THE SOURCES AND NATURE OF LONG-TERM MEMORY IN THE BUSINESS CYCLE

by

Joseph G. Haubrich
Andrew W. Lo

(5-89)

RODNEY L. WHITE CENTER FOR FINANCIAL RESEARCH
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6367

THE SOURCES AND NATURE OF LONG–TERM MEMORY

IN THE BUSINESS CYCLE

Joseph G. Haubrich* and Andrew W. Lo**

First Draft: May 1988
Latest Revision : January 1989

This paper examines the stochastic properties of aggregate macroeconomic time series from the standpoint of fractionally integrated models, and focuses on the persistence of economic shocks. We develop a simple macroeconomic model that exhibits long-term dependence, a consequence of aggregation in the presence of real business cycles. We derive the relation between properties of fractionally integrated macroeconomic time series and those of microeconomic data, and discuss how fiscal policy may alter their stochastic behavior. To implement these results empirically, we employ a test for fractionally integrated time series based on the Hurst-Mandelbrot rescaled range. This test is robust to short-term dependence, and is applied to quarterly and annual real GNP to determine the sources and nature of long-term dependence in the business cycle.

*Department of Finance, Wharton School, University of Pennsylvania.
**Sloan School of Management, Massachusetts Institute of Technology and NBER.

We thank Don Andrews, Pierre Perron, Fallaw Sowell, and seminar participants at Columbia University, the NBER Summer Institute, the Penn Macro Lunch Group, and the University of Rochester for helpful comments. Research support from the National Science Foundation (Grant No. SES–8520054), the Rodney L. White Fellowship at the Wharton School (Lo), the John M. Olin Fellowship at the NBER (Lo), and the University of Pennsylvania Research Foundation (Haubrich) is gratefully acknowledged.

1. Introduction.

Questions about the persistence of economic shocks currently occupy an important place in macroeconomics. Most controversy has centered on whether aggregate time series are better approximated by fluctuations around a deterministic trend, or by a random walk plus a stationary or temporary component. The empirical results from these studies are mixed, perhaps because the time series representation of output may belong to a much wider class of processes than previously considered. In particular, earlier studies have ignored the class of fractionally integrated processes, which serves as a compromise between the two stochastic models usually considered, and which also exhibits an interesting type of long range dependence. This new approach also accords well with the classical NBER business cycle program exemplified by Wesley Claire Mitchell, who urged examination of trends and cycles at all frequencies.

Economic life does not proceed smoothly: there are good times, bad times, alternating periods of prosperity and depression, a "rhythmical fluctuation in activity." Recurrent downturns and crises take place roughly every 3 to 5 years and thus seem part of a non-periodic cycle. Studying such cycles in detail has been the main activity of twentieth century macroeconomics. Even so, isolating cycles of these frequencies has been difficult because the data evince many other cycles of longer and shorter duration. Wesley Mitchell (1927, p. 463) remarks "Time series also show that the cyclical fluctuations of most [not all] economic processes occur in combination with fluctuations of several other sorts: secular trends, primary and secondary, seasonal variations, and irregular fluctuations." Properly removing these other influences has always been controversial. No less an authority than Irving Fisher (1925) considered the business cycle a myth, akin to runs of luck at Monte Carlo. In a similar vein, Slutzky (1937) suggested that cycles arose from smoothing procedures used to create the data.

A similar debate is now taking place. The standard methods of removing a linear or exponential trend assume implicitly that business cycles are fluctuations around a trend. Other work [e.g. Nelson and Plosser (1982)] challenges this and posits stochastic trends similar to random walks, highlighting the distinction between temporary and permanent changes. Since empirically the cyclical or temporary component is small relative to the fluctuation in the trend component [the random walk part], business cycles look more like Fisher's myth. This is important for forecasting purposes because permanent changes [as in the case of a random walk] today have a large effect many periods later, whereas temporary changes [as in stationary fluctuations around a trend] have small future effects. The large random walk component also provides evidence against some theoretical models of aggregate output. Models that focus on monetary or aggregate demand disturbances as a source of transitory fluctuations cannot explain much output variation; supply side or other models must be invoked [e.g. Nelson and Plosser (1982), Campbell and Mankiw (1987)].

However, the recent studies posit a misleading dichotomy. In stressing trends versus random walks, they overlook earlier work by Mitchell [quoted above], Adelman, and Kuznets, who have stressed correlations in the data intermediate between secular trends and transitory fluctuations. In the language of the interwar NBER, they are missing Kondratiev, Kuznets and Juglar cycles. In particular, in the language of modern time series analysis, the search for the stochastic properties of aggregate variables has ignored the class of fractionally integrated processes. These stochastic processes, midway between white noise and a random walk, exhibit a long run dependence no finite ARMA model can mimic, yet lack the permanent effects of an ARIMA process. They show promise of explaining the lower frequency effects, the long swings emphasized by Kuznets (1930), Adelman (1965) or the effects which persist from one business cycle to the next. Since long memory models exhibit dependence without the permanent effects of an ARIMA process, they may not be detected by standard methods of fitting Box-Jenkins models. This calls for a more direct investigation of this alternative class of stochastic processes.

This paper examines the stochastic properties of aggregate output from the standpoint of fractionally integrated models. We introduce this type of process in Section 2, reviewing its main properties, advantages, and weaknesses. Section 3 develops a simple macroeconomic model that exhibits long-term dependence. Section 4 employs a new test for fractional integration in time series to search for long-term dependence in the data. Though related to a test of Hurst's and Mandelbrot's, it is robust to short-term dependence. We conclude in Section 5.

2. Review of Fractional Techniques in Statistics.

A random walk can model time series that look cyclic but non-periodic. The first differences of that series [or in continuous time, the derivative] should then be white noise. This is an example of the common intuition that differencing [differentiating] a time series makes it "rougher," whereas summing [integrating] it makes it "smoother." Many macroeconomic time series look like neither a random walk nor white noise, suggesting that some compromise or hybrid between the random walk and its integral may be useful. Such a concept has been given content through the development of the fractional calculus, i.e., differentiation and integration to non-integer orders.[1] The fractional integral of order between 0 and 1 may be viewed as a filter that smooths white noise to a lesser degree than the ordinary integral; it yields a series that is rougher than a random walk but smoother than white noise. Granger and Joyeux (1980) and Hosking (1981) develop the time series implications of fractional differencing in discrete time. For expositional purposes we review the more relevant properties in Sections 2.1 and 2.2.

2.1. Fractional Differencing.

Perhaps the most intuitive exposition of fractionally differenced time series is via their infinite-order autoregressive and moving-average representations. Let X_t satisfy:

$$(1-L)^d X_t \quad = \quad \epsilon_t \tag{2.1}$$

where ϵ_t is white noise, d is the degree of differencing, and L denotes the lag operator. If $d=0$ then X_t is white noise, whereas X_t is a random walk if $d=1$. However, as Granger and Joyeux (1980) and Hosking (1981) have shown, d need not be an integer. From the binomial theorem, we have the relation:

$$(1-L)^d \quad = \quad \sum_{k=0}^{\infty}(-1)^k \binom{d}{k} L^k \tag{2.2}$$

[1] The idea of fractional differentiation is an old one, dating back to an oblique reference by Leibniz in 1695, but the subject lay dormant until the 19th century when Abel, Liouville, and Riemann, developed it more fully. Extensive applications have only arisen in this century; see, for example, Oldham and Spanier (1974). Kolmogorov (1940) was apparently the first to notice its applications in probability and statistics.

where the binomial coefficient $\binom{d}{k}$ is defined as:

$$\binom{d}{k} \equiv \frac{d(d-1)(d-2)\cdots(d-k+1)}{k!} \tag{2.3}$$

for any real number d and non-negative integer k.[2] From (2.2) the AR representation of X_t is apparent:

$$A(L)X_t = \sum_{k=0}^{\infty} A_k L^k X_t = \sum_{k=0}^{\infty} A_k X_{t-k} = \epsilon_t \tag{2.4}$$

where $A_k \equiv (-1)^k \binom{d}{k}$. The AR coefficients are often re-expressed more directly in terms of the gamma function:

$$A_k \equiv (-1)^k \binom{d}{k} = \frac{\Gamma(k-d)}{\Gamma(-d)\Gamma(k+1)} . \tag{2.5}$$

By manipulating (2.1) mechanically, X_t may also be viewed as an infinite-order MA process since:

$$X_t = (1-L)^{-d}\epsilon_t = B(L)\epsilon_t \qquad B_k = \frac{\Gamma(k+d)}{\Gamma(d)\Gamma(k+1)} . \tag{2.6}$$

The particular time series properties of X_t depend intimately on the value of the differencing parameter d. For example, Granger and Joyeux (1980) and Hosking (1981) show that when d is less than $\frac{1}{2}$, X_t is stationary; when d is greater than $-\frac{1}{2}$, X_t is invertible. Although the specification in (2.1) is a fractional integral of pure white noise, the extension to fractional ARIMA models is clear.[3]

The AR and MA representations of fractionally differenced time series have many applications, and illustrate the central properties of fractional processes, particularly long-term dependence. The MA coefficients B_k tell the effect of a shock k periods ahead, and

[2] When d is an integer, (2.3) reduces to the better-known formula for the binomial coefficient $\frac{d!}{k!(d-k)!}$. We follow the convention that $\binom{d}{0} = 1$ and $\binom{0}{0} = 0$.

[3] See Hosking (1981) for further details.

indicate the extent to which current levels of the process depend on past values. How fast this dependence decays furnishes valuable information about the process. Using Stirling's approximation, we have:

$$B_k \approx \frac{k^{d-1}}{\Gamma(d)} \tag{2.7}$$

for large k. Comparing this with the decay of an AR(1) process highlights a central feature of fractional processes: they decay hyperbolically, at rate k^{d-1}, rather than at the exponential rate of ρ^k for an AR(1). For example, compare in Figure 1 the autocorrelation function of the fractionally differenced series $(1-L)^{0.475}X_t = \epsilon_t$ with the AR(1) $X_t = 0.9X_{t-1} + \epsilon_t$. Although they both have first-order autocorrelations of 0.90, the AR(1)'s autocorrelation function decays much more rapidly. Figure 2a plots the impulse-response functions of these two processes. At lag 1 the MA-coefficients of the fractionally differenced series and the AR(1) are 0.475 and 0.900 respectively; at lag 10 they are 0.158 and 0.349, and at lag 100 they are 0.048 and 0.000027. The persistence of the fractionally differenced series is apparent at the longer lags. Alternatively, we may ask what value of an AR(1)'s autoregressive parameter will, for a given lag, yield the same impulse-response as the fractionally differenced series (2.1). This value is simply the k-th root of B_k and is plotted in Figure 2b for various lags when $d = 0.475$. For large k, this autoregressive parameter must be very close to unity.

These representations also show how standard econometric methods can fail to detect fractional processes, necessitating the methods of Section 4. Although a high order ARMA process can mimic the hyperbolic decay of a fractionally differenced series in finite samples, the large number of parameters required would give the estimation a poor rating from the usual Akaike or Schwartz Criteria. An explicitly fractional process, however, captures that pattern with a single parameter d. Granger and Joyeux (1980) and Geweke and Porter-Hudak (1983) provide empirical support for this by showing that fractional models often out-predict fitted ARMA models.

The lag polynomials $A(L)$ and $B(L)$ provide a metric for the persistence of X_t. Suppose X_t represents GNP, which falls unexpectedly this year. How much should this change a forecast of future GNP? To address this issue, define C_k as the coefficients of the lag polynomial $C(L)$ that satisfies the relation $(1-L)X_t = C(L)\epsilon_t$, where the process X_t is given by (2.1). One measure used by Campbell and Mankiw (1987) and Diebold and

Rudebusch (1988) is:

$$\lim_{k \to \infty} B_k = \sum_{k=0}^{\infty} C_k = C(1) . \tag{2.8}$$

For large k, the value of B_k measures the response of X_{t+k} to an innovation at time t, a natural metric for persistence.[4] From (2.7), it is immediate that for $0 < d < 1$, $C(1) = 0$, and asymptotically there is no persistence in a fractionally differenced series, even though the autocorrelations die out very slowly.[5] This holds true not only for $d < \frac{1}{2}$ (the stationary case), but also for $\frac{1}{2} < d < 1$, when the process is nonstationary.

