working paper
department
of economics

WHAT DO WE LEARN FROM UNIT ROOTS
IN MACROECONOMIC TIME SERIES?
by
Danny Quah

Number 469 October 1987

massachusetts
institute of
technology

50 memorial drive
cambridge, mass. 02139

WHAT DO WE LEARN FROM UNIT ROOTS IN MACROECONOMIC TIME SERIES?

by

Danny Quah

Number 469 October 1987

What Do We Learn from Unit Roots in Macroeconomic Time Series?

Danny Quah*
October 1987.
MIT Economics Department
and NBER.

* Department of Economics, MIT, Cambridge MA 02139. I thank Olivier Blanchard, Stanley Fischer, James Poterba and Julio Rotemberg for useful comments on earlier versions of this paper. I also thank John Taylor, Mark Watson and an anonymous referee for suggestions and criticisms that have helped to focus the discussion. The hospitality of the MIT Statistics Center is gratefully acknowledged.

Abstract

It is often argued that the presence of a unit root in aggregate output implies that there is no "business cycle": the economy does not return to trend following a disturbance. This paper makes this notion precise, but then develops a simple aggregative model where this relation is contradicted. In the model, output both has a unit root, and displays repeated short-run fluctuations around a deterministic trend. Some summary statistical evidence is presented that suggests the phenomena described in the paper is not without empirical basis.

Much macroeconomic research has focused on the evidence for and the implications of unit roots in aggregate output. Well-known examples of this work include the contributions of John Campbell and N. Gregory Mankiw (1987), Peter Clark (1987), John Cochrane (1987), Steven Durlauf and Peter Phillips (1986), Charles Nelson and Charles Plosser (1982), and Mark Watson (1986). Many have taken away from this discussion the view that business cycles, relative to secular changes, are small and insignificant: output does not fluctuate around trend.

This paper has two main objectives. The first is pedagogical: it makes technically precise the relation between the presence of unit roots and the phenomenon of (lack of) cycle around trend in a way more explicit than is currently available in the macroeconomic literature. Most studies almost instinctually take the key characteristic to be the implied path of an impulse response function. (See for instance Campbell and Mankiw, 1987. In the empirical literature, a notable exception is Francis Diebold and Glenn Rudebusch, 1987.) The impulse response may be a sensible statistic to examine for many purposes, but it is ill-suited to represent the absence of cycle about trend. The reason is that the presence of fluctuations about trend is a property of the sample path behavior of a time series, not of its impulse response function. The latter is at best a conditional expectation, and in general not even that, but only a linear least-squares predictor. Nevertheless, there is a precise sense in which the presence of unit roots is inconsistent with fluctuations about trend: average cycle length is infinite when a linear time series process has a unit root.

The first objective therefore serves to sharpen discussion on what exactly are the implications of a process having a unit root. That the points made here are important is illustrated by the second goal of the paper. We construct a simple model of employment growth and learning-by-doing. In the model, the conditional expectation of output is at least as explosive as that for a unit root process; however the sample path for output *always* displays well-behaved finite length cycles about trend. Thus the finding that estimated ARMA representations for output have a unit root should not be identified with a view that there are no business cycles about trend. Some summary statistical evidence for aggregate output in the US is also presented that suggests that the distinction here is not empirically irrelevant.

The plan of the paper is as follows. Section I is a technical section that relates the presence of unit roots to observable sample path behavior. This contains arguments that would normally be omitted, except for the fact that impulse response functions have been so convincing to so many. Section II presents a simple economic model where unit roots are consistent with well-behaved finite length cycles about trend. Further necessary technical results are also developed here. Section III reports informal summary statistics that suggest some empirical basis for the main message of Section II. The paper concludes with a brief Section IV. Technical proofs are presented in an Appendix.

I. Unit Roots and Sample Paths.

For the most part, this section will formally consider a pure random walk with drift. The main conclusions carry over to where the first difference is a serially correlated stationary sequence. Throughout the discussion, the reader should keep in mind the distinction between difference stationary and trend stationary sequences emphasized by Nelson and Plosser (1982) and Campbell and Mankiw (1987).

For $t \geq 1$, let $Y(t)$ be an observed sequence generated by

$$Y(t) = \beta + Y(t-1) + \epsilon(t),$$

where ϵ is a zero mean covariance stationary sequence, with spectral density bounded away from zero. The sequence Y is therefore difference-stationary, in the terminology of Nelson and Plosser (1982).

