

# Conseqüències econòmètriques del grau de memòria en dades temporals

Ernest Pons Fanals

**ADVERTIMENT.** La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX ([www.tesisenxarxa.net](http://www.tesisenxarxa.net)) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

**ADVERTENCIA.** La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR ([www.tesisenred.net](http://www.tesisenred.net)) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

**WARNING.** On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX ([www.tesisenxarxa.net](http://www.tesisenxarxa.net)) service has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized neither its spreading and availability from a site foreign to the TDX service. Introducing its content in a window or frame foreign to the TDX service is not authorized (framing). This rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

QA  
280  
.P65  
1998

# CONSEQÜÈNCIES ECONOMÈTRIQUES DEL GRAU DE MEMÒRIA EN DADES TEMPORALS

Ernest Pons Fanals

Tesi Doctoral dirigida pel Dr. Jordi Suriñach i Caralt en el marc del programa de doctorat "Economia i Territori. Mètodes Quantitatius" de la Universitat de Barcelona.

Departament d'Econometria, Estadística i Economia Espanyola.

Barcelona, Octubre de 1998.

C.B.B. de la Universitat de Barcelona  
Departament d'Econometria, Estadística i  
Economia Espanyola  
Tel. 40101000

**Capítol 7**  
**ESTIMACIÓ SEMIPARAMÈTRICA EN SÈRIES AMB MEMÒRIA**  
**LLARGA**





## 7.1. Introducció

Al marge de l'estimació paramètrica de models prou flexibles per contenir memòria llarga com són els models ARFIMA, hi ha una altra aproximació ben diferent per la realització d'inferència al voltant de la memòria llarga. Es tracta d'una aproximació semiparamètrica que no requereix una modelització específica de l'hipotètic procés generador de les dades sinó que es basa sols en algun supòsit sobre les seves característiques. L'avantatge d'aquesta aproximació és una major generalitat mentre que com inconvenient principal té l'inconvenient de tota estimació semiparamètrica, que almenys a nivell teòric és pitjor que una estimació paramètrica sempre que el model especificat sigui correcte.

Com es tracta d'hipòtesis més generals, no cal especificar un model  $ARFIMA(p,d,q)$  concret, per tant es poden evitar els problemes de biaix que presenta l'estimador MV degut als errors derivats d'estimar un model mal especificat.

De fet, creiem que el desenvolupament i l'aplicació d'aquests mètodes no és aliè als problemes que presenta l'estimació paramètrica ja analitzats al capítol anterior. Dit d'una altra manera, si es disposés d'estimadors paramètrics amb bones propietats i en particular sense gaire biaix, és molt probable que la literatura s'hagués interessat molt menys per l'aproximació semiparamètrica.

Una primera diferència entre els diferents mètodes d'estimació semiparamètrica proposats radica en la manera d'especificar el model semiparamètric. Així, a grans trets es poden distingir tres vies per fer-ho:

1. A través del concepte d'integració fraccional. En aquest cas, es suposa que la sèrie observada és integrada fraccionalment i es pretén estimar el paràmetre  $d$  sense haver d'especificar un model ARFIMA concret.
2. A través de l'estructura de la densitat espectral. Consisteix a suposar que la densitat espectral presenta unes determinades característiques. La hipòtesi d'aquest tipus més habitual passa per considerar que al voltant de la freqüència zero aquesta densitat espectral es pot representar com  $f(\omega) = \omega^{-2\delta} L(\omega)$  on  $L(\cdot)$  és una funció de variació suau ja que a partir d'aquesta hipòtesi es pot estimar<sup>1</sup> l'Índex del grau de memòria  $\delta$ .
3. A través de l'estructura de les autocovariàncies. A diferència de l'aproximació anterior, en aquest cas s'especifiquen unes certes característiques per les autocovariàncies del procés, per exemple que<sup>2</sup>  $\gamma_k \sim g(k)k^{2\delta-1}$  on  $g(\cdot)$  és una funció de variació suau a l'infinit. Aquest plantejament també permet obtenir estimacions del paràmetre  $\delta$ .

Les dues primeres aproximacions han concentrat el major interès i de fet, bona part de la recerca en l'àmbit de l'estimació semiparamètrica del grau de memòria es basa en usar el periodograma a les baixes freqüències, idea suggerida per primer cop per Granger i Joyeux (1980) i desenvolupada després per altres autors. Al llarg dels apartats següents es presenten algunes propostes que poden situar-se dins de cada una d'aquestes tres aproximacions citades.

A més dels mètodes d'estimació presentats cal destacar que a l'apartat 7.5 es presenten algunes modificacions de l'estimador proposat per Granger i Joyeux (1980) que permeten obtenir millors estimacions.

---

<sup>1</sup> Vegi's l'apartat 3.4 de la tesi.

<sup>2</sup> S'ha usat una parametrització diferent de l'habitual per que el paràmetre en les diferents aproximacions tingui la mateixa interpretació.

També és important destacar que una de les aplicacions principals d'aquests mètodes semiparamètrics és la possibilitat de realitzar algun tipus de contrast al voltant del valor del paràmetre d'integració fraccional  $d$  (o del paràmetre  $\delta$ ). En aquest sentit, recordi's que una de les conclusions més importants a la que s'ha arribat al capítol 5 de la tesi és la de que un dels problemes més grans que presenta la utilització pràctica de models ARFIMA és la dificultat per detectar a partir del correlograma si una certa sèrie temporal ha de ser modelitzada usant un model ARMA o bé un model ARFIMA.

Per tant, una alternativa a la utilització del correlograma per identificar el tipus de memòria és la d'usar alguna estimació de  $d$  o  $\delta$  per fer-ne algun contrast sobre el seu valor. En particular, té molt interès contrastar la hipòtesi nul·la  $d=0$  (o  $\delta=0$ ) contra la hipòtesi alternativa  $d \neq 0$  (o  $\delta \neq 0$ ), contrast conegut habitualment sota el nom de contrast d'integració fraccional.

En aquest sentit, es dedica l'apartat 7.6 a la comparació dels diferents mètodes d'estimació presentats al llarg del capítol per analitzar amb quin d'aquests estimadors s'obté un contrast més potent.

## 7.2. Estimació semiparamètrica de l'ordre d'integració fraccional

### 7.2.1. El mètode proposat per Janacek (1981)

Es tracta de la primera proposta per a l'estimació del paràmetre d'integració fraccional sense necessitat de suposar una parametrització explícita del procés generador de les dades. Aquest mètode té en comú amb altres mètodes d'estimació semiparamètrics, i aquesta és la seva principal aportació, el fet de basar-se en la representació en el domini de les freqüències del procés.

Concretament, sota el supòsit que les dades  $X_1, \dots, X_T$  s'han generat segons un procés que compleix  $(1-L)^d X_t = Y_t$  on  $Y_t$  és un procés amb memòria curta (densitat espectral acotada), la densitat espectral d' $X_t$  es pot representar com:

$$f(\omega) = |1 - e^{-i\omega}|^{-2d} f_Y(\omega) \quad (7.1)$$

on  $f_Y(\omega)$  és acotada. Per aïllar el paràmetre d'interès, és a dir  $d$ , n'hi ha prou amb prendre logaritmes de l'expressió (7.1):

$$\log f(\omega) = -d \log[2(1 - \cos \omega)] + \log f_Y(\omega) \quad (7.2)$$

L'avantatge d'aquesta representació es troba en el fet que el terme  $-2 \log[2(1 - \cos \omega)]$  es pot simplificar gràcies a la següent propietat:

$$-\frac{1}{2} \log[2(1 - \cos \omega)] = \sum_{k=1}^{\infty} \frac{1}{k} \cos(k\omega) \quad (7.3)$$

de manera que si es defineix una nova funció  $W(\omega) = -\frac{1}{2} \log[2(1 - \cos \omega)]$  es pot representar el logaritme de la densitat espectral de qualsevol procés integrat fraccionalment com:

$$\log f(\omega) = 2dW(\omega) + \log f_Y(\omega) \quad (7.4)$$

Com  $Y_t$  és un procés amb memòria curta, la seva densitat espectral es pot representar en forma de sèrie de Fourier, és a dir:

$$f_Y(\omega) = a_0 + a_1 \cos(\omega) + a_2 \cos(2\omega) + \dots \quad (7.5)$$

de manera que es compleix:

$$\frac{1}{\pi} \int_0^{\pi} W(\lambda) \log f(\lambda) d\lambda = d \sum_{k=1}^{\infty} \frac{1}{k^2} + \sum_{k=1}^{\infty} \frac{a_k}{2k} \quad (7.6)$$

Janacek (1981) argumenta que la successió de coeficients  $\alpha_k = (ka_k)/2$  recull informació important sobre el tipus de memòria de les dades. Concretament demostra que si el procés  $X_t$  és de memòria curta, els coeficients  $\alpha_k$  decreixen de forma exponencial mentre que si  $d > 0$ , els coeficients  $\alpha_k$  estan formats per un terme decreixent i per un terme constant proporcional a  $d$ .

Quan es disposa d'una sèrie d'observacions  $X_1, \dots, X_T$ , es poden estimar els coeficients  $a_k$  com:

$$\tilde{a}_k = \frac{1}{2m} \sum_{j=0}^m \log I(\omega_j) \cos(j\omega_j) \quad (7.7)$$

on  $I(\omega_j)$  és el periodograma en cada una de les freqüències harmòniques  $\omega_j$ . Per altra banda, si es vol trobar una estimació del paràmetre  $d$ , Janacek (1981) proposa usar l'equivalent mostral de l'expressió (7.6). De fet, si es disposa d'una estimació consistent de la densitat espectral d' $X_t$ ,  $\tilde{f}(\omega)$ , la següent integral:

$$S = \frac{1}{\pi} \int_0^{\pi} W(\lambda) \log \tilde{f}(\lambda) d\lambda \quad (7.8)$$

es distribueix aproximadament com una normal amb esperança:

$$E(S) = d \sum_{k=1}^{\infty} \frac{1}{k^2} + \sum_{k=1}^{\infty} \frac{a_k}{2k} \quad (7.9)$$

Per tant com els quocients  $a_k/(2k)$  tendeixen cap a zero, es pot usar com estimació de  $d$  el següent:

$$\tilde{d} = \frac{\left( S - \sum_{k=0}^M \frac{\tilde{a}_k}{k} \right)}{\sum_{k=0}^M k^{-2}} \quad (7.10)$$

### 7.2.2. El mètode proposat per Geweke i Porter-Hudak (1983)

Aquests autors proposen un mètode senzill d'estimar  $d$  en un procés fraccionalment integrat. Suposi's que es compleix  $(1-L)^d X_t = Y_t$  de manera que  $Y_t$  és estacionària amb densitat espectral no negativa, acotada i continua a l'interval  $[-\pi, \pi]$ . La densitat espectral de  $X_t$  és<sup>3</sup>:

<sup>3</sup> Expressió equivalent a (7.1) ja que:

$$f_x(\omega) = \{4\sin^2(\omega/2)\}^{-d} f_Y(\omega) \quad (7.11)$$

Noti's com aquest model tan general conté varis dels models presentats al capítol 4 com el soroll gaussià fraccional, els models ARFIMA, el model EXPFI, i a més, el model paramètric general de Fox i Taqu (1986).

En tot cas, el mètode d'estimació de Geweke i Porter-Hudak (1983) es pot interpretar com una modificació dels mètodes paramètrics proposats per Mohr (1981), Graf (1983) i Kashyap i Eom (1988).

Aquests autors proposen un estimador del paràmetre  $d$  sense necessitat d'especificar les característiques del procés  $Y_t$  gràcies al següent argument. Si es prenen logaritmes a l'expressió anterior s'obté:

$$\log f_x(\omega) = \log f_Y(0) - d\{4\sin^2(\omega/2)\} + \log \frac{f_Y(\omega)}{f_Y(0)} \quad (7.12)$$

i al substituir  $\omega$  per cada una de les freqüències de Fourier  $\omega_j = 2\pi j/T$  s'obté:

$$\log f_x(\omega_j) = \log f_Y(0) - d\{4\sin^2(\pi j/T)\} + \log \frac{f_Y(\omega_j)}{f_Y(0)} \quad (7.13)$$

Com l'espectre teòric no és conegut es pot introduir a cada costat d'aquesta igualtat el periodograma:

$$I(\omega_j) = \frac{1}{2\pi T} \left| \sum_{t=1}^n X_t e^{i\omega_j t} \right|^2 \quad (7.14)$$

de manera que s'obté:

$$\log I(\omega_j) = \log f_Y(0) - d\{4\sin^2(\pi j/T)\} + \log \frac{I(\omega_j)}{f_x(\omega_j)} + \log \frac{f_Y(\omega_j)}{f_Y(0)} \quad (7.15)$$

$$|1 - e^{-i\omega}|^2 = [1 - \cos(\omega) - i\sin(\omega)][1 - \cos(\omega) + i\sin(\omega)] = 2(1 - \cos(\omega)) = 4\sin^2(\omega/2)$$

Observi's que si la freqüència  $\omega_j$  és prou propera a zero, l'últim terme d'aquesta expressió és molt inferior a la resta de manera que és certa la següent aproximació:

$$\log I(\omega_j) \sim \log f_Y(0) - d \{4 \sin^2(\pi j / T)\} + \log \frac{I(\omega_j)}{f_X(\omega_j)} \quad (7.16)$$

Llavors per estimar  $d$  els autors proposen plantejar el següent model de regressió:

$$\log I(\omega_j) = \alpha + \beta R_j + \varepsilon_j \quad (7.17)$$

on:

$$R_j = -\log\{4 \sin^2(\pi j / T)\} \quad (7.18)$$

i:

$$\varepsilon_j = \log\left\{\frac{I(\omega_j)}{f_X(\omega_j)}\right\} \quad (7.19)$$

Així, a partir de (7.17), és possible estimar  $d$  per Mínims Quadrats Ordinaris, estimador que es coneix habitualment sota el nom d'estimador GPH:

$$\tilde{d} = \frac{\sum_{j=1}^m \left[ \log I(\omega_j) - \frac{1}{m} \sum_{j=1}^m \log I(\omega_j) \right] (R_j - \bar{R})}{\sum_{j=1}^m (R_j - \bar{R})^2} \quad (7.20)$$

En tot cas, és important observar que les propietats d'aquest estimador del paràmetre  $d$  depenen de les característiques estocàstiques d'aquest últim terme,  $\varepsilon_j$ , que juga el paper de terme de pertorbació en el model de regressió.

Així, si el procés  $Y_t$  fos soroll blanc, llavors el seu espectre  $f_Y(\omega)$  seria constant. Com les ordenades del periodograma són independents<sup>4</sup>, llavors les ordenades del periodograma

<sup>4</sup> Sols les de les freqüències fonamentals, que són les habitualment utilitzades.

normalitzat  $I(\omega_j)/f_Y(\omega_j)$  també són independents. Per tant, queda garantit que si  $Y_t$  és soroll blanc el mètode MQO proporciona un bon estimador de  $d$ .

En canvi, en general, quan  $Y_t$  no és soroll blanc, el seu espectre no té perquè ser constant de manera que no és certa la hipòtesi de independència entre els residus  $\varepsilon_j$ .

De totes maneres, tot i que el supòsit de independència no és cert en general, Geweke i Porter-Hudak argumenten que per  $d < 0$  els coeficients de l'expansió com mitjana mòbil són absolutament sumables i que el Teorema 3 de Hannan (1973, pàg. 169) permet demostrar que les ordenades del periodograma normalitzat són asimptòticament independents i idènticament distribuïdes com una exponencial amb esperança unitària.

Si així fos,  $\varepsilon_j$  seria asimptòticament *i.i.d.* amb esperança  $-C$  (on  $C$  és la constant d'Euler  $C=0.57721..$ ) i variància  $\pi^2/6$  i aquesta conclusió donaria fonament a usar la regressió anterior sense necessitat de suposar que  $Y_t$  és soroll blanc<sup>5</sup>. Aquest argument també ha estat defensat per Kashyap i Eom (1988), Mohr (1981), Graf (1983) i Brockwell i Davis (1992).

Però a Hurvich i Beltrao (1993) s'analitza la distribució asimptòtica del conjunt de valors  $I(\omega_j)/f_Y(\omega_j)$  per  $j=1, \dots, m$  per  $m$  fixat i es demostra que quan  $d \neq 0$  el límit de  $E(I(\omega_j)/f_Y(\omega_j))$  depèn tant de  $d$  com de  $j$ . Per tant, l'argument de Geweke i Porter-Hudak no és correcte.

Més concretament, l'esperança de les ordenades del periodograma normalitzat es poden expressar de manera exacta com<sup>6</sup>:

$$E \left[ \frac{I(\omega_j)}{f(\omega_j)} \right] = \frac{1}{f(\omega_j)} \int_{-\pi}^{\pi} F_T(\omega_j - \lambda) f(\lambda) d\lambda \quad (7.21)$$

<sup>5</sup> Pel cas  $d > 0$ , n'hi hauria prou amb estimar  $d$  a partir de les diferències de la variable enlloc de les dades originals.

<sup>6</sup> Vegis Priestley (1981, pàg. 419).



on  $F_T(\cdot)$  és el nucli de Fejer:

$$F_T(\omega) = \frac{1}{2\pi T} \frac{\sin^2(T\omega/2)}{\sin^2(\omega/2)} \tag{7.22}$$

Atès que el periodograma és un estimador asimptòticament sense biaix de la densitat espectral, és raonable pensar que aquesta esperança tendeixi asimptòticament a la unitat. De fet, Yajima (1989) demostra que per a qualsevol freqüència es compleix:

$$\lim_{T \rightarrow \infty} E \left[ \frac{I(\omega)}{f(\omega)} \right] = 1 \tag{7.23}$$

independentment del valor de  $d$ . El problema no considerat pels autors anteriorment és que pel càlcul del periodograma a les freqüències de Fourier, si es manté el subíndex  $j$ , quan canvia  $T$  també canvia la freqüència  $\omega_j$ .

De fet, Hurvich i Beltrao (1993a) demostren algunes propietats molt interessants a partir de l'expressió (7.21). En primer lloc, el límit d'aquesta esperança és:

$$L_j(d) = \lim_{T \rightarrow \infty} E \left[ \frac{I(\omega_j)}{f(\omega_j)} \right] = \frac{2}{\pi} \int_{-\infty}^{\infty} \frac{\sin^2(\omega/2)}{(2\pi j - \lambda)^2} \left| \frac{\lambda}{2\pi j} \right|^{-2d} d\lambda \tag{7.24}$$

Observi's que en el cas particular en que  $d=0$ , es compleix que  $L_j(d)=1$  i, per tant en aquest cas, l'argument que justifica la utilització de l'estimador GPH és correcte. Però quan  $d \neq 0$ , pot ser que  $L_j(d) \neq 1$  i per tant, l'estimador GPH és asimptòticament esbiaixat.

A més, a Hurvich i Beltrao (1993a) es demostra que tampoc és cert que asimptòticament les ordenades del periodograma normalitzades siguin independents. Pel que fa a la seva distribució asimptòtica es pot comprovar que:

$$\frac{I(\omega_j)}{f(\omega_j)} \Rightarrow \left\{ \frac{L_j(d)}{2} - L_j^*(d) \right\} Z_1^2 + \left\{ \frac{L_j(d)}{2} + L_j^*(d) \right\} Z_2^2 \tag{7.25}$$

on  $Z_1$  i  $Z_2$  són variables normals estàndards i  $L_j^*(d)$  és:

$$L_j^*(d) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\sin^2(\omega/2)}{(2\pi j - \omega)(2\pi j + \omega)} \left| \frac{\omega}{2\pi j} \right|^{-2d} d\omega \quad (7.26)$$

De fet, a Robinson (1995) també es demostra, usant un altre mètode, que aquestes ordenades normalitzades no són asimptòticament ni independents ni idènticament distribuïdes.

En tot cas, un aspecte molt interessant pel que fa a les propietats de l'estimador proposat per Geweke i Porter-Hudak, és analitzar el comportament de  $E(\varepsilon_j)$ . Concretament, en el seu treball es demostra que aquesta esperança és asimptòticament superior a  $-C$  i és una funció decreixent de  $j$  per valors fixats de  $d$ . Per tant, l'estimador anterior, per  $m$  fixat, és d'esperar que presenti un biaix positiu.

Tot i aquests problemes, l'estimador GPH és, amb diferència, l'estimador del paràmetre  $d$  que ha estat més usat, fins al moment, en l'anàlisi de variables econòmiques degut a la seva simplicitat. Entre aquestes aplicacions es poden destacar Cheung (1993), Cheung i Lai (1993), Diebold i Rudebusch (1989), Geweke i Porter-Hudak (1984), Porter-Hudak (1990), Shea (1991) i Tschering (1993).

En tot cas, un altre aspecte fonamental a l'hora d'aplicar aquest estimador és l'elecció de  $m$ . Diebold i Rudebusch (1989) proposen usar  $m=T^{1/2}$  mentre que Sowell (1992) argumenta que l'elecció de  $m$  s'ha de basar en la longitud dels cicles de menor període rellevants a l'anàlisi de la component de llarg termini. Una altra possibilitat interessant és escollir  $m$  de manera que la variància de l'estimador sigui aproximadament igual a  $\pi^2/6$ .

Entre els avantatges de l'estimador GPH cal assenyalar que es tracta d'un estimador molt senzill i que a més, és robust front a la no normalitat. De totes maneres, pels motius ja comentats, perd el seu atractiu quan  $Y_t=(1-L)^d X_t$ , presenta autocorrelació. Pel que fa al seu funcionament amb mostres finites, a Agiakloglou, Newbold i Wohar (1992) es realitza un

experiment de Monte Carlo que demostra que l'estimador GPH presenta un seriós biaix i és molt ineficient quan  $Y_t = (1-L)^d X_t$  és un AR(1) o un MA(1).

Hassler (1993a) proposa una generalització de l'estimador GPH, concretament:

$$\tilde{d}(l, m) = \frac{1}{S(l, m)} \sum_{j=l+1}^m (R_j - \bar{R}) \ln I(\omega_j) \quad (7.27)$$

on  $S(l, m)$  és el següent sumatori:

$$S(l, m) = \sum_{j=l+1}^m (R_j - \bar{R})^2 \quad (7.28)$$

Observi's que amb aquesta notació, l'estimador  $\tilde{d}(l, m)$  és precisament l'estimador GPH. A partir d'exercicis de simulació, Hassler (1993a) mostra que la regressió del periodograma ha de començar amb  $l=1$  ja que si  $l>1$  això no fa més que augmentar l'error quadràtic mitjà.

A més, també es mostra que si  $d < -0.5$  l'estimador presenta un gran biaix i, per tant, proposa estimar el paràmetre  $d$  a partir de les dades originals i mai a partir de les dades en diferències. De totes maneres, en el mateix treball mostra que l'elecció d' $m$  és encara un problema delicat a la pràctica ja que no és clar que l'elecció d'usar  $m = T^{1/2}$  sigui la millor.

Per altra banda, Hassler (1993a) demostra que si  $X_t$  compleix l'equació:

$$(1-L)^d \phi(L) X_t = \theta(L) \varepsilon_t \quad (7.29)$$

llavors el biaix de l'estimador GPH es pot aproximar per:

$$E\{\tilde{d} - d\} = \frac{\sum_{j=0}^m (R_j - \bar{R}) \left\{ \log|\theta(e^{i\omega_j})|^2 - \log|\phi(e^{i\omega_j})|^2 \right\}}{\sum_{j=0}^m (R_j - \bar{R})^2} \quad (7.30)$$

expressió que demostra que la presència de coeficients autoregressius o mitjana mòbil grans i provoquen un biaix important.

Hassler (1993c) ha mostrat evidència empírica també de que amb processos AR(1) aquesta aproximació del biaix és força bona i això el porta a proposar el càlcul de l'estimador  $\tilde{d}(1, m)$  amb diferents valors d' $m$  i si aquestes estimacions es mantenen estables a mesura que augmenta  $m$ , això es pot interpretar com evidència de que no hi ha biaix derivat de la presència dels paràmetres ARMA.

### 7.2.3. El mètode proposat per Robinson (1995)

Un cop demostrats els problemes que presenta l'estimador GPH i seguint la generalització proposada per Hassler (1993a), a Robinson (1995) es proposa una extensió multivariant de l'estimador GPH.

Atès que no s'utilitzarà aquest estimador al llarg de la tesi, se'l presenta de manera molt breu sense entrar en gaire detalls. Concretament, suposi's un procés multivariant de dimensió  $G$  de manera que cada un dels processos  $X_{g_t}$  per  $g=1, \dots, G$  té una densitat espectral que compleix

$$f_g(\omega) \sim C_g \omega^{-2d_g} \quad \text{si} \quad \omega \rightarrow 0+ \quad (7.31)$$

i que el periodograma de cada sèrie és:

$$I_g(\omega) = \frac{1}{2\pi T} \left| \sum_{t=1}^T X_{g_t} e^{i\omega t} \right|^2 \quad (7.32)$$

La generalització consisteix a fixar un valor  $J \geq 1$ , i estimar els valors  $d_g$  a partir dels models de regressió següents:

$$H_{g_j}^{(J)} = c_g^{(J)} - d_g (2 \log \omega_j) + U_{g_j}^{(J)} \quad (7.33)$$

on:

$$H_{gk}^{(J)} = \log \left\{ \sum_{j=1}^J I_g(\omega_{k+j-J}) \right\} \quad (7.34)$$

per  $g=1 \dots G$  i  $k=l+J, \dots, l+2J, \dots, m$ .

A partir de l'equació (7.33), es pot aplicar MQO per trobar estimadors dels paràmetres  $d_g$  i  $c_g$ . De fet, si es certa la hipòtesi de que cada procés  $X_{gk}$  té per densitat espectral (7.31), es compleix que:

$$c_g^{(J)} = \log C_g + \psi(J) \quad (7.35)$$

on  $\psi(z)$  és la funció digamma:

$$\psi(z) = \frac{\partial \log \Gamma(z)}{\partial z} \quad (7.36)$$

La qüestió fonamental a analitzar en aquest cas és el comportament de les variables inobservables  $U_{gk}^{(J)}$  ja que les seves característiques determinen les propietats de l'estimador.

Com ja és ben conegut, per a que l'estimador MQO sigui el millor estimador lineal i no esbiaixat n'hi ha prou amb que els vectors:

$$U_j^{(J)} = (U_{1j}^{(J)}, \dots, U_{Gj}^{(J)})' \quad (7.37)$$

siguin incorrelacionats i homoscedàstics amb esperança zero, però no cal que els diferents elements de  $U_j^{(J)}$  siguin incorrelacionats.

Si es pren  $J=1$ , es tracta d'una generalització multivariant de GPH. De fet, a l'hora d'estimar  $d$  a partir d'una regressió sobre el periodograma és equivalent usar<sup>7</sup>  $-\log(4\sin^2(\omega/2))$  o bé usar  $-2\log(\omega)$ .

---

<sup>7</sup> Vegi's la nota a peu de pàgina n° 3.

A partir dels residus estimats del model, es pot calcular la matriu de variàncies i covariàncies dels estimadors. Un possible estimador de la matriu de variàncies i covariàncies dels residus, suggerida en el mateix treball és el següent:

$$\begin{aligned}\tilde{\Omega}^{(J)} &= \frac{J}{m-l} \sum_j \tilde{U}_j^{(J)} \tilde{U}_j^{(J)'} = \\ &= \frac{J}{m-l} \sum_j \left[ H_{g_j}^{(J)} - \tilde{c}_g^{(J)} - \tilde{d}_g(2 \log \omega_j) \right] \left[ H_{g_j}^{(J)} - \tilde{c}_g^{(J)} - \tilde{d}_g(2 \log \omega_j) \right]'\end{aligned}\quad (7.38)$$

Pel que fa a les propietats dels estimadors  $\tilde{c}^{(J)}$  i  $\tilde{d}^{(J)}$ , a Robinson (1995) es demostra el següent teorema:

**Teorema 7.1.** Suposi's que les següents hipòtesis són certes:

1) Existeixen  $C_g \in (0, \infty)$ ,  $d_g \in (-1/2, 1/2)$  i  $\alpha \in (0, 2]$  de manera que les densitats espectrals  $f_g(\omega)$  compleixen:

$$f_g(\omega) = C_g \omega^{-2d} + O(\omega^{\alpha-2d_g}) \quad \omega \rightarrow 0+, \quad g = 1, \dots, G \quad (7.39)$$

2) En un entorn  $(0, \varepsilon)$  de l'origen, les densitat espectral  $f_g(\omega)$  són diferenciables i es compleix que:

$$\left| \frac{\partial}{\partial \omega} f_{gh}(\omega) \right| = O(\omega^{-1-d_g-d_h}) \quad \omega \rightarrow 0+, \quad g, h = 1, \dots, G \quad (7.40)$$

3) Existeix un coeficient  $\beta \in [0, 2]$  tal que es compleix:

$$\left| \kappa_{gh}(\omega) - \kappa_{gh}(0) \right| = O(\omega^\beta) \quad \omega \rightarrow 0+, \quad g < h = 2, \dots, G \quad (7.41)$$

on  $\kappa_{gh}(\omega)$  és la coherència entre les sèries  $g$ -èssima i  $h$ -èssima.

4) La matriu  $\kappa(0)$  és no singular on  $\kappa(\omega)=[\kappa_{gh}(\omega)]$  és la matriu formada per les coherències entre cada parell de processos.

5)  $X_t$  és un procés amb densitat normal.

6) Quan  $T \rightarrow \infty$  es compleix que:

$$\frac{m^{1/2} \log m}{l} + \frac{l(\log T)^2}{m} + \frac{m^{1+1/2\min(\alpha,\beta)}}{T} \rightarrow 0 \tag{7.42}$$

Llavors, sota aquestes hipòtesis, quan  $T \rightarrow \infty$  es compleix:

$$\begin{bmatrix} \frac{m^{1/2}}{\log T} (\tilde{c}^{(J)} - c^{(J)}) \\ 2m^{1/2} (\tilde{d}^{(J)} - d^{(J)}) \end{bmatrix} \Rightarrow N \left[ 0, J \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \otimes \Omega^{(J)} \right] \tag{7.43}$$

Finalment, cal fer vèries observacions importants en relació a aquest teorema:

- Per aprofitar l'avantatge de la simplicitat del mètode MQO, pel que fa a l'estimació del paràmetre  $d$  i usant la notació més habitual en models de regressió, el teorema anterior afirma que:

$$\left[ \sum_j (2 \log \omega_j)^2 - \frac{J}{m-1} \left( \sum_j 2 \log \omega_j \right)^2 \right]^{1/2} \tilde{\Omega}^{(J)-1/2} (\tilde{d}^{(J)} - d) \Rightarrow N(0, I_G) \tag{7.44}$$

- Per cada una de les sèries  $X_{g^*}$ , es té que  $2m^{1/2} (\tilde{d}_g^{(J)} - d_g) \Rightarrow N(0, J\psi'(J))$ . Així, quan  $J=1$ ,  $J\psi'(J)=\pi^2/6=1.645$ . Per calcular que passa per altres valors de  $J$  pot usar-se la següent fórmula de recurrència:  $\psi'(J+1)=\psi'(J)-J^{-2}$ . Així, per  $J=2$  s'obté  $J\psi'(J)=1.289$ , per  $J=3$   $J\psi'(J)=1.185$  i així successivament tendint cap a zero.

- Per realitzar inferència sobre un sol dels paràmetres  $d_g$  es pot usar com variància  $J\psi'(J)$  enlloc dels elements de la diagonal de la matriu estimada de variàncies i covariàncies del teorema anterior,  $\tilde{\Omega}^{(J)}$ , però si es desitja realitzar inferència sobre més d'un dels paràmetres, cal usar aquesta matriu.

### 7.3. Estimació semiparamètrica a partir de la densitat espectral

Recordi's que a l'apartat 3.7 s'han definit els conceptes de memòria curta i memòria llarga a partir de la densitat espectral d'una sèrie temporal. Per tant, sembla interessant la possibilitat de determinar de manera semiparamètrica el tipus de memòria d'una certa variable a partir d'estimar aquesta densitat espectral. En aquest sentit, s'ha justificat a l'apartat 3.7 que no és gens restrictiu suposar que si el procés estocàstic  $X_t$  presenta memòria llarga, llavors la seva densitat espectral es pot representar com:

$$f(\omega) = |\omega - \omega_0|^{-2\delta} L(\omega) \quad (7.45)$$

on  $L(\omega)$  compleix unes certes condicions de regularitat<sup>8</sup> i  $0 \leq \delta < 1/2$ . Per tant, centrant-nos en la freqüència zero, es pot suposar que aquesta densitat es pot representar com  $f(\omega) = \omega^{-2\delta} L(\omega)$  on la funció  $L(\omega)$  és prou regular per garantir que quan  $\omega$  tendeix a zero,  $f(\omega)$  tendeix cap a infinit.

Recordi's de l'apartat 3.4 que s'ha definit aquest coeficient  $\delta$  sota el nom d'Índex del grau de memòria. A partir d'aquesta representació, per conèixer el tipus de memòria sols cal trobar-ne una estimació del coeficient  $\delta$ . Si  $\delta=0$ , llavors la sèrie té memòria curta, mentre que si  $\delta>0$ , llavors la sèrie té memòria llarga.

---

<sup>8</sup> Vegi's l'apartat 3.7 de la tesi.



En un seguit d'articles, Robinson analitza diverses aproximacions a l'estimació del paràmetre  $\delta$  sense necessitat de suposar una representació específica de la funció  $L(\omega)$ . A Robinson (1994a) es basa en usar la funció de distribució associada a la densitat espectral:

$$F(\omega) = \int_0^{\omega} f(\lambda) d\lambda \quad (7.46)$$

A la pràctica, quan la densitat  $f(\omega)$  és desconeguda, també ho és la funció  $F(\omega)$ . De totes maneres, a la literatura sobre sèries temporals hi ha bastants aportacions sobre el problema de l'estimació de la funció  $F(\omega)$ . Probablement la més popular d'aquestes aproximacions es basa en substituir la densitat espectral teòrica  $f(\omega)$  pel periodograma  $I(\omega)$  i calcular<sup>9</sup>:

$$\tilde{F}(\omega) = \int_0^{\omega} I(\lambda) d\lambda \quad (7.47)$$

Observi's que per avaluar aquesta expressió caldria usar el valor de periodograma en totes les freqüències, no sols en les freqüències fonamentals. Recordi's que per una freqüència  $\omega$  qualsevol es defineix el periodograma d'una sèrie  $X_1, \dots, X_T$  amb esperança  $\mu$  en aquesta freqüència com:

$$I(\omega) = \frac{1}{2\pi T} \left| \sum_{i=1}^T (X_i - \mu) e^{i\omega} \right|^2 \quad (7.48)$$

Habitualment, es calcula el periodograma suposant  $\mu=0$  ja que el resultat és el mateix per a totes les freqüències excepte la freqüència zero. Per tant, com en la majoria de casos l'esperança  $\mu$  és desconeguda és millor calcular el periodograma com:

$$I(\omega) = \frac{1}{2\pi T} \left| \sum_{i=1}^T X_i e^{i\omega} \right|^2 \quad (7.49)$$

---

<sup>9</sup> Vegi's Ibragimov (1963).

En aquest cas, és millor usar enlloc de l'estimador (7.47), una expressió més senzilla que, a més, és invariant front a variacions de  $\mu$ :

$$\tilde{F}(\omega) = \frac{2\pi}{T} \sum_{j=1}^{\left[\frac{\omega T}{2\pi}\right]} I(\omega_j) \quad (7.50)$$

Aquest estadístic, tot i que no és gens habitual a l'anàlisi economètrica, si ho és a l'anàlisi teòrica de sèries temporals. Així, per exemple, a Brillinger (1975) es proposa usar el quocient  $\frac{\tilde{F}(\omega_m)}{\omega_m}$  com estimador de  $f(0)$  quan  $f(0)$  és finit (és a dir, seguint les definicions de la tesi, quan la sèrie té memòria curta).

Per poder usar l'estadístic (7.50) com a eina per l'estimació del paràmetre  $\delta$ , cal fer algun supòsit addicional sobre la densitat espectral de la sèrie analitzada. Concretament, a Robinson (1994a) es parteix del supòsit de que aquesta densitat compleix<sup>10</sup>:

$$f(\omega) \underset{\omega \rightarrow 0}{\sim} L\left(\frac{1}{\omega}\right) \omega^{-2\delta} \quad (7.51)$$

on la funció  $L(\cdot)$  és de variació regular a l'infinit amb  $\delta \in (0, 1/2)$ . Això garanteix, en particular que la densitat espectral  $f(\omega)$  és de variació regular en la freqüència zero. Suposant algunes condicions de regularitat addicionals sobre la sèrie<sup>11</sup>  $X_t$ , a Robinson (1994a) es demostra que:

<sup>10</sup> A l'article original, es suposa que  $f(\omega) \underset{\omega \rightarrow 0}{\sim} L\left(\frac{1}{\omega}\right) \omega^{1-2H}$  seguint la notació popularitzada per Hurst (1951) però al llarg d'aquest apartat s'ha adaptat la notació per fer-la coherent amb les definicions introduïdes al llarg de la tesi.

<sup>11</sup> Concretament ha de complir les següents condicions:

1) Existeix pel procés estocàstic  $X_t$ , una representació de Wold:

$$X_t = \mu + \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} \quad \text{amb} \quad \sum_{j=0}^{\infty} \psi_j^2 < \infty$$

que compleix:

(i)  $E(\varepsilon_t, \varepsilon_u) = 0 \quad t > u$

$$\frac{\tilde{F}(\omega_m)}{F(\omega_m)} \xrightarrow{p} 1 \text{ quan } T \rightarrow \infty \tag{7.52}$$

sempre que  $m$  compleixi  $\frac{1}{m} + \frac{m}{T} \xrightarrow{T \rightarrow \infty} 0$ , dit d'una altra manera, sempre que  $m$  tendeixi cap a infinit però de manera menys ràpida que  $T$ .

Aprofitant la propietat (7.52) i les característiques de la funció  $L(\omega)$ , es pot comprovar que per qualsevol valor  $q \in (0, 1)$  es compleix:

$$\frac{F(q\omega)}{F(\omega)} \underset{\omega \rightarrow 0}{\sim} q^{1-2\delta} \tag{7.53}$$

A partir d'aquí és natural suggerir com estimador de l'índex  $\delta$  la següent expressió:

$$\tilde{\delta}_{m,q} = \frac{1}{2} - \frac{1}{2 \log(q)} \log \left\{ \frac{\tilde{F}(q \omega_m)}{\tilde{F}(\omega_m)} \right\} \tag{7.54}$$

En aquest sentit, a Robinson (1994a) es demostra que per qualsevol valor  $q \in (0, 1)$  aquest estimador és consistent sempre que es compleixin les condicions anteriors.

A Lobato i Robinson (1996) es complementen els resultats anteriors i se'n deriva la distribució asimptòtica de l'estimador (7.54). Concretament, aquesta distribució és diferent segons el valor del paràmetre  $\delta$ :

$$(ii) E(\varepsilon_r, \varepsilon_s, \varepsilon_t, \varepsilon_u) = \begin{cases} \sigma^4 & r = s > t = u \\ 0 & \text{altrament} \end{cases}$$

(iii) Existeix una variable aleatoria  $e$  tal que qualsevol  $\nu > 0$  hi ha un  $K < 1$  de manera que  $E(e^2) < \infty$  i  $P(|e_r| > \nu) \leq KP(e > \nu)$

(iv) Es compleix que  $\frac{1}{T} \sum_{i=1}^T E(\varepsilon_i | \varepsilon_s, s < i) \xrightarrow{p} \sigma^2$  quan  $T \rightarrow \infty$ .

- Si  $0 < \delta < 1/4$  llavors l'estimador és asimptòticament normal i:

$$\sqrt{m}(\tilde{\delta}_{m,q} - \delta) \Rightarrow N\left(0, \frac{(1+q^{-1} - 2q^{-2\delta})(1-2\delta)}{2(\log q)^2(1-4\delta)}\right) \quad (7.55)$$

- Si  $1/4 < \delta < 1/2$  l'estimador no és asimptòticament normal amb una distribució força més complicada.

Per altra banda, les propietats estadístiques de l'estimador depenen de l'elecció dels paràmetres  $m$  i  $q$ . Sobre aquest aspecte, a Robinson (1994b) es presenta la manera d'escollir el paràmetre  $m$  per tal de minimitzar l'error quadràtic mitjà de l'estimador, mentre que a Lobato i Robinson (1996) es discuteix l'elecció òptima del paràmetre  $q$ . A més, a Delgado i Robinson (1994) i Lee i Robinson (1993) se'n poden trobar aplicacions.

En un altre article, i imposant algunes restriccions addicionals sobre la densitat espectral de la sèrie analitzada, es proposa un estimador semiparamètric alternatiu també basat en el domini de les freqüències. Així, a Robinson (1995b) es recupera una proposta de Künsch (1987) de la que en demostra algunes propietats. Així, suposi's que els valors observats d'una certa variable s'han generat a partir d'un procés estocàstic amb una densitat espectral que compleix:

$$f(\omega) \underset{\omega \rightarrow 0}{\sim} C\omega^{-2\delta} \quad (7.56)$$

En aquest cas, Robinson (1995b) proposa buscar el valor  $\delta$  que fa mínima la funció següent:

$$Q(\delta) = \log\left[\frac{1}{m} \sum_{j=1}^m \omega_j^\delta I(\omega_j)\right] - \delta \frac{1}{m} \sum_{j=1}^m \log \omega_j \quad (7.57)$$

Un cop estimat el paràmetre  $\delta$  com  $\tilde{\delta}$  es pot calcular una estimació de la constant  $C$  com:

$$\tilde{C} = \frac{1}{m} \sum_{j=1}^m \omega_j^{\tilde{\delta}} I(\omega_j) \quad (7.58)$$

A Robinson (1995b) es demostra, sota algunes hipòtesi addicionals sobre la funció de densitat espectral (7.56) que si  $\frac{1}{m} + \frac{m}{T} \xrightarrow{T \rightarrow \infty} 0$  llavors els estimadors (7.57) i (7.58) són consistents amb ordre de convergència igual que l'estimador semiparamètric anterior, és a dir  $m^{1/2}$ . En tot cas, l'avantatge d'aquest nou estimador és que es demostra, també sota certes condicions de regularitat, que és asimptòticament normal i que a més, la variància asimptòtica de  $\tilde{\delta}$  no depèn del valor real del paràmetre.

Una altra aproximació similar a l'anterior es pot trobar a Robinson (1994c) a partir dels treballs de Hannan (1963), Robinson (1972) i Thomson (1986). Novament, si s'aproxima la densitat espectral al voltant de la freqüència zero per  $f(\omega) = C\omega^{-2\delta}$  a l'interval  $(0, \varepsilon)$ , té sentit calcular estimacions d'aquests paràmetres a partir de la minimització de la següent aproximació a la versemblança:

$$L(C, \delta) = \frac{1}{T} \sum_{j=1}^m \left\{ \log(\tilde{C}\omega_j^{-2\delta}) + \frac{I(\omega_j)}{C\omega_j^{-\delta}} \right\} \tag{7.59}$$

on  $m$  és l'índex màxim tal que  $\omega_j < \varepsilon$ . En aquest cas, es pot comprovar com els valors dels paràmetres  $C$  i  $\delta$  que minimitzen aquesta funció poden calcular-se trobant en primer lloc  $\tilde{\delta}$  que fa mínima la següent expressió:

$$m \log \left( \sum_{j=1}^m I(\omega_j) \omega_j^{2\delta} \right) - 2\delta \sum_{j=1}^m \log \omega_j \tag{7.60}$$

i després calculant l'estimador de  $C$  com:

$$\tilde{C} = \frac{1}{m} \sum_{j=1}^m I(\omega_j) \omega_j^{2\tilde{\delta}} \tag{7.61}$$

Segons Robinson (1994c), es pot justificar que aquest estimador és consistent i asimptòticament normal per un  $\varepsilon$  fix combinant les tècniques de Fox i Taqu (1986), Robinson (1972) i Thomson (1986).

Pel que fa a les propietats de totes i cada una d'aquestes aproximacions l'autor d'aquesta tesi no té coneixement de que es disposi pel moment de resultats que facin referència a les propietats d'aquests estimadors en mostres finites i, per tant, aquest és un aspecte sobre el que cal continuar la recerca.

#### 7.4. Estimació semiparamètrica a partir de les autocovariàncies

En tot cas, com ja s'ha discutit al capítol 3, el supòsit de que densitat espectral es comporta al voltant de la freqüència zero com  $f(\omega) = C\omega^{-2\delta}$  és similar a la condició en termes d'autocovariàncies de que<sup>12</sup>:

$$\gamma_k \underset{k \rightarrow \infty}{\sim} Dk^{2\delta-1} \quad (7.62)$$

Per tant, una altra possibilitat suggerida a Robinson (1994c) i sobre la que l'autor d'aquesta tesi no té constància que s'hagi analitzat fins al moment consisteix a usar estimadors de les autocovariàncies com a base per l'estimació de l'Índex del grau de memòria.

Un argument a favor d'aquesta aproximació es troba en el fet de que aquests estimadors són consistents, mentre que el periodograma és un estimador inconsistent de la densitat espectral.

Així, si hom pren logaritmes de l'expressió (7.62), s'obté que:

$$\log(\gamma_k) \underset{k \rightarrow \infty}{\sim} \log D + (2\delta - 1) \log k \quad (7.63)$$

Aquesta aproximació suggereix que es podrien calcular estimacions del paràmetre  $\delta$  a partir d'un model de regressió lineal simple on com valors de la variable endògena s'utilitzin els valors  $\log(\gamma_k)$  i com valors de la variable exògena els valors  $\log(k)$ :

$$\log(\gamma_k) = \alpha + \beta \log k + \varepsilon_k \quad (7.64)$$

<sup>12</sup> La relació entre ambdues condicions ja ha estat comentada a l'apartat 3.7 de la tesi.

Com l'aproximació és millor com major sigui el valor del retard  $k$ , és millor a l'hora de calcular estimacions d' $\alpha$  i  $\beta$ , usar les últimes  $m$  autocovariàncies. Així, usant les autocovariàncies estimades, l'estimador MQO de  $\beta$  és:

$$\tilde{\beta} = \frac{\sum_{j=T-m}^{T-1} \left( \log \tilde{\gamma}_j - \frac{1}{m} \sum_{j=T-m}^{T-1} \log \tilde{\gamma}_j \right) \left( \log j - \frac{1}{m} \sum_{j=T-m}^{T-1} \log j \right)}{\sum_{j=T-m}^{T-1} \left( \log j - \frac{1}{m} \sum_{j=T-m}^{T-1} \log j \right)^2} \quad (7.65)$$

i, per tant, l'estimador del paràmetre  $\delta$  és:

$$\tilde{\delta} = \frac{1}{2} + \frac{\sum_{j=T-m}^{T-1} \left( \log \tilde{\gamma}_j - \frac{1}{m} \sum_{j=T-m}^{T-1} \log \tilde{\gamma}_j \right) \left( \log j - \frac{1}{m} \sum_{j=T-m}^{T-1} \log j \right)}{2 \sum_{j=T-m}^{T-1} \left( \log j - \frac{1}{m} \sum_{j=T-m}^{T-1} \log j \right)^2} \quad (7.66)$$

De totes maneres, cal destacar que per que a que sigui possible usar aquest estimador cal que totes les autocovariàncies estimades siguin positives quan  $j \geq T-m$ .

Una altra possibilitat també a partir de les últimes autocovariàncies estimades, molt menys elaborada i suggerida per Robinson (1990) consisteix a trobar els valors dels paràmetres  $D$  i  $\delta$  que fan mínima la següent funció:

$$Q(D, \delta) = \sum_{j=T-m}^{T-1} (\tilde{\gamma}_j - Dj^{2\delta-1})^2 \quad (7.67)$$

Una altra alternativa consisteix a usar una extensió semiparamètrica de l'estimador MDE presentat a l'apartat 6.10, el proposat per Tieslau *et al.* (1996). Recordi's que sota el supòsit de que les dades s'han generat a partir d'un model ARFIMA(0, $d$ ,0) aquest estimador es defineix com aquell que fa mínima la funció següent:

$$S(d) = [\tilde{\rho} - \rho(d)]' W[\tilde{\rho} - \rho(d)] \quad (7.68)$$

on  $W$  són un conjunt de ponderacions.

Basant-se en aquesta observació, en el mateix article es proposa estendre aquest estimador al cas en que es suposa que les dades s'han generat a partir d'un model ARFIMA( $p,d,q$ ). En tot cas, les propietats de l'estimador que s'obtindrà a través d'aquesta extensió vindran molt condicionades per les propietats de les autocorrelacions mostrals.

Concretament, a partir dels resultats de Hosking (1984) que deriva les distribucions asimptòtiques de les autocovariàncies mostrals d'un soroll blanc fraccional, Tieslau *et al.* (1996) avaluen l'eficiència asimptòtica de MDE. Tal com observa Hosking (1984) i demostra formalment Dahlhaus (1988, 1989) la presència de diferents tipus de memòria fa que les autocorrelacions mostrals convergeixin de manera diferent segons el valor del paràmetre  $d$ .

Així, quan les dades s'han generat a través d'un model ARFIMA( $0,d,0$ ) amb  $-0.5 < d < 0.25$  l'estadístic MDE convergeix a una distribució normal amb una velocitat  $T^{1/2}$  mentre que per  $d=0.25$  MDE també convergeix a una normal però amb una velocitat igual a  $\{T/\log(T)\}^{1/2}$ . En canvi, per  $0.25 < d < 0.5$ , MDE convergeix a una distribució no estandar a una velocitat  $T^{1/2-d}$ .

A Tieslau *et al.* (1996) també s'avalua l'eficiència teòrica del estimador quan  $X_t$  és ARFIMA( $0,d,0$ ) per diferents valors de  $d$  tals que  $-0.5 < d < 0.25$ , i per diferents valors de  $k$  i  $T$  i es comprova que la pèrdua d'eficiència per l'augment de  $k$  és molt important.

A més, si s'utilitza el mateix mètode per estimar  $d$  en models ARFIMA( $p,d,q$ ) que en models ARFIMA( $0,d,0$ ) aquest estimador introdueix un biaix molt important. Fins i tot el biaix no desapareix fins i tot quan  $k$  tendeix a infinit ja que les autocorrelacions d'un model ARFIMA( $p,d,q$ ) són diferents d'un ARFIMA( $0,d,0$ ) per tots els retards.

