Journal of Monetary Economics 29 (1992) 277–302. North-Holland

Modeling long-run behavior with the fractional ARIMA model

Fallaw Sowell*
Carnegie Mellon University, Pittsburgh, PA 15213, USA

Received December 1990, final version received July 1991

When modeling long-run behavior, fractional ARIMA models can give insights unobtainable with the nonfractional ARIMA models. As an application, the deterministic trend and unit root with drift models are nested in the fractional ARIMA model. This allows testing between the two models based on estimated parameter values. This test is applied to postwar US quarterly real GNP. The test concludes that GNP is consistent with both models. The estimated fractional parameter is significantly smaller than reported in Diebold and Rudebusch (1989). The difference is explained by bias in the previous estimates. Relationships with the cumulative impulse response and spectral density at frequency zero are noted.

1. Introduction

The parametric approach to investigating the long-run behavior of time series consists of estimating a parametric model for the series and relying on the long-run implications of the estimated model. The primary advantage is the precision gained by focusing the information in the series through the parameter estimates. A drawback is that the parameter estimates are sensitive to the class of models considered and may be misleading because of misspecification. The possibility of misspecification with parametric models can never be settled conclusively. However, the problem can be addressed by considering larger classes of models. This is the approach of the current paper. The fractional ARIMA model is used to model the long-run behavior of a time series. The fractional differencing parameter allows the capturing of

*I would like to thank Neil Ericsson, John Geweke, Albert Marcet, Jim Stock, Stan Zin, and an anonymous referee for helpful comments. I also benefited from discussion with seminar participants at Duke University, the Board of Governors of the Federal Reserve System, the Institute for Mathematics and its Applications conference on Time Series, and Queens University. Financial support by Carnegie Mellon University through a faculty development grant is gratefully acknowledged.

long-run behavior without problems commonly associated with ARMA models. As an example of its usefulness, the fractional ARIMA model is used to nest a deterministic trend model and a unit root with drift model, which allows testing between the two models based on the estimated fractional difference parameter. This test is then applied to postwar quarterly real GNP. It appears the data are consistent with both models.

The next section discusses inherent problems associated with modeling long-run behavior with ARMA models and how the fractional ARIMA model can avoid these problems. Section 3 explains how the fractional ARIMA model can nest the deterministic trend and the unit root with drift models. This allows testing between these models based on the estimated fractional differencing parameter. Section 4 considers the behavior of a fractionally integrated series at frequency zero. This highlights the difference between the fractional differencing parameter and estimates of the spectral density at frequency zero. In section 5 the proposed test is applied to postwar quarterly US real GNP. Econometric issues and the interpretation of the estimates are also discussed in section 5.

How the current results compare to previous results in the literature are discussed is section 6. Fractional ARIMA models for output were considered in Diebold and Rudebusch (1989) [D&R]. Unfortunately, the approach was limited by not considering the possibility that the output series followed a deterministic trend model.[1] In section 6, it is shown that the estimation procedure used in D&R produces significantly biased estimates for a reasonable parametric model of real GNP. This has two important implications. First, researchers modeling the long-run behavior of time series should realize that the procedure used in D&R can produce significant bias. Second, economists interested in the time series model for output should realize that the parameter estimates reported in D&R should be viewed with caution.

2. ARMA versus fractional ARIMA for modeling the long run

The most commonly used class of parametric time series models is the ARMA(p,q) model

$$\begin{aligned}&\left(1+\phi_1 L+\phi_2 L^2+\cdots+\phi_p L^p\right)x_t\\&=\left(1+\theta_1 L+\theta_2 L^2+\cdots+\theta_q L^q\right)\varepsilon_t,\end{aligned}$$

where it is assumed that the roots of $(1+\phi_1 z+\phi_2 z^2+\cdots+\phi_p z^p)$ and $(1+\theta_1 z+\theta_2 z^2+\cdots+\theta_q z^q)$ are outside the unit circle. Unfortunately, this

[1] The unit root was the only model considered. The hypothesis tests conducted were one-sided tests of a unit root null hypothesis.

class of models is not well suited to model the long-run behavior of time series. One problem is that the parameter values that capture the long run are near the boundary of the parameter space, where the asymptotic distributions are poor approximations to sampling distributions which makes inference difficult. To model positive long-run dependence, a root of the autoregressive polynomial must approach the unit circle. Similarly, to model negative long-run dependence, a root of the moving average polynomial must approach the unit circle.

A second problem in using ARMA models to capture the long-run behavior of a series is that if an AR or MA parameter does capture the long-run behavior of a series it imposes significant restrictions on the short-run behavior of the series. For example, it is not possible with a single AR parameter to model strong positive correlation at cycles of length greater than 15 years without modeling strong positive correlation at (say) cycles of length 10 years.[2] The same problem occurs in modeling negative long-run dependence in a series.

A third problem in relying on ARMA models to capture the long run concerns model selection in small samples. As pointed out in Cochrane (1988) maximum likelihood (asymptotically) chooses parameter values to minimize the difference between the periodogram of the data and the spectral density of the parametric model weighted at different frequencies. The result is that the fit of the long-run behavior of the series may be sacrificed to obtain a better fit of the short-run behavior. There is no way to direct the fit of an AR or an MA parameter to the long-run characteristics of a series, even though a researcher may be investigating long-run behavior.

One approach that avoids these problems is to rely on nonparametric estimations techniques. Unfortunately, given the sample sizes of most economic time series the nonparametric estimates are too imprecise to give meaningful restrictions. An alternative approach, considered in this paper, is to estimate a more general class of parametric models which is less susceptible to these problems.

It is possible to choose a parameterization where one parameter's primary effect concerns the long run and whose effect on the short run is limited. Such a class of models is the fractional ARIMA model. The general model will be written

$$\begin{aligned}&\left(1+\phi_1 L+\phi_2 L^2+\cdots+\phi_p L^p\right)(1-L)^d x_t\\&=\left(1+\theta_1 L+\theta_2 L^2+\cdots+\theta_q L^q\right)\varepsilon_t.\end{aligned}$$

[2]This problem can be reduced by considering a highly parameterized model. An MA root near the unit circle can offset some of the power at higher frequencies. However, this leads to a different problem of near root cancellation, and such a model is rarely judged optimal by a model selection criteria.

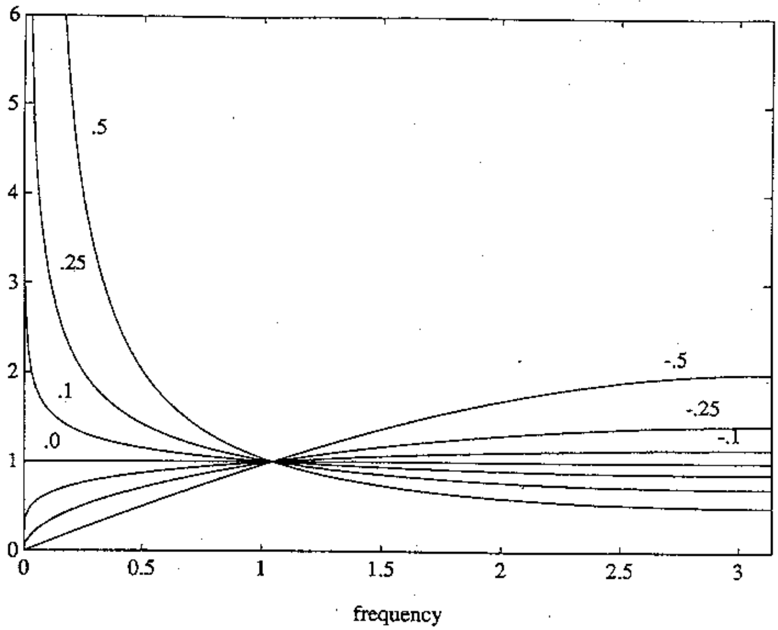

Fig. 1. Spectral density of the fractional ARIMA(0, d, 0) model $(1-L)^d x_t = \varepsilon_t$ for various value of d. The innovation variance is set to one.

It will be assumed that $d < 0.5$ and that the roots of $(1 + \phi_1 z + \phi_2 z^2 + \cdots + \phi_p z^p)$ and $(1 + \theta_1 z + \theta_2 z^2 + \cdots + \theta_q z^q)$ are outside the unit circle. In this model, the short-run behavior of the series can be captured by the ARMA parameters and the long-run behavior can be modeled by the fractional differencing parameter, which can reduce the problems discussed in the previous section. The ability of the fractional differencing parameter to model the long-run behavior of a time series was discussed in Diebold and Rudebusch (1989).