From these calculations, it is apparent that the long run dependence of fractional processes relates to the slow decay of the autocorrelations, not to any permanent effect. This distinction is important; an IMA(1,1) can have small but positive persistence, but the coefficients will never mimic the slow decay of a fractional process.

The long-term dependence of fractionally differenced time series forces us to modify some conclusions about decomposing time series into "permanent" and "temporary" components. Although Beveridge and Nelson (1981) show that non-stationary time series may always be expressed as the sum of a random walk and a stationary process, the stationary component may exhibit long range dependence. This suggests that the temporary component of the business cycle may be transitory only in the mathematical sense and is, for all practical purposes, closer to what we think of as a long non-periodic cycle.

The presence of fractional differencing has yet another implication for the Beveridge-Nelson (BN) approach. Nelson and Plosser (1982) argue that deterministic detrending may overstate the importance of the business cycle if secular movements follow a random walk. However, if the cylical component is an AR(1) with a nearly unit root, McCallum (1986) shows that the BN decomposition will assign too much variance to the permanent

[4] But see Cochrane (1988) and Quah (1988) for opposing views.

[5] There has been some confusion in the literature on this point. Geweke and Porter-Hudak (1983) argue that $C(1) > 0$. They correctly point out that Granger and Joyeux (1980) have made an error, but then incorrectly claim that $C(1) = 1/\Gamma(d)$. If our equation (2.7) is correct, then it is apparent that $C(1) = 0$ [which agrees with Granger (1980) and Hosking (1981)]. Therefore, the focus of the conflict lies in the approximation of the ratio $\Gamma(k+d)/\Gamma(k+1)$ for large k. We have used Stirling's approximation. However, a more elegant derivation follows from the functional analytic definition of the gamma function as the solution to the following recursive relation [see, for example, Iyanaga and Kawada (1980 Section 179.A)]:

$$\Gamma(x+1) = x\Gamma(x)$$

and the conditions:

$$\Gamma(1) = 1 \qquad \lim_{n \to \infty} \frac{\Gamma(x+n)}{n^x \Gamma(n)} = 1 .$$

or trend part and understate the contribution of the temporary component.[6] But suppose the temporary component follows the fractionally differenced time series (2.1). A decomposition in the spirit of BN, i.e., differencing the series the appropriate [fractional] number of times, leaves no temporary component. However, taking first-differences of the series as Plosser and Schwert (1978) suggest will obviously *overstate* the importance of the cyclical component. Therefore, the presence of fractional noise may reverse McCallum's result.

2.2. Spectral Representation.

The spectrum, or spectral density [denoted $f(\omega)$] of a time series specifies the contribution each frequency makes to the total variance. Granger (1966) and Adelman (1965) have pointed out that most aggregate economic time series have a "typical spectral shape" where the spectrum increases dramatically as the frequency approaches zero [$f(\omega) \to \infty$ as $\omega \to 0$]. Most of the power or variance seems concentrated at low frequencies, or long periods. However, prewhitening or differencing the data often leads to "over-differencing" or "zapping out the low frequency component", and often replaces the peak by a dip at 0. Fractional differencing yields an intermediate result. The spectra of fractional processes exhibit peaks at 0 [unlike the flat spectrum of an ARMA process], but ones not so sharp as the random walk's. A fractional series has a spectrum richer in low frequency terms, and more persistence. We illustrate this by calculating the spectrum of fractionally integrated white noise, and also present several formulas needed later on.

Given $X_t = (1-L)^{-d}\epsilon_t$, the series is clearly the output of a linear system with a white noise input, so that the spectrum of X_t is:[7]

$$f(\omega) \;=\; \frac{1}{|1-z|^{2d}}\frac{\sigma^2}{2\pi} \qquad \text{where} \qquad z \equiv e^{-i\omega}, \quad \sigma^2 \equiv E[\epsilon_t^2] \,. \tag{2.9}$$

The identity $|1-z|^2 = 2[1-\cos(\omega)]$ implies that for small ω we have:

$$f(\omega) \;=\; c\omega^{-2d}\,, \qquad c \equiv \frac{\sigma^2}{2\pi}\,. \tag{2.10}$$

This approximation encompasses the two extremes of a white noise [or a finite ARMA

[6] That an AR(1) with coefficient 0.99 is in any sense "temporary" is arguable.
[7] See Chatfield (1984, Chapters 6 and 9).

process] and a random walk. For white noise, $d = 0$, and $f(\omega) = c$, while for a random walk, $d = 1$ and the spectrum is inversely proportional to ω^2. A class of processes of current interest in the statistical physics literature, called $1/f$ noise, matches fractionally integrated noise with $d = \frac{1}{2}$.

3. A Simple Macroeconomic Model with Long-Term Dependence.

Over half a century ago, Wesley Claire Mitchell (1927 p.230) wrote that "We stand to learn more about economic oscillations at large and about business cycles in particular, if we approach the problem of trends as theorists, than if we confine ourselves to strictly empirical work." Indeed, gaining insights beyond stylized facts requires guidance from theory. Models of long range dependence may provide organization and discipline in the construction of economic paradigms of growth and business cycles. They can guide future research by predicting policy effects, postulating underlying causes, and suggesting new ways to analyze and combine data. Ultimately, examining the facts serves only as a prelude. Economic understanding requires more than a consensus on the Wold representation of GNP; it demands a falsifiable model based on the tastes and technology of the actual economy.

Thus, before testing for long run dependence, we develop a simple model where aggregate output exhibits long run dependence. It presents one reason that macroeconomic data might show the particular stochastic structure for which we test. It also shows that models can restrict the fractional differencing properties of time series, so that our test holds promise for distinguishing between competing theories. Furthermore, the maximizing model presented below connects long-term dependence to central economic concepts of productivity, aggregation, and the limits of the representative agent paradigm.

3.1. A Simple Real Model.

One plausible mechanism for generating long run dependence in output, which we will mention here and not pursue, is that production shocks themselves follow a fractionally integrated process. This explanation for persistence follows that used by Kydland and Prescott (1982). In general, such an approach begs the question, but in the present case evidence from geophysical and meteorological records suggests that many economically important shocks have long run correlation properties. Mandelbrot and Wallis (1969), for example, find long run dependence in rainfall, riverflows, earthquakes and weather [measured by tree rings and sediment deposits].

A more satisfactory model explains the time series properties of data by producing them despite white noise shocks. This section develops such a model with long run dependence, using a linear quadratic version of the real business cycle model of Long and Plosser (1983) and aggregation results due to Granger (1980). In our multi-sector model the output of each industry (or island) will follow an AR(1) process. Aggregate output with N sectors will not follow an AR(1) but rather an ARMA(N,N-1). This makes dynamics with even a moderate number of sectors unmanageable. Under fairly general conditions, however, a simple fractional process will closely approximate the true ARMA specification.

Consider a model economy with many goods and a representative agent who chooses a production and consumption plan. The infinitely lived agent inhabits a linear quadratic version of the real business cycle model. The agent has a lifetime utility function of $U = \sum \beta^t u(C_t)$ where C_t is an Nx1 vector denoting period-t consumption of each of the N goods in our economy. Each period's utility function $u(C_t)$ is given by:

$$u(C_t) \quad = \quad C_t'\iota \; - \; \frac{1}{2}C_t'BC_t \tag{3.1}$$

where ι is an Nx1 vector of ones. In anticipation of the aggregation considered later, we assume B to be diagonal so that $C_t'BC_t = \sum b_{ii}C_{it}^2$. The agents face a resource constraint: total output Y_t may be either consumed or saved, thus:

$$C_t \; + \; S_t\iota \quad = \quad Y_t \tag{3.2}$$

where the i,j-th entry S_{ijt} of the NxN matrix S_t denotes the quantity of good j invested in process i at time t, and it is assumed that any good Y_{jt} may be consumed or invested. Output is determined by the random linear technology:

$$Y_t \quad = \quad AS_t \; + \; \epsilon_t \tag{3.3}$$

where ϵ_t is a (vector) random production shock whose value is realized at the beginning of period $t+1$. The matrix A consists of the input-output parameters a_{ij}. To focus on long-term dependence we restrict A's form. Thus, each sector uses only its own output as input, yielding a diagonal A matrix and allowing us to simplify notation by defining $a_i \equiv a_{ii}$. This might occur, for example, with a number of distinct islands producing different goods. To further simplify the problem, all commodities are perishable and capital depreciates at a rate of 100 percent. Since the state of the economy in each period is fully specified by that period's output and productivity shock, it is useful to denote that vector $Z_t \equiv [Y_t' \;\; \epsilon_t']'$.

Subject to the production function and the resource constraints (3.2) and (3.3), the agent maximizes expected lifetime utility:

$$\underset{\{S_t\}}{\text{Max}}\, E[U|Z_t] \;=\; \underset{\{S_t\}}{\text{Max}}\, E\left[\, \sum_{\tau=t}^{\infty} \beta^{\tau-t} u(Y_t - S_t\iota) \,\middle|\, Z_t \right] \tag{3.4}$$

where we have substituted for consumption in (3.4) using the budget equation (3.2). This maps naturally into a dynamic programming formulation, with a value function $V(Z_t)$ and optimality equation:

$$V(Z_t) \;=\; \underset{\{S_t\}}{\text{Max}} \left\{ u(Y_t - S_t\iota) \;+\; \beta E[V(Z_{t+1})|Z_t] \right\}. \tag{3.5}$$

With quadratic utility and linear production, it is straightforward to discover and verify the form of $V(Z_t)$:

$$V(Y,\epsilon) \;=\; q'Y \;+\; Y'PY \;+\; R \;+\; E[\epsilon' T \epsilon] \tag{3.6}$$

where q and R denote Nx1 vectors and P and T are NxN matrices, whose entries are fixed constants given by the matrix Riccati equation that results from the recursive definition of the value function.[8] Given the value function, the first order conditions of the optimality equation (3.5) yield the chosen quantities of consumption and investment/savings and, for the example presented here, have the following closed form solutions:

[8] See Sargent (1987, Chapter 1) for an excellent exposition.

$$S_{it} = \frac{b_i}{b_i - 2\beta P_i a_i^2} Y_{it} + \frac{\beta q_i a_i - 1}{b_i - 2\beta P_i a_i^2} \tag{3.7}$$

$$C_{it} = \frac{2\beta P_i a_i^2}{2\beta P_i a_i^2 - b_i} + \frac{\beta q_i a_i - 1}{2\beta P_i a_i^2 - b_i} \tag{3.8}$$

where:

$$P_i \equiv b_i \left[\frac{a_i - \sqrt{(1 + 4\beta) a_i^2 - 4}}{4\beta a_i} \right] \tag{3.9}$$

The simple form of the optimal consumption and investment decision rules comes from the quadratic preferences and the linear production function. Two qualitative features bear emphasizing. First, higher output today will increase both current consumption and current investment, thus increasing future output. Even with 100 percent depreciation, no durable commodities, and i.i.d. production shocks, the time-to-build feature of investment induces serial correlation. Second, the optimal choices do not depend on the uncertainty present. This certainty equivalence feature is clearly an artifact of the linear-quadratic combination.

The time series of output can now be calculated from the production function (3.1) and decision rule (3.7). Quantity dynamics then come from the difference equation:

$$Y_{it+1} = \frac{a_i b_i}{b_i - 2\beta P_i a_i^2} Y_{it} + K_i + \epsilon_{it+1} \tag{3.10}$$

or

$$Y_{it+1} = \alpha_i Y_{it} + K_i + \epsilon_{it+1} \tag{3.11}$$

where K_i is some fixed constant. The key qualitative property of quantity dynamics summarized by (3.11) is that output Y_{it} follows an AR(1) process. Higher output today implies higher output in the future. That effect dies off at a rate that depends on the parameter α_i, which in turn depends on the underlying preferences and technology.