Aquests problemes suggereixen com a solució usar els quocients d'autocovariàncies. De fet, si  $\gamma_k$  és l'autocovariància de un procés ARIMA ( $0,d,0$ ), les autocovariàncies de qualsevol procés ARIMA ( $p,d,q$ ) tenen en el límit la forma  $\gamma_k^* \approx [\theta(1)/\phi(1)]\gamma_k$ . Així, tot i que les



autocovariàncies en ambdós model siguin diferents, el rati de decreixement d'aquestes autocovariàncies sols depèn del paràmetre que mesura el comportament a llarg termini, no de la dinàmica a curt termini.

A partir d'aquesta idea, hom podria plantejar-se com estimació trobar aquell valor de  $d$  que minimitza la següent funció<sup>13</sup>:

$$S^*(d) = [\tilde{\rho}^* - \rho^*(d)]' W [\tilde{\rho}^* - \rho^*(d)] \quad (7.69)$$

on el vector  $\rho^*(d)$  està format per quocients de coeficients d'autocorrelació:

$$\rho^*(d) = \left( \frac{\rho_{k+1}}{\rho_k} \quad \frac{\rho_{k+2}}{\rho_k} \quad \dots \quad \frac{\rho_{k+n}}{\rho_k} \right)' \quad (7.70)$$

En aquest cas, a diferència de l'estimador MDE paramètric per models ARFIMA(0,d,0), no està clar quina és la matriu de ponderacions  $W$  que cal usar.

Tot i que la proposta original està pensada per l'estimació del paràmetre  $d$  en models ARFIMA, creiem que la mateixa idea es podria aplicar a l'estimació semiparamètrica del paràmetre  $\delta$  sense necessitat d'especificar el tipus de model, sols a partir del supòsit que les autocovariàncies es poden aproximar per (7.62). En aquest cas, el vector  $\rho^*(d)$  és de la forma:

$$\rho^*(\delta) = \left( \frac{(k+1)^{2\delta-1}}{k^{2\delta-1}} \quad \frac{(k+2)^{2\delta-1}}{k^{2\delta-1}} \quad \dots \quad \frac{(k+n)^{2\delta-1}}{k^{2\delta-1}} \right)' \quad (7.71)$$

Observi's com en aquest cas, la funció a minimitzar no depèn del paràmetre  $C$ , sols de l'índex del grau de memòria  $\delta$ .

<sup>13</sup> La notació usada és pròpia ja que a Tieslau et al. (1997) no s'analitza aquesta proposta amb gaire detall.

En tot cas, val la pena destacar que totes aquestes propostes no són més que això, propostes, ja que encara no s'han desenvolupat el suficient com per conèixer les seves propietats estadístiques. Són idees que obren altres vies de recerca per l'estimació semiparamètrica del grau de memòria. De totes maneres, creiem que tots aquests estimadors es veurien afectats pels problemes de biaix mostrats al capítol 5 de la tesi.

Al marge d'aquestes propostes, hi ha un enfocament ben diferent que si s'ha usat bastant per mesurar el grau de memòria. Es tracta d'una modificació de l'estadístic proposat per Hurst (1951) i presentat a l'apartat 2.5.

Recordi's que a Hurst (1951) es proposa modelitzar la presència de memòria llarga a través de la següent especificació:

$$\gamma_k \sim Ak^{2(H-1)} \quad (7.72)$$

on  $H \in (1/2, 1)$  mesura el grau de memòria. Observi's que l'expressió (7.72) és equivalent a (7.62) amb  $\delta = H - 1/2$  i  $D = A$ . Per tant, la proposat de Hurst (1951) és totalment equivalent a la utilitzada en aquest apartat.

En aquest context, i tal com s'ha analitzat al capítol 3, l'estadístic rang reescalat es defineix com:

$$Q_T = \frac{1}{S_T} \left[ \text{Max}_{1 \leq k \leq T} \sum_{j=1}^k (X_j - \bar{X}) - \text{Min}_{1 \leq k \leq T} \sum_{j=1}^k (X_j - \bar{X}) \right] \quad (7.73)$$

on  $S_T$  és l'estimador habitual de la desviació estàndard:

$$S_T = \sqrt{\frac{1}{T} \sum_{i=1}^T (X_i - \bar{X})^2} \quad (7.74)$$

El fet de proposar l'estadístic  $Q_T$  es deu a que, sota la hipòtesi (7.72), es compleix la següent propietat:

$$\log[E(Q_T)] \approx \text{constant} + H[\log(T)] = \text{constant} + \left(d + \frac{1}{2}\right)[\log(T)] \quad (7.75)$$

i per tant, això suggereix estimar el paràmetre  $\delta$  com:

$$\tilde{\delta} = \frac{\log(Q_T)}{\log(T)} - \frac{1}{2} \quad (7.76)$$

En un conjunt d'articles de Mandelbrot, Taqqu i Wallis es demostra la superioritat d'aquest tipus d'anàlisi sobre els mètodes convencionals per detectar la presència de memòria llarga, com pot ser l'anàlisi de les autocorrelacions, els quocients de variàncies i les descomposicions espectrals. Per exemple, a Mandelbrot i Wallis (1969) es demostra com l'estadístic R/S permet detectar la presència de memòria llarga en dades altament no normals amb curtosi i/o asimetria molt importants. De fet, Mandelbrot (1972, 1975) demostra la convergència quasi-segura de R/S en processos estocàstics amb variància infinita, propietat que no compleixen ni els quocients de variàncies ni la descomposició espectral que necessiten que la variància sigui finita

Altres aspectes de robustesa d'aquest estadístic es poden trobar a Fox i Taqqu (1979), i Mandelbrot (1972) afirma que mentre l'anàlisi espectral sols detecta cicles periòdics, l'estadístic R/S també detecta cicles no periòdics. A més, a Mandelbrot i Wallis (1968, 1969) i Davies i Harte (1987) es presenten mètodes alternatius per estimar  $\delta$  també a partir de la relació (7.75).

Pel que fa a la distribució de l'estadístic  $Q_T$ , Lo (1991) demostra que sota la hipòtesi de soroll blanc,  $T^{1/2}Q_T$  convergeix en distribució al rang d'un moviment brownià estàndard, que la seva esperança és  $(\pi/2)^{1/2}$  i la seva desviació estàndard és  $[(\pi/2)(\pi-3)/3]^{1/2}$ .

De totes maneres, si les dades observades s'han generat per un procés amb memòria llarga més complexa que conté també elements de memòria curta, per exemple un model ARFIMA( $p,d,q$ ), el comportament en mostra finita de l'estadístic  $Q_T$  s'hi veu afectat. Per aquest motiu, molts dels primers investigadors en aquesta àrea es mostren preocupats pels problemes que pot presentar aquest estadístic quan les dades han estat generades per una combinació d'un procés I(0) i un procés amb llarga memòria.

Dintre d'aquest grup de treball són destacables els de Anis i Lloyd (1976) que analitzen el biaix de l'estadístic  $Q_T$  en mostra petita, mentre que Mandelbrot (1972, 1975), Mandelbrot i Wallis (1968), Davies i Harte (1987), Aydogan i Booth (1988) i Lo (1991) n'analitzen la robustesa de front a la presència simultània de components de memòria curta i components de memòria llarga.

La principal conclusió d'aquests treballs és que qualsevol diferència entre el comportament de les dades i l'associat amb un soroll blanc, fa augmentar la probabilitat d'acceptar la presència de memòria llarga, per tant, afecta al tamany del contrast.

Com a solució a aquest problema Siddiqui (1976) proposa modificar l'estadístic  $Q_T$  substituint la variància de la sèrie amb un estimador robust de la variància a llarg termini de la sèrie, idea que després incorpora Lo (1991). Concretament, a Lo (1991) es proposa contrastar la presència de memòria llarga usant l'estadístic anomenat sovint com rang reescalat modificat o ajustat (o MRR) i que es defineix com:

$$Q^*_T = \frac{1}{\tilde{\sigma}_T(q)} \left[ \max_{1 \leq u \leq T} \sum_{t=1}^u (X_t - \bar{X}) - \min_{1 \leq u \leq T} \sum_{t=1}^u (X_t - \bar{X}) \right] \quad (7.77)$$

on  $\tilde{\sigma}_T(q)$  és una estimació robusta de la variància a llarg termini:

$$\tilde{\sigma}_T^2(q) = c_0 + 2 \sum_{j=1}^q w_j(q) c_j \quad (7.78)$$

i  $w_j(q)$  són els pesos de la finestra de Barlett<sup>14</sup>:

$$w_j(q) = 1 - \frac{j}{q+1}, \quad j < q \quad (7.79)$$

<sup>14</sup> Un aspecte important arribats a aquest punt fora estudiar els avantatges i inconvenients d'utilitzar altres finestres.

En el context dels tests d'arrels unitàries, Phillips (1987) ha mostrat que  $\tilde{\sigma}_T^2(q)$  és consistent si  $q \sim O(T^{1/4})$ . Lo (1991) analitza la distribució de l'estadístic  $Q^*_T$  i tabula valors crítics per la seva utilització per contrastar la presència de memòria llarga.

En el mateix treball es demostra que l'estadístic MRR és consistent contra un bon nombre de processos amb llarga memòria gaussians i estacionaris. De totes maneres l'elecció del paràmetre  $q$  a utilitzar suposa també un problema. En tot cas, diferents exercicis de simulació realitzats no permeten ser gaire optimistes sobre aquest procediment ja que, com mostren Anis i Lloyd (1976), l'estadístic MRR està esbiaixat en mostres finites degut a la necessitat d'estimar la variància de la sèrie.

Tot i que s'acostuma a usar la finestra de Bartlett, es podrien usar altres finestres. Concretament, la finestra de Parzen és òptima en el sentit de error quadràtic mig asimptòtic i la de Bartlett-Priestley minimitza l'error quadràtic mig relatiu.

En general, però, Hausser (1997) demostra que les propietats de l'estadístic depenen més de l'amplitud de la finestra que del tipus de finestra. Una solució a aquest problema és usar el procediment proposat per Andrews (1991) de selecció automàtica d'amplitud de la finestra derivat sota el supòsit de que el procés generador de les dades és conegut.

De totes maneres, a l'anàlisi econòmica s'hi poden trobar varies aplicacions de l'estadístic MRR. Algunes d'elles es poden trobar a Greene i Fielitz (1977), Booth *et al.* (1982), Helms *et al.* (1984), Kaen i Rosenman (1986) i Aydogan i Booth (1988).

Finalment, a l'apartat 7.6 de la tesi es comparen les propietats d'aquest i d'altres estadístics proposats a la literatura per realitzar inferència sobre l'Índex  $\delta$ .

## 7.5. Algunes extensions de l'estimador GPH

### 7.5.1. Extensió a partir d'una estructura autoregressiva

Entre tots els mètodes d'estimació proposats un dels més populars i sens dubte, el més usat en les aplicacions pràctiques fins al moment és el mètode proposat per Geweke i Porter-Hudak basat en el periodograma. De totes maneres, aquest mètode presenta un problema important en la mesura que no permet estimar de manera simultània el paràmetre  $d$  associat a la memòria llarga i els paràmetres ARMA associats a la memòria curta i per tant, és fàcil que es produeixin distorsions.

Per fer-ne una anàlisi més acurada de les propietats d'aquest estimador s'ha realitzat un ampli exercici de simulació a partir del qual s'obtenen algunes conclusions força importants. A partir de models ARFIMA(1, $d$ ,1) amb diferents paràmetres s'han generat sèries amb diferents tamanyos mostrals i s'han calculat estimacions del biaix i de l'error quadràtic mitjà (EQM) en cada una d'aquestes situacions.

Recordi's que aquest mètode es basa en la igualtat següent:

$$\log I(\omega_j) = \log f_Y(0) - d\{4\sin^2(\pi j / T)\} + \log \frac{I(\omega_j)}{f_X(\omega_j)} + \log \frac{f_Y(\omega_j)}{f_Y(0)} \quad (7.80)$$

i consisteix a estimar per MQO la següent aproximació a la igualtat (7.80):

$$\log I(\omega_j) \sim \log f_Y(0) - d\{4\sin^2(\pi j / T)\} + \log \frac{I(\omega_j)}{f_X(\omega_j)} \quad (7.81)$$

sota el supòsit de que l'últim terme de (7.80) pren valors molt propers a zero.

A l'Annex 7.1 es presenten els resultats d'aproximar el biaix i l'EQM d'aquest estimador a partir de la generació de 5000 sèries de cada model analitzat. En cada cas, s'han usat diferents

valors de  $m$ , per valorar la importància de l'elecció d'aquest paràmetre en les propietats de l'estimador.

Els resultats de la simulació permeten treure les següents conclusions, molt similars a les que s'obtenen a Cheung (1993b):

- El biaix de l'estimador GPH no depèn gairebé del valor del paràmetre  $d$ . En canvi, a mesura que augmenta el valor del paràmetre autoregressiu  $\phi$  aquest biaix augmenta. De fet, quan  $\phi < 0.8$ , el biaix no és gaire important, però quan  $\phi \geq 0.8$  el biaix és molt important.
- En relació al paràmetre mitjana mòbil  $\theta$  passa una cosa similar, quan  $\theta < 0.8$ , el biaix no és gaire important, mentre que si  $\theta \geq 0.8$  aquest es torna força important.

Si hom vol trobar un mètode d'estimació que permeti solucionar aquests problemes, cal analitzar amb més cura la causa d'aquest biaix tan important. Concretament, el terme de la igualtat (7.80) que s'elimina perquè a freqüències baixes és proper a zero, té conseqüències importants en el comportament del test a mesura que augmenta  $T$ .

Si el terme de pertorbació no presenta independència, per exemple quan  $X_t$  és un MA(1) o un AR(1), a mesura que augmenta  $T$  l'estimador presenta cada cop un biaix major ja que a mesura que ens allunyem de la freqüència zero, majors modificacions presenta l'espectre i el terme eliminat s'allunya del valor zero. Així, valors dels paràmetres de memòria curta com  $\theta = 0.5$  o  $\phi = 0.5$  ja introdueixen un biaix en l'estimador quan augmenta  $m$ .

Al marge del comentat a l'apartat anterior, hi ha un altra biaix independent de  $m$  derivat de que el terme de pertorbació no és soroll blanc. Aquest també està associat amb la presència de termes MA(1) o AR(1) però sobretot per paràmetres de la memòria curta grans com poden ser

$\theta=0.9$  o  $\phi=0.9$ . Aquest biaix és independent front a variacions de  $m$ , i també és gairebé independent del paràmetre  $d$  real.

El biaix produït per l'augment del paràmetre  $m$  no és fàcil d'eliminar, però el biaix més important, l'introduït per la presència de coeficients autoregressius o mitjana mòbil molt importants, pot corregir-se si s'utilitza una expansió més precisa de l'aproximació (7.81). En aquest sentit, es proposa a continuació una modificació de l'estimador GPH que permet eliminar en part el biaix introduït pels coeficients autoregressiu o mitjana mòbil.

Per introduir aquesta modificació, suposi's que la sèrie observada conté paràmetres associats amb la presència de memòria curta i de memòria llarga de manera simultània. El cas més senzill en que es produeix aquesta situació és el d'un model ARFIMA(1,d,0). En aquest cas, la densitat espectral de la sèrie és:

$$f_x(\omega) = \frac{\sigma_\varepsilon^2}{2\pi} \left\{ 4 \sin^2(\omega/2) \right\}^{-d} \left\{ 1 + \phi^2 - 2\phi \cos(\omega) \right\}^{-1} \quad (7.82)$$

Per tant, si hom sap amb seguretat que una sèrie s'ha generat a partir d'un model ARFIMA(1,d,0), com pel Teorema de Taylor es compleix la següent propietat:

$$\log\{1 + \phi^2 - 2\phi \cos(\omega_j)\} \approx \phi^2 - 2\phi \cos(\omega_j) \quad (7.83)$$

pot estimar-se el paràmetre  $d$  per MQO a través del següent model de regressió:

$$\log I(\omega_j) = \alpha + \beta \log\{4 \sin^2(\omega_j/2)\} + \gamma \cos(\omega_j) + \varepsilon_j \quad (7.83)$$

fent després  $\tilde{d} = -\tilde{\beta}$ . De totes maneres, aquest mètode sols és aplicable per models ARFIMA(1,d,0) i sempre que sigui conegut que aquest és el procés generador de dades.

Observi's que l'avantatge del model (7.83) es troba en la presència de dos regressors, un associat a la component de memòria llarga, i l'altra associat amb la memòria curta. Per tant, tot i que el mètode anterior no sigui aplicable en general, creiem que aquesta idea, la d'incloure en el model de regressió a estimar un factor addicional que permeti recollir, en



part, les variacions del periodograma derivades de factors de memòria curta, permetria reduir el biaix en l'estimació del paràmetre  $d$ .

Més concretament, suposi's que la sèrie  $X_t$  s'ha generat segons un model ARFIMA( $p,d,q$ ) amb  $p>0$ . En aquest cas, es pot expressar  $X_t$  com:

$$(1-L)^d(1-\phi L)X_t = Z_t \quad (7.84)$$

on  $Z_t$  segueix un model ARFIMA( $p-1,0,q$ ).

De la condició (7.84) es dedueix que la densitat espectral de les dades observades es pot expressar com:

$$f_X(\omega) = \{4\sin^2(\omega/2)\}^{-d} \{1 + \phi^2 - 2\phi \cos(\omega)\}^{-1} f_Z(\omega) \quad (7.85)$$

Si el polinomi autoregressiu d' $X_t$ , tot i tenir ordre  $p$ , no presenta més d'una arrel superior a 0.8, i les arrels del polinomi mitjana mòbil no són massa importants, és d'esperar que l'espectre  $f_Z(\omega)$ , tot i no ser el d'un soroll blanc, pugui utilitzar-se ara el mateix argument de Geweke i Porter-Hudak (1983) per estimar el paràmetre  $d$  per MQO, sense que la presència d'altres factors de memòria curta no modelats de manera específica provoquin un biaix en aquest estimador.

De fet, no cal suposar que la sèrie segueix un ARFIMA( $p,d,q$ ) amb  $p>0$  perquè si  $p=0$ , l'expressió (7.85) és certa amb  $\phi=0$ . Per altra banda, si el problema del biaix prové de la presència d'un coeficient mitjana mòbil superior a 0.8, també fora possible trobar una aproximació a (7.85) a partir de calcular  $(1-\theta L)^{-1}$ .

Per tant, utilitzar l'expressió (7.85) no suposa la imposició de cap hipòtesi gaire restrictiva sobre el procés generador de dades. Simplement, s'inclou un factor addicional per a que reculli, en cas de que les dades continguin paràmetres associats a la memòria curta que

puguin provocar distorsions l'estimació de  $d$ , la influència d'aquests paràmetres sobre el periodograma de les dades observades.

Així, prenent logaritmes de l'expressió (7.85) s'obté:

$$\begin{aligned} \log f_x(\omega) = & \log f_z(0) - d \log\{4\sin^2(\omega/2)\} - \\ & - \log\{1 + \phi^2 - 2\phi \cos(\omega)\} + \log \frac{f_z(\omega)}{f_z(0)} \end{aligned} \quad (7.86)$$

Com l'espectre teòric no és conegut és millor introduir a cada costat d'aquesta igualtat el periodograma de manera que s'obté, per cada una de les freqüències fonamentals:

$$\begin{aligned} \ln I(\omega_j) = & \ln f_z(0) - d\{4\sin^2(\pi j/T)\} - \ln\{1 + \phi^2 - 2\phi \cos(2\pi j/T)\} + \\ & + \ln \frac{I(\omega_j)}{f_z(\omega_j)} + \ln \frac{f_z(\omega_j)}{f_z(0)} \end{aligned} \quad (7.87)$$

Usant el mateix argument que a Geweke i Porter-Hudak (1983) s'arriba a la següent aproximació:

$$\begin{aligned} \ln I(\omega_j) \sim & \ln f_z(0) - d\{4\sin^2(\pi j/T)\} - \ln\{1 + \phi^2 - 2\phi \cos(2\pi j/T)\} + \\ & + \ln \frac{I(\omega_j)}{f_z(\omega_j)} \end{aligned} \quad (7.88)$$

i, tenint el compte la propietat (7.83), pot estimar-se el paràmetre  $d$  usant el següent model de regressió ampliat:

$$\ln I(\omega_j) = \ln f_z(0) - d \ln\{4\sin^2(\pi j/T)\} + \gamma \cos(2\pi j/T) + \varepsilon_j \quad (7.89)$$

és a dir:

$$\ln I(\omega_j) = \alpha + \beta R_j + \gamma C_j + \varepsilon_j \quad (7.90)$$

on:

$$R_j = -\ln\{4\sin^2(\pi j/T)\} \quad (7.91)$$

i:

$$C_j = \cos(2\pi j/T) \quad (7.92)$$

Sobre aquest mètode, que anomenarem com GPHA, cal fer les següents observacions:

- Es podria ampliar la regressió amb més termes afegint altres cosinus per recollir millor els paràmetres a curt termini.
- Si es vol minimitzar el biaix que es produeix al usar freqüències més llunyanes de l'origen, cal usar un ordre  $m$  petit. De totes maneres, si  $m$  és petit i s'inclouen més termes al model (7.89), pot presentar-se un problema de multicolinealitat important.

De totes maneres, l'aproximació utilitzada:

$$\ln\{1 + \phi^2 - 2\phi \cos(2\pi j/T)\} \approx \phi^2 - 2\phi \cos(2\pi j/T) \quad (7.93)$$

sols és una bona aproximació per valors del paràmetre autoregressiu propers a zero i això pot fer que la capacitat de l'estimador GPHA de reduir el biaix de l'estimador GPH no sigui gaire important per paràmetres autoregressius grans.

Com a solució a aquest problema, es proposa a continuació una altra extensió de l'estimador GPH, basada en la mateixa idea, però evitant la utilització de l'aproximació (7.93). Aquest nou mètode, que anomenarem estimador GPHB, es basa en l'estimació d'un model AR(1) per a les dades observades com a etapa prèvia a l'estimació de  $d$ .

Naturalment, l'estimador  $\tilde{\phi}$  que s'obtingui pot presentar un biaix molt important si  $d \neq 0$ . De totes maneres, no cal interpretar el valor obtingut com a estimador de  $\phi$ , sinó únicament incloure'l en l'aproximació (7.88), de manera que:

$$\begin{aligned} \log I(\omega_j) \sim \log f_z(0) - d\{4\sin^2(\pi j/T)\} - \\ - \log\{1 + \tilde{\phi}^2 - 2\tilde{\phi} \cos(2\pi j/T)\} + \log \frac{I(\omega_j)}{f_z(\omega_j)} \end{aligned} \quad (7.94)$$

Ara, pot estimar-se el paràmetre  $d$  a partir del següent model de regressió:

$$\ln I(\omega_j) + \ln\{1 + \tilde{\phi}^2 - 2\tilde{\phi} \cos(2\pi j/T)\} = \ln f_y(0) - d \ln\{4\sin^2(\pi j/T)\} + \varepsilon_j \quad (7.95)$$

és a dir, estimant per MQO el paràmetre  $\beta$  al model:

$$Y_j = \alpha + \beta R_j + \varepsilon_j \quad (7.96)$$

on:

$$Y_j = \log I(\omega_j) + \log\{1 + \tilde{\phi}^2 - 2\tilde{\phi} \cos(2\pi j/T)\} \quad (7.97)$$

i fent després  $\tilde{d} = -\tilde{\beta}$ .

Òbviament, el problema d'aquest mètode es troba en el fet d'usar un estimador del paràmetre  $\phi$  esbiaixat i que aquest biaix pot afectar a les propietats de  $\tilde{d}$ .

Per valorar correctament els avantatges i inconvenients de cada una d'aquestes dues propostes, es presenten a l'Annex 7.2 els resultats d'una simulació en que es comparen els estimadors GPHA i GPHB amb dos estimadors semiparamètrics de  $d$  que es poden qualificar de *clàssics*, l'estimador GPH i l'estimador proposat per Robinson (1994a) i presentat a l'apartat 7.3, basat en l'Índex del grau de memòria,  $\tilde{\delta}_{m,q}$ .

A partir dels resultats de l'Annex 7.2 es pot arribar a les següents conclusions:

- Les característiques de l'estimador  $\tilde{\delta}_{m,q}$  són molt sensibles a l'elecció del paràmetre  $q$ , però si s'escull aquest coeficient de manera correcta, presenta en general, un biaix i un EQM lleugerament inferior al de l'estimador GPH.
- Pel que fa a l'estimador GPHB, tot i que presenta un EQM lleugerament inferior al de l'estimador GPH en els casos en que aquest últim presenta més dificultats, hi ha alguns casos, bàsicament amb la presència de coeficients mitjana mòbil negatius molt grans, en que l'estimador GPHB presenta un biaix negatiu molt important mentre que l'estimador GPH funciona força bé.
- Pel que fa a l'estimador GPHA tot i que presenta un EQM lleugerament superior al de l'estimador GPH en els casos en que aquest funciona bé, quan hi ha coeficients autoregressius o mitjana mòbil positius i grans, s'obté un EQM clarament inferior usant

GPHA, tot i que la diferència en termes de biaix no és gaire important. De fet, en aquestes situacions, l'estimador  $\tilde{\delta}_{m,q}$  té un biaix clarament inferior, tot i que al comparar l'EQM que presenten és millor l'estimador GPHA.

En definitiva, l'estimador GPHA tot i que manté la simplicitat de l'estimador GPH permet, gràcies a l'extensió del model de regressió, obtenir estimacions més precises, sobretot quan els paràmetres autoregressiu i mitjana mòbil no són propers a la unitat. Quan aquests són propers a la unitat és millor usar GPHB o el mètode proposat per Robinson (1994a). Per altra banda, tot i que l'estimador GPHB presenta millores en relació a GPH en aquestes situacions, és pitjor en d'altres i, per tant, no sembla una alternativa interessant.

De totes maneres, creiem que la manera d'aprofitar al màxim els avantatges d'aquestes extensions consisteix a aplicar de manera simultània els diferents estimadors analitzats en aquesta simulació per comparar-ne els resultats tenint en compte els comentaris anteriors.

### **7.5.2. Extensió a partir del model de Bloomfield**

A més de les extensions de l'estimador GPH proposades a l'apartat anterior, hi ha una altra possibilitat bastant natural d'avançar en la millor de l'estimació del paràmetre  $d$  que, de manera sorprenent, no ha estat considerada fins al moment a la literatura. En tot cas, es presenta a continuació aquesta proposta de manera breu i es deixa plantejada com a línia futura de recerca.

Es tracta de basar l'anàlisi en els models EXPFI enlloc dels models ARFIMA. De fet, s'ha demostrat a l'apartat 6.13 de la tesi que el model EXP té un grau de representativitat similar al dels models ARMA i per tant, si una variable  $X_t$  és FI( $d$ ), enlloc de suposar que segueix un model ARFIMA( $p,d,q$ ) pot usar-se com a supòsit alternatiu que segueix un model EXPFI( $p,d$ ).

L'avantatge d'aquest enfocament es troba en que llavors pot estendre's el model de regressió en el domini de les freqüències usat a GPH d'una manera bastant natural per trobar estimacions precises de  $d$ . Suposi's que  $X_t$  és FI( $d$ ), de manera que  $Y_t = (1-L)^d X_t$  sols presenta memòria curta i per tant, pot suposar-se que la densitat espectral d' $Y_t$  és:

$$f_Y(\omega) = \frac{\sigma^2}{2\pi} \exp\left\{2 \sum_{k=1}^p \beta_k \cos(k\omega)\right\} = f_Z(\omega) \exp\left\{2 \sum_{k=1}^p \beta_k \cos(k\omega)\right\} \quad (7.98)$$

on  $Z_t$  és soroll blanc.

Per tant, la densitat espectral d' $X_t$  és:

$$\begin{aligned} \log f_X(\omega) &= -d \log\{4 \sin^2(\omega/2)\} + \log f_Y(\omega) = \\ &= -d \log\{4 \sin^2(\omega/2)\} + \log f_Z(\omega) + 2 \sum_{k=1}^p \beta_k \cos(k\omega) \end{aligned} \quad (7.99)$$

Reordenant de manera adequada l'expressió (7.99) s'obté una variació de la relació que s'usa com a base en l'estimador GPH:

$$\log f_X(\omega) = \log f_Z(0) - d \log\{4 \sin^2(\omega/2)\} + 2 \sum_{k=1}^p \beta_k \cos(k\omega) + \log \frac{f_Z(\omega)}{f_Z(0)} \quad (7.100)$$

Ara, fent servir les ordenades del periodograma s'obté:

$$\log I(\omega_j) \sim \log f_Z(0) - d \log\{4 \sin^2(\pi j/T)\} + 2 \sum_{k=1}^p \beta_k \cos(k\omega_j) + \log \frac{I(\omega_j)}{f_Z(\omega_j)} \quad (7.101)$$

Per tant, per obtenir una estimació del paràmetre  $d$ , n'hi ha prou amb estimar per MQO el següent model de regressió:

$$\log I(\omega_j) = \alpha + \beta R_j + \sum_{k=1}^p \beta_k \cos(k\omega_j) + \varepsilon_j \quad (7.102)$$

Observi's, per tant, que l'estimador GPHA no és més que un cas particular d'aquest últim estimador on es fixa  $p=1$ .

En tot cas, i atès que l'objectiu central de la tesi no és avançar en el desenvolupament de mètodes d'estimació del paràmetre d'integració fraccional, es deixa aquesta proposta com una línia oberta de recerca, però creiem que és una alternativa força interessant a l'estimació màxim versemblant quan no es disposa d'informació addicional sobre el tipus de model a especificar.

## 7.6. La utilització dels estimadors semiparamètrics per realitzar inferència sobre el tipus de memòria

Una de les principals aplicacions d'aquests mètodes semiparamètrics és la possibilitat de realitzar algun tipus de contrast al voltant del valor del paràmetre d'integració fraccional  $d$ . Recordi's del capítol 5 de la tesi que un dels problemes més grans que presenta la utilització de models ARFIMA és la dificultat detectar si una certa sèrie temporal ha de ser modelada usant un model ARMA, un model ARIMA o bé un model ARFIMA.

Dit d'una altra manera, si hom basa l'anàlisi d'una certa variable  $X_t$  en un model el més general possible i suposa que els valors d'aquesta variable han estat generats a partir d'un model ARFIMA( $p, d, q$ ):

$$\phi(L)(1-L)^d X_t = \theta(L)\varepsilon_t \quad (7.103)$$

té força interès decidir si  $d=0$ , si  $d=1$ , si  $0 < d < 1/2$  o bé si  $1/2 \leq d < 1$ , perquè en cada una d'aquestes situacions el tractament de les dades ha de ser diferent. Una de les conclusions que es dedueixen dels resultats presentats al llarg de la tesi és que la decisió més problemàtica atenent a les conseqüències que té equivocar-se és discriminar entre  $d=0$  i  $d \neq 0$ .

Això explica que a la literatura sobre sèries temporals s'hagi manifestat tant interès pel que s'anomenen contrastos d'integració fraccional, és a dir contrastar la hipòtesi nul·la  $H_0: d=0$  contra la alternativa  $H_1: d \neq 0$ .

Naturalment, tot i que tot mètode d'estimació del paràmetre d'integració fraccional  $d$  pugui servir per realitzar aquest contrast, sempre serà millor un mètode semiparamètric d'estimació que un mètode paramètric, ja que al imposar un menor nombre de restriccions, el ventall de situacions considerades sota la hipòtesi alternativa és major. En altres paraules, tant l'estimador GPH com l'estadístic MRR semblen interessants per aquest objectiu.

De fet, a la literatura economètrica s'hi poden trobar varis treballs en que s'ha usat, preferentment l'estimador GPH per contrastar hipòtesis sobre el paràmetre<sup>15</sup>  $d$  i es pretén valorar l'eficàcia tant d'aquest com d'altres contrastos.

Però no es pot fer aquesta valoració sense tenir en compte altres estadístics proposats de manera específica per contrastar si es pot rebutjar que  $d=0$ . En particular, a Robinson (1991) es desenvolupa un test de multiplicadors de Lagrange que tot i ser vàlid en un context molt més ampli, es pot particularitzar a l'objectiu aquí plantejat.

Concretament, i seguint a Cheung (1993), per contrastar si  $X_t$  és soroll blanc contra la alternativa de que segueixi un model ARFIMA(0,d,0) es pot usar algun dels dos estadístics següents:

$$\lambda_1 = T^{1/2} \left( \sum_{j=1}^{T-1} j^{-2} \right)^{-1/2} \sum_{j=1}^{T-1} \frac{C_j}{jC_0} \quad (7.104)$$

o:

---

<sup>15</sup> Un exemple clàssic de la utilització de l'estimador GPH per a la realització d'inferència sobre el paràmetre  $d$  és Diebold i Rudebusch (1989).



$$\lambda_2 = T^{1/2} \left( \sum_{j=1}^{T-1} j^{-2} K_j \right)^{-1/2} \sum_{j=1}^{T-1} \frac{C_j}{j} \quad (7.105)$$

on els coeficients  $C_j$  són les autocovariàncies mostrals d'ordre  $j$ :

$$C_j = \frac{1}{T} \sum_{i=1}^{T-1} (X_i - \bar{X})(X_{i+j} - \bar{X}) \quad (7.106)$$

i els coeficients  $K_j$  són:

$$K_j = \frac{1}{T} \sum_{i=1}^{T-1} (X_i - \bar{X})^2 (X_{i+j} - \bar{X})^2 \quad (7.107)$$

Sota la hipòtesi nul·la de que  $X_t$  sigui soroll blanc, els dos estadístics es distribueixen asimptòticament com una normal estàndard però  $\lambda_2$  té l'avantatge respecte a  $\lambda_1$  que és robust front a la presència d'heteroscedasticitat.

Òbviament, el problema d'aquests estadístics és que es basen en unes hipòtesis nul·la i alternativa massa restrictives. Com a solució a això, a Agliakloglou i Newbold (1994) es desenvolupa un test de multiplicadors de Lagrange per contrastar si  $X_t$  segueix un model ARMA( $p, q$ ) contra la hipòtesi alternativa de que segueixi un model ARFIMA( $p, d, q$ )<sup>16</sup>.

En tot cas, cal destacar que aquest test parteix de la base d'uns valors pels paràmetres autoregressius i mitjana mòbil fixats. Així, un cop s'ha estimat un model ARMA( $p, q$ ), el test permet contrastar si es pot rebutjar que  $d=0$ , en favor d'un model ARFIMA( $p, d, q$ ) amb els mateixos paràmetres. Per a l'aplicació pràctica d'aquest test, es proposen dos procediments alternatius:

<sup>16</sup> Aquest test és similar als proposats per Godfrey (1979) i Poskitt i Tremayne (1980) per contrastar la necessitat d'incloure un paràmetre autoregressiu o un paràmetre mitjana mòbil addicional.

1) El primer anomenat test REG es basa en les següents etapes:

- Calcular estimacions de les innovacions del model a partir dels paràmetres estimats:

$$\tilde{\varepsilon}_t = \frac{\tilde{\theta}(L)}{\tilde{\phi}(L)} X_t \quad (7.108)$$

- Calcular dos noves sèries,  $W_t$  i  $Z_t$  a partir de:

$$\tilde{\theta}(L)W_t = X_t \quad \text{i} \quad \tilde{\theta}(L)Z_t = \tilde{\varepsilon} \quad (7.109)$$

- Fixar un valor  $m$  i estimar el següent model de regressió auxiliar usant  $t=m+1, \dots, T$ .

$$\tilde{\varepsilon}_t = \sum_{i=1}^p \beta_i W_{t-i} + \sum_{j=1}^q \gamma_j Z_{t-j} + \delta \left( \sum_{j=1}^m j^{-1} \tilde{\varepsilon}_{t-j} \right) + u_t \quad (7.110)$$

- Contrastar la hipòtesi  $\delta=0$  usant el  $t$ -estadístic habitual. Si es rebutja que  $\delta=0$ , llavors cal rebutjar el model ARMA( $p,q$ ) en favor del model ARFIMA( $p,d,q$ ).

2) El segon anomenat test AUTO es basa en les següents etapes:

- Calcular estimacions de les innovacions del model a partir dels paràmetres estimats com en el test REG a partir de (7.108).
- Fixar  $m$  i calcular els  $m$  primers coeficients d'autocorrelació d'aquests residus:

$$r_j = \frac{1}{T} \sum_{j=1}^{T-1} (\tilde{\varepsilon}_t - \bar{\varepsilon})(\tilde{\varepsilon}_{t+j} - \bar{\varepsilon}) \quad (7.111)$$

- Calcular la matriu  $W$  com:

$$W = \frac{1}{T(T+2)} LVL \quad (7.112)$$

on  $L$  és una matriu diagonal de dimensió  $(m \times m)$  on l' $i$ -èssim element és  $(T-i)^{1/2}$  i la matriu  $V$  és una matriu de dimensió  $(m \times m)$  està formada per les variàncies i covariàncies asimptòtiques

dels coeficients<sup>17</sup>  $r_j$  multiplicades per  $T$  utilitzant els valors estimats pels paràmetres autoregressius i mitjana mòbil.

- Calcular l'estadístic  $Z$  com:

$$Z = (h'Wh)^{-1/2} \sum_{j=1}^m j^{-1} r_j \quad (7.113)$$

on  $h$  és un vector de dimensió  $(m \times 1)$  on el  $j$ -èssim element és  $j^{-1}$ . L'estadístic  $Z$  es distribueix sota la hipòtesi nul·la asimptòticament com una normal estàndard.

Per valorar el funcionament tant de l'estimador GPH i l'estadístic MRR com dels tests de multiplicadors de Lagrange s'ha realitzat un exercici de Monte Carlo que permeti treure'n conclusions pràctiques. Per fer aquesta valoració s'han analitzat dos aspectes, tant el tamany empíric de cada un dels contrastos com la potència.

Per analitzar el tamany s'han generat 1000 sèries de 100 observacions a partir de models ARMA(1,1):

$$(1 - \phi L)X_t = (1 - \theta L)\varepsilon_t \quad (7.114)$$

amb diferents combinacions de valors dels paràmetres  $\phi$  i  $\theta$ , mentre que per analitzar la potència s'han generat 1000 sèries de 100 observacions a partir de models ARFIMA(0, $d$ ,0) amb diferents valors del paràmetre  $d$ .

Als quadres 7.7 i 7.8 es recullen alguns dels resultats més rellevants obtinguts en aquestes simulacions usant els següents contrastos<sup>18</sup>:

<sup>17</sup> A McLeod (1978) s'hi poden trobar expressions explícites per aquestes variàncies i covariàncies.

<sup>18</sup> En una sèrie de treballs no publicats, Andersson i Gredenhoff (1997 i 1998) suggereixen millorar la potència d'alguns d'aquests contrastos de integració fraccional usant la tècnica del Bootstrap, possibilitat no considerada en aquesta tesi.

- **MRR**: S'ha comparat el valor de l'estadístic  $Q_T$  definit a (7.77) amb els valors crítics tabulats per Lo (1991). El valor de  $q$  s'ha fixat seguint el criteri de Andrews (1991), és a dir<sup>19</sup>:

$$q = \left[ \left( \frac{3T}{2} \right)^{1/2} \left( \frac{2\tilde{\rho}}{1-\tilde{\rho}^2} \right)^{2/3} \right] \quad (7.115)$$

- **GPH**: S'ha estimat el paràmetre  $d$  usant l'estimador (7.19) i fixant  $m=T^{1/2}$  com proposen Geweke i Porter-Hudak (1983). A continuació s'ha contrastat la hipòtesi nul·la que  $d=0$  a través del  $t$ -estadístic habitual.
- **LM**: S'ha calculat l'estadístic  $\lambda_1$  definit a (7.104) i se l'ha comparat amb els valors d'una distribució normal estàndard.
- **REG1**: S'ha contrastat la hipòtesi  $\delta=0$  usant el  $t$ -estadístic habitual en el model de regressió (7.110). A falta de propostes sobre el valor d' $m$ , s'ha usat<sup>20</sup>  $m=10$ . A l'hora d'analitzar la potència del contrast, com s'ha generat  $X_t$  a partir d'un model ARFIMA(0, $d$ ,0) s'ha suposat que al ajustar un model ARMA( $p,q$ ) s'ha optat per un ARMA(0,0) i per tant s'ha contrastat com hipòtesi nul·la que  $X_t$  fos un soroll blanc contra la alternativa ARFIMA(0, $d$ ,0) amb  $d \neq 0$ .
- **REG2**: El supòsit en que es basa el test *REG1*, pot semblar massa restrictiu ja que si el PGD és ARFIMA(0, $d$ ,0) és probable que al ajustar un model ARMA( $p,q$ ) s'esculli de manera errònia un ARMA(1,0). Per això, com alternativa a l'hora d'analitzar la potència del mètode REG, tot i que el PGD sigui ARFIMA(0, $d$ ,0) s'ha estimat un model AR(1) i s'ha

---

<sup>19</sup> Aquesta expressió és el valor òptim per paràmetre de truncament d'un finestra de retard si les dades s'han generat segons un model AR(1). Si es suposa que el procés generador de les dades és un altre, llavors el valor òptim de  $q$  és un altre. Per altres possibilitats vegi's Andrews (1991).

<sup>20</sup> S'ha provat altres valors d' $m$  però tot i que el test pot millorar o empitjorar segons aquest valor, això no afecta a les comparacions entre els diferents tests analitzats.

contrastat la hipòtesi nul·la que  $X_t$  segueixi un model ARFIMA(1, $d$ ,0) contra la alternativa d'un model ARFIMA(1, $d$ ,0).

- **AUTO1**: S'ha calculat l'estadístic  $Z$  definit a (7.113) i se l'ha comparat amb els valors d'una distribució normal estàndard. Al igual que en el *REG1*, a l'hora d'analitzar la potència del contrast s'ha calculat aquest estadístic de manera adequada per contrastar com hipòtesi nul·la que  $X_t$  fos un soroll blanc contra la alternativa ARFIMA(0, $d$ ,0) amb  $d \neq 0$ . Novament, s'ha fixat  $m=10$ .
- **AUTO2**: S'ha usat el mateix procediment que a *AUTO1* però novament, com a *REG2*, s'ha estimat un model AR(1) i s'ha contrastat la hipòtesi nul·la que  $X_t$  segueixi un model ARFIMA(1, $d$ ,0) contra la alternativa d'un model ARFIMA(1, $d$ ,0).
- **GPHA**: S'ha estimat el paràmetre  $d$  a partir del model de regressió ampliat definit a (7.90) i s'ha contrastat la hipòtesi nul·la que  $d=0$  a través del  $t$ -estadístic habitual. Com en el cas GPH, s'ha usat també  $m=T^{1/2}$ .
- **GPHB**: S'ha estimat el paràmetre  $d$  a partir del model de regressió ampliat definit a (7.97) i s'ha contrastat la hipòtesi nul·la que  $d=0$  a través del  $t$ -estadístic habitual, també usant  $m=T^{1/2}$ .

Dels resultats que s'inclouen en els quadres 7.7 i 7.8, es poden treure les següents conclusions importants:

- El tests MRR i GPH són robusts front a components ARMA, no així el test LM
- El test MRR, però, té molt poca potència.

- Els tests REG1 i AUTO1 tenen una gran potència però es basen en contrastar la hipòtesi soroll blanc.

Quadre 7.7. Percentatge de rebutjos de l'hipòtesi nul·la  $d=0$ . PGD: ARMA(1,1)

$T$	$\phi$	$\theta$	MRR	GPH	LM	REG1	REG2	AUTO1	AUTO2	GPHA	GPHB
100	0.0	0.0	0.07	0.05	0.06	0.05	0.06	0.05	0.05	0.05	0.05
	-0.9	0.0	0.04	0.06	0.98	0.06	0.05	0.05	0.05	0.05	0.06
	-0.5	0.0	0.02	0.04	0.96	0.05	0.04	0.05	0.06	0.05	0.06
	0.5	0.0	0.02	0.07	0.93	0.06	0.08	0.08	0.09	0.05	0.06
	0.9	0.0	0.01	0.68	1.00	0.07	0.09	0.09	0.11	0.13	0.04
	0.0	-0.9	0.02	0.31	0.91	0.09	0.07	0.08	0.09	0.10	0.28
	0.0	-0.5	0.02	0.06	0.89	0.08	0.06	0.06	0.06	0.05	0.06
	0.0	0.5	0.03	0.05	0.86	0.05	0.04	0.05	0.05	0.05	0.05
	0.0	0.9	0.04	0.05	0.76	0.05	0.05	0.04	0.05	0.05	0.05
200	0.0	0.0	0.06	0.05	0.18	0.04	0.06	0.05	0.05	0.05	0.05
	-0.9	0.0	0.00	0.04	0.99	0.05	0.05	0.05	0.05	0.04	0.04
	-0.5	0.0	0.01	0.04	0.94	0.05	0.04	0.05	0.06	0.05	0.04
	0.5	0.0	0.04	0.06	1.00	0.07	0.08	0.08	0.09	0.05	0.05
	0.9	0.0	0.02	0.61	1.00	0.06	0.09	0.09	0.12	0.12	0.05
	0.0	-0.9	0.02	0.28	0.95	0.08	0.07	0.08	0.08	0.10	0.26
	0.0	-0.5	0.02	0.06	0.91	0.07	0.06	0.06	0.05	0.05	0.05
	0.0	0.5	0.03	0.05	0.89	0.05	0.04	0.05	0.05	0.05	0.05
	0.0	0.9	0.01	0.05	0.81	0.05	0.05	0.04	0.05	0.05	0.05

Quadre 7.8. Percentatge de rebutjos de l'hipòtesi nul·la  $d=0$ . PGD: ARFIMA(0, $d$ ,0). Potència.

$T$	$d$	MRR	GPH	LM	REG1	REG2	AUTO1	AUTO2	GPHA	GPHB
100	0.0	0.05	0.05	0.05	0.05	0.04	0.06	0.05	0.06	0.06
	0.1	0.04	0.06	0.24	0.18	0.07	0.25	0.08	0.06	0.07
	0.2	0.06	0.13	0.37	0.31	0.11	0.37	0.13	0.12	0.08
	0.3	0.09	0.24	0.78	0.75	0.19	0.83	0.22	0.22	0.19
	0.4	0.14	0.31	1.00	0.94	0.24	0.97	0.27	0.34	0.27
200	0.0	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
	0.1	0.08	0.08	0.36	0.19	0.09	0.24	0.10	0.12	0.09
	0.2	0.12	0.16	0.54	0.35	0.17	0.39	0.17	0.19	0.16
	0.3	0.23	0.29	0.94	0.79	0.24	0.86	0.26	0.31	0.26
	0.4	0.26	0.42	1.00	0.96	0.29	0.98	0.31	0.38	0.31

- Quan es contrasta la integració fraccional amb hipòtesi nul·la més flexible com és el cas de REG2 i AUTO2, la potència és molt inferior i és comparable a la de GPH.
- Això suggereix, tot i els problemes, usar GPH ja que és més senzill.

- Finalment, les modificacions de l'estadístic GPH proposades, sobretot GPHA permeten millorar les seves propietats ja que es solucionen les distorsions en el tamany del contrast. De fet, si se'l compara únicament amb els altres mètodes que permeten solucionar els problemes de tamany, és a dir, REG2 i AUTO2, s'obté que tot i ser més complicats, GPHA presenta una potència comparable.

La conclusió final és que per realitzar un contrast d'integració fraccional, el millor és utilitzar el contrast de multiplicadors de Lagrange proposat a Agiakloglou i Newbold (1994) o bé l'estimador GPHA proposat a la tesi que és més fàcil d'aplicar però gaudeix d'una menor fonamentació teòrica.

## 7.7. Consideracions finals

Un cop comprovats al capítol 6 els problemes que presenta l'estimació de sèries amb memòria llarga pel mètode de la màxima versemblança degut a les dificultats en l'especificació correcta dels models ARFIMA, s'han analitzat al llarg del capítol 7 diferents propostes semiparamètriques amb la finalitat de veure si per aquesta via s'obtenen millors resultats.

En general, tot i l'enorme recerca que s'ha generat al voltant d'aquest problema, tant pel que fa a l'estimació del valor del paràmetre  $d$ , com al disseny de contrastos d'integració fraccional, no es disposa ara per ara d'un mètode d'estimació que funcioni de manera prou satisfactòria en sèries temporals que no siguin gaire llargues.

En tot cas, s'ha proposat a la tesi una extensió de l'estimador presentat per Geweke i Porter-Hudak (1983) que presenta millors propietats i permet contrastar si  $d=0$  sense que es produeixin distorsions en el tamany del contrast. A més, s'ha comprovat que aquest nou contrast té

potència similar a la del contrast de multiplicadors de Lagrange proposat per Agiakloglou i Newbold (1994).

Finalment, s'ha suggerit també que cal avançar en l'estimació semiparamètrica de l'ordre d'integració fraccional a partir dels models EXPFI com alternativa a la utilització dels models ARFIMA.



## Annex 7.1. Propietats en mostra finita de l'estimador GPH

Entre tots els mètodes d'estimació proposats un dels més populars i sens dubte, el més usat en les aplicacions pràctiques fins al moment és el mètode proposat per Geweke i Porter-Hudak basat en el periodograma. De totes maneres, aquest mètode presenta un problema important en la mesura que no permet estimar de manera simultànea el paràmetre  $d$  associat a la memòria llarga i els paràmetres ARMA associats a la memòria curta i per tant, és fàcil que es produeixin distorsions.

A les pàgines següents es presenten alguns dels resultats d'un ampli exercici de simulació amb l'objectiu de valorar correctament aquests problemes. Concretament, s'han generat diferents models ARFIMA(1, $d$ ,1):

$$(1 - L)^d (1 - \phi L) X_t^* = (1 - \theta L) \varepsilon_t, \quad (\text{A7.1})$$

i per cada un d'aquests models s'han generat 5000 sèries de 100, 200 300 i 400 observacions. Per a cada una d'aquestes situacions s'ha estimat l'esperança i la desviació estàndard de l'estadístic:

$$\tilde{d} = \frac{\sum_{j=1}^m \left[ \log I(\omega_j) - \frac{1}{m} \sum_{j=1}^m \log I(\omega_j) \right] (R_j - \bar{R})}{\sum_{j=1}^m (R_j - \bar{R})^2} \quad (\text{A7.2})$$

on  $I(\omega_j)$  són les ordenades del periodograma en les freqüències fonamentals i  $R_j$  és:

$$R_j = -\log\{4\sin^2(\pi j / T)\} \quad (\text{A7.3})$$

A partir d'aquests resultats, es poden calcular el biaix associat a l'estimador GPH:

$$\text{Biaix}(\tilde{d}) = d - E(\tilde{d}) \quad (\text{A7.4})$$

i l'error quadràtic mitjà:

$$\text{Biaix}(\tilde{d}) = E[(d - \tilde{d})^2] = [\text{Biaix}(\tilde{d})]^2 + \text{var}(\tilde{d}) \quad (\text{A7.5})$$

per valorar les seves propietats com estimador.

