Reprinted from: Baldev Raj (ed.) (1989), *Advances in Econometrics and Modeling*. Needham, Mass.: Kluwer Academic Publishers. *pp.* 29-45.

CHAPTER 2

RANDOM WALKS VERSUS FRACTIONAL INTEGRATION: POWER COMPARISONS OF SCALAR AND JOINT TESTS OF THE VARIANCE-TIME FUNCTION

FRANCIS X. DIEBOLD *
Board of Governors of the Federal Reserve System
Division of Research and Statistics
20th and Constitution Ave., NW
Washington, DC 20551
U.S.A.

ABSTRACT. A class of tests for the detection of deviations from random-walk behavior in observed time series is examined. The tests are based on the variance-time function, which maps integers k into the variance of k-th differences of a time series. Both simple and joint null hypotheses are considered, and exact finite-sample critical values are tabulated. The power of the tests against fractionally-integrated alternatives, which are argued to have interesting variance-time function interpretations and potential importance in economics, is evaluated.

1. Introduction

Consider an observed time series $\{x_t\}_{t=0}^{T}$. Denote the variances of the k-th differenced series $\{\Delta_1 x\}, \{\Delta_2 x\}, \ldots, \{\Delta_k x\}, \ldots, \{\Delta_K x\}$ by $\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2, \ldots, \sigma_K^2$, respectively. Then, under the random walk null hypothesis, we have:

$$\sigma_2^2 = 2\sigma_1^2$$
$$\sigma_3^2 = 3\sigma_1^2$$
$$\cdots$$
$$\sigma_K^2 = K\sigma_1^2\,,$$

or:

$$\frac{2\sigma_1^2}{\sigma_2^2} = \frac{3\sigma_1^2}{\sigma_3^2} = \cdots = \frac{K\sigma_1^2}{\sigma_K^2} = 1\,.$$

It is well known that if a time series follows a random walk, then the variance of its k-th difference is a linear function of k, i.e., the variance-time function grows linearly. Conversely, if a time series is white noise, then the variance-time function is horizontal at $2\sigma^2$. It may also be shown that the properties of the variance-time functions of random walk and white noise processes also hold for I(1) and I(0) processes, respectively, for large k, as emphasized in Cochrane (1988).[1] In other words, I(1) (e.g., ARIMA) processes have variance-time functions which eventually grow linearly in k, and I(0) (e.g., ARMA) processes have variance-time functions which become flat.

These facts have been exploited at least since Working (1949) in attempts to determine the nature of asset price fluctuations. More recently, authors such as Campbell and Mankiw (1987), Cochrane (1988), Fama and French (1988), Huizinga (1986), Lo and MacKinlay (1988), and Poterba and Summers (1987) have used the variance-time function or closely related constructs to examine long-run mean reversion in both real and financial variables. Explicit hypothesis tests regarding the shape of this function have been proposed and used. The tests are nondirectional, in that they are not directed against a particular alternative, and the null hypotheses are simple, as opposed to composite. It is hoped that the nondirectional nature of the tests will yield power against a variety of (unspecified) alternatives.

The goals of this paper are modest. We focus on the random walk null hypothesis, as opposed to the more general null hypothesis of a unit root in a higher-ordered autoregressive lag-operator polynomial. While the random walk null is obviously too restrictive for some applications (e.g., explorations of the properties of aggregate output fluctuations), it may be quite appropriate for others, particularly those related to asset-price dynamics. We focus on similarly simple alternative hypotheses of fractional integration.

In Section 2 we introduce the class of fractionally-integrated time-series processes and study its properties in terms of the variance-time function, which can grow at increasing or decreasing rates. We motivate this result from a number of perspectives, note that it cannot be achieved with finite ARIMA representations, and argue that the fractionally-integrated process may be useful in macroeconomics and financial economics. In Section 3 we propose a *joint* test for random-walk behavior, which makes use of multiple points of the variance-time function, and we contrast it to the non-joint tests that have appeared in the literature. Exact finite-sample fractiles are tabulated under a normality assumption. It is hoped, of course, that the joint test will have greater power than its non-joint counterparts; this is investigated in Section 4, where the power properties of the various simple and joint tests are evaluated against a range of fractionally-integrated alternatives. The paper therefore extends the work of Lo and MacKinlay (1987), by providing a power evaluation of new as well as existing tests against what may prove to be a useful class of alternatives. In Section 5 we offer our conclusions.

2. Fractionally-Integrated ARIMA Processes

In this section we introduce the class of fractionally-integrated time-series models and provide a brief review of their properties in order to fix ideas and establish notation.[2] In the subsequent Monte Carlo power comparisons of scalar and joint tests of the variance-time function, the alternatives are fractionally integrated. This choice is not accidental; we argue that such processes possess long-memory properties likely to make them useful for modeling both real economic series — like aggregate output, and asset prices — like exchange rates. They provide generalized approximations to low-frequency components in economic time series; in particular, the knife-edged 'unit root' phenomenon arises as a special, and potentially restrictive, case. Consider a simple generalization of a random walk:

$$(1-L)^d x_t = \varepsilon_t , \tag{1}$$

where d takes values in the real, as opposed to integer, numbers. The process is stationary and invertible if $d \in (-1/2, 1/2)$; since d need not take integer value, we refer to the process as *fractionally integrated*. Clearly $d = 1$ yields a random walk, while $d = 0$ corresponds to white noise.[3]

We call the process (1) a *pure fractional noise*, in order to distinguish it from its natural generalization - the fractionally integrated ARIMA (ARFIMA) process:

$$\Phi(L)(1-L)^d x_t = \Theta(L)\varepsilon_t \tag{2}$$

where:

$$\Phi(L) = 1 - \phi_1 L - \cdots - \phi_p L^p$$

$$\Theta(L) = 1 - \theta_1 L - \cdots - \theta_q L^q$$

$$\varepsilon_t \sim (0, \sigma_\varepsilon^2),$$

all roots of $\Phi(L)$ and $\Theta(L)$ lie outside the unit circle, and d takes values in the real numbers. In this paper we are concerned only with deviations from random walk behavior that can be represented as pure fractional noise; such alternatives are natural against the random-walk null. The fractional difference operator $(1-L)^d$ may be expanded as:

$$(1-L)^d = 1 - dL + \frac{d(d-1)}{2!}L^2 - \frac{d(d-1)(d-2)}{3!}L^3 + \cdots \tag{3}$$

Econometricians typically have considered only integer values of d; writing the model as in (1), however, makes clear the arbitrary nature of the integer d restriction. The possibility of a graduated range of persistence effects may be entertained in a natural (and general) way by allowing for fractional integration.[4]

The intuition of fractional integration emerges clearly in the frequency domain. A series $\{x_t\}$ displays long memory if its pseudo-spectrum increases without limit as angular frequency tends to zero:

$$\lim_{\lambda \to 0} f_x(\lambda) = \infty .$$

ARFIMA processes have pseudo-spectra that behave like λ^{-2d} as $\lambda \to 0$. I(1) processes emerge as a special case, corresponding to $d = 1$; their pseudo-spectra behave like λ^{-2} near the origin. Note, however, the wider range of spectral behavior near the origin that becomes possible when the 'integer d' restriction is relaxed, which gives the ARFIMA class the potential to provide superior approximations to low-frequency dynamics.

