The Sources and Nature of Long-Term Memory in Aggregate Output

by Joseph G. Haubrich and Andrew W. Lo

Joseph G. Haubrich is an economic advisor at the Federal Reserve Bank of Cleveland and Andrew W. Lo is the Harris & Harris Group Professor at the MIT Sloan School of Management. This research was partially supported by the MIT Laboratory for Financial Engineering, the National Science Foundation (SES–8520054, SES–8821583, SBR–9709976), 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). The authors 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, and Arzu Ilhan for research assistance

Introduction

Questions about the persistence of economic shocks currently occupy an important place in economics. Much of the recent controversy has centered on "unit roots," determining whether aggregate time series are better approximated by fluctuations around a deterministic trend or by a random walk plus a stationary component. The mixed empirical results reflect the general difficulties in measuring low-frequency components. Persistence, however, has richer and more relevant facets than the asymptotic behavior at the heart of the unit root debate. In particular, fractionally differenced stochastic processes parsimoniously capture an important type of long-range dependence midway between the quick decay of an ARMA process and the infinite persistence of a random walk. Fractional differencing allows something of a return to the classical NBER business cycle program exemplified by Wesley Claire Mitchell, who urged examinations of stylized facts at all frequencies.

Though useful in areas such as international finance (Diebold, Husted, and Rush [1991] and Baillie, Bollerslev, and Mikkelsen [1993]), fractionally differenced processes have had less success in macroeconomics. For GDP at least, it is hard to estimate the appropriate fractional parameter with any precision. One promising technique (Geweke and Porter-Hudak [1983] and Diebold and Rudebusch [1989]) also has serious small-sample bias, which limits its usefulness (Agiakloglou, Newbold, and Wohar [1993]). Although both Deibold and Rudebusch (1989) and Sowell (1992) find point estimates that suggest long-term dependence, they cannot reject either extreme of finite-order ARMA or a random walk. The estimation problems raise again the question posed by Christiano and Eichenbaum (1990) about unit roots: do we know and do we care? In this paper we provide an affirmative answer to both questions.

Do we know? Applying the modified rescaled range (R/S) statistic confronts the data with a test at once both more precise and more robust than previous estimation techniques. The R/S statistic has shown its versatility and usefulness in a variety of different contexts (Lo [1991] and Haubrich [1993]). Operationally, we can determine what we know about our tests for long-range dependence by using Monte Carlo simulations of their size and power. Not surprisingly, with typical macroeconomic sample sizes we cannot distinguish between fractional exponents of 1.000 and 0.999, but we can distinguish between exponents of 0 and 0.333.

Do we care? Persistence matters directly for making predictions and forecasts. It matters in a more subtle way when making econometric inferences. This is especially true for fractionally differenced processes, which though stationary, are not "strong-mixing" and so in a well-defined probabilistic sense behave very differently than more standard ARMA processes. Furthermore, because the optimal filter depends on the characteristics of the underlying process (Christiano and Fitzgerald [1999] and Baxter and King [1999]), long-term persistence (sometimes called long memory) matters even for estimates of higher frequency objects such as business cycle properties. While for some purposes persistence is less important than how agents decompose shocks into permanent and temporary components (Quah [1990]), the results still depend on the persistence of the time series decomposed and on the persistence of the temporary component. In addition, the univariate approach, in contrast to Quah, has the advantage that it does not assume agents observe more than the econometrician.

Finally, beyond the value of purely statistical explorations, a compelling reason to search for fractional differencing comes from economic theory. We show how fractional differencing arises as a natural consequence of aggregation in real business cycle models. This holds out the promise of more specific guidance in looking for long-range dependence, and conversely, adds another criterion to judge and calibrate macroeconomic models.

This paper examines the stochastic properties of aggregate output from the standpoint of fractionally integrated models. We introduce this type of process in section I, reviewing its main properties, advantages, and weaknesses. Section II develops a simple macroeconomic model that exhibits long-range dependence. Section III employs the modified rescaled range statistic to search for long-range dependence in the data. We conclude in section IV.

I. Review of Fractional Techniques in Statistics

Macroeconomic time series look like neither a random walk nor white noise, suggesting that some compromise or hybrid between white noise and its integral may be useful. Such a concept has been given content through the development of the fractional calculus, that is, differentiation and integration to noninteger 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 this section.

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

$$(1)\quad (1-L)^d X_t = \varepsilon_t\,,$$

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. Using the binomial theorem, the AR representation of X_t becomes:

$$(2)\quad A(L)\,X_t = \sum_{k=0}^{\infty} A_k$$

$$(3)\quad L^k X_t = \sum_{k=0}^{\infty} A_k X_{t-k} = \varepsilon_t\,,$$

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

$$(4)\quad A^k \equiv (-1)^k \binom{d}{k} = \frac{\Gamma(k-d)}{\Gamma(-d)\Gamma(k+1)}\,.$$

By manipulating equation (1) mechanically, X_t may also be viewed as an infinite-order MA process since:

$$(5)\quad X_t = (1-L)^{-d}\varepsilon_t = B(L)\,\varepsilon_t$$

$$B_k = \frac{\Gamma(k+d)}{\Gamma(d)\Gamma(k+1)}\,.$$

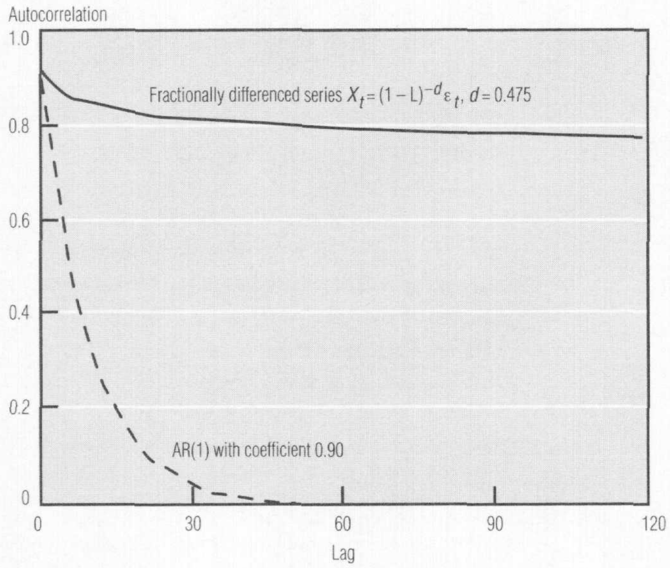
The particular time-series properties of X_t depend intimately on the value of the differencing parameter d. For example, Granger and Joyeux (1980) and Hosking (1981) show that when d is less than $\frac{1}{2}$, X_t is stationary; when d is greater than $-\frac{1}{2}$, X_t is invertible. Although the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

The AR and MA representations of fractionally differenced time series have many applications and illustrate the central properties of fractional processes, particularly long-range dependence. The MA coefficients, B_k, give the effect of a shock k periods ahead and indicate the extent to which current levels of the process depend on past values.

