1; 2, 0) model, its finite-sample distribution of the ADF-GLS$^{\tau}$ statistic, $f_{IRS}(\hat{\tau}_1)$, is plotted in Figure 3 along with those of the ARIMA(2, 1, 1), ARMA(2, 0) and Murray and Nelson's (2002) model. In this figure, the dashed lines denote the corresponding distributions $f_{TS}(\hat{\tau}_1)$, $f_{DS}(\hat{\tau}_1)$, and $f_{MN}(\hat{\tau}_1)$ while the solid line represents the distribution for the proposed IRS model. As the figure shows, the $\hat{\tau}_{1,\text{sample}}$ is located close to the center part of $f_{IRS}(\hat{\tau}_1)$. More precisely, the p-value,

$$P(\hat{\tau}_1 \leq \hat{\tau}_{1,\text{sample}} | f_{IRS}(\hat{\tau}_1)),$$

is 61.96%. This result indicates that, among the models considered in Figure 3, it is most likely to obtain a sample estimate of $\hat{\tau}_{1,\text{sample}}$ of -3.8265 when the true data generating process of U.S. GDP is the assumed IRS(1; 2, 0) model. Moreover, it is of interest to compare the simulation results of Murray and Nelson's model and the IRS model. Murray and Nelson's (2002) model is essentially a DS model, and hence the permanent shocks prevail in each period of the sample. In contrast, the proposed IRS model is neither a DS nor a TS process but a mixture of the two, which hence allows its random innovations to have permanent and transitory effects in different periods. Although both of the models can capture the heterogeneity in the data, the p-value of Murray and Nelson's model is much lower than that of the IRS model (18.08% vs. 61.96%). The huge difference in the p-values may be attributable to the two models' different treatment of permanent shocks: the permanent shock appears in each point of the sample period with certainty in the former (i.e., the Murray and Nelson's model) but not in the latter (i.e., IRS model). Consequently, our result suggests that the rejection rates of the unit-root tests are quite sensitive to treatment of the big events (such as financial crisis, the Great Depression and World Wars) as permanent or transitory innovations.

3. Finite-Sample Distributions of the Lumsdaine-Papell Statistic

In Figure 4 we plot the finite-sample distributions of the Lumsdaine-Papell test statistic: $f_{\mathrm{IRS}}(\hat{\tau}_2)$, $f_{\mathrm{TB}}(\hat{\tau}_2)$ and $f_{\mathrm{MN}}(\hat{\tau}_2)$, where $\hat{\tau}_{2,\mathrm{sample}}=-7.2032$ denotes the actual sample value obtained in Section 2.[4] As the figure shows, the shaded area under $f_{\mathrm{MN}}(\hat{\tau}_2)$ and to the left of $\hat{\tau}_{2,\mathrm{sample}}$ is quite small, corresponding to a p-value of

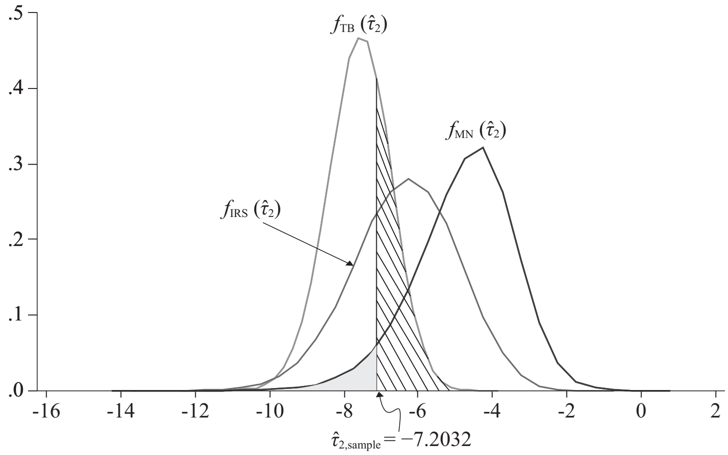
$$\mathrm{P}(\hat{\tau}_2 \leq \hat{\tau}_{2,\mathrm{sample}} | f_{\mathrm{MN}}(\hat{\tau}_2))=4.53\%.$$