The simple output dynamics for a single industry or island neither mimics business cycles nor exhibits long-run dependence. However, aggregate output, the sum across all sectors, will show such dependence, which we now demonstrate by applying the aggregation results of Granger (1980,1988). It is well-known that the sum of two series, X_t and Y_t, each AR(1) with independent error, is an ARMA(2,1) process. Simple induction then implies that the sum of N independent AR(1) processes with distinct parameters has an ARMA(N,N-1) representation. With over six million registered businesses in America (CEA, 1988), the dynamics can be incredibly rich, and the number of parameters unmanageably huge. The common response to this problem is to pretend that many different firms (islands) have the same AR(1) representation for output, which reduces the dimensions of the aggregate ARMA process. This "canceling of roots" requires identical autoregressive parameters. An alternative approach reduces the scope of the problem by showing that the ARMA process approximates a fractionally integrated process, and thus summarizes the many ARMA parameters in a parsimonious manner. Though we consider the case of independent sectors, dependence is easily handled.

Consider the case of N sectors, with the productivity shock for each serially uncorrelated and independent across islands. Furthermore, let the sectors differ according to the productivity coefficient a_i. This implies differences in α_i, the autoregressive parameter for sector i's output Y_{it}. One of our key results is that under some distributional assumptions on the α_i's aggregate output Y_t^a follows a fractionally integrated process, where:

$$Y_t^a \equiv \sum_{i=1}^{N} Y_{it} \,. \tag{3.12}$$

To show this, we approach this problem from the frequency domain and apply spectral methods which often simplify problems of aggregation.[9] Let $f(\omega)$ denote the spectrum [spectral density function] of a random variable, and let $z = e^{-i\omega}$. From the definition of the spectrum as the Fourier transform of the autocovariance function, the spectrum of Y_{it} is:

$$f_i(\omega) = \frac{1}{|1-\alpha_i z|^2} \frac{\sigma_i^2}{2\pi} \,. \tag{3.13}$$

[9] See Theil (1954).

Similarly, independence implies that the spectrum of Y_t^a is

$$f(\omega) = \sum_{i=1}^{N} f_i(\omega) . \tag{3.14}$$

The α_i's measure an industry's average output for given input. This attribute of the production function can be thought of as a drawing from nature, as can the variance of the productivity shocks ϵ_{it} for each sector. Thus, it makes sense to think of the a_i's as independently drawn from a distribution $G(a)$ and the α_i's drawn from $F(\alpha)$. Provided that the ϵ_{it} shocks are independent of the distribution of α_i's, the spectral density of the sum can be written as:

$$f(\omega) = \frac{N}{2\pi} E[\sigma^2] \cdot \int \frac{1}{|1 - \alpha z|^2} dF(\alpha) \tag{3.15}$$

If the distribution $F(\alpha)$ is discrete, so that it takes on m ($< N$) values, Y_t^a will be an ARMA(m,$m-1$) process. A more general distribution leads to a process no finite ARMA model can represent. To further specify the process, take a particular distribution for F, in this case a variant of the beta distribution.[10] In particular, let α^2 be distributed as beta(p,q), which yields the following density function for α:

$$dF(\alpha) = \begin{cases} \frac{2}{\beta(p,q)} \alpha^{2p-1} (1-\alpha^2)^{q-1} d\alpha & 0 \le \alpha \le 1 \\ 0 & \text{otherwise} \end{cases} \tag{3.16}$$

with $p, q > 0$.[11] Obtaining the Wold representation of the resulting process requires a little more work. First note that:

$$1/|1 - \alpha z|^2 = \frac{1}{2(1-\alpha^2)} \left[\frac{1+\alpha z}{1 - \alpha z} + \frac{1 + \alpha \bar{z}}{1 - \alpha \bar{z}} \right] \tag{3.17}$$

where $\bar{z}$ denotes the complex conjugate of z, and the terms in parentheses can be further expanded by long division. Substituting this expansion and the beta distribution (3.16) into the expression for the spectrum and simplifying [using the relation $z + \bar{z} = 2\cos(\omega)$] yields:

$$f(\omega) = \int_0^1 \left[2 + 2 \sum_{k=1}^{\infty} \alpha^k \cos(k\omega) \right] \frac{2}{\beta(p,1)} \alpha^{2p-1} (1 - \alpha^2)^{q-2} d\alpha . \tag{3.18}$$

[10] Granger (1980) conjectures that the particular distribution is not essential.
[11] For a discussion of the variety of shapes the beta distribution takes as p and q vary, see Johnson and Kotz (1970).

Then the coefficient of $\cos(k\omega)$ is:

$$\int_0^1 \frac{2\alpha^k}{\beta(p,q)} \alpha^{2p+k-1}(1-\alpha^2)^{q-2} d\alpha \ . \tag{3.19}$$

Since the spectral density is the Fourier transform of the autocovariance function, (3.19) is the k-th autocovariance of Y_t^a. Furthermore, the integral defines a beta function, so (3.19) simplifies to $\beta(p+k/2, q-1)/\beta(p,q)$. Dividing by the variance gives the autocorrelation coefficients which reduce to

$$\rho(k) = \frac{\Gamma(p+q-1)}{\Gamma(p)} \frac{\Gamma(p+\frac{k}{2})}{\Gamma(p+\frac{k}{2}+q+1)} \tag{3.20}$$

which, again using the result from Stirling's approximation $\Gamma(a+k)/\Gamma(b+k) \approx k^{a-b}$, is proportional (for large lags) to k^{1-q}. Thus, aggregate output Y_t^a follows a fractionally integrated process of order $d = 1 - \frac{q}{2}$. Furthermore, as an approximation for long lags, this does not necessarily rule out interesting correlations at higher, e.g. business cycle, frequencies. Similarly, co-movements can arise as the fractionally integrated income process may induce fractional integration in other observed time series. This has arisen from a maximizing model given tastes and technologies.[12]

In principle, all parameters of the model may be estimated, from the distribution of production function parameters to the variance of output shocks. Though to our knowledge no one has explicitly estimated the distribution of production function parameters, many people have estimated production functions across industries.[13] One of the better recent studies disaggregates to 45 industries.[14] For our purposes, the quantity closest to α_i is the value-weighted intermediate product factor share. Using a translog production function, this gives the factor share of inputs coming from industries, excluding labor and capital. These range from a low of 0.07 in radio and TV advertising to a high of 0.811 in petroleum and coal products. Thus, even a small amount of disaggregation reveals a large dispersion and suggests the plausibility and significance of the simple model presented in this section.

3.2. Fiscal Policy and Welfare Implications.

Taking a policy perspective raises two natural questions about the fractional properties of national income. First, will fiscal or monetary policy change the degree of long-term

[12] Two additional points are worth emphasizing. First, the beta distribution need not be over (0,1) to obtain these results, only over $(a,1)$. Second, it is indeed possible to vary the α_i's so that α_i has a beta distribution.

[13] Leontief, in his classic study (1976) reports own-industry output coefficients for 10 sectors: how much an extra unit of food will increase food production. These vary from 0.06 (fuel) to 1.24 (other industries).

[14] Jorgenson, Gollop and Fraumeni (1987).

dependence? Friedman and Schwartz (1982), for example, point out that long run income cycles correlate with long run monetary cycles. Secondly, does long-term dependence have welfare implications? Do agents care that they live in such a world?

In the basic Ramsey-Solow growth model, as in its stochastic extensions, taxes affect levels of output and capital but not growth rates,[15] so tax policy would not affect the fractional properties. However, two alternative approaches suggest richer possibilities. Recall that fractional noise arises by aggregating many autoregressive processes. Fiscal policy may not change the coefficients of each process, but a tax policy can alter the distribution of total output across individuals, effectively changing the fractional properties of the aggregate. Secondly, endogenous growth models often allow tax policy to affect growth rates[16] : the tax reduces investment in research and future growth. Hence, the autoregressive parameters of individual firm's output could change with policy, and with them aggregate income.

Unfortunately, implementing either approach with even a modicum of realism would be quite complicated. In the dynamic stochastic growth model, taxation drives a wedge between private and social returns, resulting in a suboptimal equilibrium. This eliminates methods which exploit the Pareto-optimality of competitive equilibrium, such as dynamic programming. Characterizing solutions requires simulation methods because no closed forms have been found.[17] Thus, it seems clear that fiscal policy can impact upon fractional properties, but also that explicitly calculating the impact would take this paper too far afield and should best be left for future research.

People who predict output or forecast sales will care about the fractional nature of output, but fractional processes can have normative implications as well. Following Lucas (1987), this section estimates the welfare costs of economic instability under different regimes. We can decide if people care whether their world is fractional. For concreteness, let the typical household consume C_t, evaluating this via a utility function:

$$U = E\left[\sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\sigma}}{1-\sigma}\right] = \sum_{t=0}^{\infty} \frac{\beta^t}{1-\sigma} E[C_t^{1-\sigma}] \,. \tag{3.21}$$

Also assume:

$$\ln C_t = (1+\lambda) \sum_{k=0}^{\infty} \phi_k L^k \eta_t \tag{3.22}$$

[15] See Atkinson and Stiglitz (1980).
[16] For example, Romer (1986) and King, Plosser and Rebelo (1987).
[17] See King, Plosser and Rebelo (1987) and Baxter (1988).

where $\eta_t \equiv \ln \epsilon_t$. The λ term measures compensation for variations in the process $\phi(L)$. With η_t normally distributed with mean 0 and variance 1, the compensating fraction λ between two processes ϕ and ψ is:

$$1+\lambda \quad = \quad \exp\left[\frac{1}{2}(1-\sigma)\sum_{k=0}^{\infty}(\psi_k^2-\phi_k^2)\right] . \tag{3.23}$$

Evaluating this using a realistic $\sigma = 5$, again comparing an AR(1) with $\rho = 0.9$ and fractional process of order 1/4, we find that $\lambda = -0.99996$ [this number looks larger than those in Lucas because the process is in logs rather than in levels].[18] For comparison, this is the difference between an AR(1) with ρ of 0.90 and one with ρ of 0.95. This calculation provides a rough comparison only. When feasible, welfare calculations should use the model generating the processes, as only it will correctly account for important specifics, such as labor supply or distortionary taxation.

[18] We calculate this using (2.7) and the Hardy-Littlewood approximation for the resulting Riemann Zeta Function, following Titchmarsh, 1951, sec. 4.11.

4. R/S Analysis of Real Output.

The results of Section 3 show that simple aggregation may be one source of long–term dependence in the business cycle. In this section we employ a method for detecting long memory and apply it to real GNP. The technique is based on a simple generalization of a statistic first proposed by the English hydrologist Harold Edwin Hurst (1951), which has subsequently been refined by Mandelbrot (1972, 1975) and others.[19] Our generalization of Mandelbrot's statistic [called the "rescaled range" or "range over standard deviation" or R/S] enables us to distinguish between short and long run dependence, in a sense to be made precise below.

We define our notions of short and long memory and present the test statistic in Section 4.1. In Section 4.2 we present the empirical results for real GNP Section 4.3; we find long-term dependence in log-linearly detrended output, but considerably less dependence in the growth rates. To interpret these results, we perform several Monte Carlo experiments under two null and two alternative hypotheses and report these results in Section 4.3.

4.1. The Rescaled Range Statistic.

To develop a method of detecting long memory, we must be precise about the distinction between long-term and short-term statistical dependence. One of the most widely used concepts of short-term dependence is the notion of "strong-mixing" due to Rosenblatt (1956), which is a measure of the decline in statistical dependence of two events separated by successively longer spans of time. Heuristically, a time series is strong-mixing if the maximal dependence between any two events becomes trivial as more time elapses between them. By controlling the rate at which the dependence between future events and those of the distant past declines, it is possible to extend the usual laws of large numbers and central limit theorems to dependent sequences of random variables. Such mixing conditions have been used extensively by White (1982), White and Domowitz (1984), and Phillips (1987) for example, to relax the assumptions that ensure consistency and asymptotic normality of various econometric estimators. We adopt this notion of short-term dependence as part of our null hypothesis. As Phillips (1987) observes, these conditions are satisfied by a great many stochastic processes, including all Gaussian finite-order stationary ARMA models. Moreover, the inclusion of a moment condition also allows for heterogeneously distributed

[19] See Mandelbrot and Taqqu (1979) and Mandelbrot and Wallis (1968, 1969a–c).

sequences [such as those exhibiting heteroscedasticity], an especially important extension in view of the non-stationarities of real GNP.

