PART One

Dynamics of Commodity Price Behavior

At times the prices of many commodities display volatile behaviors. Since agricultural commodities and minerals, such as crude oil and metals, are among the fundamental inputs of our economies on the production and/or the consumption side, price volatility causes disruptions and can lead to crises. An improved understanding of the dynamics of price behavior is therefore highly desirable from a policy as well as from a consumer and supplier perspective.

Part One consists of four chapters, each written from a different point of view. In Chapter 1, the recently retired chief economist of Arcelor-Mittal and two colleagues from academia present a new method for estimating long memory processes from small samples, a common problem in industry, where forecasts frequently have to be made from very short series. This contribution provides a theoretically sound and interesting solution to a practical problem.

While the first chapter takes an industry and firm perspective, the second chapter analyzes time-series data to study the link between commodity price developments and business cycles. The question asked many times is whether commodity prices lead inflation or inflation leads commodity prices. The answer is not immediately visible from looking at the data, because trends can be obscured by short-run occurrences. This chapter's analysis offers a method to uncover the true trend and provides evidence that, on balance, commodity prices are procyclical. The exceptions are the price of gold, which is countercyclical, and the price of sugar, which is acyclical.

Chapter 3 also studies the connection between inflation and commodity prices. The author uses a recently developed procedure to test the possible presence of nonlinearity in the comovement of commodity prices and the consumer price index. The results reveal interdependences between the different price series, with policy implications, for example, on how to combat inflation.

Chapter 4 also is focused on macroeconomic issues, but the issue of interest turns from domestic policy to the world market. The chapter deals with the relationship between the dollar and the oil price. The real price of the oil in every currency depends on a variety of factors, including OPEC policy. The price of crude oil is in

U.S. dollars, but most of the imports of the largest oil-producing member countries originate in the euro zone or in Japan. Hence, the devaluation of the dollar lowers the purchase power of OPEC member countries, which they try to regain by adjusting the price of oil upward.

Together, these four chapters provide models, data, results, and insights that enhance our understanding of the dynamics of commodity price behavior. They use the most current models and techniques in time-series analysis and illustrate their application. The chapters complement each other by providing information at different levels of aggregation while dealing with the same general subject. Because of their mixed background in terms of professional experiences and geographic location, the authors also bring different perspectives to their respective tasks.

CHAPTER 1

Indirect Inference and Long Memory

A New Truncated-Series Estimation Method

Armand Sadler, Jean-Baptiste Lesourd and Vêlayoudom Marimoutou

INTRODUCTION

Long-memory processes are an important and even fundamental advance in time-series modeling. More precisely, the so-called autoregressive fractionally integrated moving average (ARFIMA) model has been introduced by Granger and Joyeux (1980) and Hosking (1981). It is a generalization of the ARIMA model, which is a short memory process, by allowing the differencing parameter d to take any real value. The goal of this specification is to capture *parsimoniously* long-run multipliers that decay very slowly, which amounts to modeling long memories in a time series. ARFIMA processes, however, are associated with hyperbolically decaying autocorrelations, impulse response weights, and spectral density function exploding at zero frequency. As noted by Brockwell et al. (1998), while a long memory process can always be approximated by an ARMA(p, q) process, the orders p and q required to achieve a good approximation may be so large as to make parameter estimation extremely difficult. In any case, this approximation is not possible with small samples.

ARFIMA processes are defined as follows in their canonical form:

$$\Phi(L)(1-L)^d y_t = \mu + \Theta(L)\varepsilon_t, \quad \varepsilon_t : iid(0, \sigma^2) \tag{1.1}$$

where $d \in (-0.5, 0.5)$ is the fractional difference operator and μ can be any deterministic function of time. If μ is zero, this process is called fractionally differenced autoregressive moving average (e.g., Fuller (1996)). The *iid* (independent and identically distributed) assumption is the strongest assumption; it implies mixing, that is, conditions on the dependence of the sequence. For a stationary sequence, mixing implies ergodicity (restrictions on the dependence of the sequence). Ergodic processes are not necessarily mixing; mixing conditions are stronger than ergodicity. For details, see White (1984) and Rosenblatt (1978).

For general overviews on long memory processes, surveys, and results, we refer the reader to Baillie (1996); Brockwell and Davis (1998); Fuller (1996); Gouriéroux and Monfort (1995); Gourieroux and Jasiak (1999); Hamilton (1994); Jasiak (1999, 2000); Lardic and Mignon (1999); Maddala and Kim (1998); and Sowell (1990) as well as to the discussions and comments by Bardet (1999), Bertrand (1999), Gourieroux (1999), Jasiak (1999), Lardic and Mignon (1999), Prat (1999), Renault (1999), Taqqu (1999), and Truong-Van (1999). Concerning recent research on the topic of long memory, we refer the reader to Andrews and Guggenberger (2003), Andrews and Sun (2004), and Davidson and Terasvirta (2002). Note also the presentation of a new stationarity test for fractionally integrated processes by Dolado, Gonzalo, and Mayoral (2002). Among the most important papers concerning estimation techniques for these ARFIMA model are Fox and Taqqu (1986), Geweke and Porter-Hudak (1983), Li and McLeod (1986), and Sowell (1992a). Tests for long memory across a variety of commodity spot and futures prices can be found in Barkoulas, Labys, and Onochie (1997, 1999) as well as in Cromwell et al. (2000).

The methods for estimating d, the long-range dependence parameter, can be summarized in three classes:

1. The heuristic methods [the Hurst (1951) method, the Lo (1989, 1991) method, the Higuchi (1988) method]
2. The semiparametric methods [Geweke and Porter-Hudak (GPH) (1983) method, the Robinson (1983, 1995a, 1995b) estimation methods]
3. The maximum likelihood methods [the exact maximum likelihood method, the Whittle (1951) approximate maximum likelihood method]

For a comparison of these classes of estimators, refer to Boutahar et al. (2005).

The estimation of fractional integration exponents leads to significant problems in some cases. In the case of small samples, as often encountered with industrial data, it is even impossible. Long-memory estimations often are performed with financial time series with large numbers of observations (5,000 observations and more are not uncommon). However, small samples of 50 to 100 observations are the order of magnitude usually encountered in industrial forecasting problems. In such cases the need for a consistent and precise estimation technique is of great interest. Thus, we motivate the need for a new estimator for the long-memory parameter by the small sample sizes often encountered in practice. Why should we care about long memory in those situations? For instance, one could argue that from a forecasting perspective, long memory starts to make a difference only when forecasting over long horizons. In situations when you only have a few observations available, you would not forecast too many steps ahead. The reply to this comment covers three aspects:

1. What really matters in time series analysis is the span, not the number of observations. Fifty yearly observations on apparent steel use in a region have another informational content than 5,000 real-time observations over a short period of time on some financial stock index.
2. Many industry sectors are producing medium- and long-range forecasts based on a relatively small number of yearly or quarterly observations. A steel producer planning to invest in a new rolling mill or a new greenfield facility can not wait for a long time series before making a decision but has to work with the actually available data.

3. We have come to believe from our past studies that for transfer function models (models with explanatory variables), long memory does not even exist. Detected long memory always followed some misspecification of the actual model. If a model is correctly specified, long memory should disappear. In this sense, we look at long memory as a specification test.

By using indirect inference to adjust for the bias, is the computational burden increasing? The answer is no. We suggest the use of our reference tables to correct for the bias. To our knowledge, since the work of Li and McLeod (1986), no new estimation techniques that are valid for small samples have been proposed in the econometric literature. Moreover, Li and McLeod developed an estimation technique based on truncating the power series defining the process after about 50 terms.

In this chapter, we propose a completely different approach based on low-order truncation (after about five terms). Li and McLeod considered their truncated model as approximating the true model, whereas we explicitly consider our low-order truncated model as an instrumental model that is necessarily biased. The bias is corrected by an indirect inference technique, through minimizing a distance function.

This chapter aims at defining an estimation technique of the fractional integration exponent d for comparatively small samples. Its asymptotical properties are based on a result established by Mira and Escribano (2000) about the almost sure consistency of a nonlinear least square (NLS) estimator. The hypotheses used by these authors are shown to apply to our particular case of truncated series. A new method for identification and estimation of these truncated series is developed and applied to steel consumption time series as well as to the analysis of atmospheric carbon dioxide (CO_2) concentrations derived from in situ air measurements at Mauna Loa Observatory, Hawaii.

ALMOST SURE CONSISTENCY OF THE NLS ESTIMATOR FROM OUR TRUNCATED MODEL

We consider the simplest ARFIMA process, also called fractionally differenced (or integrated) white noise (see, e.g., Fuller (1996) or Brockwell et al. (1998)):

$$(1-L)^d\, y_t = e_t \tag{1.2}$$

with $e_t \sim iid$, or $y_t + \sum_{j=1}^{\infty} \kappa_j(d) y_{t-j} = e_t$.

$$\kappa_j(d) = [\Gamma(j+1)\Gamma(-d)]^{-1}\,\Gamma(j-d) = \prod_{i=1}^{j} i^{-1}(i-1-d)$$

Γ is the gamma function and $d \in (-0.5, 0.5)$. If ξ is not an integer, then $\xi = n+\phi,\ \phi \in (0,1)$: $\Gamma(\xi) = (\phi+n-1)(\phi+n-2)\ldots(\phi+1)\phi\Gamma(\phi)$. We define the truncated version of this model by:

$$y_t + \sum_{j=1}^{r} \kappa_j\,(d) y_{t-j} = e_t \tag{1.3}$$

In addition, we relax the *iid* assumption for e_t and replace it by the less restrictive α-mixing assumption, thus allowing for some heteroscedasticity (see White (1984) for details). In the appendix we show that the parameter d of model (1.3) can be consistently estimated and that the true fractional parameter can be estimated by indirect inference.