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 nineteenth 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), who also present an extensive historical discussion. Kolmogorov (1940) was apparently the first to notice its applications in probability and statistics.

FIGURE 1

Autocorrelation Functions
ρ_j for $(1-L)^{-.475}\,\varepsilon_t$



SOURCE: Authors' calculations.

How fast this dependence decays furnishes valuable information about the process. Using Stirling's approximation, we have:

$$(6)\ B_k \approx \frac{k^{d-1}}{\Gamma(d)}$$

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 $(1-L)^{-0.475}X_t = \varepsilon_t$ with the AR(1) $X_t = 0.9X_{t-1} + \varepsilon_t$. Although they both have first-order autocorrelations of 0.90, the AR(1)'s autocorrelation function decays much more rapidly.

These representations also show how standard econometric methods can fail to detect fractional processes, necessitating the methods of section III. 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 outpredict fitted ARMA models.

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. From equation 6, it is immediate that for $0 < d < 1$, $lim_{k\to\infty} B_k = 0$, and asymptotically there is no persistence in a fractionally differenced series, even though the autocorrelations die out very slowly.[2]

This holds true not only for $d < \frac{1}{2}$ (the stationary case), but also for $\frac{1}{2} < d < 1$ (the nonstationary case).

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 spectrum, or spectral density (denoted $f(\omega)$) of a fractionally differenced process reflects these properties. It exhibits a peak at 0 (unlike the flat spectrum of an ARMA process), but one not as sharp as the random walk's. Given $X_t = (1-L)^{-d}\varepsilon_t$, the series is clearly the output of a linear system with a white-noise input, so that the spectrum of X_t is:

$$(7)\ f(\omega) = \frac{1}{|1-z|^{2d}}\frac{\sigma^2}{2\pi} \quad \text{where}$$

$$z \equiv e^{-i\omega}, \qquad \sigma^2 \equiv E[\varepsilon_t^2].$$

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

$$(8)\ f(\omega) = c\,\omega^{-2d}, \qquad c \equiv \frac{\sigma^2}{2\pi}.$$

This encompasses the two extremes of a white noise (or a finite ARMA) 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 physics literature, called $1/f$ noise, matches fractionally integrated noise with $d = \frac{1}{2}$.

■ 2 There has been some confusion in the literature on this point. Geweke and Porter-Hudak (1983) argue that $\lim_{k\to\infty} B_k > 0$, which, in their terminology, is expressed as $C(1) > 0$. 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 6 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:

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

II. A Simple Macroeconomic Model with Long-Term Dependence

Economic insight requires more than a consensus on the Wold representation of GDP; it demands a falsifiable model based on the tastes and technology of the actual economy. As Wesley Claire Mitchell (1927, p. 230) wrote, "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."

Thus, before testing for long-run dependence, we develop a simple model where aggregate output exhibits long-run dependence. This model 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-range dependence to central economic concepts of productivity, aggregation, and the limits of the representative agent paradigm.

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 instance, find long-run dependence in rainfall, riverflows, earthquakes, and weather (measured by tree rings and sediment deposits).[3]

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).[4] In our multisector 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 quadratic utility, with a lifetime utility function of $U = \sum \beta^t u(C_t)$, where C_t is an $N \times 1$ 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

$$(9) \quad u(C_t) = C_t' \iota - \tfrac{1}{2} C_t' B C_t ,$$

where ι is an $N \times 1$ vector of ones. In anticipation of the aggregation considered later, we assume B to be diagonal so that $C_t' B C_t = \sum b_{ii} C_{it}^2$. The agents face a resource constraint: total output Y_t may be either consumed or saved, thus:

$$(10) \quad C_t + S_t \iota = Y_t ,$$

where the i,j-th entry S_{ijt} of the $N \times N$ 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 a random linear technology:

$$(11) \quad Y_t = AS_t + \varepsilon_t ,$$

where A is the matrix of input–output coefficients a_{ij}, and ε_t is a (vector) random production shock, whose value is realized at the beginning of period $t + 1$. To focus on long-range 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.

In this case, the dynamic programming problem of solving for optimal consumption and investment policies reduces to the familiar optimal stochastic linear regulator problem (see Sargent [1987], section 1.8, for an excellent exposition). Given the simple diagonal form of the A matrix, which corresponds to an assumption that each sector uses only its own output as input, the problem simplifies even further.

■ 3 For a related mechanism creating fractional intergration by aggregating shocks of differing duration, see Parke (1999). Abadir and Talmain (undated) use aggregation over heterogeneous firms in a setting of monopolistic competition.

■ 4 Dupor (1999) is skeptical of the ability of multisector models to match aggregate time series data but does not consider long-range dependence.

The chosen quantities of consumption and investment/savings have the following closed-form solutions:

$$(12)\ S_{it} = \frac{b_i}{b_i - 2\beta P_i a_i^2} Y_{it} + \frac{\beta q_i a_i - 1}{b_i - 2\beta P_i a_i^2}$$

$$(13)\ C_{it} = \frac{2\beta P_i a_i^2}{2\beta P_i a_i^2 - b_i} Y_{it} + \frac{\beta q_i a_i - 1}{2\beta P_i a_i^2 - b_i},$$

where:

$$(14)\ P_i \equiv b_i \left[\frac{a_i - \sqrt{(1+4\beta)a_i^2 - 4}}{4\beta a_i}\right],$$

and q_i are fixed constants given by the matrix Riccati equation that results from the recursive definition of the value function.