The fractional ARIMA(p, d, q) model is less susceptible to the problems encountered by the nonfractional ARMA(p, q) models. The model is nonstationary for $d \geq 0.5$. However, long-range dependence is associated with all nonzero $d > 0$, which allows capturing the long-run behavior without being 'close to the boundary' of the parameter space. This long-run dependence is achieved with less restrictions on the higher frequency behavior of the time series. This can be seen by the relatively flat spectral density over higher frequencies; see fig. 1. Finally, when investigating the long-run behavior of a time series the problems associated with model selection with nonfractional ARIMA models can be avoided by only considering models that include the fractional differencing parameter. When investigating the long-run behavior of a time series, a model should be considered that allows the long-run behavior to be captured.

Fractional ARIMA models have advantages over nonfractional ARIMA models particularly when the long-run behavior of the series is being studied. But how can they be used to answer important economic questions? One method presented in Diebold and Rudebusch (1989) used the estimated fractional differencing parameters to estimate the impulse response functions. Though constrained by imprecise estimation techniques, the approach of Diebold and Rudebusch is an example of how fractional ARIMA models can give new insights into economic problems. A different procedure, presented below, uses the fractional ARIMA models to test between different models of long-run behavior.

3. The fractional ARIMA model and models of trend

An important question in the past decade has been the trend behavior of economic and financial data. Attention has been primarily focused on testing between the deterministic time trend model and the unit root with drift model. Because this question is concerned with the long-run behavior of a time series, the fractional ARIMA model should be useful.

The general deterministic trend model is

$$x_t = \mu t + e_t,$$

where e_t is stationary. This implies that there exists a Wold representation

$$e_t = b(L)\varepsilon_t = \sum_{j=0}^{\infty} b_j \varepsilon_{t-j},$$

where $\varepsilon_t \sim (0, \sigma_\varepsilon^2)$ is uncorrelated and $\sum_{j=0}^{\infty} b_j^2 < \infty$. The additional assumption is made that $b(1) = \sum_{j=0}^{\infty} b_j \neq 0$. The series in levels is nonstationary; however, a stationary model is achieved by first differencing the series. The resulting model is

$$\Delta x_t = \mu + (1-L) \sum_{j=0}^{\infty} b_j \varepsilon_{t-j}. \tag{1}$$

The general unit root with drift model is

$$x_t = \mu + x_{t-1} + u_t,$$

where u_t is stationary. This implies that there exists a Wold representation

$$u_t = a(L)\varepsilon_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j},$$

where $\varepsilon_t \sim (0, \sigma_\varepsilon^2)$ is uncorrelated and $\sum_{j=0}^{\infty} a_j^2 < \infty$. The additional assumption is made that $a(1) = \sum_{j=0}^{\infty} a_j \neq 0$. The series in levels is nonstationary; however, a stationary model can be achieved by first differencing the series. The resulting model is

$$\Delta x_t = \mu + u_t = \mu + \sum_{j=0}^{\infty} a_j \varepsilon_{t-j}. \tag{2}$$

The only distinguishing characteristic between the first difference of the deterministic time trend model [eq. (1)] and the first difference of the unit root with drift model [eq. (2)] is the term $(1-L)$ in the first difference of the deterministic trend series. This is the well-known result that first differencing a deterministic trend imposes a unit root in the moving average representation. In terms of the fractional integration parameter, the first difference of the deterministic trend series is integrated of order -1.0 and the first difference of the unit root model is integrated of order 0.0. This suggests testing between the models by estimating the best fractional ARIMA model for the first differenced series. If the fractional difference parameter estimate is around -1.0, the data support the deterministic trend model; if the parameter estimate is around 0.0, the data support the unit root with drift model.

When attention is limited to the deterministic trend model and the unit root model and it is assumed that the error's Wold representation does not sum to zero, then a necessary and sufficient condition for the first difference of the series to be integrated of order 0.0 is that its spectral density at frequency zero is not zero. Similarly, when attention is restricted to the unit root model with drift and the deterministic trend models, a necessary and sufficient condition for the first difference of the series to be integrated of order -1.0 is that its spectral density is zero at frequency zero. This means that the test of the order of integration of a series is testing the same *implication* as tests of the spectral density at frequency zero. Though the implication being tested is the same, the tests are different. Tests which use the estimated value of the spectral density at frequency zero focus the information in the data into the estimated level of the spectral density at frequency zero. But, tests which use the fractional differencing parameter focus the sample information on the behavior of the series at higher frequencies. The next section highlights the difference between tests based on the spectral density at frequency zero and the fractional differencing parameter by considering the behavior of the spectral density of a fractionally integrated series.

4. Fractional integrated model at frequency zero

Using the fractional difference parameter to characterize the long-run behavior of a time series is different from statistics that estimate the spectral density at frequency zero. The fractional differencing parameter is not identified by the level of its spectral density at frequency zero. For every fractional ARIMA(p, d, q) model with AR and MA roots outside the unit circle the level of the spectral density at frequency zero satisfies

$$f_x(0) = \begin{cases} 0, & d < 0, \\ \infty, & 0 < d. \end{cases}$$

Similarly, the fractional differencing parameter is not identified by the derivative of the spectral density at frequency zero,

$$f'_x(0) = \begin{cases} 0, & d \leq -\frac{1}{2}, \\ -\infty, & -\frac{1}{2} < d < 0, \\ \infty, & 0 < d. \end{cases}$$

The same behavior at frequency zero is associated with many different values of the fractional differencing parameter, hence the fractional differencing parameter is not identified by its behavior at frequency zero.

The distinction between the spectral density at frequency zero and the fractional differencing parameter can be understood by considering the moving average representation of a fractionally integrated series. For the fractionally integrated model $(1-L)^d x_t = \varepsilon_t$ the Wold representation is

$$x_t = C(L)\varepsilon_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j} = \sum_{j=0}^{\infty} \frac{\Gamma(d+j)}{\Gamma(d)\Gamma(j+1)} \varepsilon_{t-j}, \tag{3}$$

and as j increases the coefficients decay at a hyperbolic rate, which is indexed by d, i.e., as $j \to \infty$,

$$c_j \sim \frac{1}{\Gamma(d)} j^{d-1}. \tag{4}$$

[This can be contrasted with a geometric decay for an ARMA$(0 \neq p < \infty, q < \infty)$ model with roots outside the unit circle.] The cumulative impulse response is defined to be the total impact that a unit innovation in the error process has on the level of the series

$$C(1) = \sum_{j=0}^{\infty} c_j,$$

which is related to the spectral density at frequency zero by $f_x(0) = C(1)^2\sigma_\varepsilon^2$. While the cumulative impulse response function captures the ultimate outcome of a unit innovation, the fractional differencing parameter captures the behavior of the series as it approaches this ultimate outcome. In terms of an economic system the cumulative impulse response is the total impact that a unit innovation has on the system, while the fractional differencing parameter characterizes the rate at which the impact of the unit innovation dies out. The fractional differencing parameter is not identified by the ultimate impact of an innovation, but is identified by the decay of the system's response to the innovation. The same argument can be made in the frequency domain. The spectral density at frequency zero is the variance associated with infinitely long cycles and the fractional differencing parameter indexes the behavior of the series as the infinitely long cycle is approached; see fig. 1.

One criticism of the cumulative impulse response function and the variance ratio statistic is that these statistics attempt to estimate the impact of an innovation in the infinite future. The point is that given a finite sample, the future behavior of the series is uncertain and this uncertainty increases with cycle length. The greatest uncertainty is associated with the behavior of the time series at the infinitely long cycle, which is what the cumulative impulse response function and the variance ratio statistic attempt to estimate. Estimates of the fractional differencing parameter avoid this criticism. The fractional differencing parameter is not identified by the behavior of the series at the infinitely long cycle. It is identified by the behavior at shorter cycles. A fractional differencing parameter imposes restrictions over all frequencies, with most of its restrictions concentrated at the low frequencies again see fig. 1.

