Symposium in Honour of Dr. Walter C. Labys
West Virginia University
May 7 2007

Indirect inference and long memory: A new truncated-series estimation method

Armand Sadler

ARCELOR, Luxembourg

Jean-Baptiste Lesourd

GREQAM-CNRS, Université de la Méditerranée, Aix-Marseille, France

and

Vêlayoudom Marimoutou

GREQAM-CNRS, Université de la Méditerranée, Aix-Marseille, France[1]

Preliminary and incomplete draft: not for citation or attribution

[1] We are grateful to Russell Davidson, Frank Giarratani and Walter Labys for helpful discussions and comments. Computations were performed in RATS and we thank Tom Maycock for support. Responsibility for errors rests with the authors.

Abstract

This paper proposes a truncated version of ARFIMA models where the number of autoregressive lags is small. Two results are derived. First, we prove that the NLS estimator of the parameter d^a from a truncated ARFIMA model is, under the assumption that this is the true model, almost surely consistent. Second, we show by Monte Carlo experiments that the fractional parameter d of the untruncated specification can be consistently estimated by NLS on the basis of the truncated version of the model. Next, an identification and estimation procedure for the truncated long memory process is proposed. The practical applications are related to steel consumption.

Keywords: fractional integration, ARFIMA models, almost sure consistency, NLS, steel consumption.

Classification JEL: C1, C5, F1.

1. Introduction

Long-memory processes are an important and even fundamental advance in time series modelling. 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 modelling long memories in a time series. ARFIMA processes, however, are associated with hyperbolically decaying autocorrelations and 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:

$$\Phi(L)(1-L)^d y_t = \mu + \Theta(L)\varepsilon_t, \varepsilon_t \sim iid(0,\sigma^2) \qquad (1)$$

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' (see e.g. Fuller (1996)).

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

The methods for estimating d, the long-range dependence parameter, can be summarized in three classes: the heuristic methods (the Hurst (1951) method, the Lo (1989, 1991) method, the Higuchi (1988) method), the semi parametric methods (GPH 1983) method, the Robinson (1994,1995a, 1995b) estimation methods) and the maximum likelihood methods (the exact maximum likelihood method, the Whittle (1951) approximate maximum likelihood method). For a comparison of these various classes of estimators we refer to Boutahar et al. (2005).

The estimation of fractional integration exponents leads to significant problems in some cases. It is even impossible in the case of small samples, as with industrial data. Long-memory estimations often concern financial time series with usually large numbers of observations (5,000 observations and even much more are not uncommon). However, smaller samples such as samples with 50 to 100 observations as an order of magnitude are usually encountered in the case of industrial forecasting problems and in such cases, the need for a consistent and precise estimation technique would be 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 do we actually care about long memory in those situations? For instance, one could argue that from a forecasting perspective long memory only starts to make a difference when forecasting over long horizons, but that in situations when you only have a few observations available, you wouldn’t be forecasting too many steps ahead anyway. Our answer is threefold: Firstly, 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 five thousand real time observations over a short period of time on some financial stock index. Secondly, a lot of industrial industry sectors are producing medium and long range forecasts based on a relatively small number of yearly or quarterly observations. A steel producer who is investing in a new rolling mill or a new Greenfield facility can not wait for a long time series in order to take a decision. His experts have to work with the actually available data. Thirdly, we have the feeling from our past studies that 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 burden, is the computational burden increasing? The answer is No. Indeed, we suggest to simply use our reference tables to correct for the bias.

At least to our knowledge, since the previous work of Li and McLeod (1986) such estimation techniques valid for small samples have not been proposed so far in the econometric literature. Moreover, Li and McLeod develop an estimation technique based on truncating the power series defining the process after about 50 terms.

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

This paper aims at defining an estimation technique of the fractional integration exponent *d* for comparatively small samples. It's asymptotical properties are based on a result established by Escribano and Mira (2000) about the almost sure consistency of an 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.

Our paper is organised as follows. The second section defines and discusses the theoretical background. More precisely, we show that the auxiliary parameter is a consistent estimator of the true underlying parameter for the truncated model, by proving that the hypotheses of Escribano and Mira apply to our truncated series. In the third section, we discuss a new algorithm which identifies and estimates the truncated long memory model. Section 4 is devoted to Monte-Carlo simulation and estimation of an ARFIMA (0,d,0) model for which we compute the true characteristics after 10,000 to 50,000 simulations for different values of d. These true parameters are compared to the estimations and the bias is corrected via the method of indirect inference. In section 5, the proposed methodology is illustrated by estimations of a bivariate model for apparent steel consumption in different parts of the world. A second illustration concerns a test of a specification for the so called steel intensity curve. Section 6 develops our conclusions.

2. Almost sure consistency of the NLS estimator from our trancated 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)):

$$(1-L)^d\, y_t = e_t\,. \tag{2}$$

with $e_t \sim iid$, or

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

where

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

where Γ is the gamma function and $d \in$ (-0.5, 0.5).

If ξ is not an integer: $\xi = n + \phi$, $0 < \phi < 1$: $\Gamma(\xi) = (\phi+n-1)\,(\phi+n-2)\,\ldots\,(\phi+1)\phi\,\Gamma(\phi)$.

We define the truncated version of this model by:

$$y_t + \sum_{j=1}^{r} \kappa_j(d) y_{t-j} = e_t. \qquad (3)$$

In addition, we relax the *iid* assumption for e_t and replace it by the less restrictive α-mixing assumption[2], thus allowing for some heteroscedasticity (see White, 1984, for details). We will show in appendix 1 that the parameter d of model (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.

3. 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 commun 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

[2] The *i.i.d.* assumption is the strongest assumption; *i.i.d.* implies mixing (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, so that mixing conditions are stronger than ergodicity. For details, see WHITE (1984) and ROSENBLATT (1978).

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)[3].

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

$$\Phi(L)\left[y_t + \sum_{j=1}^{r} \kappa_j(d) y_{t-j}\right] = \Theta(L)\varepsilon_t \tag{5}$$

or

$$\Phi(L)\nabla^d y_t = \Theta(L)\varepsilon_t \tag{6}$$

Define

$$u_t = y_t + \sum_{j=1}^{r} \kappa_j(d) y_{t-j} \tag{7}$$

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

$$\Phi(L)[u_t] = \Theta(L)\varepsilon_t \tag{8}$$

Let

$$x_t = \{\Theta(L)\}^{-1}\ \Phi(L)y_t \tag{9}$$

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

$$\nabla^d x_t = \{\Theta(L)\}^{-1}\Phi(\mathrm{L})\nabla^{\mathrm{d}}\mathrm{y}_{\mathrm{t}} = \varepsilon_{\mathrm{t}}. \tag{10}$$

d is estimated in 4 steps :

1. Start the algorithm by setting d = 0 in (5) and estimating the ARMA parameters by the Gauss-Newton algorithm.

---

3 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 $d < |0.5|$.

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 (7) with the d estimated in step 2.

4. Reestimate the ARMA parameters from (8) 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.

<u>Informal proof of the convergence of the above estimation procedure</u>

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 is formula (5) 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.

4. Monte Carlo simulation and indirect inference estimation of an Arfima (0,d,0)

In this section, we use a methodology called 'indirect inference' in order to demonstrate the usefulness of our approach. This methodology, introduced in the literature by Smith (1993), Gouriéroux, Monfort and Renault (1993) and Gallant and Tauchen (1996) is nowadays largely used in applied econometric research. The idea is to draw 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 (1996) for a detailed description of the methodology.

