WHAT IS FRACTIONAL INTEGRATION?

William R. Parke*

Abstract—A simple construction that will be referred to as an ***error-duration model*** is shown to generate fractional integration and long memory. An error-duration representation also exists for many familiar ARMA models, making error duration an alternative to autoregression for explaining dynamic persistence in economic variables. The results lead to a straightforward procedure for simulating fractional integration and establish a connection between fractional integration and common notions of structural change. Two examples show how the error-duration model could account for fractional integration in aggregate employment and in asset price volatility.

I. Introduction

A GROWING body of empirical evidence supports the notion that important economic data series might be fractionally integrated.[1] Fractional integration and, more generally, long memory can produce data that appear to be stationary, but that nonetheless have high-order autocorrelations that are too large to be accounted for by a parsimonious ARMA model. The basic fractionally integrated process can be expressed as $(1 - L)^d y_t = \eta_t$, where η_t is i.i.d. and $0 < d < 1$. Although $(1 - L)^d$ is well defined algebraically for fractional d, it is not clear what sort of economic process might generate such data.[2] It is known that finite-order autoregressive and moving-average approximations for fractionally integrated processes require extremely long lags to achieve any kind of accuracy.

This paper introduces the error-duration representation for fractionally integrated processes. It has two goals. First, it provides a simple mechanism that generates fractional integration in a way that might plausibly be part of an economic process. Second, it provides a little additional insight into what happens at the critical point $d = 1/2$ on the boundary between stationarity and nonstationarity.

Two examples show how aggregate employment and asset price volatility might be fractionally integrated. Aggregate employment is shown to be fractionally integrated if a few firms have surprisingly long lifetimes given the turnover in new firms. Asset price volatility is shown to be fractionally integrated if a few asset positions last longer than would be predicted by the lifetimes of typical positions.

The basic mechanism for an error-duration model is a sequence of shocks of stochastic magnitude and stochastic duration. The variable observed in a given period is the sum of those shocks that survive to that point, and the distribution of the durations of the shocks determines whether or not the process is fractionally integrated. Fractional integration requires that a small percentage of the shocks have long durations, where that concept is defined rigorously in this paper.

The plan of this paper is as follows. Sections II, III, and IV describe the error-duration model, show how it can induce long memory, and give a specific version that generates fractional integration. Section V then shows how the error-duration mechanism can naturally induce long memory in aggregate employment and in asset price volatility. Section VI extends the discussion to continuous time. Section VII considers some implications for estimation and structural breaks. Section VIII summarizes the theme of the discussion. An appendix gives a method for simulating error-duration models.

II. The Error Duration Model

The structure for an error-duration model is as follows. Let $\{\epsilon_t, t = 1, 2, \ldots\}$ be a series of i.i.d. shocks with mean zero and finite variance σ^2. The error ϵ_s has a stochastic duration $n_s \geq 0$, surviving from period s until period $s + n_s$. Let $g_{s,t}$ be an indicator function for the event that error ϵ_s survives to period t. That is, $g_{s,t} = 1$ for $t \leq s + n_s$, and $g_{s,t} = 0$ for $t > s + n_s$. Assume that ϵ_s and $g_{s,t}$ are independent for all $t \geq s$. Let p_k be the probability of the event that ϵ_s survives until period $s + k$; that is, $p_k = P(g_{s,s+k} = 1)$. Assume that $p_0 = 1$ and that the sequence of probabilities $\{p_k, k = 0, 1, 2, \ldots\}$ is monotone non-increasing. The realization y_t is the sum of all errors ϵ_{t-i}, $i = 0, 1, 2, \ldots$ that survive until period t:

$$y_t = \sum_{s=-\infty}^{t} g_{s,t}\epsilon_s. \tag{1}$$

The survival probabilities $\{p_0, p_1, p_2, \ldots\}$ are the fundamental parameters of this error-duration (ED) representation of y_t. The autocovariances γ_k, given in proposition (1) below, are the link that establishes a correspondence between error-duration models and other representations.

PROPOSITION (1). If they exist, the autocovariances of y_t are given by

$$\gamma_k = \sigma^2 \sum_{j=k}^{\infty} p_j, \tag{2}$$

Received for publication September 22, 1998. Revision accepted for publication June 8, 1999.

* University of North Carolina, Chapel Hill.

I would like to thank John Geweke, Clive Granger, Miguel Herce, and Ming Liu for helpful comments on an earlier draft of this paper. Ken Hightower and Alicia Robb provided valuable research assistance. Very helpful comments were received from the editors and anonymous referees.