5. The trend behavior of postwar quarterly real GNP

Much has been published concerning 'What is the correct model of the trend behavior of real GNP?'. The two primary competitors are the deterministic time trend model and the unit root with drift model. These two models of the trend are thought to apply to the natural log of real GNP. To distinguish between these two models the test proposed in section 3 will be applied to quarterly real US GNP from 1947:I to 1989:IV.

Table 1
2ln(L), *AIC*, *SIC*; first difference of the log of quarterly real GNP (s.a.); Citibase 1947:I–1989:IV; 16 fractional ARMA models and 16 ARMA models.

Number of AR parameters (p)	Number of MA parameters (q) 0	1	2	3
	Fractional ARMA			
0	1083.436 1081.436 1078.294	1085.865 1081.865 1075.582	1096.057 1090.057 1080.632	1100.848 1092.848 1080.281
1	1095.024 1091.024 1084.741	1095.198 1089.198 1079.773	1100.911 1092.911 1080.345	1102.092 1092.092 1076.384
2	1095.468 1089.468 1080.043	1097.328 1089.328 1076.761	1101.748 1091.748 1076.039	1103.053 1091.053 1072.203
3	1100.405 1092.405 1079.839	1100.467 1090.467 1074.759	1105.695 **1093.695** 1074.845	1105.721 1091.721 1069.730
	ARMA			
0	1065.658 1065.658 1065.658	1083.242 1081.242 1078.100	1095.959 1091.959 **1085.675**	1097.987 1091.987 1082.562
1	1090.567 1088.567 1085.425	1092.068 1088.068 1081.784	1097.256 1091.256 1081.831	1098.033 1090.033 1077.466
2	1093.472 1089.472 1083.189	1094.964 1088.964 1079.539	1101.185 1093.185 1080.618	1102.362 1092.362 1076.654
3	1097.269 1091.269 1081.844	1098.936 1090.936 1078.370	1102.649 1092.649 1076.941	1103.186 1091.186 1072.336

5.1. The parameter estimates

Thirty-two different models were estimated for the first difference of the log of real GNP; sixteen ARMA(p, q) models (i.e., $d = 0$), where $p = 0, 1, 2, 3$, $q = 0, 1, 2, 3$, and the corresponding sixteen fractional ARIMA models. Each model was estimated by maximum likelihood assuming that the errors are normally distributed. The exact likelihood functions were evaluated using the procedures presented in Sowell (1992) and McLeod (1979) for the fractional ARIMA models and the nonfractional ARMA models, respectively.[3] For

[3]The loglikelihood functions were maximized using the DFP routine in GQOPT. Numerical derivatives were used in the optimization and the asymptotic standard errors are based on the 'H' matrix from the DFP algorithm.

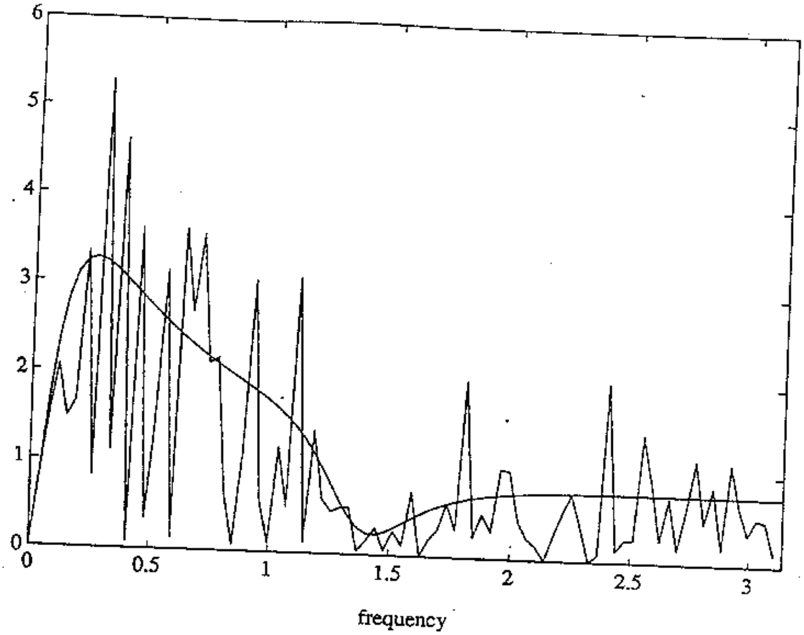

Fig. 2. Spectral density of the estimated fractional ARIMA(3, d, 2) model plotted over the periodogram for the first difference of the log of quarterly real GNP, 1947:I–1989:IV.

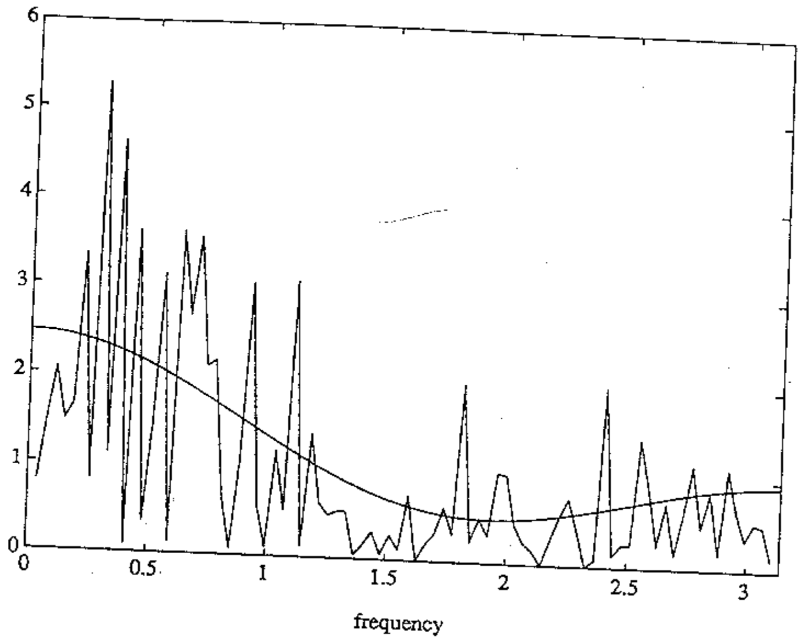

Fig. 3. Spectral density of the estimated ARMA(0, 2) model plotted over the periodogram for the first difference of the log of quarterly real GNP, 1947:I–1989:IV.

each model, 2 times the loglikelihood function ($2\mathscr{L}$), the *AIC*, and the *SIC* are reported in table 1. The *AIC* chose the fractional ARIMA(3, d, 2) model, while the *SIC* chose the nonfractional ARMA(0, 2). The decision on which of these two models to consider is settled by looking at how well the estimated models explain the long-run behavior of the data. To determine this, the implied spectral densities were plotted over the periodogram for the data in fig. 2 and fig. 3. Fig. 3 demonstrates how the point estimates of the ARMA(0, 2) model miss the behavior of the series at a low frequencies; in particular the eight periodogram ordinates[4] approaching zero show a downward pattern but the estimated spectral density is actually increasing. The data appear to be explained best by the fractional ARIMA(3, d, 2).[5] The estimated model is

$$\Big(1 - \underbrace{1.18}_{(0.397)} L + \underbrace{0.93}_{(0.317)} L^2 - \underbrace{0.51}_{(0.195)} L^3\Big)(1 - L)^{\overbrace{-0.59}^{}}_{(0.345)} \Delta x_t$$

$$= \Big(1 - \underbrace{0.29}_{(0.132)} L + \underbrace{0.81}_{(0.112)} L^2\Big)\varepsilon_t .$$

where $\hat{\sigma}^2 = 8.413 \times 10^{-5}$, $\mathscr{L} = 552.847$, and asymptotic standard errors are reported below each estimated parameter. The parameter estimates for all 16 fractional models are reported in table 2, and the parameter estimates for the 16 nonfractional models are reported in table 3. The restriction $d = 0.0$ was imposed in the estimation of the ARMA(3, 2) model. The likelihood function value for this restricted model was 551.325. The restriction $d = -1.0$ was imposed indirectly in the estimation of the ARMA(3, 3) model. The parameter estimates reported in table 3 show that the estimated ARMA(3, 3) contains a unit root in its MA process. The likelihood function value for the ARMA(3, 3) model was 551.593.