The initial model (*M*) is (2) and the approximated one (M^a) is (3).

The estimator is obtained by minimising (4) by nonlinear least squares.

Equation (4) 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} …. 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 towards the true parameter *d* because $f(\underline{y_t}, d) \neq f^a(\underline{y_t}, d)$.

First we will show hereafter 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 (3) with 6 lags and with large samples. d^a reported in row 2 in Tables 1-5 is the arithmetic mean of 10000 (respectively 50000) parameters estimated from each simulation.[4]

Table 1 *n* = number of observations = 1000 10000 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	-0.188	-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 2 *n* = number of observations = 500 50000 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 3 *n* = number of observations = 100 10000 simulations

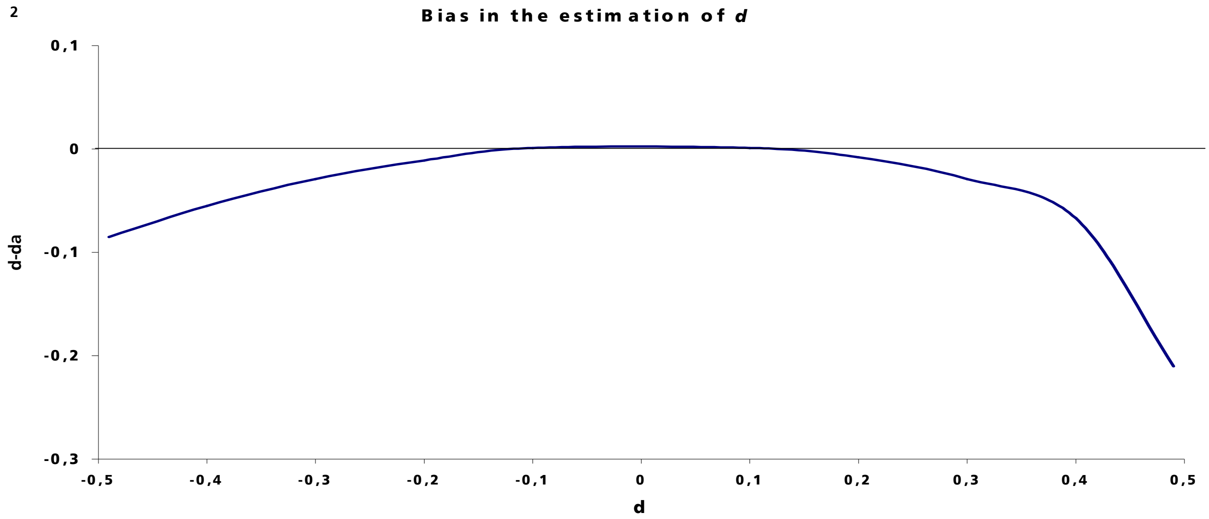
True *d*	-0.49	-0.4	-0.3	-0.2	-0.1	0.1	0.2	0.3	0.4	0.49
d^a	-0.41	-0.35	-0.28	-0.20	-0.11	0.08	0.19	0.31	0.44	0.69
Std error	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.12	0.10
Stat T	0.7	0.4	0.2	0.0	-0.1	-0.2	-0.1	0.1	0.3	2.0

[4] We observe an important bias when approaching the non-stationary case for positive *d*. The gain in precision is significant for increasing n.

Table 4	n = number of observations = 50								10000 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

Figure 2 below visualises the bias as a function of d :

Fig 2



We test now the sensitivity of the estimation as a function of the lag r.

Table 5 shows the results for $r = 3$.

Table 5	n = number of observations = 100									10000 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	-0.379	-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

It follows from Table 1 to 5 that the estimated fractional parameter d^a does not differ significatively 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 corrected. This is done by minimising the following distance (see Gouriéroux and Monfort (1996), Gouriéroux, Monfort and Renault (1993) and Smith (1990)):

$$d_{aSn}\ (\Omega) = \underset{d \in (-0.5,\ 0.5)}{\arg\min}\ [d_{an} - d_{aSn}(d)]'\ \Omega\ [d_{an} - d_{aSn}(d)] \tag{11}$$

with

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

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

In short, d_{an} is the estimator of d obtained by maximising the instrumental criterion in the case of a given sample of interest whereas $d_{aSn}(d)$ is the arithmetic mean of the maximisation of the same instrumental criterion for the S simulated samples. So, $d_{aSn}(d)$ is what we report in the tables 1 to 5. In order to maximise the asymptotic covariance matrix of an M-estimator like $d_{aSn}(\Omega)$, Gouriéroux and al. have shown that the optimal choice of Ω is:

$$\Omega^* = J_0\ (I_0 - K_0)^{-1}\ J_0.[5] \tag{12}$$

Gouriéroux and al. (1993, page S98) 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, (11) is simplified to

$$d_{aSn}(\Omega) = \arg\min\ [d_{an} - d_{aSn}(d)]^2 \tag{13}$$

[5] We use in (12) the notation proposed in Dridi and Renault (2000). We do not reproduce the formulas in (12) because they are not used in this paper. We only remark that in the absence of additional exogenous variables, like in our case here, $K_0 = 0$ (see Gouriéroux and al., 1996, page 83).

$d \in$ (-0.5, 0.5)

or simply to

$$d_{aSn}(\Omega) = \arg\min_{d \in (-0.5,\, 0.5)} \left| [d_{an} - d_{aSn}(d)] \right| \tag{14}$$

We will show now 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 6 and 7.