Al quadre A7.1 es presenta un resum dels models analitzats i els quadres on es poden trobar els resultats.

Quadre A7.1. Resultats presentat de la simulació<sup>21</sup>

<i>T</i>	<i>Biaix</i>	<i>EQM</i>
100	Quadre A7.2	Quadre A7.6
200	Quadre A7.3	Quadre A7.7
300	Quadre A7.4	Quadre A7.8
400	Quadre A7.5	Quadre A7.9

A l'apartat 7.5.1 de la tesi es presenten les principals conclusions que es poden treure d'aquest exercici de simulació.

<sup>21</sup> Per  $T=300$  i  $T=400$ , sols es presenten els resultats per  $d < 0.45$ .

Quadre A7.2. Biaix de l'estimador GPH en el model  $(1-\phi L)(1-L)^d X_t = (1-\theta L)\varepsilon_t$  amb  $T=100$ .

$d$	$m$	$\phi=0.0$	$\phi=0.0$	$\phi=0.9$	$\phi=0.0$	$\phi=0.0$	$\phi=0.5$	$\phi=0.9$	$\phi=0.5$	$\phi=0.9$
		$\theta=0.0$	$\theta=0.5$	$\theta=0.9$	$\theta=0.5$	$\theta=0.9$	$\theta=0.0$	$\theta=0.0$	$\theta=0.0$	$\theta=0.0$
0.00	5	0.01	0.02	0.02	-0.01	-0.27	0.05	0.53	0.01	0.02
	15	0.00	0.03	0.03	-0.16	-0.58	0.17	0.76	-0.03	-0.03
	25	0.00	0.09	0.10	-0.27	-0.68	0.28	0.83	-0.08	-0.09
	35	0.00	0.16	0.20	-0.34	-0.73	0.36	0.86	-0.16	-0.19
0.05	5	-0.02	-0.01	-0.01	-0.06	-0.34	0.02	0.50	-0.03	-0.04
	15	0.00	0.03	0.03	-0.17	-0.60	0.17	0.76	-0.03	-0.04
	25	0.00	0.08	0.09	-0.28	-0.70	0.28	0.82	-0.09	-0.10
	35	0.00	0.16	0.19	-0.35	-0.75	0.35	0.85	-0.16	-0.20
0.10	5	0.00	0.01	0.01	-0.04	-0.36	0.04	0.53	0.00	0.00
	15	0.00	0.03	0.03	-0.17	-0.62	0.17	0.76	-0.04	-0.04
	25	0.00	0.08	0.09	-0.27	-0.71	0.28	0.82	-0.09	-0.10
	35	0.00	0.16	0.20	-0.34	-0.76	0.35	0.84	-0.16	-0.19
0.15	5	0.01	0.02	0.02	-0.03	-0.37	0.04	0.52	0.00	0.00
	15	0.02	0.05	0.05	-0.15	-0.62	0.19	0.76	-0.02	-0.03
	25	0.01	0.10	0.11	-0.26	-0.71	0.29	0.82	-0.07	-0.09
	35	0.01	0.17	0.21	-0.34	-0.76	0.36	0.85	-0.15	-0.19
0.20	5	0.01	0.02	0.02	-0.02	-0.39	0.05	0.51	0.00	0.00
	15	0.01	0.04	0.04	-0.16	-0.65	0.18	0.75	-0.02	-0.04
	25	0.01	0.09	0.10	-0.27	-0.74	0.28	0.81	-0.08	-0.10
	35	0.00	0.16	0.20	-0.34	-0.78	0.35	0.83	-0.16	-0.20
0.25	5	0.06	0.07	0.07	0.02	-0.39	0.10	0.54	0.05	0.05
	15	0.02	0.06	0.06	-0.15	-0.65	0.20	0.76	-0.01	-0.02
	25	0.01	0.10	0.11	-0.26	-0.74	0.30	0.81	-0.07	-0.09
	35	0.01	0.18	0.21	-0.34	-0.78	0.37	0.83	-0.15	-0.19
0.30	5	0.02	0.03	0.03	-0.03	-0.47	0.06	0.52	0.01	0.00
	15	0.02	0.06	0.06	-0.15	-0.68	0.19	0.74	-0.01	-0.02
	25	0.02	0.10	0.11	-0.26	-0.75	0.29	0.79	-0.07	-0.09
	35	0.01	0.17	0.21	-0.34	-0.79	0.36	0.80	-0.15	-0.20
0.35	5	0.05	0.06	0.06	0.01	-0.42	0.08	0.54	0.05	0.03
	15	0.02	0.05	0.06	-0.16	-0.67	0.19	0.72	-0.02	-0.03
	25	0.02	0.10	0.12	-0.26	-0.74	0.30	0.77	-0.07	-0.09
	35	0.02	0.18	0.21	-0.33	-0.78	0.37	0.79	-0.15	-0.19
0.40	5	0.02	0.03	0.03	-0.03	-0.48	0.06	0.47	0.01	0.00
	15	0.03	0.06	0.06	-0.15	-0.67	0.19	0.67	-0.01	-0.02
	25	0.02	0.11	0.12	-0.26	-0.74	0.30	0.72	-0.06	-0.09
	35	0.02	0.18	0.21	-0.34	-0.78	0.36	0.73	-0.15	-0.19
0.45	5	0.05	0.05	0.05	-0.01	-0.45	0.08	0.43	0.04	0.03
	15	0.05	0.08	0.08	-0.14	-0.64	0.21	0.60	0.00	-0.01
	25	0.04	0.12	0.13	-0.25	-0.70	0.31	0.64	-0.05	-0.08
	35	0.02	0.18	0.21	-0.23	-0.69	0.30	0.60	-0.15	-0.19

Quadre A7.2 (Continuació). Biaix de l'estimador GPH en el model  $(1-\phi L)(1-L)^d X_t = (1-\theta L)\varepsilon_t$  amb  $T=100$ .

$d$	$m$	$\phi=0.0$ $\theta=0.0$	$\phi=0.0$ $\theta=0.5$	$\phi=0.9$ $\theta=0.9$	$\phi=0.0$ $\theta=0.5$	$\phi=0.0$ $\theta=0.9$	$\phi=0.5$ $\theta=0.0$	$\phi=0.9$ $\theta=0.0$	$\phi=0.5$ $\theta=0.0$	$\phi=0.9$ $\theta=0.0$
0.55	5	0.00	0.01	0.01	-0.04	-0.57	0.04	0.45	0.00	-0.01
	15	0.00	0.04	0.04	-0.18	-0.76	0.18	0.66	-0.03	-0.04
	25	0.00	0.09	0.10	-0.29	-0.82	0.29	0.69	-0.09	-0.11
	35	0.00	0.16	0.19	-0.36	-0.85	0.35	0.70	-0.16	-0.21
0.60	5	0.03	0.03	0.03	-0.01	-0.51	0.06	0.45	0.02	0.02
	15	0.01	0.04	0.04	-0.18	-0.77	0.17	0.63	-0.03	-0.04
	25	0.00	0.09	0.10	-0.29	-0.83	0.29	0.66	-0.09	-0.11
	35	0.00	0.16	0.20	-0.37	-0.86	0.35	0.66	-0.18	-0.22
0.65	5	0.04	0.04	0.04	0.00	-0.55	0.07	0.48	0.03	0.03
	15	0.01	0.04	0.05	-0.17	-0.78	0.18	0.61	-0.02	-0.04
	25	0.01	0.09	0.10	-0.28	-0.84	0.28	0.62	-0.08	-0.11
	35	0.00	0.17	0.20	-0.36	-0.86	0.35	0.62	-0.16	-0.21
0.70	5	0.03	0.03	0.03	-0.02	-0.57	0.06	0.45	0.02	0.02
	15	0.02	0.06	0.06	-0.16	-0.76	0.19	0.57	-0.01	-0.03
	25	0.01	0.10	0.11	-0.27	-0.83	0.28	0.56	-0.08	-0.10
	35	0.01	0.17	0.20	-0.35	-0.86	0.34	0.56	-0.16	-0.21
0.75	5	0.02	0.03	0.03	-0.01	-0.52	0.05	0.43	0.02	0.02
	15	0.03	0.06	0.06	-0.15	-0.77	0.19	0.52	-0.01	-0.02
	24	0.02	0.10	0.11	-0.26	-0.83	0.27	0.52	-0.06	-0.08
	35	0.02	0.17	0.20	-0.34	-0.86	0.34	0.50	-0.15	-0.20
0.80	5	0.03	0.03	0.03	-0.01	-0.51	0.06	0.38	0.02	0.02
	15	0.02	0.05	0.05	-0.15	-0.76	0.17	0.46	-0.01	-0.02
	25	0.01	0.09	0.10	-0.26	-0.83	0.25	0.45	-0.07	-0.09
	35	0.01	0.16	0.19	-0.34	-0.86	0.31	0.44	-0.15	-0.19
0.85	5	0.00	0.01	0.01	-0.03	-0.51	0.03	0.36	0.00	0.00
	15	0.02	0.05	0.05	-0.14	-0.75	0.15	0.40	-0.01	-0.02
	25	0.02	0.09	0.10	-0.25	-0.82	0.23	0.38	-0.06	-0.08
	35	0.02	0.15	0.18	-0.32	-0.85	0.28	0.36	-0.13	-0.18
0.90	5	0.06	0.06	0.06	0.03	-0.49	0.09	0.37	0.06	0.05
	15	0.02	0.04	0.05	-0.13	-0.74	0.14	0.35	-0.01	-0.02
	25	0.02	0.08	0.09	-0.24	-0.82	0.21	0.33	-0.06	-0.08
	35	0.02	0.14	0.16	-0.31	-0.84	0.25	0.31	-0.13	-0.17
0.95	5	-0.02	-0.01	-0.01	-0.04	-0.48	0.01	0.24	-0.02	-0.02
	15	0.01	0.03	0.03	-0.12	-0.73	0.11	0.26	-0.01	-0.02
	25	0.01	0.06	0.07	-0.22	-0.81	0.17	0.24	-0.05	-0.07
	35	0.01	0.11	0.12	-0.29	-0.84	0.19	0.22	-0.12	-0.15
1.00	5	0.00	0.01	0.01	-0.02	-0.44	0.03	0.25	0.00	0.00
	15	0.00	0.02	0.02	-0.12	-0.72	0.09	0.22	-0.02	-0.03
	25	0.00	0.05	0.06	-0.21	-0.79	0.14	0.18	-0.05	-0.07
	35	0.00	0.08	0.10	-0.27	-0.82	0.15	0.16	-0.12	-0.15

Quadre A7.3. Biaix de l'estimador GPH en el model  $(1-\phi L)(1-L)^d X_t = (1-\theta L)\varepsilon_t$  amb  $T=200$ .

$d$	$m$	$\phi=0.0$ $\theta=0.0$	$\phi=0.0$ $\theta=0.5$	$\phi=0.9$ $\theta=0.9$	$\phi=0.0$ $\theta=0.5$	$\phi=0.0$ $\theta=0.9$	$\phi=0.5$ $\theta=0.0$	$\phi=0.9$ $\theta=0.0$	$\phi=0.5$ $\theta=0.0$	$\phi=0.9$ $\theta=0.0$
0.00	5	0.00	0.00	0.00	0.00	-0.13	0.01	0.25	0.00	0.01
	15	0.00	0.01	0.01	-0.05	-0.39	0.06	0.56	0.00	0.00
	25	0.00	0.01	0.02	-0.12	-0.54	0.12	0.68	-0.02	-0.03
	35	-0.01	0.03	0.04	-0.18	-0.62	0.18	0.74	-0.04	-0.05
0.05	5	-0.01	-0.01	-0.01	-0.02	-0.17	0.00	0.26	-0.01	-0.01
	15	-0.01	0.00	0.00	-0.07	-0.44	0.05	0.56	-0.01	-0.02
	25	0.00	0.02	0.02	-0.12	-0.56	0.12	0.68	-0.02	-0.03
	35	0.00	0.04	0.04	-0.18	-0.64	0.18	0.74	-0.04	-0.04
0.10	5	0.01	0.01	0.01	0.00	-0.21	0.01	0.26	0.00	0.00
	15	0.01	0.02	0.02	-0.05	-0.46	0.07	0.57	0.00	0.00
	25	0.01	0.03	0.03	-0.11	-0.59	0.13	0.68	-0.01	-0.02
	35	0.00	0.04	0.05	-0.18	-0.66	0.19	0.74	-0.03	-0.04
0.15	5	0.02	0.02	0.02	0.00	-0.21	0.02	0.27	0.00	0.01
	15	0.00	0.01	0.01	-0.06	-0.48	0.06	0.56	-0.01	-0.01
	25	0.00	0.02	0.02	-0.12	-0.60	0.13	0.68	-0.02	-0.02
	35	0.00	0.04	0.04	-0.18	-0.67	0.18	0.74	-0.04	-0.05
0.20	5	0.04	0.04	0.04	0.02	-0.21	0.05	0.29	0.03	0.02
	15	0.01	0.02	0.02	-0.05	-0.50	0.07	0.57	0.00	-0.01
	25	0.01	0.03	0.03	-0.12	-0.62	0.13	0.69	-0.01	-0.02
	35	0.01	0.05	0.05	-0.18	-0.68	0.19	0.74	-0.03	-0.04
0.25	5	0.03	0.03	0.03	0.02	-0.19	0.03	0.26	0.02	0.02
	15	0.01	0.02	0.02	-0.05	-0.49	0.07	0.57	0.00	0.00
	25	0.01	0.03	0.03	-0.11	-0.62	0.13	0.68	-0.01	-0.02
	35	0.01	0.05	0.05	-0.17	-0.68	0.19	0.74	-0.03	-0.04
0.30	5	0.02	0.02	0.02	0.01	-0.23	0.03	0.27	0.02	0.02
	15	0.02	0.03	0.03	-0.04	-0.49	0.08	0.57	0.01	0.01
	25	0.02	0.04	0.04	-0.11	-0.61	0.14	0.67	-0.01	-0.01
	35	0.01	0.05	0.06	-0.17	-0.68	0.20	0.73	-0.03	-0.04
0.35	5	0.03	0.03	0.03	0.02	-0.24	0.03	0.26	0.03	0.03
	15	0.03	0.04	0.04	-0.03	-0.50	0.09	0.56	0.02	0.02
	25	0.02	0.04	0.04	-0.10	-0.62	0.14	0.66	0.00	-0.01
	35	0.02	0.06	0.06	-0.16	-0.69	0.20	0.71	-0.02	-0.03
0.40	5	0.04	0.04	0.04	0.03	-0.23	0.05	0.27	0.04	0.03
	15	0.04	0.05	0.05	-0.02	-0.50	0.10	0.55	0.03	0.03
	25	0.02	0.04	0.05	-0.10	-0.61	0.15	0.64	0.00	-0.01
	35	0.02	0.06	0.07	-0.16	-0.67	0.21	0.69	-0.02	-0.03
0.45	5	0.08	0.08	0.08	0.06	-0.25	0.09	0.26	0.08	0.07
	15	0.04	0.05	0.05	-0.03	-0.49	0.10	0.51	0.03	0.02
	25	0.02	0.04	0.05	-0.10	-0.61	0.14	0.59	0.00	-0.01
	35	0.02	0.06	0.07	-0.16	-0.67	0.21	0.68	-0.02	-0.03

Quadre A7.3 (Continuació). Biaix de l'estimador GPH en el model  $(1-\phi L)(1-L)^d X_t = (1-\theta L)\varepsilon_t$  amb  $T=200$ .

$d$	$m$	$\phi=0.0$ $\theta=0.0$	$\phi=0.0$ $\theta=0.5$	$\phi=0.9$ $\theta=0.9$	$\phi=0.0$ $\theta=0.5$	$\phi=0.0$ $\theta=0.9$	$\phi=0.5$ $\theta=0.0$	$\phi=0.9$ $\theta=0.0$	$\phi=0.5$ $\theta=0.0$	$\phi=0.9$ $\theta=0.0$
0.55	5	0.02	0.02	0.02	0.01	-0.27	0.03	0.26	0.02	0.02
	15	0.00	0.01	0.01	-0.06	-0.58	0.06	0.53	-0.01	-0.01
	25	-0.01	0.02	0.02	-0.13	-0.70	0.12	0.62	-0.03	-0.04
	34	0.00	0.04	0.04	-0.18	-0.74	0.18	0.66	-0.04	-0.05
0.60	5	0.02	0.02	0.02	0.01	-0.28	0.03	0.25	0.02	0.02
	15	0.01	0.02	0.02	-0.05	-0.57	0.07	0.52	0.00	0.00
	25	0.00	0.02	0.03	-0.12	-0.69	0.13	0.61	-0.02	-0.02
	35	0.01	0.05	0.05	-0.18	-0.74	0.19	0.64	-0.03	-0.04
0.65	5	0.00	0.00	0.00	-0.01	-0.28	0.01	0.24	0.00	0.00
	15	0.01	0.02	0.02	-0.05	-0.57	0.07	0.51	0.01	0.00
	25	0.00	0.02	0.03	-0.12	-0.69	0.13	0.58	-0.02	-0.02
	35	0.01	0.05	0.05	-0.18	-0.75	0.19	0.61	-0.03	-0.04
0.70	5	0.01	0.02	0.02	0.00	-0.27	0.03	0.25	0.01	0.01
	15	0.02	0.03	0.03	-0.04	-0.56	0.08	0.49	0.01	0.01
	25	0.02	0.04	0.04	-0.11	-0.68	0.14	0.55	-0.01	-0.01
	35	0.01	0.05	0.06	-0.17	-0.74	0.19	0.57	-0.03	-0.04
0.75	5	0.01	0.02	0.02	0.00	-0.26	0.02	0.24	0.01	0.01
	15	0.02	0.03	0.03	-0.04	-0.56	0.07	0.46	0.01	0.00
	24	0.02	0.03	0.03	-0.10	-0.67	0.12	0.50	0.00	-0.01
	35	0.02	0.06	0.06	-0.17	-0.74	0.19	0.52	-0.02	-0.03
0.80	5	0.06	0.06	0.06	0.05	-0.20	0.06	0.25	0.05	0.05
	15	0.02	0.03	0.03	-0.03	-0.55	0.07	0.41	0.01	0.01
	25	0.03	0.04	0.05	-0.09	-0.66	0.13	0.45	0.01	0.00
	35	0.02	0.06	0.06	-0.16	-0.73	0.18	0.45	-0.02	-0.02
0.85	5	0.04	0.04	0.04	0.04	-0.20	0.05	0.21	0.04	0.04
	15	0.03	0.04	0.04	-0.02	-0.53	0.08	0.37	0.02	0.02
	25	0.03	0.04	0.05	-0.09	-0.65	0.13	0.39	0.01	0.00
	35	0.02	0.05	0.06	-0.15	-0.72	0.17	0.39	-0.01	-0.02
0.90	5	-0.01	-0.01	-0.01	-0.02	-0.25	0.00	0.17	-0.01	-0.02
	15	0.01	0.02	0.02	-0.03	-0.52	0.06	0.31	0.01	0.01
	25	0.02	0.03	0.04	-0.09	-0.65	0.11	0.32	0.00	0.00
	35	0.02	0.05	0.05	-0.14	-0.71	0.15	0.32	-0.01	-0.02
0.95	5	0.02	0.02	0.02	0.01	-0.20	0.02	0.19	0.02	0.02
	15	0.00	0.01	0.01	-0.04	-0.51	0.04	0.27	-0.01	-0.01
	25	0.01	0.02	0.02	-0.09	-0.63	0.09	0.27	-0.01	-0.01
	35	0.01	0.04	0.04	-0.14	-0.70	0.12	0.26	-0.02	-0.03
1.00	5	-0.01	-0.01	-0.01	-0.02	-0.18	-0.01	0.12	-0.01	-0.01
	15	0.00	0.00	0.00	-0.04	-0.48	0.03	0.20	0.00	0.00
	25	0.00	0.01	0.01	-0.08	-0.61	0.06	0.19	-0.01	-0.02
	35	-0.01	0.01	0.02	-0.13	-0.68	0.09	0.18	-0.03	-0.03

Quadre A7.4. Biaix de l'estimador GPH en el model  $(1-\phi L)(1-L)^d X_t = (1-\theta L)\varepsilon_t$  amb  $T=300$ .

$d$	$m$	$\phi=0.0$	$\phi=0.0$	$\phi=0.9$	$\phi=0.0$	$\phi=0.0$	$\phi=0.5$	$\phi=0.9$	$\phi=-0.5$	$\phi=-0.9$
		$\theta=0.0$	$\theta=0.5$	$\theta=0.9$	$\theta=0.5$	$\theta=0.9$	$\theta=0.0$	$\theta=0.0$	$\theta=0.0$	$\theta=0.0$
0.00	5	0.00	0.00	0.00	-0.01	-0.08	0.00	0.14	0.00	0.00
	15	-0.01	-0.01	0.00	-0.04	-0.32	0.02	0.42	-0.01	-0.02
	25	-0.01	0.00	0.00	-0.07	-0.45	0.06	0.56	-0.02	-0.02
	35	0.00	0.01	0.02	-0.11	-0.54	0.10	0.64	-0.02	-0.02
0.05	5	-0.01	-0.01	-0.01	-0.02	-0.12	-0.01	0.13	-0.02	-0.02
	15	0.00	0.00	0.00	-0.02	-0.32	0.03	0.43	0.00	0.00
	25	0.00	0.01	0.01	-0.06	-0.46	0.07	0.56	-0.01	-0.01
	35	0.00	0.02	0.02	-0.10	-0.55	0.11	0.64	-0.02	-0.02
0.10	5	0.01	0.01	0.01	0.01	-0.12	0.02	0.15	0.01	0.01
	15	0.00	0.01	0.01	-0.02	-0.35	0.03	0.43	0.00	0.00
	25	-0.01	0.00	0.00	-0.07	-0.49	0.06	0.56	-0.02	-0.02
	35	-0.01	0.01	0.01	-0.11	-0.58	0.10	0.63	-0.03	-0.03
0.15	5	0.01	0.01	0.01	0.01	-0.12	0.02	0.15	0.01	0.00
	15	0.01	0.02	0.02	-0.02	-0.38	0.04	0.44	0.01	0.01
	25	0.01	0.02	0.02	-0.06	-0.50	0.08	0.57	0.00	0.00
	35	0.01	0.03	0.03	-0.10	-0.58	0.12	0.65	-0.01	-0.01
0.20	5	0.02	0.02	0.02	0.01	-0.13	0.02	0.16	0.02	0.01
	15	0.01	0.02	0.02	-0.02	-0.38	0.05	0.44	0.01	0.01
	25	0.01	0.02	0.02	-0.05	-0.51	0.08	0.57	0.00	0.00
	35	0.01	0.02	0.03	-0.10	-0.59	0.11	0.65	-0.01	-0.02
0.25	5	0.00	0.00	0.00	-0.01	-0.17	0.00	0.13	0.00	-0.01
	15	0.02	0.02	0.02	-0.01	-0.37	0.05	0.44	0.02	0.02
	25	0.01	0.02	0.02	-0.05	-0.51	0.08	0.57	0.00	0.00
	35	0.01	0.03	0.03	-0.10	-0.59	0.12	0.64	-0.01	-0.01
0.30	5	0.03	0.03	0.04	0.02	-0.14	0.04	0.18	0.03	0.03
	15	0.03	0.03	0.03	0.00	-0.38	0.05	0.44	0.02	0.02
	25	0.01	0.02	0.03	-0.05	-0.52	0.08	0.57	0.00	0.00
	35	0.02	0.04	0.04	-0.09	-0.59	0.13	0.65	0.00	0.00
0.35	5	0.03	0.03	0.03	0.03	-0.13	0.04	0.17	0.03	0.03
	15	0.03	0.03	0.03	-0.01	-0.39	0.05	0.43	0.02	0.02
	25	0.02	0.02	0.03	-0.05	-0.52	0.08	0.55	0.01	0.00
	35	0.02	0.04	0.04	-0.09	-0.59	0.13	0.63	0.00	0.00
0.40	5	0.07	0.07	0.07	0.06	-0.12	0.07	0.20	0.06	0.06
	15	0.04	0.04	0.04	0.01	-0.38	0.07	0.43	0.03	0.03
	25	0.04	0.05	0.05	-0.03	-0.50	0.10	0.56	0.03	0.02
	35	0.03	0.05	0.05	-0.08	-0.58	0.14	0.62	0.01	0.01
0.45	5	0.09	0.09	0.09	0.08	-0.11	0.10	0.21	0.09	0.09
	15	0.05	0.05	0.05	0.02	-0.38	0.08	0.41	0.05	0.04
	25	0.04	0.05	0.05	-0.03	-0.50	0.10	0.52	0.03	0.02
	35	0.03	0.05	0.09	-0.08	-0.58	0.14	0.60	0.01	0.01

Quadre A7.5. Biaix de l'estimador GPH en el model  $(1-\phi L)(1-L)^d X_t = (1-\theta L)\varepsilon_t$  amb  $T=400$ .

$d$	$m$	$\phi=0.0$	$\phi=0.0$	$\phi=0.9$	$\phi=0.0$	$\phi=0.0$	$\phi=0.5$	$\phi=0.9$	$\phi=0.5$	$\phi=0.9$
		$\theta=0.0$	$\theta=0.5$	$\theta=0.9$	$\theta=0.5$	$\theta=0.9$	$\theta=0.0$	$\theta=0.0$	$\theta=0.0$	$\theta=0.0$
0.00	5	-0.01	-0.01	-0.01	-0.01	-0.07	-0.01	0.08	-0.01	-0.01
	15	0.01	0.01	0.01	-0.01	-0.24	0.02	0.33	0.01	0.00
	25	0.00	0.01	0.01	-0.04	-0.37	0.04	0.47	0.00	0.00
	35	0.00	0.01	0.01	-0.07	-0.46	0.07	0.56	-0.01	-0.01
0.05	5	0.02	0.02	0.01	0.01	-0.08	0.02	0.10	0.01	0.01
	15	0.01	0.01	0.01	-0.01	-0.25	0.02	0.33	0.00	0.00
	25	0.00	0.01	0.01	-0.04	-0.39	0.04	0.47	0.00	0.00
	35	0.00	0.01	0.01	-0.07	-0.48	0.07	0.56	-0.01	-0.01
0.10	5	0.02	0.02	0.02	0.01	-0.07	0.02	0.10	0.01	0.01
	15	-0.01	0.00	0.00	-0.02	-0.29	0.01	0.32	-0.01	-0.01
	25	0.00	0.00	0.00	-0.04	-0.42	0.04	0.47	-0.01	-0.01
	35	0.00	0.01	0.01	-0.07	-0.50	0.07	0.56	-0.01	-0.01
0.15	5	0.03	0.03	0.03	0.02	-0.08	0.03	0.11	0.02	0.02
	15	0.01	0.01	0.01	-0.01	-0.29	0.03	0.34	0.01	0.01
	25	0.01	0.01	0.01	-0.03	-0.42	0.05	0.48	0.00	0.00
	35	0.01	0.02	0.02	-0.06	-0.51	0.08	0.56	0.00	-0.01
0.20	5	0.04	0.04	0.04	0.03	-0.08	0.04	0.12	0.04	0.03
	15	0.01	0.01	0.01	0.00	-0.30	0.03	0.34	0.01	0.01
	25	0.01	0.01	0.01	-0.03	-0.44	0.05	0.47	0.00	0.00
	35	0.01	0.02	0.02	-0.06	-0.51	0.08	0.57	0.00	0.00
0.25	5	0.02	0.02	0.01	0.01	-0.10	0.02	0.10	0.01	0.01
	15	0.03	0.04	0.04	0.01	-0.29	0.05	0.35	0.03	0.03
	25	0.02	0.03	0.03	-0.02	-0.43	0.06	0.49	0.01	0.01
	35	0.02	0.03	0.03	-0.05	-0.51	0.09	0.57	0.01	0.01
0.30	5	0.04	0.04	0.04	0.04	-0.08	0.05	0.14	0.04	0.04
	15	0.02	0.02	0.03	0.00	-0.30	0.04	0.34	0.02	0.02
	25	0.01	0.02	0.02	-0.03	-0.44	0.05	0.47	0.01	0.01
	35	0.01	0.02	0.03	-0.05	-0.52	0.08	0.56	0.00	0.00
0.35	5	0.07	0.07	0.07	0.07	-0.02	0.08	0.16	0.07	0.07
	15	0.02	0.02	0.02	0.01	-0.30	0.04	0.34	0.02	0.02
	25	0.02	0.03	0.03	-0.02	-0.43	0.06	0.48	0.02	0.02
	35	0.02	0.03	0.03	-0.05	-0.52	0.09	0.56	0.01	0.01
0.40	5	0.02	0.03	0.03	0.02	-0.08	0.03	0.11	0.02	0.02
	15	0.03	0.03	0.03	0.01	-0.30	0.04	0.34	0.02	0.02
	25	0.03	0.04	0.04	-0.01	-0.43	0.07	0.47	0.03	0.02
	35	0.03	0.04	0.04	-0.04	-0.52	0.09	0.56	0.02	0.01
0.45	5	0.04	0.04	0.04	0.04	-0.09	0.05	0.12	0.04	0.04
	15	0.05	0.05	0.05	0.02	-0.30	0.06	0.33	0.04	0.04
	25	0.03	0.04	0.04	-0.01	-0.43	0.07	0.45	0.03	0.03
	35	0.03	0.04	0.04	-0.01	-0.52	0.09	0.56	0.02	0.01



Quadre A7.6. EQM de l'estimador GPH en el model  $(1-\phi L)(1-L)^d X_t = (1-\theta L)\varepsilon_t$  amb  $T=100$ .

$d$	$m$	$\phi=0.0$ $\theta=0.0$	$\phi=0.0$ $\theta=0.5$	$\phi=0.9$ $\theta=0.9$	$\phi=0.0$ $\theta=0.5$	$\phi=0.0$ $\theta=0.9$	$\phi=0.5$ $\theta=0.0$	$\phi=0.9$ $\theta=0.0$	$\phi=0.5$ $\theta=0.0$	$\phi=0.9$ $\theta=0.0$
0.00	5	0.39	0.39	0.39	0.38	0.42	0.39	0.66	0.40	0.37
	15	0.06	0.06	0.06	0.08	0.41	0.09	0.64	0.06	0.06
	25	0.03	0.04	0.04	0.10	0.51	0.11	0.72	0.03	0.04
	35	0.02	0.05	0.06	0.14	0.57	0.15	0.75	0.04	0.06
0.05	5	0.36	0.36	0.37	0.37	0.43	0.36	0.61	0.37	0.38
	15	0.06	0.06	0.06	0.08	0.44	0.08	0.64	0.06	0.06
	25	0.03	0.03	0.04	0.11	0.54	0.10	0.71	0.04	0.04
	35	0.02	0.04	0.06	0.14	0.60	0.14	0.74	0.05	0.06
0.10	5	0.42	0.41	0.41	0.40	0.46	0.41	0.68	0.41	0.41
	15	0.06	0.06	0.06	0.09	0.46	0.09	0.64	0.06	0.06
	25	0.03	0.04	0.04	0.11	0.56	0.11	0.70	0.04	0.04
	35	0.02	0.05	0.06	0.14	0.61	0.14	0.73	0.05	0.06
0.15	5	0.37	0.38	0.38	0.36	0.43	0.38	0.67	0.36	0.37
	15	0.05	0.05	0.06	0.07	0.46	0.09	0.63	0.05	0.05
	25	0.03	0.04	0.04	0.10	0.55	0.11	0.71	0.03	0.04
	35	0.02	0.05	0.06	0.13	0.61	0.15	0.73	0.04	0.06
0.20	5	0.44	0.45	0.45	0.41	0.54	0.44	0.68	0.43	0.43
	15	0.06	0.06	0.06	0.08	0.49	0.09	0.61	0.06	0.06
	25	0.03	0.04	0.04	0.10	0.58	0.11	0.68	0.03	0.04
	35	0.02	0.04	0.06	0.14	0.64	0.14	0.71	0.04	0.06
0.25	5	0.38	0.37	0.37	0.37	0.48	0.38	0.70	0.38	0.37
	15	0.05	0.05	0.05	0.07	0.49	0.09	0.63	0.05	0.05
	25	0.03	0.04	0.04	0.10	0.58	0.11	0.68	0.03	0.04
	35	0.02	0.05	0.06	0.13	0.64	0.15	0.71	0.04	0.05
0.30	5	0.34	0.35	0.35	0.35	0.54	0.35	0.61	0.35	0.34
	15	0.06	0.06	0.06	0.08	0.52	0.09	0.60	0.06	0.05
	25	0.03	0.04	0.04	0.10	0.61	0.11	0.65	0.03	0.04
	35	0.02	0.05	0.06	0.14	0.66	0.15	0.67	0.04	0.06
0.35	5	0.40	0.40	0.40	0.40	0.52	0.40	0.68	0.41	0.40
	15	0.06	0.06	0.06	0.08	0.51	0.10	0.59	0.06	0.06
	25	0.03	0.04	0.04	0.10	0.58	0.12	0.62	0.03	0.04
	35	0.02	0.05	0.06	0.13	0.63	0.15	0.64	0.04	0.06
0.40	5	0.40	0.39	0.39	0.40	0.61	0.39	0.61	0.40	0.39
	15	0.06	0.06	0.06	0.08	0.52	0.10	0.51	0.06	0.06
	25	0.03	0.04	0.04	0.10	0.58	0.12	0.56	0.03	0.04
	35	0.02	0.05	0.06	0.13	0.64	0.15	0.57	0.04	0.06
0.45	5	0.36	0.35	0.35	0.37	0.49	0.35	0.53	0.37	0.37
	15	0.06	0.06	0.06	0.07	0.46	0.10	0.44	0.05	0.05
	25	0.03	0.04	0.05	0.09	0.52	0.12	0.46	0.03	0.03
	35	0.02	0.06	0.07	0.12	0.57	0.16	0.48	0.04	0.05

Quadre A7.6 (Continuació). EQM de l'estimador GPH en el model  $(1-\phi L)(1-L)^d X_t = (1-\theta L)\varepsilon_t$  amb  $T=100$ .

$d$	$m$	$\phi=0.0$	$\phi=0.0$	$\phi=0.9$	$\phi=0.0$	$\phi=0.0$	$\phi=0.5$	$\phi=0.9$	$\phi=0.5$	$\phi=0.9$
		$\theta=0.0$	$\theta=0.5$	$\theta=0.9$	$\theta=0.5$	$\theta=0.9$	$\theta=0.0$	$\theta=0.0$	$\theta=0.0$	$\theta=0.0$
0.55	5	0.39	0.38	0.38	0.39	0.71	0.39	0.56	0.39	0.38
	15	0.06	0.06	0.06	0.09	0.64	0.09	0.49	0.06	0.06
	25	0.03	0.04	0.04	0.11	0.70	0.11	0.51	0.04	0.04
	35	0.02	0.05	0.06	0.15	0.74	0.15	0.51	0.05	0.07
0.60	5	0.41	0.40	0.40	0.40	0.63	0.40	0.58	0.40	0.40
	15	0.05	0.05	0.05	0.08	0.64	0.08	0.46	0.05	0.05
	25	0.03	0.04	0.04	0.11	0.72	0.11	0.47	0.03	0.04
	35	0.02	0.04	0.06	0.15	0.77	0.14	0.47	0.05	0.07
0.65	5	0.38	0.38	0.38	0.40	0.71	0.39	0.60	0.39	0.40
	15	0.05	0.05	0.05	0.08	0.66	0.08	0.43	0.05	0.05
	25	0.03	0.03	0.04	0.11	0.73	0.11	0.42	0.03	0.04
	35	0.02	0.05	0.06	0.15	0.76	0.14	0.42	0.04	0.06
0.70	5	0.35	0.35	0.35	0.35	0.66	0.35	0.54	0.35	0.36
	15	0.05	0.06	0.06	0.08	0.64	0.09	0.38	0.05	0.05
	25	0.03	0.04	0.04	0.10	0.72	0.11	0.37	0.03	0.04
	35	0.02	0.04	0.06	0.14	0.76	0.13	0.36	0.04	0.06
0.75	5	0.39	0.40	0.39	0.40	0.67	0.39	0.54	0.39	0.39
	15	0.05	0.05	0.05	0.07	0.64	0.08	0.34	0.05	0.05
	24	0.03	0.04	0.04	0.10	0.72	0.10	0.32	0.03	0.04
	35	0.02	0.05	0.06	0.14	0.76	0.13	0.30	0.04	0.06
0.80	5	0.36	0.36	0.37	0.36	0.63	0.37	0.47	0.36	0.36
	15	0.05	0.06	0.06	0.08	0.64	0.08	0.28	0.05	0.05
	25	0.03	0.04	0.04	0.10	0.72	0.09	0.26	0.03	0.04
	35	0.02	0.04	0.05	0.13	0.76	0.11	0.25	0.04	0.06
0.85	5	0.35	0.35	0.35	0.35	0.64	0.35	0.49	0.35	0.35
	15	0.05	0.06	0.06	0.07	0.62	0.08	0.23	0.05	0.05
	25	0.03	0.04	0.04	0.09	0.70	0.08	0.21	0.03	0.04
	35	0.02	0.04	0.05	0.13	0.74	0.10	0.19	0.04	0.06
0.90	5	0.38	0.38	0.38	0.39	0.63	0.38	0.48	0.39	0.39
	15	0.05	0.06	0.06	0.08	0.61	0.07	0.20	0.06	0.06
	25	0.03	0.04	0.04	0.09	0.70	0.08	0.17	0.04	0.04
	35	0.02	0.04	0.05	0.12	0.73	0.09	0.15	0.04	0.05
0.95	5	0.33	0.33	0.33	0.32	0.60	0.34	0.37	0.32	0.32
	15	0.05	0.05	0.05	0.07	0.60	0.07	0.15	0.05	0.05
	25	0.03	0.03	0.03	0.08	0.68	0.06	0.12	0.03	0.03
	35	0.02	0.03	0.04	0.11	0.73	0.06	0.11	0.04	0.05
1.00	5	0.29	0.29	0.29	0.30	0.55	0.30	0.34	0.29	0.29
	15	0.05	0.05	0.05	0.07	0.58	0.06	0.12	0.05	0.05
	25	0.02	0.03	0.03	0.07	0.65	0.05	0.09	0.03	0.03
	35	0.02	0.03	0.03	0.10	0.70	0.05	0.08	0.03	0.05

Quadre A7.7. EQM de l'estimador GPH en el model  $(1-\phi L)(1-L)^d X_t = (1-\theta L)\varepsilon_t$  amb  $T=200$ .

$d$	$m$	$\phi=0.0$	$\phi=0.0$	$\phi=0.9$	$\phi=0.0$	$\phi=0.0$	$\phi=0.5$	$\phi=0.9$	$\phi=0.5$	$\phi=0.9$
		$\theta=0.0$	$\theta=0.5$	$\theta=0.9$	$\theta=0.5$	$\theta=0.9$	$\theta=0.0$	$\theta=0.0$	$\theta=0.0$	$\theta=0.0$
0.00	5	0.37	0.36	0.37	0.39	0.35	0.37	0.44	0.38	0.37
	15	0.06	0.06	0.06	0.06	0.23	0.06	0.37	0.06	0.06
	25	0.02	0.03	0.03	0.04	0.33	0.04	0.48	0.03	0.03
	35	0.02	0.02	0.02	0.05	0.42	0.05	0.56	0.02	0.02
0.05	5	0.38	0.39	0.39	0.37	0.32	0.39	0.43	0.38	0.38
	15	0.05	0.05	0.05	0.05	0.25	0.05	0.36	0.05	0.05
	25	0.03	0.03	0.03	0.04	0.36	0.04	0.49	0.03	0.03
	35	0.02	0.02	0.02	0.05	0.44	0.05	0.56	0.02	0.02
0.10	5	0.38	0.38	0.38	0.38	0.39	0.38	0.45	0.38	0.38
	15	0.05	0.05	0.05	0.05	0.27	0.05	0.37	0.05	0.05
	25	0.02	0.02	0.03	0.04	0.38	0.04	0.49	0.03	0.03
	35	0.02	0.02	0.02	0.05	0.46	0.05	0.56	0.02	0.02
0.15	5	0.41	0.39	0.39	0.39	0.39	0.39	0.46	0.39	0.38
	15	0.06	0.06	0.06	0.06	0.29	0.06	0.37	0.06	0.05
	25	0.03	0.03	0.03	0.04	0.39	0.04	0.49	0.03	0.03
	35	0.02	0.02	0.02	0.05	0.48	0.05	0.56	0.02	0.02
0.20	5	0.37	0.37	0.36	0.40	0.41	0.37	0.45	0.38	0.39
	15	0.06	0.06	0.06	0.06	0.31	0.06	0.38	0.06	0.06
	25	0.03	0.03	0.03	0.04	0.41	0.04	0.50	0.03	0.03
	35	0.02	0.02	0.02	0.05	0.49	0.05	0.57	0.02	0.02
0.25	5	0.38	0.38	0.38	0.38	0.39	0.39	0.47	0.38	0.38
	15	0.05	0.05	0.05	0.05	0.30	0.06	0.37	0.05	0.05
	25	0.02	0.03	0.03	0.04	0.41	0.04	0.49	0.02	0.02
	35	0.02	0.02	0.02	0.05	0.49	0.06	0.56	0.02	0.02
0.30	5	0.35	0.34	0.34	0.37	0.38	0.34	0.42	0.36	0.36
	15	0.06	0.06	0.06	0.06	0.30	0.07	0.39	0.06	0.06
	25	0.03	0.03	0.03	0.04	0.41	0.05	0.48	0.03	0.03
	35	0.02	0.02	0.02	0.05	0.48	0.06	0.55	0.02	0.02
0.35	5	0.39	0.39	0.39	0.39	0.43	0.39	0.43	0.39	0.39
	15	0.06	0.06	0.06	0.06	0.31	0.07	0.37	0.06	0.06
	25	0.03	0.03	0.03	0.04	0.42	0.05	0.46	0.03	0.03
	35	0.02	0.02	0.02	0.05	0.49	0.06	0.53	0.02	0.02
0.40	5	0.39	0.38	0.38	0.39	0.40	0.38	0.48	0.39	0.39
	15	0.06	0.06	0.06	0.05	0.30	0.06	0.36	0.05	0.05
	25	0.03	0.03	0.03	0.04	0.41	0.05	0.44	0.03	0.03
	35	0.02	0.02	0.02	0.04	0.48	0.06	0.50	0.02	0.02
0.45	5	0.39	0.40	0.40	0.38	0.38	0.39	0.43	0.39	0.39
	15	0.05	0.05	0.05	0.05	0.29	0.06	0.33	0.05	0.05
	25	0.03	0.03	0.03	0.04	0.40	0.05	0.39	0.03	0.03
	35	0.02	0.02	0.02	0.04	0.46	0.06	0.44	0.02	0.02

Quadre A7.7 (Continuació). EQM de l'estimador GPH en el model  $(1-\phi L)(1-L)^d X_t = (1-\theta L)\varepsilon_t$  amb  $T=200$ .

$d$	$m$	$\phi=0.0$ $\theta=0.0$	$\phi=0.0$ $\theta=0.5$	$\phi=0.9$ $\theta=0.9$	$\phi=0.0$ $\theta=0.5$	$\phi=0.0$ $\theta=0.9$	$\phi=0.5$ $\theta=0.0$	$\phi=0.9$ $\theta=0.0$	$\phi=0.5$ $\theta=0.0$	$\phi=0.9$ $\theta=0.0$
0.55	5	0.40	0.40	0.40	0.40	0.46	0.40	0.47	0.41	0.40
	15	0.05	0.05	0.05	0.06	0.39	0.06	0.33	0.05	0.05
	25	0.03	0.03	0.03	0.04	0.51	0.04	0.41	0.03	0.03
	35	0.02	0.02	0.02	0.05	0.57	0.05	0.47	0.02	0.02
0.60	5	0.41	0.41	0.41	0.41	0.47	0.41	0.45	0.41	0.41
	15	0.05	0.05	0.05	0.06	0.38	0.06	0.33	0.05	0.05
	25	0.03	0.03	0.03	0.04	0.50	0.04	0.40	0.03	0.03
	35	0.02	0.02	0.02	0.05	0.57	0.05	0.44	0.02	0.02
0.65	5	0.39	0.39	0.39	0.39	0.47	0.39	0.45	0.39	0.39
	15	0.06	0.06	0.06	0.06	0.39	0.06	0.31	0.06	0.06
	25	0.03	0.03	0.03	0.04	0.51	0.04	0.37	0.03	0.03
	35	0.02	0.02	0.02	0.05	0.58	0.05	0.40	0.02	0.02
0.70	5	0.33	0.33	0.33	0.34	0.41	0.33	0.40	0.33	0.33
	15	0.06	0.06	0.06	0.06	0.37	0.06	0.29	0.06	0.06
	25	0.03	0.03	0.03	0.04	0.49	0.05	0.34	0.03	0.03
	35	0.02	0.02	0.02	0.05	0.57	0.05	0.36	0.02	0.02
0.75	5	0.37	0.37	0.37	0.37	0.44	0.37	0.42	0.37	0.38
	15	0.06	0.06	0.06	0.06	0.37	0.06	0.27	0.06	0.06
	24	0.03	0.03	0.03	0.04	0.49	0.04	0.30	0.03	0.03
	35	0.02	0.02	0.02	0.05	0.56	0.05	0.31	0.02	0.02
0.80	5	0.39	0.39	0.39	0.39	0.43	0.38	0.42	0.39	0.39
	15	0.05	0.05	0.05	0.06	0.36	0.06	0.23	0.05	0.05
	25	0.03	0.03	0.03	0.04	0.47	0.04	0.25	0.03	0.03
	35	0.02	0.02	0.02	0.04	0.55	0.05	0.25	0.02	0.02
0.85	5	0.37	0.38	0.38	0.38	0.40	0.39	0.40	0.37	0.37
	15	0.05	0.05	0.05	0.05	0.33	0.06	0.20	0.05	0.05
	25	0.03	0.03	0.03	0.03	0.45	0.04	0.20	0.02	0.02
	35	0.02	0.02	0.02	0.04	0.54	0.04	0.20	0.02	0.02
0.90	5	0.37	0.37	0.37	0.36	0.40	0.37	0.36	0.36	0.36
	15	0.05	0.05	0.05	0.05	0.32	0.05	0.16	0.05	0.05
	25	0.02	0.03	0.02	0.03	0.45	0.03	0.16	0.02	0.02
	35	0.02	0.02	0.02	0.04	0.53	0.04	0.16	0.02	0.02
0.95	5	0.34	0.34	0.34	0.35	0.42	0.33	0.38	0.34	0.34
	15	0.05	0.05	0.05	0.05	0.32	0.05	0.14	0.05	0.05
	25	0.03	0.02	0.02	0.04	0.44	0.03	0.13	0.03	0.03
	35	0.02	0.02	0.02	0.04	0.51	0.03	0.12	0.02	0.02
1.00	5	0.31	0.32	0.32	0.31	0.40	0.33	0.33	0.31	0.31
	15	0.05	0.05	0.05	0.05	0.29	0.05	0.11	0.05	0.05
	25	0.02	0.02	0.02	0.03	0.41	0.03	0.09	0.02	0.02
	35	0.02	0.02	0.02	0.04	0.49	0.03	0.09	0.02	0.02

Quadre A7.8. EQM de l'estimador GPH en el model  $(1-\phi L)(1-L)^d X_t = (1-\theta L)\varepsilon_t$  amb  $T=300$ .

$d$	$m$	$\phi=0.0$ $\theta=0.0$	$\phi=0.0$ $\theta=0.5$	$\phi=0.9$ $\theta=0.9$	$\phi=0.0$ $\theta=0.5$	$\phi=0.0$ $\theta=0.9$	$\phi=0.5$ $\theta=0.0$	$\phi=0.9$ $\theta=0.0$	$\phi=0.5$ $\theta=0.0$	$\phi=0.9$ $\theta=0.0$
0.00	5	0.37	0.36	0.36	0.37	0.30	0.36	0.38	0.37	0.37
	15	0.06	0.05	0.05	0.06	0.16	0.05	0.23	0.06	0.06
	25	0.03	0.03	0.03	0.03	0.24	0.03	0.34	0.03	0.03
	35	0.02	0.02	0.02	0.03	0.32	0.03	0.43	0.02	0.02
0.05	5	0.41	0.41	0.42	0.40	0.39	0.41	0.41	0.41	0.42
	15	0.05	0.05	0.05	0.05	0.16	0.05	0.23	0.05	0.05
	25	0.03	0.03	0.03	0.03	0.25	0.03	0.35	0.03	0.03
	35	0.02	0.02	0.02	0.03	0.33	0.03	0.43	0.02	0.02
0.10	5	0.38	0.38	0.38	0.37	0.35	0.39	0.41	0.37	0.37
	15	0.06	0.06	0.06	0.06	0.18	0.06	0.24	0.06	0.05
	25	0.03	0.03	0.03	0.03	0.27	0.03	0.34	0.03	0.03
	35	0.02	0.02	0.02	0.03	0.36	0.03	0.42	0.02	0.02
0.15	5	0.35	0.35	0.35	0.38	0.40	0.35	0.37	0.36	0.36
	15	0.05	0.05	0.05	0.05	0.19	0.05	0.25	0.05	0.05
	25	0.03	0.03	0.03	0.03	0.28	0.03	0.35	0.03	0.03
	35	0.02	0.02	0.02	0.03	0.36	0.03	0.44	0.02	0.02
0.20	5	0.36	0.35	0.35	0.37	0.35	0.35	0.37	0.36	0.37
	15	0.06	0.06	0.06	0.06	0.20	0.06	0.25	0.06	0.06
	25	0.03	0.03	0.03	0.03	0.29	0.03	0.35	0.03	0.03
	35	0.02	0.02	0.02	0.03	0.37	0.03	0.43	0.02	0.02
0.25	5	0.41	0.42	0.42	0.40	0.41	0.41	0.43	0.41	0.40
	15	0.06	0.06	0.06	0.06	0.20	0.06	0.25	0.06	0.06
	25	0.03	0.03	0.03	0.03	0.29	0.03	0.35	0.03	0.03
	35	0.02	0.02	0.02	0.03	0.37	0.03	0.43	0.02	0.02
0.30	5	0.37	0.37	0.38	0.37	0.35	0.38	0.42	0.37	0.38
	15	0.05	0.05	0.05	0.05	0.20	0.06	0.25	0.05	0.05
	25	0.02	0.02	0.02	0.03	0.29	0.03	0.35	0.02	0.02
	35	0.02	0.02	0.02	0.02	0.37	0.03	0.43	0.02	0.02
0.35	5	0.38	0.38	0.38	0.38	0.38	0.38	0.42	0.38	0.38
	15	0.05	0.05	0.05	0.05	0.20	0.06	0.24	0.05	0.05
	25	0.03	0.03	0.03	0.03	0.30	0.03	0.34	0.03	0.03
	35	0.02	0.02	0.02	0.03	0.37	0.03	0.42	0.02	0.02
0.40	5	0.37	0.37	0.37	0.37	0.37	0.37	0.40	0.37	0.37
	15	0.05	0.06	0.06	0.05	0.19	0.06	0.24	0.05	0.05
	25	0.03	0.03	0.03	0.03	0.28	0.04	0.34	0.03	0.03
	35	0.02	0.02	0.02	0.02	0.36	0.04	0.41	0.02	0.02
0.45	5	0.36	0.37	0.37	0.35	0.36	0.37	0.39	0.36	0.35
	15	0.05	0.05	0.05	0.05	0.19	0.06	0.22	0.05	0.05
	25	0.03	0.03	0.03	0.03	0.28	0.04	0.30	0.03	0.03
	35	0.02	0.02	0.02	0.02	0.35	0.04	0.36	0.02	0.02

Quadre A7.9. EQM de l'estimador GPH en el model  $(1-\phi L)(1-L)^d X_t = (1-\theta L)\varepsilon_t$  amb  $T=400$ .