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 independent and identically distributed (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 an artifact of the linear-quadratic combination.

The time series of output can now be calculated from the production function, equation (11), and the decision rule, equation (12). Quantity dynamics then come from the difference equation:

$$(15)\ Y_{it+1} = \frac{a_i b_i}{b_i - 2\beta P_i a_i^2} Y_{it} + K_i + \varepsilon_{it+1}$$

or

$$(16)\ Y_{it+1} = \alpha_i Y_{it} + K_i + \varepsilon_{it+1},$$

where α_i is a function of the utility parameters and of a_i, the input-output coefficient of the industry, and K_i is some fixed constant. The key qualitative property of quantity dynamics summarized by equation (16) 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 demonstrate here 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, 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 (or islands) have the same AR(1) representation for output, which reduces the dimension of aggregate ARMA process. This "cancelling of roots" requires identical autoregressive parameters. An alternative approach, due to Granger, 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 sectors. 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:

$$(17)\ Y_t^a \equiv \sum_{i=1}^{N} Y_{it}.$$

To show this, we approach the problem from the frequency domain and apply spectral methods, which often simplify problems of aggregation.[5] 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:

$$(18)\ f_i(\omega) = \frac{1}{|1 - \alpha_i z|^2} \frac{\sigma_i^2}{2\pi}.$$

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

$$(19)\ f_i(\omega) = \sum_{i=1}^{N} f_i(\omega).$$

■ 5 See Theil (1954).

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 as 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:

$$(20)\ f_i(\omega) = \frac{N}{2\pi} E[\sigma^2] \cdot \int \frac{1}{|1-\alpha_i z|^2} dF(\alpha).$$

If the distribution $F(\alpha)$ is discrete, so that it takes on $m(< N)$ values, Y_t^a will be an ARMA $(m,m\text{–}1)$ 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.[6] In particular, let α^2 have a beta distribution $\beta(p,q)$, which yields the following density function for α:

$$(21)\ dF(\alpha) = \begin{cases} \frac{2}{\beta(p,q)} \alpha^{2p-1}(1-\alpha^2)^{q-1} d\alpha, & 0 \le \alpha \le 1; \\ 0, & \text{otherwise.} \end{cases}$$

with $p,q > 0$.[7] Obtaining the Wold representation of the resulting process requires a little more work. First note that:

$$(22)\ 1/|1-\alpha z|^2 = \frac{1}{2(1-\alpha^2)} \left[\frac{1+\alpha z}{1-\alpha z} + \frac{1+\alpha \bar{z}}{1-\alpha \bar{z}} \right],$$

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, equation (21), into the expression for the spectrum and simplifying (using the relation $z + \bar{z} = 2\cos(\omega)$) yields:

$$(23)\ f(\omega) = \int_0^1 [2 + 2 \sum_{k=1}^{\infty} \alpha^k \cos(k\,\omega)].$$

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

$$(24)\ \int_0^1 \frac{2\alpha^k}{\beta(p,q)} \alpha^{2p-1}(a-\alpha^2)^{q-1} d\alpha.$$

Since the spectral density is the Fourier transform of the autocovariance function, equation (24) is the k-th autocovariance of Y_t^a. Furthermore, because the integral defines a beta function, equation (24) simplifies to $\beta(p + k/2, q-1)/\beta(p,q)$. Dividing by the variance gives the autocorrelation coefficients, which reduce to

$$(25)\ \rho(k) = \frac{\Gamma(p+q-1)}{\Gamma(p)} \frac{\Gamma(p+\frac{k}{2})}{\Gamma(p+\frac{k}{2}+q+1)},$$

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 frequencies, such as those of the business cycle. Similarly, co-movements can arise, as the fractionally integrated income process may induce fractional integration in other observed time series. 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 a_i's so that α_i has a beta distribution.

In principle, all parameters of the model may be estimated, from the distribution of production function parameters to the variance of output shocks. Empirical estimates of production function parameters (such as those in Jorgenson, Gollop, and Fraumeni [1987]) reveal a large dispersion, suggesting the plausibility and significance of the simple model presented in this section.

Although the original motivation of our real business cycle model was to illustrate how long-range dependence could arise naturally in an economic system, our results have broader implications for general macroeconomic modeling. They show that moving to a multiple-sector real business cycle model introduces not unmanageable complexity, but qualitatively new behavior that can be quite manageable. Our findings also show that calibrations aimed at matching only a few first and second moments can similarly hide major differences between models and the data, missing long-range dependence properties. While widening the theoretical horizons of the paradigm, they therefore also widen the potential testing of such theories.

6 Granger (1980) conjectures that this particular distribution is not essential.

7 For a discussion of the variety of shapes the beta distribution takes as p and q vary, see Johnson and Kotz (1970).

III. R/S Analysis of Real Output

The results of section II show that simple aggregation may be one source of long-range dependence in the business cycle. In this section we employ a method for detecting long memory and apply it to real GDP. 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.[8]

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 below. Then we present the empirical results for real GDP; we find long-range 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.

The Rescaled Range Statistic

We test for fractional differencing using Lo's modification of the modified rescaled range (R/S) statistic. In particular, we define short-range dependence as Rosenblatt's (1956) concept of "strong-mixing," 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-range 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 sequences (such as those exhibiting heteroscedasticity), an especially important extension in view of the nonstationarities of real GDP.

Fractionally differenced models, however, possess autocorrelation functions that decay at much slower rates than those of weakly dependent processes and violate the conditions of strong mixing. More formally, let X_t denote the first difference of log-GDP; we assume that:

$$(26)\quad X_t = \mu + \varepsilon_t ,$$

where μ is an arbitrary but fixed parameter. For the null hypothesis H, assume that the sequence of disturbances, $\{\varepsilon_t\}$, satisfies the following conditions:

(A1) $E[\varepsilon_t] = 0$ *for all t.*

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

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

(A4) $\{\varepsilon_t\}$ *is strong-mixing with mixing coefficients* α_k *that satisfy:*[9]

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

Condition (A1) is standard. Conditions (A2)–(A4) allow dependence and heteroskedasticity, but prevent them from being too dominant. Thus, short-range dependent processes such as finite-order ARMA models are included in this null hypothesis, as are models with conditional heteroskedasticity. Unlike the statistic used by Mandelbrot, the modified R/S statistic is robust to short-range dependence. A more detailed discussion of these conditions may be found in Phillips (1987) and Lo (1991).