In the next section, we show how to identify and estimate the truncated long memory process.

IDENTIFICATION AND ESTIMATION OF THE TRUNCATED LONG MEMORY PROCESS

We define the combined consumption model (CCM) as a transfer function model including long memory. The starting point is either a cointegration relationship between variables having a common stochastic trend or a stable relationship between stationarized variables. In the case of structural breaks, the break may be in level, in slope, or in both. Care has to be taken with the specification because, as pointed out by Diebold and Inoue (2001), long memory and structural breaks are easily confused.

The CCM is thus aiming at a parsimonious representation of reality by focusing on a few key explanatory variables, an ARMA part in order to take account of short memory and a fractional parameter representing long memory. With this definition, the estimated parameter d will always lie in the open interval (–0.5, 0.5). An estimated parameter d out of that range is an indication that the series have not been correctly stationarized because the process is only both stationary and invertible if $d < 10.51$.

If long memory is specified by a truncated version of the model, the CCM can be estimated easily. The next estimation procedure follows the outline proposed by Hosking (1981), except that we change the order of the steps and estimate the combined model. To illustrate the estimation procedure, let us start with the ARFIMA model

$$\Phi(L)\left[y_t + \sum_{j=1}^{r} \kappa_j(d)\, y_{t-j}\right] = \Theta(L)\,\varepsilon_t \tag{1.4}$$

or

$$F(L)\,\nabla^d y_t = \Theta(L)\,\varepsilon_t \tag{1.5}$$

Define

$$u_t = y_t + \sum_{j=1}^{r} \kappa_j(d) y_{t-j} \tag{1.6}$$

so that $\{u_t\}$ is an ARIMA $(p, 0, q)$ process.

$$\Phi(L)[u_t] = \Theta(L)\,\varepsilon_t \tag{1.7}$$

Let

$$x_t = \{\Theta(L)\}^{-1}\,\Phi(L)\,y_t \tag{1.8}$$

so that x_t is a truncated ARIMA (0, d, 0) process because

$$\nabla^d x_t = \{\Theta(L)\}^{-1}\,\Phi(L)\,\nabla^d y_t = \varepsilon_t \tag{1.9}$$

d is estimated in four steps:

1. Start the algorithm by setting $d = 0$ in (1.4) and estimating the ARMA parameters by the Gauss-Newton algorithm.
2. Take the residuals (z_t) from the equation in step 1 and check if the series displays long memory; in other words, estimate d.
3. Calculate u_t from equation 1.6 with the d estimated in step 2.
4. Reestimate the ARMA parameters from equation 1.7 and check for convergence. If not converged, reestimate d and go to step 3.

Adding additional exogenous explanatory variables poses no problem in this estimation procedure. There is generally no convergence problem in applying this procedure, except when the sample size is really too small.

INFORMAL PROOF OF THE CONVERGENCE OF THE ABOVE ESTIMATION PROCEDURE

In practice, we are running a conditional loop. The instructions in the loop are executed repeatedly until a specified condition is true. In our case, the condition is that the distance between successive values of the respective parameters in successive runs gets arbitrarily small (the Cauchy criterion of convergence of a sequence). Suppose the true DGP (data generating process) is given by formula 1.4 and that d is positive. By setting $d = 0$ in step 1, the ARMA parameters are capturing *partly* the impact of a missing explanatory variable and are distorted. But the ARMA specification cannot capture long memory. Thus, the residuals (z_t) in step 2 are not *iid* and the estimated d in this step is necessarily positive, given our assumption. In step 4, the reestimated ARMA parameters are closer to reality as we are taking into account the new estimated d. By proceeding further in this way, the distortions are becoming smaller and smaller. The procedure is converging.

Monte Carlo Simulation and Indirect Inference Estimation of an ARFIMA (0,*d*,0)

In this section, we use a methodology called indirect inference to demonstrate the usefulness of our approach. This methodology, introduced by Gouriéroux, Monfort, and Renault (1993), Smith (1993), and Gallant and Tauchen (1996) is nowadays largely used in applied econometric research. The idea is to draw a simulation-based inference on generally intractable structural models through an instrumental model, conceived as easier to handle. We refer the reader to Gouriéroux and Monfort (1997) for a detailed description of the methodology.

The initial model (M) is equation 1.2 and the approximated one (M^a) is equation 1.3. The estimator is obtained by minimizing equation A.1 (see appendix) by nonlinear least squares.

Equation A.1 may be written

$$Q_n(d) = n^{-1} \sum_{t=1}^{n} f^a(\underline{y_t}, d),$$

whereas the initial model is

$$Q_n(d) = n^{-1} \sum_{t=1}^{n} f(\underline{y_t}, d)$$

with $\underline{y_t}$ denoting the present and past values y_t, y_{t-1}, y_{t-2}, and so on, of the process y. Let d be the true value of the parameter and d^a the estimated parameter of the instrumental model. The estimated parameter d^a does not generally converge toward the true parameter d because $f(\underline{y_t}, d) \neq f^a(\underline{y_t}, d)$.

First we will show that the asymptotic bias is a function of d. We simulated different ARFIMA (0,d,0) models by using a RATS program written by Schoen (1997). The fractionally integrated parameter d (the true d) is estimated by considering the estimator d^a obtained by minimizing formula 1.3 with six lags and with large samples. d^a, reported in row 2 in Tables 1.1 to 1.5, is the arithmetic mean of 10,000 (respectively 50,000) parameters estimated from each simulation. We observe an important bias when approaching the nonstationary case for positive d. The gain in precision is significant for increasing n.

A visual representation of the relationship between the value of the true d and the estimated value d^2 is provided in Figure 1.1. It shows that the bias is a function of d.

We next test the sensitivity of estimates as a function of the lag, r. Table 1.5 shows simulation results when $r = 3$.

It follows from Tables 1.1 to 1.5 that the estimated fractional parameter d^a does not differ significantly from the true d, except for values of d approaching the extreme points of the open interval (–0.5,0.5). Nevertheless, the bias has to be

TABLE 1.1 Bias as a Function of d with 1,000 Observations and 10,000 Simulations

True d	**–0.49**	**–0.4**	**–0.3**	**–0.2**	**–0.1**	**0.1**	**0.2**	**0.3**	**0.4**	**0.49**
d^a	–.405	–0.345	–0.271	–0.188	–0.098	0.101	0.212	0.333	0.474	0.706
Std Error	0.035	0.036	0.036	0.037	0.037	0.037	0.037	0.036	0.035	0.029
Stat T	2.4	1.5	0.8	0.3	0.05	0.03	0.3	0.9	2.1	7.5

TABLE 1.2 Bias as a Function of *d* with 500 Observations and 50,000 Simulations

True *d*	–0.49	–0.4	–0.3	–0.2	–0.1	0.1	0.2	0.3	0.4	0.49
d^a	–0.405	–0.345	–0.271	–0.189	–0.101	0.099	0.208	0.329	0.467	0.70
Std Error	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.04
Stat T	1.7	1.1	0.6	0.2	–0.02	–0.02	0.2	0.6	1.3	5.3

TABLE 1.3 Bias as a Function of *d* with 100 Observations and 10,000 Simulations

True *d*	–0.49	–0.4	–0.3	–0.2	–0.1	0.1	0.2	0.3	0.4	0.49
d^a	–0.41	–0.35	–0.28	–0.20	–0.11	0.08	0.19	0.31	0.44	0.69
Std Error	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.10
Stat T	0.7	0.4	0.2	0.0	–0.1	–0.2	–0.1	0.1	0.3	2.0

TABLE 1.4 Bias as a Function of *d* with 50 Observations and 10,000 Simulations

True *d*	–0.49	–0.4	–0.3	–0.2	–0.1	0.1	0.2	0.3	0.4	0.49
d^a	–0.416	–0.360	–0.291	–0.213	–0.130	0.060	0.163	0.280	0.421	0.680
Std Error	0.18	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.18	0.15
Stat T	0.4	0.2	0.05	–0.07	–0.2	–0.2	–0.2	–0.1	0.1	1.3

TABLE 1.5 Estimates of d^a as a Function of Lag when $r = 3$

True *d*	–0.49	–0.4	–0.3	–0.2	–0.1	0.1	0.2	0.3	0.4	0.49
d^a	–0.379	–0.326	–0.262	–0.187	–0.104	0.088	0.200	0.326	0.478	0.751
Std Error	0.12	0.12	0.12	0.12	0.125	0.125	0.124	0.121	0.115	0.088
Stat T	0.9	0.6	0.3	0.1	–0.03	–0.1	0.0	0.2	0.7	3.0

n = number of observations = 100; 10,000 simulations.