In the time domain, fractional integration imparts 'long memory,' which is associated with significant dependence between observations widely separated in time.[5] The usual ARMA process is a short-memory process, and the autocorrelations decline exponentially:

$$\rho_X(\tau) \sim r^\tau, \quad 0 < r < 1, \quad \tau \to \infty .$$

For the ARFIMA process (2), the autocorrelation function has a slower hyperbolic decline:

$$\rho_X(\tau) \sim \tau^{2d-1}, \quad \tau \to \infty .$$

Additional time-domain motivation is achieved by considering the behavior of expanding sums, which we denote S_T, of T contiguous observations on a pure fractionally integrated process. It is easy to see that the growth of the variance of such sums depends on d, such that

$$var(S_T) = O(T^{1+2d}) .$$

Thus, for example, if $d = 0$ so that S_T is a random walk, then the variance grows at the familiar rate $O(T)$.[6] This result has direct implications for the behavior of fractionally-integrated processes in terms of the variance-time function: its growth behavior is $O(k^{2d-1})$. Thus for example if $d < 1/2$, then the variance-time function becomes flat, while if $1/2 < d < 1$ or $1 < d < 3/2$, then the variance-time function eventually grows at decreasing and increasing rates, respectively.

The time-domain behavior of fractionally integrated processes is also nicely illustrated by the calibration of k-step-ahead prediction intervals, for increasing k. The behavior of such intervals for the common trend-stationary and difference-stationary cases is strikingly different and well known.[7] In particular, prediction intervals for trend-stationary processes eventually become constant around trend, while those for difference-stationary processes grow continuously around drift, at the rate $O(k^{1/2})$. The uncertainty associated with forecasts of fractionally integrated processes, on the other hand, grows at rate $O(k^{d-1/2})$, which can be faster or slower than the I(1) case, providing a natural generalization. The point is simply that while many economic series do appear to have long memory, it needn't be associated with a unit autoregressive root.[8] Thus, the uncertainty associated with k-step ahead forecasts might reasonably be expected to be continuously increasing in k, but at a rate different from that associated with a unit root.

3. Joint Diagnostic Tests Based on the Variance-Time Function

The variance-time function can be exploited to obtain tests for random walk behavior. Simple scalar asymptotic tests of the individual points of the variance-time function:

$$\sigma_k^2 = k\sigma_1^2 , \quad k = 1, 2, \ldots, K \tag{4}$$

under the null hypothesis of linearity have recently been proposed by Cochrane (1988) and Lo and MacKinlay (1988), and their finite-sample distributions have been tabulated under a normality assumption by Diebold (1988). The test statistics are given by:

$$R(k) = \frac{k\hat{\sigma}_1^2}{\hat{\sigma}_k^2} , \quad k = 2, 3, \ldots, K .$$

If drift is assumed to be zero we use:

$$\hat{\sigma}_k^2 = \frac{1}{(T-k+1)} \sum_{t=k}^{T} (x_t - x_{t-k})^2 , \quad k = 1, 2, \ldots, K , \tag{5}$$

and denote the resulting statistic $R1(k)$; if drift is estimated we use:

$$\hat{\sigma}_k^2 = \frac{1}{(T-k+1)} \sum_{t=k}^{T} (x_t - x_{t-k} - k\hat{\mu})^2 , \quad k = 1, 2, \ldots, K , \tag{6}$$

where:

$$\hat{\mu} = \frac{1}{T} \sum_{t=1}^{T} (x_t - x_{t-1}) ,$$

and we denote the resulting statistic $R2(k)$. Fractiles of $R1(k)$ are presented in Table I for various (T, k) combinations, and fractiles of $R2(k)$ for the same (T, k) combinations are given in Diebold (1988) and reproduced for convenience in Table II.[9] The sample size, T, in all tables corresponds to the number of available first-differenced observations. Thus, the 'levels' sample contains $T+1$ observations. The tabulated critical values correspond to use of differencing intervals of $k = 1, 2, 4, 8, 16$ and 32; sample sizes were accordingly chosen to be divisible by 32. Interpolation may be used to obtain critical values for other sample sizes. Note also that the variance estimators in this paper are not corrected for finite-sample bias; for our purposes, since we are tabulating exact finite-sample distributions, such corrections are unnecessary.

TABLE I. Fractiles of *R* 1

R1(2)										
FRACTILES:	0.005	0.025	0.050	0.100	0.250	0.750	0.900	0.950	0.975	0.995
T:										
64	0.756	0.799	0.825	0.860	0.921	1.092	1.193	1.260	1.321	1.461
96	0.790	0.831	0.855	0.881	0.935	1.075	1.152	1.201	1.246	1.339
128	0.815	0.854	0.870	0.897	0.943	1.064	1.128	1.172	1.212	1.303
160	0.827	0.865	0.884	0.907	0.949	1.058	1.115	1.149	1.184	1.251
192	0.843	0.876	0.893	0.915	0.952	1.053	1.104	1.137	1.166	1.231
256	0.861	0.889	0.907	0.924	0.959	1.044	1.089	1.115	1.138	1.188
512	0.898	0.919	0.932	0.946	0.970	1.031	1.059	1.078	1.094	1.125
1024	0.927	0.942	0.950	0.961	0.979	1.021	1.042	1.055	1.068	1.093
2048	0.946	0.959	0.965	0.972	0.985	1.015	1.029	1.038	1.045	1.062
4096	0.961	0.970	0.975	0.980	0.990	1.011	1.021	1.027	1.033	1.043
R1(4)										
FRACTILES:	0.005	0.025	0.050	0.100	0.250	0.750	0.900	0.950	0.975	0.995
T:										
64	0.580	0.654	0.699	0.758	0.865	1.197	1.400	1.547	1.709	2.030
96	0.634	0.709	0.748	0.798	0.888	1.157	1.308	1.413	1.512	1.727
128	0.672	0.744	0.777	0.819	0.901	1.133	1.259	1.347	1.440	1.606
160	0.706	0.760	0.792	0.834	0.911	1.119	1.228	1.301	1.370	1.505
192	0.720	0.778	0.810	0.848	0.917	1.108	1.208	1.273	1.338	1.444
256	0.749	0.806	0.833	0.867	0.928	1.088	1.175	1.230	1.281	1.383
512	0.815	0.858	0.878	0.902	0.947	1.059	1.119	1.154	1.186	1.248
1024	0.865	0.894	0.909	0.929	0.963	1.041	1.082	1.106	1.130	1.176
2048	0.901	0.925	0.936	0.949	0.973	1.028	1.056	1.072	1.088	1.114
4096	0.927	0.946	0.953	0.963	0.981	1.021	1.039	1.050	1.062	1.082
R1(8)										
FRACTILES:	0.005	0.025	0.050	0.100	0.250	0.750	0.900	0.950	0.975	0.995
T:										
64	0.426	0.522	0.581	0.655	0.813	1.376	1.775	2.069	2.379	3.075
96	0.497	0.584	0.636	0.710	0.842	1.285	1.566	1.772	1.976	2.465
128	0.546	0.629	0.676	0.734	0.858	1.232	1.460	1.618	1.785	2.128
160	0.584	0.651	0.699	0.758	0.869	1.207	1.399	1.528	1.655	1.956
192	0.607	0.677	0.720	0.775	0.878	1.182	1.359	1.482	1.597	1.825
256	0.644	0.713	0.750	0.800	0.894	1.152	1.300	1.397	1.494	1.673
512	0.730	0.787	0.816	0.853	0.919	1.099	1.198	1.264	1.319	1.418
1024	0.792	0.839	0.863	0.890	0.942	1.066	1.133	1.175	1.214	1.285
2048	0.848	0.881	0.900	0.922	0.958	1.046	1.091	1.117	1.142	1.191
4096	0.888	0.915	0.929	0.943	0.970	1.033	1.061	1.080	1.097	1.129
R1(16)										
FRACTILES:	0.005	0.025	0.050	0.100	0.250	0.750	0.900	0.950	0.975	0.995
T:										
64	0.296	0.385	0.450	0.543	0.765	1.748	2.567	3.254	3.992	5.871
96	0.366	0.461	0.525	0.607	0.795	1.518	2.074	2.473	2.943	3.985
128	0.414	0.504	0.560	0.641	0.812	1.410	1.840	2.132	2.471	3.218
160	0.451	0.535	0.592	0.667	0.827	1.352	1.703	1.959	2.210	2.828
192	0.481	0.566	0.621	0.694	0.835	1.314	1.618	1.833	2.056	2.541
256	0.528	0.612	0.659	0.724	0.853	1.257	1.507	1.676	1.834	2.204
512	0.631	0.703	0.742	0.792	0.890	1.162	1.318	1.424	1.514	1.736
1024	0.714	0.770	0.804	0.844	0.915	1.106	1.204	1.277	1.338	1.460
2048	0.783	0.830	0.853	0.886	0.940	1.072	1.140	1.181	1.221	1.303
4096	0.842	0.877	0.896	0.919	0.956	1.049	1.094	1.123	1.148	1.200