In contrast to the "short memory" of weakly dependent [i.e., strong-mixing] processes, natural phenomena often display long-term memory in the form of non-periodic cycles. This has lead several authors, most notably Mandelbrot, to develop stochastic models that exhibit dependence even over very long time spans. The fractionally integrated time series models of Mandelbrot and Van Ness (1968), Granger and Joyeux (1980), and Hosking (1981) are examples of these. Operationally, such models possess autocorrelation functions that decay at much slower rates than those of weakly dependent processes, and violate the conditions of strong-mixing. To detect long-term dependence [also called "strong dependence"], Mandelbrot suggests using the range over standard deviation (R/S) statistic, also called the "rescaled range," which was developed by Hurst (1951) in his studies of river discharges. The R/S statistic is the range of partial sums of deviations of a time series from its mean, rescaled by its standard deviation. In several seminal papers, Mandelbrot demonstrates the superiority of R/S to more conventional methods of determining long-run dependence [such as autocorrelation analysis and spectral analysis].[20]

In testing for long memory in output, we employ a modification of the R/S statistic that is robust to weak dependence. In Lo (1988), a formal sampling theory for the statistic is obtained by deriving its limiting distribution analytically using a functional central limit theorem.[21] We use this statistic and its asymptotic distribution for inference below.

Let X_t denote the first-difference of log-GNP; we assume that:

$$X_t = \mu + \epsilon_t \tag{4.1}$$

where μ is an arbitrary but fixed parameter. Whether or not X_t exhibits long-term memory depends on the properties of $\{\epsilon_t\}$. As our null hypothesis H, we assume that the sequence of disturbances $\{\epsilon_t\}$ satisfies the following conditions:

[20] See Mandelbrot (1972, 1975), Mandelbrot and Taqqu (1979), and Mandelbrot and Wallis (1968, 1969a–c).

[21] This statistic is asymptotically equivalent to Mandelbrot's under independently and identically distributed observations, however Lo (1988) shows that the original R/S statistic may be significantly biased toward rejection when the time series is short-term dependent. Although aware of this bias, Mandelbrot (1972, 1975) did not correct for it since his focus was on the relation of the R/S statistic's logarithm to the logarithm of the sample size, which involves no statistical inference; such a relation clearly is unaffected by short-term dependence.

$(A1) \quad E[\epsilon_t] = 0$ for all t.

$(A2) \quad \sup_t E[|\epsilon_t|^\beta] < \infty$ for some $\beta > 2$.

$(A3) \quad \sigma^2 = \lim_{n\to\infty} E\left[\frac{1}{n}\left(\sum_{j=1}^{n} \epsilon_j\right)^2\right]$ exists and $\sigma^2 > 0$.

$(A4) \quad \{\epsilon_t\}$ is strong-mixing with mixing coefficients α_k that satisfy:[22]

$$\sum_{k=1}^{\infty} \alpha_k^{1-\frac{2}{\beta}} < \infty .$$

Condition (A1) is standard. Conditions $(A2)$ through $(A4)$ are restrictions on the maximal degree of dependence and heterogeneity allowable while still permitting some form of the law of large numbers and the [functional] central limit theorem to obtain. Note that we have not assumed stationarity. Although condition $(A2)$ rules out infinite variance marginal distributions of ϵ_t such as those in the stable family with characteristic exponent less than 2, the disturbances may still exhibit leptokurtosis via time-varying conditional moments [e.g. conditional heteroscedasticity]. Moreover, since there is a trade-off between conditions $(A2)$ and $(A4)$, the uniform bound on the moments may be relaxed if the mixing coefficients decline faster than $(A4)$ requires.[23] For example, if we require to ϵ_t have finite absolute moments of all orders [corresponding to $\beta \to \infty$], then α_k must decline faster than $1/k$. However, if we restrict ϵ_t to have finite moments only up to order 4, then α_k must decline faster than $1/k^2$. These conditions are discussed at greater length by Phillips (1987), to which we refer interested readers.

Conditions $(A1) - (A4)$ are satisfied by many of the recently proposed stochastic models of persistence, such as the stationary AR(1) with a near-unit root. Although the distinction between dependence in the short versus long runs may appear to be a matter

[22] Let $\{\epsilon_t(\omega)\}$ be a stochastic process on the probability space $(\Omega, \mathcal{F}, P)$ and define:

$$\alpha(\mathcal{A}, \mathcal{B}) \equiv \sup_{\{A\in\mathcal{A}, B\in\mathcal{B}\}} |P(A\cap B) - P(A)P(B)| \qquad \mathcal{A}\subset\mathcal{F}, \mathcal{B}\subset\mathcal{F}$$

The quantity $\alpha(\mathcal{A},\mathcal{B})$ is a measure of the dependence between the two σ-fields $\mathcal{A}$ and $\mathcal{B}$ in $\mathcal{F}$. Denote by $\mathcal{B}_s^t$ the Borel σ-field generated by $\{\epsilon_s(\omega), \ldots, \epsilon_t(\omega)\}$, i.e., $\mathcal{B}_s^t \equiv \sigma(\epsilon_s(\omega), \ldots, \epsilon_t(\omega)) \subset \mathcal{F}$. Define the coefficients α_k as:

$$\alpha_k \equiv \sup_j \alpha(\mathcal{B}_{-\infty}^j, \mathcal{B}_{j+k}^{\infty}) .$$

Then $\{\epsilon_t(\omega)\}$ is said to be strong-mixing if $\lim_{k\to\infty} \alpha_k = 0$. For further details, see Rosenblatt (1956), White (1984), and the papers in Eberlein and Taqqu (1986).

[23] See Herrndorf (1985). Note that one of Mandelbrot's (1972) arguments in favor of R/S analysis is that finite second moments are not required. This is indeed the case if we are interested only in the almost sure convergence of the statistic. However, since we wish to derive its *limiting distribution* for purposes of inference, a stronger moment condition is needed.

of degree, strongly dependent processes behave so differently from weakly dependent time series that our dichotomy seems most natural. For example, the spectral densities of strongly dependent processes are either unbounded or zero at frequency zero. Their partial sums do not converge in distribution at the same rate as weakly dependent series. And graphically, their behavior is marked by cyclic patterns of all kinds, some that are virtually indistinguishable from trends.[24]

To construct the modified R/S statistic, consider a sample $X_1, X_2, \ldots, X_n$ and let $\bar{X}_n$ denote the sample mean $\frac{1}{n}\sum_j X_j$. Then the modified re-scaled range statistic, which we shall call Q_n, is given by:

$$Q_n \equiv \frac{1}{\hat{\sigma}_n(q)}\left[\underset{1\le k\le n}{\text{Max}} \sum_{j=1}^{k}\left(X_j - \bar{X}_n\right) - \underset{1\le k\le n}{\text{Min}} \sum_{j=1}^{k}\left(X_j - \bar{X}_n\right)\right] \tag{4.2}$$

where

$$\begin{aligned}\hat{\sigma}_n^2(q) &\equiv \frac{1}{n}\sum_{j=1}^{n}(X_j - \bar{X}_n)^2 + \frac{2}{n}\sum_{j=1}^{q}\omega_j(q)\left\{\sum_{i=j+1}^{n}(X_i - \bar{X}_n)(X_{i-j} - \bar{X}_n)\right\} \\ &= \hat{\sigma}_x^2 + 2\sum_{j=1}^{q}\omega_j(q)\hat{\gamma}_j \qquad \omega_j(q) \equiv 1 - \frac{j}{q+1} \qquad q < n.\end{aligned} \tag{4.3}$$

and $\hat{\sigma}_x^2$ and $\hat{\gamma}_j$ are the usual sample variance and autocovariance estimators of X. Q_n is the range of partial sums of deviations of X_j from its mean, normalized by an estimator of the partial sum's standard deviation divided by n. The estimator $\hat{\sigma}_n(q)$ involves not only sums of squared deviations of X_j, but also its weighted autocovariances up to lag q; the weights $\omega_j(q)$ are those suggested by Newey and West (1987), and always yield a positive estimator $\hat{\sigma}_n^2(q)$.[25] Theorem 4.2 of Phillips (1987) demonstrates the consistency of $\hat{\sigma}_n(q)$ under the following conditions:

[24] See Mandelbrot (1972) for further details.
[25] $\hat{\sigma}_n^2(q)$ is also an estimator of the spectral density function of X_t at frequency zero, using a Bartlett window.

$(A2')$ $\sup_t E[|\epsilon_t|^{2\beta}] < \infty$ for some $\beta > 2$.

$(A5)$ As n increases without bound, q also increases without bound such that $q \sim o(n^{1/4})$.

The choice of the truncation lag q is a delicate matter. Although q must increase with [but at a slower rate than] the sample size, Monte Carlo evidence suggests that when q becomes large relative to the number of observations, asymptotic approximations may fail dramatically.[26] However, q cannot be chosen too small otherwise the effects of higher-order autocorrelations may not be captured. The choice of q is clearly an empirical issue and must therefore be chosen with some consideration of the data at hand.

If the observations are independently and identically distributed with variance σ_ϵ^2, our normalization by $\hat{\sigma}_n(q)$ is asymptotically equivalent to normalizing by the usual standard deviation estimator $s_n = [\frac{1}{n}\sum_j (X_j - \bar{X}_n)^2]^{1/2}$. The resulting statistic, which we call $\tilde{Q}_n$, is precisely the one proposed by Hurst (1951) and Mandelbrot (1972):

$$\tilde{Q}_n \equiv \frac{1}{s_n}\left[\operatorname*{Max}_{1\le k\le n} \sum_{j=1}^{k}\left(X_j - \bar{X}_n\right) - \operatorname*{Min}_{1\le k\le n} \sum_{j=1}^{k}\left(X_j - \bar{X}_n\right)\right]. \tag{4.4}$$

To perform statistical inference with the standardized re-scaled range $Q_n/\sqrt{n}$, we require its distribution. Although its finite-sample distribution is not apparent, a large-sample approximation may be obtained.[27] To derive the limiting distribution of the $Q_n/\sqrt{n}$, we consider the behavior of the following standardized sum:

$$W_n(\tau) \equiv \frac{1}{\sigma\sqrt{n}} S_{[n\tau]} \qquad \tau \in [0,1] \tag{4.5}$$

where S_k denotes the partial sum $\sum_{j=1}^{k} \epsilon_j$ and $[n\tau]$ is the greatest integer less than or equal to $n\tau$. The sample paths of $W_n(\tau)$ are elements of the function space $\mathcal{D}[0,1]$, the space of all real-valued functions on $[0,1]$ that are right-continuous and possess finite left limits. Under certain conditions, it may be shown that $W_n(\tau)$ converges weakly to a Brownian motion $W(\tau)$ on the unit interval, and that well-behaved functionals of $W_n(\tau)$ converge

[26] See, for example, Lo and MacKinlay (1988).

[27] This is derived in Lo (1988); for expositional completeness, we restate those results here without proof.

weakly to the same functionals of Brownian motion.[28] Following convention, we use the symbol '$\Rightarrow$' to denote weak convergence. Since we shall make extensive use of these two results in deriving the limiting distribution of Q_n, we state them here for reference:

Lemma 4.1. *[Herrndorf (1984)] If $\{\epsilon_t\}$ satisfies assumptions (A1)–(A4) then as n increases without bound, $W_n(\tau) \Rightarrow W(\tau)$.*

Lemma 4.2. *[Continuous Mapping Theorem] If $W_n(\tau) \Rightarrow W(\tau)$ and h is a functional on $\mathcal{D}[0,1]$ that is continuous except on a set $D_h \subset \mathcal{D}[0,1]$ of Wiener-measure zero, i.e. $P(W \in D_h) = 0$, then $h(W_n) \Rightarrow h(W)$.*

Using these results, we may derive the limiting distribution of the modified rescaled range in four easy steps, which are summarized in:[29]

Theorem 4.1. *If $\{\epsilon_t\}$ satisfies assumptions $(A1)$, $(A2')$, $(A3)-(A5)$, then as n increases without bound we have:*

$$(a) \qquad \frac{1}{\hat{\sigma}_n(q)\sqrt{n}} \sum_{j=1}^{k} \left(X_j - \bar{X}_n\right) \;\Rightarrow\; W(\tau) \;-\; \tau W(1) \;\equiv\; W^{\circ}(\tau)\;.$$

$$(b) \qquad \underset{1\leq k\leq n}{\mathrm{Max}}\, \frac{1}{\hat{\sigma}_n(q)\sqrt{n}} \sum_{j=1}^{k} \left(X_j - \bar{X}_n\right) \;\Rightarrow\; \underset{0\leq\tau\leq 1}{\mathrm{Max}}\; W^{\circ}(\tau) \;\equiv\; M^{\circ}$$

$$(c) \qquad \underset{1\leq k\leq n}{\mathrm{Min}}\, \frac{1}{\hat{\sigma}_n(q)\sqrt{n}} \sum_{j=1}^{k} \left(X_j - \bar{X}_n\right) \;\Rightarrow\; \underset{0\leq\tau\leq 1}{\mathrm{Min}}\; W^{\circ}(\tau) \;\equiv\; m^{\circ}$$

$$(d) \qquad \frac{1}{\sqrt{n}} Q_n \;\Rightarrow\; M^{\circ} \;-\; m^{\circ} \;\equiv\; V\;.$$

Part (a) of Theorem 4.1 follows from Lemmas 4.1, 4.2 and Theorem 4.2 of Phillips (1987), and shows that the partial sum of deviations of X_t from its mean converges to the celebrated

[28] See Billingsley (1968, 1971) for further details.