[1] Evidence of long memory has been found in traditional business-cycle indicators such as aggregate economic activity (Diebold & Rudebusch, 1989; Sowell, 1992) and price indices (Geweke & Porter-Hudak, 1983; Baillie, Chung, et al., 1996). There is also strong evidence of long memory in asset price and exchange rate volatility (Andersen & Bollerslev, 1997; Andersen et al., 1999; Baillie, Bollerslev, et al., 1996; Breidt, et al., 1998; Ding et al., 1993). Baillie (1996) provides an excellent survey of the literature on fractional integration and long memory.

[2] The list of explanations for $I(d)$ processes is not long. Granger (1980) shows that an $I(d)$ variable could arise from aggregation of an infinite number of AR(1) variables with their parameters distributed over a range. For $d < 1/2$, Liu (1995) explains stock market volatility in terms of a regime-switching process. Section VI of this paper discusses some continuous-time explanations.

The Review of Economics and Statistics, November 1999, 81(4): 632–638

which implies

$$p_k = (\gamma_k - \gamma_{k+1})/\sigma^2. \quad (3)$$

If they exist, the autocovariances of $(1 - L)\, y_t$ are given by

$$\gamma_k = \sigma^2(p_k - p_{k-1}), \quad (4)$$

where the p_k's generate y_t, but the γ_k's apply to $(1 - L)\, y_t$.

Proof: If the autocovariances of y_t exist, consider terms in

$$\gamma_k = E\left[\sum_{i=0}^{\infty} g_{t-i,t}\epsilon_{t-i}\right]\left[\sum_{j=k}^{\infty} g_{t-j,t-k}\epsilon_{t-j}\right]. \quad (5)$$

By assumption, ϵ_{t-i} and ϵ_{t-j} are independent for $i \neq j$, which implies that terms of the form $E(g_{t-i}\epsilon_{t-i}g_{t-j}\epsilon_{t-j})$ for $i \neq j$ are zero. Survival of ϵ_{t-i} to period t clearly implies survival to period $t - k$ for $t - i < t - k \leq t$ (that is, $g_{t-i,t} = 1$ implies $g_{t-i,t-k} = 1$). Terms of the form $E(g^2_{t-i,t}\epsilon^2_{t-i})$ simplify because $E(g^2_{t-i,t}) = V(g_{t-i,t}) + E(g_{t-i,t})^2 = p_i(1 - p_i) + p_i^2 = p_i$, $E(\epsilon^2_{t-i}) = \sigma^2$, and ϵ_{t-i} and $g_{t-i,t}$ are independent.

If the autocovariances of $(1 - L)y_t$ exist, consider $\text{cov}\,(y_t - y_{t-1}, y_{t-k} - y_{t-k-1})$. The only error that can have a nonzero impact on $y_t - y_{t-1}$ and on $y_{t-k} - y_{t-k-1}$ is ϵ_{t-k}, which becomes nonzero between periods $t - k - 1$ and $t - k$ and can return to zero between periods $t - 1$ and t. The probability that $g_{t-k,t-1} = 1$ and $g_{t-k,t} = 0$ is $p_{k-1} - p_k$, and, in that event, $(y_t - y_{t-1})(y_{t-k} - y_{t-k-1}) = -\epsilon^2_{t-k}$. □

Several properties of an error-duration process do not require any further structure on the survival probabilities. If the parameter

$$\lambda = \sum_{i=1}^{\infty} p_i \quad (6)$$

is finite, then the variance of the process is given by $\sigma^2(1 + \lambda)$, the expected number of errors surviving at any point equals λ, the mean duration of any particular error is given by

$$\lambda = \sum_{i=0}^{\infty} i(p_i - p_{i+1}),$$

and the first-order autocorrelation is $\rho_1 = \lambda/(1 + \lambda)$. If λ is not finite, then the expected number of errors surviving at any point is infinite and y_t is nonstationary. This condition for nonstationarity is much weaker than the condition for a unit root, which is that all errors last forever (or $p_k \equiv 1$).

III. Long Memory

Whether or not an error-duration model generates a process with long memory also depends on the survival probabilities. The McLeod-Hipel (1978) definition of long memory is that $\lim_{n\to\infty} \Sigma_{j=-n}^{n}\gamma_j$ is not finite.

PROPOSITION (2). The process y_t has long memory if $\lim_{n\to\infty} \Sigma_{k=1}^{n} kp_k$ is not finite.