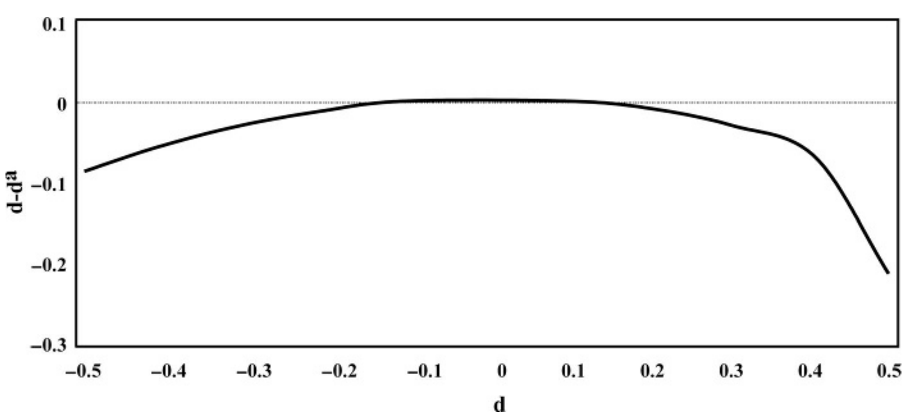

FIGURE 1.1 Bias of d^a as a Function of *d*

corrected. This is done by minimizing the following distance (Gouriéroux, Monfort and Renault (1993); Gouriéroux and Monfort (1997); and Smith (1993)):

$$d_{aSn}(\Omega) = \arg\min\,[d_{an} - d_{aSn}(d)]'\,\Omega\,[d_{an} - d_{aSn}(d)] \tag{1.10}$$

$$d \in (-0.5, 0.5) \quad \text{with} \quad d_{an} = \arg\min Q_n(d) = n^{-1}\sum_{t=1}^{n} f^a(y_t, d),$$

$$d_{aSn}(d) = (1/S)\sum_{s=1}^{S} d_{an}(d)$$

In short, d_{an} is the estimator of d obtained by maximizing the instrumental criterion in the case of a given sample of interest whereas $d_{aSn}(d)$ is the arithmetic mean of the maximization of the same instrumental criterion for the S simulated samples. So, $d_{aSn}(d)$ is what we report in Tables 1.1 to 1.5. In order to maximize the asymptotic covariance matrix of an M-estimator like $d_{aSn}(\Omega)$, Gouriéroux et al. have shown that the optimal choice of Ω is:

$$\Omega^* = J_0\,(I_0 - K_0)^{-1}\,J_0 \tag{1.11}$$

We use the notation proposed in Dridi and Renault (2000). In the absence of additional exogenous variables, as in the case here, $K_0 = 0$ (Gouriéroux et al., 1996). Gouriéroux et al. (1993) noted that the efficiency gain obtained by using the optimal estimator is negligible (and that for practical applications they) only consider the estimator based on $\Omega = Id'$. Thus, (10) is simplified to

$$d_{aSn}(\Omega) = \arg\min\,[d_{an} - d_{aSn}(d)]^2 \tag{1.12}$$

$$d \in (-0.5, 0.5)$$

Suppose the following simple analytical form for the relationship between the true d and $d_{aSn}(d)$: $d = d_{aSn}(d) - 0.1$.

$$d_{aSn}(\Omega) = \arg\min\left[d_{an}^2 - 2d_{an}(d + 0.1) + (d + 0.1)^2\right] \tag{1.13}$$

By minimizing the above distance with respect to d we obtain:

$$-2d_{an} + 2\,(d + 0.1) = 0 \text{ and } d = d_{an} - 0.1. \tag{1.14}$$

As the analytical form of the relationship is unknown, we correct on the basis of Tables 1.1 to 1.4.

We will show that this procedure works remarkably well by simulating different ARFIMA (0, *d, 0*) series with $n = 500$ observations. The corresponding fractional parameter is then estimated by indirect inference (II) and by the method of Geweke and Porter-Hudak (GPH). The results are reported in Tables 1.6 and 1.7.

The results from the two simulations are presented in Figure 1.2. Note that the results diverge toward the endpoints of the (–0.5, +0.5) open interval and that the

TABLE 1.6 Estimation by Indirect Inference

True d	–0.49	–0.4	–0.3	–0.2	–0.1	0.1	0.2	0.3	0.4	0.49
Estimated d	–0.463	–0.379	–0.285	–0.191	–0.096	0.103	0.212	0.321	0.444	0.558
Std error	0.053	0.053	0.054	0.054	0.054	0.054	0.054	0.052	0.049	0.038
Bias	0.027	0.021	0.015	0.009	0.004	0.003	0.012	0.021	0.044	0.068

n = number of observations = 500; 1000 replications.

TABLE 1.7 Estimation by GPH

True d	–0.49	–0.4	–0.3	–0.2	–0.1	0.1	0.2	0.3	0.4	0.49
Estimated d	–0.387	–0.352	–0.267	–0.197	–0.107	0.106	0.217	0.317	0.49	0.79
Std Error	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.16	0.17	0.17
Bias	0.1	0.05	0.03	0.003	–0.07	0.006	0.017	0.017	0.09	0.30

n = number of observations = 500; 1000 replications.

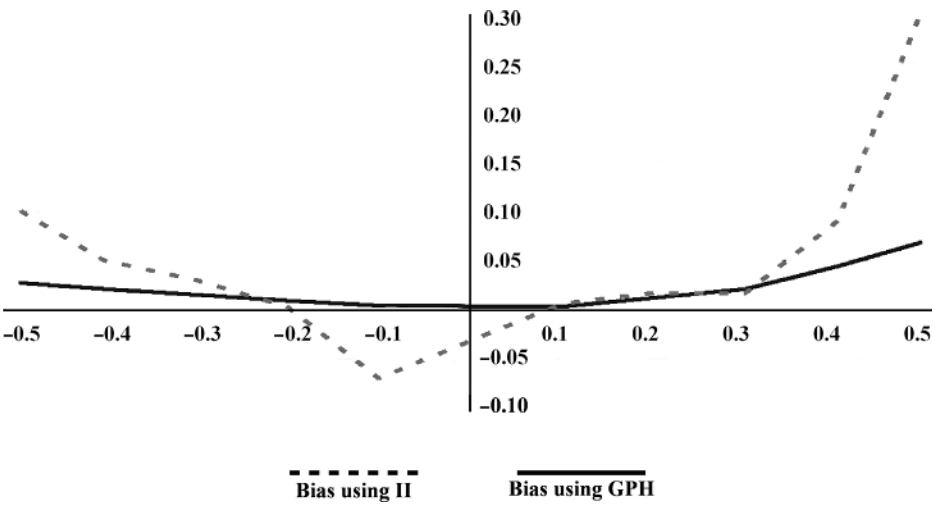

FIGURE 1.2 Bias of d^a for the Two Methods

indirect inference (II) method generates better estimates than the method suggested by GPH. The standard error and the bias are much smaller for the II method.

APPLICATIONS

In this section, we apply the estimation procedure described in the previous section to the analysis of apparent steel consumption (ASC) in the European Union 15, North America, Japan, and China.

Apparent Steel Consumption and Industrial Production

The starting point is the demand equation from the traditional standard commodity model (SCM) (see G. Adams (1996)). This demand equation is obtained from the first-order conditions of cost minimization by the firm. The explicit demand function for factor x_i for a given firm may be written:

$$x_i = h_i\,(p_1, p_2, \ldots, p_i, \ldots, p_n, q) \tag{1.15}$$

where x_i = demand of commodity i
q = the production of the firm
p_i = price of commodity i

By aggregating over the total number of firms in a country, the demand function of commodity i may be written:

$$D_t = D\,(P_t, PS_t, Y_t) \tag{1.16}$$

where P = price of commodity i
PS = price of competing commodities
Y = production of the sectors consuming commodity i

The commodity steel (x_i) is widely used in most production sectors so that it is reasonable to replace Y with the index of industrial production (IP) of the country. This approach has been taken by Afrasiabi, Moallem, and Labys (1991) in their study of the demand for copper, zinc, and lead. The properties of the global demand function of a given production factor are generally:

$$\partial D/\partial P < 0, \partial D/\partial/PS > 0,\ \partial D/\partial Y > 0 \tag{1.17}$$

For this exercise, prices have been removed from the equations, mainly because of a common stochastic trend in the aluminum price and steel price series. Table 1.8 summarizes the descriptive statistics of the endogenous and exogenous series. Figures 1.3 and 1.4 show the growth patterns of these series over the period 1974 to 2006.

Data come from three sources: Global Insight, Inc., the European Commission, and the International Iron and Steel Institute (IISI). The period studied were the years

TABLE 1.8 Descriptive Statistics, 1974–2006

Series	Obs	Mean	Std Error	Minimum	Maximum	$H_0 = I(1)$
LEU15	33	4.78	0.13	4.5	5.0	accept
LAMERNOR	33	4.75	0.15	4.4	5.0	accept
LJAPAN	33	4.28	0.13	3.9	4.5	accept
LCHINE	33	4.23	0.82	2.8	5.8	accept
IPEU15	33	1.6	2.6	–6.3	6.9	reject
IPAMNOR	33	2.6	3.7	–8.7	9.1	reject
IPJAPAN	33	2.0	5.0	–11.0	11.1	reject
IPCHINE	33	12.1	4.7	0.7	21.6	reject