$d$	$m$	$\phi=0.0$	$\phi=0.0$	$\phi=0.9$	$\phi=0.0$	$\phi=0.0$	$\phi=0.5$	$\phi=0.9$	$\phi=0.5$	$\phi=0.9$
		$\theta=0.0$	$\theta=0.5$	$\theta=0.9$	$\theta=0.5$	$\theta=0.9$	$\theta=0.0$	$\theta=0.0$	$\theta=0.0$	$\theta=0.0$
0.00	5	0.37	0.36	0.36	0.37	0.32	0.36	0.36	0.37	0.37
	15	0.05	0.05	0.05	0.05	0.11	0.05	0.16	0.05	0.05
	25	0.02	0.02	0.02	0.03	0.17	0.03	0.25	0.02	0.03
	35	0.02	0.02	0.02	0.02	0.24	0.02	0.33	0.02	0.02
0.05	5	0.39	0.38	0.38	0.39	0.35	0.37	0.37	0.39	0.38
	15	0.05	0.05	0.05	0.05	0.11	0.05	0.16	0.05	0.05
	25	0.03	0.02	0.02	0.03	0.18	0.03	0.24	0.03	0.03
	35	0.01	0.01	0.01	0.02	0.26	0.02	0.33	0.02	0.02
0.10	5	0.39	0.39	0.40	0.40	0.37	0.40	0.41	0.40	0.40
	15	0.06	0.06	0.06	0.06	0.14	0.06	0.16	0.06	0.06
	25	0.03	0.03	0.03	0.03	0.20	0.03	0.24	0.03	0.03
	35	0.02	0.02	0.02	0.02	0.27	0.02	0.33	0.02	0.02
0.15	5	0.37	0.37	0.37	0.39	0.37	0.37	0.37	0.38	0.39
	15	0.05	0.05	0.05	0.06	0.14	0.05	0.17	0.05	0.06
	25	0.03	0.03	0.03	0.03	0.21	0.03	0.25	0.03	0.03
	35	0.02	0.02	0.02	0.02	0.28	0.02	0.33	0.02	0.02
0.20	5	0.40	0.40	0.40	0.40	0.37	0.39	0.41	0.40	0.39
	15	0.05	0.05	0.05	0.05	0.14	0.05	0.17	0.05	0.05
	25	0.03	0.03	0.03	0.03	0.22	0.03	0.25	0.03	0.03
	35	0.02	0.02	0.02	0.02	0.28	0.02	0.34	0.02	0.02
0.25	5	0.36	0.36	0.36	0.36	0.37	0.36	0.37	0.35	0.35
	15	0.05	0.05	0.05	0.05	0.13	0.05	0.17	0.05	0.05
	25	0.02	0.02	0.02	0.02	0.21	0.03	0.26	0.02	0.02
	35	0.02	0.02	0.02	0.02	0.28	0.02	0.34	0.01	0.01
0.30	5	0.40	0.39	0.39	0.40	0.40	0.39	0.43	0.40	0.40
	15	0.06	0.06	0.06	0.05	0.14	0.06	0.18	0.06	0.06
	25	0.03	0.03	0.03	0.03	0.21	0.03	0.25	0.03	0.03
	35	0.02	0.02	0.02	0.02	0.28	0.02	0.33	0.02	0.02
0.35	5	0.37	0.38	0.37	0.37	0.37	0.38	0.41	0.37	0.37
	15	0.05	0.05	0.05	0.05	0.14	0.05	0.17	0.05	0.05
	25	0.02	0.02	0.02	0.02	0.21	0.03	0.25	0.02	0.02
	35	0.02	0.02	0.02	0.02	0.29	0.02	0.33	0.02	0.02
0.40	5	0.37	0.37	0.37	0.38	0.36	0.37	0.39	0.38	0.39
	15	0.05	0.05	0.05	0.05	0.14	0.05	0.17	0.05	0.06
	25	0.03	0.03	0.03	0.03	0.21	0.03	0.25	0.03	0.03
	35	0.02	0.02	0.02	0.02	0.28	0.03	0.33	0.02	0.02
0.45	5	0.36	0.36	0.36	0.36	0.37	0.36	0.38	0.36	0.36
	15	0.05	0.06	0.06	0.05	0.14	0.06	0.17	0.05	0.05
	25	0.03	0.03	0.03	0.03	0.21	0.03	0.23	0.03	0.03
	35	0.02	0.02	0.02	0.02	0.28	0.03	0.29	0.02	0.02

## Annex 7.2. Anàlisi dels estimadors GPHA i GPHB

A l'apartat 7.5.1 de la tesi s'han proposat dues extensions de l'estimador GPH amb la finalitat de millorar-ne les seves propietats quan les dades s'han generat a través d'un model ARFIMA( $p,d,q$ ) enlloc d'un model ARFIMA(0, $d$ ,0) ja que els resultats de l'Annex 7.1 mostren com la presència de coeficient autoregressiu o mitjana mòbil en el procés generador de les dades generen un biaix molt important de l'estimador GPH.

A les pàgines següents es presenten alguns dels resultats d'un exercici de simulació en que es comparen aquests estimadors que s'han anomenat com GPHA i GPHB amb l'estimador GPH i l'estimador proposat per Robinson (1994a) i presentat a l'apartat 7.3 de la tesi. Concretament, s'han generat diferents models ARFIMA(1, $d$ ,1):

$$(1-L)^d(1-\phi L)X_t = (1-\theta L)\varepsilon_t \quad (\text{A7.6})$$

i per cada un d'aquests models s'han generat 5000 sèries de 100 observacions. Per a cada una d'aquestes situacions s'ha estimat l'esperança i la desviació estàndard de quatre estimadors semiparamètrics diferents del paràmetre  $d$ .

Per una banda, i com a mètodes *clàssics* d'estimació semiparamètrica s'han usat l'estimador GPH i l'estimador de l'Índex del grau de memòria proposat a Robinson (1994a)<sup>22</sup>:

$$\tilde{\delta}_{m,q} = \frac{1}{2} - \frac{1}{2 \log(q)} \log \left\{ \frac{\tilde{F}(q \omega_m)}{\tilde{F}(\omega_m)} \right\} \quad (\text{A7.7})$$

amb  $q=0.2, 0.4, 0.6$  i  $0.8$ .

---

<sup>22</sup> Recordi's que, en models ARFIMA, aquest índex  $\delta$  coincideix amb  $d$ .

Per altra banda, s'han aplicat els estimadors GPHA i GPHB proposats a la tesi. Recordi's que l'estimador GPHA consisteix a estimar per MQO el paràmetre  $\beta$  en el següent model de regressió:

$$\log I(\omega_j) = \alpha + \beta R_j + \gamma C_j + \varepsilon_j \quad (\text{A7.8})$$

on la variable  $R_j$  és:

$$R_j = -\log\{4\sin^2(\pi j / T)\} \quad (\text{A7.9})$$

i la variable  $C_j$ :

$$C_j = \cos(2\pi j / T) \quad (\text{A7.10})$$

Pel que fa a l'estimador GPHB, consisteix a estimar per MQO el paràmetre  $\beta$  en el següent model de regressió lineal simple:

$$\log I(\omega_j) + \log\{1 + \tilde{\phi}^2 - 2\tilde{\phi} \cos(2\pi j / T)\} = \alpha + \beta R_j + \varepsilon_j \quad (\text{A7.11})$$

Per cada un d'aquests quatre estimadors s'han calculat estimacions de la seva esperança i desviació estàndard, per tal d'obtenir-ne el biaix i l'error quadràtic mitjà. Al quadre A7.10 es presenta un resum dels models analitzats i els quadres on es poden trobar els resultats.

**Quadre A7.10. Models analitzats en l'exercici de simulació**

$\phi$	$\theta$	<i>Biaix</i>	<i>EQM</i>
0.0	0.0	Quadre A7.11	Quadre A7.20
0.0	-0.5	Quadre A7.12	Quadre A7.21
0.0	-0.9	Quadre A7.13	Quadre A7.22
0.0	0.5	Quadre A7.14	Quadre A7.23
0.0	0.9	Quadre A7.15	Quadre A7.24
0.5	0.0	Quadre A7.16	Quadre A7.25
0.9	0.0	Quadre A7.17	Quadre A7.26
-0.5	0.0	Quadre A7.18	Quadre A7.27
-0.9	0.0	Quadre A7.19	Quadre A7.28

A l'apartat 7.5.1 de la tesi es presenten les principals conclusions que es poden treure d'aquest exercici de simulació.



Quadre A7.11. Biaix dels estimadors semiparamètrics en el model  $(1 - L)^d X_t = \varepsilon_t$ , amb  $T=100$ .

$d$	$m$	$GPH$	$q=0.2$	$q=0.4$	$q=0.6$	$q=0.8$	$GPHA$	$GPHB$
0.00	5	0.01	0.21	-0.48	-0.31	-0.25	0.03	0.01
	15	0.00	-0.17	-0.10	-0.08	-0.06	0.02	0.00
	25	0.00	-0.09	-0.06	-0.04	-0.03	0.00	0.00
	35	0.00	-0.06	-0.04	-0.03	-0.02	0.00	0.00
0.05	5	-0.02	0.21	-0.52	-0.33	-0.29	0.00	-0.02
	15	0.00	-0.19	-0.11	-0.09	-0.07	-0.01	-0.01
	25	0.00	-0.10	-0.07	-0.05	-0.05	0.00	-0.02
	35	0.00	-0.07	-0.05	-0.04	-0.04	0.00	-0.03
0.10	5	0.00	0.21	-0.52	-0.33	-0.26	-0.05	0.00
	15	0.00	-0.20	-0.11	-0.08	-0.07	0.00	-0.02
	25	0.00	-0.11	-0.07	-0.06	-0.04	-0.01	-0.04
	35	0.00	-0.08	-0.05	-0.04	-0.04	-0.01	-0.05
0.15	5	0.01	0.24	-0.51	-0.34	-0.26	0.01	0.01
	15	0.02	-0.19	-0.10	-0.09	-0.08	0.02	-0.01
	25	0.01	-0.10	-0.07	-0.06	-0.05	0.02	-0.05
	35	0.01	-0.07	-0.06	-0.05	-0.05	0.02	-0.08
0.20	5	0.01	0.24	-0.56	-0.35	-0.28	-0.09	0.00
	15	0.01	-0.20	-0.12	-0.10	-0.08	0.00	-0.03
	25	0.01	-0.11	-0.08	-0.06	-0.06	0.01	-0.08
	35	0.00	-0.08	-0.06	-0.06	-0.06	0.01	-0.12
0.25	5	0.06	0.27	-0.51	-0.33	-0.28	0.08	0.05
	15	0.02	-0.20	-0.12	-0.10	-0.08	0.04	-0.04
	25	0.01	-0.11	-0.08	-0.07	-0.06	0.03	-0.11
	35	0.01	-0.09	-0.07	-0.06	-0.07	0.02	-0.16
0.30	5	0.02	0.32	-0.53	-0.37	-0.31	0.11	0.00
	15	0.02	-0.22	-0.14	-0.12	-0.11	0.04	-0.06
	25	0.02	-0.13	-0.10	-0.09	-0.09	0.03	-0.15
	35	0.01	-0.11	-0.09	-0.08	-0.08	0.03	-0.21
0.35	5	0.05	0.36	-0.55	-0.37	-0.30	0.04	0.02
	15	0.02	-0.23	-0.15	-0.13	-0.12	0.04	-0.10
	25	0.02	-0.15	-0.11	-0.10	-0.09	0.02	-0.19
	35	0.02	-0.12	-0.10	-0.09	-0.09	0.02	-0.26
0.40	5	0.02	0.48	-0.58	-0.38	-0.31	-0.03	-0.01
	15	0.03	-0.25	-0.18	-0.16	-0.14	0.03	-0.13
	25	0.02	-0.17	-0.13	-0.12	-0.11	0.03	-0.24
	35	0.02	-0.14	-0.12	-0.11	-0.11	0.03	-0.32
0.45	5	0.05	0.61	-0.57	-0.38	-0.33	0.07	-0.01
	15	0.05	-0.26	-0.19	-0.17	-0.15	0.06	-0.18
	25	0.04	-0.18	-0.14	-0.13	-0.12	0.05	-0.30
	35	0.03	-0.15	-0.13	-0.12	-0.12	0.05	-0.38

Quadre A7.12. Biaix dels estimadors semiparamètrics en el model  $(1 - L)^d X_t = (1 + 0,5L)\varepsilon_t$ , amb  $T=100$ .

$d$	$m$	$GPH$	$q=0.2$	$q=0.4$	$q=0.6$	$q=0.8$	$GPHA$	$GPHB$
0.00	5	0.02	0.21	-0.48	-0.30	-0.24	0.01	0.00
	15	0.03	-0.16	-0.07	-0.04	-0.01	0.02	-0.07
	25	0.09	-0.04	0.02	0.07	0.11	0.00	-0.10
	35	0.16	0.04	0.11	0.17	0.22	-0.03	-0.09
0.05	5	-0.01	0.21	-0.52	-0.33	-0.28	-0.01	-0.04
	15	0.03	-0.17	-0.08	-0.05	-0.02	-0.01	-0.10
	25	0.08	-0.05	0.01	0.05	0.08	-0.01	-0.14
	35	0.16	0.02	0.09	0.14	0.19	-0.03	-0.13
0.10	5	0.01	0.21	-0.52	-0.33	-0.26	-0.05	-0.03
	15	0.03	-0.18	-0.09	-0.05	-0.03	0.00	-0.12
	25	0.08	-0.07	0.00	0.03	0.07	-0.02	-0.17
	35	0.16	0.01	0.07	0.13	0.17	-0.04	-0.16
0.15	5	0.01	0.24	-0.51	-0.34	-0.26	0.00	-0.03
	15	0.05	-0.18	-0.08	-0.06	-0.04	0.02	-0.14
	25	0.10	-0.06	-0.01	0.03	0.06	0.01	-0.20
	35	0.17	0.00	0.06	0.11	0.15	-0.01	-0.20
0.20	5	0.02	0.24	-0.56	-0.35	-0.28	-0.08	-0.04
	15	0.04	-0.19	-0.10	-0.07	-0.05	0.00	-0.18
	25	0.09	-0.07	-0.02	0.02	0.04	0.00	-0.25
	35	0.16	-0.01	0.05	0.09	0.12	-0.02	-0.25
0.25	5	0.07	0.27	-0.51	-0.32	-0.28	0.08	0.00
	15	0.06	-0.18	-0.10	-0.07	-0.05	0.04	-0.21
	25	0.10	-0.08	-0.03	0.00	0.03	0.02	-0.29
	35	0.18	-0.02	0.03	0.07	0.10	-0.01	-0.29
0.30	5	0.03	0.33	-0.52	-0.37	-0.30	0.11	-0.06
	15	0.06	-0.21	-0.12	-0.09	-0.07	0.04	-0.25
	25	0.10	-0.10	-0.05	-0.03	0.00	0.02	-0.33
	35	0.17	-0.05	0.00	0.04	0.06	0.00	-0.33
0.35	5	0.06	0.36	-0.55	-0.36	-0.30	0.05	-0.06
	15	0.05	-0.22	-0.14	-0.11	-0.09	0.04	-0.31
	25	0.10	-0.12	-0.07	-0.04	-0.02	0.02	-0.38
	35	0.18	-0.07	-0.02	0.01	0.04	-0.01	-0.38
0.40	5	0.03	0.48	-0.58	-0.38	-0.31	-0.03	-0.13
	15	0.06	-0.24	-0.16	-0.13	-0.11	0.03	-0.35
	25	0.11	-0.14	-0.09	-0.06	-0.05	0.02	-0.43
	35	0.18	-0.09	-0.05	-0.02	0.00	0.00	-0.43
0.45	5	0.05	0.62	-0.57	-0.38	-0.32	0.08	-0.16
	15	0.08	-0.25	-0.17	-0.15	-0.13	0.06	-0.41
	25	0.12	-0.15	-0.11	-0.09	-0.07	0.04	-0.48
	35	0.19	-0.11	-0.07	-0.05	-0.03	0.02	-0.47

Quadre A7.13. Biaix dels estimadors semiparamètrics en el model  $(1 - L)^d X_t = (1 + 0,9L)\varepsilon_t$ , amb  $T=100$ .

<i>d</i>	<i>m</i>	<i>GPH</i>	<i>q=0.2</i>	<i>q=0.4</i>	<i>q=0.6</i>	<i>q=0.8</i>	<i>GPHA</i>	<i>GPHB</i>
0.00	5	0.02	0.21	-0.48	-0.30	-0.24	0.01	-0.02
	15	0.03	-0.15	-0.06	-0.04	0.00	0.02	-0.13
	25	0.10	-0.03	0.03	0.08	0.13	-0.01	-0.17
	35	0.20	0.05	0.13	0.20	0.27	-0.05	-0.14
0.05	5	-0.01	0.21	-0.52	-0.33	-0.28	-0.02	-0.06
	15	0.03	-0.17	-0.08	-0.05	-0.01	-0.01	-0.15
	25	0.09	-0.04	0.02	0.07	0.10	-0.01	-0.21
	35	0.19	0.04	0.11	0.18	0.23	-0.05	-0.18
0.10	5	0.01	0.21	-0.52	-0.32	-0.26	-0.05	-0.04
	15	0.03	-0.18	-0.08	-0.04	-0.02	0.00	-0.18
	25	0.09	-0.06	0.01	0.05	0.09	-0.02	-0.24
	35	0.20	0.02	0.09	0.16	0.21	-0.06	-0.21
0.15	5	0.01	0.24	-0.51	-0.34	-0.26	0.00	-0.05
	15	0.05	-0.17	-0.08	-0.05	-0.03	0.02	-0.20
	25	0.11	-0.05	0.00	0.04	0.07	0.00	-0.26
	35	0.21	0.02	0.08	0.13	0.18	-0.03	-0.24
0.20	5	0.02	0.24	-0.56	-0.35	-0.28	-0.08	-0.07
	15	0.04	-0.19	-0.09	-0.06	-0.04	0.00	-0.24
	25	0.10	-0.07	-0.01	0.03	0.05	0.00	-0.32
	35	0.20	0.00	0.06	0.11	0.15	-0.04	-0.29
0.25	5	0.07	0.27	-0.51	-0.32	-0.27	0.07	-0.04
	15	0.06	-0.18	-0.10	-0.06	-0.04	0.04	-0.27
	25	0.11	-0.07	-0.02	0.01	0.04	0.02	-0.35
	35	0.21	-0.01	0.04	0.09	0.12	-0.02	-0.33
0.30	5	0.03	0.33	-0.52	-0.37	-0.30	0.11	-0.10
	15	0.06	-0.21	-0.12	-0.09	-0.07	0.04	-0.32
	25	0.11	-0.10	-0.04	-0.02	0.01	0.02	-0.39
	35	0.21	-0.04	0.01	0.05	0.09	-0.02	-0.37
0.35	5	0.06	0.36	-0.55	-0.36	-0.30	0.05	-0.11
	15	0.06	-0.22	-0.13	-0.11	-0.09	0.04	-0.38
	25	0.12	-0.12	-0.06	-0.03	-0.01	0.01	-0.44
	35	0.21	-0.06	-0.01	0.03	0.06	-0.02	-0.41
0.40	5	0.03	0.48	-0.58	-0.38	-0.31	-0.02	-0.18
	15	0.06	-0.24	-0.16	-0.13	-0.11	0.03	-0.42
	25	0.12	-0.14	-0.08	-0.06	-0.04	0.02	-0.48
	35	0.21	-0.09	-0.04	-0.01	0.02	-0.01	-0.45
0.45	5	0.05	0.62	-0.57	-0.38	-0.32	0.08	-0.21
	15	0.08	-0.25	-0.17	-0.15	-0.13	0.06	-0.46
	25	0.13	-0.15	-0.11	-0.08	-0.07	0.04	-0.52
	35	0.22	-0.11	-0.06	-0.04	-0.02	0.00	-0.49

Quadre A7.14. Biaix dels estimadors semiparamètrics en el model  $(1 - L)^d X_t = (1 - 0,5L)\varepsilon_t$ , amb  $T=100$ .

<i>d</i>	<i>m</i>	<i>GPH</i>	<i>q=0.2</i>	<i>q=0.4</i>	<i>q=0.6</i>	<i>q=0.8</i>	<i>GPHA</i>	<i>GPHB</i>
0.00	5	-0.01	0.19	-0.50	-0.33	-0.30	0.06	-0.01
	15	-0.16	-0.29	-0.27	-0.31	-0.33	-0.01	-0.14
	25	-0.27	-0.32	-0.39	-0.44	-0.48	-0.08	-0.19
	35	-0.34	-0.39	-0.48	-0.54	-0.54	-0.13	-0.20
0.05	5	-0.06	0.20	-0.54	-0.37	-0.33	0.02	-0.05
	15	-0.17	-0.30	-0.28	-0.31	-0.33	-0.04	-0.14
	25	-0.28	-0.33	-0.40	-0.45	-0.51	-0.09	-0.21
	35	-0.35	-0.39	-0.48	-0.53	-0.56	-0.14	-0.22
0.10	5	-0.04	0.19	-0.54	-0.36	-0.30	-0.03	-0.03
	15	-0.17	-0.31	-0.28	-0.29	-0.32	-0.04	-0.14
	25	-0.27	-0.33	-0.38	-0.44	-0.46	-0.10	-0.21
	35	-0.34	-0.39	-0.47	-0.51	-0.55	-0.14	-0.22
0.15	5	-0.02	0.22	-0.52	-0.36	-0.29	0.00	-0.02
	15	-0.15	-0.29	-0.26	-0.29	-0.32	-0.02	-0.13
	25	-0.26	-0.31	-0.37	-0.43	-0.47	-0.07	-0.20
	35	-0.34	-0.37	-0.46	-0.52	-0.55	-0.12	-0.22
0.20	5	-0.02	0.23	-0.57	-0.38	-0.32	-0.09	-0.02
	15	-0.16	-0.30	-0.27	-0.29	-0.31	-0.03	-0.13
	25	-0.27	-0.31	-0.36	-0.41	-0.46	-0.08	-0.21
	35	-0.34	-0.37	-0.45	-0.51	-0.55	-0.12	-0.24
0.25	5	0.02	0.25	-0.53	-0.36	-0.32	0.10	0.02
	15	-0.15	-0.29	-0.26	-0.27	-0.30	0.00	-0.13
	25	-0.26	-0.30	-0.35	-0.41	-0.44	-0.06	-0.21
	35	-0.34	-0.36	-0.44	-0.50	-0.54	-0.12	-0.25
0.30	5	-0.03	0.31	-0.56	-0.41	-0.35	0.10	-0.03
	15	-0.15	-0.32	-0.27	-0.28	-0.31	-0.01	-0.13
	25	-0.26	-0.32	-0.36	-0.41	-0.45	-0.06	-0.22
	35	-0.34	-0.37	-0.44	-0.49	-0.52	-0.11	-0.27
0.35	5	0.01	0.35	-0.57	-0.39	-0.33	0.03	0.01
	15	-0.16	-0.32	-0.28	-0.29	-0.31	0.00	-0.14
	25	-0.26	-0.31	-0.35	-0.39	-0.42	-0.07	-0.23
	35	-0.33	-0.36	-0.42	-0.47	-0.51	-0.12	-0.28
0.40	5	-0.03	0.46	-0.61	-0.41	-0.34	-0.05	-0.02
	15	-0.15	-0.33	-0.29	-0.30	-0.32	-0.02	-0.14
	25	-0.26	-0.32	-0.35	-0.38	-0.41	-0.07	-0.24
	35	-0.34	-0.36	-0.42	-0.47	-0.51	-0.11	-0.31
0.45	5	-0.01	0.59	-0.60	-0.42	-0.36	0.03	-0.01
	15	-0.14	-0.33	-0.29	-0.30	-0.31	0.01	-0.14
	25	-0.25	-0.32	-0.34	-0.37	-0.40	-0.06	-0.26
	35	-0.32	-0.35	-0.39	-0.44	-0.47	-0.11	-0.33

Quadre A7.15. Biaix dels estimadors semiparamètrics en el model  $(1-L)^d X_t = (1-0,9L)\varepsilon_t$ , amb  $T=100$ .

$d$	$m$	$GPH$	$q=0.2$	$q=0.4$	$q=0.6$	$q=0.8$	$GPHA$	$GPHB$
0.00	5	-0.27	0.05	-0.69	-0.62	-0.67	-0.06	-0.27
	15	-0.58	-0.74	-0.81	-0.87	-0.87	-0.35	-0.55
	25	-0.68	-0.82	-0.88	-0.88	-0.86	-0.49	-0.59
	35	-0.73	-0.85	-0.86	-0.85	-0.80	-0.56	-0.57
0.05	5	-0.34	0.04	-0.75	-0.67	-0.68	-0.15	-0.33
	15	-0.60	-0.77	-0.83	-0.88	-0.88	-0.40	-0.57
	25	-0.70	-0.84	-0.90	-0.90	-0.91	-0.52	-0.62
	35	-0.75	-0.86	-0.88	-0.85	-0.83	-0.59	-0.60
0.10	5	-0.36	0.00	-0.78	-0.69	-0.68	-0.22	-0.35
	15	-0.62	-0.78	-0.84	-0.86	-0.88	-0.43	-0.59
	25	-0.71	-0.85	-0.89	-0.90	-0.86	-0.56	-0.64
	35	-0.76	-0.86	-0.88	-0.83	-0.83	-0.61	-0.60
0.15	5	-0.36	0.00	-0.78	-0.69	-0.68	-0.20	-0.35
	15	-0.62	-0.78	-0.83	-0.88	-0.89	-0.43	-0.59
	25	-0.71	-0.84	-0.89	-0.90	-0.88	-0.55	-0.64
	35	-0.76	-0.86	-0.88	-0.86	-0.84	-0.61	-0.61
0.20	5	-0.39	-0.01	-0.85	-0.74	-0.73	-0.31	-0.39
	15	-0.65	-0.80	-0.84	-0.88	-0.89	-0.48	-0.62
	25	-0.74	-0.85	-0.89	-0.89	-0.90	-0.58	-0.66
	35	-0.78	-0.87	-0.88	-0.87	-0.86	-0.63	-0.63
0.25	5	-0.39	0.01	-0.83	-0.74	-0.76	-0.18	-0.39
	15	-0.65	-0.80	-0.84	-0.86	-0.88	-0.47	-0.63
	25	-0.74	-0.85	-0.89	-0.90	-0.89	-0.58	-0.66
	35	-0.78	-0.87	-0.89	-0.87	-0.87	-0.64	-0.64
0.30	5	-0.47	0.03	-0.89	-0.80	-0.82	-0.19	-0.47
	15	-0.68	-0.82	-0.85	-0.87	-0.90	-0.51	-0.65
	25	-0.75	-0.86	-0.89	-0.91	-0.90	-0.61	-0.68
	35	-0.79	-0.88	-0.89	-0.87	-0.85	-0.66	-0.66
0.35	5	-0.42	0.08	-0.87	-0.75	-0.74	-0.27	-0.42
	15	-0.67	-0.80	-0.83	-0.87	-0.89	-0.49	-0.64
	25	-0.74	-0.84	-0.88	-0.88	-0.89	-0.60	-0.67
	35	-0.78	-0.86	-0.87	-0.87	-0.85	-0.65	-0.64
0.40	5	-0.48	0.16	-0.92	-0.78	-0.74	-0.41	-0.48
	15	-0.67	-0.81	-0.84	-0.86	-0.87	-0.54	-0.65
	25	-0.74	-0.84	-0.87	-0.87	-0.87	-0.62	-0.67
	35	-0.78	-0.86	-0.87	-0.87	-0.86	-0.66	-0.66
0.45	5	-0.45	0.26	-0.88	-0.77	-0.75	-0.29	-0.45
	15	-0.64	-0.77	-0.79	-0.82	-0.82	-0.50	-0.61
	25	-0.70	-0.79	-0.82	-0.83	-0.83	-0.58	-0.64
	35	-0.73	-0.81	-0.82	-0.82	-0.81	-0.62	-0.62

Quadre A7.16. Biaix dels estimadors semiparamètrics en el model  $(1 - L)^d(1 - 0,5L)X_t = \varepsilon_t$ , amb  $T=100$ .

$d$	$m$	$GPH$	$q=0.2$	$q=0.4$	$q=0.6$	$q=0.8$	$GPHA$	$GPHB$
0.00	5	0.05	0.22	-0.46	-0.27	-0.21	0.01	0.01
	15	0.17	-0.07	0.05	0.10	0.15	0.06	0.01
	25	0.28	0.09	0.18	0.24	0.27	0.10	0.02
	35	0.36	0.17	0.25	0.29	0.31	0.14	0.02
0.05	5	0.02	0.23	-0.50	-0.30	-0.25	-0.01	-0.03
	15	0.17	-0.09	0.03	0.09	0.13	0.02	-0.02
	25	0.28	0.07	0.16	0.21	0.23	0.09	-0.02
	35	0.35	0.15	0.23	0.26	0.28	0.14	-0.03
0.10	5	0.04	0.23	-0.49	-0.30	-0.22	-0.04	-0.02
	15	0.17	-0.10	0.02	0.08	0.12	0.03	-0.06
	25	0.28	0.05	0.14	0.18	0.21	0.08	-0.07
	35	0.35	0.13	0.20	0.24	0.25	0.13	-0.07
0.15	5	0.04	0.25	-0.49	-0.31	-0.22	0.00	-0.03
	15	0.19	-0.11	0.02	0.06	0.09	0.05	-0.09
	25	0.29	0.04	0.12	0.16	0.19	0.11	-0.11
	35	0.36	0.11	0.18	0.21	0.22	0.16	-0.12
0.20	5	0.05	0.26	-0.54	-0.33	-0.25	-0.08	-0.05
	15	0.18	-0.12	-0.01	0.04	0.07	0.03	-0.15
	25	0.28	0.02	0.10	0.14	0.16	0.10	-0.17
	35	0.35	0.09	0.15	0.18	0.19	0.15	-0.17
0.25	5	0.10	0.28	-0.49	-0.30	-0.25	0.07	-0.04
	15	0.20	-0.12	-0.02	0.03	0.06	0.08	-0.19
	25	0.30	0.01	0.08	0.11	0.13	0.12	-0.22
	35	0.37	0.07	0.12	0.15	0.16	0.17	-0.22
0.30	5	0.06	0.34	-0.51	-0.34	-0.27	0.10	-0.11
	15	0.19	-0.16	-0.05	0.00	0.03	0.08	-0.25
	25	0.29	-0.02	0.05	0.08	0.09	0.12	-0.27
	35	0.36	0.04	0.09	0.11	0.12	0.17	-0.27
0.35	5	0.08	0.37	-0.53	-0.34	-0.28	0.05	-0.14
	15	0.19	-0.17	-0.07	-0.03	0.00	0.07	-0.31
	25	0.30	-0.05	0.02	0.05	0.06	0.11	-0.32
	35	0.37	0.01	0.06	0.08	0.09	0.17	-0.31
0.40	5	0.06	0.48	-0.56	-0.36	-0.28	-0.02	-0.22
	15	0.19	-0.20	-0.10	-0.06	-0.03	0.06	-0.36
	25	0.30	-0.08	-0.01	0.01	0.03	0.12	-0.37
	35	0.36	-0.03	0.02	0.04	0.04	0.17	-0.36
0.45	5	0.08	0.62	-0.55	-0.36	-0.30	0.08	-0.28
	15	0.21	-0.21	-0.12	-0.09	-0.06	0.10	-0.42
	25	0.31	-0.10	-0.05	-0.02	-0.01	0.14	-0.42
	35	0.37	-0.06	-0.02	0.00	0.01	0.18	-0.40

Quadre A7.17. Biaix dels estimadors semiparamètrics en el model  $(1-L)^d(1-0,9L)X_t = \varepsilon_t$ , amb  $T=100$ .

$d$	$m$	$GPH$	$q=0.2$	$q=0.4$	$q=0.6$	$q=0.8$	$GPHA$	$GPHB$
0.00	5	0.53	0.43	-0.14	0.07	0.16	0.32	0.14
	15	0.76	0.26	0.37	0.40	0.42	0.63	0.10
	25	0.83	0.38	0.43	0.44	0.45	0.71	0.08
	35	0.86	0.42	0.45	0.45	0.46	0.76	0.06
0.05	5	0.50	0.44	-0.19	0.04	0.13	0.27	0.04
	15	0.76	0.23	0.33	0.36	0.38	0.59	0.05
	25	0.82	0.34	0.39	0.40	0.41	0.71	0.04
	35	0.85	0.38	0.41	0.41	0.42	0.76	0.03
0.10	5	0.53	0.43	-0.19	0.03	0.11	0.29	0.00
	15	0.76	0.20	0.30	0.33	0.34	0.62	0.01
	25	0.82	0.30	0.35	0.36	0.36	0.71	-0.01
	35	0.84	0.34	0.36	0.37	0.37	0.75	-0.01
0.15	5	0.52	0.44	-0.22	-0.01	0.09	0.30	-0.09
	15	0.76	0.17	0.26	0.29	0.30	0.61	-0.05
	25	0.82	0.27	0.31	0.32	0.32	0.71	-0.04
	35	0.85	0.30	0.32	0.33	0.33	0.76	-0.04
0.20	5	0.51	0.44	-0.26	-0.03	0.04	0.30	-0.15
	15	0.75	0.13	0.22	0.24	0.26	0.59	-0.09
	25	0.81	0.23	0.26	0.27	0.28	0.70	-0.08
	35	0.83	0.26	0.27	0.28	0.28	0.75	-0.08
0.25	5	0.54	0.45	-0.26	-0.05	0.02	0.31	-0.19
	15	0.76	0.10	0.18	0.20	0.21	0.61	-0.12
	25	0.81	0.19	0.22	0.23	0.23	0.71	-0.10
	35	0.83	0.21	0.23	0.23	0.24	0.75	-0.10
0.30	5	0.52	0.50	-0.27	-0.08	0.00	0.35	-0.24
	15	0.74	0.06	0.14	0.16	0.17	0.62	-0.16
	25	0.79	0.14	0.17	0.18	0.18	0.70	-0.14
	35	0.80	0.17	0.18	0.19	0.19	0.73	-0.14
0.35	5	0.54	0.52	-0.30	-0.11	-0.04	0.33	-0.26
	15	0.72	0.02	0.09	0.11	0.12	0.60	-0.19
	25	0.77	0.10	0.12	0.13	0.13	0.68	-0.17
	35	0.79	0.12	0.13	0.14	0.14	0.71	-0.16
0.40	5	0.47	0.60	-0.36	-0.15	-0.09	0.29	-0.34
	15	0.67	-0.02	0.05	0.06	0.07	0.55	-0.24
	25	0.72	0.05	0.08	0.08	0.09	0.63	-0.22
	35	0.73	0.07	0.08	0.09	0.09	0.67	-0.22
0.45	5	0.43	0.69	-0.38	-0.19	-0.14	0.30	-0.38
	15	0.60	-0.07	0.00	0.01	0.02	0.50	-0.31
	25	0.64	0.00	0.02	0.03	0.03	0.57	-0.30
	35	0.66	0.02	0.03	0.04	0.04	0.60	-0.29

Quadre A7.18. Biaix dels estimadors semiparamètrics en el model  $(1 - L)^d(1 + 0,5L)X_t = \varepsilon_t$ , amb  $T=100$ .

<i>d</i>	<i>m</i>	<i>GPH</i>	<i>q=0.2</i>	<i>q=0.4</i>	<i>q=0.6</i>	<i>q=0.8</i>	<i>GPHA</i>	<i>GPHB</i>
0.00	5	0.01	0.20	-0.48	-0.31	-0.26	0.04	0.02
	15	-0.03	-0.19	-0.13	-0.13	-0.11	0.02	0.00
	25	-0.08	-0.15	-0.14	-0.16	-0.19	0.01	0.00
	35	-0.16	-0.19	-0.24	-0.32	-0.42	0.03	0.00
0.05	5	-0.03	0.21	-0.53	-0.34	-0.30	0.01	-0.02
	15	-0.03	-0.21	-0.14	-0.13	-0.12	-0.01	0.00
	25	-0.09	-0.15	-0.15	-0.17	-0.21	0.00	-0.01
	35	-0.16	-0.19	-0.24	-0.32	-0.42	0.03	-0.01
0.10	5	0.00	0.20	-0.52	-0.34	-0.27	-0.04	0.00
	15	-0.04	-0.21	-0.14	-0.12	-0.12	0.00	-0.01
	25	-0.09	-0.16	-0.15	-0.17	-0.19	-0.01	-0.01
	35	-0.16	-0.19	-0.24	-0.30	-0.40	0.02	-0.01
0.15	5	0.01	0.23	-0.50	-0.34	-0.26	0.01	0.01
	15	-0.02	-0.21	-0.13	-0.13	-0.12	0.02	0.01
	25	-0.07	-0.15	-0.15	-0.16	-0.19	0.03	0.00
	35	-0.15	-0.18	-0.23	-0.30	-0.39	0.05	-0.02
0.20	5	0.00	0.24	-0.56	-0.36	-0.29	-0.09	0.01
	15	-0.02	-0.22	-0.14	-0.13	-0.13	0.00	0.00
	25	-0.08	-0.16	-0.15	-0.16	-0.19	0.01	-0.01
	35	-0.16	-0.19	-0.22	-0.29	-0.39	0.04	-0.03
0.25	5	0.05	0.27	-0.52	-0.33	-0.28	0.09	0.06
	15	-0.01	-0.21	-0.14	-0.13	-0.12	0.04	0.01
	25	-0.07	-0.16	-0.15	-0.16	-0.18	0.03	-0.01
	35	-0.15	-0.18	-0.22	-0.28	-0.36	0.05	-0.05
0.30	5	0.01	0.32	-0.53	-0.38	-0.31	0.10	0.01
	15	-0.01	-0.24	-0.17	-0.15	-0.14	0.04	0.01
	25	-0.07	-0.18	-0.16	-0.18	-0.20	0.04	-0.02
	35	-0.15	-0.20	-0.23	-0.28	-0.35	0.06	-0.07
0.35	5	0.05	0.36	-0.56	-0.37	-0.31	0.04	0.05
	15	-0.02	-0.25	-0.18	-0.16	-0.16	0.04	-0.01
	25	-0.07	-0.18	-0.17	-0.17	-0.19	0.03	-0.04
	35	-0.15	-0.20	-0.22	-0.26	-0.33	0.05	-0.09
0.40	5	0.01	0.47	-0.59	-0.39	-0.31	-0.04	0.01
	15	-0.01	-0.27	-0.20	-0.18	-0.17	0.02	-0.01
	25	-0.06	-0.20	-0.18	-0.18	-0.20	0.03	-0.05
	35	-0.15	-0.21	-0.23	-0.27	-0.34	0.06	-0.13
0.45	5	0.04	0.61	-0.58	-0.39	-0.33	0.05	0.04
	15	0.00	-0.27	-0.21	-0.19	-0.18	0.06	-0.01
	25	-0.05	-0.21	-0.19	-0.19	-0.20	0.05	-0.08
	35	-0.14	-0.21	-0.22	-0.26	-0.31	0.07	-0.16



Quadre A7.19. Biaix dels estimadors semiparamètrics en el model  $(1 - L)^d(1 + 0,9L)X_t = \varepsilon_t$ , amb  $T=100$ .

$d$	$m$	$GPH$	$q=0.2$	$q=0.4$	$q=0.6$	$q=0.8$	$GPHA$	$GPHB$
0.00	5	0.02	0.20	-0.48	-0.31	-0.26	0.06	0.02
	15	-0.03	-0.20	-0.13	-0.13	-0.12	0.02	0.00
	25	-0.09	-0.16	-0.16	-0.18	-0.22	0.01	0.01
	35	-0.19	-0.22	-0.29	-0.40	-0.55	0.06	0.01
0.05	5	-0.04	0.21	-0.53	-0.34	-0.30	0.00	-0.03
	15	-0.04	-0.21	-0.14	-0.13	-0.13	-0.02	-0.01
	25	-0.10	-0.16	-0.17	-0.20	-0.25	0.00	0.00
	35	-0.20	-0.22	-0.29	-0.40	-0.54	0.04	0.00
0.10	5	0.00	0.20	-0.52	-0.33	-0.26	-0.06	0.00
	15	-0.04	-0.22	-0.15	-0.13	-0.13	-0.01	0.00
	25	-0.10	-0.17	-0.16	-0.19	-0.21	-0.01	0.00
	35	-0.19	-0.22	-0.29	-0.37	-0.52	0.04	0.01
0.15	5	0.01	0.23	-0.50	-0.34	-0.26	0.00	0.01
	15	-0.03	-0.21	-0.14	-0.14	-0.13	0.02	0.01
	25	-0.09	-0.16	-0.16	-0.19	-0.22	0.02	0.01
	35	-0.19	-0.21	-0.28	-0.37	-0.51	0.06	0.01
0.20	5	0.00	0.24	-0.56	-0.36	-0.30	-0.09	0.01
	15	-0.04	-0.23	-0.15	-0.14	-0.14	0.00	0.00
	25	-0.10	-0.17	-0.16	-0.18	-0.22	0.01	0.00
	35	-0.20	-0.22	-0.27	-0.36	-0.50	0.05	0.00
0.25	5	0.05	0.27	-0.52	-0.33	-0.28	0.10	0.05
	15	-0.02	-0.22	-0.15	-0.14	-0.13	0.02	0.01
	25	-0.09	-0.17	-0.16	-0.19	-0.21	0.03	0.01
	35	-0.19	-0.21	-0.27	-0.35	-0.48	0.06	0.01
0.30	5	0.00	0.32	-0.53	-0.39	-0.32	0.09	0.01
	15	-0.02	-0.25	-0.17	-0.16	-0.15	0.03	0.01
	25	-0.09	-0.19	-0.18	-0.20	-0.23	0.03	0.00
	35	-0.20	-0.23	-0.27	-0.34	-0.45	0.06	0.00
0.35	5	0.03	0.36	-0.56	-0.37	-0.31	0.02	0.04
	15	-0.03	-0.25	-0.18	-0.17	-0.17	0.04	0.00
	25	-0.09	-0.20	-0.18	-0.19	-0.22	0.02	0.00
	35	-0.19	-0.23	-0.26	-0.33	-0.43	0.05	0.00
0.40	5	0.00	0.47	-0.59	-0.39	-0.32	-0.05	0.01
	15	-0.02	-0.27	-0.21	-0.19	-0.18	0.02	0.01
	25	-0.09	-0.21	-0.20	-0.20	-0.22	0.02	0.01
	35	-0.19	-0.24	-0.27	-0.33	-0.43	0.06	0.00
0.45	5	0.03	0.61	-0.58	-0.39	-0.34	0.04	0.03
	15	-0.01	-0.28	-0.21	-0.20	-0.19	0.05	0.02
	25	-0.08	-0.22	-0.21	-0.21	-0.23	0.04	0.01
	35	-0.18	-0.24	-0.26	-0.32	-0.40	0.07	0.00

Quadre A7.20. EQM dels estimadors semiparamètrics en el model  $(1 - L)^d X_t = \varepsilon_t$ , amb  $T=100$ .

$d$	$m$	$GPH$	$q=0.2$	$q=0.4$	$q=0.6$	$q=0.8$	$GPHA$	$GPHB$
0.00	5	0.39	0.12	0.64	0.46	0.63	2.67	0.39
	15	0.06	0.09	0.05	0.07	0.11	0.20	0.06
	25	0.03	0.03	0.03	0.03	0.05	0.09	0.03
	35	0.02	0.02	0.02	0.02	0.04	0.06	0.02
0.05	5	0.36	0.14	0.67	0.43	0.63	2.76	0.36
	15	0.06	0.09	0.05	0.06	0.10	0.20	0.06
	25	0.03	0.03	0.03	0.03	0.05	0.09	0.03
	35	0.02	0.02	0.02	0.02	0.04	0.06	0.02
0.10	5	0.42	0.15	0.66	0.44	0.52	2.53	0.42
	15	0.06	0.10	0.05	0.05	0.08	0.21	0.06
	25	0.03	0.04	0.02	0.03	0.04	0.09	0.03
	35	0.02	0.02	0.02	0.02	0.03	0.07	0.02
0.15	5	0.37	0.17	0.61	0.42	0.47	2.75	0.37
	15	0.05	0.09	0.04	0.05	0.07	0.20	0.05
	25	0.03	0.03	0.02	0.03	0.04	0.09	0.03
	35	0.02	0.02	0.01	0.02	0.03	0.06	0.02
0.20	5	0.44	0.19	0.74	0.42	0.52	3.09	0.43
	15	0.06	0.09	0.04	0.05	0.07	0.20	0.06
	25	0.03	0.03	0.02	0.02	0.03	0.10	0.03
	35	0.02	0.02	0.01	0.01	0.03	0.06	0.03
0.25	5	0.38	0.19	0.56	0.34	0.47	2.87	0.37
	15	0.05	0.08	0.04	0.04	0.05	0.19	0.05
	25	0.03	0.03	0.02	0.02	0.03	0.08	0.03
	35	0.02	0.02	0.01	0.01	0.02	0.06	0.04
0.30	5	0.34	0.25	0.57	0.38	0.38	2.87	0.34
	15	0.06	0.09	0.04	0.04	0.05	0.20	0.05
	25	0.03	0.03	0.02	0.02	0.03	0.09	0.04
	35	0.02	0.02	0.02	0.02	0.02	0.06	0.05
0.35	5	0.40	0.30	0.64	0.34	0.38	2.95	0.39
	15	0.06	0.09	0.04	0.04	0.05	0.20	0.06
	25	0.03	0.04	0.02	0.02	0.02	0.09	0.06
	35	0.02	0.02	0.02	0.01	0.02	0.06	0.08
0.40	5	0.40	0.43	0.64	0.34	0.29	2.63	0.39
	15	0.06	0.10	0.05	0.05	0.05	0.20	0.06
	25	0.03	0.04	0.03	0.02	0.03	0.09	0.08
	35	0.02	0.03	0.02	0.02	0.02	0.06	0.11
0.45	5	0.36	0.64	0.61	0.29	0.29	2.73	0.34
	15	0.06	0.10	0.05	0.04	0.04	0.19	0.07
	25	0.03	0.04	0.03	0.02	0.02	0.09	0.11
	35	0.02	0.03	0.02	0.02	0.02	0.06	0.15

Quadre A7.21. EQM dels estimadors semiparamètrics en el model  $(1 - L)^d X_t = (1 + 0,5L)\varepsilon_t$ , amb  $T=100$ .

$d$	$m$	$GPH$	$q=0.2$	$q=0.4$	$q=0.6$	$q=0.8$	$GPHA$	$GPHB$
0.00	5	0.39	0.12	0.63	0.45	0.63	2.62	0.39
	15	0.06	0.08	0.05	0.06	0.09	0.20	0.06
	25	0.04	0.02	0.02	0.03	0.04	0.09	0.03
	35	0.05	0.01	0.02	0.04	0.06	0.06	0.02
0.05	5	0.36	0.14	0.67	0.42	0.62	2.78	0.36
	15	0.06	0.08	0.04	0.05	0.09	0.20	0.06
	25	0.03	0.02	0.02	0.02	0.04	0.09	0.04
	35	0.04	0.01	0.02	0.03	0.05	0.06	0.03
0.10	5	0.41	0.15	0.64	0.43	0.52	2.50	0.41
	15	0.06	0.09	0.05	0.04	0.06	0.20	0.07
	25	0.04	0.03	0.02	0.02	0.03	0.09	0.05
	35	0.05	0.01	0.01	0.02	0.04	0.07	0.04
0.15	5	0.38	0.17	0.62	0.43	0.47	2.76	0.37
	15	0.05	0.08	0.03	0.04	0.06	0.20	0.07
	25	0.04	0.02	0.01	0.02	0.02	0.09	0.06
	35	0.05	0.01	0.01	0.02	0.03	0.06	0.05
0.20	5	0.45	0.19	0.74	0.42	0.51	3.17	0.44
	15	0.06	0.08	0.04	0.04	0.05	0.20	0.08
	25	0.04	0.02	0.01	0.01	0.02	0.10	0.08
	35	0.04	0.01	0.01	0.01	0.02	0.07	0.07
0.25	5	0.37	0.20	0.55	0.33	0.46	2.81	0.36
	15	0.05	0.07	0.03	0.03	0.04	0.19	0.08
	25	0.04	0.02	0.01	0.01	0.01	0.08	0.10
	35	0.05	0.01	0.01	0.01	0.01	0.06	0.09
0.30	5	0.35	0.26	0.56	0.37	0.37	2.89	0.34
	15	0.06	0.09	0.04	0.03	0.03	0.20	0.10
	25	0.04	0.03	0.01	0.01	0.01	0.09	0.13
	35	0.05	0.01	0.00	0.01	0.01	0.06	0.12
0.35	5	0.40	0.30	0.63	0.33	0.38	2.94	0.38
	15	0.06	0.09	0.04	0.03	0.04	0.20	0.13
	25	0.04	0.03	0.01	0.01	0.01	0.09	0.16
	35	0.05	0.01	0.00	0.00	0.00	0.07	0.15
0.40	5	0.39	0.43	0.63	0.33	0.29	2.63	0.38
	15	0.06	0.09	0.05	0.04	0.03	0.20	0.16
	25	0.04	0.03	0.01	0.01	0.01	0.09	0.20
	35	0.05	0.02	0.01	0.00	0.00	0.06	0.19
0.45	5	0.35	0.64	0.60	0.29	0.29	2.73	0.33
	15	0.06	0.09	0.04	0.04	0.03	0.19	0.20
	25	0.04	0.03	0.02	0.01	0.01	0.09	0.24
	35	0.06	0.02	0.01	0.00	0.00	0.06	0.23

Quadre A7.22. EQM dels estimadors semiparamètrics en el model  $(1 - L)^d X_t = (1 + 0,9L)\varepsilon_t$ , amb  $T=100$ .