R1(32)

FRACTILES: T:	0.005	0.025	0.050	0.100	0.250	0.750	0.900	0.950	0.975	0.995
64	0.179	0.263	0.333	0.449	0.742	2.751	4.802	6.534	8.500	13.567
96	0.255	0.337	0.405	0.506	0.756	2.045	3.258	4.218	5.340	7.991
128	0.288	0.385	0.449	0.548	0.769	1.781	2.646	3.356	4.072	5.592
160	0.330	0.421	0.484	0.579	0.778	1.635	2.340	2.860	3.441	4.871
192	0.354	0.453	0.509	0.601	0.785	1.549	2.135	2.581	3.080	4.218
256	0.404	0.500	0.556	0.632	0.807	1.433	1.890	2.235	2.586	3.442
512	0.522	0.606	0.655	0.718	0.850	1.256	1.523	1.698	1.893	2.268
1024	0.626	0.691	0.734	0.786	0.887	1.167	1.327	1.433	1.520	1.724
2048	0.704	0.768	0.801	0.843	0.915	1.110	1.214	1.280	1.339	1.468
4096	0.781	0.826	0.855	0.885	0.938	1.073	1.140	1.184	1.228	1.303

TABLE II. Fractiles of $R2$

R2(2)

FRACTILES: T:	0.005	0.025	0.050	0.100	0.250	0.750	0.900	0.950	0.975	0.995
64	0.767	0.810	0.837	0.871	0.934	1.112	1.217	1.286	1.353	1.486
96	0.799	0.839	0.862	0.890	0.944	1.087	1.164	1.217	1.264	1.356
128	0.821	0.858	0.877	0.903	0.950	1.073	1.138	1.183	1.227	1.313
160	0.830	0.869	0.889	0.912	0.954	1.064	1.122	1.157	1.193	1.256
192	0.847	0.879	0.897	0.919	0.957	1.058	1.112	1.144	1.173	1.238
256	0.864	0.892	0.910	0.928	0.962	1.049	1.093	1.119	1.144	1.194
512	0.901	0.921	0.934	0.948	0.972	1.033	1.062	1.080	1.096	1.128
1024	0.928	0.943	0.951	0.961	0.980	1.023	1.043	1.056	1.069	1.093
2048	0.947	0.959	0.965	0.973	0.985	1.016	1.029	1.038	1.046	1.063
4096	0.962	0.970	0.975	0.980	0.990	1.011	1.021	1.028	1.033	1.043

R2(4)

FRACTILES: T:	0.005	0.025	0.050	0.100	0.250	0.750	0.900	0.950	0.975	0.995
64	0.605	0.680	0.732	0.794	0.906	1.264	1.487	1.635	1.806	2.202
96	0.651	0.730	0.770	0.820	0.916	1.193	1.353	1.467	1.573	1.785
128	0.684	0.759	0.794	0.837	0.923	1.162	1.294	1.387	1.481	1.653
160	0.716	0.773	0.807	0.847	0.928	1.141	1.254	1.328	1.400	1.542
192	0.728	0.788	0.823	0.861	0.931	1.127	1.229	1.294	1.361	1.474
256	0.757	0.815	0.842	0.877	0.939	1.101	1.187	1.244	1.296	1.407
512	0.821	0.863	0.882	0.907	0.953	1.065	1.126	1.162	1.194	1.258
1024	0.867	0.895	0.912	0.932	0.965	1.044	1.085	1.110	1.133	1.180
2048	0.901	0.926	0.937	0.951	0.974	1.030	1.058	1.074	1.090	1.117
4096	0.928	0.946	0.954	0.964	0.982	1.022	1.040	1.051	1.062	1.083

R2(8)

FRACTILES T	0.005	0.025	0.050	0.100	0.250	0.750	0.900	0.950	0.975	0.995
64	0.477	0.580	0.646	0.730	0.914	1.569	2.035	2.347	2.692	3.581
96	0.525	0.631	0.685	0.759	0.906	1.399	1.706	1.937	2.150	2.639
128	0.574	0.663	0.710	0.773	0.906	1.306	1.558	1.742	1.915	2.307
160	0.601	0.679	0.726	0.788	0.908	1.264	1.465	1.609	1.741	2.056
192	0.623	0.701	0.745	0.803	0.912	1.228	1.417	1.542	1.657	1.902
256	0.653	0.732	0.771	0.821	0.918	1.186	1.337	1.437	1.534	1.727
512	0.738	0.796	0.826	0.862	0.932	1.117	1.215	1.284	1.338	1.445
1024	0.796	0.844	0.869	0.897	0.948	1.075	1.141	1.184	1.222	1.295
2048	0.849	0.884	0.903	0.925	0.961	1.050	1.095	1.122	1.148	1.197
4096	0.890	0.916	0.930	0.945	0.972	1.035	1.063	1.082	1.099	1.132

R2(16)

FRACTILES:	0.005	0.025	0.050	0.100	0.250	0.750	0.900	0.950	0.975	0.995
T:										
64	0.381	0.507	0.588	0.712	1.007	2.313	3.374	4.219	5.082	6.832
96	0.419	0.540	0.612	0.720	0.939	1.829	2.463	2.988	3.549	4.839
128	0.454	0.556	0.629	0.723	0.919	1.620	2.101	2.454	2.833	3.646
160	0.479	0.585	0.644	0.732	0.908	1.501	1.891	2.170	2.448	3.116
192	0.509	0.614	0.670	0.750	0.906	1.431	1.758	2.007	2.244	2.777
256	0.552	0.643	0.696	0.767	0.906	1.338	1.602	1.783	1.950	2.343
512	0.651	0.720	0.762	0.816	0.915	1.195	1.365	1.470	1.564	1.810
1024	0.723	0.781	0.815	0.856	0.930	1.125	1.227	1.300	1.360	1.478
2048	0.789	0.835	0.858	0.893	0.947	1.081	1.149	1.190	1.231	1.315
4096	0.843	0.881	0.899	0.922	0.960	1.052	1.098	1.126	1.151	1.204

R2(32)