5.2. *Econometric issues and interpretations*

The consistency and asymptotic normality for the maximum likelihood estimates for the Gaussian fractional ARIMA model are presented in Dahlhaus (1989), for models where $0 < d < \frac{1}{2}$. Asymptotic properties when $d < 0$, the relevant case for this paper, are still an open question. Monte Carlo results presented in Cheung and Diebold (1990) and Sowell (1992) show that the maximum likelihood estimators have comparable small sample

[4]These ordinates are only associated with the behavior of the time series for cycles that are 5 years or longer.

[5]A likelihood ratio test of the restrictions implied by the ARMA(0, 2) on the ARIMA (3, d, 2) can be rejected at the 5 percent level of significance.

Table 2

Parameter estimates for fractional ARMA models; first difference of the log of quarterly real GNP (s.a.); Citibase 1947:I–1989:IV; t-statistics in parentheses.

Model	d	ϕ_1	ϕ_2	ϕ_3	θ_1	θ_2	θ_3
$(0,d,0)$	0.29 (4.18)						
$(0,d,1)$	0.16 (1.52)				0.16 (1.65)		
$(0,d,2)$	−0.03 (−0.31)				0.33 (2.89)	0.28 (3.84)	
$(0,d,3)$	−0.20 (−1.92)				0.48 (4.68)	0.46 (4.73)	0.25 (2.38)
$(1,d,0)$	−0.45 (−2.90)	−0.77 (−6.48)					
$(1,d,1)$	−0.38 (−2.16)	−0.74 (−6.63)			−0.05 (−0.44)		
$(1,d,2)$	−0.41 (−1.42)	−0.65 (−2.57)			0.06 (0.50)	0.23 (2.57)	
$(1,d,3)$	−0.33 (−1.73)	−0.37 (−1.04)			0.26 (1.17)	0.39 (2.60)	0.17 (1.17)
$(2,d,0)$	−0.30 (−1.27)	−0.60 (−2.44)	−0.08 (−0.79)				
$(2,d,1)$	−0.29 (−1.56)	−0.07 (−0.22)	−0.44 (−2.10)		0.52 (1.65)		
$(2,d,2)$	−0.25 (−0.79)	−0.78 (−2.80)	0.26 (1.04)		−0.23 (−0.92)	0.37 (1.44)	
$(2,d,3)$	−0.13 (−0.79)	−0.41 (−1.72)	0.39 (1.47)		0.03 (0.12)	0.63 (3.45)	0.19 (1.50)
$(3,d,0)$	−0.39 (−1.53)	−0.70 (−2.77)	−0.17 (−1.65)	0.17 (2.23)			
$(3,d,1)$	−0.34 (−1.09)	−0.75 (−2.27)	−0.11 (−0.45)	0.18 (2.21)	−0.11 (−0.25)		
$(3,d,2)$	−0.59 (−1.71)	−1.18 (−2.97)	0.93 (2.93)	−0.51 (−2.61)	−0.29 (−2.19)	0.81 (7.22)	
$(3,d,3)$	−0.52 (−1.16)	−1.16 (−3.53)	0.94 (3.56)	−0.50 (−2.65)	−0.34 (−1.08)	0.83 (6.03)	−0.04 (−0.18)

properties for positive and negative values of d. This suggests that the standard asymptotic properties apply when $d < 0$.

Because the fractionally differencing parameter captures the long-run behavior of a series, the appropriateness of asymptotic results when interpreting small sample results is questionable.[6] The small sample properties of

[6] For example, see section 6.2 of this paper.

Table 3

Parameter estimates for ARMA(p, q) models; first difference of the log of quarterly real GNP (s.a.); Citibase 1947:I–1989:IV; t-statistics in parentheses.

Model	ϕ_1	ϕ_2	ϕ_3	θ_1	θ_2	θ_3
(0, 0)						
(0, 1)				0.27 (4.48)		
(0, 2)				0.30 (4.07)	0.27 (3.88)	
(0, 3)				0.33 (3.46)	0.34 (5.40)	0.13 (1.57)
(1, 0)	−0.37 (−5.35)					
(1, 1)	−0.52 (−4.28)			−0.17 (−1.81)		
(1, 2)	−0.25 (−1.57)			0.07 (0.46)	0.24 (2.81)	
(1, 3)	−0.93 (−1.88)			−0.64 (−0.78)	−0.05 (−0.06)	−0.30 (−0.78)
(2, 0)	−0.32 (−3.96)	−0.13 (−1.63)				
(2, 1)	0.07 (0.24)	−0.29 (−2.56)		0.39 (1.34)		
(2, 2)	−0.60 (−3.60)	0.49 (4.83)		−0.30 (−2.40)	0.64 (5.70)	
(2, 3)	−0.40 (−1.85)	0.53 (3.35)		−0.09 (−0.39)	0.71 (6.42)	0.14 (1.23)
(3, 0)	−0.34 (−4.53)	−0.18 (−2.27)	0.15 (1.98)			
(3, 1)	−0.83 (−2.30)	−0.02 (−0.15)	0.22 (2.89)	−0.50 (−1.36)		
(3, 2)	−0.60 (−4.05)	0.67 (3.75)	−0.14 (−1.33)	−0.28 (−2.29)	0.79 (6.71)	
(3, 3)	−1.52 (−5.65)	0.97 (1.77)	−0.41 (−1.06)	−1.26 (−4.09)	0.90 (1.63)	−0.64 (−1.59)

the estimates must be investigated. The estimates reported above will be interpreted assuming standard asymptotic results and also based on the small sample distributions determined by Monte Carlo simulation.

The null hypotheses of interest are $d = 0.0$ (support for the unit root model) and $d = -1.0$ (support for the time trend model). Imposing the hypothesis $d = 0.0$ on the fractional ARIMA(3, d, 2) model gives the ARMA(3, 2) model. So to determine the small sample distribution of tests for

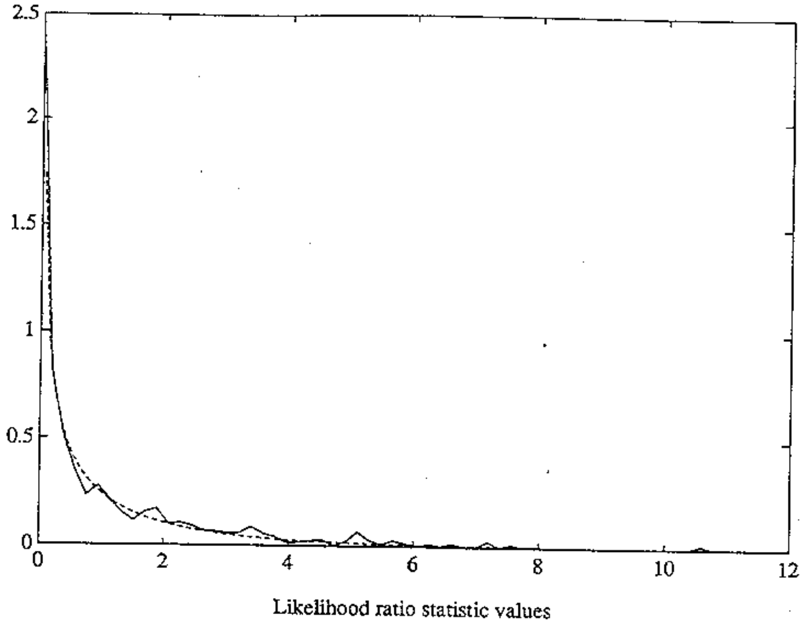

Fig. 4. Estimate of the likelihood ratio statistic's small sample density for the hypothesis H_0: $d = 0.0$. The simulated model was the ARMA(3, 2) model estimated for the first difference of the log of real US GNP. The dotted line is the chi-squared density with one degree of freedom.

the hypothesis $d = 0.0$, 1000 samples were simulated from the estimated ARMA(3, 2) model. For each sample both the fractional ARIMA(3, d, 2) and the ARMA(3, 2) model were estimated by maximum likelihood. The maximum likelihood estimates of d, the Wald test statistics for the hypothesis $d = 0.0$, and the likelihood ratio statistics of the hypothesis $d = 0.0$ were used to estimate small sample densities and empirical p-values for the observed test statistics for the real GNP series. As noted above, imposing the restriction $d = -1.0$ on the fractional ARIMA(3, d, 2) model, for this series, gives the ARMA(3, 3) model. So to determine the small sample distribution of tests for the hypothesis $d = -1.0$, 1000 samples were simulated from the estimated ARMA(3, 3) model. For each sample both the fractional ARIMA(3, d, 2) and the ARMA(3, 3) model (with a unit root in the MA process) were estimated by maximum likelihood. The maximum likelihood estimates of d, the Wald test statistics for the hypothesis $d = -1.0$, and the likelihood ratio statistics of the hypothesis $d = -1.0$ were used to estimate small sample densities and empirical p-values for the observed test statistics for the real GNP series.