For any time trend coefficient α, the deviations process

$$W(t;\alpha) = Y(t) - \alpha t = Y(0) + (\beta - \alpha)\cdot t + \sum_{j=1}^{t} \epsilon(t)$$

remains a difference stationary sequence:

$$W(t;\alpha) = (\beta - \alpha) + W(t-1;\alpha) + \epsilon(t),$$

$$W(0;\alpha) = Y(0).$$

Thus, removing any deterministic time trend from a difference stationary sequence only produces yet another difference stationary sequence. The probability law for the first-differenced sequence remains completely unchanged, except possibly in mean. The first lesson can therefore be stated as follows: *unless to begin with, there is a presumption that some difference stationary sequence should display cyclicality, arbitrary linear detrending cannot induce spurious cyclicality.* (This may of course fail if the "trend" coefficient is continually being re-chosen when a given finite sample is growing through time.)

Next, write

$$Y(t) = \beta \cdot t + \sum_{j=1}^{t} \epsilon(j) + Y(0).$$

Notice that the right hand side contains terms of different asymptotic order. The deterministic part $\beta \cdot t$, for β different from zero, is of mean square asymptotic order $O(n^3)$ whereas the stochastic component is only of mean square asymptotic order $O(n)$, where n is the sample size. Thus, the more likely avenue for deviation from cyclical behavior is the dominant deterministic trend part. However, this component is shared as well by trend-stationary sequences (data that is covariance stationary about a linear trend). Thus we will ignore this part in the subsequent discussion, and concentrate on the purely stochastic component. But then this is exactly a zero drift unit root sequence.

To fix ideas and to simplify the calculations, suppose that ϵ is an *iid* sequence, which takes the values 1 and -1 with equal probability. If anything, one would imagine that this assumption severely restricts the possible cyclical behavior of the accumulation $\sum \epsilon(j)$: once the sum gets sufficiently distant from an arbitrary starting point, it takes a special sequence of events for the process to return to its origin. If ϵ were to have infinite support (such as for a normally distributed random variable), $\sum \epsilon(j)$ can always return to its origin in one step. The assumption of serial independence makes the calculations easier; relaxing it would change none of the conclusions.

Let $W(t;\beta) = W(t) = \sum_{j=1}^{t} \epsilon(j)$: this may be interpreted as the deviation of the economy from trend. We study its sample path properties through the following set of questions. First, for $W(t) \neq 0$, what is the behavior of the sequence of conditional expectations $E[W(t+m) \mid W(t)]$, for $m \geq 1$? In words, given that the economy has deviated from trend, what are expectations regarding its future path? Next, for $W(t) = 0$, what is the probability that at time $t+m$, the economy will return to trend, $W(t+m) = 0$? What is the average waiting time until such a return to trend? The average cycle length would then be twice this average waiting time.

For the process here, the conditional expectation $E[W(t+m) \mid W(t)]$ is just $W(t)$, for all $m \geq 1$. More to the point, when $W(t) \neq 0$, the conditional expectation of future $W(t+m)$ is bounded away from zero forever. If ϵ is serially correlated, the exact pattern for the conditional expectations of future W changes somewhat. However, provided ϵ does not have a spectral density that vanishes at frequency zero, the main conclusion that the process is never expected to return to trend holds. Mean reversion, or trend reversion, in this expectation sense, only has to do with whether W has a unit root, not whether the first-difference sequence is serially correlated. Campbell and Mankiw (1987) interpret this to mean that there should be no presumption that an economy should return to its natural rate.

What is the likelihood that the economy *does* return to trend at some future time $t+m$? For the simple example here, this probability is not difficult to calculate. Starting on trend at $W(t)$ say, there is a return to trend at time $t+m = t+2n$ if $W(t+m) = W(t)$. The probability of each possible sample path with $W(t+2n) = W(t)$ is $(1/2)^n \cdot (1/2)^n$: this happens when there are n +1's and n -1's in $2n$ independent draws from the distribution of ϵ. However there is more than one such path where $W(t+2n) = W(t)$. The total number of different, and therefore mutually exclusive, paths with $W(t+2n) = W(t)$ is simply the binomial coefficient $\binom{2n}{n} = (2n!)/(n!)^2$. The probability of a return to trend in $2n$ periods is therefore:

$$P_{2n} = \binom{2n}{n} \cdot (1/2)^{2n}, \qquad \text{for } n \geq 1.$$

Defining $P_0 = 1$, the probability generating function for a return to trend is:

$$P(s) = \sum_{n=0}^{\infty} P_{2n} s^{2n},$$

where s is a dummy variable. (The reader who is unfamiliar with these kinds of calculations is referred to William Feller, 1968, Ch. 13.) But then by Newton's binomial formula,

$$P(s) = \sum_{n=0}^{\infty} \binom{2n}{n} \cdot (1/2)^{2n} \cdot s^{2n} = (1-s^2)^{-1/2}.$$

This function diverges to infinity as the variable s tends to unity. This divergence is significant because we can now state the following: *Even in the presence of a unit root, the economy returns to trend infinitely often with probability 1.* In any time period, no matter how far the economy has wandered from trend, the probability is always positive of a return to trend some time in the future.