[29] Mandelbrot (1975) derives similar limit theorems for the statistic $\tilde{Q}_n$ under the more restrictive i.i.d. assumption, in which case the limiting distribution will coincide with that of Q_n. Since we wish to expand our null hypothesis to include weakly dependent disturbances, we extend his results via the more general functional central limit theorem of Herrndorf (1984, 1985). Note, however, that Mandelbrot (1975) also derives limit theorems for $\tilde{Q}_n$ when the sequence $\{\epsilon_t\}$ is strongly dependent (e.g. fractional Brownian motion, etc.).

Brownian bridge $W^\circ(\tau)$ on the unit interval, also called "pinned" or "tied-down" Brownian motion because $W^\circ(0) = W^\circ(1) = 0$. That the limit is a Brownian bridge is not unexpected since the terms in the partial sum are deviations from the mean and must therefore sum to zero at $k = n$. Parts (*b*)–(*d*) of the theorem follow immediately by applying Lemma 4.2 to (*a*). Part (*d*) is the key result, as it allows us to perform large sample statistical inference once we obtain the distribution function for the range of the Brownian bridge. Since the joint distribution of the maximum M° and minimum m° of the Brownian bridge is well-known [see Billingsley (1968)], the distribution function of their difference may be readily obtained:[30]

Theorem 4.2. *The distribution and density functions of* $V \equiv M^\circ - \mathrm{m}^\circ$, *respectively* $F_V(v)$ *and* $f_V(v)$, *are given by:*

$$F_V(v) \;=\; \sqrt{2\pi} \sum_{k=-\infty}^{\infty} \left\{ k[1-\alpha_k^2]\phi(\alpha_k) \;-\; k\phi(\beta_k) \right\} \tag{4.6}$$

$$f_V(v) \;=\; 2\sqrt{2\pi} \sum_{k=-\infty}^{\infty} \left\{ k[\alpha_k^3 - (k+2)\alpha_k]\phi(\alpha_k) \;+\; k(k+1)\beta_k\phi(\beta_k) \right\} \tag{4.7}$$

$$\alpha_k \;\equiv\; 2kv \qquad \beta_k \;\equiv\; 2(k+1)v \qquad \phi(x) \;\equiv\; \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \; .$$

Using F_V, critical values may readily be calculated for tests of any significance level. The most commonly used values are reported in Table 1. The moments of V are also easily computed using f_V; it is straightforward to show that $E[V] = \sqrt{\frac{\pi}{2}}$ and $E[V^2] = \frac{\pi^2}{6}$, thus the mean and standard deviation of V are approximately 1.25 and 0.27 respectively. The distribution and density functions are plotted in Figure 3. Observe that the distribution is positively skewed and most of its mass falls between $\frac{3}{4}$ and 2.

[30] Feller (1951) obtains similar results under the more restrictive assumption of i.i.d. disturbances.

4.2. Empirical Results for Real Output.

We apply our test to two time series of real output: quarterly postwar real GNP from 1947:1 to 1987:4, and the annual Friedman and Schwartz (1982) series from 1869 to 1972. These results are reported in Table 2. The first row of numerical entries are estimates of the classical rescaled range $\tilde{V}_n$ which is not robust to short-term dependence. The next eight rows are estimates of the modified rescaled range $V_n(q)$ for values of q from 1 to 8. Recall that q is the truncation lag of the estimator of the spectral density at frequency zero. Reported in parentheses below the entries for $V_n(q)$ are estimates of the percentage bias of the statistic $\tilde{V}_n$, and is computed as $100 \cdot [(\tilde{V}_n / V_n(q)) - 1]$.

The first column of numerical entries in Table 2 indicate that the null hypothesis of short-term dependence for the first-difference of log-GNP cannot be rejected for any value of q. The classical rescaled range statistic also supports the null hypothesis. The results for the Friedman and Schwartz series are similar. When we log-linearly detrend real GNP, the results differ considerably. The third column of numerical entries in Table 2 show that short-term dependence may be rejected for log-linearly detrended quarterly output with values of q from 1 to 4. That the rejections are weaker for larger q is not surprising since additional noise arises from estimating higher-order autocorrelations. When values of q beyond 4 are used, we no longer reject the null hypothesis at the 5 percent level of significance. Finally, using the Friedman and Schwartz time series, we only reject with the classical rescaled range and with $V_n(1)$.

Taken together, these results confirm the unit root findings of Campbell and Mankiw (1987), Nelson and Plosser (1982), Perron and Phillips (1987), and Stock and Watson (1986). That there are more significant autocorrelations in log-linearly detrended GNP is precisely the spurious periodicity suggested by Nelson and Kang (1981). Moreover, the trend plus stationary noise model of GNP is not contained in our null hypothesis, hence our failure to reject the null hypothesis is also consistent with the unit root model.[31] To see this, observe that if log-GNP y_t were trend stationary, i.e.:

$$y_t \;=\; \alpha \;+\; \beta t \;+\; \eta_t \tag{4.8}$$

where η_t is stationary white noise, then its first-difference X_t is simply $X_t = \beta + \epsilon_t$ where $\epsilon_t \equiv \eta_t - \eta_{t-1}$. But this innovations process violates our assumption $(A3)$ and is therefore not contained in our null hypothesis.

[31] Of course, this may be the result of low power against stationary but near-integrated processes, and must be addressed by Monte Carlo experiments.

To conclude that the data support the null hypothesis because our statistic fails to reject it is, of course, premature since the size and power of our test in finite samples is yet to be determined. We perform illustrative Monte Carlo experiments and report the results in the next section.

4.3. The Size and Power of the Test.

To evaluate the size and power of our test in finite samples, we perform several illustrative Monte Carlo experiments for a sample size of 163 observations, corresponding to the number of quarterly observations of real GNP growth from 1947:2 to 1987:4.[32] We simulate two null hypotheses: independently and identically distributed increments, and increments that follow an ARMA(2,2) process. Under the i.i.d. null hypothesis, we fix the mean and standard deviation of our random deviates to match the sample mean and standard deviation of our quarterly data set: 7.9775×10^{-3} and 1.0937×10^{-3} respectively. To choose parameter values for the ARMA(2,2) simulation, we estimate the model:

$$(1 - \phi_1 L - \phi_2 L^2) y_t \;=\; \mu \;+\; (1 + \theta_1 L + \theta_2 L^2)\epsilon_t \qquad \epsilon_t \sim WN(0, \sigma_\epsilon^2) \tag{4.9}$$

using nonlinear least squares. The parameter estimates are [with standard errors in parentheses]:

$$\begin{aligned}
\phi_1 &= 0.5837 & \theta_1 &= -0.2825 \\
&\quad (0.1949) & &\quad (0.1736) \\
\phi_2 &= -0.4844 & \theta_2 &= 0.6518 \\
&\quad (0.1623) & &\quad (0.1162) \\
\mu &= 0.0072 & & \\
&\quad (0.0016) & & \\
\hat{\sigma}_\epsilon^2 &= 0.0102 & &
\end{aligned}$$

Table 3 reports the results of both null simulations.

[32] All simulations were performed in double precision on a VAX 8700 using the IMSL 10.0 random number generator DRNNOA; each experiment was comprised of 10,000 replications.

It is apparent from the "I.I.D. Null" Panel of Table 3 that the 5 percent test based on the classical rescaled range rejects too frequently. The 5 percent test using the modified rescaled range with $q = 3$ rejects 4.6 percent of the time, closer to its nominal size. As the number of lags increases to 8, the test becomes more conservative. Under the ARMA(2,2) null hypothesis, it is apparent that modifying the rescaled range by the spectral density estimator $\hat{\sigma}_n^2(q)$ is critical; the size of a 5 percent test based on the classical rescaled range is 34 percent, whereas the corresponding size using the modified R/S statistic with $q = 5$ is 4.8 percent. As before, the test becomes more conservative when q is increased.

Table 3 also reports the size of tests using the modified rescaled range when the lag length q is chosen optimally using Andrews' (1987) procedure. This data-dependent procedure entails computing the first-order autocorrelation coefficient $\hat{\rho}(1)$ and then setting the lag length to be the integer-value of $\bar{M}_n$, where:[33]

$$\bar{M}_n \equiv \left(\frac{3\hat{\alpha}n}{2}\right)^{1/3} \qquad \hat{\alpha} \equiv \frac{4\hat{\rho}^2}{(1-\hat{\rho}^2)^2}\,. \tag{4.10}$$

Under the i.i.d. null, Andrews' formula yields a 5 percent test with empirical size 6.9 percent; under the ARMA(2,2) alternative, the corresponding size is 4.1 percent. Although significantly different from the nominal value, the empirical size of tests based on Andrews' formula may not be economically important. In addition to its optimality properties, the procedure has the advantage of eliminating a dimension of arbitrariness in performing the test.

Table 4 reports power simulations under two fractionally differenced alternatives: $(1-L)^d\epsilon_t = \eta_t$ where $d = 1/3, -1/3$. Hosking (1981) has shown that the autocovariance function $\gamma_\epsilon(k)$ of ϵ_t is given by:

$$\gamma_\epsilon(k) = \frac{\Gamma(1-2d)\Gamma(d+k)}{\Gamma(d)\Gamma(1-d)\Gamma(1-d+k)}\sigma_\eta^2 \qquad d \in (-\frac{1}{2}, \frac{1}{2}) \tag{4.11}$$

Realizations of fractionally differenced time series [of length 163] are simulated by pre-multiplying vectors of independent standard normal random variates by the Cholesky-factorization of the 163 x 163 covariance matrix whose entries are given by (4.11). To calibrate the simulations, σ_η^2 is chosen to yield unit variance ϵ_t's, the $\{\epsilon_t\}$ series is then multiplied by the sample standard deviation of real GNP growth from 1947:1 to 1987:4, and to this series is added the sample mean of real GNP growth over the same sample

[33] In addition, Andrews' procedure requires weighting the autocovariances by $1 - \frac{j}{\bar{M}_n}$ $(j = 1, \ldots, [\bar{M}_n])$ in contrast to Newey and West's (1987) $1 - \frac{j}{q+1}$ $(j = 1, \ldots, q)$, where q is an integer and $\bar{M}_n$ need not be.

period. The resulting time series is used to compute the power of the rescaled range; Table 4 reports the results.

For small values of q, tests based on the modified rescaled range have reasonable power against both fractionally differenced alternatives. For example, using one lag the 5 percent test against the $d = 1/3$ alternative has 58.7 percent power; against the $d = -1/3$ alternative this test has 81.1 percent power. As the lag length is increased, the test's power declines. Note that tests based on the classical rescaled range is significantly more powerful than those using the modified R/S statistic. This, however, is of little value when distinguishing between long-term versus short-term dependence since the test using the classical statistic also has power against some stationary finite-order ARMA processes. Finally, note that tests using Andrews' truncation lag formula have reasonable power against the $d = -1/3$ alternative but are considerably weaker against the more relevant $d = 1/3$ alternative.