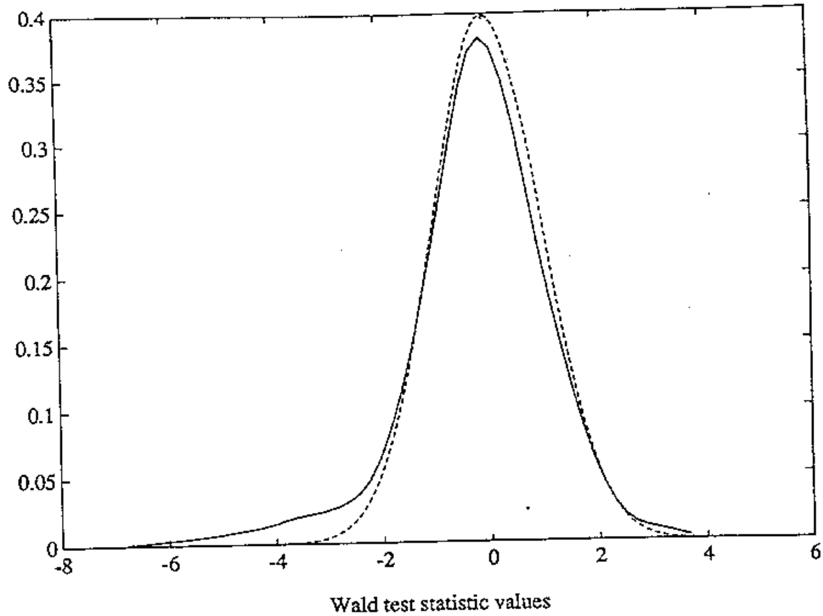

Fig. 5. Estimate of the Wald statistic's small sample density for the hypothesis H_0: $d = 0.0$. The simulated model was the ARMA(3, 2) model estimated for the first difference of the log of real US GNP. The dotted line is a standard normal density.

First consider the null hypothesis $d = 0.0$; the likelihood ratio test statistics value is 3.046. For a chi-squared distribution with one degree of freedom this implies a p-value of 8.3 percent. The estimated small sample density for this statistic is given in fig. 4. The chi-squared distribution with one degree of freedom appears to be a reasonable approximation.[7] The empirical p-value from the observed statistic is 10.1 percent. The Wald test of this hypothesis gives a test statistic value of 1.710 and for a normal distribution this implies a one-tailed p-value of 4.36 percent. The estimated small sample density for this statistic is given in fig. 5. The standard normal density does not appear to be a good approximation. The small sample density has a much longer and thicker left tail. The empirical p-value is 8.8 percent. The small sample density for maximum likelihood estimate of d is given in fig. 6. The density shows a second smaller mode apparently centered near -0.59. Using the

[7]The sample mean of the likelihood ratio statistics is 1.1035 and the standard deviation is 1.5198. The small sample density has slightly more mass on higher values than would a chi-squared density with one degree of freedom.

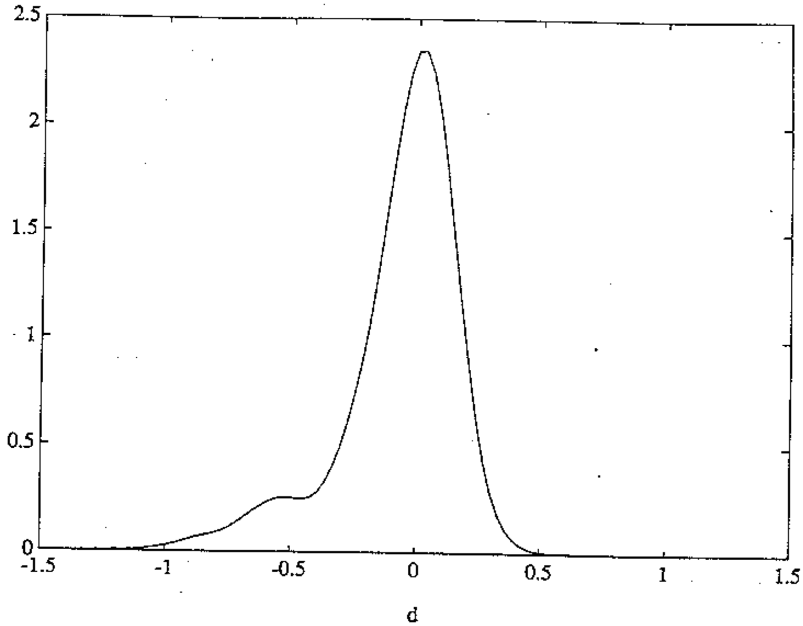

Fig. 6. Estimate of the small sample density for the maximum likelihood estimate of d, where the population parameter value is $d=0.0$. The simulated model was the ARMA(3, 2) model estimated for the first difference of the log of real US GNP.

small sample density for the maximum likelihood estimate of d the empirical p-value for the estimate of d for the observed GNP series is 4.9 percent. The evidence generally agrees that the hypothesis $d=0.0$, which indicates support of a unit root in real GNP, cannot be rejected at the 95 percent level of significance but could be rejected at the 90 percent level of significance.

Now consider the null hypothesis $d=-1.0$; the likelihood ratio test statistic value is 2.509 and for a chi-squared distribution with one degree of freedom this implies a p-value of 11.41 percent. The estimated small sample density for this statistic is given in fig. 7. The chi-squared distribution with one degree of freedom appears to be a reasonable approximation.[8] The empirical p-value from the observed statistic is 8.0 percent. The Wald test statistic for this hypothesis is 1.188. For a normal distribution this implies a

[8]The sample mean of the likelihood ratio statistics is 0.8572 and the standard deviation is 1.1816. The small sample density has slightly more mass on smaller values than would a chi-squared density with one degree of freedom.

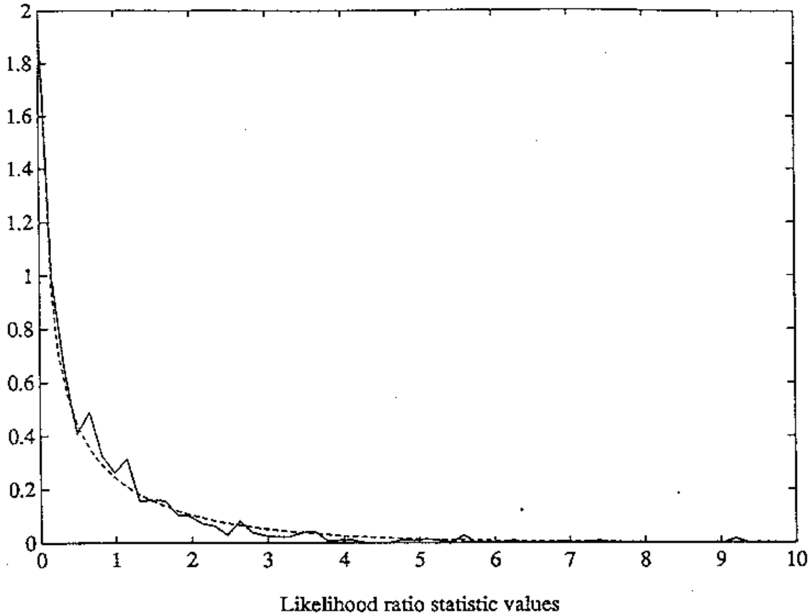

Fig. 7. Estimate of the likelihood ratio statistic's small sample density for the hypothesis H_0: $d = -1.0$. The simulated model was the ARMA(3,3) model estimated for the first difference of the log of real US GNP. The dotted line is the chi-squared density with one degree of freedom.

one-tailed p-value of 11.7 percent. The estimated small sample density for this statistic is given in fig. 8. The standard normal density does not appear to be good approximation. The small sample density has a shorter left tail and a longer right tail than a standard normal density. The empirical p-value is 16.8 percent. The small sample density for the maximum likelihood estimate of d is given in fig. 9. The density shows a thick right tail. Using the small sample density for the maximum likelihood estimate of d, the empirical p-value for the estimate of d for the observed GNP series is 9 percent. Again, the evidence generally agrees that the hypothesis $d = -1.0$, which indicates support of a time trend in real GNP, cannot be rejected at the 95 percent level of significance but could possibly be rejected at the 90 percent level of significance.