Proof: The result obtains by direct substitution of equation (2) into the definition. □

This result shows directly that the survival probabilities $p_k = k^{-2+2d}$ generate long memory for $0 < d \leq 1$. More generally, proposition (2) shows that long memory arises if

$$\frac{p_k}{c_p k^{-2+2d}} \to 1 \text{ as } k \to \infty, \quad (7)$$

for $0 < d \leq 1$ and some constant c_p. This result could also be derived using equations (3) and (4) and the characterization of long memory processes that the autocorrelations (if they exist) follow

$$\frac{\gamma_k}{c_\gamma k^{-1+2d}} \to 1 \text{ as } k \to \infty, \quad (8)$$

where c_γ is some constant and $d > 0$. Long memory thus generally requires that the survival probabilities p_k approach zero more slowly than k^{-2}. If they approach zero more slowly than k^{-1}, then y_t is long memory and nonstationary.

The conditional survival probabilities p_{k+1}/p_k converge to one for a long-memory process. To see this, note that, if p_{k+1}/p_k were bounded from above by $\phi < 1$, then p_k would go to zero faster than ϕ^k, but $\Sigma_{k=0}^{\infty} k\phi^k = \phi/(1 - \phi)^2$ is finite for $-1 < \phi < 1$. (This is the AR(1) case discussed below.) The conditional survival probabilities p_{k+1}/p_k can, however, converge to one even if the process is not long memory. If $p_k = k^\alpha$ for $\alpha < -2$, then p_{k+1}/p_k converges to one, but the summation in proposition 2 is finite.

IV. Fractional Integration

The fractionally integrated process $(1 - L)^d y_t = \eta_t$ where η_t is i.i.d. can be explained intuitively within the error-duration framework by applying proposition 1 to obtain the following result.

PROPOSITION (3). For $0 \leq d \leq 1$, the survival probabilities for an $I(d)$ process are

$$p_k = \frac{\Gamma(k + d)\Gamma(2 - d)}{\Gamma(k + 2 - d)\Gamma(d)}. \quad (9)$$

Proof: For $d < 1/2$, the autocovariances are

$$\gamma_k = \frac{\Gamma(k + d)\Gamma(1 - 2d)}{\Gamma(k + 1 - d)\Gamma(1 - d)\Gamma(d)}\sigma_0^2, \quad (10)$$

where σ_0^2 is an error variance (Granger & Joyeux, 1980). For $0 < d < 1/2$, equation (9) follows from equation (3). For $1/2 \leq d < 1$, equation (9) follows from (4) by showing that the autocovariances from (10) for $d - 1$ are the first differences of the probabilities from equation (9). Specifically, for $0 < d < 1/2$, the factors involving k yield

$$\frac{\Gamma(k+d)}{\Gamma(k+1-d)} - \frac{\Gamma(k+d+1)}{\Gamma(k+2-d)} = (1-2d)\frac{\Gamma(k+d)}{\Gamma(k+2-d)}.$$

For $1/2 \leq d < 1$, the factors involving k yield

$$\frac{\Gamma(k+d)}{\Gamma(k+2-d)} - \frac{\Gamma(k+d-1)}{\Gamma(k+1-d)} = (2d-2)\frac{\Gamma(k+d-1)}{\Gamma(k+2-d)}. \quad \square$$

These survival probabilities generate an $I(d)$ process precisely. The approximation (7) can be verified using $c_p = \Gamma(2-d)/\Gamma(d)$ and a variation of Stirling's formula $k^{2(1-d)} \cdot \Gamma(k+d)/\Gamma(k+2-d) \to 1$ as $k \to \infty$. (Abramowitz & Stegun (1972, p. 257).

The $I(1/2)$ process on the boundary between stationarity and nonstationarity is an important special case. The survival probabilities for an $I(1/2)$ process are

$$\{p_k\} = 1, 1/3, 1/5, 1/7, 1/9, \ldots.$$

The variance of y_t is infinite because the sums of this well-known series diverge. The MA representation parameters are

$$\{\psi_k\} = 1, 3/8, 15/48, 105/384, 945/3840, \ldots.$$

The variance of y_t is infinite because the sums of the squares of this series diverge. Although both representations are mathematically valid, trying to guess the next term in each series suggests that the ED representation gives a more intuitive view of the boundary between stationarity and nonstationarity.

The ED representation adds a new perspective on the difference between fractionally integrated processes and ARMA models. The $I(d)$ survival probabilities in equation (9) follow the recursion

$$p_{k+1} = \frac{k+d}{k+2-d}p_k \tag{11}$$

with $p_0 = 1$. For an AR(1) model $(1-\phi L)y_t = \eta_t$ with η_t i.i.d. and $0 \leq \phi < 1$, the autocorrelations are $\rho_k = \phi^k$, and normalizing to $p_0 = 1$ shows that the survival probabilities are $p_k = \phi^k$. These survival probabilities follow the recursion

$$p_{k+1} = \phi p_k.$$

For large k, the conditional survival probabilities p_{k+1}/p_k for the $I(d)$ model approach unity, while those for the AR(1) model are constant at ϕ.