<u>*Estimation by II*</u>

Table 6 *n* = number of observations = 500 1000 replications

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

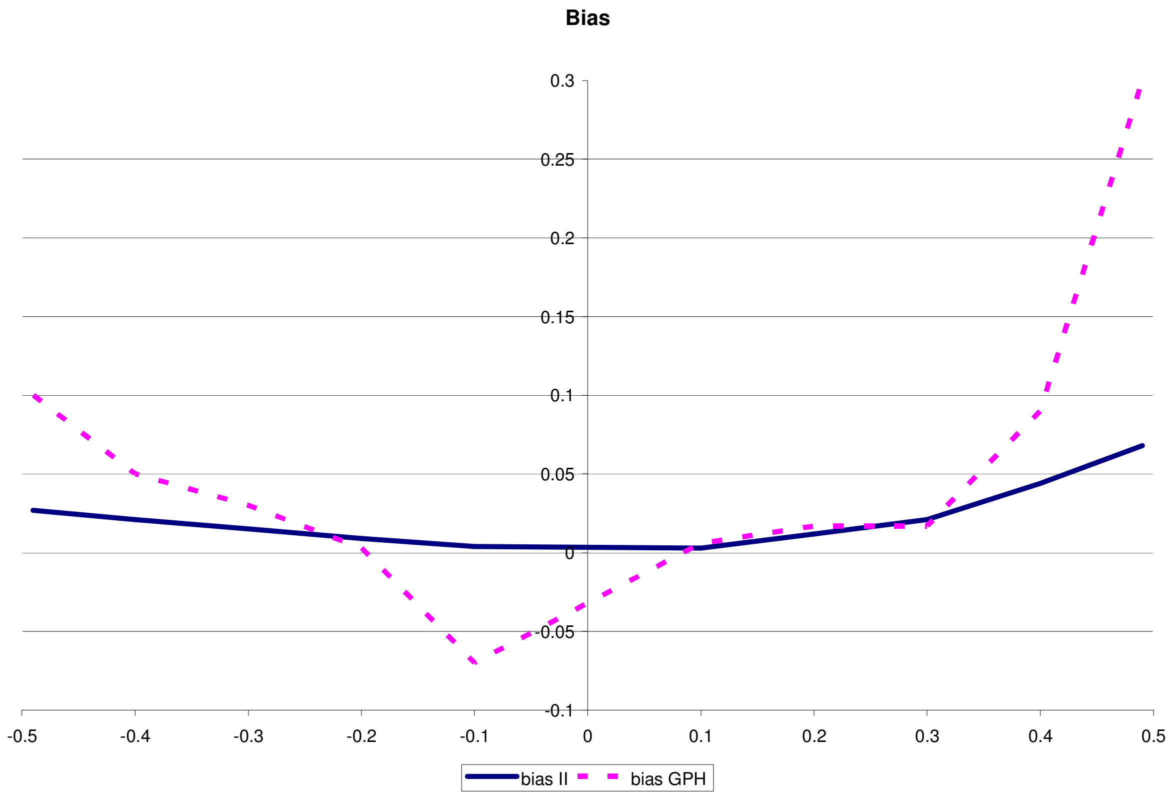
<u>*Estimation by GPH*</u>

Table 7 *n* = number of observations = 500 1000 replications

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

Figure 3 below shows the resulting bias for the two methods.

Fig 3



It follows from these simulations that the II method generates better estimates than the method suggested by GPH. The standard error and the bias are much smaller for the II method.

5. Applications: Growth patterns of apparent steel consumption

5.1 Apparent steel consumption and industrial production

In this section, we apply the estimation procedure described in section 3 to the analysis of apparent steel consumption (ASC) in the EU15, North America, Japan and China. 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:

$x_i = h_i (p_1,p_2,\ldots.p_i,\ldots.p_n,q)$

where x_i = demand of commodity i

q = the production of the firm

and p_i is the price of commodity i.

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

$D_t = D (P_t , PS_t , Y_t)$

where D = demand

P = price of commodity i

PS = price of competing commodities

Y = production of the sectors consuming commodity i.

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

$\partial D / \partial P < 0$

$\partial D / \partial PS > 0$

and $\partial D / \partial Y > 0$.

For this exercise, prices have been removed from the equations, mainly because of a common stochastic trend in the aluminium price and steel price series.

Table 1 summarizes the descriptive statistics of the endogenous and exogenous series.

Table 1 Descriptive statistics

Series	Obs	Mean	Std Error	Minimum	Maximum	Stationarity test $H_0 = I(1)$
LEU15	30	4.76	0.12	4.57	4.97	accept
LAMERNOR	30	4.73	0.14	4.42	4.98	accept
LJAPAN	30	4.27	0.14	3.94	4.53	accept
LCHINE	30	4.07	0.68	2.89	5.44	accept
IPEU15	30	1.58	2.77	-6.3	6.9	reject
IPAMNOR	30	2.47	3.93	-8.7	9.1	reject
IPJAPAN	30	1.85	5.26	-11.0	11.1	reject
IPCHINE	30	11.70	4.70	0.7	21.6	reject

Figure 4 and 5 show the growth patterns of these series over the period 1974-2003.

Figure 4

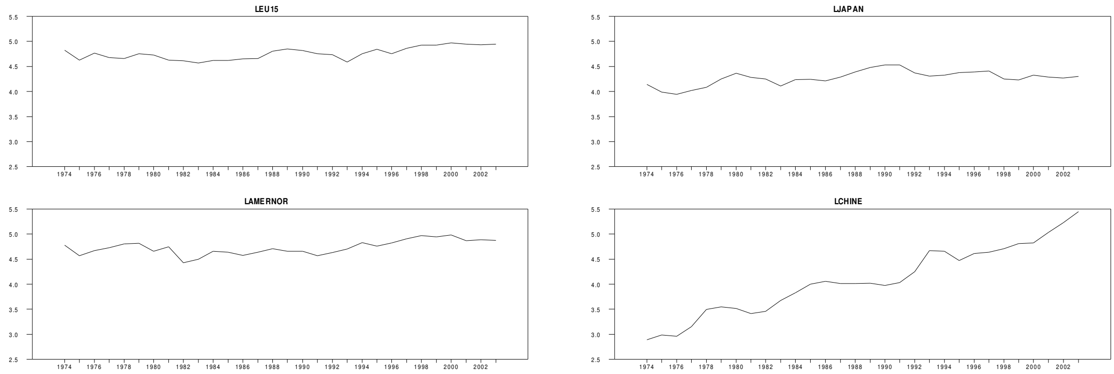
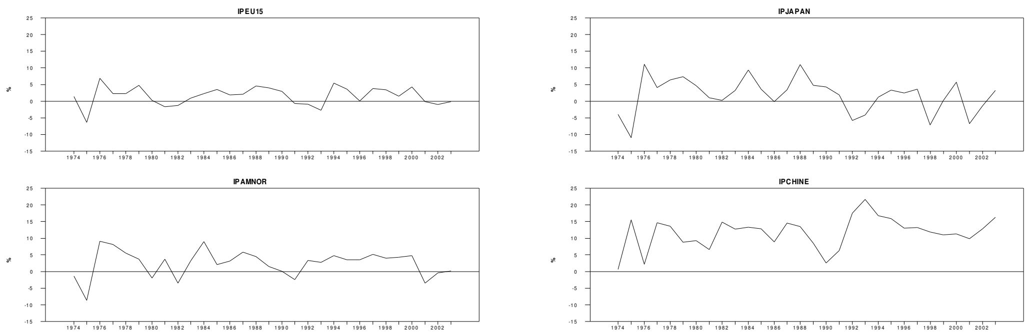


Figure 5



The most important stylized facts of these series are the following:

Firstly, the ASC series (in logs) are I (1) whereas the growth rates of IP are I (0). Secondly, 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 behaviour on inventories held by the steel consumers (merchants, steel service centres and final consumers like 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 modelling strategy proposed in the previous sections allowing taking into account exogenous explanatory variables, short memory ARMA components as well as a long memory parameter.

Tables 2 to 5 show the following results by running the CCM by country or region:

Table 2

```
------------------------------------------------------------------------
Dependent Variable DLEU15
Annual Data From 1970:01 To 2003:01
R Bar **2   0.653152
Q(8-0)                          7.427144
Significance Level of Q         0.49132416

   Variable                     Coeff          Signif
***********************************************
1.  CONSTANT              -0.040319215     0.00043955
2.  N_IPEU15{0}            0.023458940     0.00000000
1.  D                     -0.013782621     0.92928379
------------------------------------------------------------------------
```

Table 3

```
------------------------------------------------------------------------
Dependent Variable DLAMERNOR
Annual Data From 1979:01 To 2003:01
R Bar **2   0.786817
Q(6-0)                          4.028105
Significance Level of Q         0.67287283
```

Variable	Coeff	Signif
1. CONSTANT	-0.040908964	0.00523317
2. N_IPAMNOR{0}	0.021708878	0.00000201
3. N_DU{0}	-0.199495947	0.00087038
4. D	-0.092106490	0.65207685

DU = dummy 1982

Table 4

Dependent Variable DLJAPAN
Annual Data From 1979:01 To 2003:01
R Bar **2 0.639021
Q(6-3) 2.076871
Significance Level of Q 0.55660609