In contrast, the hatched area under $f_{\mathrm{TB}}(\hat{\tau}_2)$ and to the right of $\hat{\tau}_{2,\mathrm{sample}}$ is rather sizable, corresponding to a p-value of

$$\mathrm{P}(\hat{\tau}_2 \geq \hat{\tau}_{2,\mathrm{sample}} | f_{\mathrm{TB}}(\hat{\tau}_2))=32.97\%.$$

These findings suggest that it is unlikely to reach $\hat{\tau}_{2,\mathrm{sample}}$ when the true data generating process is assumed to be Murray and Nelson's model while it is very likely to reach $\hat{\tau}_{2,\mathrm{sample}}$ when the data generating process is assumed to be the TS model with two trend breaks. The finding here confirms that result of Papell and Prodan (2004, 2007), i.e., the Lumsdaine-Papell test often points toward the rejection of a unit root when in favor of a trend break process and the long-span U.S. real output data are employed. However, it is worth noticing that in Figure 4, the

FIGURE 4. FINITE-SAMPLE DISTRIBUTIONS OF THE LUMSDAINE-PAPELL STATISTIC FOR THE TREND-BREAK, THE MURRAY-NELSON AND THE IRS MODELS



[4] Since the finite-sample distributions of the Lumsdaine-Papell test statistic for the ARIMA(2, 1, 1) and ARMA(2, 0) are more extreme than Murray and Nelson's (2002) model and the TS model with two breaks, respectively, to keep the figure simple we choose not to plot them.

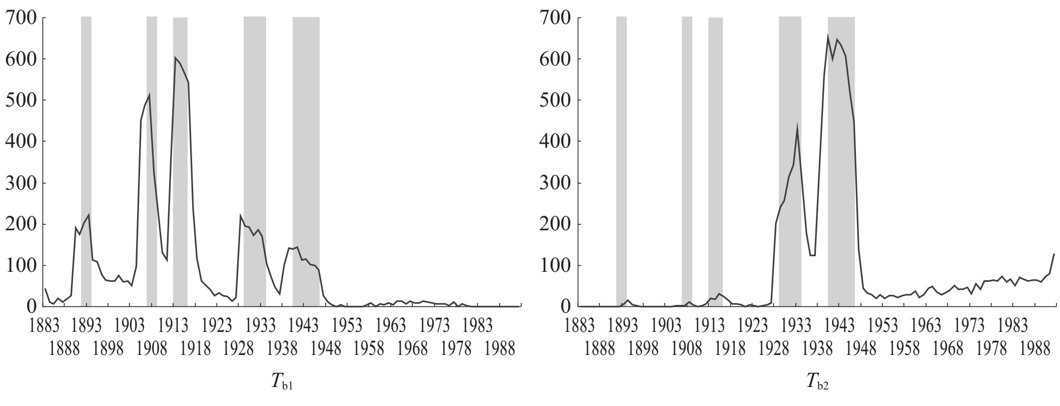
area under $f_{\mathrm{IRS}}(\hat{\tau}_2)$ and to the left of $\hat{\tau}_{2,\mathrm{sample}}$ is also sizable, which corresponds to a p—value of

$$\mathrm{P}(\hat{\tau}_2 \geq \hat{\tau}_{2,\mathrm{sample}} | f_{\mathrm{IRS}}(\hat{\tau}_2)) = 27.32\%.$$

This indicates that the Lumsdaine-Papell test is not able to discriminate between the trend-break models and the IRS model, even though it provides evidence against Murray and Nelson's DS model.[5] This result suggests that the Lumsdaine-Papell test results should be interpreted with caution. When the Lumsdaine-Papell test indicates the rejection of the unit-root null of Murray and Nelson's DS processes, as found in Papell and Prodan (2004), the result should not automatically be taken as evidence supporting the TS model with offsetting trend breaks. This is because a process very different from the TS model with breaks, such as the one described by this study's IRS model, might just as well be the true underlying output process.