FRACTILES:	0.005	0.025	0.050	0.100	0.250	0.750	0.900	0.950	0.975	0.995
T:										
64	0.455	0.632	0.745	0.939	1.460	4.412	7.281	9.603	11.926	17.481
96	0.362	0.508	0.604	0.755	1.133	3.008	4.500	5.669	6.947	9.988
128	0.366	0.488	0.584	0.714	1.015	2.386	3.492	4.286	5.217	7.138
160	0.384	0.509	0.585	0.707	0.959	2.048	2.909	3.578	4.239	5.838
192	0.411	0.531	0.609	0.712	0.940	1.864	2.561	3.120	3.657	4.987
256	0.452	0.560	0.626	0.716	0.914	1.648	2.155	2.571	2.945	3.886
512	0.554	0.640	0.694	0.759	0.903	1.346	1.629	1.824	2.025	2.440
1024	0.639	0.710	0.757	0.809	0.912	1.207	1.376	1.480	1.570	1.801
2048	0.712	0.781	0.814	0.855	0.929	1.128	1.232	1.299	1.363	1.482
4096	0.786	0.830	0.860	0.892	0.945	1.082	1.150	1.193	1.234	1.315

Due to the potential power advantages of a joint test of:

$$\frac{2\sigma_1^2}{\sigma_2^2} = \frac{3\sigma_1^2}{\sigma_3^2} = \cdots = \frac{K\sigma_1^2}{\sigma_K^2} = 1\,, \tag{7}$$

as opposed to the sequence of component 'scalar' tests (4), availability of a joint test may be useful. It is well known (e.g., Snedecor and Cochrane, 1980) that the negative of twice the log likelihood ratio for testing equality of variances from independent samples is given by:

$$J \equiv S \ln \hat{\sigma}^2 - \sum_{k=1}^{K} \left[\Psi_k \ln \left(\frac{\hat{\sigma}_k^2}{k} \right) \right],$$

where:

$$\Psi_k = F\left(\frac{T}{k}\right)$$

$$S = \sum_{k=1}^{K} \Psi_k$$

$$\hat{\sigma}^2 = \frac{1}{S} \sum_{k=1}^{K} \left[\Psi_k \left(\frac{\hat{\sigma}_k^2}{k} \right) \right],$$

$\hat{\sigma}_k^2$ is as defined earlier in (5) or (6), depending on whether drift is estimated, and $F(\cdot)$ rounds down to the nearest integer. Under the null (7) and independence of the samples from which the $\hat{\sigma}_k^2$ are calculated, $J \xrightarrow{d} \chi^2_{K-1}$. In the present context, this limiting result is invalid due to sample dependence. The null distribution is easily characterized (in small as well as large samples) by

Monte Carlo, however; tabulations corresponding to the cases of zero drift and estimated drift appear in Tables III and IV. In accordance with earlier notation, they are denoted $J1$ and $J2$. The tabulated critical values for the J statistics are for joint tests of five points on the variance-time function corresponding to use of differencing intervals of $k = 1, 2, 4, 8, 16$ and 32.[10]

TABLE III. Fractiles of $J1$

FRACTILES:	0.005	0.025	0.050	0.100	0.250	0.750	0.900	0.950	0.975	0.995
T:										
64	0.079	0.172	0.248	0.369	0.727	2.809	4.794	6.272	7.776	11.103
96	0.067	0.146	0.222	0.340	0.650	2.678	4.621	6.070	7.621	11.121
128	0.072	0.150	0.213	0.323	0.639	2.606	4.559	6.008	7.542	10.990
160	0.067	0.137	0.203	0.317	0.635	2.633	4.505	6.005	7.673	11.206
192	0.066	0.136	0.205	0.314	0.634	2.592	4.531	6.090	7.578	11.574
256	0.065	0.136	0.201	0.308	0.598	2.600	4.492	5.930	7.533	11.134
512	0.061	0.128	0.184	0.293	0.582	2.494	4.375	5.873	7.491	11.292
1024	0.051	0.126	0.183	0.286	0.580	2.460	4.440	6.027	7.587	11.179
2048	0.056	0.130	0.184	0.279	0.557	2.453	4.352	5.941	7.577	11.970
4096	0.053	0.116	0.177	0.286	0.564	2.371	4.292	5.897	7.438	11.152

TABLE IV. Fractiles of $J2$

FRACTILES:	0.005	0.025	0.050	0.100	0.250	0.750	0.900	0.950	0.975	0.995
T:										
64	0.086	0.187	0.281	0.429	0.874	3.513	5.579	7.064	8.548	12.382
96	0.071	0.161	0.242	0.382	0.776	3.244	5.259	6.920	8.647	12.006
128	0.076	0.151	0.222	0.354	0.734	3.094	5.262	6.815	8.599	12.051
160	0.073	0.153	0.232	0.352	0.714	3.046	5.116	6.768	8.291	11.954
192	0.066	0.149	0.220	0.335	0.683	2.897	4.977	6.686	8.301	12.277
256	0.069	0.147	0.209	0.326	0.658	2.786	4.917	6.487	8.120	11.711
512	0.064	0.130	0.196	0.311	0.610	2.603	4.628	6.137	7.842	11.986
1024	0.052	0.121	0.183	0.284	0.587	2.571	4.542	6.181	7.830	11.275
2048	0.057	0.126	0.185	0.278	0.562	2.496	4.425	5.993	7.630	11.985
4096	0.055	0.114	0.177	0.288	0.563	2.387	4.280	5.901	7.496	11.267

4. Monte Carlo Power Evaluation

We consider fractionally-integrated data-generating processes, as discussed above. We use d values of 0.3,0.45,0.6,0.7,0.8,0.9,1.0,1.1,1.2, and 1.3. In all cases the innovation variance σ_ε^2 is held fixed at 1.0. Sample sizes of $T = 64, 128$ and 256 are examined, with $N = 1000$ replications performed for each sample size.

Samples of size T from the ARFIMA process (1) with $d = 0.3$ and $d = 0.45$ (stationary parameter configurations) are formed as follows. First, a vector, v, consisting of $T\, N(0, 1)$ deviates is generated using IMSL subroutine GGNML. Then the desired $T \times T$ data covariance matrix (Σ) is constructed. This is simply the Toeplitz matrix formed from the autocovariances, which are given by:

$$\gamma_X(\tau) = \frac{\Gamma(1-2d)\Gamma(d+\tau)}{\Gamma(d)\Gamma(1-d)\Gamma(1-d+\tau)}\sigma_\varepsilon^2,$$

where $\Gamma(\cdot)$ is the gamma function. We next obtain the Choleski factorization of Σ, $\Sigma = PP'$, where P is lower triangular, using IMSL subroutine LUDECP. Finally the sample, x, is generated as

$x = Pv$.[11] Construction of x in this way eliminates dependence on presample startup values, which can be particularly problematic with long-memory models. For the nonstationary parameter configurations $d = 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2$, and 1.3, we generate fractional noise with parameter $d^* = d - 1$, which yields observations on $(1-L)x_t$. Then, taking $x_0 = 0$, we construct the sample $\{x_1, \ldots, x_T\}$ by cumulating.