$d$	$m$	$GPH$	$q=0.2$	$q=0.4$	$q=0.6$	$q=0.8$	$GPHA$	$GPHB$
0.00	5	0.39	0.12	0.63	0.45	0.63	2.63	0.39
	15	0.06	0.08	0.04	0.06	0.09	0.20	0.07
	25	0.04	0.02	0.02	0.03	0.05	0.09	0.05
	35	0.06	0.02	0.03	0.05	0.08	0.06	0.03
0.05	5	0.37	0.14	0.68	0.42	0.61	2.78	0.37
	15	0.06	0.08	0.04	0.05	0.08	0.20	0.08
	25	0.04	0.02	0.02	0.02	0.04	0.09	0.06
	35	0.06	0.01	0.02	0.04	0.06	0.06	0.04
0.10	5	0.41	0.15	0.64	0.42	0.52	2.51	0.41
	15	0.06	0.09	0.05	0.04	0.06	0.20	0.09
	25	0.04	0.03	0.01	0.02	0.03	0.09	0.08
	35	0.06	0.01	0.02	0.03	0.05	0.07	0.05
0.15	5	0.38	0.17	0.63	0.43	0.47	2.76	0.38
	15	0.06	0.08	0.03	0.04	0.05	0.20	0.09
	25	0.04	0.02	0.01	0.02	0.02	0.09	0.09
	35	0.06	0.01	0.01	0.02	0.04	0.06	0.07
0.20	5	0.45	0.19	0.74	0.43	0.51	3.22	0.45
	15	0.06	0.08	0.03	0.04	0.05	0.20	0.11
	25	0.04	0.02	0.01	0.01	0.02	0.09	0.12
	35	0.06	0.01	0.01	0.02	0.03	0.07	0.09
0.25	5	0.37	0.20	0.55	0.33	0.45	2.78	0.36
	15	0.05	0.07	0.03	0.03	0.03	0.19	0.11
	25	0.04	0.02	0.01	0.01	0.01	0.08	0.14
	35	0.06	0.01	0.01	0.01	0.02	0.06	0.11
0.30	5	0.35	0.26	0.56	0.36	0.37	2.89	0.35
	15	0.06	0.08	0.04	0.03	0.03	0.20	0.14
	25	0.04	0.02	0.01	0.01	0.01	0.09	0.17
	35	0.06	0.01	0.00	0.01	0.01	0.06	0.14
0.35	5	0.40	0.30	0.63	0.33	0.38	2.95	0.38
	15	0.06	0.09	0.04	0.03	0.03	0.20	0.18
	25	0.04	0.03	0.01	0.01	0.01	0.10	0.21
	35	0.06	0.01	0.00	0.00	0.01	0.07	0.18
0.40	5	0.39	0.43	0.63	0.33	0.30	2.65	0.39
	15	0.06	0.09	0.04	0.04	0.03	0.20	0.21
	25	0.04	0.03	0.01	0.01	0.01	0.09	0.24
	35	0.06	0.01	0.00	0.00	0.00	0.06	0.21
0.45	5	0.35	0.64	0.60	0.29	0.29	2.70	0.35
	15	0.06	0.09	0.04	0.03	0.03	0.19	0.25
	25	0.05	0.03	0.02	0.01	0.01	0.09	0.29
	35	0.07	0.02	0.01	0.00	0.00	0.06	0.25

Quadre A7.23. EQM dels estimadors semiparamètrics en el model  $(1 - L)^d X_t = (1 - 0,5L)\varepsilon_t$ , amb  $T=100$ .

$d$	$m$	$GPH$	$q=0.2$	$q=0.4$	$q=0.6$	$q=0.8$	$GPHA$	$GPHB$
0.00	5	0.38	0.12	0.66	0.47	0.70	2.85	0.38
	15	0.08	0.15	0.13	0.20	0.32	0.20	0.08
	25	0.10	0.13	0.19	0.26	0.39	0.10	0.07
	35	0.14	0.17	0.26	0.34	0.43	0.08	0.06
0.05	5	0.37	0.14	0.70	0.50	0.71	2.79	0.37
	15	0.08	0.15	0.13	0.18	0.31	0.20	0.07
	25	0.11	0.14	0.20	0.27	0.41	0.10	0.07
	35	0.14	0.17	0.26	0.34	0.44	0.08	0.07
0.10	5	0.40	0.14	0.67	0.47	0.56	2.52	0.40
	15	0.09	0.15	0.13	0.16	0.26	0.20	0.08
	25	0.11	0.14	0.18	0.26	0.35	0.10	0.07
	35	0.14	0.17	0.25	0.31	0.41	0.09	0.07
0.15	5	0.36	0.16	0.62	0.42	0.52	2.77	0.36
	15	0.07	0.14	0.11	0.15	0.24	0.19	0.07
	25	0.10	0.12	0.17	0.24	0.34	0.09	0.07
	35	0.13	0.16	0.24	0.31	0.41	0.07	0.07
0.20	5	0.41	0.18	0.72	0.45	0.59	2.93	0.41
	15	0.08	0.14	0.11	0.15	0.23	0.19	0.07
	25	0.10	0.12	0.16	0.22	0.31	0.10	0.07
	35	0.14	0.15	0.23	0.30	0.40	0.08	0.07
0.25	5	0.37	0.19	0.61	0.38	0.54	2.95	0.37
	15	0.07	0.13	0.10	0.13	0.19	0.19	0.07
	25	0.10	0.12	0.15	0.21	0.28	0.09	0.07
	35	0.13	0.15	0.22	0.28	0.37	0.07	0.08
0.30	5	0.35	0.24	0.63	0.42	0.43	2.92	0.35
	15	0.08	0.15	0.11	0.13	0.18	0.20	0.07
	25	0.10	0.12	0.15	0.21	0.29	0.09	0.08
	35	0.14	0.15	0.21	0.27	0.34	0.07	0.09
0.35	5	0.40	0.29	0.65	0.37	0.42	2.87	0.40
	15	0.08	0.14	0.11	0.13	0.17	0.19	0.08
	25	0.10	0.12	0.14	0.18	0.25	0.10	0.08
	35	0.13	0.14	0.19	0.25	0.32	0.08	0.09
0.40	5	0.40	0.41	0.68	0.38	0.33	2.63	0.40
	15	0.08	0.15	0.12	0.13	0.17	0.20	0.07
	25	0.10	0.13	0.14	0.17	0.23	0.10	0.08
	35	0.13	0.15	0.19	0.25	0.31	0.08	0.11
0.45	5	0.37	0.61	0.65	0.34	0.35	2.77	0.37
	15	0.07	0.15	0.11	0.12	0.15	0.19	0.07
	25	0.09	0.12	0.13	0.16	0.20	0.09	0.09
	35	0.12	0.13	0.17	0.22	0.27	0.07	0.12

Quadre A7.24. EQM dels estimadors semiparamètrics en el model  $(1 - L)^d X_t = (1 - 0.9L)\varepsilon_t$ , amb  $T=100$ .

$d$	$m$	$GPH$	$q=0.2$	$q=0.4$	$q=0.6$	$q=0.8$	$GPFA$	$GPFB$
0.00	5	0.42	0.10	0.89	0.86	1.59	2.33	0.42
	15	0.41	0.64	0.78	0.98	1.27	0.33	0.38
	25	0.51	0.72	0.85	0.89	1.03	0.36	0.40
	35	0.57	0.76	0.79	0.80	0.83	0.40	0.36
0.05	5	0.43	0.08	0.95	0.89	1.41	2.24	0.43
	15	0.44	0.67	0.81	0.99	1.25	0.36	0.40
	25	0.54	0.76	0.88	0.93	1.12	0.38	0.43
	35	0.60	0.78	0.81	0.82	0.87	0.43	0.39
0.10	5	0.46	0.09	0.99	0.91	1.34	2.20	0.46
	15	0.46	0.69	0.81	0.91	1.21	0.37	0.42
	25	0.56	0.77	0.85	0.93	0.99	0.42	0.45
	35	0.61	0.78	0.82	0.77	0.86	0.46	0.40
0.15	5	0.43	0.09	0.94	0.87	1.32	2.41	0.43
	15	0.46	0.68	0.78	0.95	1.20	0.37	0.42
	25	0.55	0.76	0.86	0.92	1.03	0.40	0.45
	35	0.61	0.77	0.82	0.82	0.89	0.45	0.40
0.20	5	0.54	0.10	1.14	1.02	1.51	2.59	0.53
	15	0.49	0.72	0.79	0.95	1.21	0.42	0.45
	25	0.58	0.76	0.86	0.90	1.04	0.44	0.47
	35	0.64	0.78	0.82	0.84	0.91	0.47	0.43
0.25	5	0.48	0.11	1.06	0.94	1.52	2.56	0.48
	15	0.49	0.71	0.79	0.90	1.13	0.41	0.46
	25	0.58	0.76	0.84	0.92	0.99	0.44	0.48
	35	0.64	0.78	0.84	0.83	0.90	0.48	0.44
0.30	5	0.54	0.14	1.16	1.04	1.46	2.71	0.54
	15	0.52	0.75	0.81	0.91	1.13	0.45	0.48
	25	0.61	0.78	0.86	0.93	1.02	0.47	0.50
	35	0.66	0.80	0.84	0.83	0.87	0.51	0.46
0.35	5	0.52	0.14	1.11	0.92	1.24	2.58	0.52
	15	0.51	0.70	0.76	0.89	1.09	0.42	0.47
	25	0.58	0.75	0.83	0.86	0.98	0.46	0.49
	35	0.63	0.76	0.80	0.81	0.86	0.49	0.44
0.40	5	0.61	0.18	1.27	0.95	1.09	2.64	0.61
	15	0.52	0.73	0.78	0.86	1.04	0.48	0.48
	25	0.58	0.74	0.80	0.82	0.92	0.48	0.49
	35	0.64	0.77	0.79	0.82	0.87	0.50	0.46
0.45	5	0.49	0.25	1.07	0.86	1.05	2.31	0.49
	15	0.46	0.65	0.68	0.77	0.89	0.40	0.43
	25	0.52	0.66	0.72	0.76	0.82	0.42	0.45
	35	0.57	0.68	0.71	0.73	0.77	0.45	0.42

Quadre A7.25. EQM dels estimadors semiparamètrics en el model  $(1 - L)^d(1 - 0,5L)X_t = \varepsilon_t$ , amb  $T=100$ .

$d$	$m$	$GPH$	$q=0.2$	$q=0.4$	$q=0.6$	$q=0.8$	$GPHA$	$GPHB$
0.00	5	0.39	0.13	0.60	0.41	0.57	2.61	0.39
	15	0.09	0.06	0.03	0.05	0.07	0.20	0.05
	25	0.11	0.03	0.05	0.07	0.08	0.10	0.02
	35	0.15	0.04	0.07	0.09	0.10	0.08	0.01
0.05	5	0.36	0.15	0.64	0.39	0.56	2.79	0.36
	15	0.08	0.05	0.03	0.04	0.06	0.20	0.05
	25	0.10	0.02	0.04	0.05	0.07	0.10	0.02
	35	0.14	0.03	0.06	0.07	0.08	0.08	0.01
0.10	5	0.41	0.16	0.61	0.39	0.47	2.55	0.41
	15	0.09	0.06	0.03	0.03	0.04	0.20	0.05
	25	0.11	0.02	0.03	0.04	0.06	0.10	0.02
	35	0.14	0.02	0.04	0.06	0.07	0.08	0.01
0.15	5	0.38	0.17	0.61	0.40	0.42	2.79	0.38
	15	0.09	0.06	0.02	0.02	0.04	0.21	0.05
	25	0.11	0.01	0.02	0.03	0.04	0.10	0.03
	35	0.15	0.02	0.04	0.05	0.05	0.09	0.02
0.20	5	0.44	0.20	0.71	0.39	0.47	3.19	0.44
	15	0.09	0.06	0.02	0.02	0.03	0.20	0.07
	25	0.11	0.01	0.02	0.02	0.03	0.10	0.05
	35	0.14	0.01	0.03	0.03	0.04	0.09	0.04
0.25	5	0.38	0.20	0.53	0.30	0.40	2.76	0.36
	15	0.09	0.05	0.02	0.01	0.02	0.19	0.07
	25	0.11	0.01	0.01	0.02	0.02	0.10	0.06
	35	0.15	0.01	0.02	0.02	0.03	0.08	0.06
0.30	5	0.35	0.26	0.54	0.34	0.33	2.88	0.35
	15	0.09	0.06	0.02	0.01	0.01	0.20	0.10
	25	0.11	0.01	0.01	0.01	0.01	0.10	0.09
	35	0.15	0.01	0.01	0.01	0.02	0.09	0.08
0.35	5	0.40	0.31	0.61	0.31	0.34	2.96	0.38
	15	0.10	0.06	0.02	0.01	0.01	0.20	0.13
	25	0.12	0.01	0.00	0.00	0.01	0.11	0.12
	35	0.15	0.00	0.00	0.01	0.01	0.09	0.11
0.40	5	0.39	0.44	0.62	0.31	0.27	2.65	0.40
	15	0.10	0.07	0.02	0.01	0.01	0.20	0.17
	25	0.12	0.02	0.00	0.00	0.00	0.11	0.15
	35	0.15	0.00	0.00	0.00	0.00	0.09	0.14
0.45	5	0.35	0.65	0.57	0.28	0.26	2.65	0.37
	15	0.10	0.07	0.02	0.01	0.01	0.19	0.21
	25	0.12	0.02	0.00	0.00	0.00	0.11	0.19
	35	0.16	0.01	0.00	0.00	0.00	0.09	0.17

Quadre A7.26. EQM dels estimadors semiparamètrics en el model  $(1 - L)^d(1 - 0,9L)X_t = \varepsilon_t$ , amb  $T=100$ .

$d$	$m$	$GPH$	$q=0.2$	$q=0.4$	$q=0.6$	$q=0.8$	$GPHA$	$GPHB$
0.00	5	0.66	0.26	0.28	0.17	0.21	2.80	0.33
	15	0.64	0.10	0.15	0.17	0.18	0.60	0.04
	25	0.72	0.15	0.18	0.20	0.20	0.60	0.02
	35	0.75	0.18	0.20	0.21	0.21	0.64	0.01
0.05	5	0.61	0.28	0.34	0.15	0.17	2.88	0.31
	15	0.64	0.08	0.12	0.14	0.15	0.55	0.03
	25	0.71	0.12	0.15	0.16	0.17	0.59	0.02
	35	0.74	0.14	0.16	0.17	0.17	0.64	0.01
0.10	5	0.68	0.28	0.27	0.13	0.16	2.58	0.33
	15	0.64	0.06	0.09	0.11	0.12	0.57	0.03
	25	0.70	0.10	0.12	0.13	0.13	0.59	0.01
	35	0.73	0.12	0.13	0.14	0.14	0.62	0.01
0.15	5	0.67	0.30	0.33	0.17	0.11	3.11	0.33
	15	0.63	0.05	0.07	0.08	0.09	0.59	0.03
	25	0.71	0.07	0.09	0.10	0.10	0.59	0.02
	35	0.73	0.09	0.10	0.11	0.11	0.64	0.01
0.20	5	0.68	0.33	0.36	0.13	0.16	3.20	0.35
	15	0.61	0.04	0.05	0.06	0.07	0.54	0.04
	25	0.68	0.05	0.07	0.07	0.08	0.58	0.02
	35	0.71	0.07	0.08	0.08	0.08	0.62	0.02
0.25	5	0.70	0.32	0.31	0.10	0.09	2.81	0.37
	15	0.63	0.02	0.04	0.04	0.05	0.58	0.05
	25	0.68	0.04	0.05	0.05	0.05	0.60	0.03
	35	0.71	0.05	0.05	0.05	0.06	0.62	0.02
0.30	5	0.61	0.38	0.27	0.10	0.06	2.97	0.34
	15	0.60	0.02	0.02	0.03	0.03	0.57	0.06
	25	0.65	0.02	0.03	0.03	0.03	0.57	0.04
	35	0.67	0.03	0.03	0.03	0.04	0.60	0.04
0.35	5	0.68	0.43	0.30	0.09	0.07	2.90	0.39
	15	0.59	0.01	0.01	0.01	0.02	0.55	0.08
	25	0.62	0.01	0.02	0.02	0.02	0.56	0.05
	35	0.64	0.01	0.02	0.02	0.02	0.58	0.05
0.40	5	0.61	0.53	0.37	0.11	0.07	2.70	0.44
	15	0.51	0.01	0.00	0.00	0.01	0.50	0.11
	25	0.56	0.00	0.01	0.01	0.01	0.49	0.08
	35	0.57	0.01	0.01	0.01	0.01	0.52	0.08
0.45	5	0.53	0.67	0.32	0.10	0.08	2.66	0.42
	15	0.44	0.02	0.00	0.00	0.00	0.43	0.15
	25	0.46	0.00	0.00	0.00	0.00	0.43	0.13
	35	0.48	0.00	0.00	0.00	0.00	0.44	0.12



Quadre A7.27. EQM dels estimadors semiparamètrics en el model  $(1 - L)^d(1 + 0,5L)X_t = \varepsilon_t$ , amb  $T=100$ .

$d$	$m$	$GPH$	$q=0.2$	$q=0.4$	$q=0.6$	$q=0.8$	$GPHA$	$GPHB$
0.00	5	0.40	0.12	0.65	0.45	0.64	2.78	0.40
	15	0.06	0.10	0.06	0.09	0.14	0.20	0.06
	25	0.03	0.05	0.05	0.06	0.12	0.09	0.03
	35	0.04	0.05	0.08	0.14	0.28	0.06	0.02
0.05	5	0.37	0.15	0.68	0.46	0.65	2.79	0.37
	15	0.06	0.10	0.06	0.07	0.13	0.20	0.06
	25	0.04	0.05	0.05	0.07	0.13	0.09	0.03
	35	0.05	0.05	0.08	0.14	0.27	0.06	0.02
0.10	5	0.41	0.15	0.64	0.45	0.52	2.54	0.41
	15	0.06	0.10	0.06	0.06	0.10	0.20	0.06
	25	0.04	0.05	0.05	0.06	0.10	0.09	0.03
	35	0.05	0.05	0.08	0.12	0.24	0.07	0.02
0.15	5	0.36	0.16	0.58	0.40	0.48	2.72	0.36
	15	0.05	0.09	0.05	0.06	0.09	0.19	0.05
	25	0.03	0.04	0.04	0.06	0.10	0.08	0.03
	35	0.04	0.05	0.07	0.12	0.23	0.06	0.02
0.20	5	0.43	0.18	0.74	0.42	0.53	3.03	0.43
	15	0.06	0.10	0.05	0.06	0.09	0.20	0.06
	25	0.03	0.04	0.04	0.05	0.08	0.09	0.03
	35	0.04	0.05	0.06	0.11	0.21	0.06	0.02
0.25	5	0.38	0.19	0.59	0.34	0.47	2.93	0.38
	15	0.05	0.09	0.05	0.05	0.07	0.19	0.05
	25	0.03	0.04	0.04	0.05	0.07	0.09	0.03
	35	0.04	0.05	0.06	0.10	0.19	0.06	0.02
0.30	5	0.35	0.25	0.58	0.40	0.39	2.93	0.35
	15	0.06	0.10	0.05	0.05	0.07	0.20	0.05
	25	0.03	0.05	0.04	0.05	0.08	0.09	0.03
	35	0.04	0.05	0.06	0.10	0.17	0.06	0.02
0.35	5	0.41	0.30	0.63	0.34	0.38	2.95	0.41
	15	0.06	0.10	0.05	0.05	0.07	0.20	0.06
	25	0.03	0.05	0.04	0.05	0.07	0.09	0.03
	35	0.04	0.05	0.06	0.09	0.15	0.07	0.02
0.40	5	0.40	0.42	0.66	0.34	0.30	2.63	0.40
	15	0.06	0.11	0.06	0.06	0.06	0.20	0.05
	25	0.03	0.06	0.04	0.05	0.06	0.09	0.03
	35	0.04	0.06	0.06	0.09	0.15	0.07	0.03
0.45	5	0.37	0.64	0.61	0.30	0.31	2.76	0.37
	15	0.05	0.10	0.06	0.05	0.06	0.19	0.05
	25	0.03	0.05	0.05	0.05	0.06	0.09	0.03
	35	0.04	0.05	0.06	0.08	0.12	0.06	0.04

Quadre A7.28. EQM dels estimadors semiparamètrics en el model  $(1 - L)^d(1 + 0,9L)X_t = \varepsilon_t$ , amb  $T=100$ .

$d$	$m$	$GPH$	$q=0.2$	$q=0.4$	$q=0.6$	$q=0.8$	$GPHA$	$GPHB$
0.00	5	0.37	0.12	0.60	0.43	0.64	2.75	0.37
	15	0.06	0.10	0.06	0.09	0.14	0.19	0.06
	25	0.04	0.05	0.05	0.07	0.14	0.09	0.03
	35	0.06	0.07	0.11	0.21	0.44	0.06	0.02
0.05	5	0.38	0.14	0.71	0.48	0.66	2.93	0.38
	15	0.06	0.10	0.06	0.08	0.13	0.20	0.06
	25	0.04	0.05	0.06	0.08	0.15	0.09	0.03
	35	0.06	0.07	0.11	0.20	0.42	0.06	0.02
0.10	5	0.41	0.15	0.65	0.44	0.50	2.54	0.41
	15	0.06	0.10	0.06	0.06	0.10	0.20	0.05
	25	0.04	0.05	0.05	0.07	0.12	0.09	0.03
	35	0.06	0.07	0.10	0.18	0.38	0.06	0.02
0.15	5	0.37	0.16	0.60	0.40	0.47	2.77	0.37
	15	0.05	0.09	0.05	0.06	0.10	0.19	0.05
	25	0.04	0.05	0.05	0.07	0.12	0.09	0.03
	35	0.06	0.06	0.10	0.17	0.36	0.06	0.02
0.20	5	0.43	0.18	0.73	0.43	0.54	2.99	0.43
	15	0.06	0.10	0.05	0.07	0.10	0.20	0.06
	25	0.04	0.05	0.05	0.06	0.10	0.09	0.03
	35	0.06	0.06	0.09	0.16	0.34	0.07	0.02
0.25	5	0.37	0.19	0.58	0.34	0.46	2.92	0.37
	15	0.05	0.09	0.05	0.05	0.07	0.19	0.05
	25	0.04	0.05	0.04	0.06	0.09	0.08	0.03
	35	0.05	0.06	0.09	0.15	0.30	0.06	0.02
0.30	5	0.34	0.25	0.58	0.40	0.39	2.82	0.34
	15	0.05	0.11	0.06	0.05	0.07	0.20	0.05
	25	0.04	0.05	0.05	0.06	0.10	0.09	0.03
	35	0.06	0.06	0.09	0.14	0.27	0.06	0.02
0.35	5	0.40	0.30	0.66	0.34	0.38	2.96	0.40
	15	0.06	0.10	0.06	0.06	0.07	0.20	0.06
	25	0.04	0.06	0.05	0.05	0.08	0.10	0.03
	35	0.06	0.06	0.08	0.13	0.24	0.07	0.02
0.40	5	0.39	0.42	0.66	0.35	0.30	2.61	0.39
	15	0.06	0.11	0.07	0.06	0.07	0.20	0.06
	25	0.04	0.06	0.05	0.06	0.08	0.09	0.03
	35	0.06	0.07	0.08	0.13	0.23	0.07	0.02
0.45	5	0.37	0.63	0.63	0.31	0.31	2.72	0.37
	15	0.05	0.11	0.06	0.06	0.06	0.19	0.05
	25	0.03	0.06	0.05	0.06	0.07	0.09	0.03
	35	0.05	0.07	0.08	0.12	0.20	0.06	0.02



**TERCERA PART**  
**CONSEQÜÈNCIES ÈCONOMÈTRIQUES**



## **Capítol 8**

# **CONSEQÜÈNCIES DEL GRAU DE MEMÒRIA PER L'ANÀLISI DE DADES TEMPORALS ECONÒMIQUES**



## 8.1. Introducció

Fins al moment, la tesi s'ha centrat únicament en dos aspectes relacionats amb el grau de memòria de les sèries temporals. D'una banda, en la definició precisa de les diferents categories de memòria en que es poden dividir les dades temporals i en les característiques principals associades a cada categoria. I, d'altra banda, en l'anàlisi de diferents possibilitats disponibles per a la modelització i estimació més o menys explícita de les dades que contenen memòria llarga, ja que per la modelització de dades amb memòria curta o bé amb memòria permanent ja es disposa de tècniques força conegudes.

Però fins al moment, l'única justificació que s'ha donat al fet de centrar la tesi en el concepte de memòria és l'evidència que certes característiques habituals de les dades temporals econòmiques no es corresponen amb les característiques dels models usats habitualment. De totes maneres, pensem que aquest no és motiu suficient per dedicar un esforç tan important a la recerca sobre aquest tema ja que, en general, en qualsevol àmbit de l'anàlisi estadística aplicada, cal usar models que no són més que una simplificació extrema de la realitat. Per exemple, la utilització dels models ARMA no pressuposa en cap moment que les dades observades han estat generades segons un procés d'aquest tipus. De fet, en el cas de les dades econòmiques, aquest és un supòsit clarament simplificador ja que les dades observades són el resultat de la combinació d'un gran nombre de factors.

Per tant, la justificació de la utilització, tant dels models ARMA o ARIMA com dels models de memòria llarga, es basa en la seva utilitat sempre que la simplificació que implica no afecti a les conclusions que s'obtenen per la seva utilització. Així, tot i que sota presència de memòria llarga els models ARMA no siguin bones aproximacions a nivell teòric del procés generador de les dades, abans d'entrar a utilitzar altres eines més sofisticades cal comprovar que els resultats d'usar aquestes noves eines són millors i que la no consideració del grau de memòria en les dades econòmiques pot portar a conclusions errònies. Dit d'una altra manera,



cal analitzar la sensibilitat de les eines més habituals front a errors d'especificació en la memòria.

En aquest sentit creiem que el diagnòstic sobre la necessitat de considerar els aspectes teòrics presentats al llarg de la tesi no pot ser general, sinó que per cada tipus d'ús que es faci de les dades temporals es poden obtenir diagnòstics particulars. Així, per exemple, les conseqüències que pot tenir obviar el tipus de memòria són ben diferents si l'objectiu és el de calcular prediccions o bé si l'objectiu és estimar un model de regressió lineal entre dos variables econòmiques temporals.

Aquesta diferent valoració de les tècniques segons quin sigui l'objectiu és un aspecte que sovint s'oblida en l'anàlisi econòmica aplicada i així, per exemple, pel fet d'usar un model  $ARIMA(p, I, q)$  per calcular prediccions d'una certa variable, s'infereix que els shocks tenen un efecte permanent en aquesta variable quan per arribar a aquest tipus de conclusió fora necessària una anàlisi més acurada.

Per tant, aquesta última part de la tesi es centra precisament en l'anàlisi d'aquestes conseqüències, és a dir, analitzar quan és important tenir en compte en l'anàlisi aplicada de sèries econòmiques els diferents tipus de memòria i quines tècniques cal usar per evitar els problemes que es poden presentar si s'apliquen les tècniques habituals.

De fet, aquestes conseqüències són tantes i tan variades que l'anàlisi detallada de totes elles supera clarament els objectius de la tesi. Per aquest motiu, no es pretén ser exhaustiu en els propers apartats sinó que es presenten algunes de les possibles conseqüències de la no consideració de la memòria llarga que creiem poden ser importants. Concretament, es posa especial èmfasi en quines eines estadístiques i econòmiques són més sensibles a la presència de memòria llarga, i es presenten alguns instruments novedosos per evitar aquestes conclusions errònies.

Òbviament, una primera distinció molt important d'aquestes conseqüències ve donada per l'àmbit de l'anàlisi. Si s'analitza una certa variable temporal de manera independent d'altres variables les possibles conseqüències de no considerar la memòria llarga estan associades a l'especificació d'un model erroni mentre que en un àmbit d'anàlisi multivariant la problemàtica és molt més complexa. El primer tipus de conseqüències serà objecte d'aquest capítol 8 mentre que al capítol 9 es presenta una introducció dels problemes d'inferència en models de regressió quan es considera la possibilitat que les variables implicades continguin memòria llarga.

Si hom centra l'anàlisi en una sola sèrie econòmica, entre els molts resultats que es poden obtenir a partir d'aquestes dades, com a mínim n'hi ha tres que semblen, almenys a nivell teòric especialment sensibles al tipus de memòria de les dades. Es tracta de: *a)* la interpretació que es faci del grau de persistència temporal que tenen els shocks en aquesta variable; *b)* de la utilització d'alguna tècnica, ja sigui paramètrica o no paramètrica per calcular prediccions; i *c)* de l'estimació d'algun tipus de senyal no observable com pot ser un senyal cicle-tendència per a la seva utilització en l'anàlisi de conjuntura econòmica o el perfil cíclic per analitzar les característiques cícliques d'una determinada economia o sector econòmic.

L'objectiu d'aquest capítol és valorar en quines situacions el fet de no considerar la possible memòria llarga pot portar a errors importants en les conclusions d'aquests tipus d'anàlisi. Per fer possible aquesta anàlisi a un nivell relativament senzill es parteix del supòsit que les dades observades s'han generat segons un model ARMA, ARIMA o ARFIMA<sup>1</sup>. Per tant, la hipòtesi de partida és que les dades observades han estat generades, un cop s'han eliminat tots els components deterministes, a través d'un model del tipus següent:

$$\phi(L)(1-L)^d X_t = \theta(L)\varepsilon_t \quad (8.1)$$

<sup>1</sup> Recordi's que els models ARMA i ARIMA són de fet casos particulars dels models ARFIMA i per tant, es suposa que les dades s'han generat segons un model ARFIMA. A les pàgines següents es repeteix sovint aquesta redundància per destacar la triple possibilitat entre models ARMA que sols contenen memòria curta, models ARFIMA amb  $d < 1$  que contenen memòria llarga i models ARIMA que contenen memòria permanent.

on es suposa que  $d \in [0, 2]$ .

S'ha seleccionat aquesta especificació perquè el model (8.1) conté com a casos particulars molts dels models considerats al llarg d'aquesta tesi. Concretament, si  $d=0$  el model (8.1) queda reduït a un model ARMA( $p, q$ ) on  $p$  i  $q$  són els ordres dels polinomis  $\phi(L)$  i  $\theta(L)$  respectivament mentre que si  $d=1$  o  $d=2$ , es tracta d'un model ARIMA( $p, 1, q$ ) o ARIMA( $p, 2, q$ ). Quan el paràmetre  $d$  no sigui un número sencer, el procés definit a (8.1) és estacionari amb memòria llarga quan  $d < 1/2$  i és no estacionari quan  $d \geq 1/2$ .<sup>3</sup>

En aquest context té interès considerar i comparar dues vies d'anàlisi basades en estratègies diferents:

a) Una primera estratègia basada en la metodologia Box-Jenkins per identificar un model ARMA o ARIMA i usar aquest model per calcular prediccions, estimar components no observables o treure conclusions sobre la persistència dels shocks, sense tenir en compte la possibilitat que la sèrie tingui memòria llarga. Així, en aquesta estratègia s'imposa que el model seleccionat ha de ser del tipus  $\phi(L)X_t = \theta(L)\varepsilon_t$ , o  $\phi(L)(1-L)X_t = \theta(L)\varepsilon_t$ .

b) Una segona estratègia basada en la utilització de models ARFIMA<sup>4</sup>. En aquest cas, s'utilitza una modelització més general que inclou els models de l'estratègia anterior, però per contra, el tractament d'aquests models és força més complicat.

---

<sup>2</sup> Es podrien considerar altres valors del paràmetre  $d$ , però com ja s'ha comentat en capítols previs, no sembla raonable pensar en variables econòmiques amb valors fora d'aquest interval.

<sup>3</sup> No es consideren altres tipus de memòria llarga per facilitar l'anàlisi posterior. En tot cas es tracta d'il·lustrar les conseqüències de tractar una sèrie temporal  $X_t$  com model ARMA o ARIMA quan aquest pot comportar-se com un ARFIMA.

<sup>4</sup> Recordi's que no és un problema gens senzill detectar el tipus de memòria d'una sèrie i, front a aquesta dificultat el segon analista usa sempre el model més general, és a dir, un model ARFIMA.

En aquest context, interessa analitzar en quins casos el primer analista pot cometre errors greus degut a que sols considera models ARMA o ARIMA. Hom pot pensar en una primera aproximació que sempre que les dades s'hagin generat segons un model ARFIMA amb  $d < 1$  el primer analista pot cometre errors greus, més greus com més diferent de zero i de la unitat sigui el paràmetre  $d$  i que per tant, si sols es tenen en compte els errors de selecció del model és millor l'aproximació en que es basa la segona estratègia.

Però s'han comentat ja extensament al llarg dels capítols anteriors les dificultats per a la identificació i l'estimació de models ARFIMA de manera que també s'han de considerar els errors d'estimació. Quan es tenen en compte els dos aspectes, ja no és clar que sigui sempre millor usar la segona estratègia. Així, per exemple, podria ser millor usar un model ARMA seguint el principi de parsimònia per calcular prediccions, tot i que les dades s'hagin generat segons un model ARFIMA, ja que s'obtenen millors estimacions dels paràmetres d'un model ARMA que d'un model ARFIMA<sup>5</sup>.

Seguint aquesta plantejament, s'estructura aquest capítol de la següent manera. En primer lloc, s'analitza quins son els resultats d'aplicar les eines habituals de la metodologia Box-Jenkins a dades que contenen memòria llarga, concretament models ARFIMA. Posteriorment, es centra l'anàlisi en les conseqüències concretes que tenen tant els errors de selecció del model com els errors d'estimació en l'anàlisi de la persistència, en el càlcul de prediccions i en l'estimació de components no observables.

---

<sup>5</sup> S'ha de tenir en compte que a l'anàlisi aplicada de sèries temporals hi ha una tercera font d'error a més de les anteriors derivada del fet que tant els models ARMA o ARIMA com els models ARFIMA no són més que una aproximació al vertader procés generador de les dades. De totes maneres, aquest és un error que cal tenir en compte a nivell conceptual però que no és possible analitzar, almenys sense fer supòsits addicionals sobre el procés generador de dades.

## 8.2. Resultat d'aplicar la metodologia Box-Jenkins a dades amb memòria llarga

A partir del treball de Box i Jenkins (1970), es disposa d'un marc adequat per a la modelització de gran varietat de sèries econòmiques a través dels models ARIMA. Aquesta metodologia s'utilitza extensament a l'anàlisi econòmica i és tremendament popular. Entre les aplicacions d'aquesta metodologia destaca la facilitat en el càlcul de prediccions òptimes i, de manera més recent, la possibilitat d'obtenir els filtres òptims per a l'estimació de components no observables.

Com és ben conegut, la filosofia d'aquesta metodologia es basa en el principi de parsimònia. Tot i que durant els anys seixanta s'havien desenvolupat models macroeconòmics molt complexes amb centenars de variables i d'equacions de comportament, en treballs com el de Nelson (1972) es mostra que amb models  $ARMA(p,d)$  d'ordres  $p$  i  $q$  petits es poden obtenir millors prediccions. La proposta de Box i Jenkins es tradueix en l'aplicació de quatre etapes:

1. Transformar les dades, si cal, per garantir que siguin estacionàries de segon ordre
2. Identificar de manera tentativa els ordres  $p$  i  $q$  del model a usar.
3. Estimar els paràmetres AR i MA del model  $ARMA(p,q)$ .
4. Usar elements de diagnòstic per determinar si el model usat és adequat o no.

Òbviament, com s'ha comentat ja al llarg de la tesi la utilització de models ARMA pressuposa que les dades contenen sols memòria curta. Cal analitzar doncs els problemes que pot presentar aquesta metodologia sota presència de memòria llarga. Concretament s'analitzen els següents aspectes:

1. El tipus de transformació usada, concretament l'ordre de diferenciació detectat amb els tests d'arrels unitàries habituals.
2. Els problemes en la identificació del model.
3. Els resultats de l'estimació del model identificat.

4. Cal analitzar si hi ha algun element de diagnòstic que permeti detectar els errors. S'analitzen els contrastos habituals de Box i Pierce (1970) i Ljung i Box (1978) a més de presentar algunes propostes alternatives a partir del periodograma presentades al capítol 5 de la tesi.

A continuació, es desenvolupen cada un d'aquests quatre aspectes.

### **8.2.1. Selecció de l'ordre d'integració de les variables econòmiques**

Una primera decisió en la metodologia habitual consisteix a escollir l'ordre d'integració de la variable que, en dades econòmiques no acostuma a ser superior a  $d=2$ . Per a aquesta decisió es disposa d'una àmplia varietat de contrastos coneguts sota el nom genèric de *contrastos d'arrels unitàries*<sup>6</sup>.

L'acció de diferenciar o no una sèrie temporal té importants implicacions sobre els resultats posteriors. Així, si es diferencia una sèrie que no conté una arrel unitària s'estarà sobrediferenciant i apareixerà en la representació mitjana mòbil del procés una arrel unitària no invertible. Observi's que en el context de models ARFIMA aquest és un risc evident ja que no hi ha cap raó que assigni una major probabilitat als valors  $d=0$  o  $d=1$  que a altres valors.

Però prèviament a quines conseqüències pot tenir aquesta sobrediferenciació, cal fixar-se en quins resultats s'obtenen de l'aplicació dels contrastos d'arrels unitàries a sèries integrades fraccionalment. Noti's que en aquest context hi ha una gran varietat de possibilitats. En tot cas fóra bo, com a mínim, que si realment  $0 < d < 1/2$  el resultat del test fos acceptar que  $d=0$ , mentre que si  $1/2 \leq d < 1$  el resultat fos  $d=1$ . En aquests casos, tot i l'error, s'encerta el caràcter d'estacionarietat o no estacionarietat de les dades.

---

<sup>6</sup> Vegi's Suriñach *et al.* (1996) per una introducció als tests d'arrels unitàries.

Alguns dels tests d'arrels unitàries més populars estan basats en la regressió següent<sup>7</sup>:

$$X_t = \rho X_{t-1} + \varepsilon_t \quad (8.2)$$

Així, el test de Dickey i Fuller (DF) (1979) es basa en contrastar si  $\rho=1$  (i per tant, la sèrie  $X_t$  és integrada) o bé si  $\rho < 1$ , que en el context habitual porta a acceptar que  $X_t$  és estacionària.

En aquest context és fonamental el resultat de Phillips (1987) que demostra que si  $X_t = X_{t-1} + \varepsilon_t$  i  $X_0=0$  llavors l'estimador MQO del paràmetre  $\rho$  té la següent distribució asimptòtica:

$$T(\tilde{\rho} - 1) \Rightarrow \frac{(1/2)\{W(1)^2 - 1\}}{\int_0^1 W(t)^2 dt} \quad (8.3)$$

Suposi's ara un cas més general en que  $X_t \sim FI(d)$  on  $d$  pot ser diferent de la unitat. En aquest cas s'obté  $\varepsilon_t = (X_t - X_{t-1}) \sim FI(d-1)$ . Si es defineix  $\delta = d-1 \neq 0$  i s'estima per MQO la regressió anterior per contrastar si les dades contenen una arrel unitària, llavors ja no té perquè ser cert el resultat de Phillips (1987).

De fet, Sowell (1990) demostra<sup>8</sup> quina és la distribució límit de  $(\tilde{\rho} - 1)$  si  $X_t = X_{t-1} + \varepsilon_t$ ,  $X_0=0$  i  $\varepsilon_t \sim I(\delta)$  amb  $-1/2 < \delta < 1/2$ . Concretament, si es descomposa aquest terme com  $(\tilde{\rho} - 1) = A_\delta + B_\delta$  on:

<sup>7</sup> La distribució de l'estimador de  $\rho$  depèn de la presència d'un terme constant o d'una tendència determinista en l'especificació de l'equació del contrast. Així, cal considerar a la pràctica altres possibilitats:

$$X_t = \mu + \rho X_{t-1} + \varepsilon_t$$

o:

$$X_t = \mu + \beta t + \rho X_{t-1} + \varepsilon_t$$

però a efectes teòrics es centra l'anàlisi sota la hipòtesi d'absència de components deterministes.

<sup>8</sup> Per arribar a aquest resultat no és aplicable la teoria desenvolupada per Phillips (1987) ja que sota integració fraccional el creixement de les sumes parcials,  $S_T = \sum X_t$ , depèn del paràmetre  $d$  però a Sowell (1990) es demostra, usant el Teorema funcional de Davydov (1970) i Taqqu (1975) que, sota les hipòtesis següents:

- (i)  $X_t \sim I(d)$  amb  $-1/2 < d < 1/2$  i esperança zero.
- (ii)  $E|\varepsilon_t|^r < \infty$  per  $r \geq \max\{4, -8d/(1+2d)\}$

$$A_\delta = \frac{\frac{1}{2} X_T^2}{\sum_{t=2}^T X_{t-1}^2} \quad \text{i} \quad B_\delta = \frac{-\frac{1}{2} \sum_{t=1}^T \varepsilon_t^2}{\sum_{t=2}^T X_{t-1}^2} \tag{8.4}$$

tenen ordres de convergència diferents i es compleixen les següents lleis de convergència en distribució:

$$TA_\delta \Rightarrow \frac{1/2 [W_\delta(1)]^2}{\int_0^1 [W_\delta(s)]^2 ds} \tag{8.5}$$

i

$$T^{1+2\delta} B_\delta \Rightarrow \frac{\left[ \frac{1}{2} + \delta \right] \frac{\Gamma(1+\delta)}{\Gamma(1-\delta)}}{\int_0^1 [W_\delta(s)]^2 ds} \tag{8.6}$$

Si  $\delta > 0$ , el terme que depèn de  $B_\delta$  convergeix a zero i la distribució asimptòtica de  $T(\tilde{\rho} - 1)$  és la distribució asimptòtica de  $TA_\delta$ . Observi's, en particular, que  $A_\delta$  és una variable aleatòria positiva. En canvi, si  $\delta < 0$ , llavors el terme que depèn de  $A_\delta$  convergeix a zero i la distribució asimptòtica de  $T^{1+2\delta}(\tilde{\rho} - 1)$  és la de  $T^{1+2\delta}B_\delta$ , que és una variable aleatòria negativa i a més té un ordre de convergència diferent. Quan  $\delta = 0$ , intervenen els dos termes i el suport de la distribució és tota la recta real.

Un altre resultat interessant és que l'estimador MQO és consistent per qualsevol valor de  $\delta$  sempre que compleixi  $-1/2 < \delta < 1/2$  però de manera sorprenent, no és cert que a mesura que augmenta l'ordre d'integració, també augmenta el rati de convergència. Recordi's que si  $X_t \sim I(0)$ , l'estimador  $\tilde{\rho}$  convergeix amb velocitat  $T^{1/2}$ , mentre que si  $X_t \sim I(1)$ , ho fa amb ordre

es demostra la següent llei de convergència en distribució:

$$\sigma_T^{-1} S_{[T\delta]} \Rightarrow W_d(s)$$

on  $W_d(\cdot)$  és un moviment brownià fraccional definit per Mandelbrot i Van Ness (1968) com (per  $-0.5 < d < 0.5$ ):

$$W_d(t) = \frac{1}{\Gamma(d+1)} \int_0^t (t-x)^d dW(x)$$



$T$ . Però dels resultats anteriors es dedueix que si una sèrie és  $FI(d)$  amb  $1 \leq d < 3/2$  l'estimador convergeix a un rati  $T$ , el mateix per qualsevol  $d$ , mentre que si  $1/2 < d \leq 1$ , l'estimador convergeix a un rati  $T^{2d-1}$ . Per tant, si  $1/2 < d < 3/4$ , la convergència és molt lenta, amb rati menor a  $T^{1/2}$ .

Per tant, si  $X_t \sim FI(d)$  amb  $d$  proper a  $1/2$ , la convergència cap a la seva distribució límit de l'estadístic de contrast és tant lenta que és d'esperar que el test de Dickey i Fuller presenti resultats molt pobres amb sèries no gaire llargues.

Aquests resultats mostren l'enorme dificultat que té contrastar la presència d'una arrel unitària si hi ha la possibilitat de presència de memòria llarga. Hom podria pensar en usar com alternativa el  $t$ -estadístic, però el mateix Sowell (1990) mostra com també aquest presenta greus problemes. Concretament, si:

$$t = \frac{\sum_{i=2}^T X_{i-1} \varepsilon_i}{S \left( \sum_{i=2}^T X_{i-1}^2 \right)^{1/2}} \quad (8.7)$$

amb:

$$S^2 = T^{-1} \sum_{i=2}^T (X_i - \tilde{\rho} X_{i-1})^2 \quad (8.8)$$

llavors, sota les mateixes hipòtesis dels teoremes anteriors,  $t \rightarrow \infty$  quan  $d < 1$  i  $t \rightarrow \infty$  quan  $d > 1$ . En definitiva, el  $t$ -estadístic sols convergeix pel cas  $d=1$ .

De totes formes, tot i aquestes dificultats, el test DF d'arrels unitàries és consistent contra hipòtesis alternatives del tipus  $FI(d)$ , però presenta problemes quan les sèries són curtes. A Diebold i Rudebusch (1991) comparen a través d'una simulació el comportament del test de Dickey i Fuller a partir de dos processos generadors de dades diferents:

$$A) (1 - \phi L)X_t = \varepsilon_t$$

$$B) (1 - L)^d X_t = \varepsilon_t$$

Noti's com els casos  $\phi=1$  i  $d=1$  corresponen a un camí aleatori i per tant a un procés  $I(1)$  mentre que els casos  $\phi=0$  i  $d=0$  corresponen a un soroll blanc i per tant  $I(0)$ . Al comparar el comportament del test DF en ambdós situacions la principal conclusió que s'obté és que la potència augmenta molt més lentament en el model B a mesura que el paràmetre s'allunya de la unitat que quan el paràmetre autoregressiu s'allunya de la unitat en el model A. Així, el test DF té propietats molt pobres a l'hora de distingir entre la hipòtesi nul·la  $I(1)$  contra hipòtesis  $FI(d)$ .

A la Taula 8.1 es pot comprovar aquesta diferència a partir dels resultats d'una simulació en que s'han generat 5000 sèries temporals de tamany  $T=100$  segons els models A i B per diferents valors dels paràmetres<sup>9</sup>  $\phi$  i  $d$ . En efecte, si s'utilitza aquest mètode per contrastar  $H_0:\phi=1$  contra  $H_A:\phi<1$  s'obté un contrast amb una gran potència, almenys quan la hipòtesi alternativa és soroll blanc. En canvi, si s'utilitza aquest mètode per contrastar  $H_0:d=1$  contra  $H_A:d<1$  (objectiu pel que no ha estat dissenyat) es converteix en un contrast amb molt poca potència.

De totes maneres, els resultats de la taula 8.1 suggereixen que fora possible usar aquest procediment per contrastar  $H_0:d=1$  contra  $H_A:d<0.5$ , convertint-se així en un contrast d'estacionarietat. Recordi's en relació amb això que una de les conclusions importants del capítol 5 de la tesi és la de que el correlograma és un bon instrument per detectar la presència de memòria permanent.

---

<sup>9</sup> Els casos aquí considerats inclouen els de Diebold i Rudebusch (1991) però no coincideixen ja que en aquell treball s'aplica el contrast DF a dues cues, és a dir  $H_0:\rho=1$  contra  $H_1:\rho\neq 1$  mentre que aquí s'ha usat a una cua, com és més habitual, és a dir  $H_0:\rho=1$  contra  $H_1:\rho< 1$ .

En aquest sentit, al test DF li passa una cosa similar ja que tot i que no és capaç de distingir  $d=1$  de  $d<1$ , si que permet detectar la presència de memòria permanent. Novament, per tant, cal interpretar el no rebuig de la hipòtesi nul·la del test DF com la necessitat de prendre diferències de la variable per convertir-la en estacionària, sense que això sigui senyal de que els shocks de la variable tinguin efectes permanents en el futur ja que és molt probable que  $d<1$ .

Quadre 8.1. Potència del test DF sota hipòtesis alternatives AR(1) o FI( $d$ ).  $T=100$ .

$\phi$	Potència	$d$	Potència
0.05	1.00	0.05	1.00
0.15	1.00	0.15	1.00
0.25	1.00	0.25	1.00
0.35	1.00	0.35	1.00
0.45	1.00	0.45	0.99
0.55	1.00	0.55	0.98
0.65	1.00	0.65	0.85
0.75	1.00	0.75	0.60
0.85	0.68	0.85	0.29
0.95	0.12	0.95	0.11

De totes maneres, aquest és un procediment dissenyat quan la hipòtesi alternativa és soroll blanc. En general, per solucionar els problemes d'autocorrelació, s'han proposat a la literatura dos correccions del contrast DF, una de paramètrica proposada per Dickey i Fuller (1979) coneguda amb el nom de contrast de Dickey i Fuller ampliat (ADF) i una correcció no paramètrica proposada per Phillips i Perron (PP) (1988).