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 R/S statistic, which we shall call Q_n, is given by:

$$(27)\quad Q_n \equiv \frac{1}{\hat{\sigma}_n(q)}\left[\mathrm{Max}_{1\le k\le n}\sum_{j=1}^{k}(X_j-\bar{X}_n) - \mathrm{Min}_{1\le k\le n}\sum_{j=1}^{k}(X_j-\bar{X}_n)\right],$$

■ 8 See Mandelbrot and Taqqu (1979) for further references.

■ 9 Let $\{\varepsilon_t(\omega)\}$ 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 $\{\varepsilon_s(\omega),\ldots,\varepsilon_t(\omega)\}$, i.e., $\mathcal{B}_s^t \equiv \sigma(\varepsilon_s(\omega),\ldots,\varepsilon_t(\omega)) \subset \mathcal{F}$. Define the coefficients α_k as:

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

Then $\{\varepsilon_t(\omega)\}$ 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).

TABLE 1(a)

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	1.474	1.862	2.098

SOURCE: Authors' calculations.

TABLE 1(b)

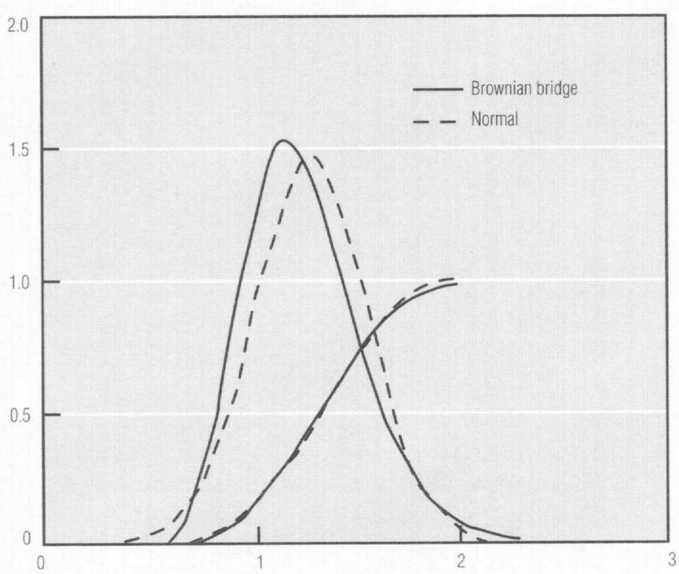
Symmetric Confidence Intervals about the Mean

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

SOURCE: Authors' calculations.

FIGURE 2

Distribution and Density Function of the Range V of a Brownian Bridge



SOURCE: Authors' calculations.

where

$$(28)\quad \hat{\sigma}_n^2(q) \equiv \frac{1}{n}\sum_{j=1}^{n}(X_j-\bar{X}_n)^2 + \frac{2}{n}\sum_{j=1}^{q}\omega_j(q)\left\{\sum_{i=j+1}^{n}(X_i-\bar{X}_n)(X_{i-j}-\bar{X}_n)\right\}$$

$$(29)\quad = \hat{\sigma}_x^2 + 2\sum_{j=1}^{q}\omega_j(q)\,\hat{\gamma}_j$$

$$\omega_j(q) \equiv 1-\frac{j}{q+1} \qquad q<n$$

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 $\hat{\sigma}_n^2(q)$.[10]

Intuitively, the numerator in equation (27) measures the memory in the process via the partial sums. White noise does not stay long above the mean: positive values are soon offset by negative values. A random walk will stay above or below 0 for a long time, and the partial sums (positive or negative) will grow quickly, making the range large. Fractional processes fall in between. Mandelbrot (1972) refers to their behavior as the "Joseph effect"—seven fat and seven lean years. The denominator normalizes not only by the variance but also by a weighted average of the autocovariances. This innovation over Hurst's (1951) R/S statistic provides the robustness to short-range dependence.

The choice of the truncation lag, q, is a delicate matter. Although q must increase with (albeit 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.[11] However, q cannot be too small or 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.

The partial sums of white noise constitute a random walk, so $Q_n(q)$ grows without bound as n increases. A further normalization makes the statistic easier to work with and interpret:

$$(30)\quad V_n(q) \equiv \frac{Q_n(q)}{\sqrt{n}}$$

■ 10 $\hat{\sigma}_n^2(q)$ is also an estimator of the special density function of X_t at frequency zero, using a Bartlett window.

■ 11 See, for example, Lo and MacKinlay (1989).