TABLE V. Sample Powers of Variance-Time Function Tests, Two-Tailed, No Estimated Drift

	d									
	0.3	0.45	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3
T=64										
R1(2)	95,98	85,91	64,76	45,58	26,36	10,18	05,10	15,23	47,57	81,87
R1(4)	01,01	97,99	81,90	59,73	30,47	11,20	05,11	18,27	52,63	83,88
R1(8)	01,01	99,01	84,93	59,74	30,45	11,19	05,11	19,27	50,60	79,85
R1(16)	01,01	98,99	76,88	48,64	23,36	09,16	05,10	16,24	42,51	70,76
R1(32)	01,01	91,97	57,74	34,48	17,27	08,14	05,11	13,20	32,40	55,63
J1	01,01	99,01	83,92	54,68	24,36	09,16	05,10	12,18	34,42	65,72
T=128										
R1(2)	01,01	99,01	92,96	74,86	45,58	17,27	05,10	26,32	74,79	98,98
R1(4)	01,01	01,01	99,01	92,97	61,75	22,33	05,10	29,39	77,84	98,98
R1(8)	01,01	01,01	01,01	94,98	65,77	24,35	05,11	29,37	75,82	95,97
R1(16)	01,01	01,01	99,01	90,96	56,70	17,28	06,11	26,36	66,73	91,94
R1(32)	01,01	01,01	97,99	78,87	41,54	14,22	06,11	22,29	56,64	81,85
J1	01,01	01,01	01,01	93,97	58,72	18,28	06,11	21,29	61,68	91,93
T=256										
R1(2)	01,01	01,01	01,01	01,01	77,86	33,43	05,10	41,53	95,98	01,01
R1(4)	01,01	01,01	01,01	01,01	90,95	39,50	05,09	53,62	97,98	01,01
R1(8)	01,01	01,01	01,01	01,01	92,96	35,50	04,09	51,62	95,97	01,01
R1(16)	01,01	01,01	01,01	01,01	89,95	31,43	05,10	46,55	91,94	99,01
R1(32)	01,01	01,01	01,01	01,01	78,89	23,35	04,09	36,44	81,86	97,98
J1	01,01	01,01	01,01	01,01	91,96	31,43	05,09	38,46	91,93	99,01

Note: Each cell of the table has two entries, separated by a comma. The first is estimated power for a 5% level test, while the second is estimated power for a 10% level test. The data-generating process is:

$$\Delta x_t = (1-L)^{d-1}\varepsilon_t \qquad \varepsilon_t \overset{iid}{\sim} N(0, 1).$$

At each Monte Carlo replication $i = 1, \ldots, N$, the processes corresponding to each of the various d values are constructed using the same vector v of random numbers, to aid in variance-reduction. Each test at each replication is assigned a 1 if rejection occurs, and 0 otherwise. After completion of the N replications the power estimates are computed as the relative rejection frequencies. The power estimates are asymptotically normally distributed around the true power p, with variance $p(1-p)/N$; thus, $\pm 1.96\,[\hat{p}(1-\hat{p})/N]^{1/2}$ provides an estimate of the approximate 95 percent confidence interval.[12]

Estimated powers for two-tailed $R1$ and $J1$ tests are presented in Table V, for which the true data-generating process (DGP) is:[13]

$$\Delta x_t = (1-L)^{d-1}\varepsilon_t \qquad (8)$$
$$\varepsilon_t \overset{iid}{\sim} N(0, 1).$$

Note first that power equals nominal size under the null of $d = 1$ for all tests and sample sizes, which must be the case since we are using exact finite-sample critical values. The power curves of all tests are asymmetric around $d = 1$; power grows more quickly for $d > 1$ than for $d < 1$. Power of all tests grows rapidly with sample size as well.

For each sample size, a consistent power pattern emerges for the five two-sided $R1(\cdot)$ tests. Power is generally highest for $R1(4)$ or $R1(8)$, with $R1(2)$, $R1(16)$ and $R1(32)$ displaying somewhat less power. The power of the two-sided joint test $J1$ is generally less than that of the best $R1$ statistic, but greater than that of the worst $R1$ statistic.

TABLE VI. Sample Powers of Variance-Time Function Tests, Lower-Tailed, No Estimated Drift

	d									
	0.3	0.45	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3
T=64										
R1(2)	00,00	00,00	00,00	00,00	00,00	01,02	05,10	22,35	57,69	87,92
R1(4)	00,00	00,00	00,00	00,00	00,00	00,01	05,11	26,38	62,73	88,93
R1(8)	00,00	00,00	00,00	00,00	00,00	00,01	05,10	26,37	60,69	85,90
R1(16)	00,00	00,00	00,00	00,00	00,00	00,01	04,10	23,32	50,61	76,83
R1(32)	00,00	00,00	00,00	00,00	00,00	00,02	05,10	18,27	40,50	62,71
J1	00,00	00,00	00,00	00,00	00,00	03,06	05,10	04,07	01,02	00,01
T=128										
R1(2)	00,00	00,00	00,00	00,00	00,00	00,01	04,09	32,48	79,86	98,99
R1(4)	00,00	00,00	00,00	00,00	00,00	00,01	05,09	39,51	84,90	98,99
R1(8)	00,00	00,00	00,00	00,00	00,00	00,00	05,10	36,49	81,88	97,98
R1(16)	00,00	00,00	00,00	00,00	00,00	00,01	05,10	34,43	73,81	94,96
R1(32)	00,00	00,00	00,00	00,00	00,00	00,00	05,10	28,40	64,72	85,89
J1	00,00	00,00	00,00	00,00	00,00	02,05	05,08	02,05	00,01	00,00
T=256										
R1(2)	00,00	00,00	00,00	00,00	00,00	00,00	06,10	53,63	98,99	01,01
R1(4)	00,00	00,00	00,00	00,00	00,00	00,00	05,10	62,73	98,99	01,01
R1(8)	00,00	00,00	00,00	00,00	00,00	00,00	04,09	61,74	97,99	01,01
R1(16)	00,00	00,00	00,00	00,00	00,00	00,00	04,09	54,66	94,97	01,01
R1(32)	00,00	00,00	00,00	00,00	00,00	00,00	04,08	44,56	86,91	98,99
J1	00,00	00,00	00,00	00,00	01,02	00,00	05,09	01,02	00,00	00,00

Note: Each cell of the table has two entries, separated by a comma. The first is estimated power for a 5% level test, while the second is estimated power for a 10% level test. The data-generating process is:

$$\Delta x_t = (1-L)^{d-1}\varepsilon_t \qquad \varepsilon_t \overset{iid}{\sim} N(0, 1).$$

The above results may appear to bode poorly for the $J1$ test, but such is not the case. In the absence of prior information on the nature of allowable deviations of d from 1.0, the $R1(\cdot)$ tests should be used in two-tailed mode, but $J1$ should always be used as a one-tailed test (specifically, upper-tailed), which yields considerable power gains. The intuition is straightforward. By virtue of the definitions of the $R1(\cdot)$ and $J1$ statistics, one-sided lower-tailed $R1(\cdot)$ tests will have power *only* against alternatives for which $d > 1$, and one-sided upper-tailed $R1(\cdot)$ tests will have power *only* against alternatives for which $d < 1$. Conversely, one-sided upper-tailed $J1$ tests will have power against *all* alternatives.[14] These assertions are clearly illustrated in Tables VI and VII. The lower-tailed $R1(\cdot)$ tests in Table VI have no power against alternatives for which $d < 1$, and the lower-tailed $J(\cdot)$ tests have no power against *any* alternatives. The upper-tailed $R1(\cdot)$ tests of Table VII have no power against alternatives for which $d > 1$, while the upper-tailed $J1$ tests have power against *all* alternatives. Comparison of the power of upper-tailed $J1$ tests and two-tailed $R1$ tests reveals the superiority of the joint test, for all d and T values.[15] To solidify these ideas consider a representative case: $d = 0.7$ and $T = 64$. To test the null that d equals one against the alternative of d not equal to one, use of the the two-sided $R1(\cdot)$ tests and the one-sided upper-tailed

$J1$ test is appropriate. From Table V, we see that the power of the $R1(\cdot)$ tests (at the 5% level) ranges from 0.34 ($R1(32)$) to 0.59 ($R1(8)$), while Table VII reports the power of $J1$ as 0.68.