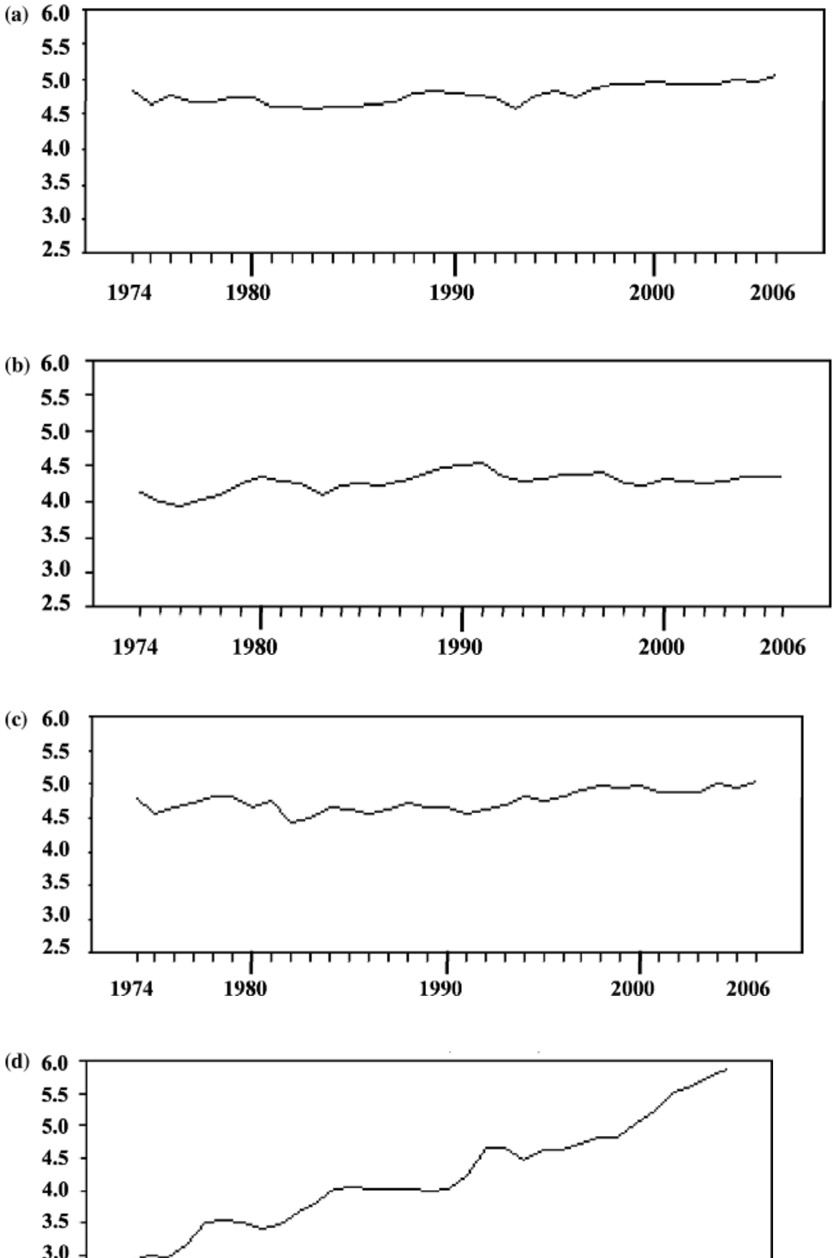

FIGURE 1.3 (a) ASC in EU15 (in logs, LEU15), 1974–2006. (b) ASC in Japan (in logs, LJAPAN), 1974–2006. (c) ASC in North America (in logs, LAMERNOR), 1974–2006. (d) ASC in China (in logs, LCHINE), 1974–2006

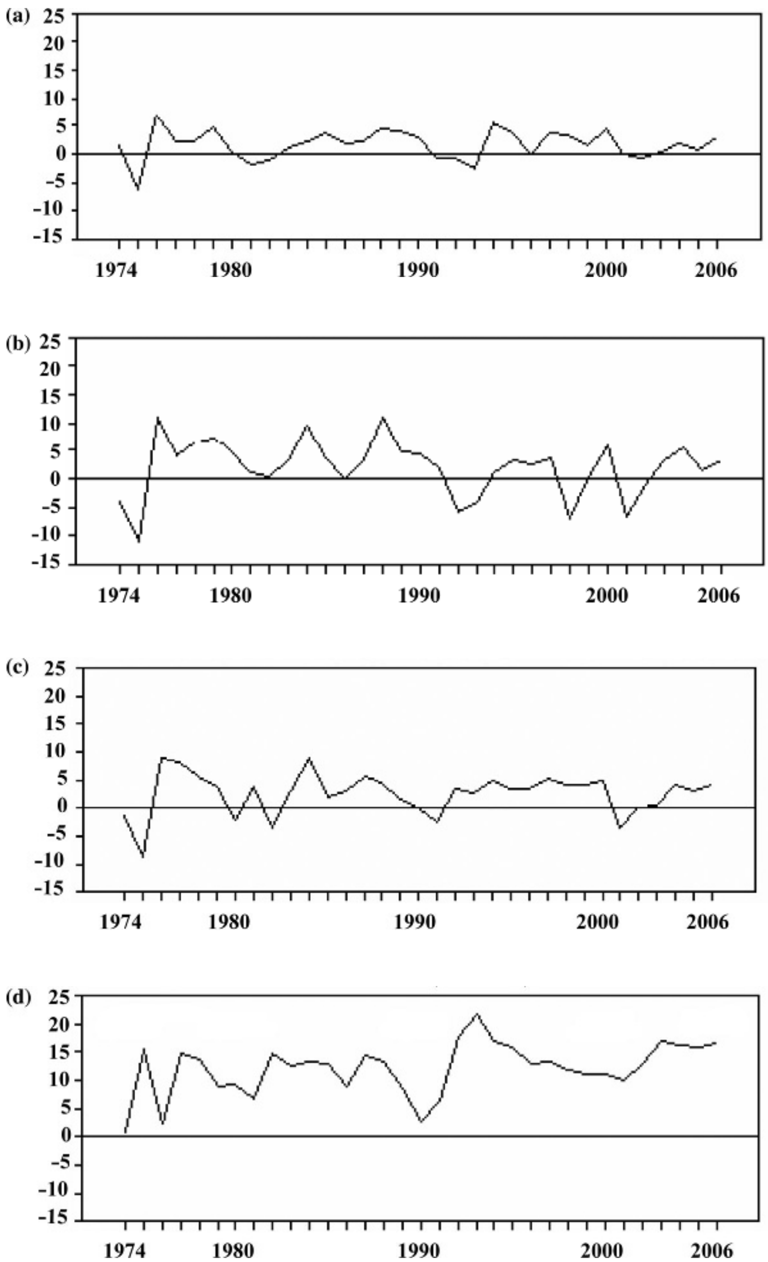

FIGURE 1.4 (a) Industrial Production in EU15 (IPEU15), 1974–2006. (b) Industrial Production in Japan (IPJAPAN), 1974–2006. (c) Industrial Production in North America (IPAMNOR), 1974–2006. (d) Industrial Production in China (IPCHINE), 1974–2006

1974 to 2006, and the data were obtained for each year. For this study we defined these variables:

- EU15 = apparent steel consumption in the European Union 15, in million metric tons
- AMERNOR = apparent steel consumption in North America (US, Canada, and Mexico), in million metric tons
- JAPAN = apparent steel consumption in Japan, in million metric tons
- CHINE = apparent steel consumption in China, in million metric tons
- LEU15 = log(EU15)
- LAMERNOR = log(AMERNOR)
- LJAPAN = log(JAPAN)
- LCHINE = log(CHINE)
- D = (1 − L); e.g., DLEU15 = (1 − L)*LEU15 = $LEU15_t - LEU15_{t-1}$
- IP = industrial production growth rate, year on year; e.g., IPEU15 = industrial production growth rate in the EU15

Figures 1.3a to 1.3d show the apparent steel consumption in the European Union (EU15), North America, Japan, and China from 1974 to 2006. The difference in trend between China and the other three economies stands out, with consumption in the latter growing rapidly to support the growth and development of its economy, whereas the other economies were growing at a slower pace and becoming less dependent on manufacturing; hence the flatter curves.

The trend differences shown in Figures 1.3a to 1.3d are also reflected in Figures 1.4a to 1.4d, which show industrial production in the same economies, but the differences are not as striking. Whereas the ASC trend lines are relatively smooth, the production trend lines have pronounced peaks and valleys.

The most important stylized facts of these series are:

- The ASC (apparent steel consumption) series (in logs) are I(1) whereas the growth rates of IP are I(0).
- All series display highly stochastic cycles. *Real* steel consumption (RC) equals ASC +/– stocks movements and is unobserved. The stochastic cycles in RC follow closely those in the IP series. ASC cycles, however, have much larger amplitude due to the speculative behavior on inventories held by the steel consumers (merchants, steel service centers, and final consumers, e.g., automotive, construction, mechanical engineering, domestic appliances, metal ware and tubes).

The specification of the statistical model, together with the constraints imposed by economic theory, raises some identification problems. The latter are solved by the rigorous modeling strategy proposed in the previous sections, taking into account exogenous explanatory variables, short-memory ARMA components, as well as a long-memory parameter.

Tables 1.9 to 1.12 show the results by running the CCM by country or region.

It follows from Tables 1.9 to 1.12 that the fractional parameter d is not significantly different from zero, so that the formulas can be simplified (see Tables 1.13 to 1.16). All processes are characterized by the property of short memory. Additional specifications not reproduced in this chapter show that price variables of steel and aluminium were not significant.