The AR(1) model with $\phi = 0.500$ and the $I(d)$ model with $d = 0.333$ illustrate the differences. Both models have first-order autocorrelations $\rho_1 = 0.500$ and average error durations $\lambda = 1.000$. The AR(1) model's survival probabilities for $k = 1, \ldots, 6$ are $p_k = 1/2^k$ or 0.500, 0.250, 0.125, 0.063, 0.031, 0.016, ... The $I(d)$ model's survival probabilities are 0.200, 0.100, 0.064, 0.045, 0.035, 0.028, ... The latter series starts out lower, but decreases much more slowly toward zero. The probabilities that an error lasts at least ten periods are $p_{10} = 0.001$ for the AR(1) process and $p_{10} = 0.015$ for the $I(d)$ process. The effect of these survival probabilities is apparent in the autocorrelation functions. For the AR(1) process, the values $\{\rho_1, \rho_{10}, \rho_{20}, \rho_{100}\}$ are $\{0.500, 10^{-3}, 10^{-6}, 10^{-30}\}$. The same autocorrelations for the $I(d)$ process are $\{0.500, 0.232, 0.184, 0.108\}$. This dramatic difference in high-order autocorrelations is due—in the context of the error-duration model—to the high conditional survival probabilities for $I(d)$ errors that survive the first few periods.

V. Examples

Having an elegant representation is not nearly as important, of course, as having an economic mechanism, and that issue has been an important cause of the mystery surrounding fractional integration. The following two examples illustrate how aggregate employment and financial asset volatility might exhibit long memory as a result of the survival probabilities for businesses and the survival probabilities for asset positions. In both cases, the focus is on providing a simple example rather than a complete, realistic model.

A. *Example 1: Aggregate Employment*

Suppose that a firm is created in period s, and $g_{s,t}$ is an indicator function for whether that firm is still in business in period t. Let ϵ_s be the effect of this firm on aggregate employment, and assume that this effect is independent of $g_{s,t}$ and constant as long as the firm is in business. Assume that aggregate employment is given by

$$E_t = \sum_{s=-\infty}^{t} g_{s,t}\epsilon_s + \overline{E},$$

TABLE 1.—SURVIVAL RATES S_k FOR U.S. BUSINESS ESTABLISHMENTS

Year k	1	2	3	4	5	6	7	8	9	10
Empirical S_k	0.812	0.652	0.538	0.461	0.401	0.357	0.322	0.292	0.266	0.246
$1.25k^{-2+2(0.65)}$	1.248	0.765	0.575	0.469	0.401	0.352	0.316	0.288	0.265	0.246
$1.40k^{-2+2(0.62)}$	1.401	0.830	0.611	0.491	0.415	0.362	0.322	0.291	0.266	0.246
S_k/S_{k-1}	0.812	0.803	0.826	0.857	0.868	0.891	0.902	0.908	0.911	0.923

Note: The figures for S_k are obtained directly from Nucci (1999, table II). They are based on all 5,727,985 active U.S. business establishments in the 1987 U.S. Bureau of the Census Standard Statistical Establishment List. Nucci measures survival rates between 1987 and 1988 conditional on the age of the firm as determined using earlier SSEL data. He then applies the conditional survival rates to a new firm to obtain the empirical survival rates shown here.

where $\overline{E}$ is a constant term.[3] The question of whether total employment is short memory or long memory depends on the survival probabilities for firms. Employment will be a long-memory process if the survival probabilities p_k for firms approach zero more slowly than k^{-2}. If that is the case, we would expect to observe a fairly rapid turnover in new businesses, but some examples of businesses that have survived over many periods.

Table 1 presents empirical survival rates for U.S. businesses. S_k is the fraction of U.S. business establishments starting in year 0 that survive to year k. These figures certainly do not favor a short-memory interpretation. The conditional survival rates S_k/S_{k-1} clearly increase as firms age. Squaring the five-year survival rate $S_5 = 0.401$ to obtain a prediction of the ten-year survival rate would be appropriate if the business survival rates implied an AR(1) model for employment, but the actual ten-year survival rate $S_{10} = 0.246$ is 50% greater than $0.401^2 = 0.161$.