El test ADF es realitza a través de la següent regressió auxiliar:

$$X_t = \mu + \beta t + \rho X_{t-1} + \sum_{i=1}^k \phi_i (1-L) X_{t-i} + u_t \quad (8.9)$$

on  $k$  és el nombre de retards usats per corregir la possible autocorrelació. En aquesta regressió es contrasta:

$$\begin{cases} H_0: \rho - 1 = 0 \\ H_1: \rho - 1 < 0 \end{cases}$$

de manera que sota la hipòtesi nul·la,  $X_t$  és I(1) i si es rebutja aquesta hipòtesi,  $X_t$  és I(0).

El test PP és una correcció no paramètrica del test DF que es basa en l'estadístic:

$$Z = T(\tilde{\rho} - 1) - \frac{1}{2}(\tilde{\sigma}_l^2 - \tilde{\sigma}_u^2)T^2 \left[ \sum_{l=2}^T (X_{l-1} - \bar{X})^2 \right]^{-1} \quad (8.10)$$

calculat a partir de l'estimació del model de regressió següent:

$$X_t = \mu + \beta t + \rho X_{t-1} + u_t \quad (8.11)$$

i on  $\tilde{\sigma}_u^2$  és la variància mostral del terme de perturbació  $u_t$ , i  $\tilde{\sigma}_l^2$  és una estimació consistent d'aquesta variància a partir de la finestra de Bartlett:

$$\tilde{\sigma}_l^2 = \frac{1}{T} \sum_{i=1}^n e_i^2 + \frac{2}{T} \sum_{j=1}^l \left(1 - \frac{j}{l-1}\right) \sum_{i=j+1}^T e_i e_{i-j} \quad (8.12)$$

Un estudi realitzat per Hassler i Wolters (1994) mostra que el test de Phillips i Perron es comporta de manera similar al DF i que el test ADF és el que presenta un pitjor comportament, sobretot a mesura que augmenta el nombre de diferències usades.

Hom pot argumentar que aquests contrastos són poc adequats per conèixer si una sèrie econòmica és integrada d'ordre 1 en el context de la integració fraccional perquè moltes d'aquestes sèries són no estacionàries de manera que en el context dels models FI( $d$ ) tenen un paràmetre d'integració superior a 1/2 i que fora millor usar un contrast on la hipòtesi nul·la sigui la de que la sèrie és I(0) contra alternativa I(1). Quan es vol contrastar aquestes hipòtesis és habitual usar la proposta de Kwiatkowski, Phillips, Smith i Shin (1992), basat en l'estimació del model:

$$X_t = \mu + \beta t + \rho X_{t-1} + u_t \quad (8.13)$$

Si es vol contrastar  $H_0: \rho=0$  contra  $H_1: \rho=1$  aquests autors demostren que el contrast dels multiplicadors de Lagrange es basa en l'estadístic:

$$LM = \left( \frac{1}{T} \sum_{i=1}^T e_i^2 \right)^{-1} \sum_{i=1}^T S_i^2 \quad (8.14)$$

on  $S_j$  és la suma parcial del residus  $e_t$ . Per evitar problemes d'autocorrelació, s'utilitza també un estimador consistent de la variància del terme de pertorbació igual que en el test PP:

$$\tilde{\sigma}_I^2 = \frac{1}{T} \sum_{t=1}^n e_t^2 + \frac{2}{T} \sum_{j=1}^l \left(1 - \frac{j}{l-1}\right) \sum_{t=j+1}^T e_t e_{t-j} \quad (8.15)$$

de manera que l'estadístic és:

$$LM = (\tilde{\sigma}_I^2)^{-1} \sum_{t=1}^T S_t^2 \quad (8.16)$$

Lee i Schmidt (1996) analitzen el comportament del test KPSS que està dissenyat en principi per contrastar la hipòtesi nul·la  $I(0)$  contra  $I(1)$ . El resultat és que el KPSS és consistent contra qualsevol hipòtesi  $I(d)$  de manera que permet distingir la presència de memòria curta i de memòria llarga. Els mateixos autors mostren que sota hipòtesis  $I(d)$ , l'estadístic en que es basa aquest test convergeix a funcions de moviments brownians que són extensions naturals dels moviments brownians usats per MacNeill (1978) i Schmidt i Phillips (1992). Per altra banda, Lo (1991) mostra en un exercici de simulació mostra que la potència d'aquest test és similar a la de l'estadístic MRR analitzat a l'apartat 7.6 de la tesi, per distingir el comportament  $I(0)$  del comportament  $I(d)$ .

En tot cas, per poder comparar aquests contrastos en mostra finita, s'ha realitzat un exercici de simulació a partir de 5000 sèries generades a partir de models ARFIMA( $l, d, l$ ):

$$(1-L)^d (1-\phi L) X_t = (1+\theta L) \varepsilon_t \quad (8.17)$$

amb diferents valors dels paràmetres  $d$ ,  $\phi$  i  $\theta$ . Als quadres 8.2, 8.3 i 8.4 es presenten els resultats d'aplicar els tests DF, ADF, PP i KPSS al model (8.17) amb  $\phi=\theta=0$ , és a dir, a models ARFIMA(0, $d$ ,0) i a l'Annex 8.1 es presenten alguns resultats més per altres valors de  $\phi$  i  $\theta$ .

Les principals conclusions que es poden treure d'aquesta simulació són:

1. El contrast DF té una potència que augmenta més lentament a l'entorn de models FI( $d$ ) que a l'entorn de models AR(1).

**Quadre 8.2.** Percentatge de rebutjos de la hipòtesi nul·la en els test DF, ADF, PP i KPSS.  
 Model ARFIMA(1,d,1) amb  $\phi=0$ ,  $\theta=0$ ,  $0 \leq d < 0.5$  i  $T=100$ .

	k	d=0.00	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
<b>DF</b>		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.99
<b>ADF</b>	2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.97	0.95	0.93
	4	1.00	1.00	1.00	1.00	0.99	0.97	0.91	0.77	0.73	0.68
	6	1.00	1.00	0.99	0.98	0.94	0.86	0.75	0.58	0.56	0.51
	8	0.99	0.98	0.96	0.91	0.83	0.73	0.60	0.45	0.41	0.32
	10	0.97	0.93	0.88	0.81	0.71	0.60	0.48	0.35	0.33	0.31
	12	0.91	0.85	0.78	0.69	0.59	0.49	0.39	0.29	0.28	0.26
	14	0.84	0.77	0.68	0.58	0.50	0.42	0.33	0.24	0.23	0.22
	16	0.75	0.66	0.60	0.50	0.43	0.36	0.28	0.21	0.21	0.20
<b>PP</b>	2	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.97	0.95
	4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.97	0.95
	6	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.97	0.95
	8	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.97	0.96
	10	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.98	0.96
	12	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.98	0.97
	14	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.97
	16	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.98
<b>KPSS</b>	2	0.05	0.09	0.15	0.21	0.28	0.36	0.41	0.50	0.54	0.57
	4	0.05	0.08	0.12	0.17	0.22	0.29	0.34	0.40	0.42	0.43
	6	0.05	0.07	0.10	0.15	0.19	0.24	0.29	0.34	0.35	0.36
	8	0.05	0.06	0.09	0.13	0.16	0.21	0.25	0.30	0.31	0.31
	10	0.04	0.05	0.08	0.12	0.15	0.18	0.22	0.26	0.26	0.26
	12	0.04	0.05	0.07	0.10	0.13	0.16	0.20	0.21	0.21	0.21
	14	0.04	0.05	0.06	0.09	0.11	0.14	0.17	0.17	0.17	0.17
	16	0.03	0.04	0.06	0.08	0.10	0.13	0.13	0.13	0.13	0.13

Quadre 8.3. Percentatge de rebutjos de la hipòtesi nul·la en els test DF, ADF, PP i KPSS.  
 Model ARFIMA(1,d,1) amb  $\phi=0$ ,  $\theta=0$ ,  $0.5 < d \leq 1$  i  $T=100$ .

	k	d=0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00
DF		0.98	0.93	0.85	0.75	0.60	0.42	0.29	0.19	0.11	0.06
ADF	2	0.86	0.76	0.65	0.54	0.41	0.29	0.20	0.15	0.09	0.06
	4	0.61	0.52	0.42	0.35	0.27	0.20	0.14	0.11	0.08	0.06
	6	0.46	0.39	0.32	0.26	0.20	0.16	0.12	0.10	0.07	0.06
	8	0.36	0.30	0.25	0.21	0.18	0.14	0.11	0.08	0.07	0.05
	10	0.29	0.25	0.21	0.18	0.15	0.13	0.09	0.09	0.07	0.05
	12	0.25	0.21	0.18	0.16	0.14	0.12	0.09	0.08	0.07	0.06
	14	0.21	0.19	0.17	0.15	0.13	0.11	0.09	0.08	0.07	0.06
	16	0.20	0.17	0.16	0.14	0.13	0.10	0.09	0.08	0.07	0.06
PP	2	0.92	0.85	0.75	0.65	0.50	0.35	0.24	0.17	0.10	0.06
	4	0.92	0.84	0.73	0.63	0.49	0.34	0.24	0.17	0.10	0.06
	6	0.92	0.85	0.74	0.63	0.49	0.34	0.24	0.17	0.10	0.06
	8	0.93	0.86	0.75	0.65	0.50	0.36	0.24	0.17	0.10	0.07
	10	0.93	0.87	0.76	0.66	0.52	0.36	0.25	0.17	0.11	0.06
	12	0.94	0.88	0.78	0.68	0.53	0.38	0.26	0.18	0.11	0.07
	14	0.95	0.89	0.79	0.69	0.54	0.38	0.27	0.18	0.11	0.06
	16	0.96	0.90	0.80	0.71	0.56	0.39	0.27	0.19	0.11	0.06
KPSS	2	0.59	0.66	0.71	0.76	0.79	0.83	0.86	0.88	0.90	0.92
	4	0.45	0.51	0.57	0.61	0.64	0.69	0.73	0.76	0.79	0.83
	6	0.37	0.42	0.46	0.51	0.54	0.60	0.65	0.67	0.70	0.75
	8	0.31	0.36	0.40	0.44	0.46	0.53	0.57	0.60	0.64	0.69
	10	0.26	0.30	0.34	0.38	0.41	0.47	0.52	0.55	0.59	0.64
	12	0.21	0.25	0.28	0.33	0.35	0.42	0.46	0.49	0.54	0.59
	14	0.17	0.21	0.24	0.28	0.31	0.36	0.41	0.44	0.49	0.54
	16	0.13	0.17	0.20	0.23	0.26	0.31	0.37	0.39	0.44	0.50

**Quadre 8.4.** Percentatge de rebutjos de la hipòtesi nul·la en els test DF, ADF, PP i KPSS.  
 Model ARFIMA(1,d,1) amb  $\phi=0$ ,  $\theta=0$ ,  $1 \leq d < 1.5$  i  $T=100$ .

	k	d=1.00	1.05	1.10	1.15	1.20	1.25	1.30	1.35	1.40	1.45
<b>DF</b>		0.06	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
<b>ADF</b>	2	0.06	0.03	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00
	4	0.06	0.04	0.02	0.02	0.01	0.01	0.01	0.01	0.00	0.00
	6	0.06	0.04	0.03	0.02	0.02	0.01	0.01	0.01	0.01	0.01
	8	0.05	0.05	0.03	0.03	0.02	0.02	0.02	0.02	0.01	0.01
	10	0.05	0.05	0.04	0.03	0.03	0.02	0.02	0.02	0.02	0.01
	12	0.06	0.05	0.04	0.03	0.03	0.02	0.03	0.02	0.02	0.02
	14	0.06	0.05	0.05	0.03	0.03	0.02	0.03	0.02	0.02	0.02
	16	0.06	0.06	0.05	0.03	0.04	0.03	0.03	0.03	0.03	0.03
<b>PP</b>	2	0.06	0.03	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00
	4	0.06	0.04	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00
	6	0.06	0.04	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00
	8	0.07	0.04	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00
	10	0.06	0.04	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00
	12	0.07	0.04	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00
	14	0.06	0.04	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00
	16	0.06	0.03	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00
<b>KPSS</b>	2	0.92	0.93	0.95	0.96	0.97	0.98	0.99	0.99	0.99	1.00
	4	0.83	0.84	0.87	0.90	0.91	0.93	0.95	0.96	0.97	0.99
	6	0.75	0.77	0.80	0.84	0.85	0.87	0.90	0.92	0.94	0.97
	8	0.69	0.71	0.75	0.79	0.81	0.84	0.86	0.89	0.92	0.95
	10	0.64	0.66	0.70	0.75	0.77	0.81	0.83	0.87	0.90	0.94
	12	0.59	0.62	0.65	0.72	0.74	0.77	0.80	0.85	0.88	0.93
	14	0.54	0.58	0.61	0.68	0.71	0.74	0.77	0.83	0.86	0.92
	16	0.50	0.53	0.57	0.64	0.67	0.71	0.75	0.81	0.84	0.90



2. La potència del PP és similar al DF.
3. En canvi, el pitjor contrast és clarament l'ADF. De fet, a mesura que augmenta el nombre de diferències usades,  $k$ , la potència del contrast es redueix ràpidament. Això suggereix que una possible via per la detecció de memòria llarga pot ser la d'analitzar el comportament de l'estadístic ADF per diferents valors de  $k$ , ja que en el cas de models AR(1) no es produeix aquest fenomen tant accentuat.
4. Pel que fa al KPSS, presenta també una potència molt baixa.

Però aquesta poca potència tant dels tests d'arrels unitàries en que la hipòtesi nul·la és que la sèrie és I(0) com que la sèrie és I(1) pot aprofitar-se com a mètode per detectar la presència d'integració fraccional a través de les contradiccions entre ambdós tipus de contrastos. Així, si una sèrie fos generada segons el model  $(1-L)^d X_t = \varepsilon_t$ , amb  $0.25 < d < 0.75$ , és molt probable que la conclusió del test KPSS sigui la de que la sèrie és I(0) mentre que els tests ADF o PP porten a concloure que la sèrie és I(1). Això es degut a que no hi ha prou evidència per rebutjar cap de les dues hipòtesi. Per tant, les contradiccions entre aquests tests s'han d'interpretar amb cura ja que poden ser un indicatiu d'integració fraccional.

Per valorar amb quina freqüència s'obtenen aquestes contradiccions amb dades fraccionalment integrades, s'han generat 5000 sèries de 100 observacions d'un model ARFIMA(0,d,0) i s'han aplicat els test ADF i PP de manera simultània al test KPSS i s'han comptabilitzat en quants casos la conclusió dels dos tests era la mateixa i en quants casos era contradictòria.

**Quadre 8.5.** Resultats de l'aplicació simultània de tests d'arrels unitàries amb hipòtesi nul·la I(1) (ADF o PP) conjuntament amb el test KPSS<sup>10</sup>. Model ARFIMA(1,d,1) amb  $\phi=0$ ,  $\theta=0$  i  $0 \leq d < 0.5$ .

	<i>k</i>		<i>d=0.00</i>	<i>0.05</i>	<i>0.10</i>	<i>0.15</i>	<i>0.20</i>	<i>0.25</i>	<i>0.30</i>	<i>0.35</i>	<i>0.40</i>	<i>0.45</i>	
ADF	2	I(0)	0.95	0.91	0.85	0.79	0.72	0.64	0.59	0.49	0.38	0.23	
		I(1)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.09	0.26
		Cont.	0.05	0.09	0.15	0.21	0.28	0.36	0.41	0.49	0.53	0.51	
	4	I(0)	0.95	0.92	0.88	0.83	0.78	0.70	0.62	0.49	0.34	0.19	
		I(1)	0.00	0.00	0.00	0.00	0.00	0.01	0.04	0.11	0.20	0.34	
		Cont.	0.05	0.08	0.12	0.17	0.22	0.29	0.34	0.40	0.46	0.46	
	6	I(0)	0.95	0.93	0.89	0.84	0.78	0.68	0.57	0.42	0.28	0.16	
		I(1)	0.00	0.00	0.00	0.01	0.02	0.06	0.10	0.18	0.25	0.36	
		Cont.	0.05	0.07	0.10	0.15	0.20	0.26	0.33	0.41	0.47	0.48	
	8	I(0)	0.95	0.93	0.88	0.81	0.72	0.61	0.49	0.34	0.24	0.14	
		I(1)	0.00	0.00	0.01	0.03	0.05	0.09	0.14	0.20	0.26	0.33	
		Cont.	0.05	0.07	0.11	0.16	0.23	0.30	0.38	0.46	0.50	0.53	
10	I(0)	0.93	0.89	0.83	0.74	0.64	0.52	0.41	0.29	0.20	0.11		
	I(1)	0.01	0.01	0.03	0.05	0.07	0.11	0.15	0.20	0.25	0.31		
	Cont.	0.06	0.09	0.15	0.21	0.29	0.37	0.44	0.51	0.55	0.58		
12	I(0)	0.88	0.82	0.75	0.65	0.55	0.45	0.34	0.25	0.17	0.10		
	I(1)	0.01	0.02	0.04	0.06	0.08	0.11	0.15	0.19	0.23	0.28		
	Cont.	0.10	0.15	0.22	0.30	0.37	0.44	0.51	0.56	0.61	0.63		
PP	2	I(0)	0.95	0.91	0.85	0.79	0.72	0.64	0.59	0.50	0.41	0.27	
		I(1)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.04	0.19
		Cont.	0.05	0.09	0.15	0.21	0.28	0.36	0.41	0.50	0.55	0.54	
	4	I(0)	0.95	0.92	0.88	0.83	0.78	0.71	0.66	0.60	0.50	0.34	
		I(1)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.04	0.16
		Cont.	0.05	0.08	0.12	0.17	0.22	0.29	0.34	0.40	0.46	0.50	
	6	I(0)	0.95	0.93	0.90	0.85	0.81	0.76	0.71	0.65	0.57	0.39	
		I(1)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.14
		Cont.	0.05	0.07	0.10	0.15	0.19	0.24	0.29	0.34	0.40	0.47	
	8	I(0)	0.96	0.94	0.91	0.87	0.84	0.79	0.75	0.69	0.62	0.44	
		I(1)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.12
		Cont.	0.05	0.06	0.09	0.13	0.16	0.21	0.25	0.30	0.36	0.44	
10	I(0)	0.96	0.95	0.92	0.88	0.85	0.82	0.78	0.73	0.66	0.48		
	I(1)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.10	
	Cont.	0.04	0.05	0.08	0.12	0.15	0.18	0.22	0.27	0.32	0.42		
12	I(0)	0.96	0.95	0.93	0.90	0.87	0.84	0.81	0.76	0.70	0.51		
	I(1)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.09	
	Cont.	0.04	0.05	0.07	0.10	0.13	0.16	0.20	0.24	0.29	0.40		

<sup>10</sup> A la fila nomenada com "I(0)" hi ha la proporció de casos en que els dos contrastos porten a la conclusió de que la sèrie és I(0), a la segona fila hi ha la proporció de casos amb coincidència dels dos contrastos en l'altre sentit, mentre que a la tercera fila hi ha la proporció de casos on els resultats d'ambdós contrastos són contradictoris.

**Quadre 8.6.** Resultats de l'aplicació simultània de tests d'arrels unitàries amb hipòtesi nul·la I(1) (ADF o PP) conjuntament amb el test KPSS<sup>11</sup>. Model ARFIMA(1,d,1) amb  $\phi=0$ ,  $\theta=0$  i  $0.5 < d \leq 0.5$ .

	<i>k</i>		<i>d</i> =0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00
ADF	2	I(0)	0.37	0.29	0.23	0.17	0.13	0.09	0.06	0.05	0.03	0.02
		I(1)	0.10	0.18	0.29	0.39	0.51	0.63	0.72	0.78	0.84	0.89
		Cont.	0.53	0.53	0.49	0.44	0.36	0.28	0.22	0.17	0.13	0.09
	4	I(0)	0.37	0.31	0.25	0.19	0.15	0.12	0.08	0.06	0.05	0.03
		I(1)	0.22	0.30	0.39	0.45	0.52	0.60	0.67	0.72	0.76	0.80
		Cont.	0.41	0.40	0.37	0.36	0.32	0.28	0.24	0.22	0.19	0.16
	6	I(0)	0.33	0.28	0.23	0.18	0.15	0.11	0.09	0.07	0.05	0.04
		I(1)	0.24	0.32	0.38	0.44	0.49	0.56	0.61	0.64	0.68	0.73
		Cont.	0.42	0.40	0.39	0.38	0.37	0.33	0.31	0.29	0.26	0.22
	8	I(0)	0.30	0.24	0.20	0.17	0.15	0.11	0.09	0.07	0.06	0.04
		I(1)	0.24	0.30	0.35	0.40	0.43	0.50	0.55	0.58	0.62	0.68
		Cont.	0.46	0.45	0.45	0.43	0.42	0.39	0.36	0.35	0.32	0.27
10	I(0)	0.26	0.21	0.18	0.15	0.13	0.11	0.08	0.07	0.06	0.04	
	I(1)	0.22	0.27	0.31	0.35	0.38	0.45	0.50	0.53	0.58	0.63	
	Cont.	0.52	0.52	0.52	0.49	0.48	0.44	0.42	0.40	0.36	0.32	
12	I(0)	0.23	0.19	0.16	0.14	0.13	0.11	0.08	0.07	0.06	0.05	
	I(1)	0.19	0.23	0.26	0.31	0.34	0.41	0.45	0.49	0.53	0.58	
	Cont.	0.58	0.58	0.58	0.54	0.54	0.48	0.47	0.44	0.41	0.37	
PP	2	I(0)	0.39	0.31	0.25	0.19	0.14	0.11	0.07	0.05	0.04	0.02
		I(1)	0.05	0.12	0.21	0.30	0.44	0.58	0.69	0.76	0.84	0.88
		Cont.	0.56	0.57	0.54	0.51	0.42	0.31	0.24	0.19	0.13	0.09
	4	I(0)	0.51	0.43	0.36	0.29	0.23	0.17	0.12	0.09	0.06	0.03
		I(1)	0.05	0.10	0.20	0.27	0.38	0.52	0.62	0.68	0.75	0.80
		Cont.	0.44	0.46	0.45	0.44	0.38	0.31	0.27	0.23	0.19	0.17
	6	I(0)	0.59	0.51	0.44	0.36	0.28	0.20	0.15	0.11	0.07	0.04
		I(1)	0.04	0.09	0.17	0.24	0.34	0.46	0.56	0.61	0.67	0.73
		Cont.	0.37	0.40	0.40	0.40	0.38	0.34	0.29	0.28	0.26	0.23
	8	I(0)	0.65	0.57	0.50	0.42	0.33	0.24	0.17	0.12	0.08	0.05
		I(1)	0.03	0.07	0.14	0.20	0.29	0.41	0.50	0.55	0.61	0.68
		Cont.	0.32	0.36	0.36	0.38	0.37	0.35	0.33	0.32	0.31	0.27
10	I(0)	0.70	0.62	0.55	0.46	0.37	0.27	0.19	0.14	0.09	0.05	
	I(1)	0.02	0.05	0.12	0.18	0.26	0.37	0.45	0.51	0.57	0.63	
	Cont.	0.28	0.32	0.33	0.36	0.37	0.36	0.36	0.36	0.35	0.32	
12	I(0)	0.75	0.67	0.59	0.50	0.40	0.29	0.21	0.15	0.09	0.06	
	I(1)	0.02	0.04	0.10	0.16	0.23	0.33	0.41	0.46	0.52	0.58	
	Cont.	0.24	0.29	0.31	0.34	0.37	0.38	0.38	0.39	0.38	0.36	

<sup>11</sup> A la fila nomenada com "I(0)" hi ha la proporció de casos en que els dos contrastos porten a la conclusió de que la sèrie és I(0), a la segona fila hi ha la proporció de casos amb coincidència dels dos contrastos en l'altre sentit, mentre que a la tercera fila hi ha la proporció de casos on els resultats d'ambdós contrastos són contradictoris.

Als quadres 8.5 i 8.6 es presenta el percentatge de cada tipus de conclusions segons el valor del paràmetre  $d$  amb que s'han generat les dades. Observi's que, com era d'esperar, a mesura que el valor de  $d$  s'allunya dels valors 0 i 1, el percentatge de contradiccions és molt elevat. De fet, quan el valor del paràmetre  $d$  és proper a 1/2, s'arriben a produir més d'un 50% de contradiccions. Per tant, es poden interpretar la presència de contradiccions en els contrastos d'arrels unitàries més habituals com a símptoma d'integració fraccional.

Per altra banda, tot i que no coneixem evidència prèvia sobre aquest aspecte, també és important valorar quin és el comportament del contrast de Phillips i Ouliaris (1988) (PO) en el context de sèries  $FI(d)$ .

Recordi's que aquest contrast es basa en contrastar la invertibilitat del terme de perturbació un cop s'ha diferenciat la variable. Així, si s'ha produït una sobrediferenciació, la variable diferenciada presenta una mitjana mòbil no invertible i per tant, presenta un zero a l'espectre. Més concretament, es pot resumir l'aplicació del contrast de la següent manera:

- Diferenciar la variable i centrar-ne la sèrie diferenciada:  $e_t = Y_t - \bar{Y}_t$  on  $Y_t = X_t - X_{t-1}$
- Calcular  $\tilde{\sigma}_u^2$  i  $\tilde{\sigma}_l^2$  estimacions de la variància i de la variància a llarg termini respectivament a partir de:

$$\tilde{\sigma}_u^2 = \frac{1}{T} \sum_{t=1}^n e_t^2 \tag{8.18}$$

i:

$$\tilde{\sigma}_l^2 = \frac{1}{T} \sum_{t=1}^n e_t^2 + \frac{2}{T} \sum_{j=1}^l W(j,l) \sum_{t=j+1}^T e_t e_{t-j} \tag{8.19}$$

on  $W(j,l)$  és una finestra de retards<sup>12</sup>.

<sup>12</sup> Es diu que la funció  $W(j,l)$  és una finestra de retards (lag window) si compleix:

- 1)  $|W(k,l)| \leq 1$ .
- 2)  $W(0,l) = 1$ .
- 3)  $W(k,l) = W(-k,l)$ .
- 4)  $W(k,l) = 0$  si  $k \geq l$ .

- Calcular l'estadístic  $r^2 = \tilde{\sigma}_l^2 / \tilde{\sigma}_u^2$  que es distribueix asimptòticament com una normal
- Finalment, aplicar la següent regla de decisió:

a) Si  $r^2 + r^2 z_\alpha / \sqrt{l} < 0.05$  es rebutja la hipòtesi nul·la de presència d'una arrel unitària.

b) Si  $r^2 - r^2 z_\alpha / \sqrt{l} > 0.05$  no es rebutja la hipòtesi nul·la i  $X_t$  és integrable de primer ordre.

on  $z_\alpha$  és el valor que acumula la probabilitat  $(1-\alpha)$  d'obtenir un valor inferior en la distribució normal estàndard.

Cal destacar que per aplicar el contrast de Phillips i Ouliaris cal usar alguna finestra de retards concreta que pot influir en les propietats del contrast. La més habitual és la finestra de Bartlett o triangular definida segons:

$$W(j,l) = \left(1 - \frac{j}{l-1}\right) \quad (8.20)$$

cas en que la variància a llarg termini (8.19) coincideix amb la variància (8.15) usada en el contrast de Phillips i Perron.

Als quadres 8.7 i 8.8 es presenten els resultats d'una simulació similar a les anteriors per poder valorar l'eficàcia del contrast PO a partir de la finestra de Bartlett quan les dades s'han generat a partir del model  $(1-\phi L)(1-L)^d X_t = \varepsilon_t$  amb  $T=100$ . Observi's com el contrast presenta una potència gairebé nul·la, degut a que el contrast és no concloent en molts casos, ja que  $r^2 + r^2 z_\alpha / \sqrt{l} > 0.05$  i  $r^2 - r^2 z_\alpha / \sqrt{l} < 0.05$ .

---

De fet, tota finestra de retards és l'equivalent en el domini temporal d'una finestra espectral definida en el domini de les freqüències.





Quadre 8.9. Percentatge de rebutjos de la hipòtesi nul·la d'integrabilitat (potència) amb el contrast PO amb la finestra Rectangular si  $(1-\phi L)(1-L)^d X_t = \varepsilon_t$  amb  $T=100$ .

<i>d</i>	$\phi=0$				$\phi=0.5$				$\phi=0.9$			
	<i>l</i> =5	6	7	8	5	6	7	8	5	6	7	8
0.00	0.49	0.54	0.53	0.56	0.32	0.41	0.47	0.50	0.01	0.03	0.05	0.06
0.05	0.50	0.51	0.52	0.52	0.31	0.39	0.44	0.47	0.01	0.02	0.03	0.04
0.10	0.51	0.51	0.49	0.49	0.25	0.32	0.41	0.44	0.00	0.01	0.02	0.03
0.15	0.47	0.52	0.48	0.49	0.20	0.31	0.38	0.43	0.00	0.01	0.01	0.01
0.20	0.44	0.52	0.46	0.49	0.18	0.25	0.30	0.36	0.00	0.00	0.01	0.01
0.25	0.44	0.47	0.46	0.47	0.14	0.19	0.27	0.32	0.00	0.00	0.00	0.01
0.30	0.46	0.42	0.42	0.48	0.08	0.16	0.22	0.28	0.00	0.00	0.00	0.00
0.35	0.39	0.39	0.43	0.44	0.07	0.12	0.19	0.24	0.00	0.00	0.00	0.00
0.40	0.33	0.38	0.39	0.42	0.05	0.10	0.14	0.18	0.00	0.00	0.00	0.00
0.45	0.29	0.32	0.37	0.37	0.02	0.05	0.09	0.13	0.00	0.00	0.00	0.00
0.55	0.23	0.29	0.32	0.33	0.01	0.03	0.08	0.11	0.00	0.00	0.00	0.00
0.60	0.19	0.23	0.28	0.32	0.01	0.03	0.05	0.10	0.00	0.00	0.00	0.00
0.65	0.14	0.18	0.21	0.25	0.01	0.02	0.03	0.07	0.00	0.00	0.00	0.00
0.70	0.10	0.13	0.15	0.19	0.00	0.01	0.02	0.05	0.00	0.00	0.00	0.00
0.75	0.06	0.09	0.12	0.14	0.00	0.00	0.01	0.03	0.00	0.00	0.00	0.00
0.80	0.04	0.06	0.07	0.10	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00
0.85	0.02	0.03	0.05	0.07	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00
0.90	0.01	0.02	0.03	0.05	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
0.95	0.01	0.01	0.02	0.03	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
1.00	0.00	0.01	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00



Quadre 8.10. Percentatge de no rebuig de la hipòtesi nul·la d'integrabilitat amb el contrast PO amb la finestra Rectangular si  $(1-\phi L)(1-L)^d X_t = \varepsilon_t$  amb  $T=100$ .

$d$	$\phi=0$				$\phi=0.5$				$\phi=0.9$			
	$l=5$	6	7	8	5	6	7	8	5	6	7	8
0.00	0.00	0.05	0.11	0.14	0.04	0.15	0.17	0.21	0.79	0.86	0.85	0.84
0.05	0.00	0.05	0.10	0.16	0.05	0.16	0.21	0.23	0.86	0.92	0.92	0.90
0.10	0.00	0.05	0.12	0.17	0.06	0.21	0.26	0.28	0.93	0.95	0.93	0.93
0.15	0.00	0.06	0.13	0.17	0.12	0.24	0.29	0.28	0.97	0.97	0.97	0.96
0.20	0.01	0.05	0.13	0.16	0.18	0.33	0.36	0.37	0.99	0.99	0.98	0.98
0.25	0.00	0.06	0.14	0.19	0.23	0.38	0.42	0.43	0.99	1.00	0.99	0.98
0.30	0.01	0.10	0.18	0.20	0.36	0.50	0.52	0.51	1.00	1.00	1.00	1.00
0.35	0.02	0.09	0.19	0.21	0.45	0.57	0.56	0.57	1.00	1.00	1.00	1.00
0.40	0.02	0.12	0.20	0.25	0.59	0.67	0.64	0.63	1.00	1.00	1.00	1.00
0.45	0.03	0.15	0.24	0.27	0.67	0.75	0.75	0.72	1.00	1.00	1.00	1.00
0.55	0.06	0.20	0.29	0.32	0.77	0.83	0.80	0.76	1.00	1.00	1.00	1.00
0.60	0.10	0.26	0.35	0.38	0.87	0.89	0.86	0.82	1.00	1.00	1.00	1.00
0.65	0.16	0.36	0.45	0.47	0.92	0.93	0.90	0.87	1.00	1.00	1.00	1.00
0.70	0.22	0.48	0.56	0.55	0.96	0.96	0.93	0.92	1.00	1.00	1.00	1.00
0.75	0.38	0.59	0.66	0.68	0.98	0.98	0.97	0.94	1.00	1.00	1.00	1.00
0.80	0.52	0.70	0.72	0.73	0.99	0.99	0.97	0.96	1.00	1.00	1.00	1.00
0.85	0.64	0.79	0.81	0.82	0.99	0.99	0.98	0.97	1.00	1.00	1.00	1.00
0.90	0.75	0.86	0.87	0.86	1.00	1.00	0.99	0.99	1.00	1.00	1.00	1.00
0.95	0.85	0.93	0.93	0.91	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00
1.00	0.94	0.96	0.96	0.95	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00

El motiu d'aquest comportament tan deficient es troba en el biaix que es produeix amb la finestra de Bartlett a l'estimar la densitat espectral a la freqüència zero. De fet, a Sansó *et al.* (1997) a partir d'una anàlisi detallada de l'estadístic de contrast es conclou que la utilització de finestres espectrals introdueix un biaix molt important que invalida la seva aplicació per contrastar si  $d=1$ .

En el mateix treball es comprova, a partir de l'anàlisi de diferents classes de finestres espectrals que sols la finestra Rectangular definida com  $W(j,l)=1$  si  $j<l$ , i la finestra de Daniell definida com:

$$W(j,l) = \frac{\sin(j\pi/l)}{(j\pi, l)} \quad (8.21)$$

presenten una proporció de casos amb resultats concloents acceptable.

Per aquest motiu, s'han analitzat també aquestes finestres a partir d'un exercici de simulació i s'ha comprovat com la rectangular és la que presenta millors resultats. Aquests resultats es recullen als quadres 8.9 i 8.10, i pot comprovar-se com, tot i que els resultats són ara molt millors que amb la finestra de Bartlett, s'obté una potència molt inferior a la que s'obté amb el contrast DF, no perquè el contrast porti a concloure que no es rebutja la hipòtesi nul·la, sinó que encara en molts casos no és concloent.

Finalment, Robinson (1994) proposa un test alternatiu per contrastar la hipòtesi de la presència d'una arrel unitària i altres formes de no estacionarietat basat en un procés generador de les dades prou general per a que contingui un major número d'hipòtesis alternatives, és a dir, un contrast dissenyat específicament per discriminar entre els diferents valors del paràmetre  $d$ .

Concretament, es proposen un test que és localment de màxima potència en relació a un conjunt molt ampli d'alternatives amb l'avantatge que els valors crítics asimptòtics segueixen

una distribució jhi-quadrat. Com a hipòtesis alternatives es consideren els processos integrats fraccionalment ja sigui en la freqüència zero com en les freqüències estacionals.

Suposi's com a punt de partida la següent representació per la variable  $X_t$ :

$$\varphi(L)X_t = u_t \quad (8.22)$$

on  $u_t$  pot ser soroll blanc o presentar memòria curta, com per exemple un model ARMA.

El test proposat permet contrastar hipòtesis del tipus:

$$\varphi(L) = (1-L)^{d_1} (1+L)^{d_2} \prod_{j=3}^m (1-2\cos\omega_j L + L^2)^{d_j} \quad (8.23)$$

contra alternatives del tipus:

$$\varphi(L) = (1-L)^{d_1+\delta_1} (1+L)^{d_2+\delta_2} \prod_{j=3}^m (1-2\cos\omega_j L + L^2)^{d_j+\delta_j} \quad (8.24)$$

Per fer-ho, es contrasta la hipòtesi nul·la:

$$H_0: \delta_1 = \delta_2 = \dots = \delta_m = 0$$

contra el seu complementari. Per exemple, en el cas més senzill usat aquí com exemple, suposi's que es vol contrastar el paràmetre de diferenciació fraccional  $H_0: d=d_0$  en el model:

$$(1-L)^d X_t = u_t \quad (8.25)$$

Per contrastar-ho cal calcular:

$$\tilde{u}_t = (1-L)^{d_0} X_t \quad (8.26)$$

on prèviament, és necessari suposar una certa estructura de dependència per  $u_t$ . Suposi's que es vol usar aquest mètode per un contrast sobre el valor de  $d$  en el model (8.25) suposant que  $u_t$  és soroll blanc. En aquest cas,  $u_t$  té un espectre de la forma:

$$f_u(\omega; \sigma^2) = \frac{\sigma^2}{2\pi} \quad (8.27)$$

Aquest espectre s'estima com és ben conegut a través del periodograma:

$$I(\omega_j) = \frac{1}{2\pi T} \left| \sum_{t=1}^T \tilde{u}_t e^{\frac{2\piijt}{T}} \right|^2 \quad (8.28)$$

Robinson (1994) proposa basar el contrast en el següent estadístic:

$$\tilde{a} = -\frac{2\pi}{T} \sum_{j=1}^{T-1} \log |2 \sin(\pi j/T)| I(\omega_j) \quad (8.29)$$

Observi's com l'estadístic sols depèn de les dades a través dels residus que s'obtenen filtrant amb el valor que es desitja contrastar  $d=d_0$ .

Un cop normalitzat aquest estadístic, es demostra que l'estadístic:

$$\tilde{S} = \left( \frac{T}{\tilde{b}} \right)^{1/2} \frac{\tilde{a}}{\tilde{\sigma}^2} \quad (8.30)$$

convergeix en distribució, sota la hipòtesi nul·la, cap a una normal estàndard. Els factors de normalització són un estimador de la variància de les dades i:

$$\tilde{b} = \frac{2}{T} \sum_{j=1}^{T-1} \log^2 |\sin(\pi j/T)| \quad (8.31)$$

En particular, aquest mètode pot usar-se per realitzar un contrast d'arrels unitàries utilitzant  $d_0=1$  amb l'avantatge de que el conjunt d'hipòtesis alternatives considerades inclou models els casos en que  $(1-L)^d X_t = \varepsilon_t$  amb  $d \neq 1$ . Per comparar el funcionament d'aquest contrast en mostra finita amb els ja analitzats, s'ha realitzat un exercici de simulació equivalent als anteriors.

Així, s'han generat 5000 sèries de 100 observacions segons el procés  $(1-L)^d X_t = \varepsilon_t$ , segons diferents valors del paràmetre  $d$ , s'ha calculat per cada una d'aquestes sèries l'estadístic  $\tilde{S}$  i s'ha calculat en quin percentatge de casos s'ha rebutjat la hipòtesi nul·la que  $d=1$ . Al quadre 8.11 es comparen aquests percentatges amb els que s'obtenen al aplicar el test proposat per Dickey i Fuller (1979).

**Quadre 8.11.** Percentatge de rebutjos de la hipòtesi nul·la  $d=1$  a partir de l'estadístic  $\tilde{S}$  i a partir del contrast DF quan  $(1-L)^d X_t = \varepsilon_t$  amb  $T=100$ .

$d$	$\tilde{S}$	DF
0.40	1.00	1.00
0.45	1.00	0.99
0.55	1.00	0.98
0.60	1.00	0.93
0.65	1.00	0.85
0.70	0.95	0.75
0.75	0.89	0.60
0.80	0.78	0.42
0.85	0.58	0.29
0.90	0.37	0.19
0.95	0.20	0.11
1.00	0.08	0.06

La conclusió d'aquesta comparació és clara, en el sentit que el test proposat per Robinson (1994) és molt més potent que el test DF en el context de models ARFIMA(0,d,0).

Tot i que s'ha presentat el contrast de Robinson sota la hipòtesi de soroll blanc, aquest es pot ampliar per tenir en compte processos més complexos. Així, si es suposa que les dades  $u_t$  no són soroll blanc sinó que es permet una certa estructura de dependència, cal especificar la forma funcional de l'espectre en funció d'un vector de  $q$  paràmetres que nomenarem  $\theta$ .

$$f_u(\omega; \sigma^2) = \frac{\sigma^2}{2\pi} g(\omega; \theta) \quad (8.32)$$

Per exemple, en el cas d'un AR(1), s'obté:

$$g(\omega; \theta) = (1 + \theta^2 - 2\theta \cos \omega)^{-1} \quad (8.33)$$

Ara cal estimar aquests paràmetres i Robinson (1994) proposa escollir-los a partir de la minimització de:

$$\tilde{\sigma}^2(\theta) = \frac{2\pi}{T} \sum_{j=1}^{T-1} \frac{I(\omega_j)}{g(\omega_j; \theta)} \quad (8.34)$$

Ara l'estadístic es basa en:

$$\tilde{\alpha} = -\frac{2\pi}{T} \sum_{j=1}^{T-1} \log|2\sin(\pi j/T)| \frac{I(\omega_j)}{g(\omega_j; \tilde{\theta})} \quad (8.35)$$

i per calcular el factor de normalització cal definir un vector columna  $m$  format pels  $(T-1)$  termes  $\log|2\sin \pi j/T|$  i una matriu  $M$  de dimensió  $(T-1) \times q$  formada en cada fila pels elements:

$$\frac{\partial \log g(\omega_j; \tilde{\theta})}{\partial \theta} \quad (8.36)$$

i calcular aquest factor com:

$$\tilde{b} = \frac{2}{T} \{m' m - m' M (M' M)^{-1} M' m\} \quad (8.37)$$

Pel que fa a la variància es calcula com:

$$\tilde{\sigma}^2 = \tilde{\sigma}^2(\tilde{\theta}) \quad (8.38)$$

Quan hi ha regressors o termes deterministes, n'hi ha prou amb usar com estimador de  $u_t$ :

$$\tilde{u}_t = (1-L)^{d_0} X_t - \tilde{\beta}' Z_t \quad (8.39)$$

on:

$$w_t = (1-L)^{d_0} Z_t \quad (8.40)$$

i:

$$\tilde{\beta} = \left( \sum_{t=1}^T w_t w_t' \right)^{-1} \sum_{t=1}^T w_t (1-L)^{d_0} X_t \quad (8.41)$$

Observi's que un dels avantatges d'aquest enfocament, és que no cal restringir-se al fet d'usar models ARMA com modelització de la memòria curta sinó que és possible suposar qualsevol tipus de forma funcional per la densitat espectral. Per altra banda, es pot plantejar també el contrast en el domini temporal, però llavors l'estadístic no té una formulació tant compacte. A més, donat que s'ha definit el concepte de memòria llarga en termes de la densitat espectral, d'aquesta manera és més clara la interpretació del contrast.

Tal i com suggereix el seu autor, també es pot substituir el factor de normalització per qualsevol altre que tingui el mateix límit en probabilitat, per exemple, quan la funció espectral és coneguda com pot ser el cas soroll blanc, llavors es pot agafar directament el límit:

$$\tilde{b} = \frac{1}{\pi} \int_{-\pi}^{\pi} (\log|2\sin(\omega/2)|)^2 d\omega = \frac{\pi^2}{6} \quad (8.42)$$

Per tant, és d'esperar que el contrast sigui clarament més potent que els analitzats abans, sempre que el model considerat per  $u_t$  sigui prou general com per incloure el procés generador de dades.

Per tant, la conclusió final d'aquesta revisió dels tests d'arrels unitàries en el context de la integrabilitat fraccional no pot ser altre que la recomanació de que, davant la sospita de presència de memòria llarga, hom no pot limitar-se a usar els contrastos habituals d'arrels unitàries dissenyats pensant en la hipòtesi alternativa autoregressiva, sinó que cal usar també algun contrast dissenyat específicament contra hipòtesis alternatives que continguin memòria llarga. En aquest sentit, l'evidència disponible mostra que el contrast proposat per Robinson (1994) presenta molt bones propietats.

### 8.2.2 Identificació i estimació del model ARMA

Suposi's que es disposa d'una realització d'un procés estocàstic  $X_t$  que compleix la següent equació estocàstica:

$$(1 - L)^d X_t = \varepsilon_t, \quad (8.43)$$

on  $\varepsilon_t$  és soroll blanc, de manera que, en la terminologia habitual  $X_t \sim \text{ARFIMA}(0, d, 0)$ . Estem interessats a conèixer en aquest cas quin fora el resultat d'intentar ajustar un model ARIMA a aquestes dades.

A l'apartat anterior s'ha comprovat com per  $0.5 \leq d \leq 1$  és molt probable acceptar que la sèrie conté una arrel unitària i per tant, que hom es decideixi a diferenciar les dades, de manera que l'objectiu és trobar un model ARMA per  $Y_t = (1-L)X_t = X_t - X_{t-1}$ . A partir de la definició (8.43) de  $X_t$ , s'obté que<sup>13</sup>:

$$\begin{aligned}
 Y_t &= (1-L)^{1-d} \varepsilon_t = \varepsilon_t + (d-1)\varepsilon_{t-1} + \frac{d(d-1)}{2} \varepsilon_{t-2} + \frac{d(d-1)(d+1)}{6} \varepsilon_{t-3} + \dots = \\
 &= \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots = \sum_{j=1}^{\infty} \psi_j \varepsilon_{t-j}
 \end{aligned}
 \tag{8.44}$$

per tant,  $Y_t$  segueix un procés MA( $\infty$ ) amb coeficients que tendeixen ràpidament cap a zero de manera que és molt probable que en aquesta etapa d'identificació del model s'acabi escollint un model MA(1) per la sèrie en diferències.

Al quadre 8.12 es pot comprovar aquest ràpid decreixement dels coeficients del model (8.44) i als gràfics 8.1 i 8.2 es recullen, com exemple, les funcions d'autocorrelació simple i parcial<sup>14</sup> de  $Y_t = (1-L)X_t$  quan  $X_t \sim \text{ARFIMA}(0, d, 0)$  amb  $d = 0.45$ . Noti's com, efectivament, a partir d'aquests gràfics és d'esperar l'elecció d'un model MA(1) per  $Y_t$ .

Aquesta observació dona una possible explicació al fet que moltes variables econòmiques s'acaben aproximant a través del model línies aèries, fenomen que podria ser degut a que aquestes dades presenten memòria llarga, o bé són no estacionàries amb  $d < 1$ .

<sup>13</sup> Aplicant que l'expansió del binomi  $(1-x)^a$  és:

$$(1-x)^a = \sum_{k=0}^{\infty} (-1)^k \binom{a}{k} x^k = 1 - aL + \frac{a(a-1)}{2} x^2 - \frac{a(a-1)(a-2)}{6} x^3 + \dots$$

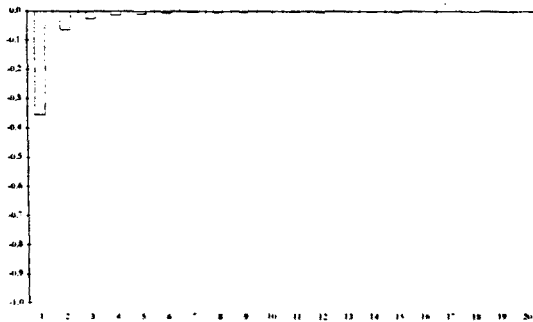
<sup>14</sup> Si  $(1-L)^d X_t = \varepsilon_t$  amb  $d < 1/2$ , la funció d'autocorrelació parcial (FAP) de  $X_t$  és  $\alpha(k) = d/(k-d)$ .



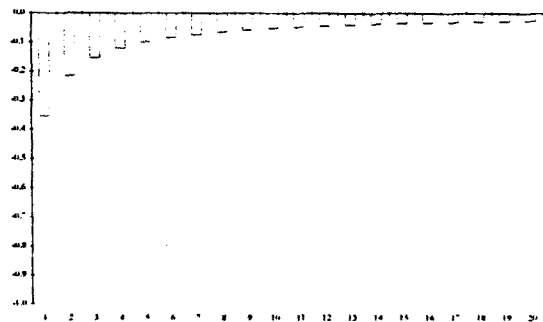
Quadre 8.12. Coeficients  $\psi_k = (-1)^k \binom{1-d}{k}$  per alguns valors de  $d$ .

k	d=0.2	d=0.45
0	1.000	1.000
1	-0.800	-0.355
2	-0.080	-0.063
3	-0.032	-0.014
4	-0.018	-0.009
5	-0.011	-0.006
6	-0.008	-0.004
7	-0.006	-0.003
8	-0.005	-0.002
9	-0.004	-0.002
10	-0.003	-0.001

Gràfic 8.1. FAS de la primera diferència de  $X_t$  quan  $(1-L)^d X_t \sim \varepsilon_t$  amb  $d=0.45$ .



Gràfic 8.2. FAP de la primera diferència de  $X_t$  quan  $(1-L)^d X_t \sim \varepsilon_t$  amb  $d=0.45$ .



De totes maneres, aquesta elecció depèn del nombre d'observacions disponibles en la sèrie analitzada en la mesura en que aquest nombre determina els intervals de confiança dels coeficients de correlació estimats; però per tamany mostrals de 100 o 200 observacions, habituals en l'anàlisi de variables econòmiques mensuals, s'obtenen uns coeficients de la FAS no significatius a partir del segon retard.

Fins al moment s'ha suposat que prèviament a l'especificació d'un model ARMA s'ha optat per diferenciar la sèrie. Però encara sabent el risc de sobrediferenciació i plantejant com alternatives els dos models següents<sup>15</sup>:

<sup>15</sup> Noti's que per  $d < 1$  els models A i B presenten característiques molt diferents en relació a la seva memòria. Mentre el model A presenta reversió a la mitjana, en el model B els shocks que afecten a la variable tenen

A)  $(1-L)^d X_t = \varepsilon_t$

B)  $(1-L)X_t = (1+\theta L)\varepsilon_t$ , amb  $\theta=d-1$ .

no sembla que les eines habituals en l'anàlisi de sèries econòmiques siguin capaces de distingir un model de l'altre. De fet aquesta decisió acostuma a basar-se en tres classes d'instruments:

- 1) En les característiques de les autocorrelacions estimades a partir de les observacions de la variable, de manera que un decreixement lent del correlograma s'associa amb la necessitat de diferenciar les dades. Als gràfics 8.3 i 8.4 es pot comprovar com aquests correlogrames<sup>16</sup> són molt semblants pels models A i B<sup>17</sup>, tot i que el model A no requereix diferenciació i el model B sí. A més, cal citar també que al capítol 5 de la tesi s'ha mostrat com els correlogrames són estimacions molt esbiaixades quan les dades presenten memòria llarga.
- 2) D'altra banda, és habitual usar els tests d'arrels unitàries i, tal com s'ha comprovat a l'apartat anterior, és fàcil concloure que  $X_t$  és I(1) quan està integrat fraccionalment amb  $d < 1$ .