TABLE 1.9 European Union (EU15)

1975–2006	CONSTANT	N_IPEU15{0}	*d*
DLEU15	–0.0341	0.0258	–0.1067
(Significance)	(0.0005)	(0.0000)	(0.4623)
$\overline{R}^2 = 0.72$		Q(8-0) = 2.34	
		(0.97)	

TABLE 1.10 North America

1979–2006	CONSTANT	N_IPAMNOR{0}	Dummy (1982)	*d*
DLAMERNOR	–0.042	0.023	–0.192	–0.124
(Significance)	(0.005)	(0.000)	(0.001)	(0.503)
$\overline{R}^2 = 0.76$		(7-0) = 8.82		
		(0.26)		

TABLE 1.11 Japan

1979–2006	CONSTANT	N_IPJAPAN{0}	AR{1}	MA{1}	MA{2}	*d*
DLJAPAN	−0.018	0.015	−0.369	0.521	0.514	−0.004
(Significance)	(0.213)	(0.000)	(0.337)	(0.142)	(0.011)	(0.976)
$\overline{R}^2 = 0.63$		Q(7-3) = 3.02				
		(0.55)				

TABLE 1.12 China

1979–2006	CONSTANT	N_IPCHINE{0}	Dummy (1994–1995)	*d*
DLCHINE	–0.133	0.020	–0.254	–0.082
(Significance)	(0.027)	(0.000)	(0.001)	(0.586)
$\overline{R}^2 = 0.49$		Q(7-0) = 11.74		
		(0.109)		

TABLE 1.13 Reestimated Combined Consumption, European Union (EU15)

1975–2006	CONSTANT	N_IPEU 15{0}
DLEU15	–0.036	0.026
(Significance)	(0.0004)	(0.0000)
$\overline{R}^2 = 0.72$	Q(8-0) = 1.46	
	(0.99)	

JB = 1.67, Hansen (1992) = 0.32, DF = –5.8.

TABLE 1.14 Reestimated Combined Consumption, North America

1975–2006	CONSTANT	N_IPAMNOR{0}	Dummy (1982)
DLAMERNOR	–0.040	0.020	–0.210
(Sign)	(0.003)	(0.000)	(0.001)
$\bar{R}^2 = 0.75$		Q(8-0) = 11.3	
		(0.19)	

JB = 1.40, Hansen (1992) = 0.21, DF = –6.5.

TABLE 1.15 Reestimated Combined Consumption, Japan

1976–2006	CONSTANT	N_IPJAPAN{0}	AR{1}	MA{1}	MA{2}
DLJAPAN	–0.036	0.015	–0.570	0.637	0.510
(Sign)	(0.0211)	(0.000)	(0.044)	(0.02)	(0.009)
$\bar{R}^2 = 0.50$		Q(7-3) = 8.50			
		(0.07)			

JB = 6.91, Hansen (1992) = 0.10, DF = –6.5.

TABLE 1.16 Reestimated Combined Consumption, China

1975–2006	CONSTANT	N_IPCHINE{0}	Dummy (1994–1995)
DLCHINE	–0.117	0.018	–0.278
(Sign)	(0.034)	(0.000)	(0.000)
$\bar{R}^2 = 0.46$		Q(8-0) = 20.36	
		(0.009)	

JB = 0.167, Hansen (1992) = 0.05, DF = –4.3.

Tables 1.13 to 1.16 show the reestimated equations with the fractional parameter dropped and the analysis of the residuals from these equations as well as a test of the stability of the parameter linked to industrial production.

The results of this model are compatible with economic theory. For the EU15, North America, and China, the stochastic cycle in the ASC series seems to be entirely captured by the cyclical pattern of the explanatory variable. In Japan, however, additional ARMA parameters are needed to explain the cycle in ASC. The analysis of the residuals shows that the latter are *iid* normal, except in the case of Japan.

Steel Intensity Curve (SI Curve)

The so-called SI curve is part of the rich history of studies related to materials use in economic systems. Main references to this literature are found in Sadler (2003). The first SI curve was constructed in the late 1960s by the Committee on Economic Studies of the IISI (International Iron and Steel Institute). It relates the evolution of SI (the ratio of apparent steel consumption to gross domestic product [GDP]) to the level of economic development of a country as measured by GDP per capita (IISI 1974).

There are five stages in the development of SI:

1. Very low level before economic takeoff
2. Rapid rise
3. Leveling-off stage
4. Decline
5. Stabilization

The development at the first two stages of SI is due to changes in the economic structure of a country, mainly increases in the shares of investments and manufacturing production.

The decline at the fourth stage results from the changes in the relative importance of activity of steel-using sectors in total economic activity (Swip [Steel Weighted Industrial Production Index]/GDP) and a decline of specific steel consumption defined as "apparent steel consumption/Swip."

Apparent steel consumption (ASC) of a country A is defined as production + imports – exports. So ASC equals real steel consumption +/– stock movements. Production and trade figures are based on a broad definition of steel industry products as compiled by the IISI, including ingots and semifinished products, tubes and tube fittings, single-strand wire, railway wheels, tires, and axles.

The IISI and the Organization for Economic Cooperation and Development (OECD) (H. Duisenberg 1985) proposed the following formula to estimate the SI curve:

$$SI = f - (a - bx)\, e^{-cx},$$

where $x =$ GDP/capita with $SI > 0$, $x > 0$, and $b > 0$

The formula above largely reproduces the theoretical SI curve.

Figures 1.5 to 1.7 show the SI patterns over the period 1960 to 2005. The results are in Table 1.17.

The long memory parameter is not significant. We conclude that there is no (additional) long memory. In other words, no important "other" explanatory variables are omitted in the above-specified models.

Atmospheric Concentrations

In this application, we test for the presence of a long memory parameter in a series of atmospheric CO_2 concentrations (*ppmco*2) derived from in situ air measurements at Mauna Loa Observatory, Hawaii. The period covers monthly data from March 1958 to January 2007 (Keeling et al. 2005).

Figure 1.8 shows the growth pattern of CO_2 concentrations over the period March 1958 to January 2007.

The Hodrick-Prescott filter was used to extract the trend. Further, the series *ppmco*2 is stationarized by taking logs and applying a seasonal differencing filter to the latter. The series tested for long memory is labeled *lppm*1. The spectrum (Figure 1.9) suggests the presence of long memory.

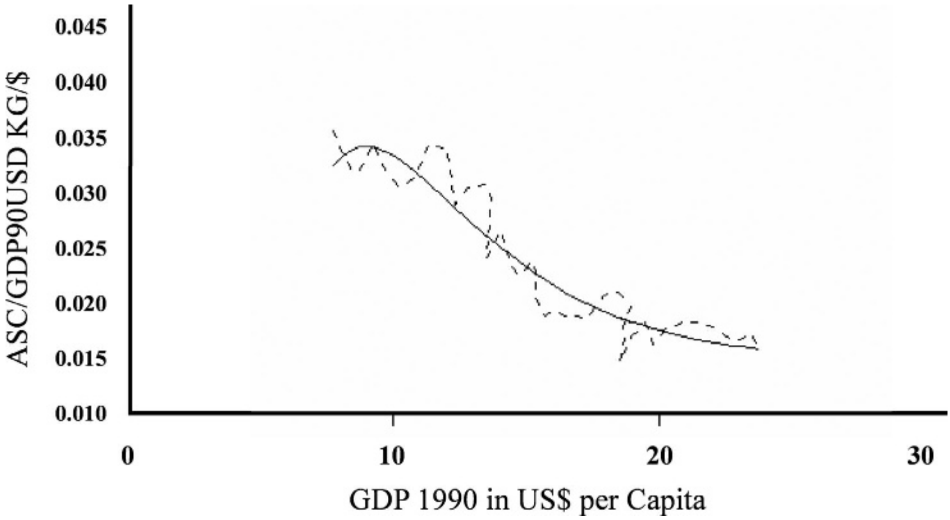

FIGURE 1.5 Steel Intensity Curve, EU15

We use our instrumental model with six lags to estimate the long memory parameter. The results are reported in Table 1.18.

From Table 1.2, we correct for bias and derive an estimated d of about 0.46. Testing for structural change in the first and second half of the sample revealed an unstable parameter. The Chow test for the Gauss-Newton regression is $F(6{,}563) = 2.58$ with significance level 0.017. Splitting the sample in two halves, we estimated with the instrumental model a d of 0.578 for the first half and a d of 0.641 for the second half. We test the model by running intrasample forecasts over the observations 451 to 588 on the basis of an estimate of d, using the first 450 observations. Figure 1.10 shows that the model forecasts well.

FIGURE 1.6 Steel Intensity Curve, Japan

FIGURE 1.7 Steel Intensity Curve, USA

TABLE 1.17 SI Curve

	f	a	b	c	d
EU15	0.011	3.423	0.434	0.435	0.129
(Significance)	(0.000)	(0.31)	(0.28)	(0.000)	(0.35)
Japan	0.010	37.880	3.24	0.438	0.116
(Significance)	(0.000)	(0.35)	(0.34)	(0.000)	(0.441)
USA	0.013	0.187	0.029	0.199	0.154
(Significance)	(0.000)	(0.247)	(0.153)	(0.000)	(0.314)

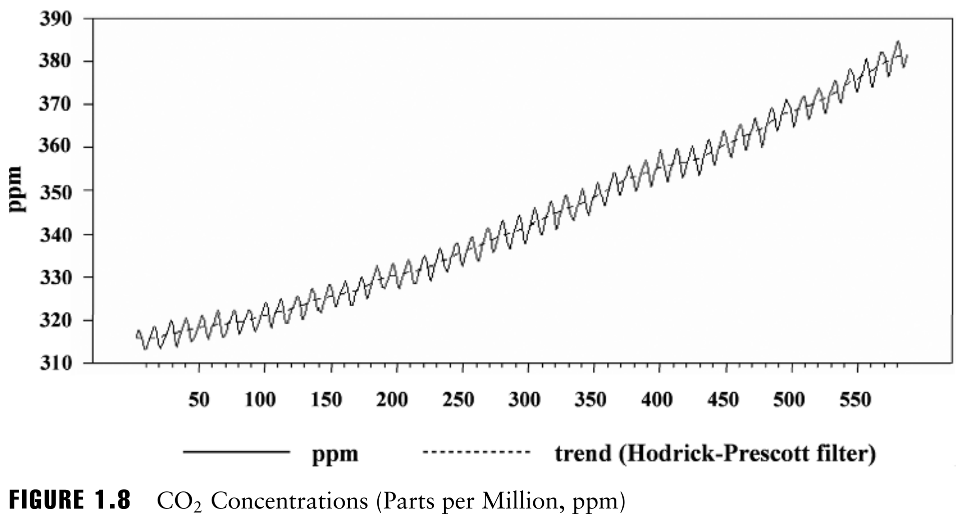

FIGURE 1.8 CO_2 Concentrations (Parts per Million, ppm)