We also find in our simulations that the Lumsdaine-Papell test tends to identify break points incorrectly when the data are actually generated from the assumed IRS model. In our IRS simulations, the output shocks during the periods of the Great Depression and the two World Wars are assumed to be transitory. However, when the Lumsdaine-Papell test is conducted, permanent trend breaks are spuriously identified during these periods. Figure 5 shows the results of permanent break points identified by the Lumsdaine-Papell test when the data are generated from the assumed IRS model with the shaded areas denoting the periods of transitory shocks. In this figure, it can be seen that more than 60% of breakpoints (64.45% for T_{b_1} and 61.36% for T_{b_2}) are located in the shaded areas, indicating that the problem of spurious identification of the trend breaks points is rather serious.

FIGURE 5. FREQUENCIES OF THE ESTIMATED BREAK POINTS IDENTIFIED BY THE LUMSDAINE-PAPELL TEST: $DU_{1,t}$ (Left) and $DU_{2,t}$ (Right)



To check for the robustness of our results, we have redone our simulations using the unit-root test of Zivot and Andrews (1992) in the case of a single break. We also conduct Rudebusch's (1993) bootstrap procedure using the unit-root test of Papell and Prodan (2004)

[5] As discussed before, such a difference could be attributed to the uncertainty of the state of output shock in the IRS model.

with restricted structural change (i.e., two offsetting structural changes). We find no qualitative differences between these results and those discussed above. In other words, these two tests are still unable to distinguish between the trend-break models and the empirical IRS model. These tests also tend to identify break points incorrectly when the data are generated from the IRS model.

In sum, our simulation results of unit-root test statistics raise doubt about the appropriateness of modeling U.S. real GDP as a pure DS process or a modified DS process, as in Murray and Nelson (2002). However, the evidence against the DS modeling of the U.S. GDP cannot be automatically assumed to support the TS modeling of the series. This is because, based on the unit-root tests examined, little can be said about the relative likelihood of the specific TS/trend-breaks and IRS models of the U.S. GDP considered above. Moreover, our simulations indicate that the Lumsdaine-Papell tests tend to spuriously identify transitory shocks as permanent trend breaks. This casts doubt on Papell and Prodan's (2004) contention that the U.S. real GDP follows a TS process with permanent breaks occurring during the 1929-1946 period.

V. *Conclusion*

Instead of exploring the uncertainty about the existence of a unit root in the long-span U.S. real GDP series as has been done in previous studies, in this study, we investigate the uncertainty about the state of output shock period by period by using the IRS model in which the effect of a shock may be permanent or transitory in different time periods. By applying the IRS model to 1876-2008 U.S. real GDP data, we find that the output shocks in the periods of the 1893 depression, the 1907 financial panic, the two World Wars and the Great Depression are likely to have had a large but transitory effect, whereas the output shocks in the remaining periods are likely to have had a permanent effect. More specifically, for the whole sample period there are nineteen years (around fourteen percent of the sample) where the real GDP shocks are identified as transitory. Our results reveal the importance of both permanent and transitory shocks as the source of U.S. GDP fluctuations and suggest that the long-span real GDP is neither a unit-root series nor a TS series. This finding is in sharp contrast with the assertions of traditional TS, broken TS and DS models.

Our simulations also show that the unit-root tests, while producing evidence unfavorable to the existence of a unit root, are not able to discriminate between TS/trend-break models and the assumed IRS model. In addition, the sample estimates of the unit-root test statistics often have values close to the center section of the finite-sample distribution of the statistics generated from the assumed IRS model. This, together with our estimation results, indicates that the importance to output fluctuations may not lie on the presence of an exact unit root, but on the uncertainty about the state of output shocks and the identification of the nature of the state. Finally, in our Monte Carlo experiment investigating the sequential breakpoint selection tests of Lamsdaine and Papell (1997), we find that the tests tend to spuriously identify permanent trend breaks in periods where the data have actually been generated by transitory shocks according to the assumed IRS model, especially in the period from 1929-1946. This raises strong doubt about Papell and Prodan's (2004) argument that the U.S. real GDP follows a TS process with permanent breaks occurring during the 1929-1946 period.

References

Barro, R.J. (1981), "Output Effects of Government Purchases," Journal of Political Economy 89, pp.1086–1121.

Beaudry, P. and G. Koop (1993), "Do Recessions Permanently Change Output? " *Journal of Monetary Economics* 31, pp.149–163.

Ben-David, D. and D.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, pp.453–475.