TABLE VII. Sample Powers of Variance-Time Function Tests, Upper-Tailed, No Estimated Drift

	d									
	0.3	0.45	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3
T=64										
R1(2)	98,99	91,96	76,87	58,73	36,52	17,28	05,10	01,03	00,00	00,00
R1(4)	01,01	99,01	90,96	73,85	47,63	20,32	06,11	01,03	00,00	00,00
R1(8)	01,01	01,01	93,98	74,87	45,61	19,31	06,11	01,03	00,01	00,00
R1(16)	01,01	99,01	88,96	64,81	36,53	16,27	06,11	02,04	00,01	00,00
R1(32)	01,01	97,99	74,88	48,66	27,42	14,23	06,11	02,05	01,02	00,01
J1	01,01	01,01	92,97	68,81	35,51	13,23	05,10	14,21	41,50	71,78
T=128										
R1(2)	01,01	01,01	96,99	86,93	58,72	27,39	05,11	00,01	00,00	00,00
R1(4)	01,01	01,01	01,01	97,99	75,87	33,49	05,11	00,01	00,00	00,00
R1(8)	01,01	01,01	01,01	98,99	77,89	35,48	06,10	00,02	00,00	00,00
R1(16)	01,01	01,01	01,01	96,98	70,83	28,44	06,11	02,03	00,00	00,00
R1(32)	01,01	01,01	01,01	87,96	54,72	22,36	06,11	01,03	00,01	00,00
J1	01,01	01,01	01,01	97,99	72,83	26,38	06,11	26,34	68,76	93,95
T=256										
R1(2)	01,01	01,01	01,01	98,01	86,92	43,56	05,10	00,00	00,00	00,00
R1(4)	01,01	01,01	01,01	01,01	95,98	50,66	04,09	00,00	00,00	00,00
R1(8)	01,01	01,01	01,01	01,01	96,99	50,65	05,09	00,00	00,00	00,00
R1(16)	01,01	01,01	01,01	01,01	95,99	43,60	05,10	00,00	00,00	00,00
R1(32)	01,01	01,01	01,01	01,01	89,95	35,52	05,10	00,00	00,00	00,00
J1	01,01	01,01	01,01	01,01	96,99	42,55	05,09	46,56	93,96	01,01

Note: Each cell of the table has two entries, separated by a comma. The first is estimated power for a 5% level test, while the second is estimated power for a 10% level test. The data-generating process is:

$$\Delta x_t = (1-L)^{d-1}\varepsilon_t \qquad \varepsilon_t \overset{iid}{\sim} N(0,1).$$

Variations on the above themes are explored in Tables VIII-XI.[16] The $R2(\cdot)$ and $J2$ are evaluated in Table VIII, while the DGP is still the no-drift model (8). As expected, there is a consistent (but very slight) power loss for all tests in Table VIII, since a drift has been needlessly estimated. The result is of practical importance: since only a *slight* power loss occurs, it is clear that little penalty is incurred when drift is needlessly estimated.

In Table IX, the true DGP displays drift:

$$\Delta x_t = 1.0 + (1-L)^{d-1}\varepsilon_t \qquad (9)$$

$$\varepsilon_t \overset{iid}{\sim} N(0,1),$$

but the no-drift $R1(\cdot)$ and $J1$ test statistics are used. It is at once apparent that severe penalties, in terms of departures of empirical from nominal test sizes, are incurred when drift is incorrectly assumed to be zero.

In Table X the power properties of the estimated-drift statistics $R2(\cdot)$ and $J2$ are evaluated for DGP (9); thus, the scenario corresponds to the correct inclusion of drift. As expected, power is qualitatively the same as in Table V, which corresponds to the correct exclusion of drift.[17]

Finally, in Table XI, the effects of violation of the normality assumption on empirical test size are investigated. The DGP is:

$$\Delta x_t = (1-L)^{d-1} \quad [\,sign(\varepsilon_t) * (\varepsilon_t^2)\,]$$

$$\varepsilon_t \overset{iid}{\sim} N(0,1),$$

which has leptokurtic (but symmetric) innovations. All tests appear quite robust; empirical size stays very close to nominal size.

TABLE VIII. Sample Powers of Variance-Time Function Tests, Two-Tailed, Estimated Drift

	d									
	0.3	0.45	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3
T=128										
R2(2)	01,01	99,01	88,95	70,82	42,56	13,24	05,09	21,30	66,73	95,97
R2(4)	01,01	01,01	98,99	88,93	54,70	17,29	05,10	25,35	73,81	96,97
R2(8)	01,01	01,01	99,01	88,94	54,69	16,27	05,09	24,33	68,76	92,95
R2(16)	01,01	01,01	98,99	83,91	43,60	14,22	04,09	19,28	54,64	82,87
R2(32)	01,01	01,01	90,96	64,78	31,45	09,17	04,10	15,23	41,51	65,73
J2	01,01	01,01	01,01	91,95	55,69	15,24	05,10	08,14	32,41	72,80

Note: Each cell of the table has two entries, separated by a comma. The first is estimated power for a 5% level test, while the second is estimated power for a 10% level test. The data-generating process is:

$$\Delta x_t = (1-L)^{d-1}\varepsilon_t \qquad \varepsilon_t \overset{iid}{\sim} N(0,1).$$

TABLE IX. Empirical Sizes of Variance-Time Function Tests, No Estimated Drift

	Upper-tailed	Lower-tailed	Two-sided
T=128			
R1(2)	00,00	01,01	01,01
R1(4)	00,00	01,01	01,01
R1(8)	00,00	01,01	01,01
R1(16)	00,00	01,01	01,01
R1(32)	00,00	01,01	01,01
J1	01,01	00,00	01,01

Note: Each cell of the table has two entries, separated by a comma. The first is estimated power for a 5% level test, while the second is estimated power for a 10% level test. The data-generating process is:

$$\Delta x_t = 1.0 + (1-L)^{d-1}\varepsilon_t \qquad \varepsilon_t \overset{iid}{\sim} N(0,1).$$

TABLE X. Sample Powers of Variance-Time Function Tests, Two-Tailed Tests, Estimated Drift

	d									
	0.3	0.45	0.6	0.7	0.8	0.9	1.0	1.1	1.2	1.3
T=128										
R2(2)	01,01	99,99	91,96	72,83	44,57	16,25	05,10	21,29	66,74	94,96
R2(4)	01,01	01,01	99,99	89,95	56,70	17,28	04,09	26,36	73,81	96,97
R2(8)	01,01	01,01	99,01	89,95	56,70	16,26	06,10	26,34	68,75	92,95
R2(16)	01,01	01,01	98,99	81,91	42,58	15,23	08,11	20,30	53,65	83,88
R2(32)	01,01	01,01	91,96	63,79	30,46	10,18	04,10	14,22	39,50	66,74
J2	01,01	01,01	01,01	91,95	55,69	15,26	05,10	07,13	30,41	72,80

Note: Each cell of the table has two entries, separated by a comma. The first is estimated power for a 5% level test, while the second is estimated power for a 10% level test. The data-generating process is:

$$\Delta x_t = 1.0 + (1-L)^{d-1}\varepsilon_t \qquad \varepsilon_t \overset{iid}{\sim} N(0,1).$$

TABLE XI. Empirical Sizes of Variance-Time Function Tests, No Estimated Drift

	Upper-tailed	Lower-tailed	Two-tailed
T=256			
R1(2)	06,10	05,10	06,11
R1(4)	05,11	05,10	06,10
R1(8)	06,11	04,10	05,10
R1(16)	05,11	04,08	05,10
R1(32)	05,10	05,10	04,09
J1	04,09	05,09	04,09

Note: Each cell of the table has two entries, separated by a comma. The first is estimated power for a 5% level test, while the second is estimated power for a 10% level test. The data-generating process is:
$\Delta x_t = (1-L)^{d-1}[\,sign(\varepsilon_t)*(\varepsilon_t^2)\,] \qquad \varepsilon_t \overset{iid}{\sim} N(0, 1)$.