But this is a relatively weak sense in which shocks to a unit root economy are not permanent. To see why this is, we turn to the final question posed above. What is the average time between returns to trend?

Let the probability of a first return to trend in $2n$ periods be denoted Q_{2n}. Then the sequence $\{Q_{2n},\ n \geq 1\}$ satisfies:

$$P_{2n} = \sum_{j=1}^{n} Q_{2j} P_{2(n-j)}, \qquad \text{for } n \geq 1.$$

The right hand side is simply the sum of probabilities of n mutually exclusive events. For each j, the representative term $Q_{2j} P_{2(n-j)}$ is the probability of a first return to trend in j periods, followed by another return in $2(n-j)$ periods. Taken together, this collection of disjoint events is simply the event that there is a return to trend in $2n$ periods, the probability of which has been computed above.

Multiplying both sides by s^{2n}, summing over $n \geq 1$, and utilizing a fundamental property of convolutions, the equation above becomes:

$$P(s) - 1 = Q(s)P(s),$$

where $Q(s)$ has been defined as $\sum_{n=1}^{\infty} Q_{2n}s^{2n}$. The average waiting time for a return to trend is then

$$\begin{aligned} E\left[2n \,|\, W(t+2n) = W(t) = 0,\ W(t+j) \neq 0,\ 1 \leq j < 2n\right] &= \sum_{n=1}^{\infty} 2n \cdot Q_{2n} \\ &= \lim_{s \to 1} \frac{dQ}{ds}. \end{aligned}$$

But, we also have:

$$Q(s) = 1 - P(s)^{-1} = 1 - (1 - s^2)^{1/2},$$

which implies

$$\frac{dQ}{ds} = s(1 - s^2)^{-1/2} \longrightarrow +\infty \text{ as } s \longrightarrow 1.$$

Thus, *the average length of a cycle in a unit root economy is infinite.*

Therefore, if business cycles are defined to be fluctuations about trend, a unit root economy will not have business cycles of average length equal to the 50 months given by conventional wisdom.

In summary, this section has made precise the sense in which unit roots and business cycles may be viewed as inconsistent. As far as I know, this is the first discussion where these different notions of persistence and business cycles have been explicitly related to the presence of unit roots. (Although again, in the empirical literature, Diebold and Rudebusch, 1987, have taken steps in this direction.) Typically, it is taken for granted that the conditional expectation, or more accurately, the impulse response function is a sufficient measure for persistence. That the distinction between sample path behavior and that of the conditional expectation sequence is important is highlighted in the analysis of the next section.

II. A Model with Unit Roots and (Finite Length) Business Cycles.

This section develops a simple model where aggregate output is linearly represented by a unit (or explosive) root process, but nevertheless output fluctuations about a deterministic trend have finite average length. The model itself is not complicated; however its dynamic properties will need to be studied by using ideas that may be unfamiliar. Thus we first analyze a sequence of examples to build intuition.

There are two papers related to the discussion of this section. James Hamilton (1987) presents a careful empirical investigation of stochastic nonlinearities in aggregate output. His model is substantively different from that in this paper, although both share the feature that there is a discreteness in output growth. Technically, we will show below that a random sequence can be *strictly stationary* with fluctuations of finite average length, even if it has the conditional expectations behavior of a unit root process. Daniel Nelson (1987), in a different context, has also constructed strictly stationary unit root sequences. The mechanism that produces stationarity in his work are quite different from that used here. Taken together however, this suggests that there may be other examples of stochastic sequences that have a unit root but nevertheless remain strictly stationary.

The examples here draw on a generalization of a process that first appeared (as far as I know) in work by Olivier Blanchard (1979). He does not establish any properties for the special case; we do so now for the generalization.

A. Some Technical Results.

Let $X(t)$ evolve as

$$X(t) = \gamma(t)X(t-1) + \eta(t)$$

where (γ, η) is *iid*, with γ independent of η, $E\gamma = \Gamma$, $E\eta = 0$. Suppose also that $|\Gamma| \geq 1$, while

$$\Pr(\gamma(1) = 0) > 0.$$

This can be rewritten as follows:

$$X(t) = \Gamma X(t-1) + [(\gamma(t) - \Gamma)\, X(t-1) + \eta(t)]$$

$$\Rightarrow X(t) = \Gamma X(t-1) + \epsilon(t)$$

Notice that while not independent, $\epsilon(t)$ turns out to be uncorrelated with $X(t-1)$, and in fact, is orthogonal to all lagged X's. This fact allows us to conclude that the equation immediately above is a regression, and that Γ is the population regression coefficient. Thus $X(t)$ is a stochastic process that has explosive roots; when $\Gamma = 1$, X has a unit root.