---

efectes permanents. Si a més  $d < 1/2$ , llavors la diferència entre els models s'accentua ja que el model A és estacionari de segon ordre i el model B presenta una tendència estocàstica.

<sup>16</sup> Wichern (1973) demostra que si  $(1-L)X_t = (1+\theta L)\varepsilon_t$ , llavors l'esperança de la funció d'autocorrelació estimada a través de l'estimador habitual  $\tilde{\rho}_k = \left[ \sum (X_t - \bar{X})^2 \right]^{-1} \sum (X_t - \bar{X})(X_{t+k} - \bar{X})$  és:

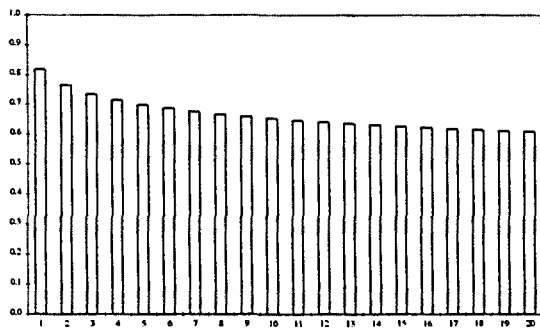
$$E[\tilde{\rho}_k] = (T-k) \left\{ (1+\theta)^2 (T^2 - 1 + 2k^2 - 4kT) + 6\theta \right\} \left[ T(T-1) \left\{ (T+1)(1+\theta)^2 - 6\theta \right\} \right]^{-1}$$

mentre que les autocorrelacions d'un model ARFIMA(0,d,0) són  $\rho_k = \frac{\Gamma(k+d)\Gamma(1-d)}{\Gamma(k-d+1)\Gamma(d)}$ .

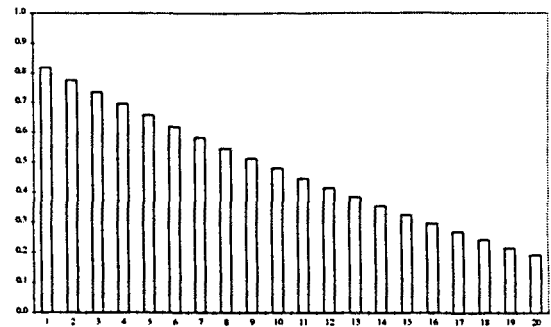
<sup>17</sup> Quan  $d > 1/2$ , el model A no és estacionari però es produeix la mateixa confusió encara que en aquest cas cal usar  $E(\tilde{\rho}_k)$  enlloc de  $\rho_k$ .

3) Tampoc sembla que els test *portmanteau* de validació del model estimat basats en els residus, com poden ser per exemple els contrastos de Box i Pierce (1970) o Ljung i Box (1978), siguin capaços de detectar el problema de sobrediferenciació. Concretament, a l'apartat següent s'aporta evidència a través d'un exercici de simulació d'aquests problemes.

**Gràfic 8.3.** FAS de la sèrie  $X_t$  en el model A amb  $d=0.45$ .



**Gràfic 8.4.**  $E(\tilde{\rho}_k)$  de la sèrie  $X_t$  en el model B amb  $\theta=-0.55$ .



Fins al moment s'ha considerat sols la dificultat per determinar amb eines definides en el domini temporal la memòria o persistència d'una sèrie temporal a la freqüència zero, però a les freqüències estacionals passa exactament el mateix. Així, seguint la proposta de Porter-Hudak (1990) té també interès considerar un procés integrat fraccionalment a les freqüències estacionals:

$$(1-L^s)^d X_t = \varepsilon_t \quad (8.45)$$

En aquest cas, si  $d < 1$  l'aplicació de l'operador  $(1-L^s)$  porta també a sobrediferenciar les dades i a obtenir el següent procés<sup>18</sup>:

$$\begin{aligned} (1-L^s)X_t &= (1-L^s)^{1-d} \varepsilon_t = \\ &= \varepsilon_t + (d-1)\varepsilon_{t-s} + \frac{d(d-1)}{2}\varepsilon_{t-2s} + \frac{d(d-1)(d+1)}{6}\varepsilon_{t-3s} + \dots \end{aligned} \quad (8.46)$$

<sup>18</sup> Ja que  $(1-x^s)^a = \sum_{k=0}^{\infty} (-1)^k \binom{a}{k} x^{sk} = 1 - aL + \frac{a(a-1)}{2}x^{2s} - \frac{a(a-1)(a-2)}{6}x^{3s} + \dots$

Aquestes observacions són coherents amb els resultats de Porter-Hudak (1990) que mostra com les sèries mensuals d'agregats monetaris M1, M2 i M3 es poden modelitzar força bé a partir del model (8.45) amb  $s=12$ , i s'obtenen prediccions amb una qualitat comparable a les que s'obtenen usant el model línies aèries<sup>19</sup>.

De manera anàloga al que passa amb els models A i B, els estadístics habituals en el domini temporal no permeten distingir sèries temporals obtingudes segons el procés  $(1-L^s)^d X_t = \varepsilon_t$  de sèries generades a partir de  $(1-L^s)X_t = (1+\theta L^s)\varepsilon_t$  amb  $\theta=d-1$ .

Però a la pràctica, la majoria de sèries temporals econòmiques presenten força memòria tant a la freqüència zero com a les freqüències estacionals. De fet, a Porter-Hudak (1990) es suggereix la possibilitat d'usar un model que inclogui de manera simultània diferències fraccional regulars i estacionals<sup>20</sup>:

$$(1-L)^{d_1} (1-L^{12})^{d_2} X_t = \varepsilon_t \quad (8.47)$$

tot i que la utilització d'aquest tipus de model no és de moment gens fàcil ja que presenta uns problemes molt importants en la identificació i estimació.

Usant arguments similars als anteriors, si s'obté una sèrie temporal a partir d'un procés estocàstic definit per l'equació (8.47), és molt probable que a partir dels estadístics habituals es decideixi la utilització dels operadors diferència regular i diferència estacional. En aquest cas, es dedueix a partir de (8.47) que:

<sup>19</sup> El nom de model línies aèries prové de la sèrie temporal que Box i Jenkins (1970) utilitzen com exemple per presentar i discutir les característiques d'aquest model, un model ARFIMA(0,1,1)(0,1,1)<sub>s</sub>. De fet, les sèries d'agregats monetaris en són exemples molt típics de variables econòmiques que s'acaben aproximant habitualment amb el model línies aèries.

<sup>20</sup> Es tracta d'un cas particular dels models SARFIMA  $(1-L)^{d_1} (1-L^s)^{d_2} \phi(L)X_t = \theta(L)\varepsilon_t$ , introduïts al capítol 4 de la tesi.

$$(1-L)(1-L^s)X_t = \left[ (1+\theta_1 L)(1+\theta_2 L^s) + \sum_{j=s+1}^{\infty} \zeta_j L^j \right] \varepsilon_t \quad (8.48)$$

on els coeficients  $\zeta_k$  són prou propers a zero per ser considerats com no significatius.

Per tant, la recerca d'un model ARIMA portarà probablement a l'estimació d'un model línies aèries:

$$(1-L)(1-L^s)X_t = (1+\theta_1 L)(1+\theta_2 L^s)\varepsilon_t \quad (8.49)$$

on és d'esperar que s'obtinguin valors estimats per  $\theta_1$  y  $\theta_2$  propers a  $d_1-1$  i  $d_2-1$  respectivament.

Efectivament, als gràfics 8.5 i 8.6 es comparen els correlogrames d'ambdós models, el (8.47) i el (8.49) usant  $s=12$ ,  $d_1=d_2=0.4$ ,  $\theta_1=d_1-1=-0.6$  i  $\theta_2=d_2-1=-0.6$ ; i es pot comprovar com a partir d'aquests estadístics és impossible distingir ambdós models<sup>21</sup>.

A continuació, s'ha realitzat un exercici de simulació molt senzill com a complement de l'anàlisi teòrica anterior. En una primera etapa s'han comparat els models A i B:

A)  $(1-L)^d X_t = \varepsilon_t$ ,

B)  $(1-L)X_t = (1+\theta L)\varepsilon_t$ , amb  $\theta=d-1$ .

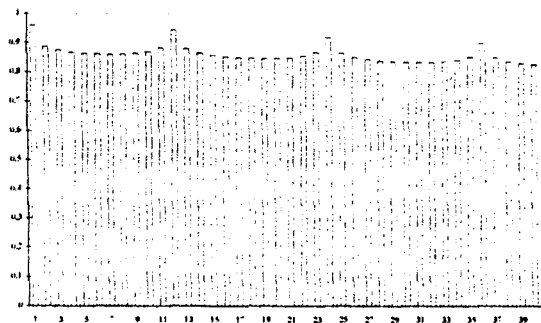
Per fer-ho, s'han generat un seguit de sèries temporals de longitud  $T=250$  a partir d'un procés estocàstic  $X_t \sim \text{ARFIMA}(0, d, 0)$  per a diferents valors del paràmetre  $d$  de l'interval  $[0, 1]$ . Per a generar aquestes sèries cal prèviament calcular les autocovariàncies de  $X_t$  que es poden obtenir a partir de la següent expressió:

<sup>21</sup> Pel moment no es disposa d'una expressió explícita pel càlcul de les autocovariàncies d'un procés estocàstic definit a partir de  $(1-L)^{d_1}(1-L^s)^{d_2}X_t = \varepsilon_t$ . En el Teorema 8.2 de l'Annex 8.2 es desenvolupa un mètode que permet obtenir de manera aproximada però molt precisa aquestes autocovariàncies.

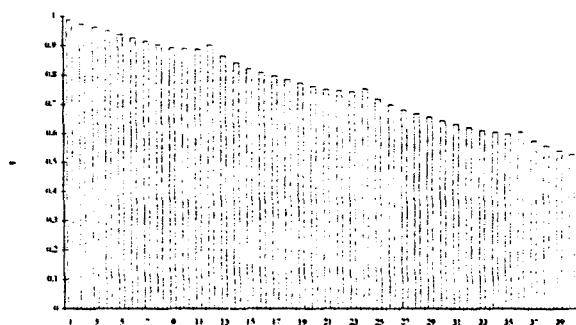
$$\gamma(k) = \frac{(-1)^k \Gamma(1-2d)}{\Gamma(1-d+k)\Gamma(1-d-k)} \sigma_\varepsilon^2 \tag{8.50}$$

Un cop es disposa de les autocovariàncies del procés a generar, es construeix la matriu de covariàncies  $\Sigma$ , que és una matriu de Toeplitz. A continuació, el mètode de Tcholeski permet trobar una matriu  $P$  triangular inferior que compleix  $\Sigma=PP'$ . Finalment, si  $U$  és un vector  $N(0,1)$  aleshores el vector  $X=PU$  té per matriu de covariàncies  $cov(X)=cov(PUUP')=PP'=\Sigma$  que era l'objectiu que es pretenia assolir.

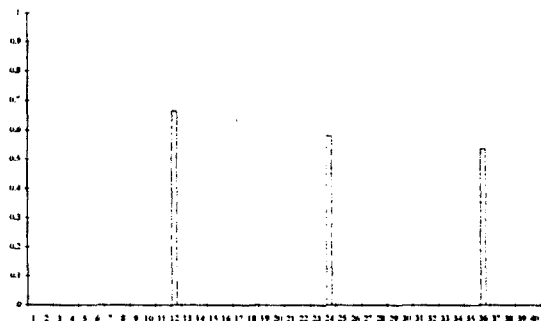
**Gràfic 8.5.** FAS de la sèrie  $X_t$  generada com  $(1-L)^{d_1}(1-L^s)^{d_2}X_t=\varepsilon_t$  amb  $d_1=d_2=0.4$ .



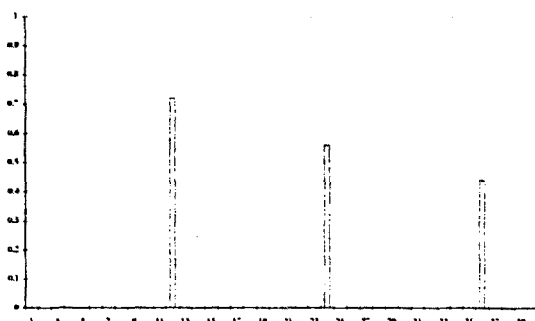
**Gràfic 8.6.**  $E(\tilde{\rho}_k)$  de la sèrie  $X_t$  generada a partir de (8.49) amb  $\theta_1=\theta_2=-0.6$ .



**Gràfic 8.7.** FAS de  $(1-L)X_t$  amb  $X_t$  complint  $(1-L)^{d_1}(1-L^s)^{d_2}X_t=\varepsilon_t$  amb  $d_1=d_2=0.4$ .



**Gràfic 8.8.**  $E(\tilde{\rho}_k)$  de  $(1-L)X_t$  amb  $X_t$  complint (8.49) amb  $\theta_1=\theta_2=-0.6$ .



El mètode anterior sols és vàlid per a processos estacionaris, és a dir, quan  $d < 1/2$ . Per a generar sèries integrades fraccionalment amb  $1/2 \leq d \leq 1$  cal generar prèviament una sèrie segons el model

$Z_t \sim \text{ARFIMA}(0, d^*, 0)$  amb  $d^* = d - 1$  i a continuació es calcula de manera recursiva  $X_t = Z_t + X_{t-1}$ , de manera que la sèrie  $X_t$  és integrada fraccionalment amb paràmetre  $1 + d^* = d$ .

Per a cada valor del paràmetre  $d$  s'han generat 1000 sèries i en cada cas s'ha estimat un model MA(1) per  $(1-L)X_t$ . Al quadre 8.13 es recullen la mitjana i la variància de les 1000 estimacions del paràmetre  $\theta$  que segons els resultats analítics anteriors, s'hauria de distribuir al voltant del valor  $\theta = d - 1$ .

Quadre 8.13. Característiques de  $\tilde{\theta}$  en el model  $\Delta X_t = (1 + \theta L)\varepsilon_t$ .

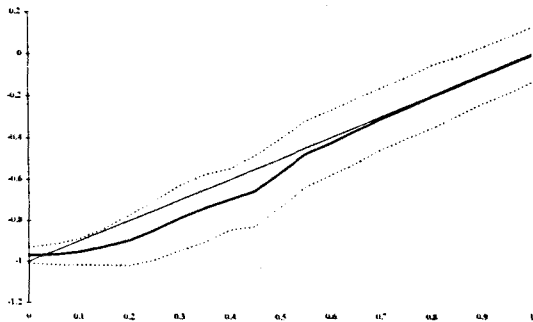
d	$\theta = d - 1$	Mitjana	Variància
0.00	-1.00	-0.97	0.0004
0.05	-0.95	-0.96	0.0006
0.10	-0.90	-0.95	0.0010
0.15	-0.85	-0.93	0.0020
0.20	-0.80	-0.90	0.0038
0.25	-0.75	-0.85	0.0052
0.30	-0.70	-0.79	0.0063
0.35	-0.65	-0.74	0.0067
0.40	-0.60	-0.70	0.0056
0.45	-0.55	-0.66	0.0074
0.55	-0.45	-0.48	0.0063
0.60	-0.40	-0.43	0.0061
0.65	-0.35	-0.37	0.0060
0.70	-0.30	-0.31	0.0056
0.75	-0.25	-0.26	0.0054
0.80	-0.20	-0.21	0.0050
0.85	-0.15	-0.16	0.0045
0.90	-0.10	-0.11	0.0045
0.95	-0.05	-0.05	0.0045
1.00	0.00	0.00	0.0044

Efectivament, del quadre 8.13 es dedueix que es compleix aquesta propietat tot i que per valors del paràmetre  $d < 1/2$  l'estimador presenta un cert biaix positiu. De fet, el perill de sobrediferenciar les dades és molt superior per valors d'integració fraccional superiors a  $1/2$  que per valors inferiors.

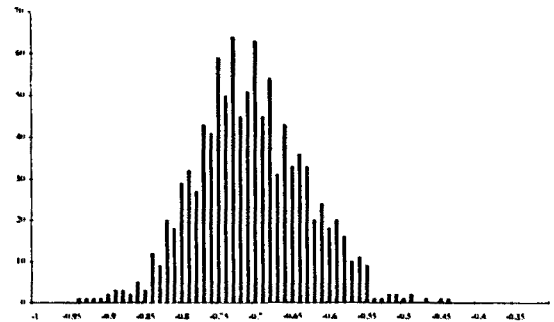
Al gràfic 8.9 en el que s'ha representat, a més de la mitjana de l'estimador  $\tilde{\theta}$ , un interval de confiança calculat a partir de la variància estimada<sup>22</sup>, es pot veure clarament aquest biaix. Com exemple, al gràfic 8.10 es presenta la distribució de  $\tilde{\theta}$  quan  $d=0.3$ .

A continuació, s'ha realitzat un exercici de simulació similar a l'anterior però generant ara sèries temporals integrades fraccionalment tant a la freqüència zero com a les freqüències estacionals segons el procés  $(1-L)^{d_1}(1-L^{12})^{d_2}X_t = \varepsilon_t$ . Per a generar aquestes sèries també s'ha usat la descomposició de Choleski però ara a partir d'una aproximació als valors de les autocovariàncies<sup>23</sup>.

Gràfic 8.9.  $E(\tilde{\theta})$  en funció de  $d$  al estimar el model  $\Delta X_t = (1 + \theta L)\varepsilon_t$  quan  $(1-L)^d X_t = \varepsilon_t$ .



Gràfic 8.10. Distribució de  $\tilde{\theta}$  al estimar el model  $\Delta X_t = (1 + \theta L)\varepsilon_t$  si  $(1-L)^d X_t = \varepsilon_t$  amb  $d=0.3$ .



<sup>22</sup> Al gràfic 8.9, la línia més gruixuda recull la mitjana del valor estimat a les simulacions; en línia fina el valor esperat a partir de l'anàlisi teòric anterior, és a dir,  $\theta_1=d_1-1$  i  $\theta_2=d_2-1$ ; i en línia discontinua l'interval de confiança a partir de  $\pm 2 \text{ var}(\tilde{\theta})$

<sup>23</sup> Vegi's el Teorema 8.2 de l'Annex 8.2.



Al quadre 8.14, on es recullen alguns dels resultats de la simulació, es pot comprovar com, efectivament, les estimacions dels paràmetres mitjana mòbil que s'obtenen si s'especifica un model línies aèries es distribueixen al voltant dels valors  $d_1-1$  i  $d_2-1$ .

**Quadre 8.14.** Característiques de  $\tilde{\theta}_1$  i  $\tilde{\theta}_2$  al estimar el model  $\Delta\Delta_t X_t = (1 + \theta_1 L) (1 + \theta_2 L^2) \varepsilon_t$ , amb dades generades segons el procés  $(1 - L)^{d_1} (1 - L^2)^{d_2} X_t = \varepsilon_t$ .

$d_1$	$d_2$	$\theta_1 = d_1 - 1$	Distribució de $\tilde{\theta}_1$		$\theta_2 = d_2 - 1$	Distribució de $\tilde{\theta}_2$	
			Mitjana	Variància		Mitjana	Variància
0.00	0.00	-1.00	-0.97	0.0015	-1.00	-0.97	0.0018
	0.40		-0.96	0.0027		-0.66	0.0021
	0.80		-0.96	0.0033		-0.20	0.0023
0.20	0.00	-0.80	-0.95	0.0041	-1.00	-0.96	0.0035
	0.40		-0.96	0.0051		-0.65	0.0037
	0.80		-0.95	0.0058		-0.22	0.0038
0.40	0.00	-0.60	-0.63	0.0081	-1.00	-0.95	0.0052
	0.40		-0.65	0.0083		-0.65	0.0054
	0.80		-0.62	0.0086		-0.20	0.0060
0.60	0.00	-0.40	-0.45	0.0064	-1.00	-0.94	0.0046
	0.40		-0.44	0.0067		-0.63	0.0047
	0.80		-0.44	0.0066		-0.21	0.0049
0.80	0.00	-0.20	-0.21	0.0045	-1.00	-0.94	0.0032
	0.40		-0.21	0.0047		-0.62	0.0034
	0.80		-0.20	0.0048		-0.20	0.0035

### 8.2.3. Validació del model ARIMA estimat

Com a complement de l'anàlisi anterior, s'han analitzat els contrastos més habituals per validar l'adequació del model estimat a les dades observades. Concretament, s'han analitzat els contrastos proposats per Box i Pierce (1970) i per Ljung i Box (1978). El quadre 8.15 mostra el percentatge de casos en que es rebutja la hipòtesi nul·la que les primeres  $k$  autocorrelacions dels residus són nul·les. Observi's com el nombre de casos en que es rebutja  $H_0$  és clarament inferior al que cabria esperar tenint en compte que s'ha produït un error d'especificació.

Quadre 8.15. Percentatge de rebuig de la hipòtesi nul·la als contrastos de BP i LB<sup>24</sup>.

<i>d</i>	<i>Contrast de Box-Pierce</i>			<i>Contrast de Ljung-Box</i>		
	<i>k=2</i>	<i>k=6</i>	<i>k=10</i>	<i>k=2</i>	<i>k=6</i>	<i>k=10</i>
0.05	0.19	0.05	0.03	0.07	0.00	0.00
0.15	0.40	0.24	0.25	0.31	0.05	0.03
0.25	0.55	0.45	0.34	0.48	0.15	0.11
0.35	0.61	0.54	0.48	0.52	0.14	0.17
0.45	0.38	0.28	0.25	0.33	0.07	0.12
0.55	0.26	0.44	0.33	0.24	0.08	0.18
0.65	0.40	0.40	0.37	0.33	0.06	0.09
0.75	0.22	0.26	0.19	0.18	0.04	0.04
0.85	0.14	0.12	0.09	0.11	0.02	0.03
0.95	0.03	0.03	0.02	0.04	0.03	0.03

Així, atès que els instruments estadístics habituals basats en el domini temporal no són capaços de distingir de manera adequada el tipus de memòria de certs models, es proposa usar algun instrument basat en el domini de les freqüències.

De fet, en aquest domini, els models A i B:

A)  $(1 - L)^d X_t = \varepsilon_t$ ,

B)  $(1 - L)X_t = (1 + \theta L)\varepsilon_t$ , amb  $\theta = d - 1$ .

presenten característiques ben diferents ja que mentre l'espectre del model A és (sempre que es compleixi  $-1/2 < d < 1/2$ ):<sup>25</sup>

<sup>24</sup> Box i Pierce (1970) proposen contrastar la hipòtesi nul·la que  $\rho_\varepsilon(1) = \dots = \rho_\varepsilon(k) = 0$  a través de l'estadístic  $Q(k) = T \sum_{j=1}^k \tilde{\rho}_\varepsilon^2(j)$  que sota la hipòtesi nul·la es distribueix asimptòticament com una  $\chi^2$  amb  $k - m$  graus de llibertat on  $m$  és el número de paràmetres del model. Ljung i Box (1978) proposen usar l'estadístic  $Q^*(k) = T(T + 2) \sum_{j=1}^k (T - j)^{-1} \tilde{\rho}_\varepsilon^2(j)$  que sota la hipòtesi nul·la segueix la mateixa distribució asimptòtica que l'estadístic anterior.

<sup>25</sup> Si  $d \geq 1/2$ ,  $X_t$  no és estacionari i per tant no té sentit parlar d'espectre. En tot cas, si s'usa el concepte de pseudo-espectre l'expressió (8.51) és vàlida per qualsevol valor del paràmetre  $d$ .

$$f_A(\omega) = \frac{\sigma^2}{2\pi(2 - 2\cos\omega)^d} \quad (8.51)$$

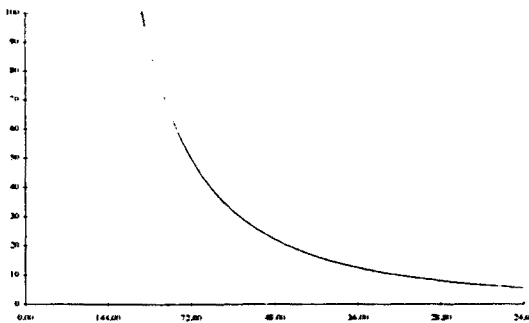
i el pseudo-espectre associat al model B és:

$$f_B(\omega) = \frac{\sigma^2(1 + \theta^2 - 2\theta\cos\omega)}{2\pi(2 - 2\cos\omega)} \quad (8.52)$$

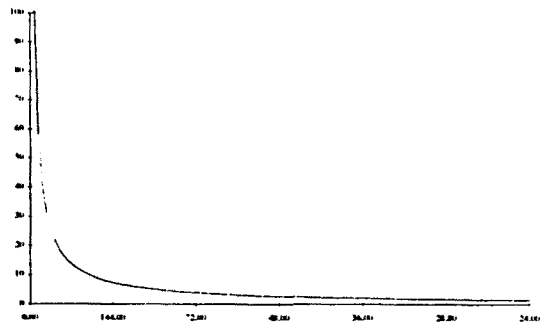
El perfil gràfic d'aquestes dues expressions aparentment tan similars, és extraordinàriament diferent. De fet, sols tenen en comú un clar predomini de les baixes freqüències similar al que mostren moltes de les variables econòmiques, però aquest predomini és molt major en el model B que en el model A.

Per apreciar aquesta diferència, als gràfics 8.11 i 8.12 es presenten exemples de les funcions (8.51) i (8.52), és a dir les densitats espectrals dels models A i B per valors  $d=0.45$  i  $\theta=-0.55$ , respectivament sota el supòsit d'una variància del soroll blanc igual a la unitat. Per poder comparar aquests gràfics cal tenir en compte tres aspectes importants:

**Gràfic 8.11.** Espectre d' $X_t$  per períodes a  $[24, \infty)$  en el model A amb  $d=0.45$ .



**Gràfic 8.12.** Espectre d' $X_t$  per períodes  $[24, \infty)$  en el model B amb  $\theta=-0.55$ .



- D'una banda, que el pseudo-espectre del model B no és integrable i per tant delimita una àrea infinita, mentre que l'espectre del model A, al ser  $d < 0$ , delimita una àrea finita que coincideix amb la variància d' $X_t$ .
- Tenint en compte l'anterior, s'han representat aquests dos gràfics amb la mateixa escala vertical per a fer possible la comparació.

- A més, donat que gairebé tota la potència d'aquests espectres es concentra al voltant de la freqüència zero, s'han representat en aquests gràfics els períodes compresos a l'interval  $[24, \infty)$ , és a dir, les freqüències de l'interval  $[0, \pi/12]$ .

Recordi's que els correlogrames d'aquests dos models representats als gràfics 8.3 i 8.4 són molt similars i en canvi, els espectres dels gràfics 8.11 i 8.12 són ben diferents. Aquest exemple n'és una prova més de que el domini temporal no és l'adequat per distingir el tipus de memòria que presenten les sèries temporals i que per fer aquesta distinció cal usar el domini de les freqüències.

De fet, la densitat espectral d'un procés estocàstic descriu com es distribueix la variació total de la sèrie temporal entre les diferents freqüències de tal manera que la presència d'una major o menor memòria en una determinada freqüència es reflexa en una major potència de la densitat espectral en dita freqüència. En aquest sentit, el gràfic 8.12 presenta una major concentració al voltant de la freqüència zero (els períodes més grans) ja que conté memòria permanent, mentre que el gràfic 8.11 tot i també tenir la potència al voltant de la freqüència zero, es tracta d'un espectre que delimita una àrea finita i per tant és estacionari tot i que amb memòria llarga.

Per això, qualsevol instrument que permeti mesurar el grau de concentració de la potència de l'espectre d'una sèrie en unes poques freqüències sembla una bona mesura de la seva memòria. En aquest sentit, es pot usar algun dels instruments proposats al capítol 5 per mesurar la memòria de les sèries temporals, per exemple la mediana del periodograma estandaritzat  $Me_X = \text{Mediana}\{z_j\}$  on:

$$z_j = \frac{1}{\sum (X_t - \bar{X})^2} \left| \sum_{t=1}^T X_t e^{-\frac{2i\pi j t}{T}} \right|^2 \quad (8.53)$$

Recordi's que un soroll blanc té un periodograma estandaritzat distribuït segons una exponencial amb esperança i variància unitària. Per tant, la mediana en aquest cas és

$Me_X = \log(2) = 0.69$ . Naturalment, a mesura que la memòria de la sèrie augmenta, l'estadístic  $Me_X$  es redueix.

**Quadre 8.16.** Resum dels valors de l'estadístic  $Me_X$  en els models A i B

Model A			Model B		
d	Mitjana	Variància	$\theta=d-1$	Mitjana	Variància
0.00	0.689	0.0041	-1.00	0.689	0.0041
0.05	0.681	0.0042	-0.95	0.649	0.0058
0.10	0.674	0.0040	-0.90	0.536	0.0130
0.15	0.655	0.0045	-0.85	0.431	0.0170
0.20	0.628	0.0044	-0.80	0.339	0.0180
0.25	0.590	0.0048	-0.75	0.260	0.0145
0.30	0.549	0.0057	-0.70	0.208	0.0114
0.35	0.495	0.0061	-0.65	0.170	0.0092
0.40	0.451	0.0061	-0.60	0.141	0.0064
0.45	0.473	0.0094	-0.55	0.113	0.0045
0.55	0.268	0.0064	-0.45	0.080	0.0024
0.60	0.215	0.0054	-0.40	0.071	0.0019
0.65	0.171	0.0041	-0.35	0.061	0.0015
0.70	0.138	0.0034	-0.30	0.053	0.0011
0.75	0.103	0.0023	-0.25	0.046	0.0008
0.80	0.082	0.0016	-0.20	0.043	0.0007
0.85	0.062	0.0010	-0.15	0.038	0.0005
0.90	0.047	0.0007	-0.10	0.033	0.0004
0.95	0.036	0.0004	-0.05	0.032	0.0004
1.00	0.029	0.0003	0.00	0.030	0.0003

Com exemple, al quadre 8.16 es recullen els valors mitjans d'aquest estadístic calculats a partir de 1000 sèries de 100 observacions generades a partir dels models A i B. Així, en relació a la confusió entre els models A i B, el quadre 8.18 mostra com aquest estadístic pren valors bastant diferents per ambdós models, i per tant es pot usar fins i tot com a contrast, almenys per valors del paràmetre  $d$  propers a  $1/2$ .

En definitiva, aquests resultats corroboren els obtinguts al capítol 5 en referència a la identificació del tipus de memòria. És a dir, amb les eines estadístiques habituals és molt difícil decidir quin és el model més correcte, un model ARIMA o bé un model ARFIMA. En tot cas, aquesta elecció haurà de basar-se en els objectius pretesos.

Finalment, enlloc de decidir *a priori* sobre l'elecció d'ajustar un model ARFIMA o bé un model ARIMA, hi ha també l'opció d'estimar per separat ambdós tipus de models i fer-ne una selecció *a posteriori*. Per fer aquesta selecció *a posteriori*, un cop estimats els dos models es poden usar algun dels criteris de selecció de models més habituals<sup>26</sup>:

a) El criteri d'informació d'Akaike (AIC)<sup>27</sup>:

$$AIC = T \ln(\tilde{\sigma}_\epsilon^2) + 2m \quad (8.54)$$

on  $T$  és el número d'observacions incloses a la sèrie analitzada,  $\tilde{\sigma}_\epsilon^2$  és la variància estimada de les innovacions del model ajustat a la sèrie i  $m$  és el número de paràmetres estimats en el model.

b) El criteri d'informació d'Akaike corregit (AICc)<sup>28</sup>:

$$AICc = T \ln(\tilde{\sigma}_\epsilon^2) + T \frac{1 + m/T}{1 - (2 + m)/T} \quad (8.55)$$

c) El criteri d'informació de Schwarz (SIC)<sup>29</sup>:

$$SIC = T \ln(\tilde{\sigma}_\epsilon^2) + m \ln(T) \quad (8.56)$$

Per decidir quin d'aquests tres criteris funciona millor a l'hora de discriminar entre models ARIMA i models ARFIMA, s'ha realitzat una simulació en la que s'han generat 1000 sèries temporals de longitud  $T=100$  segons un model ARFIMA(1,d,1) per diferents paràmetres  $\phi$ ,  $\theta$  i

<sup>26</sup> La decisió d'incloure el paràmetre  $d$  en el model estimat no pot basar-se en la significació d'aquest paràmetre, donades les dificultats que presenta l'estimació d'aquest paràmetre, sobretot quan s'inclouen simultàniament en el model una diferència fraccional i termes autoregressius.

<sup>27</sup> Vegi's Akaike (1973).

<sup>28</sup> Vegi's Hurvich i Tsai (1989).

<sup>29</sup> Vegi's Schwarz (1978).

d. Per cada una d'aquestes sèries s'han estimat models ARMA(1,1), ARIMA(1,1,1) i ARFIMA(1,d,1) i s'han usat els estadístics AIC, AICc i SIC per seleccionar un dels tres models. Al quadre 8.17 es recullen els resultats d'aquesta simulació per alguns valors dels paràmetres.

Dels resultats del quadre 8.17 es dedueixen les següents conclusions:

- El percentatge de casos en que el model seleccionat és el correcte no és gaire gran, sobretot quan  $0 < d < 1$  però és proper al zero o a la unitat. En aquests casos, com era d'esperar és més probable la selecció d'un model ARMA o ARIMA respectivament.
- El percentatge de casos en que es selecciona el model ARFIMA augmenta molt quan  $0.4 \leq d \leq 0.6$ .
- En general, el criteri SIC és el que porta al major número d'encerts en la selecció del model.

Quadre 8.17. Proporció de casos en que s'ha seleccionat cada model usant AIC, AICc i SIC a partir de sèries generades segons un ARFIMA(1,d,1) amb  $T=100$ .

$d$	$\phi$	$\theta$	AIC			AICc			SIC		
			ARMA	ARIMA	ARFIMA	ARMA	ARIMA	ARFIMA	ARMA	ARIMA	ARFIMA
0.0	0.0	0.0	0.86	0.05	0.09	0.88	0.04	0.08	0.77	0.08	0.15
	0.5	0.0	0.80	0.08	0.12	0.80	0.06	0.14	0.55	0.12	0.33
	0.0	0.5	0.84	0.07	0.09	0.83	0.06	0.11	0.60	0.12	0.28
0.2	0.0	0.0	0.72	0.12	0.16	0.71	0.11	0.18	0.72	0.09	0.19
	0.5	0.0	0.59	0.16	0.25	0.61	0.15	0.24	0.60	0.14	0.26
	0.0	0.5	0.54	0.09	0.37	0.57	0.05	0.38	0.51	0.07	0.42
0.4	0.0	0.0	0.17	0.32	0.51	0.24	0.28	0.48	0.26	0.23	0.51
	0.5	0.0	0.13	0.35	0.52	0.14	0.37	0.49	0.07	0.39	0.54
	0.0	0.5	0.21	0.35	0.44	0.17	0.38	0.45	0.11	0.41	0.48
0.6	0.0	0.0	0.04	0.58	0.38	0.13	0.48	0.39	0.02	0.61	0.37
	0.5	0.0	0.20	0.65	0.15	0.25	0.54	0.21	0.23	0.54	0.23
	0.0	0.5	0.43	0.48	0.09	0.42	0.51	0.07	0.25	0.61	0.14
0.8	0.0	0.0	0.08	0.75	0.17	0.15	0.67	0.18	0.23	0.65	0.12
	0.5	0.0	0.07	0.79	0.14	0.14	0.75	0.11	0.16	0.69	0.15
	0.0	0.5	0.07	0.81	0.12	0.02	0.83	0.15	0.10	0.74	0.16

En tot cas, aquests resultats confirmen que en presència d'integració fraccional amb  $0.5 < d < 1$  i per tant, amb memòria permanent finita, és molt probable cometre un error de sobrediferenciació que porti a la introducció de paràmetres mitjana mòbil.

### 8.3. Conseqüències per a la mesura de la persistència dels *shocks* en variables econòmiques

Tal i com ja s'ha comentat en diferents moments al llarg de la tesi, bona part de l'interès o preocupació que els analistes de les dades temporals econòmiques han mostrat pel concepte de persistència es troba en les implicacions econòmiques que té l'acceptació d'un cert model com a representatiu del seu comportament. En particular, quan s'accepta que una variable  $X_t$  no és estacionària i cal prendre una diferència per a que ho sigui, s'interpreta habitualment que els shocks que es puguin produir en aquesta variable tindran un efecte permanent en el temps, conclusió que segons quina sigui la variable objecte d'anàlisi pot portar a importants conclusions de tipus econòmic i, fins i tot, de tipus polític.

Però s'ha demostrat a la tesi que aquesta és una conclusió errònia ja que si una variable és  $FI(d)$  amb  $1/2 < d < 1$  i, per tant, no estacionària, és molt probable que les eines habituals ens indiquin la necessitat de diferenciar-la però la variable presenta reversió a la mitjana, és a dir, els shocks no tenen un efecte permanent en el temps.

Naturalment, el concepte de persistència en economia és molt proper al concepte de memòria tal com s'ha presentat en aquesta tesi. La diferència més important es troba en que, a partir de la definició proposada a la tesi, la memòria és una característica no paramètrica de manera que no té una interpretació econòmica directa. Sols permet discriminar entre diferents tipus de memòria, fins i tot permet discriminar dintre del mateix grup més o menys persistència però no permet interpretar aquesta persistència. De totes maneres, els diferents instruments presentats tant al capítol 2 (instruments més clàssics) com al capítol 5 (instruments més novedosos) per mesurar el grau de memòria poden servir com mesures de la persistència.

Ja s'han comentat prèviament els problemes que hi ha per contrastar a nivell empíric a partir d'observacions  $X_1, \dots, X_T$  la hipòtesi que  $d=1$  contra  $d \neq 1$  o bé contra  $d < 1$  en el context de models ARFIMA. De manera que, tot i que un procés  $X_t$  sigui integrat fraccionalment amb  $d < 1$  i, per tant, amb memòria permanent finita, és probable que no sigui possible rebutjar la hipòtesi



d'integrabilitat d'ordre 1. Així, les dificultats apuntades en molts treballs per decidir amb un nombre reduït d'observacions si  $d=0$  o  $d=1$  en el context de la integració (no fraccional) queden multiplicades en el context de la integració fraccional.

En tot cas, creiem que és important acceptar sempre aquesta limitació i defensem la utilització de mesures no paramètriques com les apuntades al capítol 5 per aproximar el grau de memòria. Aquestes mesures tenen el problema obvi de que no permeten interpretar en termes econòmics aquesta persistència però almenys permeten establir comparacions entre la memòria de diferents processos i a més, tal i com s'ha il·lustrat en el mateix capítol, donen eines per validar l'adequació dels models paramètrics usats.

Les decisions entre memòria permanent o transitòria (i també entre estacionarietat o no estacionarietat) són poc potents en termes purament estadístics, a no ser que es disposi de gran nombre d'observacions, i creiem que haurien de basar-se a més de criteris estadístics en criteris de Teoria Econòmica.

A més, defensem que la decisió d'usar un model ARIMA com a punt de partida pel càlcul de prediccions, per l'extracció de senyals o bé per decidir el mètode d'estimació de models econòmics no ha d'implicar l'acceptació que la variable presenta memòria permanent infinita sinó que, tal i com es desenvolupa més endavant, pot ser millor a efectes pràctics i en certes ocasions restringir l'anàlisi a un model ARIMA tot i que hi hagi la possibilitat que les dades puguin ser modelitzades millor amb un model ARFIMA.

Aquest és un aspecte que ha de ser analitzat amb més detall en un futur, però es pretén destacar en aquest punt que es tracta d'aspectes diferents així com defensar que en ocasions pot ser interessant usar diferents models per les mateixes dades quan els objectius de l'anàlisi són diferents.

## 8.4. Conseqüències pel càlcul de prediccions

### 8.4.1. Introducció

Tal com ja s'ha comentat al llarg de la tesi, en la mesura que el tipus de memòria d'una sèrie temporal determina l'estructura de la seva dependència temporal, a l'hora de realitzar prediccions per aquestes dades s'ha de tenir en compte el tipus de memòria. Així, una memòria més important està associada a una dependència major de cada valor respecte als valors passats de la variable. Òbviament, si no es té en compte aquest aspecte, s'obtenen estimacions ineficients.

Formalment, recordi's que sota el supòsit d'estacionarietat, el predictor lineal òptim a partir dels valors  $X_1, \dots, X_T$  pel valor d' $X_{T+1}$  es pot calcular com una combinació lineal d'aquests valors<sup>30</sup>:

$$\tilde{X}_{T+1} = \phi_{T1} X_T + \dots + \phi_{TT} X_1 \quad (8.57)$$

on els coeficients  $\phi_{Tj}$  es poden calcular amb l'algorisme Durbin-Levinson<sup>31</sup>.

Aquest mètode és, a més, fàcilment extensible a la predicció per períodes més allunyats. Així, per calcular el predictor òptim a partir dels valors  $X_1, \dots, X_T$  pel valor de la variable en el moment  $T+k$ , es pot usar l'expressió (8.57) substituint les dades no observades per les seves prediccions, és a dir:

$$\tilde{X}_{T+k} = \phi_{T1} \tilde{X}_{T+k-1} + \dots + \phi_{Tk-1} \tilde{X}_{T+1} + \phi_{Tk} X_T + \dots + \phi_{TT} X_k \quad (8.58)$$

<sup>30</sup> L'anàlisi següent es centra en el concepte de predictor òptim entès com el predictor sense biaix i que minimitza l'error quadràtic mitjà de manera que, sota aquest punt de vista i usant resultats bàsics d'Estadística i el supòsit d'estacionarietat, el predictor òptim d' $X_{T+k}$  a partir de  $X_1, \dots, X_T$  és  $\tilde{X}_{T+k} = E(X_{T+k} | X_1, \dots, X_T)$ .

<sup>31</sup> Vegi's l'Annex 6.1 per una síntesi d'aquest algorisme. També és possible calcular les prediccions a partir del valor de les innovacions del procés enlloc de a partir dels valors de la variable usant l'algorisme de la innovació presentat al mateix Annex. De totes maneres, almenys a efectes expositius és més interessant interpretar les prediccions en termes dels valors observats de la variable.

Una de les causes de la popularitat dels models ARMA radica en les facilitats que suposa la seva utilització pel càlcul de prediccions ja que no cal calcular explícitament els coeficients de l'equació (8.57) sinó que es poden calcular de manera més senzilla les prediccions a partir dels coeficients autoregressius i mitjana mòbil del model ARMA<sup>32</sup>.

De totes maneres, el punt de partida d'aquest apartat és el perill que suposa usar un model ARMA o ARIMA per calcular prediccions d'una certa variable econòmica quan aquests models no recullen de manera adequada el tipus de memòria d'aquesta variable. Així, si la variable conté memòria llarga i s'utilitza un model ARIMA o ARMA, les prediccions que s'obtenen no tenen perquè presentar cap de les característiques desitjables en termes del biaix i de la variància.

De totes maneres també s'ha de tenir en compte que el càlcul de prediccions en el cas dels models ARFIMA és bastant més complexa que en el cas dels models ARMA. Per tant, no es pot descartar *a priori* la possibilitat que pugui ser millor usar un model ARMA que un model ARFIMA per calcular prediccions, fins i tot d'una variable que s'ha generat a partir d'un procés ARFIMA. Si a més es tenen en compte les especials dificultats que presenta actualment l'estimació de models ARFIMA, això no fa més que reforçar aquesta idea.

La diferència fonamental a efectes de predicció entre models ARMA i models ARFIMA es troba en que aquests últims no es poden representar a partir de polinomis en el retard  $L$  amb un nombre finit de coeficients sinó que calen infinits coeficients mentre que a la pràctica es disposa sols d'un nombre finit d'observacions de la variable i per tant, cal recórrer a algun tipus d'aproximació.

Per il·lustrar aquest problema, pensi's en el model més senzill que conté memòria llarga, un model ARFIMA(0,d,0). Si  $(1-L)^d X_t = \varepsilon_t$ , hom podria pensar en representar aquest procés en

---

<sup>32</sup> Vegi's el Capítol 5 de Box i Jenkins (1976) per una presentació dels mètodes de predicció en models ARMA i ARIMA.

forma autoregressiva i aplicar els principis de la predicció en models AR però aquesta representació és :

$$\begin{aligned} X_t &= dX_{t-1} + \frac{1}{2}d(1-d)X_{t-2} + \frac{1}{6}d(1-d)(2-d)X_{t-3} + \dots + \varepsilon_t = \\ &= \sum_{j=1}^{\infty} \frac{-\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} X_{t-j} + \varepsilon_t \end{aligned} \quad (8.59)$$

de manera que no està clar com cal calcular la predicció, per exemple, d' $X_{T+j}$  a partir de les observacions  $X_1, \dots, X_T$ . Per trobar-ne la solució a aquest problema una opció és la de calcular els coeficients de la projecció (8.57). De fet, en el cas d'un model ARFIMA(0,d,0) es pot comprovar com els coeficients d'aquesta projecció són<sup>33</sup>:

$$\phi_{Tj} = -\binom{T}{j} \frac{\Gamma(j-d)\Gamma(T-d-j+1)}{\Gamma(-d)\Gamma(T-d+1)} \quad (8.60)$$

i per tant, el problema de la predicció en models ARFIMA(0,d,0) està resolt a partir d'aquest resultat. El problema és que no es disposa d'expressions equivalents a (8.60) per models ARFIMA més generals. De fet, tenint en compte els problemes plantejats en capítols anteriors a l'hora de calcular les autocovariàncies i autocorrelacions dels models ARFIMA, no sembla que sigui possible trobar expressions explícites per aquests coeficients excepte en casos molt concrets. En definitiva, que per avançar en el càlcul de prediccions a partir de models ARFIMA cal avançar per altres vies.

Però abans de continuar per aquestes altres vies, cal destacar un aspecte addicional que obliga a posar especial èmfasi en la distinció dels models ARMA o ARIMA i els models ARFIMA pel que fa al càlcul de prediccions. Així, a més de les diferències pel que fa al mètode de càlcul, les prediccions basades en uns o altres mètodes presenten característiques molt diferents pel que fa a variància d'aquestes prediccions i per tant a la seva precisió.

<sup>33</sup> Per una demostració vegi's Brockwell i Davis (1992).

Si es nota amb  $E_T X_{T+k}$  el valor esperat d' $X_{T+k}$  quan es disposa d'informació sobre la variable fins al moment  $T$ , és a dir:

$$E_T X_{T+k} = E(X_{T+k} | X_1, \dots, X_T) \quad (8.61)$$

la variància del predictor òptim de  $X_{T+k}$  a partir de les dades  $X_1, \dots, X_T$  és:

$$\xi_k = \text{var}(X_{T+k} - E(X_{T+k})) \quad (8.62)$$

El comportament d'aquesta variància com funció de l'horitzó de predicció  $k$ , està molt relacionada amb la *funció variància-temps*  $\sigma_k^2$  introduïda a l'apartat 2.3. Basant-se en els resultats presentats al capítol 4 sobre el comportament de les sumes parcials en els models ARFIMA, Diebold i Lindner (1996) analitzen el comportament asimptòtic d'aquestes variàncies.

Concretament, si un procés  $X_t$  és soroll blanc o  $I(0)$ , es compleix que  $\xi_k = O(k^0) = O(1)$ . Per tant la variància dels errors de predicció està acotada, és a dir, no augmenta indefinidament. En canvi, si  $X_t$  és un camí aleatori o és  $I(1)$ , es compleix  $\xi_k = O(k)$ . Observi's que en aquest cas, la variància no està acotada sinó que tendeix cap a infinit però menys ràpidament que  $k$  ja que el quocient  $\xi_k/k$  està acotat.

De fet, ja s'ha comentat anteriorment que la limitació dels models ARIMA es troba en que sols permeten memòria curta o memòria permanent però no permeten situacions a mig camí entre les dues, és a dir, memòria llarga estacionària o bé memòria transitòria no estacionària. En termes de predicció, això es tradueix en que les prediccions sols poden ser d'ordre 0 o bé d'ordre 1.

Els processos integrats fraccionalment permeten una major varietat de situacions ja que en el cas estacionari, quan  $-1/2 < d < 1/2$ , es compleix  $\xi_k = O(1)$  mentre que en el cas no estacionari s'obté  $\xi_k = O(k^{2d-1})$ .

Per tant, és important tenir en compte la diferent precisió d'aquestes prediccions a l'hora d'interpretar-les i valorar-les en termes econòmics. Per això, en el context plantejat al llarg de la tesi de diferents tipus de memòria, és important el contrast de si les dades tenen memòria permanent o transitòria, sense que això sigui equivalent a ser integrades o no, perquè en un cas  $\xi_k = O(1)$  mentre que en l'altre  $\xi_k = O(k^{2d-1})$ .

En tot cas, i fins on arriba la nostra informació, encara hi ha varies preguntes sense resposta sobre molts aspectes relacionats amb la predicció amb dades que són susceptibles de presentar memòria llarga. Entre aquests temes en destaquen tres als que es pretén donar resposta al llarg d'aquest capítol.

- En primer lloc, i atès que per calcular prediccions en models ARFIMA cal usar algun tipus d'aproximació finita a un model que és infinit, no està clar quin mètode de càlcul és el més idoni.
- Per altra banda, és important delimitar la importància que pot tenir en el tema de les prediccions el fet de no considerar la possibilitat que les variables econòmiques continguin memòria llarga.
- Finalment, i tenint en compte els problemes en l'estimació dels models ARFIMA no està clar quina és la millor estratègia a seguir quan hom necessita calcular prediccions d'una variable econòmica sobre la que hi ha la sospita que tot i no ser integrable presenta una singularitat en la densitat espectral.

A partir d'aquesta perspectiva, als apartats 8.4.2, 8.4.3 i 8.4.4 es desenvolupen algunes aportacions amb l'objectiu de donar resposta a aquestes preguntes.

#### 8.4.2. Mètodes alternatius de càlcul de prediccions en models ARFIMA

Després de presentada una visió global de la problemàtica del tipus de memòria en el càlcul de prediccions, l'objectiu fonamental de l'apartat consisteix a respondre a la pregunta de com calcular prediccions en models ARFIMA. Ja s'ha comentat que l'origen del problema es troba en el fet que el terme  $(1-L)^d$  no té representació com a polinomi finit i per tant, cal trobar-ne una aproximació finita.