Variable	Coeff	Signif
1. CONSTANT	-0.017521214	0.28201632
2. AR{1}	-0.381654565	0.28202371
3. MA{1}	0.546609450	0.08259418
4. MA{2}	0.596710346	0.00573050
5. N_IPJAPAN1{0}	0.016339220	0.00000003
6. D	0.029015419	0.86292841

Table 5

Dependent Variable DLCHINE
Annual Data From 1981:01 To 2003:01
R Bar **2 0.495960
Q(5-2) 6.300162
Significance Level of Q 0.09788570

Variable	Coeff	Signif
1. CONSTANT	0.075951866	0.38604062
2. AR{1}	0.911105730	0.00150374
3. AR{2}	-0.430654521	0.07080872
4. N_IPCHINE{0}	0.005657863	0.34880283
5. N_DUU{0}	-0.376526406	0.00075481
6. D	-0.142007362	0.44871532

DUU = dummy 1994 and 1995

It follows from the above tables 2 to 5 that the fractional parameter *d* is not significantly different from zero so that the formulas can be much simplified (see Tables 6 to 9). All the processes are characterized by the property of short memory. Additional specifications not reproduced in this paper show that price variables of steel and aluminium were not significant.

Tables 6 to 9 show the reestimated equations where the fractional parameter is dropped, the analysis of the residuals from these latter equations as well as a test of the stability of the parameter linked to industrial production.

Table 6

```
------------------------------------------------------------------------
Dependent Variable DLEU15
Annual Data From 1975:01 To 2003:01
R Bar **2   0.729972
Q(7-0)                        3.448876
Significance Level of Q          0.84061047

   Variable                           Coeff         Signif
*************************************************
1.  CONSTANT               -0.037212351   0.00053388
2.  N_IPEU15{0}              0.025975583   0.00000000

Statistics on Series RESIDS DLEU15
Sample Mean    -0.0000000000          Variance              0.001887
Standard Error    0.0434374278          SE of Sample Mean      0.008066
t-Statistic          -0.00000             Signif Level (Mean=0) 1.00000000
Skewness           -0.43798              Signif Level (Sk=0)   0.36128003
Kurtosis            -0.10416              Signif Level (Ku=0)   0.91950548
Jarque-Bera        0.94025              Signif Level (JB=0)   0.62492288
HANSEN STABILITY TEST  L = 0.28072
Dickey-Fuller Test with  0 Lags
T-test statistic =       -6.88022
----------------------------------------------------------------------------------------------------------
```

Table 7

```
--------------------------------------------------------------------
Dependent Variable DLAMERNOR
Annual Data From 1975:01 To 2003:01
R Bar **2   0.785746
Q(7-0)                        7.460554
Significance Level of Q          0.38255010
```

Variable	Coeff	Signif
1. CONSTANT	-0.039229279	0.00231140
2. N_IPAMNOR1{0}	0.019189962	0.00000003
3. N_DU{0}	-0.212277569	0.00040713

DU = dummy 1982

Statistics on Series RESIDS DLAMERNOR

Sample Mean	-0.0000000000	Variance	0.002240
Standard Error	0.0473295335	SE of Sample Mean	0.008789
t-Statistic	-0.00000	Signif Level (Mean=0)	1.00000000
Skewness	-0.49000	Signif Level (Sk=0)	0.30707886
Kurtosis	-0.53774	Signif Level (Ku=0)	0.60185341
Jarque-Bera	1.50989	Signif Level (JB=0)	0.47003625

HANSEN STABILITY TEST COEFFICIENT IPAMNOR L = 0.135

Dickey-Fuller Test with 0 Lags

T-test statistic = -5.54703

Table 8

Dependent Variable DLJAPAN

Annual Data From 1976:01 To 2003:01

R Bar **2 0.505237

Q(7-3) 8.048299

Significance Level of Q 0.08982487

Variable	Coeff	Signif
1. CONSTANT	-0.037952245	0.02902140
2. AR{1}	-0.574982710	0.04631024
3. MA{1}	0.640174596	0.02598579
4. MA{2}	0.540711365	0.00891143
5. N_IPJAPAN1{0}	0.015734453	0.00000028

Statistics on Series RESIDS DLJAPAN

Sample Mean	0.00753775862	Variance	0.002918
Standard Error	0.05401668432	SE of Sample Mean	0.010208
t-Statistic	0.73840	Signif Level (Mean=0)	0.46664014
Skewness	-0.85235	Signif Level (Sk=0)	0.08145010
Kurtosis	1.21041	Signif Level (Ku=0)	0.25070094
Jarque-Bera	5.09960	Signif Level (JB=0)	0.07809732

HANSEN STABILITY TEST ON COEFFICIENT IPJAPAN : L = 0.116

Dickey-Fuller Test with 0 Lags

T-test statistic = -6.12990

Table 9

Dependent Variable DLCHINE
Annual Data From 1977:01 To 2003:01
R Bar **2 0.484841
Q(6-2) 11.183267
Significance Level of Q 0.02457976

Variable	Coeff	Signif
1. CONSTANT	-0.098172972	0.19202645
2. AR{1}	0.349852355	0.14259152
3. AR{2}	-0.129410738	0.54939170
4. N_IPCHINE{0}	0.017304284	0.00487069
5. N_DUU{0}	-0.318959134	0.00185916

DUU = dummy 1994 and 1995

Statistics on Series RESIDS DLCHINA
Sample Mean -0.0000000000 Variance 0.007956
Standard Error 0.0891988252 SE of Sample Mean 0.017166
t-Statistic -0.00000 Signif Level (Mean=0) 1.00000000
Skewness 0.33515 Signif Level (Sk=0) 0.50200471
Kurtosis 0.19093 Signif Level (Ku=0) 0.85947721
Jarque-Ber 0.54648 Signif Level (JB=0) 0.76090867
Dickey-Fuller Test with 0 Lags
T-test statistic = -5.12488
HANSEN STABILITY TEST : L = 0.058

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

5.2 The 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 for this literature are found in SADLER (2003).

Concerning steel, the first SI-curve has been constructed in the late 1960s by the Committee on Economic Studies of the IISI. It relates the evolution of SI (the ratio of apparent steel consumption to GDP) to the level of economic development of a country as measured by GDP per capita (IISI, 1974).

There are four stages in the development of SI:

a very low level before economic take-off

a rapid rise

a levelling off stage

a decline

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[6] / 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 semi-finished products, tubes and tube fittings, single strand wire, railway wheels , tyres and axles.