When $\Pr(\gamma(1) = 0) > 0$, as we assume here, it will be convenient to call X a *clinging process*. The reason for this is that no matter how small this probability is, we have the following property:

Theorem 1 (Stationarity). *A clinging process is strictly stationary, and has mean zero and infinite variance.*

Proof: Appendix.

While X has a linear regression representation that appears explosive (in Box-Jenkins terminology, it does not have a stationary ARMA representation), it is actually stable: its joint unconditional distributions are time-invariant and are independent of initial conditions. Further, as is clear from the arguments in the Proof, X displays fluctuations about its mean that have finite average length.

Next consider any distributed lag function of X, where the lag distribution is stable (i.e., the lag distribution has no zeroes inside the unit circle). The resulting random sequence is again strictly stationary, and again, has an 'explosive' linear autoregressive representation. But now the transformed clinging sequence will display much richer dynamics as well.

Without a distributed lag, X displays much sharper returns to its mean, than movements away from the mean. Thus a sample path analysis may uncover evidence of asymmetry. However, conditional on $X(t)$ and $\gamma(t)$, $X(t+1)$ is symmetrically distributed provided that η is symmetrically distributed. Since the effects of initial conditions vanish eventually, provided that the accumulated sequence in η is dominant, X will also unconditionally have a symmetric distribution. Next, a clinging sequence will display cycles (in the sense of returns to a neighborhood of its mean) persistently and infinitely often. As mentioned above, these cycles turn out to have finite average length. If a time trend is added to a clinging sequence, this example would reconcile a number of seemingly contradictory findings: J. Bradford DeLong and Lawrence Summers (1985) find that after detrending, business cycles are actually 'symmetric' in unconditional distribution. In opposition to this, Salih Neftçi (1984) concludes that certain features of economic fluctuations do display a kind of asymmetry. Campbell and Mankiw (1987) and Nelson and Plosser (1982) find evidence for a unit root in linear representations for output; however it also a wide-spread belief that fluctuations are persistently recurring events. If the clinging model is correct, these statements are not inconsistent.

The conclusion from this discussion is that the presence of unit roots alone is not evidence against well-behaved finite length business cycles. Further, in the sense that the joint distributions are independent of initial conditions, disturbances do not persist indefinitely even in the presence of unit roots.

The question that remains is the following: Is there a plausible economic mechanism that will produce a clinging sequence for output in equilibrium? We turn to this next.

B. A Simple Economic Model.

The model to be described here relates employment, learning-by-doing, and output growth. Output is produced with two factors of production: capital K and labor N. In any time period t, the measured stock of labor comprises skilled and unskilled employees:

$$N(t) = N_0(t) + N_1(t).$$

Skilled labor N_0 is labor that has been employed for at least one period. Unskilled labor N_1 comprises new entrants. We make the extreme assumption that only skilled labor is productive; unskilled labor is in training and is completely unproductive in the apprenticeship period.

Skilled labor evolves as follows:

$$N_0(t) = \min[N(t-1) - N_2(t),\ 0],$$

where $N_2(t)$ is the labor withdrawn in period t. This would include factors such as retirements, a reduction of market-wide skills due to the introduction of new technology, or possibly sectoral shifts between industries, so that the average skill level falls. This equation states that skilled labor this period is equal to skilled labor last period minus any decumulation due to labor withdrawals; at worst, skilled labor falls to zero.

For simplicity, we assume N_2 and N_1 have the same expectation, so that on average, the number of new entrants just equals the number of withdrawals. Also assume that both N_2 and N_1 are serially independent and to rule out uninteresting degeneracies, we assume N_2 and N_1 have strictly positive variances. (The arguments below become a little more complicated and subtle if these don't hold, but the main substantive conclusions can be made to remain intact.)

We then have the following result:

Theorem 2 (Eventual Loss of all Skilled Labor). *For an arbitrary initial quantity of skilled labor $N_0(0) > 0$, the probability is positive that at some period $t > 0$, there is a total loss of skilled labor: $N_0(t) = 0$.*

Proof: Appendix.