The support for a long-memory interpretation can be assessed by fitting the approximation $p_k = c_p\, k^{-2+2d}$ from equation (7) to these survival rates. The value for d can be calculated by inverting the relation $p_n/p_m = (n/m)^{-2+2d}$ to obtain[4]

$$d = 1 + \frac{1}{2}\frac{\log\,(p_n/p_m)}{\log\,(n/m)}. \tag{12}$$

Using the empirical estimates S_5 and S_{10} of p_5 and p_{10} in equation (12) produces the estimate $d = 0.65$. Using S_7 and S_{10} produces the estimate $d = 0.62$. Table 1 shows the predicted survival probabilities for $d = 0.65$ and $d = 0.62$ with c_p chosen to match S_{10}. The former function fits S_5 through S_{10} to within 0.006, and the latter function fits S_7 through S_{10} to within 0.001. The survival rates for firms are thus quite consistent with an error-duration explanation for long memory in employment.

B. *Example 2: Asset Price Volatility*

A second example shows how the survival probabilities for financial asset positions can lead to long memory in asset price volatility. Let $g_{s,t}$ be an indicator function for the event that an asset position taken in period s lasts until period t. Suppose that the conditional probability that an asset position survives an additional period increases as the position ages (perhaps following equation (7) or (11)).

While a complete model of asset pricing with heterogeneous agents is beyond the scope of this paper, a simple structure serves to illustrate the basic idea. Assume that

- The market price p_t is an average of the expectations of agents currently holding positions in that market.
- Prices will change even given no change in the set of open asset positions because the traders' information sets change over time.
- Informed traders lower volatility and uninformed traders increase volatility because agents with better information have smaller changes in their expectations. (On average, their forecasts are closer to the perfect foresight price that all expectations are converging toward.)

The volatility of returns v_t will be a function of the average quality of information of agents currently holding positions in the market:

$$\ln\,(v_t) = \sum_{s=-\infty}^{t} g_{s,t}\epsilon_s,$$

where ϵ_s measures the quality of an agent's information.[5] The log-volatility will be short memory if the current age of an asset position has no effect on the probability that it will survive for one more period. The log-volatility will be long memory if the survival probabilities of asset positions approach zero more slowly than k^{-2}. If the latter case prevails, then we would expect to see a high volume of

[3] The assumption that ϵ_s has mean zero is analytically convenient, but not essential. A zero mean is not implausible, especially if total employment is determined by the efficiency with which a target amount of output is produced. While some firms add to employment, others, by virtue of improved technology or increased capital, decrease employment.

If $E(\epsilon_t) = \mu$, then y_t can be viewed as the sum of two processes: one generated by errors $\epsilon_t - \mu$ and one generated by μ. If either of these processes has long memory, then y_t will have long memory, so the analysis for the zero mean errors $\epsilon_t - \mu$ will be sufficient to establish the existence of long memory.

[4] Equation (12) reflects an important point about time scaling. Suppose the time interval is changed from annual to quarterly so that the probability p_k of surviving k years becomes known as the probability p_{4k} of surviving $4k$ quarters. The ratios p_{4m}/p_{4n} and $4m/4n$ do not change, and the d obtained from equation (12) is invariant to the time scaling.

[5] This discussion assumes that volatility at time t is a reasonable concept. For example, if the period of observation for v_t is a day, then v_t would be computed from high-frequency price changes within that day. Andersen et al. (1999) provide an illustration of this idea.

short-term trading and a small minority of positions that are held long-term.

VI. Continuous Time

The primary goal of this paper is to present the most intuitive possible construction that matches the discrete time $I(d)$ autocorrelations precisely. The employment and asset price volatility examples, however, naturally raise the important question of whether the long-memory property will disappear for large numbers of firms or asset positions. That issue is most conveniently addressed in continuous time. This section will consider some continuous-time extensions to the error-duration model and then review some known results that imply that employment and asset price volatility can exhibit long memory even if there are thousands of businesses or asset positions.

If the condition that y_t is observed at discrete time intervals is retained, allowing errors to originate in continuous time does not present any particular complications. To keep the probability structure simple, it is important to have independence among three things: the errors ϵ_s, the durations n_s, and the mechanism that creates new errors.[6] The simplest step in that direction is to assume that y_t is observed at times $\{t\} = 1, 2, 3, \ldots$, but the errors ϵ_s originate at times $\{s\} = 1/m, 2/m, 3/m, \ldots$, where m is a positive integer. The proof strategy of proposition 1 yields

$$\gamma_k = \sum_{j=0}^{\infty} p_{k+j/m}$$

for the case that the autocovariances exist, where p_k is now defined for fractions of the observation time interval. If the errors originate stochastically in continuous time at a constant rate μ per unit of time (this is known as a Poisson process (Haight, 1967, p. 23)), then the summation above is replaced by the integral

$$\gamma_k = \int_{t=k}^{\infty} p_t \mu \, dt.$$

The survival probability function $p_k = \partial\gamma_k/\partial k$, where γ_k is given by equation (10), will then produce an $I(d)$ process.