TABLE 2

R/S Analysis of Real GDP

Series	$\tilde{V}_n$	$V_n(1)$	$V_n(2)$	$V_n(3)$	$V_n(4)$	$V_n(5)$	$V_n(6)$	$V_n(7)$	$V_n(8)$
US Log First-Difference	1.092	0.973	0.933	0.934	0.971	1.032	1.082	1.115	1.139
Percentage Bias of $\tilde{V}_n$		(6.0)	(8.2)	(8.1)	(6.1)	(2.9)	(0.5)	(–1.0)	(–2.1)
S Log Detrended	2.374	1.741	1.479	1.337	1.252	1.198	1.160	1.134	1.116
Percentage Bias of $\tilde{V}_n$		(16.8)	(26.7)	(33.2)	(37.7)	(40.8)	(43.1)	(44.7)	(45.8)
CAN Log First-Difference	1.254	1.116	1.042	1.018	1.024	1.045	1.073	1.096	1.132
Percentage Bias of $\tilde{V}_n$		(6.0)	(9.7)	(11.0)	(10.7)	(9.5)	(8.1)	(7.0)	(5.2)
CAN Log Detrended	3.410	2.458	2.048	1.813	1.660	1.552	1.472	1.410	1.360
Percentage Bias of $\tilde{V}_n$		(17.8)	(29.1)	(37.2)	(43.3)	(48.2)	(52.2)	(55.5)	(58.3)
GER Log First-Difference	1.357	1.185	1.159	1.158	1.176	1.203	1.235	1.278	1.325
Percentage Bias of $\tilde{V}_n$		(7.0)	(8.2)	(8.3)	(7.4)	(6.2)	(4.8)	(3.1)	(1.2)
GER Log Detrended	4.241	3.052	2.539	2.242	2.044	1.903	1.796	1.712	1.643
Percentage Bias of $\tilde{V}_n$		(17.9)	(29.2)	(37.5)	(44.0)	(49.3)	(53.7)	(57.4)	(60.6)
UK Log First-Difference	1.051	0.907	0.851	0.853	0.887	0.920	0.961	0.993	1.031
Percentage Bias of $\tilde{V}_n$		(7.7)	(11.1)	(11.0)	(8.9)	(6.9)	(4.6)	(2.9)	(1.0)
UK Log Detrended	4.637	3.327	2.760	2.431	2.213	2.055	1.933	1.837	1.757
Percentage Bias of $\tilde{V}_n$		(18.1)	(29.6)	(38.1)	(44.8)	(50.2)	(54.9)	(58.9)	(62.5)
GDP Log First-Difference	1.391	1.201	1.107	1.066	1.057	1.069	1.087	1.109	1.132
Percentage Bias of $\tilde{V}_n$		(7.6)	(12.1)	(14.3)	(14.7)	(14.1)	(13.1)	(12.0)	(10.9)
GDP Log Detrended	5.612	3.999	3.290	2.873	2.592	2.387	2.229	2.103	2.000
Percentage Bias of $\tilde{V}_n$		(18.5)	(30.6)	(39.8)	(47.1)	(53.3)	(58.7)	(63.4)	(67.5)

NOTE: US, CAN, GER, and UK refer to the annual Maddison series for those countries 1870–1994. GDP refers to the quarterly U.S. series 1947:QI–1999:QI. The classical rescaled range $\tilde{V}_n$ and the modified rescaled range $V_n(q)$ are reported. Under a null hypothesis of short-range dependence, 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 rows are computed as $[\tilde{V}_n / V_n(q)^{1/2} - 1] \bullet 100$ and are estimates of the bias of the classical R/S statistic in the presence of short-term dependence.

SOURCE: Authors' calculations.

The limiting distribution of $V_n(q)$ is derived by Lo (1991), and its most commonly-used values are reported in tables 1(a) and 1(b). Table 1(a) reports the fractiles of the limiting distribution while table 1(b) reports the symmetric confidence intervals about the mean. The moments of the limiting distribution are also easily computed using f_V, the density of the random variable V to which $V_n(q)$ converges in distribution; 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 2. Observe that the distribution is positively skewed, and most of its mass falls between $\frac{3}{4}$ and 2.

Empirical Results for Real Output

We apply our test to several time series of real output: quarterly U.S. postwar real GDP from 1947:QI to 1999:QI, and the annual Maddison (1995) OECD series for the United States, Canada, Germany, and the United Kingdom from 1870 to 1994. These results are reported in table 2. Entries in the first numerical column are estimates of the classical rescaled range, $\tilde{V}_n$, which is not robust to short-range dependence. The next eight columns 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 $V_n(q)$ is an estimate of the percentage bias of the statistic $\tilde{V}_n$, which is computed as $100 \cdot [(\tilde{V}_n / V_n(q)) - 1]$.

The first row of numerical entries in table 2 indicate that the null hypothesis of short-range dependence for the first-difference of log-GDP cannot be rejected for any value of q. The classical rescaled range statistic also supports the null hypothesis, as do the results for the Maddison series. On the other hand, when we log-linearly detrend real GDP, the results differ considerably. Looking at the results for the annual data in table 2 shows that short-range dependence may be rejected for log-linearly detrended output using the classical statistic for the United States and with q values from 1 to 2 for Canada, 1 to 5 for Germany, and 1 to 6 for the United Kingdom. For quarterly U.S. data, short-term dependence is rejected for all q up to 8. That the rejections are weaker for larger q is not surprising since additional noise arises from estimating higher-order autocorrelations.

The values reported in table 2 are qualitatively consistent with other empirical investigations of fractional processes in GNP, such as Diebold and Rudebusch (1989) and Sowell (1992). For first-differences, the R/S statistic falls below the mean, suggesting a negative fractional exponent, or in level terms, an exponent between 0 and 1. Furthermore, though earlier papers produce point estimates, they do not lead to a rejection of the hypothesis of short-term dependence because of imprecise estimates. For example, the 2 standard deviation error bounds for two point estimates of Diebold and Rudebusch (1989), $d = 0.9$ and 0.52, are [0.42, 1.38] and [–0.06, 1.10], respectively.

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 GDP is precisely the spurious periodicity suggested by Nelson and Kang (1981). Moreover, the trend plus stationary noise model of GDP is not contained in our null hypothesis; hence our failure to reject the null hypothesis is also consistent with the unit root model.[12] To see this, observe that if log-GDP y_t were trend stationary, that is,

$$(31) \quad y_t = \alpha + \beta t + \eta_t$$

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

Sowell (1992) has used estimates of d to argue that the trend-stationary model is correct. Following the lead of Nelson and Plosser (1982), Sowell checks if the d parameter for the first-differenced series is close to 0 as the unit root specification suggests, or close to –1 as the trend-stationary specification suggests. His estimate of d is in the general range of –0.6 to 0.2, providing some evidence that the trend-stationary interpretation is correct. Even in his case though, the standard errors tend to be rather large, on the order of 0.3. Although our procedure yields no point estimate of d, our results do seem to rule out the trend-stationary case.

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.

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 208 observations, corresponding to the number of quarterly observations of real GDP growth from 1947:QII to 1999:QI.[13] 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: 8.221×10^{-3} and 1.0477×10^{-2}, respectively. To choose parameter values for the ARMA(2,2) simulation, we estimate the model:

$$(32) \quad (1 - \phi_1 L - \phi_2 L^2) y_t = \mu + (1 + \theta_1 L + \theta_2 L^2)\varepsilon_t \quad \varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$$

using nonlinear least squares. The parameter estimates are (with standard errors in parentheses):

$$\hat{\phi}_1 = \underset{(0.1678)}{1.3423}, \quad \hat{\theta}_1 = \underset{(0.1839)}{1.0554}$$

$$\hat{\phi}_2 = \underset{(0.1198)}{-0.7065}, \quad \hat{\theta}_2 = \underset{(0.1377)}{-0.5200}$$

$$\hat{\mu} = \underset{(0.0008)}{0.0082}, \quad \hat{\sigma}_\varepsilon = 0.0097$$

■ **12** Of course, this may be the result of low power against stationary but near-integrated processes, and it must be addressed by Monte Carlo experiments.