This is not a deep result: it says that it will be unusual to come upon situations where the stock of labor is always productive. One suspects that small deviations from the assumptions here will leave the major conclusion unmodified. Notice that the result allows the variation in withdrawals N_2 to be arbitrarily small. If labor entrants on average exceed withdrawals, then the result may not hold, as there is then a positive drift in the stock of labor. However for population growth rate not too different from zero, the difference must be small, so that equality of the two means is a good approximation. Further, most would agree that aggregate demographic factors cannot be too much related to economic fluctuations. Thus a nongrowing labor force is probably the natural assumption to use here.

The rest of the model economy is trivial. Let the production function be such that (the logarithm of) output is described by $Y(t) = N_0(t)\,K(t) + \eta(t)$, where η is a technology shock. Capital decays completely after one period; however there is no consumption, and all output is invested, so that $K(t) = Y(t-1)$.

Thus the model displays high persistence in output when $N_0(t)$ exceeds 1. Given a technically skilled work force, output evolves as a highly positively correlated sequence. Productivity shocks that increase output this period, $\eta(t) > 0$, tend to increase output persistently in the near future. Increased output implies a higher capital stock; with skilled labor, this increase in output is transmitted to increased output in future periods as well. By Theorem 2 however, no matter how small the fluctuations in withdrawals from the labor force, at some stage, the economy is sure to lose all its skilled labor. When this happens, the accumulated capital is "useless", except to train labor for the next period. In this case, the effects of productivity shocks are completely transitory.

The model therefore quite easily produces a process for output that is (close to) a clinging sequence. The crucial property that the random coefficient on lagged output periodically realizes as zero is borne by our assumptions on the process followed by skilled labor. The model has quite reasonable assumptions; we see that no very special feature is needed to produce persistently recurring short-lived fluctuations despite the possible existence of a unit root linear representation. Slight variations in the assumptions on capital decay can be accommodated by the fact that distributed lag transformations of a clinging sequence leave its key properties unchanged. The fact that the loss of *all* of the skilled labor force is technically a key factor in producing the result should not detract from the plausibility of the effects here. (As an example on a smaller scale, consider a department's loss of a senior econometrician. Even if he or she is one among a few and the loss occurs only for a short period of time, it may sufficiently slow the flow of graduate students in that field that the loss is felt for several generations hence.)

In summary, we describe again informally the characteristics of the model. Productivity shocks both have a persistent long-run effect (on average), and also an eventually transitory effect. The persistence comes from a feedback between capital and output. Increased output leads to a high capital stock. With a highly skilled labor force, this produces even higher output in the next period. The transitoriness comes from the fact that the slight (exogenous) variation in labor withdrawals and entrants leads periodically to a completely unskilled labor force that needs some length of time to become re-trained. Given an economy-wide low-skilled labor force, no amount of beneficial productivity shocks is immediately persistent in increased output. Under these circumstances, productivity shocks have a transitory effect. Of course, this effect is again only temporary for the skilled labor force will come online again sometime in the future. Thus in a low output period, the capital stock is only apparently unproductive; it serves to tool up the labor force for subsequent periods. This example has therefore embedded the dynamics of a clinging sequence in an economic setting: we see that it is in fact quite plausible to model output as a clinging sequence.

The next section is empirical: it attempts to obtain some evidence on whether the kinds of effects described here are present in actual aggregate measures of output.

III. Some Empirical Evidence.

The clinging sequence above produces a meaningful distinction between 'hard' and 'soft' unit roots. The first kind of unit root describes the processes of Section I; the second describes those of Section II. We now present some summary statistical evidence on distinguishing the two.

Let $Y(t)$ be (the natural logarithm of) some measure of aggregate output. The null hypothesis is that Y contains a hard unit root:

$$Y(t+1) = \beta + Y(t) + A(L)^{-1}\eta(t+1), \qquad \text{for } t \geq 0,$$

where η is *iid* $\mathcal{N}(0, \sigma_\eta^2)$. Defining $W(t; Y(0), \beta) = Y(t) - Y(0) - \beta\, t$, this can be rewritten:

$$W(t+1; Y(0), \beta) = W(t; Y(0), \beta) + A(L)^{-1}\eta(t+1)$$

Next, let $X(t) = A(L)W(t)$, so that:

$$X(t; Y(0), \beta) = X(t-1; Y(0), \beta) + \eta(t).$$

The alternative hypothesis that Y contains a soft unit root is most concisely expressed as follows. In the definition of W, substitute β_0 in place of $Y(0)$, where β_0 is a free parameter. Then write

$$X(t) = \gamma(t)\, X(t-1) + \eta(t),$$

where γ, η are pairwise and serially independent, η is $\mathcal{N}(0, \sigma_\eta^2)$, $E\gamma = 1$, $\Pr(\gamma = 0) = p \geq 0$, and X is defined exactly as above.