5. Concluding Remarks

It is argued that the class of fractionally-integrated processes may prove useful in empirical economics, due to its ability to approximate a wide range of low-frequency dynamics, and the power properties of tests based on the variance-time function against fractionally-integrated alternatives are examined. All test comparisons are performed using exact finite-sample fractiles, which are presented in tabular form. A new joint test is proposed and found to be more powerful than currently popular tests based on scalar variance ratios. Finally, some preliminary evidence indicates that the variance-time tests may display robustness to fat-tailed innovations.

If a particular time series *does* in fact possess long memory, but not a unit root, it is natural to ask whether a researcher would be able to detect such deviation from unit-root behavior using conventional tests.[18] Formally, the problem amounts to determining the power properties of various unit-root tests against fractionally-integrated alternatives. One such test has been examined in the present paper, for the simplest null (random walk) against a very simple alternative (pure fractional integration). Others, such as the Dickey-Fuller tests and their relatives, are examined in Diebold and Rudebusch (1988b).

Finally, we note that the spectral procedure of Geweke and Porter-Hudak (1983) for estimating (posssibly) fractionally-integrated models holds promise as a unit root test against fractional as well as nonfractional alternatives. The semiparametric nature of the first-stage d estimate makes such an approach particularly attractive — consistent and asymptotically normal estimates of d are obtained *independent* of the potentially infinite-dimensional nuisance parameter in $\Phi^{-1}(L)\Theta(L)$.[19]

Notes

* The author expresses appreciation to two referees for constructive and useful comments. He would also like to thank Anil Bera, Rob Engle, Tarhan Feyzioglu, Roger Koenker, Jim Nason, Marc Nerlove, Fallaw Sowell, Pravin Trivedi, Jim Stock, Sam Yoo, and especially Glenn Rudebusch, all of whom provided useful input, but who are not to be held responsible for the outcome. The views expressed are solely those of the author; they do not necessarily reflect those of the Federal Reserve System or its staff.

1. I(1) and I(0) denote, respectively, integrated processes of order one and zero. I(1) processes are commonly referred to as 'difference stationary', or 'homogeneous nonstationary of order one,' and are made stationary by taking a first difference. The leading example of an I(1) process is the finite-ordered ARIMA, while the stationary finite-ordered ARMA is a commonly-encountered I(0) process.

For purposes of this paper, we use I(1) to denote a finite-ordered ARIMA $(p, 1, q)$ process with a positive, real, unit autoregressive root.

2. For in-depth treatments of various aspects of these processes, see Granger and Joyeux (1980), Hosking (1981), Geweke and Porter-Hudak (1983), Li and McLeod (1986), Brockwell and Davis (1987), Sowell (1987), and Diebold and Nerlove (1989), *inter alia*.
3. The range $1/2 < d < 3/2$ may be of special interest in economics, due to the local generalization of unit-root behavior that it permits. Such processes are stationary in first differences, since the first-differentiated series will be fractionally integrated of order d^*, where $-1/2 < d^* < 1/2$.
4. Note, however, that the spectrum of the first difference of a fractionally-integrated series is 0 if $d < 1$ and ∞ if $d > 1$. In the time domain, this corresponds to the fact that the sum of (infinitely many) coefficients of the moving-average representation of the first difference of a fractionally-integrated series is 0 if $d < 1$ and ∞ if $d > 1$. Thus, fractional integration allows for richer cumulative impulse-response effects only at finite horizons — but these are, of course, the horizons of greatest economic interest.
5. In particular, fractionally-integrated processes are *not* strong mixing. Thus, the assumptions underlying much of the asymptotic theory recently popularized in econometrics (e.g., White (1984), *inter alia*) do not hold. Under suitable regularity conditions, however, they are stationary, ergodic, regular, and weak mixing, so that (weaker) asymptotic results are available. See Graf, et al. (1984), Samarov and Taqqu (1988), and Gourieroux et al. (1987).
6. Stock (1988) develops a class of unit-root tests based upon these ideas.
7. See Dickey, Bell and Miller (1986) and Stock and Watson (1988) for nice expositions.
8. Furthermore, standard unit root tests (e.g., the Dickey-Fuller tests and their relatives) may have low power against fractionally-integrated alternatives, as argued by Diebold and Rudebusch (1988b).
9. The fractiles given in Diebold (1988) are based on 25000 replications and are therefore somewhat more accurate than those given here, which are based on 10000 replications. A detailed description of the procedures used to tabulate the various test statistics is given in the appendix.
10. Different sets of points on the variance-time function may be jointly tested by first temporally aggregating the data to the desired degree.
11. Quick calculation verifies that $cov(x) = PP' = \Sigma$.
12. For $N = 1000$, the maximum width of the estimated confidence interval (occurring when $p = 1/2$) is ± 0.03.
13. Use of the $R1$ and $J1$ tests exploits the knowledge that drift is not present; in practice such information may be uncommon. Alternative scenarios are subsequently explored, such as allowance for a nonexisting drift, or failure to allow for an existing drift.
14. This is analogous to the standard χ^2 test, which could be used if the samples were independent.
15. If, however, prior information is available indicating that a one-sided alternative (either $d < 1$ or $d > 1$) is appropriate, then maximal power may be attained by using the appropriate one-sided $R1(\bullet)$ test. Such prior information is rarely available.
16. Tables VIII-XI report results for two-sided tests only, for comparison with Table V. It should be kept in mind that, in practice, the $J1$ and $J2$ tests would never be used in two-sided mode. Our intent is comparison *of relative powers*, however, so that,for example, comparison of the power of $J2$ in Table VIII and $J1$ in Table V *does* convey useful information.
17. Power is slightly reduced in Table X, however, due to the loss of one degree of freedom in estimating the drift.
18. Note, for example, that the Dickey-Fuller tests allow for fractional integration neither under the null nor under the alternative.
19. For an application to aggregate output dynamics, see Diebold and Rudebusch (1988a).

Appendix: Details of Numerical Procedure

The simulation for each of the test statistics is executed as follows. Consider first $R1(k)$, the scalar test statistic for the no-drift case. A sample of T $N(0,1)$ deviates is generated by IMSL subroutine GGNML; these are the values of $\Delta_1 x_t$. A distributional assumption is of course necessary for finite-sample tabulation. In some economic contexts the normality assumption may be inappropriate, but judicious choice of sampling frequency will usually enable its approximate satisfaction. For example, while daily stock returns are known to be leptokurtic, monthly returns are approximately normal. Furthermore, as argued in the text, the test sizes appear robust to innovation non-normality.

The level series is obtained by cumulating the $\Delta_1 x_t$ series from an initial value of 0. Then $\hat{\sigma}_1^2$ is calculated, imposing the zero-drift restriction. Next, the data are k-th differenced and $\hat{\sigma}_k^2$ is calculated, again imposing the zero-drift restriction, and the test statistic is formed. This is repeated 10000 times, whereupon the resulting sequence of 10000 values of the test statistic is ordered and the fractiles extracted. This is repeated for all of the various (T,k) pairs that are tabulated. An identical procedure is followed when drift is allowed, except that the sample k-variances are for data centered around an estimated mean, as discussed in the text. The true (but unknown) mean, μ, is maintained at 0. The joint test statistics $J1$ and $j2$ are tabulated in similar fashion, using k values of 2, 4, 8, 16 and 32.