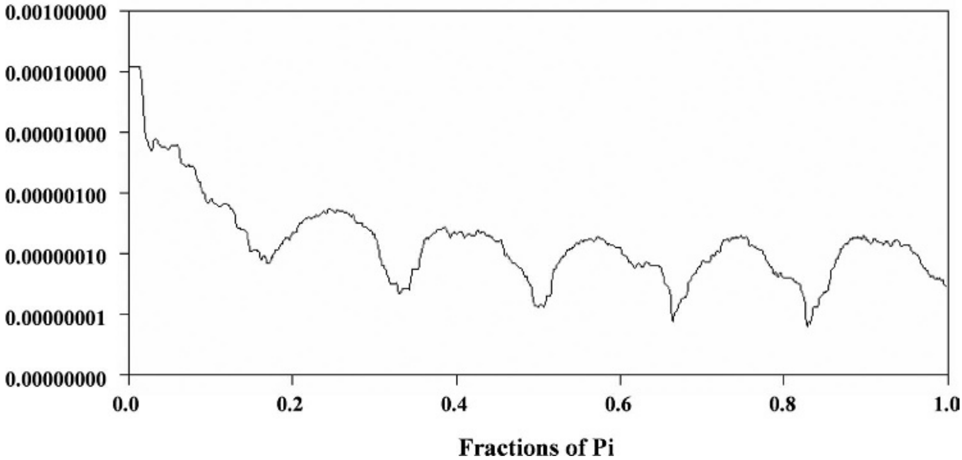

FIGURE 1.9 Long Memory Test, Spectrum lppm1

TABLE 1.18 CO_2 Concentrations, 1958(3)–2007(1)

	c	d
Lppm1	0.0005	0.626
	(0.00000000)	(0.00000000)
$\overline{R}^2 = 0.67$, *Log Likehood* $= 3071$		

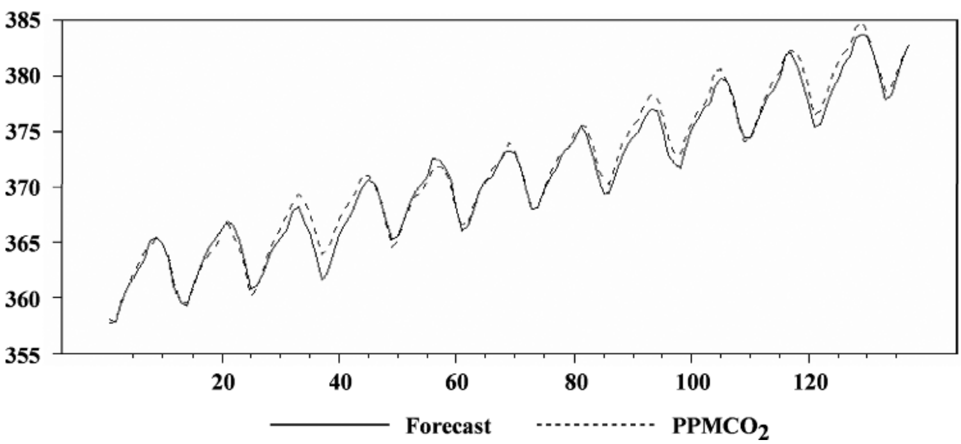

FIGURE 1.10 PPMCO2 Intrasample Forecast 451–588

CONCLUSIONS

Fractional integration is an important issue in modern time-series analysis. Traditional ARMA models, insofar as they are parsimonious, do not accurately describe situations where the long memory component of the impulse-response coefficients is predominant. Of course, long memory could be approximated arbitrarily well with a suitably large-order ARMA representation, but this is of little help in the case of small samples. Care must be taken to correctly stationarize the original time series. Long memory and structural change are easily confused. The concept of long memory leads in a natural way to the detection of stable relationships for stationary series, the so-called copersistence.

The problem with fractional integration, however, lies in the estimation techniques of the parameters. In order to simplify these techniques, we propose a truncated version of the fractionally integrated model that has the advantage of being easy to estimate and that captures parsimoniously the growth pattern of processes displaying impulse-response coefficient decaying at a much slower rate than those for stationary ARMA processes.

In this chapter we show that the number of autoregressive lags in this truncation can be chosen in the short range, from 2 to 6. We derived two results. First, under the assumption that the truncated model is the true model, the NLS estimator d^* of the parameter d of this model is consistent. This result is obtained under rather general assumptions. Specifically, we relax the *iid* assumption for e_t and replace it with the less restrictive α-mixing assumption. Second, using Monte Carlo experiments, we show that the fractional parameter (of the nontruncated model) can be consistently estimated by NLS and indirect inference on the basis of the simple truncated model. In our applications related to apparent steel consumption (annual data), we found no evidence for the presence of long memory. However, we found stochastic cycles and a significant impact of IP, particularly in the EU15, North America, and Japan, confirming economic theory. In the case of China, the exogenous variable IP has less explanatory power. Concerning the SI curves in the EU15, the United States, and Japan, there was no evidence for long memory, so that we conclude that there is no misspecification, in other words, no important explanatory variables are missing in the specification.

For a series of atmospheric CO_2 concentrations derived from in situ air measurements at Mauna Loa Observatory, Hawaii, we found a high persistence with a significant positive long memory parameter.

APPENDIX: PROOF OF CONSISTENCY OF THE ESTIMATOR D^*

We assume that model 1.3 is the true model and show that the NLS estimator d^* of the parameter d from model 1.3 is consistent.

First, we justify the choice of the NLS estimator. It follows from assumption OP (optimand) in Gallant and White (1988) that the methodology proposed hereafter in order to prove almost sure consistency allows us to consider the class of M estimators, which are defined as solutions to an optimization problem, such as NLS estimators, maximum likelihood (ML) estimators, and generalized method of moments (GMM) estimators. The unified theory of these estimators was developed originally in Hansen (1982). We use the NLS estimator for the following reasons.

The ML approach is primarily a large-sample approach (see Davidson and Mackinnon (1993), p. 247). The same argument holds for the GMM approach. As claimed by Bates (1990), the method of instrumental variables is inherently a large-sample estimation method based as it is on the law of large numbers and the central limit theorem. Of course, the GMM allows us to deal efficiently with heteroskedasticity if the latter is of a known form. This however is generally not the case. So we rely only on NLS. However, we propose to correct for heteroskedasticity by computing a consistent estimate of the covariance matrix as in White (1980). This correction does not affect the coefficients themselves, only their standard errors. Of course, if the form of heteroskedasticity is known, this latter approach will not be as efficient as weighted least squares.

Note that the robust errors approach is also a way to check the quality of the Monte Carlo simulation of ARFIMA processes. For example, we simulated an ARFIMA (0, 0.2, 0) with *iid* errors and estimated the fractional parameter d on the basis of the truncated model 1.3 with $r = 6$ lags. The estimated d was 0.206 with a standard error of 0.009 without the robust errors correction of the covariance matrix, while the standard error was 0.0091 if this correction is taken into account.

Let $\{y_t\}_{t=1}^{n}$ be a process generated by equation 1.3, and we desire an estimator of d. Consider $\hat{d}$ solution of

$$\hat{d} = \arg\min Q_n(d) = n^{-1} \sum_{t=1}^{n} \left[y_t + \sum_{j=1}^{r} \kappa_j(d) y_{t-j} \right]^2 \tag{A.1}$$

r is a constant that may be chosen in practice.

We specify the nonlinear autoregressive distributed lag model 1.3 in companion form. Let us define the p vectors $Y_t = [y_t, \ldots, y_{t-p+1}]'$, $V_t = [e_t, 0, \ldots, 0]'$, and the p^2 matrix $\boldsymbol{B}^*$ by

$$B^* = \begin{pmatrix} 0 & 0 & \vdots & 0 & 0 \\ 1 & 0 & \vdots & 0 & 0 \\ \ldots & \ldots & \ddots & 0 & \ldots \\ 0 & 0 & \vdots & 1 & 0 \end{pmatrix} \tag{A.2}$$

Define also the p vector $F\,(Y_{t-1}, d^*) = [\,f(y_{t-1}, \ldots, y_{t-p}, d^*), 0, \ldots, 0]'$. Thus, (3) can be rewritten as

$$Y_t = B^* Y_{t-1} + F\,(Y_{t-1}, d^*) + V_t \tag{A.3}$$

We will now prove the almost sure (a.s.) consistency of the nonlinear least squares estimator d^*. To do this, we apply theorem 3.5 from Mira and Escribano (2000) by checking that the assumptions (*MD, MX, CT, LR,* and *LN*) they used to derive the consistency result are satisfied in the case of model 1.3. Their approach is based on Gallant and White (1988), the seminal paper on estimation of and inference for nonlinear dynamic models, with the main advantage that they are able to write explicit assumptions related to a nonlinear model, such as moment conditions and conditions on the nonlinear function. They show (lemma 3.4) that assumptions *MD* to *LN* imply *near epoch dependence*, *r-integrability uniformly in t*, *s-domination* and the *Lipschitz-L_1 condition a.s.*, and thus consistency.