■ **13** Simulations were performed on a DEC Alphaserver 2100 4/275 using a Gauss random number generator; each experiment comprised 10,000 replications.

TABLE 3

Finite Sample Distribution of the Modified R/S Statistic under IID and ARMA (2,2) Null Hypotheses for the First-Difference of Real Log GNP

q	Min.	Max.	Mean	S.D.	Size 1%–Test	Size 5%–Test	Size 10%–Test
IID Null							
0	0.527	2.468	1.175	0.266	0.002	0.030	0.061
1.5	0.525	2.457	1.171	0.253	0.015	0.069	0.121
1	0.548	2.342	1.177	0.258	0.001	0.027	0.052
2	0.548	2.251	1.180	0.251	0.001	0.024	0.054
3	0.555	2.203	1.183	0.245	0.000	0.021	0.052
4	0.572	2.156	1.187	0.240	0.000	0.020	0.050
5	0.592	2.098	1.190	0.234	0.000	0.018	0.046
6	0.622	2.058	1.193	0.228	0.000	0.015	0.044
7	0.637	2.031	1.197	0.223	0.000	0.012	0.041
8	0.657	1.981	1.200	0.218	0.000	0.010	0.038
ARMA (2,2) Null							
0	0.654	2.864	1.411	0.314	0.025	0.152	0.245
6.8	0.610	2.200	1.177	0.229	0.009	0.041	0.084
1	0.570	2.473	1.227	0.269	0.004	0.039	0.083
2	0.533	2.269	1.134	0.246	0.001	0.014	0.035
3	0.517	2.190	1.094	0.233	0.000	0.007	0.021
4	0.522	2.139	1.086	0.226	0.000	0.006	0.016
5	0.543	2.101	1.099	0.223	0.000	0.005	0.017
6	0.562	2.035	1.123	0.221	0.000	0.006	0.020
7	0.587	2.011	1.149	0.220	0.000	0.007	0.024
8	0.620	1.995	1.171	0.218	0.000	0.008	0.029

NOTE: The Monte Carlo experiments under the two null hypotheses are independent and consist of 10,000 replications each, for a sample size $n = 208$. Parameters of the i.i.d. simulations were chosen to match the sample mean and variance of quarterly real GNP growth rates from 1947:QII to 1999:QI; 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 noninteger lag value corresponds to the average (across replications) lag value used according to Andrews's (1991) optimal lag formula. Standard errors for the empirical size may be computed using the usual normal approximation; they are 9.95×10^{-4}, 2.18×10^{-3}, and 3.00×10^{-3} for the 1, 5, and 10 percent tests, respectively.
SOURCE: Authors' calculations.

Table 3 reports the results of both null simulations. It is apparent from the i.i.d. null panel of table 3 that when serial correlation is not a problem, the classical and modified rescaled range statistics perform similarly. The 5 percent test using the classical statistic rejects 3 percent of the time: the modified R/S with $q = 4$ rejects 2 percent of the time. As the number of lags increases to 8, the test becomes more conservative. Under the ARMA(2,2) null hypothesis, however, it is apparent that modifying the R/S by the spectral density estimator $\hat{\sigma}_n^2(q)$ is critical; the size of a 5 percent test based on the classical R/S is 15.2 percent, whereas the corresponding size using the modified R/S statistic with $q = 1$ is 3.9 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's (1991) 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:[14]

$$(33)\quad \bar{M}_n \equiv \left(\frac{3\hat{\alpha}n}{2}\right)^{1/3} \qquad \hat{\alpha} \equiv \frac{4\hat{\rho}^2}{(1-\hat{\rho}^2)^2}.$$

Under the i.i.d. null hypothesis, Andrews's 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's 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.

14 In addition, Andrews's 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.

TABLE 4

Power of the Modified R/S Statistics under a Gaussian Fractionally Differenced Alternative with Differencing Parameters $d = 1/3, -1/3$

q	Min.	Max.	Mean	S.D.	Power 1%–Test	Power 5%–Test	Power 10%–Test
$d = 1/3$							
0	0.888	5.296	2.551	0.673	0.722	0.842	0.890
6.1	0.665	2.569	1.577	0.305	0.039	0.193	0.310
1	0.825	4.112	2.149	0.528	0.511	0.680	0.758
2	0.752	3.497	1.936	0.452	0.355	0.543	0.638
3	0.712	3.126	1.799	0.403	0.244	0.427	0.535
4	0.687	2.877	1.701	0.367	0.156	0.339	0.446
5	0.675	2.616	1.630	0.344	0.097	0.268	0.379
6	0.669	2.469	1.571	0.321	0.051	0.203	0.3171
7	0.666	2.350	1.523	0.302	0.020	0.148	0.256
8	0.663	2.294	1.481	0.281	0.007	0.095	0.196
$d = -1/3$							
0	0.339	1.009	0.583	0.095	0.917	0.979	0.993
3.9	0.467	1.598	0.814	0.132	0.257	0.525	0.671
1	0.398	1.136	0.673	0.108	0.698	0.890	0.944
2	0.443	1.282	0.741	0.117	0.474	0.736	0.848
4	0.518	1.468	0.844	0.129	0.167	0.430	0.594
5	0.550	1.499	0.884	0.132	0.091	0.312	0.467
6	0.576	1.573	0.922	0.136	0.046	0.2134	0.358
7	0.613	1.633	0.957	0.139	0.021	0.138	0.263
8	0.596	1.578	0.989	0.143	0.011	0.092	0.190

NOTE: The Monte Carlo experiments under the two alternative hypotheses are independent and consist of 10,000 replications each, for sample size $n = 208$. Parameters of the simulations were chosen to match the sample mean and variance of quarterly real GDP growth rates from 1947:QII to 1999:QI. Entries in the column labeled "q" indicate the number of lags used to compute the R/S statistic; a lag of 0 corresponds to Mandelbrot's classical range, and a noninteger lag value corresponds to the average (across replications) lag value used according to Andrews's (1991) optimal lag formula.
SOURCE: Authors' calculations.