The IISI and the OECD (H.Duisenberg 1985) proposed the following formula to estimate the SI Curve:

$$SI = f - (a - bx)\, e^{-cx} \quad \text{where } x = GDP / capita$$

With $SI > 0$

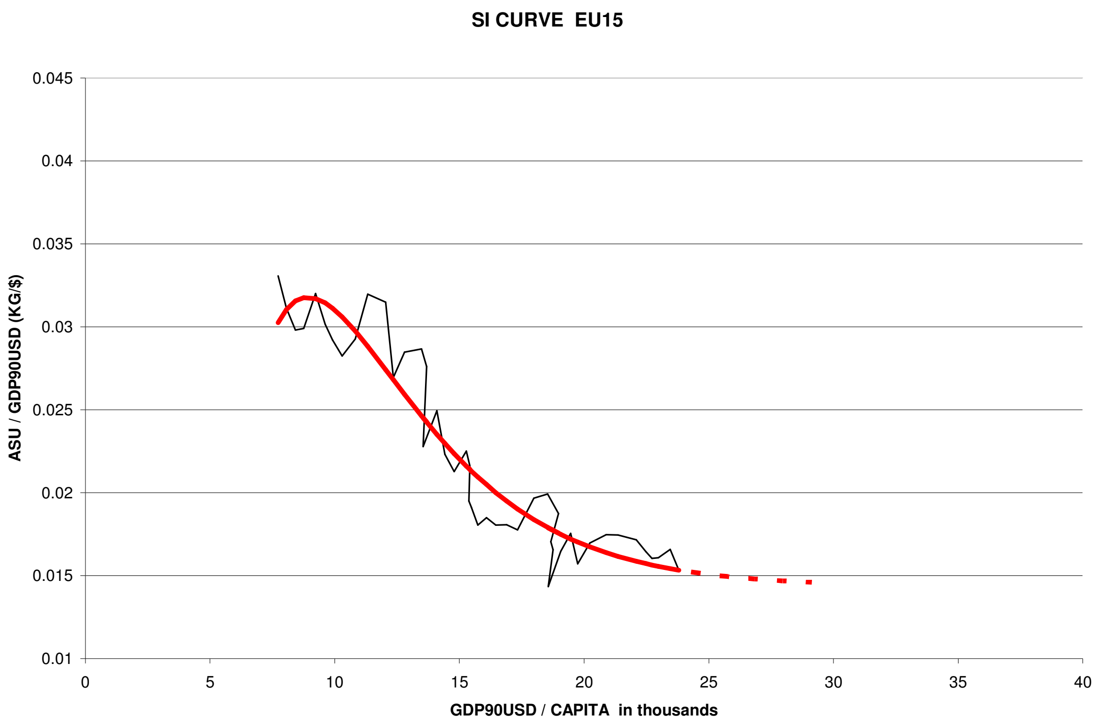
$x > 0$

$b > 0$

The formula above is largely reproducing the theoretical SI curve.

Figures 6 to 8 show the SI patterns over the period 1960-2005.

Figure 6



[6] Steel Weighted Industrial Production Index

Figure 7

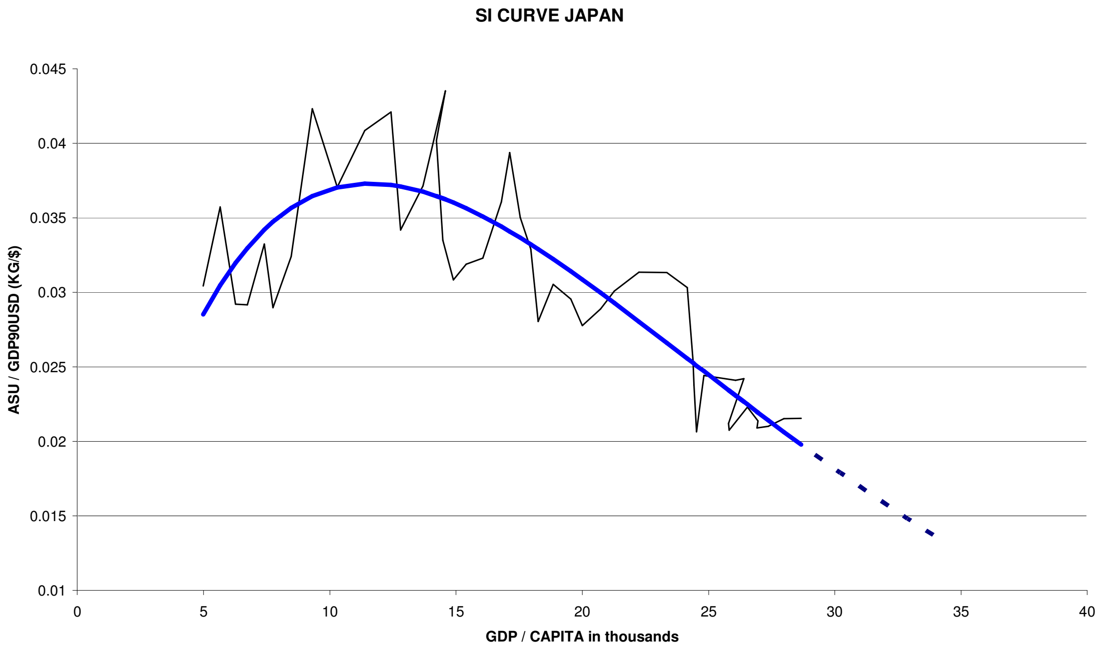
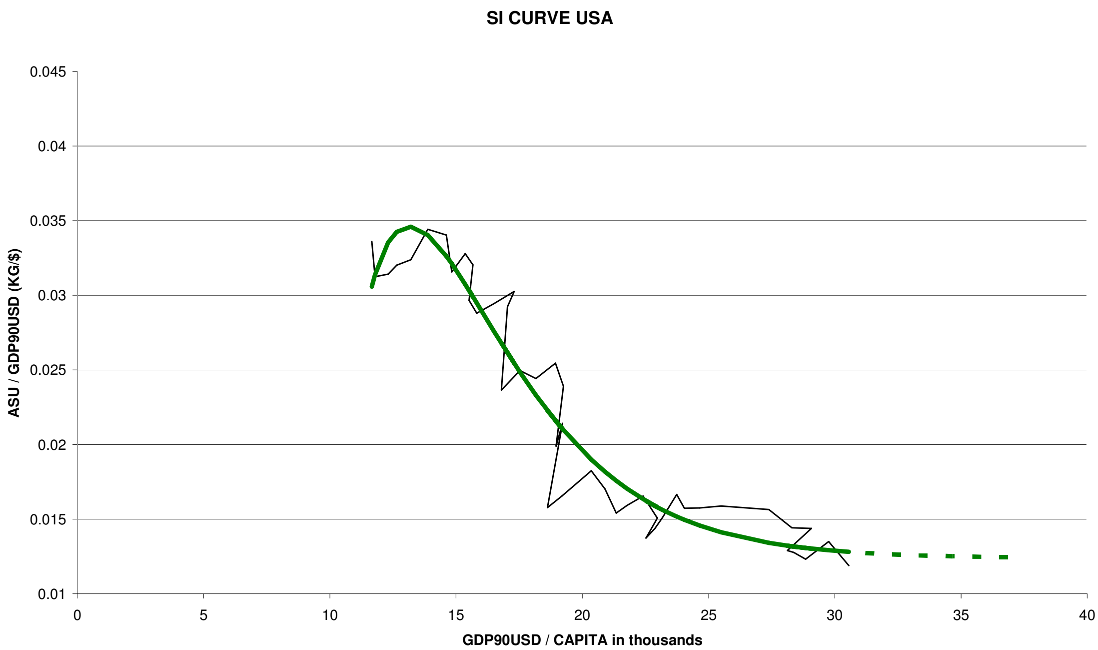