That the fractional parameter estimate was not significantly (i.e., 5 percent level) different from zero and not significantly different from -1.0 implies that the data cannot distinguish between the two models of trend. This result can be seen graphically by plotting cross-sections of the loglikelihood function. In fig. 10, three cross-sections of the loglikelihood function are plotted.

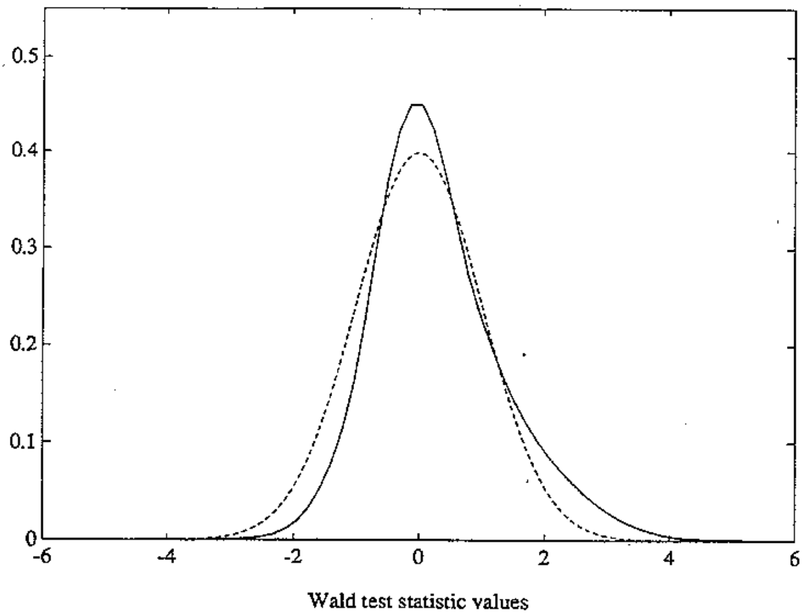

Fig. 8. Estimate of the Wald statistic's small sample density for the hypothesis H_0: $d = -1.0$. The simulated model was the ARMA(3, 3) model estimated for the first difference of the log of real US GNP. The dotted line is a standard normal density.

Each cross-section passes through the maximum for the fractional ARIMA($3, d, 2$) model. One cross-section is in the d dimension, another passes through the estimated ARMA(3, 2) model, and the third passes through the estimated ARMA(3, 3) model. The parabolic shape of the cross-section in the d dimension is in sharp contrast to the other cross-sections. The other cross-sections show a much slower decline from the maximum, which highlights the uncertainty in the data concerning the population parameter values. This explains why this series regularly fails to reject both the null hypothesis of time trend and unit root. This also highlights the danger of forcing the data to fit into one type of model or the other.

That the optimal model selected by the AIC is a fractional model shows the importance of including the fractional differencing parameter when modeling the long-run behavior of a time series. If the fractional differencing parameter were not included, the AIC would have selected the ARMA(2, 2) as in Campbell and Mankiw (1987). The parameter estimates associated with the ARMA(2, 2) model suggest strong support that the series is consistent with the unit root model. However, a comparison of the periodogram for the data and the spectral density implied by the ARMA(2, 2) model in fig. 11

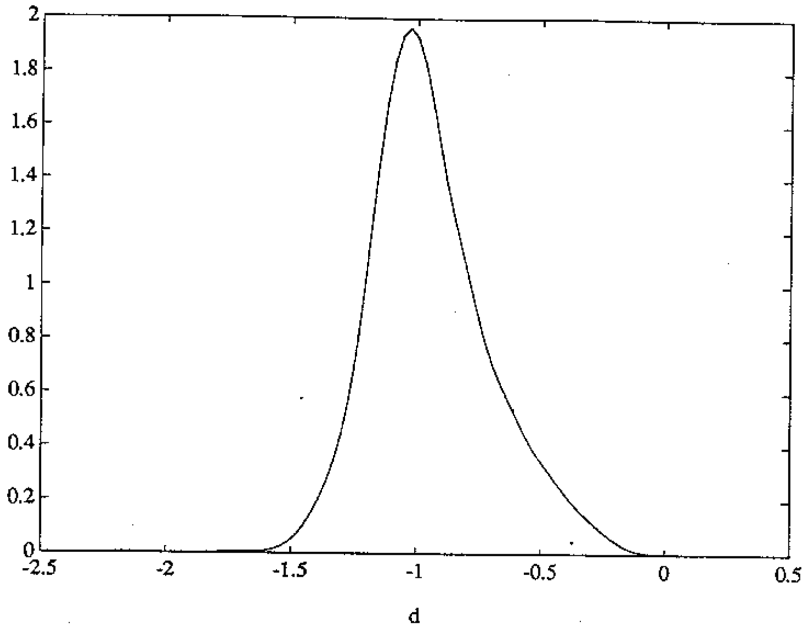

Fig. 9. Estimate of the small sample density for the maximum likelihood estimate of d, where the population parameter value is $d = -1.0$. The simulated model was the ARMA(3, 3) model estimated for the first difference of the log of real US GNP.

shows that the estimated model does a poor job of capturing the long-run behavior of the series,[9] which is the behavior being investigated. This is an example where the ARMA(p, q) models miss the long-run behavior of the series. This position is supported by the fact that when the fractional differencing parameter is included the drop in the periodogram is captured and the fractional model obtains a higher *AIC* than did all of the estimated ARMA(p, q) models.

6. Comparison to previous work

6.1. Christiano and Eichenbaum (1990)

The conclusion that postwar quarterly GNP is equally consistent with both the deterministic time trend model and the unit root with drift model is the same conclusion reached in Christiano and Eichenbaum (1990) [C&E]. A

[9]The periodogram shows a decline between frequencies 0.4 and 0.0 (cycles longer than four years). The ARMA(2, 2) model misses this long-run behavior.

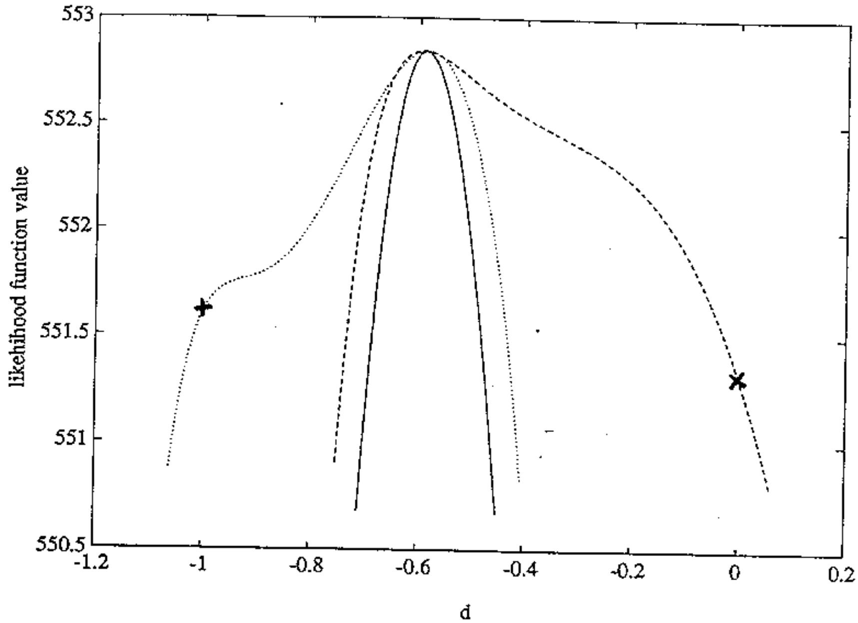

Fig. 10. Cross-sectional plots of the loglikelihood function for the first difference of the log of real US GNP. Each cross-section is through the maximum for the fractional ARIMA(3, d, 2) model. The solid line is the cross-section in the d dimension. The dotted line is the cross-section that passes though the maximum achieved by the ARMA(3, 3) model. For this time series the ARMA(3, 3) model is actually the fractional ARIMA(3, d, 2) model with the restriction $d = -1.0$, which is consistent with a time trend model. The dashed line is the cross-section that passes though the maximum achieved by the ARMA(3, 2) model. The ARMA(3, 2) model can be thought of as the fractional ARIMA(3, d, 2) model with the restriction $d = 0.0$, which is consistent with a unit root model. The + denotes the maximum achieved by the (time trend) ARMA(3, 3) model. The × denotes the maximum achieved by the (unit root) ARMA(3, 2) model.