The simulation evidence in Tables 3 and 4 suggest that our empirical results do indeed support the short-term dependence of GNP with a unit root. Our failure to reject the null hypothesis does not seem to be explicable by a lack of power against long-memory alternatives. Of course, our simulations were illustrative and by no means exhaustive; additional Monte Carlo experiments must be performed before a full assessment of the test's size and power is complete. Nevertheless our modest simulations indicate that there is little empirical evidence in favor of long-term memory in GNP growth rates. Perhaps the direct estimation of long-memory models would yield stronger results and is currently being investigated by several authors.[34]

5. Conclusion.

This paper has suggested a new approach to the stochastic structure of aggregate output. Traditional dissatisfaction with the conventional methods – from observations about the typical spectral shape of economic time series, to the discovery of cycles at all periods – calls for such a reformulation. Indeed, recent controversy over deterministic versus stochastic trends and the persistence of shocks underscores the difficulties even modern methods have of identifying the long run properties of the data.

[34] See, for example, Diebold and Rudebusch (1988), Sowell (1987a,b), and Yajima (1985,1988).

Fractionally integrated random processes provide one explicit approach to the problem of long-term dependence; naming and characterizing this aspect is the first step in studying the problem scientifically. Controlling for its presence improves our ability to isolate business cycles from trends and to assess the propriety of that decomposition. To the extent that it explains output, long-term dependence deserves study in its own right. Furthermore, Singleton (1988) has recently pointed out that dynamic macroeconomic models often link inextricably predictions about business cycles, trends, and seasonal effects. So too linked is long-term dependence: a fractionally integrated process arises quite naturally in a dynamic linear model via aggregation. This model not only predicts the existence of fractional noise, but also suggests the character of its parameters. This class of models leads to testable restrictions on the nature of long-term dependence in aggregate data, and also holds the promise of policy evaluation.

Advocating a new class of stochastic processes would be a fruitless task if its members were intractable. In fact, manipulating such processes causes few problems. We constructed an optimizing linear dynamic model that exhibits fractionally integrated noise, and provided an explicit test for such long-term dependence. Modifying a statistic of Hurst and Mandelbrot gives us a statistic robust to short-term dependence, and this modified R/S statistic possesses a well-defined limiting distribution which we have tabulated. Illustrative computer simulations indicate that this test has power against at least two specific alternative hypotheses of long-memory.

Two main conclusions arise from the empirical work and Monte Carlo experiments. First, the evidence does not support long-term dependence in GNP. Rejections of the short-term dependence null hypothesis occur only with detrended data, and is consistent with the well-known problem of spurious periodicities induced by log-linear detrending. Second, since a trend-stationary model is not contained in our null hypothesis, our failure to reject may also be viewed as supporting the first-difference stationary model of GNP, with the additional result that the resulting stationary process is weakly dependent at most. This supports and extends the conclusion of Adelman that, at least within the confines of the available data, there is little evidence of long-term dependence in the business cycle.

References

Adelman, Irma (1956): "Long Cycles: Fact or Artefact?" *American Economic Review* 55, 444-463.

Andrews, Donald (1987): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," Working Paper, Cowles Foundation, Yale University.

Baxter, Marianne (1988): "Approximating Suboptimal Dynamic Equilibria: An Euler Equation Approach" Working Paper, University or Rochester.

Beveridge, Stephen, and Charles R. Nelson (1981): "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'," *Journal of Monetary Economics* 7, 151-174.

Billingsley, Patrick (1968): *Convergence of Probability Measures.* New York: John Wiley and Sons.

———, (1971): *Weak Convergence of Measures: Applications in Probability*, CBMS–NSF Regional Conference Series in Applied Mathematics. S.I.A.M.: Philadelphia.

Campbell, John Y., and N. Gregory Mankiw (1987): "Are Output Fluctuations Transitory?" *Quarterly Journal of Economics* 102, 857-880.

Chatfield, Christopher (1984): *The Analysis of Time Series: An Introduction*, Third Edition, Chapman and Hall, New York.

Cochrane, John (1988): "How Big is the Random Walk in GNP?," *Journal of Political Economy* 96, 893–920.

Council of Economic Advisors (1988): *Economic Report of the President*, U.S. Government Printing Office.

Diebold, Francis X., and Glenn D. Rudebusch (1988): " Long Memory and Persistence in Aggregate Output" Federal Reserve Board of Governors Finance and Economics Discussion Series No. 7.

Eberlein, Ernst and Murad Taqqu (1986): *Dependence in Probability and Statistics.* Progress in Probability and Statistics 11. Birkhäuser: Boston.

Erdős, Paul and Mark Kac (1946): "On Certain Limit Theorems in the Theory of Probability," *Bulletin of the American Mathematical Society* 52, 292–302.

———, (1947): "On the Number of Positive Sums of Independent Random Variables," *Bulletin of the American Mathematical Society* 53, 1011–1020.

Feller, William (1951): "The Asymptotic Distribution of the Range of Sums of Independent Random Variables," *Annals of Mathematical Statistics* 22, 427–432.

Fisher, Irving (1925): "Our Unstable Dollar and the So-called Business Cycle", *Journal of the American Statistical Association* 20, 179-202.

Friedman, Milton, and Anna J. Schwartz (1982): *Monetary Trends in The United State the United Kingdom*, NBER Monograph, University of Chicago Press, Chicago.

Geweke, John, and Susan Porter-Hudak (1983): " The Estimation and Application of Long Memory Time series Models" *Journal of Time Series Analysis* 4, 221–238.

Granger, Clive W. J. (1969): "The Typical Spectral Shape of an Economic Variable," *Econometrica* 37, 150-161.

——— (1980): "Long Memory Relations and the Aggregation of Dynamic Models" *Journal of Econometrics* 14, 227-238.

——— (1988): "Aggregation of Time Series Variables–A Survey" Federal Reserve Bank of Minneapolis Institute for Empirical Macroeconomics, Discussion paper 1.

———, and Roselyne Joyeux (1980): "An Introduction to Long-Memory Time Series Models and Fractional Differencing", *Journal of Time Series Analysis* 1,14-29.

Herrndorf, Norbert (1984): "A Functional Central Limit Theorem for Weakly Dependent Sequences of Random Variables," *Annals of Probability* 12, 141–153.

———, (1985): "A Functional Central Limit Theorem for Strongly Mixing Sequences of Random Variables," *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 541–550.

Hosking, J.R.M. (1981): "Fractional Differencing," *Biometrika* 68, 165–176.

Hurst, Harold E. (1951): "Long Term Storage Capacity of Reservoirs," *Transactions of the American Society of Civil Engineers* 116, 770–799.

Iyanaga, Shokichi, and Yukiyosi Kawada, editors (1977): *Encyclopedic Dictionary of Mathematics*, Mathematical Society of Japan, MIT Press, Cambridge.

Jorgenson, Dale W., Frank M. Gollop, and Barbara M. Fraumeni (1987): *Productivity and U.S. Economic Growth*, Harvard Economic Studies vol 159, Harvard University Press, Cambridge.

King, Robert G., Charles I. Plosser and Sergio Rebelo (1987): "Production, Growth and Business Cycles", Working Paper, University of Rochester.

Kolmogorov, Andrei N. (1940): "Wienersche Spiralen und einige andere interessante Kurven im Hilberteschen Raum," *C. R. Doklady Acad. Sci. URSS* (n.s.) 26, 115–118.

Kuznets, Simon (1965); *Economic Growth and Structure*, Norton, New York.

Kydland, Finn, and Edward C. Prescott (1982): "Time to Build and Aggregate Economic Fluctuations", *Econometrica* 50, 1345-1370.

Leontief, Wassily W. (1976): *The Structure of the American Economy 1919–1939*, second enlarged edition, International Arts and Sciences Press Inc., White Plains.

Lo, Andrew W. (1988): "Long-Term Memory in Stock Market Prices," Working Paper, Sloan School of Management, M.I.T.

——— and A. Craig MacKinlay (1988): "The Size and Power of the Variance Ratio Test in Finite Samples: A Monte Carlo Investigation," NBER Technical Working Paper No. 66, to appear in *Journal of Econometrics*.

Long, John B. Jr. and Charles I. Plosser (1983): "Real Business Cycles", *Journal of Political Economy* 91, 39-69.

Lucas, Robert E. Jr. (1987): *Models of Business Cycles*, Yrjö Jahnsson Lectures, Basil Blackwell, New York.

Mandelbrot, Benoit (1972): "Statistical Methodology for Non-Periodic Cycles: From the Covariance to R/S Analysis," *Annals of Economic and Social Measurement* 1, 259–290.

———, (1975): "Limit Theorems on the Self-Normalized Range for Weakly and Strongly Dependent Processes," *Z. Wahrscheinlichkeitstheorie verw. Gebiete* 31, 271–285.

———, and Murad Taqqu (1979): "Robust R/S Analysis of Long-Run Serial Correlation," *Bulletin of the International Statistical Institute* 48, Book 2, 59–104.

———, and John Van Ness (1968), "Fractional Brownian Motion, Fractional Noises and Applications," *S.I.A.M. Review* 10, 422–437.

———, and James Wallis (1968): "Noah, Joseph and Operational Hydrology," *Water Resources Research* 4, 909–918.

———, (1969a): "Computer Experiments with Fractional Gaussian Noises. Parts 1, 2, 3," *Water Resources Research* 5, 228–267.

———, (1969b): "Some Long Run Properties of Geophysical Records," *Water Resources Research* 5, 321–340.

———, (1969c): "Robustness of the Rescaled Range R/S in the Measurement of Noncyclic Long Run Statistical Dependence," *Water Resources Research* 5, 967–988.

McCallum, Bennett T. (1986): "On 'Real' and 'Sticky Price' Theories of the Business Cycle", *Journal of Money, Credit, and Banking* 18, 397-414.

Mitchell, Wesley Claire (1927): *Business Cycles: the problem and its setting*, NBER Studies in Business Cycles No.1 National Bureau of Economic Research, New York.

Nelson, Charles R. and Heejoon Kang (1981): "Spurious Periodicity in Inappropriately Detrended Time Series," *Econometrica* 49, 741–751.

Nelson, Charles R. and Charles I. Plosser (1982): "Trends and Random Walks in Macroeconomic Time Series: some evidence and implications", *Journal of Monetary Economics* 10, 139-162.

Newey, Whitney K. and Kenneth D. West (1987): "A Simple Positive Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* 55, 703–705.

Oldham, Keith B., and Jerome Spanier (1974): *The Fractional Calculus*, Academic Press, New York.

Perron, Pierre and Peter C. B. Phillips (1987): "Does GNP Have a Unit Root?" *Economics Letters* 23, 139–145.

Phillips, Peter C. B. (1987): "Time Series Regression with a Unit Root," *Econometrica* 55, 277–301.

Plosser, Charles I., and G. William Schwert (1978): "Money, Income, and Sunspots: Measuring Economic Relationships and the Effects of Differencing", *Journal of Monetary Economics* 4, 637–660.

Quah, Danny (1987): "What Do We Learn From Unit Roots in Macroeconomic Time Series?" NBER Working Paper No. 2450.

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

Rosenblatt, Murray (1956): "A Central Limit Theorem and a Strong Mixing Condition," *Proceedings of the National Academy of Sciences* 42, 43–47.

Sargent, Thomas J. (1987): *Dynamic Macroeconomic Theory*, Harvard University Press, Cambridge.

Singleton, Kenneth J. (1988): "Econometric Issues in the Analysis of Equilibrium Business Cycle Models," *Journal of Monetary Economics* 21, 361–386.

Slutzky, Eugene (1937): "The Summation of Random Causes as the Source of Cyclic Processes", *Econometrica* 5, 105-146.

Sowell, Fallaw. (1987a): "Fractional Unit Root Distributions," Discussion Paper No: 87–05, Institute of Statistics and Decision Sciences, Duke University.

———, (1987b): "Maximum Likelihood Estimation of Fractionally Integrated Time Series Models," working paper, Duke University.

Stock, J. and M. Watson (1986): "Does GNP Have a Unit Root?" *Economics Letters* 22, 147–151.

Theil, Henri (1954): *Linear Aggregation of Economic Relations*, North-Holland, Amsterdam.

Titchmarsh, E.C. (1951): *The Theory of the Riemann Zeta-Function*, Claredon Press, Oxford.

White, Halbert (1980): "A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity," *Econometrica* 48, 817–838.

———, (1984): *Asymptotic Theory for Econometricians*. John Wiley and Sons, New York.

———, and I. Domowitz (1984): "Nonlinear Regression with Dependent Observations," *Econometrica* 52, 143–162.