Figure 8



1. The SI-curve in the UE15

For the RATS model we refer to appendix 2

Result:

```
Nonlinear Least Squares - Estimation by Gauss-Newton
Convergence in    14 Iterations. Final criterion was  0.0000078 <  0.0000100
Dependent Variable W
Annual Data From 1964:01 To 2005:01
Usable Observations     42      Degrees of Freedom    37
Centered R**2     0.861972       R Bar **2   0.847050
Uncentered R**2   0.991368       T x R**2      41.637
Mean of Dependent Variable       0.0158776779
Std Error of Dependent Variable 0.0041507144
Standard Error of Estimate       0.0016232939
Sum of Squared Residuals         0.0000974981
Regression F(4,37)               57.7655
Significance Level of F          0.00000000
Log Likelihood                   212.84488
Durbin-Watson Statistic          1.413501

   Variable              Coeff          Std Error        T-Stat    Signif
*******************************************************************************
1.  F                 0.0116719820 0.0006515579     17.91396  0.00000000
2.  A                 3.4230303378 3.3845507195      1.01137  0.31840822
3.  B                 0.4342127737 0.3979143366      1.09122  0.28223056
4.  C                 0.4353835800 0.0669718043      6.50100  0.00000013
5.  D                 0.1296664710 0.1383799758      0.93703  0.35554959
-------------------------------------------------------------------------------
```

We conclude that there is no long memory, in other words no important ‘other’ explanatory variables are omitted in the above specified model.

---

2. The SI-curve in the USA

Specification : see appendix 2

```
Nonlinear Least Squares - Estimation by Gauss-Newton
Convergence in    2 Iterations. Final criterion was  0.0000000 <  0.0000100
Dependent Variable W
Annual Data From 1964:01 To 2005:01
Usable Observations     42      Degrees of Freedom    37
Centered R**2     0.895304       R Bar **2   0.883986
Uncentered R**2   0.988223       T x R**2      41.505
Mean of Dependent Variable       0.0158910710
Std Error of Dependent Variable 0.0057259273
```

Standard Error of Estimate 0.0019502982
Sum of Squared Residuals 0.0001407355
Regression F(4,37) 79.1014
Significance Level of F 0.00000000
Log Likelihood 205.13683
Durbin-Watson Statistic 1.484430

Variable	Coeff	Std Error	T-Stat	Signif
1. F	0.010305862	0.000661751	15.57363	0.00000000
2. A	37.880107896	40.798150766	0.92848	0.35917851
3. B	3.244379134	3.342260926	0.97071	0.33799408
4. C	0.438624470	0.057922629	7.57259	0.00000000
5. D	0.116079400	0.149102612	0.77852	0.44180921

3. The SI-curve for Japon

Specification : see appendix 2

Nonlinear Least Squares - Estimation by Gauss-Newton
Convergence in 5 Iterations. Final criterion was 0.0000071 < 0.0000100
Dependent Variable W
Annual Data From 1965:01 To 2005:01
Usable Observations 41 Degrees of Freedom 36
Centered R**2 0.688027 R Bar **2 0.653363
Uncentered R**2 0.981418 T x R**2 40.238
Mean of Dependent Variable 0.0212724566
Std Error of Dependent Variable 0.0054200299
Standard Error of Estimate 0.0031910906
Sum of Squared Residuals 0.0003665901
Regression F(4,36) 19.8486
Significance Level of F 0.00000001
Log Likelihood 180.13270
Durbin-Watson Statistic 1.417531

Variable	Coeff	Std Error	T-Stat	Signif
1. F	0.0131687747	0.0027608328	4.76986	0.00003029
2. A	0.1877324832	0.1596368420	1.17600	0.24731447
3. B	0.0295278968	0.0202698525	1.45674	0.15385582
4. C	0.1993511970	0.0508514625	3.92026	0.00038022
5. D	0.1547483653	0.1513128314	1.02270	0.31411712

The long memory parameter is not significant.

6. Conclusion

Fractional integration is an important issue in modern time series analysis. Traditional ARMA models, in so far as they are parsimonious, do not accurately describe those 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 in order to stationarize correctly 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 co-persistence.

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 easily estimable 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 paper we show that the number of autoregressive lags in this truncation can be chosen in the small range from 2 to 6 let's say.

We derived 2 results. Firstly, 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. More specifically, we relax the *iid* assumption for e_t and replace it by the less restrictive α-mixing assumption. Secondly, we showed by Monte Carlo experiments that the fractional parameter (of the untruncated 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 the 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 US 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.

Appendix 1

Proof of the consistency of the estimator *d**

We assume that model (3) is the true model and show that the NLS estimator d^* of the parameter *d* from model (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)[7] estimators. 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 C. E. 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 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, but 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.

[7] The unified theory of these estimators was developed originally in Hansen (1982).

Note that the robust errors approach is also a way to check the quality of the Monte Carlo simulation of ARFIMA processes.[8]

Let $\{y_t\}^n_{t=1}$ be a process generated by (3) and we desire an estimator of d. Consider $\hat{d}$ solution of[9]

$$\hat{d} = \arg\min Q_n(d) = n^{-1}\sum_{t=1}^{n}\left[y_t + \sum_{j=1}^{r}\kappa_j(d)y_{t-j} \right]^2 \tag{4}$$

We specify the nonlinear autoregressive distributed lag model (3) in companion form. Let us define the p vectors $\boldsymbol{Y}_t = [y_t, \ldots, y_{t-p+1}]'$, $\boldsymbol{V}_t = [e_t, 0, \ldots, 0]'$ and the p^2 matrix $\boldsymbol{B}^*$ by

$$B^* = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}.$$

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

$$\boldsymbol{Y}_t = \boldsymbol{B}^*\boldsymbol{Y}_{t-1} + \boldsymbol{F}\ (\boldsymbol{Y}_{t-1},\ d^*) + \boldsymbol{V}_t.$$

We will now prove the almost sure consistency of the nonlinear least squares estimator d^*.

To do this, we apply theorem 3.5 from S. Mira & A. Escribano (2000)[10] by checking that their assumptions (*MD,MX,CT,LR* and *LN*) used to derive the consistency result are satisfied in the case of model (3).

---

[8] We simulated for example an ARFIMA (0,0.2,0) with *i.i.d.* errors and estimated the fractional parameter d on the basis of the truncated model (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.

[9] r is a constant that may be chosen in practice ≤ 6.

Their assumptions are as follows:

Assumption MD : Model (3) is the true model in the sense that

$$E(y_t | y_{t-1},..., y_{t-p}) \equiv f(y_{t-1},..., y_{t-p}, d^*)$$

Assumption MX (mixing) : The sequence $\{V_t\}$ is strong mixing with $\{\alpha_n\}$ 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

$$\rho(\boldsymbol{B}\nabla\boldsymbol{F}) < 1 - \varepsilon < 1$$

where $\rho(\boldsymbol{B}\nabla\boldsymbol{F})$ 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 $\|\cdot\|_s$ will verify that

$\|\boldsymbol{B}\nabla\boldsymbol{F}\|_s \equiv \delta_{BY} < 1 - \varepsilon.$

---

[10] Their approach is based on Gallant and White, 1988, (the seminal paper on estimation and inference for nonlinear dynamic models) with the main advantage that they are able to write explicit assumptions related to a nonlinear model, e.g. model (1.3), 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 B.M.Pötscher and I.R.Prucha, 1991). The latter paper provides also a set of modules which can readily be used to prove consistency of a variety of M-estimators.

$$\|\cdot\|_s \equiv (\mathrm{E}(\|\cdot\|_s^r))^{1/r} \equiv \mathrm{E}^{1/r}(\|\cdot\|_s^r)$$

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

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

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

(i) $E\|V_t\|_S^r \leq \Delta_V^{(r)}$;

(ii) $E\|V_t\|_S^r \|V_s\|_S^r \leq \Delta_{VV}^{(r)}$;

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

(i) the following inequality holds a.s. :

$$\|F(Y_t, d)\|_s \leq \delta_{CF}(\|Y_t\|_s);$$

(ii) the following inequality holds a.s. :

$$\|\nabla_d F(Y_{t-1}, d)\|_s^2 \leq \|\nabla_d f(y_{t-1}, \ldots, y_{t-p}, d))\|_s^2 \leq \delta_L(\|Y_{t-1}\|_S)^2$$

We will now check the above assumptions in the case of model (3).