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

$$(34)\quad \gamma_\varepsilon(k) = \frac{\Gamma(1-2d)\,\Gamma(d+k)}{\Gamma(d)\,\Gamma(1-d)\,\Gamma(1-d+k)}\,\hat{\sigma}_\eta^2$$

$$d \in \left(-\frac{1}{2}, \frac{1}{2}\right).$$

Realizations of fractionally differenced time series (of length 208) are simulated by premultiplying vectors of independent standard normal random variates by the Cholesky factorization of the (208 x 208) covariance matrix whose entries are given by equation (34). To calibrate the simulations, σ_η^2 is chosen to yield unit variance ε_t's, the $\{\varepsilon_t\}$ series is then multiplied by the sample standard deviation of real GDP growth from 1947:QII to 1999:QI, and to this series is added the sample mean of real GDP growth over the same sample 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 has 68 percent power against the $d = 1/3$ alternative, and 89 percent power against the $d = -1/3$ alternative. As the lag length is increased, the test's power declines.

Note that tests based on the classical rescaled range are significantly more powerful than those using the modified R/S statistic. This, however, is of little value when distinguishing between long-range versus short-range dependence since the test using the classical statistic also has power against some stationary finite-order ARMA processes. Finally, note that tests using Andrews's 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 suggests that our empirical results do indeed support the short-range dependence of GDP 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 GDP growth rates.

IV. Conclusions

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.

Fractionally integrated random processes provide one explicit approach to the problem of long-range 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-range dependence deserves study in its own right. Furthermore, Singleton (1988) has pointed out that dynamic macroeconomic models often inextricably link predictions about business cycles, trends, and seasonal effects. So, too, is long-range dependence linked: 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-range dependence in aggregate data and 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-range dependence. Modifying a statistic of Hurst and Mandelbrot gives us a statistic robust to short-range 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-range dependence in GDP—the greater power of the modified R/S test may explain why our results contradict earlier work that purported to find long-range dependence. Rejections of the short-range dependence null hypothesis occur only with detrended data, and this 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 GDP, with the additional implication that the resulting stationary process is weakly dependent at most. This supports and extends the conclusion of Adelman (1956) that, at least within the confines of the available data, there is little evidence of long-range dependence in the business cycle. Nevertheless, Haubrich (1993) finds indirect evidence for long-range dependence using aggregate consumption series, and hence the empirical relevance of long memory for economic phenomena remains an open question that deserves further investigation.

References

Abadir, Karim, and Bariel Talmain. "Aggregation, Persistence, and Volatility in a Macromodel," University of York (undated).

Adelman, Irma. "Long Cycles: Fact or Artifact?" *American Economic Review*, vol. 55, no. 3 (June 1965), pp. 444–63.

Andrews, Donald W. K. "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, vol. 59, no. 3 (May 1991), pp. 817–58.

Agiakloglou, Christos, Paul Newbold, and Mark Wohar. "Bias in an Estimator of the Fractional Difference Parameter," *Journal of Time Series Analysis*, vol. 14, no. 3 (1993), pp. 235–46.

Baillie, Richard T., Tim Bollerslev, and Hans Ole Mikkelsen. "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, vol. 74, no. 1 (September 1996), pp. 3–30.

Baxter, Marianne, and Robert G. King. "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series," *Review of Economics and Statistics*, vol 81, no. 4 (November 1999), pp. 575–93.

Campbell, John Y., and N. Gregory Mankiw. "Are Output Fluctuations Transitory?" *Quarterly Journal of Economics*, vol. 102, no. 4 (November 1987), pp. 857–80.

Christiano, Lawrence J., and Martin Eichenbaum. "Unit Roots in Real GNP: Do We Know, and Do We Care?" *Carnegie-Rochester Conference Series on Public Policy*, vol. 32, (Spring 1990), pp. 7–61.

Christiano, Lawrence J., and Terry J. Fitzgerald. "The Band Pass Filter," National Bureau of Economic Research, Working Paper No. 7257, July 1999.

Cochrane, John H. "How Big is the Random Walk in GNP?" *Journal of Political Economy*, vol. 96, no. 5 (October 1988), pp. 893–920.

Diebold, Francis X., Steven Husted, and Mark Rush. "Real Exchange Rates under the Gold Standard," *Journal of Political Economy*, vol. 99, no. 6 (December 1991), pp. 1252–71.

Diebold, Francis X., and Glenn D. Rudebusch. "Long Memory and Persistence in Aggregate Output," *Journal of Monetary Economics*, vol. 24, no. 2 (September 1989), pp. 189–209.

Dupor, Bill. "Aggregation and Irrelevance in Multi-sector models," *Journal of Monetary Economics*, vol. 43, no. 2 (April 1999), pp. 391–409.

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

Fisher, Irving. "Our Unstable Dollar and the So-Called Business Cycle," *Journal of the American Statistical Association*, vol. 20, new series no. 149 (March 1925), pp. 179–202.

Geweke, John, and Susan Porter-Hudak. "The Estimation and Application of Long Memory Time Series Models," *Journal of Time Series Analysis*, vol. 4, no. 4 (1983), pp. 221–38.

Granger, Clive W. J. "The Typical Spectral Shape of an Economic Variable," *Econometrica*, vol. 34, no. 1 (January 1966), pp. 150–61.

———. "Long Memory Relationships and the Aggregation of Dynamic Models," *Journal of Econometrics*, vol. 14, no. 2 (October 1980), pp. 227–38.

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

Granger, Clive W. J. and Roselyne Joyeux. "An Introduction to Long-Memory Time Series Models and Fractional Differencing," *Journal of Time Series Analysis*, vol. 1, no. 1 (1980), pp. 15–29.

Haubrich, Joseph G. "Consumption and Fractional Differencing: Old and New Anomalies," *Review of Economics and Statistics*, vol. 75, no. 4 (November 1993), pp. 767–72.