Yajima, Yoshihiro (1985): "On Estimation of Long-Memory Time Series Models," *Australian Journal of Statistics*, 303–320.

———, (1988): "On Estimation of a Regression Model with Long-Memory Stationary Errors," *Annals of Statistics* 16, 791-807.

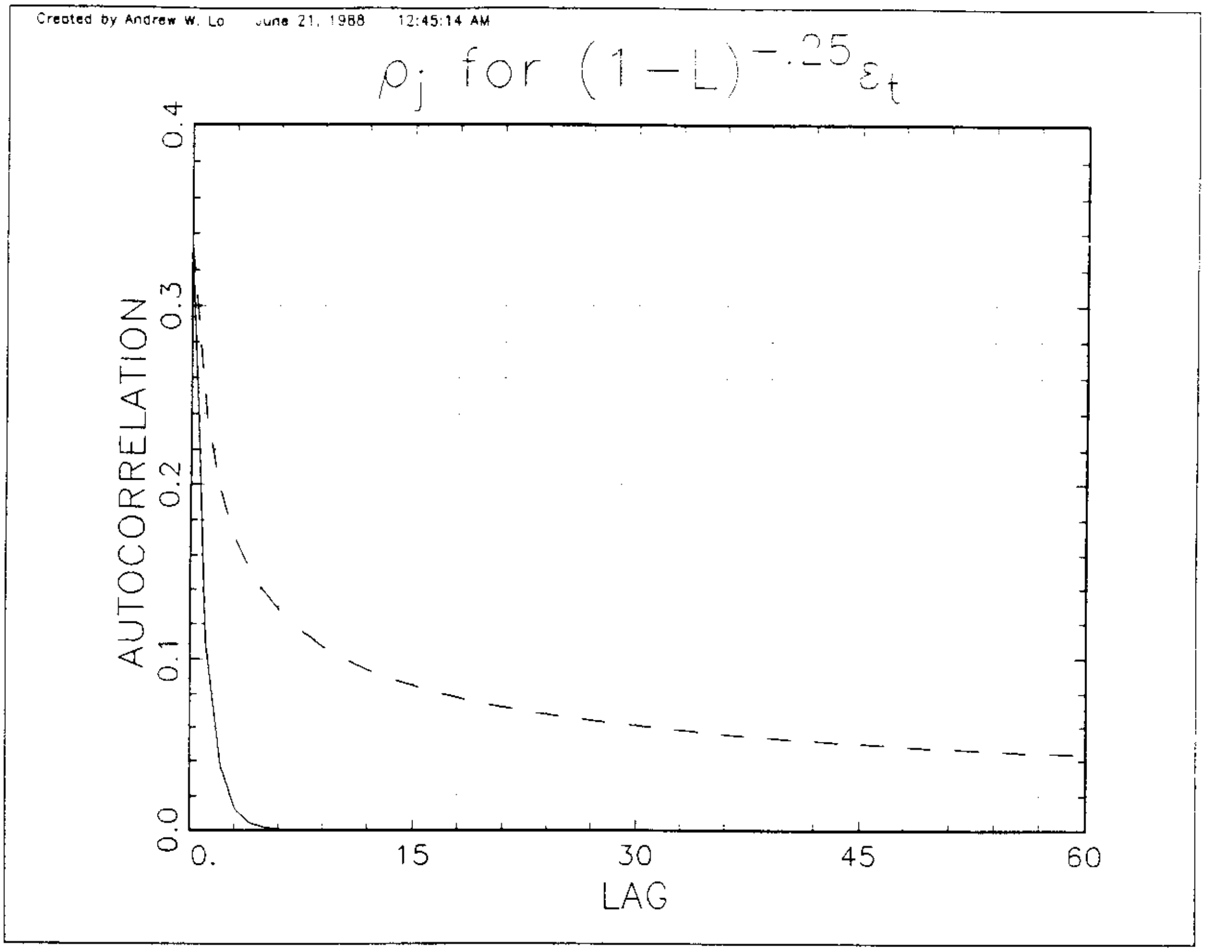

Figure 1.

Autocorrelation functions of an AR(1) with coefficient 1/3 [solid line] and a fractionally-differenced series $X_t = (1-L)^{-d}\epsilon_t$ with differencing parameter $d = 0.25$ [dashed lines]. Although both processes have a first-order autocorrelation of 1/3, the fractionally-differenced process decays much more slowly.

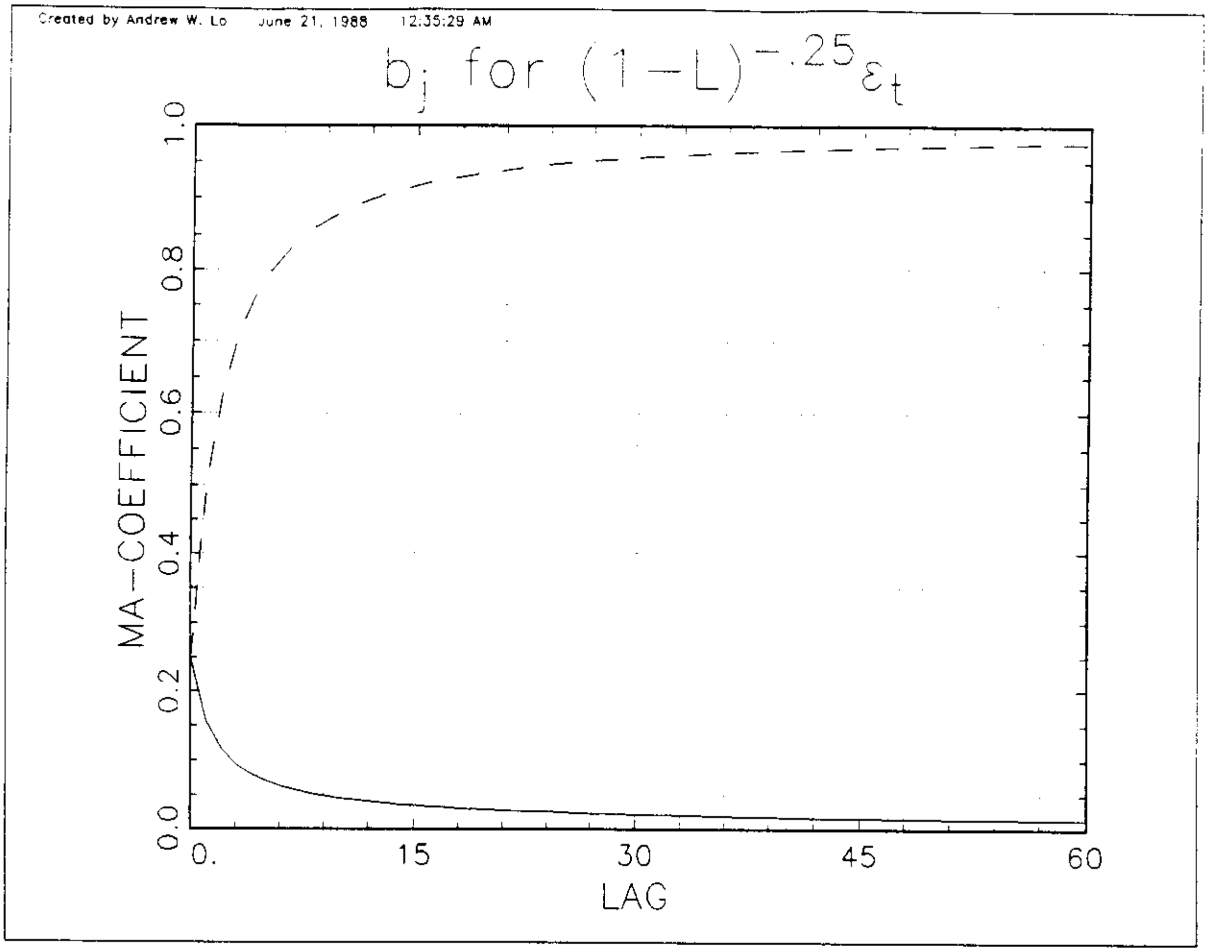

Figure 2.

Plot of the impulse response function [solid line] of the fractionally differenced time series $X_t = (1-L)^{-d}\epsilon_t$ for differencing parameter $d = 0.25$. The dashed line represents the value of an AR(1)'s autoregressive parameter required to generate the same k-th order autocorrelation as the fractionally differenced series. For large k, the autoregressive parameter must be very close to unity.

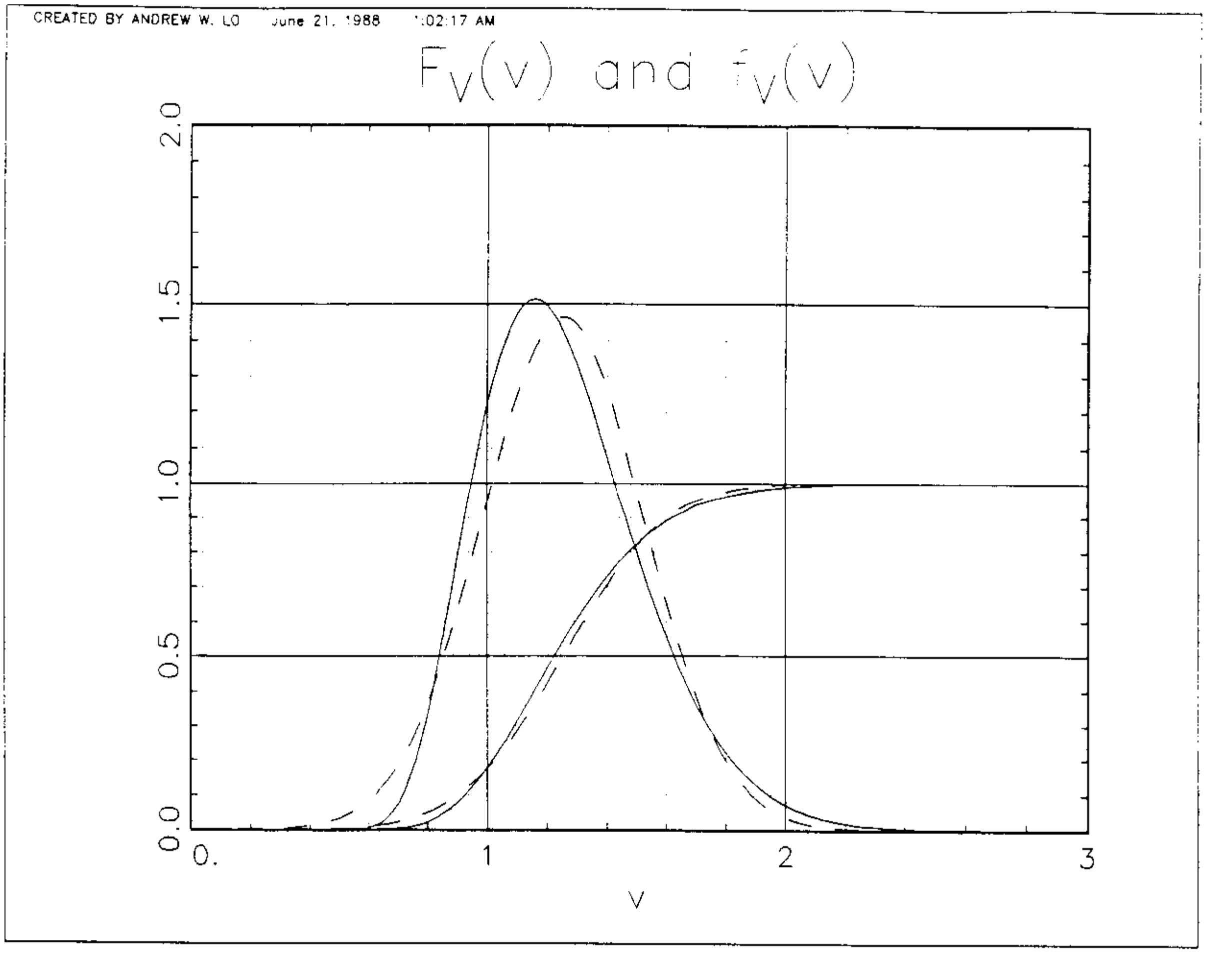

Figure 3.

Distribution and density function of the range V of a Brownian bridge. Dashed curves are the normal distribution and density functions with mean and variance equal to those of V.

Table 1a. Fractiles of the Distribution $F_V(v)$.