More-fundamental issues are involved if y_t is observed in continuous time. A variety of functional central-limit theorems describe convergence of processes (some similar to the one described here) to fractional Brownian motion, which is the continuous-time counterpart to an $I(d)$ process with $1/2 < d < 3/2$. Two lines of argument show convergence from processes involving elements resembling the indicator functions $g_{s,t}$ used here. Cioczek-Georges and Mandelbrot (1995) propose a model with micropulses with survival probabilities $p_k = k^{-2+2d}$ for $1/2 < d < 1$. The process converges to fractional Brownian motion with $1/2 < d < 1$. Taqqu et al. (1997) study a model of Ethernet packets with survival probabilities $p_k = k^{-2+2d}$ for $0 < d < 1/2$, in which the number of sources of these packets goes to infinity. The aggregate process converges to fractional Brownian motion with $1 < d < 3/2$, which has increments that are fractional white noise with $0 < d < 1/2$.

These functional central-limit theorems are not needed to establish the main results in this paper because the discrete-time fractionally integrated process $I(d)$ can be constructed precisely with an error-duration model. They do, however, contribute two important results to the discussion.[7]

First, the long-memory property survives the transition to large numbers of shocks. Indeed, the transition to large numbers of shocks is fundamental to how these functional central-limit theorems construct fractional Brownian motion. This suggests that the long-memory properties for the discrete-time examples that are considered here carry over to continuous-time markets with large numbers of agents.

Second, a wide range of structures are covered by functional central-limit theorems. Neither construction discussed above, for example, multiplies the counterpart to $g_{s,t}$ by a stochastic component ϵ_s, which suggests that the error durations are fundamental to the long-memory property and the error magnitudes are not. In particular, these continuous-time results reinforce the intuition that the fundamental feature generating long memory is tail survival probabilities proportional to k^{-2+2d} for $0 < d \leq 1$.

VII. Estimation, Generalizations, and Structural Change

The ED, AR, and MA representations for an $I(d)$ process may provide some insight into the nature of the process, but they do not lead directly to practical estimation procedures.[8] The ED representation does, however, shed some light on the properties of some estimation methodologies in the presence of long memory.

The class of ARMA and ARFIMA models is more general than the class of ED models because the autocorrelations of an ED model must satisfy two constraints. To yield a monotone series $\{p_k\}$, the autocovariance function must be monotone $\gamma_k \geq \gamma_{k+1}$. To yield $p_k \geq p_{k+1}$, it must also satisfy $\gamma_{k-1} - \gamma_k \geq \gamma_k - \gamma_{k+1}$. Estimated autocorrelations will often fail to satisfy these constraints even for an $I(d)$ process because the autocorrelations are subject to considerable sampling error, especially at the higher orders. Estimating the survival probabilities from the autocorrelation function is thus not likely to be any more effective than directly estimating the AR or MA representations of an $I(d)$ model.

[6] In discrete time, allowing multiple errors to be created at a given period s does not present difficulties if the process determining the number of those errors is independent of the errors and their durations.

[7] Although it is not of primary importance in this paper, fractional Brownian motion is self-similar, which means that the autocorrelation structure is the same for time aggregates over widely different observation frequencies. This property can be very helpful in identifying long memory if sufficient data is available to permit a range of time aggregation frequencies. (See Andersen et al. (1999) and Leland et al. (1994).)

[8] The employment and asset volatility examples do suggest that we might learn something useful about the time-series properties of aggregate data by studying survival probabilities for the underlying error mechanism.

The ED representation can be generalized to account for non-monotone autocorrelations. Suppose that the observed time series z_t is the sum of two components

$$z_t = x_t + y_t, \tag{13}$$

where x_t is an $I(d)$ process and y_t is an independent ARMA(p, q) process. The high-order autocovariances of z_t will be dominated by x_t, but the low-order autocovariances can have the flexibility of the ARMA component. The sum (13) can be expressed (in a variation on equation (1)) as

$$z_t = \sum_{s=1}^{n} \epsilon_s d_{s,t} + y_t, \tag{14}$$

where $d_{s,t}$ is one for the interval $(s, s + n_s)$ and zero otherwise, and $\{\epsilon_s\}$ is as in section II. Estimating the three sets of unknowns $\{\epsilon_s\}$, $\{n_s\}$, and the parameters of the y_t process is not possible because there are more than twice as many parameters as observations. Interpreting the errors $\{\epsilon_s\}$ as random coefficients on $\{d_{s,t}\}$ does, however, provide some insight into how long memory might affect traditional estimation techniques. Individually, each of the terms $\epsilon_s d_{s,t}$ could be thought of as representing structural change augmenting a model y_t that would otherwise explain z_t.