These assumptions are called *Lipschitz-type* assumptions. Consistency can be proved also on the basis of an *equicontinuity* assumption of the underlying functions (see Pötscher and Prucha 1991). The latter paper also provides a set of modules that can readily be used to prove consistency of a variety of M-estimators. Mira's and Escribano's (2000) assumptions are:

Assumption MD: Model 1.3 is the true model in the sense that

$$E\left(y_t | y_{t-1}, \ldots, y_{t-p}\right) \equiv f\left(y_{t-1}, \ldots, y_{t-p}, d^*\right) \tag{A.4}$$

Assumption MX (mixing): The sequence $\{V_t\}$ is strong mixing with $\{a_n\}$ of size $-v/(v-2)$ with $v > 2$.

By this assumption, we allow for some heterogeneity (some nonstationarity).

Assumption CT:

(i) For some fixed value $\varepsilon > 0$ and for all matrices $\boldsymbol{B}\nabla\boldsymbol{F}$ given by $\boldsymbol{B}\nabla\boldsymbol{F} \equiv \boldsymbol{B} + \nabla_y \boldsymbol{F}(\boldsymbol{Y}, d)$, with $\theta \in \Theta$, we have that $\rho\left(\boldsymbol{B}\nabla\boldsymbol{F}\right) < 1 - \varepsilon < 1$ *where* $\rho\left(\boldsymbol{B}\nabla\boldsymbol{F}\right)$ is the spectral radius of $\boldsymbol{B}\nabla\boldsymbol{F}$, i.e., the largest eigenvalue of the matrix $\boldsymbol{B}\nabla\boldsymbol{F}$.

Notice that for each specific matrix $\boldsymbol{B}\nabla\boldsymbol{F}$, its associated norm $\|.\|_S$ will verify that

$$\|\boldsymbol{B}\nabla\boldsymbol{F}\|_S \equiv \delta_{BY} < 1 - \varepsilon \tag{A.5}$$

$$\|.\|_s \equiv \left(E\left(\|.\|_s^r\right)\right)^{1/r} \equiv E^{1/r}\left(\|.\|_s^r\right) \tag{A.6}$$

(ii) For the norms $\|.\|_S$ and $\|.\|_2$ we have $\|B\| \leq \delta_{CB}$.

(iii) The compact parametric space Θ is such that the Jordan decomposition of the matrix $\mathbf{B}\nabla\mathbf{F}$ given in part (i), $J = M^{-1}\left(\mathbf{B}\nabla\mathbf{F}\right) M$, verifies $\|M^{-1}\|_\infty < \Delta^{-1}$ and $\|M\|_\infty < \Delta$ for some fixed values Δ and Δ^{-1}.

Assumption CN: $f\,(y_{t-1}, \ldots, y_{t-p}, d)$ is continuously differentiable in each argument, and its second-order derivatives with respect to d are continuous functions.

Assumption LR: For $r = 6$ we have

$$\mathrm{E}\,\| V_t \|_S^r \leq \Delta_V^{(r)} \tag{A.7}$$

$$E\,\| V_t \|_S^r \, \| V_s \|_S^r \leq \Delta_{VV}^{(r)} \tag{A.8}$$

Assumption LN: For the norms$\|.\|_S$ and $\|.\|_2$:

The following inequality holds a.s.:

$$\| F\,(Y_t, d)\|_S \le \delta_{CF}\,(\| Y_t \|_S) \tag{A.9}$$

The following inequality holds a.s.:

$$\|\nabla_d F\,(Y_{t-1}, d)\|_s^2 \le \left\|\nabla_d f\left(y_{t-1,\dots,} y_{t-p}, d\right)\right\|_s^2 \le \delta_L\left(\| Y_{t-1}\|_S\right)^2 \tag{A.10}$$

We will now check these assumptions in the case of model 3.

Assumption MD: Assumption *MD* is satisfied because of the specification of model 1.3.

Assumption MX (mixing): α-mixing sequences are called *strong mixing*. The quantity $\alpha\,(m)$ measures how much dependence exists between events separated by at least m time periods. By definition, y_t is a stationary time series where all ARMA components have been removed. It is therefore reasonable to assume that assumption *MX* is satisfied for the sequence $\{V_t\}$ because e_t may effectively be interpreted as an innovation.

Assumption CT(i):

$$\begin{aligned} y_t &= d y_{t-1} + d(1-d)/2 y_{t-2} + d(1-d)(2-d)/6 y_{t-3} \\ &\quad + d(1-d)(2-d)(3-d)/24 y_{t-4} + d(1-d)(2-d)(3-d)(4-d)/120 y_{t-5} \\ &\quad + d(1-d)(2-d)(3-d)(4-d)(5-d)/720 y_{t-6} + e_t \\ &\equiv f(.) \end{aligned} \tag{A.11}$$

$$\begin{array}{lll} \partial f(.)\big/\partial y_{t-1} = d & & = \kappa_1 \\ \partial f(.)\big/\partial y_{t-2} = d\,(1-d)\big/2 & & = \kappa_2 \\ \qquad\quad \vdots & & \;\vdots \\ \partial f(.)\big/\partial y_{t-6} = d\,(1-d)\,(2-d)\,(3-d)\,(4-d)\,(5-d)\big/720 & & = \kappa_6 \end{array} \tag{A.12}$$

$$B^* = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \nabla F = \begin{pmatrix} \kappa_1 & \kappa_2 & \kappa_3 & \kappa_4 & \kappa_5 & \kappa_6 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{A.13}$$

$$\boldsymbol{B}\nabla F = \begin{pmatrix} \kappa_1 & \kappa_2 & \kappa_3 & \kappa_4 & \kappa_5 & \kappa_6 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \tag{A.14}$$

Figure A.1 shows the growth pattern of the spectral radius of the matrix $\boldsymbol{B}\nabla\boldsymbol{F} = \boldsymbol{B} + \nabla\boldsymbol{F}$ as a function of d.

As the spectral radius is less than 1 for $-0.5 < d < 0.5$, we conclude that assumption CT is satisfied.

Assumption LR are restrictions as moment conditions on $\boldsymbol{V}_t$.

Assumption LN(i):

$$|f(y_{t-1}, \ldots y_{t-6}, d)| = |dy_{t-1} + \ldots + d(1-d)(2-d)(3-d)(4-d)(5-d)/720y_{t-6}|$$
$$\leq |\mathrm{d}|\,|y_{t-1}| + \ldots + |d(1-d)(2-d)(3-d)(4-d)(5-d)/$$
$$720|\,|y_{t-6}| < \sum_{i=1}^{6} |y_{t-i}| \qquad \text{(A.15)}$$

The last inequality follows from $-0.5 < d < 0.5$.

Assumption LN(ii):

$$\left|\frac{\partial f(y_{t-1}, \ldots, y_{t-6}, d)}{\partial d}\right| = \left|y_{t-1} + \frac{(1-2d)}{2}y_{t-2} + \frac{3d^2 - 6d + 2}{6}y_{t-3}\cdots\right|$$
$$\leq |y_{t-1}| + \left|\frac{1-2d}{2}\right||y_{t-2}| + \cdots \qquad \text{(A.16)}$$
$$< \sum_{i=1}^{6} |y_{t-i}|$$

Again, the last inequality follows from $-0.5 < d < 0.5$. Thus, assumption LN is satisfied.

The next theorem proves the consistency of the NLS estimator d that minimizes equation A.1.

- **Theorem 1:** Under assumptions *MD, MX, CT, CN, LR* and *LN* and the identification condition stated below, the nonlinear least squares estimator for model 1.3 converges a.s. to the true value of the parameter.
- **Proof:** See Mira and Escribano (2000).

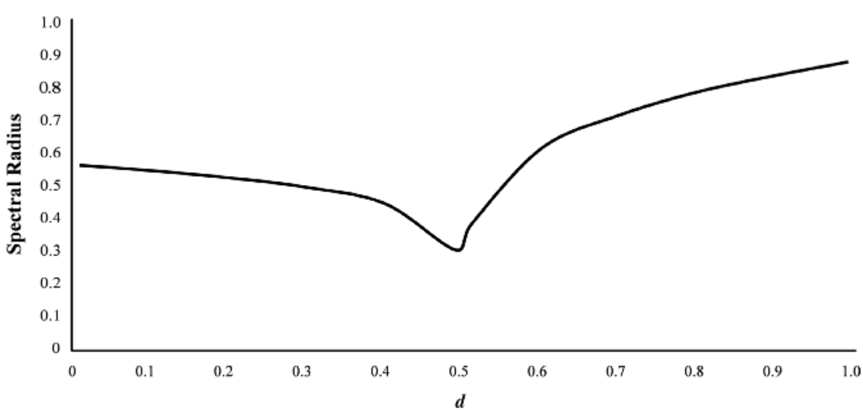

FIGURE A.1 Growth Pattern of Spectral Radius of Matrix B∇F = B + ∇F

In our case, the identification assumption is: Since the mean square error has a unique minimum at the conditional mean, and since model 1.3 is the conditional mean from assumption *MD*, the identification condition is that

$$F\,(Y_{t-1}, d^{*}) \neq F\,(Y_{t-1}, d) \;\text{ for } d^{*} \neq d \tag{A.17}$$

REFERENCES

Adams, F.G. (1996). *Modeling Commodity Markets: New Needs and New Methods*, The State of Art in Applied Econometrics, 50th Conference, Association d'Econométrie Appliquée.