comparison of the current test with the analysis in C&E demonstrates the usefulness of the fractionally integrated model. In C&E only nonfractional ARMA models were considered. As noted in section 2, these possess some inherent problems in modeling the long run, which were noted in C&E. The case made in C&E for the consistency of the GNP data with a time trend model relied on the parameter estimates of the ARMA(3, 3). Unfortunately, this model was not selected by any of the information criteria. This ARMA(3, 3) model had near root cancelation of a near unit root in the AR polynomial and a unit root in the MA polynomial. One possible explanation for this is that the data were forced to fit into an ARMA model when the data were more accurately explained by the fractional ARIMA model. The problems are noticeably absent in the estimated fractional ARIMA(3, d, 2)

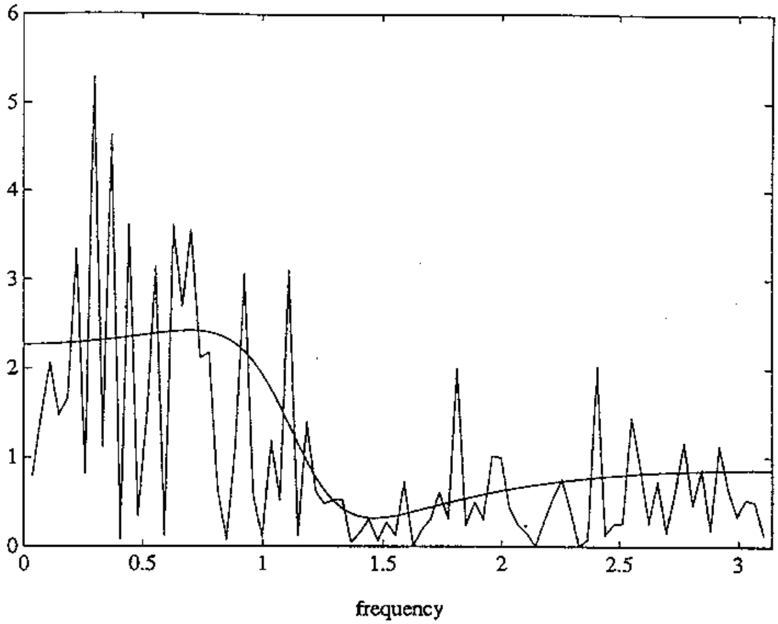

Fig. 11. Spectral density of the estimated ARMA(2, 2) model plotted over the periodogram for the first difference of the log of quarterly real GNP, 1947:I–1989:IV.

model. The fractional model was selected by the *AIC*, the AR and the MA processes did not have near root cancelation and the parameter values were not near the boundary of the parameter space. The analysis in C&E relied on extensive Monte Carlo and plots of the likelihood functions in the dimension of the spectral density at frequency zero. The usefulness of the fractional ARIMA model can be seen by noting that the same conclusions are reached in the current paper by simply estimating (fractional) ARIMA models for the series.

6.2. *Diebold and Rudebusch (1989)*

The conclusions that postwar GNP data cannot distinguish between a unit root and a time trend model is in sharp contrast to the results reported in D&R. Based on the point estimates and standard errors reported in D&R[10] the data strongly reject a time trend model and are consistent with only a unit root model. This serious disagreement in the long-run characteristics of

[10]For the sample period, 1947:I to 1987:II, the comparable fractional differencing parameter estimates, with standard errors in parentheses, were −0.1 (0.24), −0.08 (0.23), and −0.12 (0.21).

real GNP is a result of bias in the fractional differencing parameter reported in D&R.

The problem in the analysis used in D&R is that the periodogram ordinates, used in calculating the estimate of d, were influenced by the short-run dynamics of the series and gave misleading measures of the long run. To see how this occurred consider what happened when D&R used $\alpha = 0.5$ to determine the number of periodogram ordinates, which will be denoted by k. The number of periodogram ordinates used was $k = 13$. However, the thirteenth harmonic ordinate was associated with a frequency of $\lambda = (2\pi)13/161$ which corresponds to cycles of length 3.1 years. Hence, the long-run parameter was being estimated assuming no short-run influence and 'constant' long-run behavior[11] for all cycles longer than 3.1 years. Most economists would argue that this range includes the business cycle and that the assumption of no short-run influence from all cycles longer than 3.1 years is suspect. The problem is made even worse as larger values of α are considered; when $\alpha = 0.55$ the number of ordinates used was 17 which would assume no short-run influence and constant long-run influence from 2.37 years to infinity.

The estimation procedure used in D&R was presented in Geweke and Porter-Hudak (1983). The estimator exploits the fact that the log spectral density of a series that is fractionally integrated, i.e., $(1-L)^d x_t = \varepsilon_t$, is of the form

$$\log(f_x(\lambda)) = -d\log\left[4\sin^2(\lambda/2)\right] + \log(\sigma_\varepsilon^2). \tag{5}$$

This relationship between frequency and the spectral density holds for all frequencies if the series is a pure fractionally integrated series, i.e., no AR or MA terms; see fig. 1. When there are additional parameters in the model the relationship given in eq. (5) does not hold. However, it was shown in Geweke and Porter-Hudak (1983) that even if there exist short-run parameters, as the number of observations increases, the relationship holds approximately for frequencies in a neighborhood of zero. If the neighborhood shrinks at an appropriate rate with sample size, then a consistent estimator can be obtained. The GPH estimator only uses sample information through the relationship given by eq. (5). The addition of more periodogram ordinates is useful only if eq. (5) is well approximated.

To understand the impact that this has on the parameter estimates, a Monte Carlo was performed. One thousand samples were simulated from the fractional ARIMA(3, d, 2) model estimated for quarterly real GNP. For each 171 observation sample, the fractional differencing parameter d was esti-

[11]Constant long-run behavior in the sense of a single fractional parameter with no ARMA parameters.

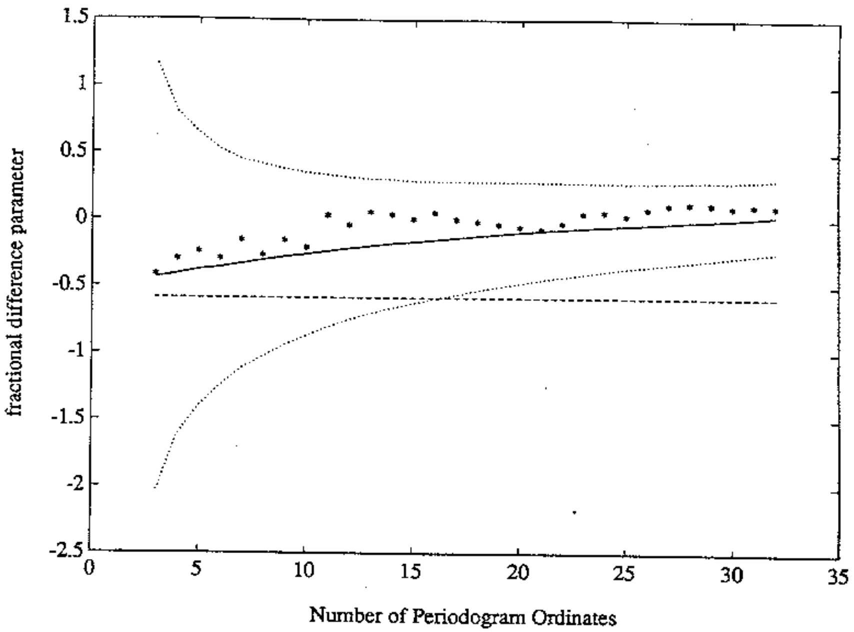

Fig. 12. Summary statistics of the small sample distributions of the GPH estimate for 1000 samples of 171 observations simulated from the estimated model for the first difference of the log of real GNP. The solid line connects the means and the dotted lines connect the two standard deviations confidence bounds. The dashed line is the population parameter value of d used in the simulations. For the actual GNP series the GPH parameter estimate is denoted by an asterisk ($*$).

mated using the first k nonzero periodogram ordinates and k was varied between 3 and 32. For each sample 30 different fractional differencing parameters were estimated. The mean and standard deviation of the 1000 estimated values were calculated and plotted in fig. 12. The solid line connects the means and the dotted lines connect two standard deviations confidence bounds. The dashed line in fig. 12 shows the population parameter value of d for the Monte Carlo. The estimated means all show bias which monotonically increases with the number of ordinates used. The standard deviations decrease as the number of periodogram ordinates increase but the cost is increased bias. The increased bias is noted by the divergence of the estimated means (the solid line) and the population value of d (the dashed line). For the parametric model that describes postwar quarterly GNP, it appears to be the case that the effect of the ARMA parameters is that the relationship given in eq. (5) is not well approximated even restricting attention to three periodogram ordinates. The relationship in eq. (5) cannot be successfully exploited with the current sample size. In Geweke and Porter-

Table 4

Monte Carlo results from 1000 samples of $T = 171$ observations; each sample simulated from the fractional ARMA$(3, d, 2)$ model estimated from postwar quarterly real GNP. For each sample the maximum likelihood estimates were calculated, and the means and standard deviations of the estimated parameters are reported.