An empirical researcher with a model y_t and suspicions about structural change is likely to be drawn toward identifying errors in equation (14) with large magnitudes ϵ_s and long durations n_s as structural change, adding dummy variables or trimming the sample period as a result. Many of the structural changes thought to be found in economics thus might be evidence in favor of the existence of fractional integration, which will tend to produce occasional errors with very long durations.

Suppose output in an economy follows a nonstationary $I(d)$ process with annual shocks and $d = 0.6$. The expected number of twentieth-century shocks surviving to the year 2001 if $d = 0.6$ is about 4.5. Suppose there were five: mass production of automobiles and trucks, the Great Depression, the Second World War, the energy price shock in the 1970s, and the development of microelectronics. Even if these are the only five historical shocks that affect the economy in 2001, they are numerous enough to be consistent with $d = 0.6$ and nonstationarity. A researcher using the post-WWII data might, of course, treat the last two shocks as structural changes and proceed to put forward a stationary model for the remaining data.[9]

[9] The prospects for correctly interpreting the data are not good. If $d = 1/2$, in the year 2001, the expected number of surviving shocks from the previous 2,000 years is 3.782. The infinite total expected number of surviving shocks would be essentially all due to shocks more than 2,000 years old. For $d = 0.485$, which implies stationarity, the expected number of surviving historical shocks in the year 2001 is 16, but the most recent 2,000 years account for only 3.278 of that total.

VIII. Summary

The error-duration models for aggregate employment and asset price volatility help to illustrate how fractional integration and long memory might naturally occur in economic models. These examples illustrate a more general difference between the various members of the ARFIMA(p,d,q) class of representations and the ED representation.

The ARFIMA(p, d, q) representation, $\phi(L)(1 - L)^d y_t = \theta(L)\eta_t$, which includes AR and MA representations as special cases, establishes a linear relationship between η_t and y_{t+k} for all k. In this framework, the effect of η_t on y_{t+k} is the same as the effect of η_{t+i} on y_{t+k+i} for any i. If any error has an effect fifty periods into the future, then all errors have the same effect fifty periods into the future. The only stochastic feature is the magnitude of the error.

The ED representation, on the other hand, assumes that errors do not fade out uniformly, but instead have stochastic durations. For a long-memory process, most errors have only short-term effects, but occasional errors have long-term effects. The aggregate employment and asset price volatility examples are convenient because it is possible to identify businesses and asset positions as model elements that have stochastic durations.

REFERENCES

Abramowitz, M. and I. A. Stegun, ***Handbook of Mathematical Functions*** (New York: Dover Publications, Inc., 1972).

Andersen, T. G., and T. Bollerslev, "Heterogeneous Information Arrivals and Returns Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns," ***Journal of Finance,*** 52 (1997), 975–1005.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys, "The Distribution of Exchange Rate Volatility," working paper (1999).

Baillie, R. T., "Long Memory Processes and Fractional Integration in Economics," ***Journal of Econometrics,*** 73 (1996), 5–59.

Baillie, R. T., T. Bollerslev, and H. O. Mikkelsen, "Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity," ***Journal of Econometrics,*** 74 (1996), 3–30.

Baillie, R. T., C.-F. Chung, and M. A. Tieslau, "Analysing Inflation by the Fractionally Integrated ARFIMA-GARCH Model," ***Journal of Applied Econometrics,*** 11 (1996), 23–40.

Breidt, F. J., N. Crato, and P. de Lima, "The Detection and Estimation of Long Memory in Stochastic Volatility," ***Journal of Econometrics,*** 83 (1998), 325–348.

Cioczek-Georges, R. and B. B. Mandelbrot, "A Class of Micropulses and Antipersistent Fractional Brownian Motion," ***Stochastic Processes and their Applications,*** 60 (1995), 1–18.

Diebold, F. X. and G. D. Rudebusch, "Long Memory and Pesistence in Aggregate Output," ***Journal of Monetary Economics,*** 24 (1989), 189–209.

Ding, Z., C. W. J. Granger, and R. F. Engle, "A Long Memory Property of Stock Market Returns and a New Model," ***Journal of Empirical Finance,*** 1 (1993), 83–106.

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

Granger, C. W. J., "Long Memory Relationships and the Aggregation of Dynamic Models," ***Journal of Econometrics,*** 14 (1980), 227–238.