In words, a filtered version of deviations from trend $Y(0) + \beta t$ is hypothesized under the null to be a pure random walk with no drift. Under the alternative, the same filtered version of trend deviations is hypothesized to be a clinging sequence with a soft unit root. The difference stationary model is strictly nested within the clinging alternative: that null hypothesis model obtains by setting the variance of γ to zero and β_0 to $Y(0)$.

Estimation under the null proceeds as follows. First difference (the log of) observed output, and estimate the regression

$$Y(t) - Y(t-1) = \beta\, A(1) - \left[L^{-1}A(L)\right]_+ \left(Y(t-1) - Y(t-2)\right) + \eta(t).$$

The model is just identified. Under our assumptions, ordinary least squares is equivalent to maximum likelihood. Estimation under the alternative is a little more complicated. We parametrize γ as the product of serially and pairwise independent Bernoulli and normal random variables:

$$\gamma(t) = b(t) \cdot \epsilon(t),$$

where b is Bernoulli with probability p of equalling unity, ϵ is $\mathcal{N}(p^{-1}, \sigma_\epsilon^2)$. Setting σ_ϵ^2 to zero and p to unity recovers the property implied by the null hypothesis that $\gamma = 1$ for all t. The likelihood function for this alternative model can be written in a way close to that of an unobserved switching regression model. Call the vector of parameters in the model θ; the likelihood of the t-th observation on $X(t)$ conditional on $X(t-1)$ is

$$\ell(X(t)\,|\,X(t-1);\,\theta) = p \cdot \phi\left(\frac{X(t) - p^{-1}\,X(t-1)}{\left(X(t-1)^2\sigma_\epsilon^2 + \sigma_\eta^2\right)^{1/2}}\right) + (1-p)\cdot\phi(X(t)/\sigma_\eta).$$

where ϕ is the density of a standard normal. The likelihood of the entire sample conditional on $X(0)$ is seen to be simply the product of these conditional likelihoods. Since X is strictly stationary under the alternative, the impact of the initial condition $X(0)$ vanishes asymptotically. Parameter estimates obtained under the null are used as starting points for the maximum likelihood observation.

Tables 1 and 2 display estimates of the model for two measures of aggregate output, quarterly GNP and monthly industrial production. Results are presented for $A(L)$ chosen to be first through third order; that is, $A(L) = \sum_{j=0}^{k} a_j L^j$ where a_0 is always 1 and k varies from 1 through 3. The GNP series is measured in 1982 dollars and industrial production is seasonally adjusted.

There are three restrictions under the null: $\beta_0 = Y(0)$, $p = 1$, and $\sigma_\epsilon^2 = 0$. Thus, twice the difference in log-likelihoods is asymptotically distributed χ^2 with three degrees of freedom. For GNP, this test-statistic takes values 2.614, 3.842, and 1.532 with marginal significance levels of 0.46, 0.28, and 0.67, respectively, when A is set to first, second, and third order lag polynomials. Therefore, the null hypothesis that GNP is difference stationary, rather than trend stationary cannot be rejected. GNP has a hard unit root.

Turning to the industrial production data in Table 2, the null hypothesis of a hard unit root is now seen to be soundly rejected. For second and third order polynomials, the differences in log-likelihoods is around 50, with associated marginal significance levels well below the standard cutoff points.

Table 1 GNP82, 1947,1 to 1986,1

(ln, $\times 10^3$)

	1-lag				2-lag				3-lag			
	OLS		MLE		OLS		MLE		OLS		MLE	
β_0	6962.7	*	6963.64	(21.27)	6962.7	*	6980.22	(22.84)	6962.7	*	6963.37	(21.34)
β	7.909	(0.61)	8.784	(0.54)	7.946	(0.47)	7.884	(0.27)	7.886	(0.72)	8.795	(0.35)
p	1.0	*	0.991	(0.11)	1.0	*	0.990	(0.12)	1.0	*	0.991	(0.11)
σ_η^2	109.29		100.07	(1.21)	107.70		94.38	(1.20)	105.56		97.21	(1.20)
σ_ε^2	0.0	*	0.0	(0.09)	0.0	*	0.00	(0.14)	0.0	*	0.00	(0.10)
a_1	-0.366	(0.08)	-0.408	(0.07)	-0.315	(0.08)	-0.369	(0.08)	-0.338	(0.08)	-0.376	(0.08)
a_2	–	–	–		-0.138	(0.08)	-0.113	(0.08)	-0.189	(0.08)	-0.176	(0.08)
a_3	–		–		–		–		+0.161	(0.08)	+0.140	(0.08)
ln L	-438.439		-437.132		-434.469		-432.548		-430.104		-429.338	

* Fixed in the Restricted Estimation. Numbers in Parentheses are numerical standard errors.