Herrndorf, Norbert. "A Functional Central Limit Theorem for Strongly Mixing Sequences of Random Variables," *Zeitschrift für Wahrscheinlichkeitstheorie verwandte Gebiete*, vol. 69, no. 4 (1985), pp. 541–50.

Hosking, J.R.M. "Fractional Differencing," *Biometrika*, vol. 68, no. 1 (April 1981), pp. 165–76.

Hurst, Harold E. "Long-Term Storage Capacity of Reservoirs," *Transactions of the American Society of Civil Engineers*, vol. 116, paper no. 2447 (1951), pp. 770–99.

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

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

Kolmogorov, Andrei N. "Wienersche Spiralen und einige andere interessante Kurven im Hilbertschen Raum," *Comptes Rendus (Doklady) de l'Académie des Sciences de l' URSS*, vol. 26, no. 2 (January 1940), pp. 115–18.

Kuznets, Simon. *Economic Growth and Structure*. New York: Norton, 1965

Kydland, Finn E., and Edward C. Prescott. "Time to Build and Aggregate Fluctuations," *Econometrica*, vol. 50, no. 6 (November 1982), pp. 1345–70.

Lo, Andrew W. "Long-Term Memory in Stock Market Prices," *Econometrica*, vol. 59, no. 5 (September 1991), pp. 1279–313.

Lo, Andrew W., and A. Craig MacKinlay. "The Size and Power of the Variance Ratio Test in Finite Samples: A Monte Carlo Investigation," *Journal of Econometrics*, vol. 40, no. 2 (February 1989), pp. 203–38.

Long, John B., Jr., and Charles I. Plosser. "Real Business Cycles," *Journal of Political Economy*, vol. 91, no. 1 (February 1983), pp. 39–69.

Lucas, Robert E., Jr. *Models of Business Cycles*. New York: Basil Blackwell, 1987.

Maddison, Angus. *Monitoring the World Economy 1820–1992*. New York: Organisation for Economic Co-operation and Development, 1995.

Mandelbrot, Benoit. "Statistical Methodology for Non-Periodic Cycles: From the Covariance to R/S Analysis," *Annals of Economic and Social Measurement*, vol. 1, no. 3 (July 1972), pp. 259–90.

———. "Limit Theorems on the Self-Normalized Range for Weakly and Strongly Dependent Processes," *Zeitschrift für Wahrscheinlichkeitstheorie verwandte Gebiete*, vol. 31, no. 2 (1975), pp. 271–85.

Mandelbrot, Benoit, and Murad Taqqu. "Robust R/S Analysis of Long-Run Serial Correlation," *Bulletin of the International Statistical Institute*, vol. 48, book 2 (1979), pp. 59–104.

Mandelbrot, Benoit, and James Wallis. "Noah, Joseph, and Operational Hydrology," *Water Resources Research*, vol. 4, no. 5 (October 1968), pp. 909–18.

———. "Some Long-Run Properties of Geophysical Records," *Water Resources Research*, vol. 5, no. 2 (April 1969), pp. 321–40.

Mitchell, Wesley Claire. *Business Cycles: The Problem and Its Setting*, NBER Studies in Business Cycles No. 1, New York: National Bureau of Economic Research, 1927.

Nelson, Charles R., and Heejoon Kang. "Spurious Periodicity in Inappropriately Detrended Time Series," *Econometrica*, vol. 49, no. 3 (May 1981), pp. 741–51.

Nelson, Charles R., and Charles I. Plosser. "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics*, vol. 10, no. 2 (September 1982), pp. 139–62.

Newey, Whitney K., and Kenneth D. West. "A Simple, Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, vol. 55, no. 3 (May 1987), pp. 703–08.

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

Parke, William R. "What is Fractional Integration?" *Review of Economics and Statistics*, vol. 81, no. 4 (November 1999), pp. 632–38.

Perron, Pierre, and Peter C. B. Phillips. "Does GNP Have a Unit Root? A Reevaluation," *Economics Letters*, vol. 23, no. 2 (1987), pp. 139–45.

Phillips, Peter C. B. "Time Series Regression with a Unit Root," *Econometrica*, vol. 55, no. 2 (March 1987), pp. 277–301.

Rosenblatt, Murray. "A Central Limit Theorem and a Strong Mixing Condition," *Proceedings of the National Academy of Sciences,* vol. 42, (1956), pp. 43–7.

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

Singleton, Kenneth J. "Econometric Issues in the Analysis of Equilibrium Business Cycle Models," *Journal of Monetary Economics*, vol. 21, no. 2/3 (March/May 1988), pp. 361–86.

Slutzky, Eugen. "The Summation of Random Causes as the Source of Cyclic Processes," *Econometrica,* vol. 5, no. 2 (April 1937), pp. 105–46.

Sowell, Fallaw. "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models," *Journal of Econometrics,* vol. 53, no 1/3 (July/September 1992), pp. 165–188.

———. "Modeling Long-Run Behavior with the Fractional ARIMA Model," *Journal of Monetary Economics*, vol. 29, no. 2 (April 1992), pp. 277–302.

———. "The Fractional Unit Root Distribution," *Econometrica*, vol. 58, no. 2 (March 1990), pp. 495–506.

Stock, James, and Mark Watson. "Does GNP Have a Unit Root?" *Economics Letters,* vol. 22, no. 2/3 (1986), pp. 147–51.

White, Halbert. "A Heteroscedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroscedasticity," *Econometrica*, vol. 48, no. 4 (May 1980), pp. 817–38.

———. *Asymptotic Theory for Econometricians.* Orlando, Fl.: Academic, 1984.

White, Halbert, and Ian Domowitz. "Nonlinear Regression with Dependent Observations," *Econometrica*, vol. 52, no.1 (January 1984), pp. 143–62.

Yajima, Yoshihiro. "On Estimation of Long-Memory Time Series Models," *Australian Journal of Statistics,* vol. 27, no. 3 (1985), pp. 303–20.

———. "On Estimation of a Regression Model with Long-Memory Stationary Errors," *Annals of Statistics*, vol. 16, no. 2 (June 1988), pp. 791–807.