Granger, C. W. J. and R. Joyeux, "An Introduction to Long-Memory Time Series Models and Fractional Differencing," ***Journal of Time Series Analysis,*** 1 (1980), 15–29.

Haight, F. A., ***Handbook of the Poisson Distribution*** (New York: John Wiley & Sons, 1967).

Leland, W. E., M. S. Taqqu, W. Willinger, and D. V. Wilson, ''On the Self-Similar Nature of Ethernet Traffic (Extended Version),'' ***IEEE/ACM Transactions on Networking,*** 2 (1994), 1–15.

Liu, M., ''Modeling Long Memory in Stock Market Volatility,'' unpublished working paper, Duke University, (1995).

McLeod, A. I. and K. W. Hipel, ''Preservation of the Rescaled Range 1: A Reassessment of the Hurst Phenomenon,'' ***Water Resources Research,*** 14 (1978), 491–508.

Nucci, A. R., ''The Demography of Business Closings,'' ***Small Business Economics,*** 12 (1999), 25–39.

Sowell, F., ''Modeling Long-Run Behavior with the Fractional ARIMA Model,'' ***Journal of Monetary Economics,*** 29 (1992), 277–302.

Taqqu, M. S., W. Willinger, and R. Sherman, ''Proof of a Fundamental Result in Self-Similar Traffic Modeling, ***Computer Communication Review,*** 27 (1997), 5–23.

Appendix A

A Simulation Algorithm

The algorithm described in this appendix generates long-memory data directly from the error-duration model. While other methods are available for special cases such as an I(d) process, the direct approach given here may prove useful in simulating less-pure processes that are likely to arise in applications to things like employment and asset price volatility where it might prove interesting to examine the implications of, for example, dependence between $g_{s,t}$ and ϵ_s.

Step 1: Error and Duration Draws. Drawing the errors ϵ_t and the durations n_t given σ^2 and $p_0, p_1, \ldots$ for $t \geq 1$ is a straightforward process. The probabilities $p_0, p_1, \ldots, p_T$ can be calculated just once using the recursion (11). The duration n_t for the error ϵ_t can be found by drawing an independent error u_t uniformly distributed on the unit interval and setting n_t equal to the index i such that $p_i \geq u_t > p_{i+1}$.

Step 2: Drawing a Presample. A presample of size K consisting of errors $\epsilon_{1-K}, \ldots, \epsilon_0$ and durations $n_{1-K}, \ldots, n_0$ can be added by extending step 1 to include the range of probabilities $p_0, p_1, \ldots, p_{T+K}$. The error ϵ_{1-j} has no effect on $y_1, \ldots, y_T$ and need not be drawn unless $n_{1-j} \geq j$. If n_{1-j} falls between j and $j + T$, then ϵ_{1-j} should be added to $y_1, \ldots, y_k$, where $k = n_{1-j} - j$. The value of K needed to achieve a good approximation could be substantial, but the following strategy minimizes the required presample size.

Step 3: A Prehistoric Mean Approximation. As $K \to \infty$, the number of errors originating before period $1 - K$ and expiring in period t converges to a Poisson variate with mean p_{K+t}.[10] That Poisson variate can be simulated using two types of random draws. The first step is determining the number of expiring errors by comparing a uniform random variable with the Poisson probabilities.[11] The second step is drawing that number of realized expiring errors from the error distribution.

The accuracy of the Poisson approximation depends on K. The number of prehistoric errors is infinite for any K, so the relevant consideration is that the probabilities of expiring in period t are small. The sum of all the probabilities from the infinite past is p_{K+t}. This total is obviously less than one, and K can easily be chosen to make p_{K+t} small. K should also be large enough to avoid the related difficulty that the numbers of errors expiring in two different periods are not absolutely independent. A slight dependence arises because, if ϵ_s expires in period t, then the same error cannot expire in period $t + 1$. Setting K equal to T has the intuitive appeal that any approximation error involving errors that have already survived T periods is not likely to be detectable in a sample of size T. Practical experience suggests that setting K equal to the sample size T is conservative in the sense that a smaller K produces nearly identical sample statistics.

This algorithm also ignores errors that originate in period $1 - K$ or before and do not expire during the sample period. This has little effect because such errors are empirically undetectable. They merely add a constant to the observed data throughout the entire sample and thus affect only the sample mean.

[10] The Poisson approximation is appropriate even though the probabilities of expiring in period t are not equal. Haight (1967, pp. 16, 118–119) gives an interesting historical account of the debate on this issue.

[11] If the probabilities of i or more expiring errors is q_i and u is a $U(0, 1)$ random draw, then the Poisson draw is j such that $q_{j+1} < u < q_j$.