Table 2 Industrial Production, 1947,1 to 1986,3

(ln, $\times 10^4$)

	1-lag				2-lag				3-lag			
	OLS		MLE		OLS		MLE		OLS		MLE	
β_0	33534.0	*	†		33534.0	*	35141.8	(241.3)	33534.0	*	35278.0	(2.37)
β	31.12	(0.40)			30.94	(0.36)	30.13	(0.99)	31.57	(0.34)	29.80	(0.53)
p	1.0	*			1.0	*	0.990	(0.72)	1.0	*	0.989	(0.73)
σ_η^2	11032				10909		6272.9	(7.28)	10904		6171.0	(7.14)
σ_ε^2	0.0	*			0.0	*	0.01	(0.03)	0.0	*	0.01	(0.04)
a_1	-0.437	(0.04)			-0.397	(0.05)	-0.499	(0.04)	-0.393	(0.05)	-0.497	(0.04)
a_2	–		–		-0.093	(0.05)	-0.099	(0.04)	-0.075	(0.05)	-0.073	(0.04)
a_3	–		–		–		–		-0.046	(0.05)	-0.071	(0.04)
ln L	-2412.20				-2404.44		-2353.25		-2399.17		-2346.70	

* Fixed in the Restricted Estimation. Numbers in Parentheses are numerical standard errors.
† Unchanged from Restricted Estimates.

Examining the results more closely, we see that individually the probability p is not significantly different from 1, and that σ_η^2 is not significantly different from zero. Thus the reason for rejection appears to come mainly from the β_0 intercept term. However, the estimates for these three parameters are correlated and so the individual t-statistics may be misleading. Testing for the significance of all three is essentially equivalent to examining the likelihood-ratio statistic, which is what we have done here. Evidently, allowing for the possibility that $1 - p$ and σ_η^2 may differ from 0 is sufficient to pin down the intercept term quite precisely. But it is one of the important implications of the hard unit root hypothesis that the intercept in the "trend" line changes for every different initial condition. More precisely, OLS estimates for this intercept diverge as the sample size grows arbitrarily large (see for example Durlauf and Phillips, 1986). Our results indicate that, contrary to this, for our sample of industrial production data, the intercept, and consequently the trend line, is in fact quite precisely estimated.

A plot of the likelihood function for industrial production (not presented here) shows a distinct and unique peak in p at 0.99. This implies a realization of $\gamma(t) = 0$ every 91 months on average. This is approximately twice the consensus length of a business cycle. However, notice that $A(L)$ is a further source for cyclical dynamics. In the second order case, this polynomial has roots of 0.65 and 0.16. These roots are well inside the unit circle, and serve to enrich the estimated cyclicality.

IV. Conclusion

This paper has undertaken two tasks. The first and pedagogical objective in Section I makes clear why unit root processes and the notion of persistently recurring business cycles are at odds with each other. The analysis in terms of the relation between sample path behavior and the presence of unit roots appears not to have been made explicit thus far in the macroeconomic literature.

That it is important to make this connection becomes clear in Section II where a simple model is constructed that displays persistently recurring fluctuations in the sample paths of output despite the presence of a unit root. Summary statistical evidence presented in Section III suggests that the effects described in the model are not without empirical basis for measures of aggregate output in the US.

Appendix.

This appendix proves the technical results in the paper.

Proof of Theorem 1. *We verify that the joint distribution of subsequences of X is invariant across time shifts. For $k \geq 1$, let $t_1, t_2, \ldots, t_k$ be some collection of integers, assumed increasing without loss of generality. We wish to show that for all integer s:*

$$F(X(t_1), X(t_2), \ldots, X(t_k)) = F(X(t_1+s), X(t_2+s), \ldots, X(t_k+s)),$$

where F denotes the joint distribution function. From the form of the clinging sequence, it suffices to establish equality of the unconditional distributions of $X(t_1)$ and $X(t_1+s)$. Let m, M be integer with $M > 1$. Iterating on $X(m) = \gamma(m)X(m-1) + \eta(m)$, we have

$$X(m) = \left(\prod_{j=m+1-M}^{m} \gamma(j)\right) X(m-M) + \eta(m) + \sum_{j=m+1-M}^{m-1} \left(\prod_{k=j+1}^{m} \gamma(k)\right) \eta(j).$$