Assumption MD : Assumption *MD* is satisfied because of the specification of model (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 follows that it is 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) :

$$y_t = dy_{t-1} + d(1-d)/2y_{t-2} + d(1-d)(2-d)/6y_{t-3} + d(1-d)(2-d)(3-d)/24y_{t-4} + d(1-d)(2-d)(3-d)(4-d)/120y_{t-5} + d(1-d)(2-d)(3-d)(4-d)(5-d)/720y_{t-6} + e_t$$

$$\equiv f(\cdot)$$

$$\partial f(\cdot)/\partial y_{t-1} = d \qquad = \kappa_1$$

$$\partial f(\cdot)/\partial y_{t-2} = d\,(1\text{-}d)/2 \qquad = \kappa_2$$

.
.
.
.
.

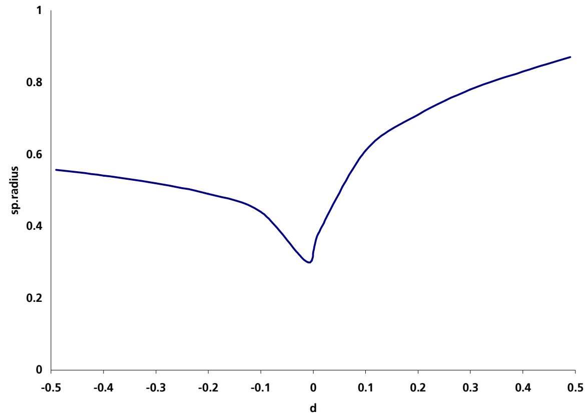
$$\partial f(\cdot)/\partial y_{t-6} = d\,(1\text{-}d)\,(2\text{-}d)\,(3\text{-}d)\,(4\text{-}d)\,(5\text{-}d)/720 \qquad = \kappa_6$$

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

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

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

Spectral radius of the matrix $B\nabla F$



As the spectral radius is < 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) :

$$\begin{aligned}\left|f(y_{t-1},...y_{t-6},d)\right| &= \left|dy_{t-1} + ... + d(1-d)(2-d)(3-d)(4-d)(5-d)/720\, y_{t-6}\right| \\ &\le \left|\mathrm{d}\right|\left|y_{t-1}\right| + ... + \left|d(1-d)(2-d)(3-d)(4-d)(5-d)/720\right|\left|y_{t-6}\right| \\ &< \sum_{i=1}^{6}\left|y_{t-i}\right|\end{aligned}$$

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

Assumption LN(ii) :

$$\begin{aligned}\left|\frac{\partial f(y_{t-1},...y_{t-6},d)}{\partial d}\right| &= \left|y_{t-1} + \frac{(1-2d)}{2}y_{t-2} + \frac{3d^2-6d+2}{6}y_{t-3}...\right| \\ &\le \left|y_{t-1}\right| + \left|\frac{1-2d}{2}\right|\left|y_{t-2}\right| + ... \\ &< \sum_{i=1}^{6}\left|y_{t-i}\right|\end{aligned}$$

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

The theorem below proves the consistency of the NLS estimator *d* that minimizes (4).

Theorem 1 : Under assumptions *MD, MX, CT, CN, LR* and *LN* and the identification condition stated below, the nonlinear least squares estimator for model (3) converges a.s. to the true value of the parameter.

Proof : See S. Mira and A. Escribano (2000)

In our case, the identification assumption is: Since the mean square error has a unique minimum at the conditional mean, and since model (3) is the conditional mean from assumption *MD*, the identification condition is that

$\boldsymbol{F}(\boldsymbol{Y}_{t-1}, d^*) \neq \boldsymbol{F}(\boldsymbol{Y}_{t-1}, d)$ for $d^* \neq d$.

Appendix 2

1. The SI in the UE15: the RATS model

```
NONLIN F A B C D
set x = 0.0
FRML H3 X = D*X{1}+D*(1-D)/2*X{2}+D*(1-D)*(2-D)/6*X{3}+D*(1-D)*(2-D)*(3-D)/24*X{4}

dec vec change(5)
compute maxchange = 1.0
clear resids nlresids
compute d = 0.0

FRML H SI = F-(A-B*E)*EXP(-C*E)
*FRML P W = F-(A-B*E)*EXP(-C*E)
INPUT F A B C
0.014 0.48 0.085 0.31 0.0
NLPAR(SUBITERATIONS=100)
NLLS(FRML=H) SI / RESIDS



compute lastd = 0.0, lastF = %beta(1),lasta = %beta(2), lastb = %beta(3), lastc = %beta(4)
dis change

until maxchange<0.001
{
 clear x resids nlresids
 set w = SI-( D*SI{1}+D*(1-D)/2*SI{2}+D*(1-D)*(2-D)/6*SI{3}+D*(1-D)*(2-D)*(3-D)/24*SI{4})

FRML P W = F-(A-B*E)*EXP(-C*E)
compute F=0.014, A=0.48, B=0.085, C=0.31
NLPAR(SUBITERATIONS=100)
NLLS(FRML=P) w / RESIDS
GROUP FMODEL4 P>>estimy4

compute F = %beta(1), A = %beta(2), B = %beta(3), C = %BETA(4)
 SET X = RESIDS
 NLLS(FRML=H3,pri) X / nlresids
 GROUP FMODEL3 H3>>ESTIMY3



compute change = ||d-lastd, F-lastF,A-lastA,B-lastB,C-LASTC||
 compute lastd=d,lastF=F,lastA=A,lastB=B, LASTC=C
 compute maxchange = %maxvalue(%abs(change))
 dis d change maxchange F A B C
 }
end until
```

Appendix 3

Definitions and sources

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

frequency = annual

period = 1974-2003

sources: Global*Insight*, European Commission, IISI

References

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

A. AFRASIABI, M. MOALLEM and W.C. LABYS, 1991, *Spectral interpretation of stock adjustment processes in mineral markets,* International Commodity Market Models, Chapman and Hall.

D. W. K. ANDREWS and P. GUGGENBERGER, 2003, *A Bias-Reduced Log-Periodogram Regression Estimator for the Long-Memory Parameter,* Econometrica, Vol. 71, No.2, 675-712.

D. W. K. ANDREWS and Y. SUN, 2004, *Adaptive Local Polynomial Whittle Estimation of Long-Range Dependance,* Econometrica, Vol. 72, No.2, 569-614.

R.T. BAILLIE, 1996, *Long Memory Processes and Fractional Integration in Econometrics,* Journal of Econometrics, 73, 5-59.

J.M. BARDET, 1999, *La mémoire longue en économie, discussion et commentaires*, Journal de la Société Française de Statistique, tome 140, n.2.