Afrasiabi, A., M. Moallem, and W.C. Labys. (1991). *Spectral Interpretation of Stock Adjustment Processes in Mineral Markets. International Commodity Market Models*. London, UK: Chapman and Hall.

Andrews, D.W.K., and P. Guggenberger. (2003). "A Bias-Reduced Log-Periodogram Regression Estimator for the Long-Memory Parameter," *Econometrica* 71.2: 675–712.

Andrews, D.W.K., and Y. Sun. (2004). "Adaptive Local Polynomial Whittle Estimation of Long-Range Dependence," *Econometrica* 72.2: 569–614.

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

Bardet, J.M. (1999). "La mémoire longue en économie, discussion et commentaires," *Journal de la Société Française de Statistique* 140.2: 49–54.

Barkoulas, J., W.C. Labys, J. Onochie. (1997). "Fractional Dynamics in International Commodity Prices," *Journal of Futures Markets* 17: 161–189.

Bates, Ch.E. (1990). "Instrumental Variables."In J. Eatwell, M. Milgate, and P. Newman, eds., *Econometrics, The New Palgrave*, pp. 113–117. New York: Norton.

———. (1999). "Long Term Memory in Commodity Futures Prices," *Financial Review* 34: 91–100.

Bertrand, P. (1999). La mémoire longue en économie, discussion et commentaires, *Journal de la Société Française de Statistique* 140.2: 55–60.

Boutahar, M., V. Marimoutou, and L. Nouira (2005). *Estimation Methods of the Long Memory Parameter: Monte Carlo Analysis and Application*, GREQAM, Université de la Méditerranée, Working paper.

Brockwell, P.J., and R.A. Davis. (1998). *Time Series: Theory and Methods*. New York: Springer Verlag.

Cromwell, J.B., W.C. Labys, and E. Kouassi. (2000). "What Color Are Commodity Prices? A Fractal Analysis." *Empirical Economics* 25: 563–580.

Davidson, J., and T. Terasvirta, eds. (2002). *Long Memory and Nonlinear Time Series*, special issue of *Journal of Econometrics* 110. 2.

Davidson, R., and J. Mackinnon. (1993). *Estimation and Inference in Econometrics*. New York: Oxford University Press.

Diebold, F.X., and A. Inoue. (2001). "Long Memory and Regime Switching," *Journal of Econometrics* 105: 131–159.

Dolado, J.J., J. Gonzalo, and L. Mayoral. (2002). "A Fractional Dickey-Fuller Test for Unit Roots," *Econometrica* 70.5: 1963–2006.

Dridi, R., and E. Renault. (2000). "Semi-Parametric Indirect Inference," *STICERD/ Econometrics*, Discussion paper No. EM/00/392.

Duisenberg, H. (1985) "Long-Term Forecasting Models for Steel Demand," Communication, 12th International Conference of the Applied Econometric Association, Saragoza, Spain.

Fox, R., and M.S. Taqqu. (1986). "Large Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series," *Annals of Statistics* 14: 517–532.
Fuller, W.A. (1996). *Introduction to Statistical Time Series*. New York: Wiley-Interscience.
Gallant, A.R., and Tauchen, G. (1996). "Which Moments to Match?" *Econometric Theory* 12: 657–681.
Gallant, A.R., and H. White. (1988). *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models*. Oxford, UK: Basil Blackwell.
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. (1999). "La mémoire longue en économie, discussion et commentaires," *Journal de la Société Française de Statistique* 140.2: 61–64.
Gourieroux, C., and J. Jasiak. (1999). *Nonlinear Persistence and Copersistence,* Research paper, http://dept.econ.yorku.ca/~jasiakj.
Gourieroux, C., and A. Monfort. (1995a). *Series Temporelles et Modèles Dynamiques*. Paris, France: Economica.
———. (1997). *Simulation-Based Econometric Methods*. Oxford, UK: Oxford University Press.
Gourieroux, C., A. Monfort, and E. Renault. (1993). "Indirect Inference," *Journal of Applied Econometrics* 8: S85–S118.
Granger, C.W.J., and R. Joyeux. (1980). "An Introduction to Long-Memory Models and Fractional Differencing," *Journal of Time Series Analysis* 1: 15–39.
Hamilton, J. D. (1994). *Time Series Analysis*. Princeton, NJ: Princeton University Press.
Hansen, B. E. (1992). "Testing for Parameter Instability in Linear Models," *Journal of Policy Modeling*, 14: 517–533
Hansen, L.P. (1982). "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica* 50: 1029–1054.
Higuchi, T. (1988). "Approach to an Irregular Time Series on the Basis of the Fractal Theory," *Physica D*, 31: 277–283.
Hosking, J.R.M. (1981). "Fractional Differencing," *Biometrika* 68: 165–176.
Hurst, H.E. (1951). "Long-Term Storage Capacity of Reservoirs," *Transactions of the American Society of Civil Engineers* 116: 770–799.
IISI. (1974). *Steel Intensity and GNP Structure*. Brussels, Belgium: International Iron and Steel Institute.
Jasiak, J. (1999). "La mémoire longue en économie, discussion et commentaires," *Journal de la Société Française de Statistique* 140.2: 65–70.
———. (1999). *Persistence in Intertrade Durations*. Research paper, http://dept.econ.yorku.ca/~jasiakj.
———. (2000). *Long Memory in Economics: A Comment*, Research paper, http://dept.econ.yorku. ca/~jasiakj/.
Keeling, C.D., T.P. Whorf, and the Carbon Dioxide Research Group. (2005). "Atmospheric CO_2 Concentrations (Ppmv) Derived from in situ Air Samples Collected at Mauna Loa Observatory, Hawaii." Scripps Institution of Oceanography (SIO) University of California La Jolla, California USA 92093-0444, May 2005 http://cdiac.ornl.gov/ftp/trends/co2/maunaloa.co2.
Lardic, S., and V. Mignon. (1999). "La mémoire longue en économie: Une revue de la littérature," *Journal de la Société Française de Statistique* 140.2: 5–48
———. (1999). "La mémoire longue en économie: Une revue de la littérature, réponse aux intervenants," *Journal de la Société Française de Statistique* 140.2: 103–108.
Li, W. K., and A. I. McLeod. (1986). "Fractional Time Series Modelling," *Biometrica* 73: 217–221.
Lo, A. (1991). "Long-Term Memory in Stock Market Prices," *Econometrica* 59: 1279–1313.

Lo, A., and C. Mackinlay. (1989). "The Size and Power of the Variance Ratio Test in Finite Samples: A Monte Carlo Investigation," *Journal of Econometrics* 40: 203–238.

Maddala, G.S., and I.M. Kim. (1998). *Unit Roots, Cointegration, and Structural Change.* Cambridge, UK: Cambridge University Press.

Mandelbroth, B.B. (1977). *Fractals: Form, Chance and Dimension*. San Francisco, CA: W.H. Freeman.

Mayoral, L. (2000). *A New Minimum Distance Estimation Procedure of ARFIMA Processes*, Working Paper Series No. 00-17, Universidad Carlos III de Madrid. Revised version (2002). Available upon request.

Mira, S., and A. Escribano. (2000). "Nonlinear Time Series Models: Consistency and Asymptotic Normality of NLS under New Conditions."In W. Barnett et al., eds., *Nonlinear Econometric Modeling in Time Series Analysis*: 119–164 New York: Cambridge University Press.

Pötscher, B.M., and I.R. Prucha. (1991). "Basic Structure of the Asymptotic Theory in Dynamic Nonlinear Econometric Models, Part I: Consistency and Approximation Concepts," *Econometric Reviews* 10.2: 125–216.

Prat, G. (1999). "La mémoire longue en économie, discussion et commentaires," *Journal de la Société Française de Statistique* 140.2: 71–77.

Renault, E. 1999, "La mémoire longue en économie, discussion et commentaires," *Journal de la Société Française de Statistique* 140.2: 79–89

Robinson, P.M. (1994). "Semiparametric Analysis of Long-Memory Time Series," *Annals of Statistics* 22: 515–539.

———. (1995). "Gaussian Semiparametric Estimation of Log Range Dependence," *Annals of Statistics* 23: 1630–1661.

———. (1995). "Log-Periodogram Regression of Time Series with Long Range Dependence," *Annals of Statistics* 23: 1048–1072.

Rosenblatt, M. (1978). "Dependence and Asymptotic Independence for Random Processes."In M. Rosenblatt, ed., *Studies in Probability Theory*. Washington, DC: Mathematical Association of America: 24–25

Sadler, A. (2003), *Demande Internationale de Produits Sidérurgiques, Croissance et Développement*, Ph.D. dissertation, Université de la Méditerranée. Marseille, France.

Smith, A.A. Jr. (1993). "Estimating Nonlinear Time-Series Models Using Simulated Vector Autoregressions," *Journal of Applied Econometrics* 8: S63–S84.

Sowell, F. (1990). "The Fractional Unit Root Distribution," *Econometrica* 58: 495–505.

———. (1992). "Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models," *Journal of Econometrics* 53: 165–188.

———. (1992). "Modeling Long Run Behavior with the Fractional ARIMA Model," *Journal of Monetary Economics* 29: 277–302.

Taqqu, M.S. (1999). "La mémoire longue en économie, discussion et commentaires," *Journal de la Société Française de Statistique* 140.2: 91–96.

Truong-Van, B. (1999). "La mémoire longue en économie, discussion et commentaires," *Journal de la Société Française de Statistique* 140.2: 97–102.

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

———. (1984). *Asymptotic Theory for Econometricians.* New York: Academic Press.

Whittle, P. (1951). *Hypothesis Testing in Time Series Analysis.* New York: Hafner Publishing Company.