If $\gamma(j)$ vanishes for some j between $m+1-M$ and m, $X(m)$ becomes independent of $X(m-M)$. In that case, $X(m)$ can be written entirely in terms of the (γ, η) partial sequence between epochs $m+1-M$ and m. Setting m to t_1 and t_1+s in turn, we see that the distributions of $X(t_1)$ and $X(t_1+s)$ are identical, conditional on a zero realization for $\gamma(j)$, for some j between $m+1-M$ and m. More explicitly, if for some j, $0 \leq j < M$, we have $\gamma(t_1 - j) = \gamma(t_1 + s - j) = 0$, then

$$F(X(t_1), X(t_2), \ldots, X(t_k)) = F(X(t_1+s), X(t_2+s), \ldots, X(t_k+s)).$$

Let p denote $\Pr(\gamma(0) = 0)$, with $p > 0$. The conditioning event

$$\{\gamma(t_1 - j) = \gamma(t_1 + s - j) = 0, \text{ for some } j \text{ between } 0 \text{ and } M-1\}$$

occurs with probability bounded from below by $1 - (1-p^2)^M$. But this tends to 1 as $M \to \infty$, hence establishing strict stationarity. To show that EX is zero, write for $n \geq m \geq 0$:

$$E\left(X(t)\,|\,\gamma(t-m), \gamma(t-m-1), \ldots, \gamma(t-n), X(t-n-1)\right) = |\Gamma|^{-m} \left(\prod_{j=m}^{n} \gamma(t-j)\right) X(t-n-1).$$

Define $\varsigma_{m,n}$ to be this conditional expectation: it is a well-defined random variable as it is simply a finite product of random variables. Since γ is iid and places positive probability on 0, for fixed m the random *sequence* $\{\varsigma_{m,n},\ n \geq m\}$ *converges almost surely to the degenerate random variable at 0. By the law of iterated expectations,* $E(\lim_{n\to\infty} \gamma_{m,n}) = 0$ *uniformly in* m *which implies* $E(X(t)) = 0$. *The infinite variance property is obvious. QED*

Proof of Theorem 2. *Suppose that contrary to the conclusion of the theorem,* $N(t) > 0$ *for all* t. *But then* N *is a zero drift random walk. Thus, from any starting point,* N *will take on arbitrarily negative values infinitely often with probability 1. This is a contradiction as* N *is restricted to be always nonnegative. QED*

FOOTNOTES.

[1] Department of Economics, MIT, Cambridge MA 02139. I thank Olivier Blanchard, Stanley Fischer, James Poterba and Julio Rotemberg for useful comments on earlier versions of this paper. I also thank John Taylor, Mark Watson and an anonymous referee for suggestions and criticisms that have helped to focus the discussion. The hospitality of the MIT Statistics Center is gratefully acknowledged.

References

Blanchard, O.J., "Speculative Bubbles, Crashes, and Rational Expectations," *Economics Letters*, Nov. 1979, *3* , 387-389.

Campbell, J. and N.G. Mankiw, "Are Output Fluctuations Transitory?" *Quarterly Journal of Economics*, 1987, forthcoming.

Clark, P., "The Cyclical Component of US Economic Activity," *Quarterly Journal of Economics*, 1987, forthcoming.

Cochrane, J.H., "How Big is the Random Walk in GNP?" University of Chicago mimeo, 1986.

DeLong, J.B. and L. Summers, "Are Business Cycles Symmetric?" Harvard University mimeo, 1985.

Diebold, F.X. and G.D. Rudebusch, "Does the Business Cycle have Duration Memory?" Board of Governors of the Federal Reserve System, Division of Research and Statistics, Special Studies Paper No. 223, August 1987.

Durlauf, S.N and P.C.B. Phillips, "Trends versus Random Walks in Time Series Analysis," Cowles Foundation Discussion Paper no. 788, Yale University, 1986.

Feller, W., *An Introduction to Probability Theory and its Applications*, vol. 1, 3rd edition, New York, John Wiley, 1968.

Hamilton, J.D., "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," University of Virginia mimeo, July 1987.

Neftçi, S.N., "Are Economic Time Series Asymmetric over the Business Cycle?" *Journal of Political Economy*, April 1984, *92*, 307-328.

Nelson, C.R. and C.I. Plosser, "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics*, September 1982, *10*, 139-162.

Nelson, D., "The Steady State Distribution of the IGARCH(1, 1) and explosive GARCH(1, 1) Processes," MIT mimeo, 1987.

Watson, M.W., "Univariate Detrending Methods with Stochastic Trends," *Journal of Monetary Economics*, July 1986, *18*, 49-75.