$P(V < v)$	.005	.025	.050	.100	.200	.300	.400	.500
v	0.721	0.809	0.861	0.927	1.018	1.090	1.157	1.223

$P(V < v)$	.543	.600	.700	.800	.900	.950	.975	.995
v	$\sqrt{\frac{\pi}{2}}$	1.294	1.374	1.473	1.620	1.747	1.862	2.098

Table 1b. Symmetric Confidence Intervals About the Mean

$P(\sqrt{\frac{\pi}{2}} - \gamma < V < \sqrt{\frac{\pi}{2}} + \gamma)$	γ
.001	0.748
.050	0.519
.100	0.432
.500	0.185

Table 2

R/S analysis of real GNP; $y^Q_{\text{Det.}}$ indicates log-linearly detrended quarterly real GNP from 1947:1 to 1987:4, and Δy^Q indicates the first-differences of the logarithm of real GNP. $y^{FS}_{\text{Det.}}$ and Δy^{FS} are defined similary for the Friedman and Schwartz series. The classical rescaled range $\tilde{V}_n$, and the modified rescaled range $V_n(q)$ are reported.[1]

	Δy^Q	Δy^{FS}	$y^Q_{\text{Det.}}$	$y^{FS}_{\text{Det.}}$
$\tilde{V}_n$	1.25	1.00	4.23*	2.83*
$V_n(1)$	1.07	0.94	3.02*	2.10*
(%–Bias)	(17.2)	(6.6)	(40.0)	(35.2)
$V_n(2)$	0.97	0.93	2.49*	1.79
(%–Bias)	(29.0)	(7.3)	(69.6)	(58.5)
$V_n(3)$	0.93	0.95	2.19*	1.62
(%–Bias)	(34.6)	(4.7)	(93.5)	(74.9)
$V_n(4)$	0.92	1.00	1.98*	1.52
(%–Bias)	(36.3)	(-0.1)	(113.5)	(86.8)
$V_n(5)$	0.92	1.07	1.83	1.45
(%–Bias)	(36.1)	(-6.4)	(130.7)	(95.7)
$V_n(6)$	0.92	1.10	1.72	1.40
(%–Bias)	(35.3)	(9.3)	(145.7)	(102.7)
$V_n(7)$	0.93	1.12	1.63	1.36
(%–Bias)	(34.4)	(-10.8)	(159.0)	(107.9)
$V_n(8)$	0.94	1.14	1.56	1.34
(%–Bias)	(33.2)	(-12.7)	(170.9)	(111.5)

[1] Under the null hypothesis H [conditions (A1), (A2'), (A3)–(A5) of the paper], the limiting distribution of $V_n(q)$ is the range of a Brownian bridge, which has a mean of $\sqrt{\pi/2}$. Fractiles are given in Table 1; the 95 percent confidence interval with equal probabilities in both tails is [0.809, 1.862]. Entries in the %-Bias columns are computed as $[(\tilde{V}_n/V_n(q))^{1/2} - 1] \cdot 100$, and are estimates of the bias of the classical R/S statistic in the presence of short-term dependence. Asterisks indicate significance at the 5 percent level.

Table 3

Finite sample distribution of the modified R/S statistic under i.i.d. and ARMA(2,2) null hypotheses for the first-difference of real log-GNP. The Monte Carlo experiments under the two null hypotheses are independent and consist of 10,000 replications each. Parameters of the i.i.d simulations were chosen to match the sample mean and variance of quarterly real GNP growth rates from 1947:1 to 1987:4; parameters of the ARMA(2,2) were chosen to match point estimates of an ARMA(2,2) model fitted to the same data set. Entries in the column labelled 'q' indicate the number of lags used to compute the R/S statistic; a lag of 0 corresponds to Mandelbrot's classical rescaled range, and a non-integer lag value corresponds to the average (across replications) lag value used according to Andrews' (1987) optimal lag formula. Standard errors for the empirical size may be computed using the usual normal approximation; they are SE(SZE01) = 9.95 x 10^-4, SE(SZE05) = 2.18 x 10^-3, SE(SZE10) = 3.00 x 10^-3.

n	q	MAX	MIN	MEAN	STD	SKEW	XKURT	PR005	PR025	PR050	PR100	PR900	PR950	PR975	PR995	SZE01	SZE05	SZE10
I.I.D. Null:																		
163	0	2.457	0.522	1.167	0.264	0.011	0.389	0.020	0.070	0.112	0.185	0.944	0.974	0.989	0.998	0.022	0.081	0.138
163	1.5	2.457	0.525	1.171	0.253	0.009	0.265	0.014	0.060	0.099	0.170	0.950	0.978	0.991	0.999	0.015	0.069	0.121
163	1	2.423	0.533	1.170	0.254	0.009	0.278	0.015	0.059	0.102	0.172	0.948	0.977	0.990	0.999	0.016	0.069	0.125
163	2	2.326	0.564	1.174	0.246	0.008	0.181	0.010	0.050	0.091	0.160	0.951	0.981	0.992	0.999	0.011	0.058	0.111
163	3	2.221	0.602	1.179	0.239	0.006	0.073	0.008	0.041	0.080	0.150	0.955	0.983	0.995	0.999	0.009	0.046	0.097
163	4	2.136	0.641	1.184	0.232	0.005	-0.019	0.006	0.032	0.068	0.137	0.959	0.987	0.996	1.000	0.006	0.036	0.082
163	5	2.087	0.645	1.189	0.225	0.005	-0.083	0.004	0.027	0.060	0.121	0.963	0.989	0.996	1.000	0.004	0.030	0.071
163	6	2.039	0.636	1.193	0.219	0.004	-0.144	0.002	0.021	0.050	0.109	0.965	0.990	0.997	1.000	0.002	0.024	0.061
163	7	1.989	0.648	1.198	0.213	0.003	-0.196	0.000	0.016	0.041	0.098	0.969	0.991	0.998	1.000	0.000	0.018	0.050
163	8	1.960	0.657	1.203	0.207	0.003	-0.256	0.000	0.014	0.033	0.086	0.972	0.993	0.999	1.000	0.000	0.015	0.040
ARMA(2,2) Null:																		
163	0	3.649	0.746	1.730	0.396	0.034	0.232	0.000	0.000	0.001	0.004	0.431	0.559	0.661	0.825	0.175	0.340	0.442
163	6.8	2.200	0.610	1.177	0.229	0.004	-0.153	0.008	0.039	0.074	0.142	0.964	0.990	0.997	1.000	0.009	0.041	0.084
163	1	3.027	0.626	1.439	0.321	0.017	0.155	0.001	0.006	0.014	0.032	0.728	0.832	0.896	0.967	0.034	0.110	0.182
163	2	2.625	0.564	1.273	0.279	0.010	0.096	0.006	0.025	0.052	0.097	0.885	0.940	0.971	0.996	0.010	0.054	0.111
163	3	2.412	0.550	1.202	0.257	0.008	0.031	0.011	0.045	0.079	0.143	0.934	0.973	0.990	0.999	0.012	0.055	0.106
163	4	2.294	0.569	1.180	0.244	0.006	-0.045	0.011	0.049	0.085	0.153	0.951	0.983	0.995	0.999	0.012	0.054	0.102
163	5	2.241	0.609	1.178	0.236	0.005	-0.121	0.010	0.045	0.081	0.146	0.958	0.987	0.997	1.000	0.010	0.048	0.093
163	6	2.181	0.616	1.180	0.229	0.004	-0.187	0.008	0.038	0.073	0.138	0.963	0.990	0.998	1.000	0.008	0.040	0.082
163	7	2.109	0.629	1.180	0.222	0.003	-0.244	0.006	0.033	0.066	0.128	0.968	0.994	0.999	1.000	0.006	0.034	0.073
163	8	2.035	0.644	1.179	0.215	0.003	-0.289	0.005	0.029	0.061	0.121	0.974	0.995	0.999	1.000	0.005	0.030	0.066

Table 4

Power of the modified R/S statistic under a Gaussian fractionally differenced alternative with differencing parameters d = 1/3, - 1/3. The Monte Carlo experiments under the two alternative hypotheses are independent and consist of 10,000 replications each. Parameters of the simulations were chosen to match the sample mean and variance of quarterly real GNP growth rates from 1947:1 to 1987:4. Entries in the column labelled 'q' indicate the number of lags used to compute the R/S statistic; a lag of 0 corresponds to Mandelbrot's classical rescaled range, and a non-integer lag value corresponds to the average (across replications) lag value used according to Andrews' (1987) optimal lag formula.

n	q	MAX	MIN	MEAN	STD	SKEW	XKURT	PR005	PR025	PR050	PR100	PR900	PR950	PR975	PR995	PWR01	PWR05	PWR10
d = 1/3:																		
163	0	4.659	0.824	2.370	0.612	0.078	-0.279	0.000	0.000	0.000	0.001	0.108	0.161	0.222	0.363	0.637	0.778	0.839
163	6.0	2.513	0.702	1.524	0.286	-0.001	-0.471	0.000	0.002	0.005	0.014	0.618	0.766	0.876	0.984	0.017	0.126	0.240
163	1	3.657	0.751	2.004	0.478	0.025	-0.421	0.000	0.001	0.001	0.003	0.233	0.321	0.413	0.584	0.416	0.587	0.680
163	2	3.140	0.721	1.811	0.409	0.010	-0.493	0.000	0.001	0.002	0.005	0.346	0.457	0.554	0.746	0.254	0.448	0.545
163	3	2.820	0.708	1.688	0.363	0.004	-0.539	0.000	0.002	0.003	0.009	0.443	0.563	0.670	0.859	0.141	0.331	0.440
163	4	2.589	0.696	1.600	0.330	0.002	-0.569	0.000	0.002	0.005	0.012	0.528	0.655	0.768	0.932	0.068	0.234	0.350
163	5	2.417	0.700	1.534	0.304	0.000	-0.589	0.000	0.002	0.007	0.016	0.597	0.736	0.845	0.974	0.027	0.158	0.271
163	6	2.297	0.700	1.482	0.282	-0.001	-0.600	0.000	0.004	0.008	0.020	0.667	0.807	0.907	0.993	0.008	0.096	0.201
163	7	2.195	0.699	1.440	0.264	-0.001	-0.602	0.000	0.004	0.009	0.023	0.727	0.868	0.948	1.000	0.001	0.056	0.141
163	8	2.107	0.694	1.405	0.249	-0.002	-0.598	0.000	0.004	0.010	0.027	0.784	0.914	0.977	1.000	0.000	0.027	0.097
d = - 1/3:																		
163	0	1.080	0.352	0.614	0.103	0.001	0.321	0.849	0.956	0.981	0.995	1.000	1.000	1.000	1.000	0.849	0.956	0.981
163	4.1	1.626	0.449	0.838	0.142	0.001	0.322	0.211	0.456	0.600	0.747	1.000	1.000	1.000	1.000	0.211	0.456	0.600
163	1	1.251	0.416	0.708	0.116	0.001	0.249	0.587	0.811	0.895	0.957	1.000	1.000	1.000	1.000	0.587	0.811	0.895
163	2	1.344	0.456	0.779	0.125	0.001	0.248	0.350	0.631	0.758	0.872	1.000	1.000	1.000	1.000	0.350	0.631	0.758
163	3	1.467	0.512	0.837	0.132	0.001	0.268	0.194	0.458	0.612	0.765	1.000	1.000	1.000	1.000	0.194	0.458	0.612
163	4	1.545	0.546	0.887	0.137	0.001	0.326	0.100	0.309	0.471	0.649	1.000	1.000	1.000	1.000	0.100	0.309	0.471
163	5	1.667	0.564	0.931	0.141	0.002	0.330	0.046	0.200	0.334	0.530	1.000	1.000	1.000	1.000	0.046	0.200	0.334
163	6	1.664	0.600	0.970	0.144	0.002	0.273	0.019	0.124	0.236	0.415	1.000	1.000	1.000	1.000	0.019	0.124	0.236
163	7	1.731	0.638	1.007	0.147	0.002	0.254	0.008	0.074	0.158	0.312	1.000	1.000	1.000	1.000	0.008	0.074	0.158
163	8	1.775	0.652	1.041	0.149	0.002	0.197	0.004	0.041	0.103	0.232	0.999	1.000	1.000	1.000	0.004	0.041	0.103