P. BERTRAND, 1999, *La mémoire longue en économie, discussion et commentaires*, Journal de la Société Française de Statistique, tome 140, n.2.

J. BARKOULAS, W.C. LABYS, J. ONOCHIE, 1997, *Fractional dynamics in international commodity prices,* Journal of Futures Markets, 17 :161-189.

J. BARKOULAS, W.C. LABYS, J. ONOCHIE, 2000, *Long term memory in commodity futures prices,* Financial Review.

M. BOUTAHAR, V. MRIMOUTOU and L. NOUIRA, 2005, *Estimation methods of the long memory parameter: Monte Carlo analysis and application,* GREQAM, Université de la Méditerranée, Working paper

PETER J. BROCKWELL and RICHARD A. DAVIS, 1998, *Time Series : Theory and Methods,* Springer.

J.B. CROMWELL, W.C. LABYS and E. KOUASSI, 2000, *What color are commodity prices? A fractal analysis*, Empirical Economics, Springer Verlag.

R. DAVIDSON and J. MACKINNON, 1993, *Estimation and Inference in Econometrics,* Oxford University Press.

J. DAVIDSON and T. TERASVIRTA, 2002, editors, *Long memory and nonlinear time series*, Journal of Econometrics, Vol. 110, No 2.

F.X. DIEBOLD and A. INOUE, 2001, *Long memory and regime switching*, Journal of Econometrics 105, 131-159.

J.J. DOLADO, J. GONZALO and L. MAYORAL, 2002, *A Fractional Dickey-Fuller Test for Unit Roots*, Econometrica, Vol. 70, No.5, 1963-2006.

RAMDAN DRIDI and ERIC RENAULT, May 2000, *SEMI-PARAMETRIC INDIRECT INFERENCE*, STICERD / Econometrics, Discussion paper No. EM/00/392.

R. FOX and M.S. TAQQU, 1986, *Large Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series,* Annals of Statistics, 14, 517-32.

W. A. FULLER, 1996, *Introduction to Statistical Time Series,* Wiley- Interscience.

A.RONALD GALLANT and HALBERT WHITE, 1988, *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models,* Basil Blackwell.

J.GEWEKE and S.PORTER-HUDAK, 1983, *The Estimation and Application of Long Memory Time Series Models,* Journal of Time Series Analysis, 4, 221-38.

C. GOURIEROUX, A. MONFORT and E. RENAULT, 1993, *Indirect Inference*, Journal of Applied Econometrics, Vol.8, S85-S118.

CHRISTIAN GOURIEROUX and ALAIN MONFORT, 1995, *Series Temporelles et Modèles Dynamiques,* Economica.

CHRISTIAN GOURIEROUX and ALAIN MONFORT, 1996, *Simulation-based Econometric Methods,* Oxford University Press.

CHRISTIAN GOURIEROUX and JOANN JASIAK, 1999, *Nonlinear Persistence and Copersistence,* Research Paper, http://dept.econ.yorku.ca/˜jasiakj.

CH. GOURIEROUX, 1999, *La mémoire longue en économie, discussion et commentaires*, Journal de la Société Française de Statistique, tome 140, n.2.

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

HL. GRAY, N.F. ZHANG and W.A. WOODWARD, 1989, *On generalized fractional processes*, Journal of Time Series Analysis 10: 233-257.

JAMES D. HAMILTON, 1994, *Time Series Analysis,* Princeton University Press.

BRUCE E. HANSEN, 1992, *Testing for Parameter Instability in Linear Models*, Journal of

Policy Modeling.

L.P. HANSEN, 1982, *Large Sample Properties of Generalized Method of Moments Estimators*, Econometrica, Vol.50.

J.R.M. HOSKING, 1981, *Fractional Differencing,* Biometrika, 68, 165-176.

T. HIGUCHI, 1988, *Approach to an irregular time series on the basis of the fractal theory,* Physica, D, 31, 277-283.

H.E. HURST, 1951, *Long-Term Storage Capacity of Reservoirs,* Transactions of the American Society of Civil Engineers, 116, 770-799.

G.J. JANACEK, 1982, *Determining the Degree for Time Series via the Log Spectrum*, Journal of Time Series Analysis, Vol. 3, pp. 177-183.

J. JASIAK, 1999, *La mémoire longue en économie, discussion et commentaires*, Journal de la Société Française de Statistique, tome 140, n.2.

JOANN JASIAK, 2000, *Long Memory in Economics : A Comment,* Research Paper, http://dept.econ.yorku.ca/˜jasiakj.

JOANN JASIAK, 1999, *Persistence in Intertrade Durations,* Reseach Paper, http://dept.econ.yorku.ca/˜jasiakj.

S. LARDIC et V. MIGNON, 1999, *La Mémoire longue en économie: une revue de la littérature*, Journal de la Société Française de Statistique, tome 140, no 2.

S. LARDIC et V. MIGNON, 1999, *La Mémoire longue en économie: une revue de la littérature*, *Réponse aux intervenants*, Journal de la Société Française de Statistique, tome 140, no 2.

W. K. LI and A. I. McLEOD, 1986, 73, *Fractional time series modelling*, Biometrica.

A. LO 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.

A. LO, 1991, *Long-Term Memory In Stock Market Prices,* Econometrica, 59, 1279-1313.

G.S.MADDALA and IN-MOO KIM, 1998, *Unit Roots, Cointegration, and Structural Change,* Cambridge.

B.B.MANDELBROT, 1977, *Fractals: Form, Chance and Dimension,* San Francisco: Freeman.

L. MAYORAL, 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.

S. MIRA & A. ESCRIBANO, 2000, *Nonlinear time series models : Consistency and asymptotic normality of NLS under new conditions,* Nonlinear Econometric modeling in time series analysis, Edited by W.Barnett et al., Cambridge.

P.PERRON, 1989, *The great crash, the oil price shock, and the unit root hypothesis,* Econometrica, Vol. 57, no 6, 1361 – 1401.

B.M.PÖTSCHER 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.

G. PRAT, 1999, *La mémoire longue en économie, discussion et commentaires*, Journal de la Société Française de Statistique, tome 140, n.2.

E. RENAULT, 1999, *La mémoire longue en économie, discussion et commentaires*, Journal de la Société Française de Statistique, tome 140, n.2.

P.M. ROBINSON (1994), *Semiparametric analysis of long-memory time series,* The Annals of Statistics, 22, 515-539.

P.M. ROBINSON (1995a), *Log-Periodogram Regression of Time Series With Long Range Dependence,* The Annals of Statistics, 23, 1048-1072.

P.M. ROBINSON (1995b), *Gaussian Semiparametric Estimation of Log Range Dependence,* The Annals of Statistics, 23, 1630-1661.

F.SOWELL, 1992a, *Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models,* Journal of Econometrics, 53, 165-188.

F.SOWELL, 1992b, *Modeling long run behavior with the fractional ARIMA model,* Journal of Monetary Economics 29, 277-302.

M.S. TAQQU, 1999, *La mémoire longue en économie, discussion et commentaires*, Journal de la Société Française de Statistique, tome 140, n.2.

B. TRUONG-VAN, 1999, *La mémoire longue en économie, discussion et commentaires*, Journal de la Société Française de Statistique, tome 140, n.2.

P. WHITTLE, 1951, *Hypothesis testing in time series analysis,* Hafner, New York

H. WHITE, 1984, *Asymptotic Theory for Econometricians*, Academic Press.