Parameter	d	ϕ_1	ϕ_2	ϕ_3	θ_1	θ_2
Population value	−0.59	−1.18	0.93	−0.51	−0.29	0.81
Mean	−0.617	−1.197	0.913	−0.444	−0.300	0.779
Std. dev.	0.252	0.321	0.323	0.201	0.215	0.168

Hudak (1983) [GPH] consistency and asymptotic normality were shown to hold for $d < 0$. The above Monte Carlo shows that for a series, of typical length for macroeconomic series, the asymptotic results are a poor approximation of small sample distributions when AR and MA terms are present.[12]

To see if this bias can explain the difference in the maximum likelihood estimates and the parameter estimates presented in D&R, for the GNP series the GPH parameter estimates were calculated for $k = 3$ to 32. Each point estimate is plotted in fig. 12 by an asterisk. The point estimates are all in the two standard deviation confidence bounds and appear consistent with the fractional ARIMA$(3, d, 2)$ model for real GNP. The bias in the GPH estimates can explain the difference in the maximum likelihood estimates and those published in D&R.

Maximum likelihood also uses information at high frequencies. The difference is that with maximum likelihood the short-run parameters are also estimated. To be sure that this is the situation the similar Monte Carlo was considered for the maximum likelihood estimation procedure. One thousand samples were simulated from the fractional ARIMA$(3, d, 2)$ model estimated for quarterly real GNP. For each sample the fractional ARIMA$(3, d, 2)$ model was estimated by the exact maximum likelihood procedure. The means and standard deviations of the 1000 estimated parameter vectors are reported in table 4. All the parameters of the model are accurately estimated with the bias of each estimate below 0.066. The estimated small sample density of the maximum likelihood estimate of d is plotted in fig. 13. The normal distribution with the same mean and standard deviation is also plotted and appears to be a good approximation to the small sample density.

The above Monte Carlo shows that when the correct model specification is known, maximum likelihood gives more accurate estimates than the estima-

[12] More extensive Monte Carlo results presented in Sowell (1992) indicate that the GPH standard errors are a good approximation to the observed standard deviation in the GPH estimate. In practice, the problem is the bias caused by short-run dynamics.

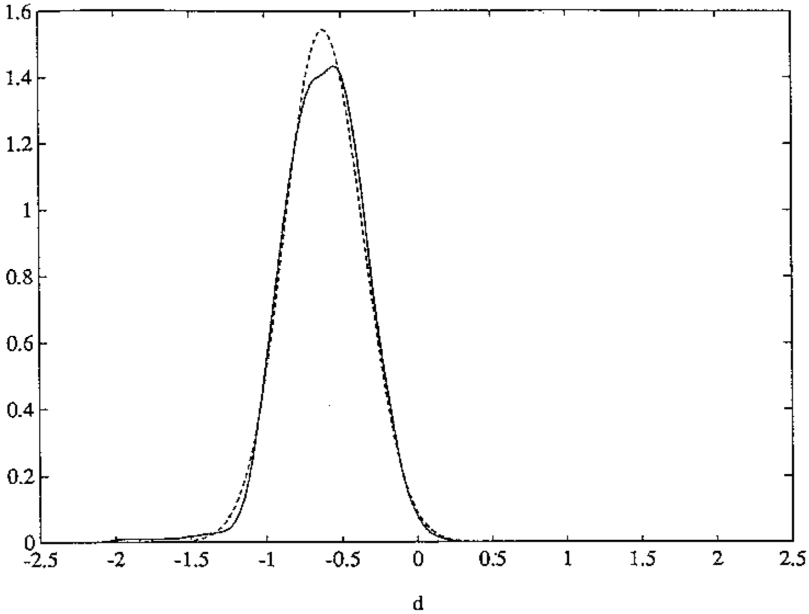

Fig. 13. Estimate of the small sample density for the maximum likelihood estimate of d, where the population parameter value is $d=-0.59$. The simulated model was the fractional ARIMA$(3,d,2)$ model estimated for the first difference of the log real US GNP. The dotted curve is a normal density with the same mean and standard deviation.

tion procedure presented in Geweke and Porter-Hudak (1983). However, in practice the specification is uncertain and must be estimated. It is still an open question how maximum likelihood performs relative to the GPH estimation procedure when the specification must be estimated.

7. Summary and conclusion

This paper has considered modeling the long-run behavior of economic time series using fractional ARIMA models. Shortcomings of nonfractional models in modeling the long-run behavior of a series are reduced or avoided by considering the fractional ARIMA models. As an application, the deterministic trend and the unit root with drift models were nested in a fractionally integrated ARIMA(p,d,q) model. This allowed testing between the two models based on estimated parameters. The only restriction between these two models concerns the absence or presence of the term $(1-L)$ in the Wold representation. One testable restriction that this implies concerns the rate at which the dependence between observations decays. This can be indexed by the fractional differencing parameter and forms the basis for a test between the two models. This test was applied to postwar US real quarterly GNP. The model which best explained the data was the fractional ARIMA(3, d, 2) model. The selection of a fractional model and the parameter estimates showed the importance of considering fractional ARIMA models when modeling the long-run behavior of a time series. The conclusion of the test is that the data are consistent with both models. Hence, researchers should be careful not to force the series to follow one model or the other. This underscores again that the failure of a statistical test to reject a model is not sufficient to conclude that the series follows the given model.

The fact that the postwar GNP series cannot distinguish between a time trend and a unit root model has important implications for theoretical models of the economy. Attention should be given to models where both the policy and theoretical implications of interest are not sensitive to the model of the trend. Ideally we would like a model which implies the same results if the trend is modeled as either a time trend or a unit root. Until such models are developed, further attention should be given to new statistical techniques which focus on discovering the long-run behavior of time series.

References

Campbell, J.Y. and G. Mankiw, 1987, Are output fluctuations transitory?, Quarterly Journal of Economics 102, 857–880.

Cheung, Y. and F.X. Diebold, 1990, On maximum-likelihood estimation of the differencing parameter of fractionally integrated noise with unknown mean, Discussion paper no. 34 (Institute for Empirical Macroeconomics, Federal Reserve Bank of Minneapolis, MN).

Christiano, L.J. and M. Eichenbaum, 1990, Unit roots in real GNP: Do we know and do we care?, Carnegie–Rochester Conference Series on Public Policy 32, 7–61.

Cochrane, J.H., 1988, How big is the random walk in GNP?, Journal of Political Economy 96, 893–920.

Dahlhaus, R., 1989, Efficient parameter estimation for self-similar processes, Annals of Statistics 17, 1749–1766.

Diebold, F.X. and G.D. Rudebusch, 1989, Long memory and persistence in aggregate output, Journal of Monetary Economics 24, 189–209.

Geweke, J.F. and S. Porter-Hudak, 1983, The estimation and application of long memory time series models, Journal of Time Series Analysis 4, 221–238.

McLeod, I., 1975, Derivation of the theoretical autocovariance function of autoregressive-moving average time series, Applied Statistics 24, 255–256.

Sowell, F.B., 1992, Maximum likelihood estimation of stationary univariate fractionally integrated time series models, Journal of Econometrics, forthcoming.