Cheung, Y.-W. and M.D. Chinn (1997), "Further Investigation of the Uncertain Unit Root in GNP," *Journal of Business & Economic Statistics* 15, pp.68–73.

Christiano, L.J. and M. Eichenbaum (1989), "Unit Roots in Real GNP: Do We Know, and Do We Care? " *Carnegie-Rochester Conference Series on Public Policy* 32, pp.7–62.

Diebold, F.X. and G.D. Rudebusch (1989), "Long Memory and Persistence in Aggregate Output," *Journal of Monetary Economics* 24, pp.189–209.

Diebold, F.X. and A.S. Senhadji (1996), "The Uncertain Unit Root in Real GNP: Comment," *The American Economic Review* 86, pp.1291–1298.

Durlauf, S.N. (1989), "Output Persistence, Economic-Structure, and the Choice of Stabilization Policy," *Brookings Papers on Economic Activity* 1, pp.69–135.

Elliott, G., T.J. Rothenberg and J.H. Stock (1996), "Efficient Tests for an Autoregressive Unit Root," *Econometrica* 64, pp.813–836.

Engel, C. and J.D. Hamilton (1990), " Long Swings in the Dollar: Are They in the Data and Do Markets Know It?" *American Economic Review* 80, pp.689–713.

Engle, R.F. (1982), "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica* 50, pp.987–1007.

Friedman, M. and A.J. Schwartz (1963), *A Monetary History of the United States, 1867-1960*, Princeton, NJ, Princeton University Press.

Granger, C.W.J. and N.R. Swanson (1997), "An Introduction to Stochastic Unit Root Processes," *Journal of Econometrics* 80, pp.35–62.

Kilian, L. and L.E. Ohanian (2002), "Unit Roots, Trend Breaks, and Transitory Dynamics: A Macroeconomic Perspective," *Macroeconomic Dynamics* 6, pp.614–632.

Kim, C.-J. (1994), "Dynamic Linear Models with Markov-Switching," *Journal of Econometrics* 60, pp.1–22.

Kuan, C.-M., Y-L. Huang and S.R. Tsay (2005), "An Unobserved-Component Model with Switching Permanent and Transitory Innovations," *Journal of Business & Economic Statistics* 23, pp.443–454.

Ljung, G.M. and G.E.P. Box (1978), "On a Measure of Lack of Fit in Time Series Models," *Biometrika* 65, pp.297–303.

Lucas, R.E. and L.A. Rapping (1969), "Real Wages, Employment, and Inflation," *Journal of Political Economy* 77, pp.721–754.

Lumsdaine, R.L. and D.H. Papell (1997), "Multiple Trend Breaks and the Unit-Root Hypothesis," *The Review of Economics and Statistics* 79, pp.212–218.

Murray, C.J. and C.R. Nelson (2000), "The Uncertain Trend in U.S. GDP," *Journal of Monetary Economics* 46, pp.79–95.

Murray, C.J. and C.R. Nelson (2002), “The Great Depression and Output Persistence,” *Journal of Money Credit and Banking* 34, pp.1090–1098.

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

Newbold, P., S. Leybourne and M.E. Wohar (2001), “Trend-Stationarity, Difference-Stationarity, or Neither: Further Diagnostic Tests with an Application to U.S. Real GNP, 1875-1993,” *Journal of Economics and Business* 53, pp.85–102.

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

Papell, D.H. and R. Prodan (2004), “The Uncertain Unit Root in US Real GDP: Evidence with Restricted and Unrestricted Structural Change,” *Journal of Money Credit and Banking* 36, pp.423–427.

Papell, D.H. and R. Prodan (2007), “Restricted Structural Change and the Unit Root Hypothesis,” *Economic Inquiry* 45, pp.834–853.

Rudebusch, G.D. (1993), “The Uncertain Unit Root in Real GNP,” *The American Economic Review* 83, pp.264–272.

Sims, C.A. (1988), “Bayesian Skepticism on Unit-Root Econometrics,” *Journal of Economic Dynamics & Control* 12, pp.463–474.

Zivot, E. and D.W.K. Andrews (1992), “Further Evidence on the Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis,” *Journal of Business & Economic Statistics* 10, pp.251–270.