Precision of the fractile estimates may be evaluated using the well-known result (e.g., Rao, 1973) that the sample fractiles are asymptotically normal. Specifically, the p-th fractile of a distribution function F is any value δ_p such that:

$$P(x \leqq \delta_p) \geqq p$$

$$P(x \geqq \delta_p) \geqq q\,,$$

where $q = 1-p$. If $F(x)$ has a density function f continuous in x, δ_p is unique, and $f(\delta_p) > 0$, then:

$$n^{1/2}(\hat{\delta}_p - \delta_p) \xrightarrow{d} X \sim N\left(0,\ \frac{p(1-p)}{[f(\delta_p)]^2}\right),$$

where n is the number of replications upon which the fractile estimates are based. The fact that the asymptotic standard error depends on the height of the unknown density function f and δ_p is inconvenient, but f may be estimated by nonparametric methods in order to obtain estimated standard errors. Alternatively, nonasymptotic distribution-free fractile confidence intervals may be obtained as in David (1981) or Rohatgi (1984). Let $X_{(1)}, \ldots, X_{(n)}$ be the order statistic for the sample of replications, and let $x = \upsilon_p$ be the population fractile which we are attempting to estimate (i.e., $F(x) = p$, $0 < p < 1$, is uniquely solved by $x = \upsilon_p$). It may then be shown that:

$$P[\upsilon_p \varepsilon (X_{(r)}, X_{(s)})] \geqq \pi(r,s,n,p) = \sum_{i=r}^{s-1} \binom{n}{i} p^i (1-p)^{n-1}\,.$$

References

Brockwell, P. J., and R. A. Davis: 1987, *Time Series: Theory and Models*, Springer-Verlag, New York, New York, U.S.A.

Campbell, J. Y., and N. G. Mankiw: 1987, 'Permanent and Transitory Components in Macroeconomic Fluctuations', *American Economic Review*, **77**, 111-117.

Cochrane, J. H.: 1988, 'How Big is the Random Walk in GNP?', *Journal of Political Economy*, **96**, 893-920.

Cochrane, J. H.: 1987, 'Spectral Density Estimates of Unit Roots', manuscript, Department of Economics, University of Chicago, Chicago, Illinois, U.S.A.

David, H. A.: 1970, *Order Statistics*, John Wiley, New York, New York, U.S.A.

Dickey, D. A.: 1976, 'Estimation and Testing of Nonstationary Time Series', Ph.D. Dissertation, Department of Statistics, Iowa State University, Ames, Iowa, U.S.A.

Dickey, D. A., W. R. Bell, and R. B. Miller: 1986, 'Unit Roots in Time Series Models: Tests and Implications', *The American Statistician*, **40**, 12-26.

Diebold, F. X.: 1988, 'Testing for Bubbles, Reflecting Barriers and Other Anomalies', *Journal of Economic Dynamics and Control*, **12**, 63-70.

Diebold, F. X., and M. Nerlove: 1989, 'Unit Roots in Economic Time Series: A Selective Survey', forthcoming in T. B. Fomby and G. F. Rhodes, (eds.), *Advances in Econometrics: Co-Integration, Spurious Regressions, and Unit Roots*, JAI Press, Greenwich, Connecticut, U.S.A.

Diebold, F. X., and G. D. Rudebusch: 1988a, 'Long Memory and Persistence in Aggregate Output', Finance and Economics Discussion Series #7, Federal Reserve Board, Washington, D.C., U.S.A.

Diebold, F. X., and G. D. Rudebusch: 1988b, 'Why Unit Root Tests on Macroeconomic Variables May be Misleading', manuscript in progress, Federal Reserve Board, Washington, D.C., U.S.A.

Fama, E. F., and K. R. French: 1988, 'Permanent and Temporary Components of Stock Prices', *Journal of Political Economy*, **96**, 246-273.

Geweke, J., and S. Porter-Hudak: 1983, 'The Estimation and Application of Long Memory Time Series Models', *Journal of Time Series Analysis*, **4**, 221-238.

Gourieroux, C., F. Maurel, and A. Monfort: 1987, 'Regression and Nonstationarity', Working Paper #8708, INSEE, Paris, France.

Graf, H., F. R. Hampel, and J. D. Tacier: 1984, 'The Problem of Unsuspected Serial Correlations', in J. Franke, W. Hardle and D. Martin (eds.), *Robust and Nonlinear Time Series Analysis (Lecture Notes in Statistics, Volume 26)*, Springer-Verlag, New York, New York, U.S.A.

Granger, C. W. J.: 1981, 'Some Properties of Time Series Data and their Use in Econometric Model Specification', *Journal of Econometrics*, **16**, 121-130.

Granger, C. W .J., and R. Joyeux: 1980, 'An Introduction to Long-Memory Time Series Models and Fractional Differencing', *Journal of Time Series Analysis*, **1**, 15-39.

Hosking, J. R. M.: 1981, 'Fractional Differencing', *Biometrika*, **68**, 165-176.

Huizinga, J.: 1986, 'An Empirical Investigation of the Long-Run Behavior of Real Exchange Rates', manuscript, Department of Economics, University of Chicago, Chicago, Illinois, U.S.A.

Li, W. K., and A. E. McLeod: 1986, 'Fractional Time Series Modelling', *Biometrika*, **73**, 217-21.

Lo, A. W., and A. C. MacKinlay: 1987, 'The Size and Power of the Variance Ratio Test in Finite Samples: A Monte Carlo Investigation', Working Paper 28-87, Rodney L. White Center for Financial Research, Department of Finance, Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania, U.S.A. Forthcoming, *Journal of Econometrics*.

Lo, A. W., and A. C. MacKinlay: 1988, 'Stock Market Prices do not Follow Random Walks: Evidence From a Simple Specification Test', *Journal of Financial Research*, **1**, 41-66.

Mandelbrot, B. B.: 1972, 'Statistical Methodology for Nonperiodic Cycles: From the Covariance to R/S Analysis', *Annals of Economic and Social Measurement*, **1**, 259-290.

Poterba, J. M., and L. H. Summers: 1987, 'Mean Reversion in Stock Prices: Evidence and Implications', manuscript, Sloan School, M.I.T., and Department of Economics, Harvard University, Cambridge, Massachusetts, U.S.A.

Rao, C. R.: 1973, *Linear Statistical Inference and Its Applications*, John Wiley, New York, New York, U.S.A.

Rohatgi, V. K.: 1984, *Statistical Inference*, John Wiley, New York, New York, U.S.A.

Samarov, A,. and M. S. Taqqu: 1988, 'On the Efficiency of the Sample Mean in Long-Memory Noise', *Journal of Time Series Analysis*, **9**, 191-200.

Shea, G. S.: 1987, 'Long-Memory Models of Interest Rates: Estimation, Forecasting and Inference for Variance Bounds on the Interest Rate Term Structure', manuscript, Department of Finance, Pennsylvania State University, University Park, Pennsylvania, U.S.A.

Sowell, F. B.: 1987, 'Maximum-Likelihood Estimation of Fractionally Integrated Time-Series Models', Research Paper 87-07, Institute of Statistics and Decision Sciences, Duke University, Durham, North Carolina, U.S.A.

Stock, J. H.: 1988, 'A Class of Tests for Integration and Cointegration', manuscript, Kennedy School, Harvard University, Cambridge, Massachusetts, U.S.A.

Stock, J. H., and M. W. Watson: 1988, 'Variable Trends in Economic Time Series', *Economic Perspectives*, **2**, 147-174.

White, H.: 1984, *Asymptotic Theory for Econometricians*, Academic Press, New York, New York, U.S.A.

Working, H.: 1949, 'The Investigation of Economic Expectations', *American Economic Review*, **39**, 164-165.