De totes maneres, hi ha una via molt intuïtiva per abordar el problema. Suposi's que es disposa de  $T$  observacions d'una variable generada segons un model ARFIMA  $\phi(L)(1-L)^d X_t = \theta(L)\varepsilon_t$ , amb  $d=1$ . En aquest cas, la sèrie de les primeres diferències de la variable  $X_t$  és estacionària i segueix un model ARMA. Per tant, sembla una bona estratègia treballar amb la sèrie  $Y_t=(1-L)X_t$ , per calcular-ne prediccions d' $Y_t$  a partir del model ARMA i, finalment, trobar les prediccions de la variable original  $X_t$ , integrant les prediccions de la variable  $Y_t$ <sup>34</sup>.

De fet, aquesta idea és fàcilment extensible als models ARFIMA amb paràmetre  $d$  qualsevol. Així, suposi's que usant observacions  $X_1, \dots, X_T$  es pretén calcular una predicció per  $T+k$  sabent que  $X_t$  s'ha generat a partir d'un model ARFIMA( $p, d, q$ ). El mètode més immediat per fer-ho consisteix a definir un nou procés  $Y_t=(1-L)^d X_t$  que es comporta segons un procés ARMA( $p, q$ ). Pel que fa a la variable  $Y_t$ , es poden usar les tècniques habituals de predicció en models ARMA per calcular prediccions de  $Y_{T+1}, \dots, Y_{T+k}$  i a partir d'aquestes es poden calcular prediccions per  $X_t$ .

Més concretament, les etapes d'aquest mètode que en direm Mètode 1 per calcular la predicció d' $X_{T+k}$  són les següents:

---

<sup>34</sup> De fet, aquesta és la idea subjacent en el mètode de calcular prediccions proposat per Box i Jenkins (1976) per models ARIMA.

1) En primer lloc, calcular  $Y_1, \dots, Y_T$  a partir de la següent aproximació<sup>35</sup>:

$$\begin{aligned}
 Y_t &= (1-L)^d X_t = X_t - dX_{t-1} + \frac{1}{2}d(1-d)X_{t-2} + \dots + \prod_{i=1}^{t-1} \frac{i-1-d}{i} X_1 = \\
 &= \sum_{j=0}^{t-1} \zeta_j X_{t-j}
 \end{aligned}
 \tag{8.63}$$

on  $\zeta_0=1$  i  $\zeta_j = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}$  per  $j>0$ .

2) Si es compleix que  $\phi(L)(1-L)^d X_t = \theta(L)\varepsilon_t$ , llavors es poden considerar els valors  $Y_1, \dots, Y_T$  com observacions del procés ARMA  $\phi(L)Y_t = \theta(L)\varepsilon_t$ . Aprofitant això, es poden calcular prediccions per  $Y_{T+1}, \dots, Y_{T+k}$  usant els coeficients dels polinomis  $\phi(L)=1-\phi_1L-\phi_2L^2-\dots-\phi_pL^p$  i  $\theta(L)=1-\theta_1L-\theta_2L^2-\dots-\theta_qL^q$ :

$$\tilde{Y}_{T+j} = \phi_1 Y_{T+j-1}^* + \dots + \phi_p Y_{T+j-p}^* + \tilde{\varepsilon}_{T+j} - \theta_1 \tilde{\varepsilon}_{T+j-1} - \dots - \theta_q \tilde{\varepsilon}_{T+j-q}
 \tag{8.64}$$

on  $Y_t^*$  és el valor d' $Y_t$  si  $1 \leq t \leq T$  o bé la predicció  $\tilde{Y}_t$  quan  $t > T$ ; i  $\tilde{\varepsilon}_t$  són estimacions de les innovacions, és a dir  $\tilde{\varepsilon}_t = Y_t - \tilde{Y}_t$  quan  $1 \leq t \leq T$  i  $\tilde{\varepsilon}_t = 0$  quan  $t > T$ .

3) A continuació, amb la informació disponible es poden calcular prediccions d'  $X_{T+1}, \dots, X_{T+k}$ . Com  $(1-L)^d X_t = Y_t$ , es compleix, pels resultats del capítol 4, la següent propietat:

$$X_t = Y_t + \sum_{j=1}^{\infty} \pi_j X_{t-j}
 \tag{8.65}$$

on:

$$\pi_j = -\zeta_j = \frac{-\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}
 \tag{8.66}$$

Tot i que (8.65) és una representació no aplicable a la pràctica ja que es tracta d'una suma infinita, pot usar-se una aproximació finita per calcular prediccions per la variable  $X_t$ :

<sup>35</sup> A partir dels resultats del capítol 4 de la tesi.



$$\tilde{X}_{t+i} = \tilde{Y}_{t+i} + \sum_{j=1}^m \pi_j X_{t+i-j}^* \quad i = 1, \dots, k \quad (8.67)$$

on  $X_{t+i-j}^*$  és el valor observat de la variable quan  $i-j \leq 0$  o bé la seva predicció si  $i-j > 0$  i  $\tilde{Y}_{t+j}$  és la predicció calculada a l'etapa 2.

Naturalment, per a que aquest mètode sigui aplicable, cal que  $m \leq T$ . En tot cas, no està clar quin és el nombre de termes  $m$  òptim per a que l'error de predicció sigui menor tot i que és d'esperar que l'aproximació serà tant millor com major sigui  $m$ , de manera que sembla raonable usar  $m=T$ .

De fet, a la literatura es poden trobar varies propostes basades en la mateixa idea de truncar el polinomi  $(1-L)^d$  en forma de polinomi finit però usant mètodes d'aproximació diferents. D'una banda, Geweke i Porter-Hudak (1983) proposen calcular directament les prediccions a partir de la representació autoregressiva infinita del model d' $X_t$ , proposta que anomenarem Mètode 2. De fet, si  $\phi(L)(1-L)^d X_t = \theta(L)\varepsilon_t$ , es pot representar el procés  $X_t$  com un procés autoregressiu infinit:

$$X_t = \sum_{j=1}^{\infty} \lambda_j X_{t-j} + \varepsilon_t \quad (8.68)$$

on els coeficients  $\lambda_j$  es poden calcular a partir del següent producte de polinomis en  $L$ :

$$\lambda(L) = (1-L)^d \phi(L) \theta(L)^{-1} \quad (8.69)$$

on cal tenir en compte que els coeficients del polinomi  $(1-L)^d$  són els coeficients  $\zeta_j$  de (8.63). Així, la predicció d' $X_{t+i}$  per  $i=1, \dots, k$  es pot obtenir a partir de la següent expressió:

$$\tilde{X}_{t+i} = \sum_{j=1}^m \lambda_j X_{t+i-j}^* \quad (8.70)$$

on  $X_{t+i-j}^*$  és el valor observat de la variable quan  $i \leq j$  o bé la seva predicció si  $i > j$ .

D'altra banda, a Peiris i Perea (1988) es desenvolupa un mètode formalment més correcte pel càlcul de les prediccions ja que se'n demostren les seves propietats. A aquest mètode l'anomenem Mètode 3.

Concretament, en aquest últim treball es demostra que si  $\tilde{Y}_t$  és el predictor òptim de  $Y_t$  quan es compleix  $(1-L)^d X_t = Y_t$ , el predictor òptim de la variable  $X_t$  compleix la següent condició<sup>36</sup>:

$$\sum_{j=0}^{k-1} \sum_{i=0}^j (-1)^i \binom{d}{i} \tilde{X}_{t+k-j} = \sum_{j=0}^{\infty} \sum_{i=j}^{j+k} (-1)^{i+1} \binom{d}{i} X_{t-j} + \sum_{j=1}^k \tilde{Y}_{t+j} \tag{8.71}$$

Aquesta condició és una condició teòrica que es refereix al supòsit abstracte en que es disposa d'una realització infinita del procés estocàstic  $X_t$ . En tot cas, l'aplicació formal d'aquesta condició quan es disposa d'un nombre finit d'observacions  $X_1, \dots, X_T$  i es pretén predir el valor de  $X_{T+k}$  dona lloc a la següent igualtat:

$$\sum_{j=0}^{k-1} \sum_{i=0}^j (-1)^i \binom{d}{i} \tilde{X}_{T+k-j} = \sum_{j=0}^{\infty} \sum_{i=j}^{j+k} (-1)^{i+1} \binom{d}{i} X_{T-j} + \sum_{j=1}^k \tilde{Y}_{T+j} \tag{8.72}$$

Aquesta condició permet calcular de manera recursiva les prediccions d' $X_{T+j}$ , per  $j=1, \dots, K$ :

$$\tilde{X}_{T+1} = \tilde{Y}_{T+1} + \sum_{j=0}^{T-1} \sum_{i=j}^{j+1} (-1)^{i+1} \binom{d}{i} X_{T-j} \tag{8.73}$$

$$\tilde{X}_{T+2} = d\tilde{X}_{T+1} + \tilde{Y}_{T+1} + \tilde{Y}_{T+2} + \sum_{j=0}^{T-1} \sum_{i=j}^{j+k} (-1)^{i+1} \binom{d}{i} X_{T-j} \tag{8.74}$$

i així successivament.

<sup>36</sup> De fet, a Peiris i Perea (1988) es demostra una propietat més general, que si  $(1-L)^d X_t = Y_t$ , es compleix:

$$\sum_{j=0}^{k-1} \sum_{i=s-1}^{(j+1)s-1} (-1)^i \binom{d}{i} \tilde{X}_{t+k-j} = \sum_{j=0}^{\infty} \sum_{i=s+j}^{(j+1)s+k-1} (-1)^{i+1} \binom{d}{i} X_{t-j} + \sum_{j=1}^k \tilde{Y}_{t+j}$$

on  $\tilde{X}_t$  i  $\tilde{Y}_t$  són els predictors sense biaix que minimitzen l'error quadràtic mitjà.

Com a aplicació particular de l'expressió (8.72), si  $\phi(L)(1-L)^d X_t = \theta(L)\varepsilon_t$ , llavors es pot aplicar un mètode similar al Mètode 1. Concretament, es poden aplicar les etapes 1 i 2 del Mètode 1 i després usar (8.72) per trobar les prediccions de la variable  $X_t$ .

En tot cas, els dos primers mètodes, especialment el primer és molt més senzill d'aplicar pel que té interès valorar quin és el guany d'usar un o altre mètode. A més, a efectes pràctics i amb mostres finites, la determinació del millor mètode d'estimació és un tema que no està encara resolt ja que els models fraccionals no admeten una representació a l'espai dels estats (*state-space*).

Per aquest motiu, s'ha realitzat una comparació de l'Error Quadràtic Mitjà (EQM) associat a cada un dels mètodes. Però per fer aquesta comparació cal tenir present que a més dels problemes que presenta el càlcul de prediccions associats al mètode d'aproximació finita usat, són també importants els problemes associats amb l'estimació dels paràmetres del model ARFIMA. De fet, tal i com ja s'ha comentat, aquests problemes són tan importants que no està clar si, tot i que una certa variable presenti memòria llarga, pot ser millor usar un model ARMA o ARIMA pel càlcul de les prediccions ja que l'estimació dels models ARMA o ARIMA presenta molts menys problemes i el càlcul de prediccions òptimes és molt senzill.

Per poder treure conclusions clares cal distingir aquestes dues fonts d'errors, és a dir, els problemes derivats de l'estimació i els problemes de l'aproximació finita a un model infinit. Per fer-ho, s'ha realitzat una simulació on s'ha suposat conegut el procés generador de les dades per evitar la necessitat d'estimar els paràmetres del model.

Per tant, l'interès d'aquesta simulació es troba en determinar quin és el millor mètode pel càlcul de prediccions quan les dades s'han generat segons un model ARFIMA sense entrar ni en els problemes d'identificació del tipus de memòria ni en els problemes d'estimació del model. Així, s'han generat 5000 sèries de 120 observacions seguint un model ARFIMA(1,d,0) amb diferents valors dels paràmetres  $d$  i  $\phi$ . D'aquestes sèries, s'han usat les

primeres 100 observacions per predir les 20 últimes usant cada un dels tres mètodes proposats més amunt.

A més, s'ha cregut interessant comparar els EQM de cada un d'aquests mètodes amb el que s'obté si es calculen les prediccions usant algun mètode no paramètric, és a dir, que no es basi en el coneixement del tipus de model. Concretament, el Mètode 4 usat es basa en aplicar l'algorisme Durbin-Levinson presentat a l'Annex 6.1, usant les autocovariàncies del model amb que s'han generat les dades.

Cal assenyalar també que, amb l'objectiu de poder fer comparacions entre els resultats que s'obtenen per diferents models, s'han generat totes les sèries a partir de models ARFIMA on la variància del terme soroll blanc no és unitària, sinó que s'ha fixat aquesta variància per a que la variància d' $X_t$  (la teòrica) sigui unitària.

Als quadres 8.18 a 8.21 es presenta un resum dels principals resultats obtinguts usant cada un dels quatre mètodes. A partir d'aquests resultats s'extreuen, com a principals conclusions, les següents:

1. En primer lloc, observi's com tots els resultats són coherents amb el fet que l'EQM ha de tendir cap a  $u$  (la variància d' $X_t$ ) a mesura que augmenta el nombre d'observacions<sup>37</sup>. Dit d'una altra manera, a mesura que es pretén predir dades més allunyades en el temps, arriba un moment que les dades disponibles no aporten informació rellevant per a aquesta predicció.

---

<sup>37</sup> Això és així per qualsevol dels tres primers mètodes però no per l'algorisme Durbin-Levinson. Això és degut a que amb aquest algorisme no es té en compte tota la informació disponible. De fer, es pot comprovar com amb aquest mètode l'EQM també tendeix asimptòticament a  $u$  però és normal trobar-se amb errors superiors per valors de  $k$  no gaire grans.

Quadre 8.18. EQM del predictor d' $X_{T+k}$  si  $(1-L)^d(1-\phi L)X_t = \varepsilon_t$ . Mètode 1.

d	$\phi$	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
0	0.0	1.041	0.974	0.954	1.011	1.053	1.038	0.965	0.997	0.936	0.960
	0.4	0.942	0.968	1.069	1.101	0.986	0.958	1.011	1.149	1.125	0.896
	0.8	0.582	0.715	0.821	0.893	0.948	0.974	1.016	0.992	1.047	1.044
0.1	0.0	0.837	0.898	0.890	0.918	0.983	0.927	0.946	0.894	0.893	0.871
	0.4	0.854	0.908	0.966	1.002	0.994	0.898	0.893	0.893	1.012	0.945
	0.8	0.436	0.563	0.691	0.805	0.876	0.938	0.944	0.966	0.962	0.977
0.2	0.0	0.820	0.940	0.891	0.912	0.889	0.857	0.895	0.922	0.968	0.813
	0.4	0.746	0.851	0.946	1.087	0.905	0.988	0.997	1.011	0.928	0.809
	0.8	0.308	0.488	0.635	0.713	0.798	0.879	0.918	0.951	0.969	0.945
0.3	0.0	0.818	0.886	0.859	0.915	0.897	0.858	0.935	0.851	0.906	0.769
	0.4	0.591	0.763	0.931	0.965	0.887	0.953	0.884	0.814	0.790	0.880
	0.8	0.208	0.395	0.517	0.622	0.738	0.790	0.828	0.862	0.871	0.834
0.4	0.0	0.617	0.691	0.684	0.684	0.771	0.783	0.718	0.767	0.753	0.657
	0.4	0.371	0.555	0.623	0.706	0.729	0.696	0.740	0.717	0.656	0.669
	0.8	0.130	0.243	0.360	0.442	0.524	0.586	0.630	0.646	0.652	0.648

Quadre 8.19. EQM del predictor d' $X_{T+k}$  si  $(1-L)^d(1-\phi L)X_t = \varepsilon_t$ . Mètode 2.

d	$\phi$	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
0	0.0	1.041	0.974	0.954	1.011	1.053	1.038	0.965	0.997	0.936	0.960
	0.4	0.942	0.968	1.069	1.101	0.986	0.958	1.011	1.149	1.125	0.896
	0.8	0.582	0.715	0.821	0.893	0.948	0.974	1.016	0.992	1.047	1.044
0.1	0.0	0.837	0.907	0.891	0.920	0.984	0.926	0.945	0.896	0.893	0.870
	0.4	0.854	0.940	0.982	1.013	0.995	0.902	0.895	0.896	1.013	0.947
	0.8	0.436	0.590	0.703	0.813	0.883	0.942	0.948	0.968	0.963	0.978
0.2	0.0	0.820	0.985	0.900	0.928	0.892	0.862	0.899	0.920	0.966	0.814
	0.4	0.746	0.944	0.978	1.105	0.913	0.994	1.001	1.011	0.929	0.809
	0.8	0.308	0.536	0.662	0.729	0.810	0.887	0.922	0.954	0.972	0.947
0.3	0.0	0.818	0.943	0.879	0.923	0.902	0.861	0.936	0.850	0.909	0.770
	0.4	0.591	0.855	0.964	0.977	0.905	0.968	0.902	0.822	0.799	0.885
	0.8	0.208	0.446	0.540	0.635	0.747	0.794	0.831	0.862	0.869	0.831
0.4	0.0	0.617	0.760	0.710	0.709	0.782	0.790	0.727	0.766	0.759	0.661
	0.4	0.371	0.658	0.665	0.738	0.744	0.711	0.747	0.725	0.654	0.667
	0.8	0.130	0.305	0.399	0.468	0.541	0.597	0.636	0.650	0.651	0.644

Quadre 8.20. EQM del predictor d' $X_{T+k}$  si  $(1-L)^d(1-\phi L)X_t = \varepsilon_t$ . Mètode 3.

d	$\phi$	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
0	0.0	1.041	0.974	0.954	1.011	1.053	1.038	0.965	0.997	0.936	0.960
	0.4	0.788	0.940	1.060	1.103	0.986	0.959	1.011	1.149	1.125	0.896
	0.8	0.359	0.590	0.754	0.856	0.933	0.960	1.012	0.997	1.050	1.044
0.1	0.0	0.837	0.898	0.890	0.918	0.983	0.927	0.946	0.894	0.893	0.871
	0.4	0.715	0.873	0.957	0.990	0.993	0.895	0.891	0.892	1.013	0.945
	0.8	0.255	0.439	0.599	0.732	0.820	0.907	0.927	0.958	0.948	0.970
0.2	0.0	0.820	0.940	0.891	0.912	0.889	0.857	0.895	0.923	0.967	0.813
	0.4	0.608	0.799	0.940	1.080	0.901	0.986	0.993	1.011	0.928	0.809
	0.8	0.185	0.355	0.519	0.627	0.730	0.828	0.883	0.922	0.942	0.926
0.3	0.0	0.818	0.886	0.859	0.915	0.897	0.858	0.935	0.851	0.905	0.768
	0.4	0.452	0.701	0.907	0.962	0.881	0.949	0.876	0.815	0.785	0.876
	0.8	0.119	0.284	0.413	0.532	0.654	0.729	0.781	0.824	0.842	0.808
0.4	0.0	0.616	0.691	0.684	0.684	0.770	0.783	0.717	0.765	0.749	0.652
	0.4	0.312	0.511	0.606	0.694	0.722	0.692	0.740	0.712	0.651	0.668
	0.8	0.074	0.167	0.278	0.358	0.441	0.509	0.558	0.591	0.613	0.615

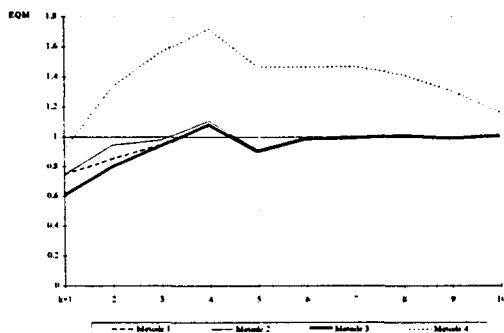
Quadre 8.21. EQM del predictor d' $X_{T+k}$  si  $(1-L)^d(1-\phi L)X_t = \varepsilon_t$ . Mètode 4.

d	$\phi$	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
0	0.0	1.303	1.186	1.213	1.205	1.239	1.175	1.074	1.156	1.094	1.072
	0.4	1.027	1.195	1.363	1.364	1.194	1.153	1.202	1.383	1.307	1.044
	0.8	0.708	0.961	1.166	1.320	1.446	1.502	1.565	1.546	1.506	1.490
0.1	0.0	1.080	1.141	1.146	1.127	1.192	1.146	1.147	1.072	1.089	1.068
	0.4	0.970	1.291	1.504	1.424	1.487	1.305	1.246	1.304	1.379	1.261
	0.8	0.602	0.840	1.090	1.346	1.528	1.658	1.690	1.700	1.665	1.658
0.2	0.0	1.217	1.324	1.291	1.330	1.270	1.243	1.280	1.318	1.312	1.111
	0.4	0.912	1.333	1.562	1.719	1.467	1.465	1.469	1.414	1.307	1.168
	0.8	0.547	0.875	1.298	1.649	1.896	2.079	2.177	2.214	2.251	2.191
0.3	0.0	1.181	1.416	1.460	1.550	1.556	1.510	1.485	1.451	1.463	1.200
	0.4	1.087	1.752	2.205	2.384	2.404	2.470	2.108	1.987	2.048	1.923
	0.8	0.457	0.748	1.054	1.440	1.814	2.094	2.332	2.484	2.492	2.435
0.4	0.0	1.070	1.217	1.357	1.499	1.599	1.612	1.500	1.484	1.330	1.259
	0.4	0.658	1.059	1.373	1.688	1.757	1.816	1.797	1.784	1.616	1.471
	0.8	0.348	0.525	0.813	1.132	1.477	1.762	2.062	2.311	2.456	2.488

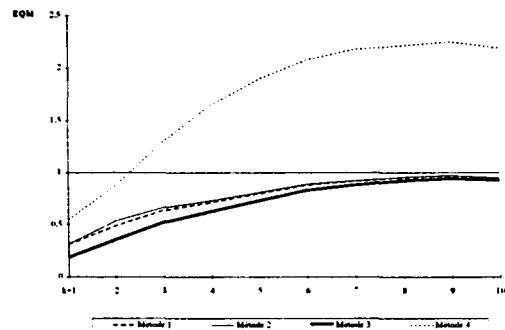
De totes maneres, observi's també que aquesta convergència cap a la unitat és molt diferent segons quin sigui el model amb que s'han generat les dades. Així, per models ARFIMA(1,0,0) amb memòria curta, l'EQM tendeix molt ràpidament cap a la unitat. Fins i tot per models ARFIMA(1, $d$ ,0) quan  $d=0.2$ , aquesta convergència és relativament ràpida sigui quin sigui el valor del paràmetre autoregressiu com reflexen els gràfics 8.11 i 8.12.

En canvi, quan  $d=0.4$  aquesta convergència és molt lenta com reflexen els gràfics 8.13 i 8.14. Això és important ja que demostra que si hom pren com objectiu predir els valors futurs d'una variable econòmica, en el cas de memòria curta no té sentit predir més enllà de 4 o 5 dades, mentre que si aquesta presenta memòria llarga, es disposa d'informació per predir bastants més dades, almenys en el cas en que el model és conegut.

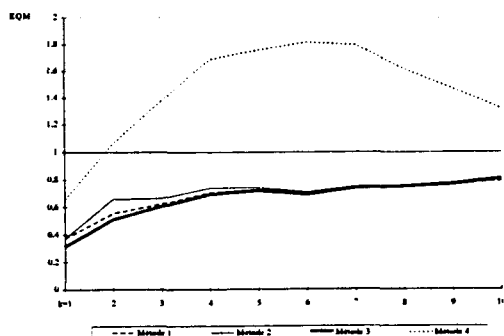
**Gràfic 8.11.** EQM del predictor d' $X_{T+k}$  si  $(1-L)^d(1-\phi L)X_t=\varepsilon_t$ , amb  $d=0.2$  i  $\phi=0.4$ .



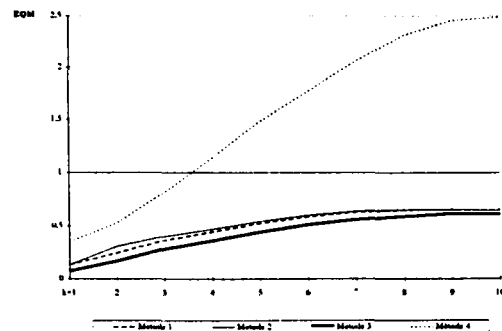
**Gràfic 8.12.** EQM del predictor d' $X_{T+k}$  si  $(1-L)^d(1-\phi L)X_t=\varepsilon_t$ , amb  $d=0.2$  i  $\phi=0.8$ .



**Gràfic 8.13.** EQM del predictor d' $X_{T+k}$  si  $(1-L)^d(1-\phi L)X_t=\varepsilon_t$ , amb  $d=0.4$  i  $\phi=0.4$ .



**Gràfic 8.14.** EQM del predictor d' $X_{T+k}$  si  $(1-L)^d(1-\phi L)X_t=\varepsilon_t$ , amb  $d=0.4$  i  $\phi=0.8$ .



2. Pel que fa al biaix de les prediccions, es pot assegurar que cap dels 4 mètodes considerats presenta biaix ja que no s'ha obtingut cap biaix estimat a partir de les 5000 rèpliques superior en valor absolut a 0.01. Per tant, les diferències entre mètodes són exclusivament en termes de la variància del predictor.
3. Es pot comprovar com el menor EQM s'obté, com era d'esperar i en tots els casos, usant el mètode 3, és a dir, el proposat per Peiris i Perea (1988).
4. Tal com era també previsible, l'algorisme Durbin-Levinson porta a obtenir errors molt superiors als altres mètodes. De fet, s'obtenen valors de l'EQM superiors a la variància de la sèrie per valors de  $k$  no gaire grans. Tot i que no es recullen a les taules de resultats, per valors de  $k$  al voltant de 15 o 20, aquest EQM tendeix a la unitat. També es pot observar com aquest mètode és especialment dolent per valors del paràmetre autoregressiu grans.
5. Els mètodes 1 i 2 presenten resultats molt similars i molt propers als resultat obtinguts amb el mètode 3. De fet, aquests dos primers mètodes tendeixen a presentar errors similars al mètode 3 a mesura que augmenta  $k$ , de manera que per valors  $k=15$  o  $k=20$  els tres mètodes són equivalents. Això suggereix que per models amb memòria llarga, si l'objectiu és calcular gran nombre de prediccions, el mètode 1 proposat a la tesi com a mètode d'especial simplicitat presenta resultats excel·lents.

#### **8.4.3. Conseqüències d'ignorar la presència de memòria llarga**

Seguint l'ordre abans esmentat, un segon nivell d'anàlisi és el de les conseqüències que pot tenir el fet d'ignorar la presència de memòria llarga cara al càlcul de prediccions.

En aquest sentit, al treball de Lyhagen (1997) es demostra que, a nivell teòric, el fet d'ignorar la presència de memòria llarga podria portar a EQM relatius de predicció molt grans. Així,



seguint la proposta de Ray (1993) es pot estudiar aquest efecte teòric de la no consideració de la memòria llarga.

Concretament, si es suposa que  $X_t \sim \text{ARFIMA}(0, d, 0)$ , és a dir,  $(1-L)^d X_t = \varepsilon_t$ , amb  $|d| < 1/2$ , té interès analitzar que passaria si hom ajusta un model  $\text{AR}(p)$  per realitzar prediccions, és a dir, si es suposa que  $\phi(L)X_t = u_t$ . Si es disposa de  $T$  dades,  $X_1, \dots, X_T$ , a partir de l'estimació dels paràmetres autoregressius es pot calcular una predicció puntual per  $X_{T+h}$  a través de la següent fórmula recursiva:

$$\tilde{X}_{T+h} = \sum_{j=1}^p \tilde{\phi}_j X_{T+h-j} \quad h = 1, \dots, k \quad (8.75)$$

Òbviament, l'error comès en aquesta estimació depèn de les propietats dels estimadors però per simplificar l'anàlisi, suposi's que els paràmetres autoregressius de l'aproximació truncada a la representació com  $\text{AR}(\infty)$  de  $X_t$ ,  $\pi(L)X_t = \varepsilon_t$  són coneguts. Recordi's que aquests paràmetres compleixen:

$$\pi_j = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} \quad (8.76)$$

Si s'usa una aproximació del polinomi autoregressiu  $\pi(L)$  usant un número finit,  $p$  de coeficients, l'error quadràtic mitjà (EQM) comès en aquesta estimació es pot calcular a través de la representació de Wold d' $X_t$ :

$$X_t = \phi^{-1}(L)u_t = \theta(L)u_t = \theta_0 u_t + \theta_1 u_{t-1} + \dots \quad (8.77)$$

de manera que:

$$EQM_A(k) = E(\tilde{X}_{t+k} - X_{t+k})^2 = \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} \theta_i \theta_j \gamma_u(i-j) \quad (8.78)$$

Observi's que  $u_t$  no és soroll blanc, sinó que  $u_t = \phi(L)X_t = \phi(L)(1-L)^{-d} \varepsilon_t$ , i, per tant,  $u_t \sim \text{ARFIMA}(0, d, p)$ . Però per cada ordre  $p$  de l'aproximació i per cada valor del paràmetre  $d$  associat a  $X_t$  es poden calcular les autocovariàncies del procés<sup>38</sup>  $u_t$ , i, per tant, es pot avaluar el

<sup>38</sup> Per fer-ho, es pot usar per exemple l'algorisme de càlcul proposat a l'apartat 6.11. que és força senzill i ràpid.

$EQM_A(k)$  per diferents valors de  $k$ . Aquest EQM es pot comparar amb el que s'obtidria si es calculessin les prediccions usant el model "correcte" pel procés. Així, usant (4.26) la representació de Wold d' $X_t$  és:

$$X_t = \psi(L)\varepsilon_t = \sum_{j=0}^{\infty} \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)} \varepsilon_{t-j} \quad (8.79)$$

Com  $\varepsilon_t$  és soroll blanc, l'EQM comès en aquest cas per predir  $X_{T+k}$  és:

$$EQM(k) = E(\tilde{X}_{T+k} - X_{T+k})^2 = \sum_{j=0}^{k-1} \psi_j^2 \quad (8.80)$$

De la comparació dels dos EQM es pot trobar quin és l'augment en percentatge d'aquest error degut al fet d'usar una aproximació autoregressiva. Els quadres 8.22 a 8.24 recullen els resultats de calcular per diferents valors dels paràmetres  $p$ ,  $d$  i  $k$  aquest percentatge<sup>39</sup>:

$$P_p^d(k) = \frac{EQM_A - EQM}{EQM} \times 100 \quad (8.81)$$

De l'anàlisi dels resultats d'aquest càlcul se'n dedueixen varies coses:

- En general, els augments de l'error de predicció poden ser molt importants pel fet d'usar eines dissenyades pensant en dades de memòria curta.
- Amb  $p=1$ , fins i tot la predicció del valor en  $T+1$  si  $d>0.1$ , l'EQM és superior al doble usant l'aproximació autoregressiva que el model correcte i pot arribar a ser fins i tot 20 vegades més gran si  $d>0.45$ .
- Com funció de  $k$ , la pèrdua relativa derivada de l'aproximació té un màxim a  $k+1$  i a partir d'aquí és decreixent quan augmenta el nombre d'observacions predites.

---

<sup>39</sup> Aquesta mesura és la proposada per Davies i Newbold (1980).

Quadre 8.22. Augment percentual de l'EQM,  $P_p^d(k)$ , al realitzar prediccions amb un model AR( $p$ ) amb  $p=1$  quan  $X_t \sim \text{ARFIMA}(0,d,0)$ .

k	d=0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
1	16.73	70.34	169.04	327.64	573.88	964.20	1087.39	1165.73	1938.52
2	18.08	82.03	212.08	440.33	821.85	1462.93	1797.33	1968.84	3373.05
3	12.50	62.94	177.86	398.84	796.10	1503.02	2014.40	2252.91	3980.25
4	9.34	49.09	145.10	340.41	710.08	1398.30	2099.89	2357.18	4255.17
5	7.48	40.34	122.50	295.36	633.32	1281.92	2337.63	3126.50	6156.55
6	6.25	34.43	106.66	262.25	573.32	1182.93	2418.21	3276.64	6555.37
7	5.39	30.16	94.97	237.22	526.58	1102.58	2286.29	3350.57	7007.59
8	4.74	26.92	85.96	217.62	489.31	1037.08	2175.36	3453.23	7568.87
9	4.24	24.37	78.78	201.79	458.82	982.75	2081.72	3591.31	8293.15
10	3.83	22.31	72.89	188.68	433.32	936.86	2001.73	3770.73	9252.23
11	3.51	20.60	67.97	177.61	411.59	897.45	1932.48	4399.17	12672.83
12	3.23	19.16	63.78	168.11	392.80	863.11	1871.77	4288.25	12423.27
13	3.00	17.93	60.17	159.85	376.35	832.84	1817.96	4189.46	12199.71
14	2.80	16.86	57.01	152.59	361.77	805.89	1769.81	4100.70	11998.06
15	2.62	15.92	54.23	146.14	348.76	781.68	1726.36	4020.32	11814.95
16	2.47	15.10	51.76	140.37	337.04	759.77	1686.88	3947.05	11647.63
17	2.34	14.36	49.55	135.17	326.41	739.81	1650.77	3879.84	11493.87
18	2.22	13.70	47.55	130.45	316.73	721.53	1617.57	3817.88	11351.85
19	2.11	13.11	45.73	126.14	307.84	704.70	1586.89	3760.47	11220.06
20	2.01	12.57	44.08	122.19	299.66	689.13	1558.42	3707.06	11097.27

Quadre 8.23. Augment percentual de l'EQM,  $P_p^d(k)$ , al realitzar prediccions amb un model AR( $p$ ) amb  $p=10$  quan  $X_t \sim \text{ARFIMA}(0,d,0)$ .

k	d=0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
1	2.41	9.91	21.74	38.01	60.54	83.64	119.67	114.11	179.67
2	2.64	11.83	28.01	52.46	88.82	129.47	194.10	192.70	314.07
3	2.76	12.95	32.02	62.42	109.62	165.17	254.99	259.75	432.86
4	2.85	13.76	35.04	70.27	126.72	195.60	308.59	320.41	542.84
5	2.91	14.39	37.49	76.86	141.51	222.65	357.39	376.80	646.86
6	2.96	14.91	39.56	82.59	154.68	247.28	402.69	430.02	746.47
7	3.00	15.35	41.37	87.68	166.64	270.05	445.26	480.77	842.59
8	3.04	15.73	42.97	92.28	177.63	291.33	485.63	529.49	935.86
9	3.07	16.07	44.40	96.48	187.84	311.38	524.16	576.51	1026.72
10	3.10	16.37	45.71	100.36	197.40	330.39	561.12	622.07	1115.51
11	3.12	16.64	46.91	103.97	206.40	348.51	596.71	666.36	1202.49
12	2.97	16.31	46.88	105.52	212.13	362.11	625.95	704.86	1281.12
13	2.82	15.83	46.38	106.00	215.87	372.63	650.44	738.69	1352.46
14	2.67	15.32	45.63	105.80	218.16	380.72	670.99	768.45	1417.23
15	2.54	14.79	44.73	105.12	219.34	386.84	688.15	794.60	1475.96
16	2.41	14.27	43.75	104.09	219.64	391.29	702.36	817.50	1529.14
17	2.30	13.77	42.73	102.81	219.24	394.34	713.99	837.48	1577.18
18	2.19	13.28	41.67	101.35	218.26	396.18	723.32	854.79	1620.45
19	2.09	12.81	40.61	99.74	216.81	396.99	730.62	869.67	1659.29
20	2.00	12.36	39.55	98.03	214.97	396.90	736.08	882.33	1694.00

**Quadre 8.24.** Augment percentual de l'EQM,  $P_p^d(k)$ , al realitzar prediccions amb un model AR( $p$ ) amb  $p=20$  quan  $X_t \sim \text{ARFIMA}(0,d,0)$ .

k	d=0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
1	1.23	5.06	10.87	19.01	30.29	41.68	59.44	61.26	99.30
2	1.35	6.05	14.04	26.29	44.50	64.58	96.47	103.49	173.60
3	1.42	6.64	16.08	31.34	55.00	82.47	126.83	139.56	239.30
4	1.46	7.07	17.63	35.35	63.67	97.78	153.61	172.24	300.17
5	1.50	7.40	18.90	38.73	71.21	111.43	178.04	202.64	357.76
6	1.52	7.68	19.99	41.69	77.95	123.90	200.77	231.38	412.94
7	1.55	7.92	20.94	44.33	84.10	135.46	222.18	258.82	466.23
8	1.57	8.13	21.78	46.74	89.78	146.30	242.53	285.19	517.95
9	1.58	8.32	22.55	48.95	95.07	156.55	261.98	310.68	568.38
10	1.60	8.49	23.25	51.00	100.05	166.29	280.69	335.41	617.68
11	1.61	8.64	23.90	52.91	104.75	175.61	298.73	359.47	666.00
12	1.63	8.78	24.50	54.71	109.22	184.54	316.20	382.94	713.46
13	1.64	8.91	25.06	56.41	113.48	193.14	333.16	405.89	760.13
14	1.65	9.04	25.59	58.02	117.56	201.44	349.65	428.37	806.10
15	1.66	9.15	26.08	59.55	121.48	209.48	365.73	450.41	851.43
16	1.67	9.25	26.55	61.02	125.25	217.27	381.42	472.06	896.17
17	1.68	9.35	27.00	62.42	128.88	224.83	396.77	493.34	940.37
18	1.68	9.45	27.42	63.76	132.39	232.19	411.79	514.29	984.06
19	1.69	9.54	27.83	65.05	135.79	239.37	426.50	534.92	1027.28
20	1.70	9.62	28.22	66.29	139.09	246.36	440.94	555.26	1070.07

- Curiosament, per  $d < 0.35$ , quan augmenta el nombre d'observacions a predir es redueix la pèrdua relativa pel fet d'usar un AR. De fet, això passa sols per  $p < 5$  (tot i que no es pot comprovar a partir de les taules presentades).
- Els guanys de passar de  $p=1$  a  $p=10$  són importants però únicament per  $k$  petit o bé quan  $d$  és gran, és a dir, en situacions especialment difícils. En canvi, per  $k$  gran i valors de  $d$  al voltant de 0.2 o 0.3 ni hi ha una millora gaire substancial. Així, per exemple, per  $k=10$  amb  $d=0.2$  es passa d'una pèrdua relativa d'un 189% a una pèrdua relativa d'un 100%.
- Fins i tot en la predicció d'una observació amb  $p=20$ , situacions amb una memòria llarga relativament poc important, porten a pèrdues relatives superiors al 50%.

Però no n'hi ha prou amb aquesta anàlisi teòrica, ja que la necessitat d'estimar els paràmetres del model usat ja sigui un ARMA o un ARFIMA i les dificultats que presenti aquesta

estimació afecten a la qualitat de les prediccions. Així, a l'apartat següent s'amplien els resultats anteriors amb els resultats d'una simulació en que s'estimen els diferents models. Això permet analitzar l'efecte que tenen les diferents fonts d'error sobre la qualitat de les prediccions en models ARMA o ARFIMA.

De totes maneres, tot i que no calgui estimar un model ARFIMA per obtenir bones prediccions de variables que contenen memòria llarga, aquests resultats corroboren la importància que té disposar d'eines que permetin detectar aquest tipus de memòria perquè usant la metodologia Box-Jenkins no és habitual, ni de bon tros, usar models amb tants coeficients autoregressius. Això és degut, a més del principi de parsimònia, al fet que tot i la presència de memòria llarga, no és fàcil obtenir més de 2 o 3 coeficients autoregressius significatius.

A més, dels resultats de l'apartat 8.2 es dedueix que és molt probable sobrediferenciar amb la metodologia habitual les dades amb memòria llarga, fet que portaria a usar paràmetres mitjana mòbil enlloc de paràmetres autoregressius tal i com seria més adequat.

En definitiva, aquests resultats mostren que és fonamental usar els instruments presentats al capítol 5 per identificar el tipus de memòria de les variables econòmiques tot i que després pugui ser interessant quan no és possible estimar amb precisió els paràmetres d'un model ARFIMA, usar altres mètodes de càlcul de les prediccions.

#### ***8.4.4. Conseqüències dels errors d'especificació en el tipus de memòria***

A partir d'un exercici de simulació, Ray (1993) conclou que en certes condicions els problemes d'estimació dels models ARFIMA superen els problemes derivats d'usar un model de memòria curta com és un model autoregressiu. De totes maneres, en aquesta simulació s'utilitza un mètode d'estimació del paràmetre d'integració fraccional que no és gaire habitual.

Concretament, a partir de sèries de 400 observacions generades segons un model  $FI(d)$  s'estimen 395 coeficients autoregressius per MQO i a continuació s'estima el paràmetre  $d$  a partir d'aquests coeficients autoregressius a través d'un mètode d'optimització no lineal. A partir d'aquests resultats es conclou que usant un model  $AR(p)$  amb  $p$  no gaire gran (8, 10, o 12) la pèrdua en termes d'EQM pel fet de no usar un model  $ARFIMA(d)$  no és gaire important si es tenen en compte els problemes en l'estimació del paràmetre  $d$ .

Però d'altra banda, i pel cas de models  $ARFIMA(0,d,0)$ , Smith i Yadav (1994) destaquen que hi ha una pèrdua potencial important pel fet d'ajustar de manera incorrecta un  $AR(p)$  quan  $d > 0$ . Per altra banda, si  $d < 0$ , la pèrdua es produeix únicament en la predicció a un període.

Per poder treure conclusions clares en relació amb aquests aspectes, es realitza a continuació una simulació en que s'apliquen diferents mètodes tant paramètrics com no paramètrics al càlcul de prediccions. Concretament, s'han generat 5000 sèries de 120 observacions segons un procés  $ARFIMA(1,d,0)$ :

$$(1 - \phi L)(1 - L)^d X_t = \varepsilon_t \quad (8.82)$$

amb diferents valors dels paràmetres  $d$  i  $\phi$ . Per cada una d'aquestes sèries s'han utilitzat cinc mètodes diferents per calcular prediccions de les últimes 20 observacions suposant conegudes les primeres 100. Més concretament, els mètodes analitzats han estat els següents:

- A. Càlcul de prediccions usant el mètode de Peiris i Perea (1988) suposant que els paràmetres  $d$  i  $\phi$  són coneguts<sup>40</sup>.
- B. Estimació d'un model  $ARMA(1,1)$  per màxim versemblança i càlcul de prediccions a partir del mètode proposat a Box i Jenkins (1970).
- C. Predicció no paramètrica via algorisme Durbin-Levinson

<sup>40</sup> S'utilitza aquest mètode pel càlcul de prediccions en models  $ARFIMA$  atès que a l'apartat anterior s'ha comprovat que n'és el millor mètode. El motiu d'aplicar un mètode de càlcul de prediccions basat en paràmetres reals i no estimats és per valorar la incidència real que té la necessitat d'estimar els paràmetres en la qualitat de les prediccions que s'obtenen.

- D. Estimació d'un model  $AR(p)$  amb  $p=10$  i càlcul de prediccions de manera recursiva a partir dels coeficients estimats.
- E. Estimació d'un model  $ARFIMA(1,d,0)$  per màxima versemblança i càlcul de prediccions usant el mètode de Peiris i Perea (1988)<sup>41</sup>.

Per cada un d'aquests mètodes, s'han calculat estimacions del biaix i de l'error quadràtic mitjà de les prediccions per  $k=1,..20$ . A partir dels resultats dels quadres 8.25 a 8.29 que en recullen els principals resultats per cada un dels 5 mètodes s'obtenen les següents conclusions:

- Usant un mètode no paramètric com l'algorisme Durbin-Levinson que no requereix de l'estimació dels paràmetres de cap model explícit, s'obtenen unes prediccions de qualitat molt inferior a les que s'obtenen per mètodes paramètrics. Per tant, la primera conclusió de caire general és que tot i el perill d'especificar un model de manera errònia, per la via paramètrica s'obtenen prediccions molt més precises, fins i tot a mig termini.
- Mitjançant l'estimació d'un model  $ARMA(1,1)$ , s'obtenen prediccions prou bones comparades amb les que s'obtenen a partir del model  $ARFIMA(1,0)$  correcte, sobretot per  $k$  petit, excepte en un cas. Concretament, els resultats que s'obtenen usant un model  $ARMA(1,1)$  són especialment negatius quan el procés generador de dades inclou una diferència fraccional amb  $d>0$  i un paràmetre autoregressiu gran (en els quadres  $\phi=0.8$ ). Això és degut a que el model no té la flexibilitat suficient per reproduir simultàniament components de memòria curta i de memòria llarga. En aquest cas, s'obtindrien clarament molts millors resultats amb un model  $ARIMA(1,1,0)$  que amb un model  $ARMA(1,1)$ .

---

<sup>41</sup> S'han estimat els paràmetres del model per màxima versemblança enlloc d'usar algun mètode d'estimació semiparamètric ja que quan el model és conegut, aquest és el mètode d'estimació que presenta millors resultats. En tot cas, per evitar els problemes derivats de la necessitat d'estimar el valor esperat de la variable, s'han generat totes les sèries amb esperança nul·la. A més, s'ha utilitzat el càlcul de prediccions aplicant el mètode de Peiris i Perea (1988) a models estimats per GPH, però els resultats han estat pitjors que usant l'estimador MV.

**Quadre 8.25.** EQM del predictor d' $X_{T+k}$  si  $(1-L)^d(1-\phi L)X_t = \varepsilon_t$ . Mètode A.

d	$\phi$	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
0.0	0.0	1.046	0.958	0.951	0.979	0.990	0.931	0.959	0.970	0.999	0.995
	0.4	0.897	0.936	0.908	0.954	0.960	0.987	1.020	1.048	0.982	1.092
	0.8	0.494	0.622	0.782	0.871	0.944	0.912	0.937	0.942	0.985	0.994
0.1	0.0	0.779	0.859	0.895	0.868	0.804	0.862	0.854	0.854	0.889	0.842
	0.4	0.766	0.966	0.811	0.882	0.840	0.866	0.898	0.852	0.845	0.884
	0.8	0.223	0.414	0.592	0.709	0.774	0.883	0.927	0.954	0.991	0.970
0.2	0.0	0.894	0.894	0.883	0.856	0.890	0.782	0.852	0.905	0.948	0.829
	0.4	0.663	0.849	0.795	0.838	0.860	0.898	0.856	0.849	0.832	0.763
	0.8	0.216	0.350	0.651	0.748	0.838	0.867	0.903	0.912	0.895	0.850
0.3	0.0	0.796	0.899	0.946	0.975	0.884	0.992	0.915	0.875	0.883	0.840
	0.4	0.421	0.753	0.897	0.913	0.924	0.913	0.889	0.897	0.858	0.810
	0.8	0.117	0.236	0.543	0.655	0.740	0.806	0.830	0.853	0.855	0.838
0.4	0.0	0.607	0.639	0.696	0.666	0.771	0.746	0.704	0.767	0.791	0.728
	0.4	0.366	0.579	0.651	0.640	0.746	0.745	0.795	0.780	0.753	0.773
	0.8	0.054	0.140	0.256	0.382	0.450	0.588	0.516	0.592	0.618	0.677

**Quadre 8.26.** EQM del predictor d' $X_{T+k}$  si  $(1-L)^d(1-\phi L)X_t = \varepsilon_t$ . Mètode B.

d	$\phi$	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
0.0	0.0	1.035	1.310	1.314	1.361	1.371	1.255	1.259	1.274	1.357	1.404
	0.4	0.808	1.074	1.111	1.112	1.082	1.132	1.180	1.180	1.115	1.220
	0.8	0.380	0.567	0.724	0.876	1.038	1.057	1.127	1.172	1.237	1.305
0.1	0.0	0.775	1.209	1.148	1.203	1.144	1.137	1.139	1.111	1.117	1.167
	0.4	0.619	0.946	0.884	0.922	0.918	0.901	0.937	0.921	0.917	0.932
	0.8	0.254	0.482	0.673	0.776	0.927	1.025	1.165	1.318	1.488	1.619
0.2	0.0	0.882	1.030	1.053	1.032	1.052	0.989	0.978	1.115	1.143	1.000
	0.4	0.528	0.743	0.766	0.837	0.864	0.912	0.889	0.897	0.881	0.824
	0.8	0.178	0.386	0.630	0.905	1.218	1.539	1.899	2.342	2.924	3.596
0.3	0.0	0.791	0.921	0.948	0.940	0.868	1.014	0.897	0.855	0.906	0.903
	0.4	0.491	0.719	0.854	0.929	0.939	0.990	0.966	0.991	0.957	0.884
	0.8	0.124	0.334	0.594	0.942	1.374	1.945	2.666	3.617	4.894	6.635
0.4	0.0	0.701	0.746	0.788	0.762	0.742	0.759	0.778	0.744	0.666	0.685
	0.4	0.373	0.626	0.800	0.884	0.953	1.014	1.095	1.105	1.118	1.099
	0.8	0.090	0.257	0.546	0.994	1.606	2.483	3.809	6.047	9.459	14.943



Quadre 8.27. EQM del predictor d' $X_{T+k}$  si  $(1-L)^d(1-\phi L)X_t = \varepsilon_t$ . Mètode C.

d	$\phi$	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
0.0	0.0	1.391	1.202	1.161	1.240	1.207	1.085	1.186	1.150	1.214	1.196
	0.4	1.096	1.301	1.237	1.138	1.183	1.226	1.259	1.271	1.150	1.311
	0.8	0.712	0.965	1.180	1.364	1.515	1.471	1.465	1.430	1.492	1.424
0.1	0.0	0.983	1.114	1.144	1.061	0.995	1.101	1.049	0.996	1.047	1.002
	0.4	0.907	1.216	1.154	1.202	1.203	1.248	1.253	1.200	1.167	1.163
	0.8	0.608	0.868	1.166	1.342	1.473	1.563	1.668	1.659	1.660	1.601
0.2	0.0	1.181	1.215	1.268	1.224	1.171	1.119	1.138	1.262	1.254	1.126
	0.4	0.890	1.245	1.349	1.416	1.371	1.466	1.406	1.371	1.326	1.154
	0.8	0.567	0.901	1.239	1.507	1.746	1.884	1.966	2.059	2.069	1.914
0.3	0.0	1.214	1.377	1.519	1.626	1.604	1.576	1.575	1.428	1.505	1.379
	0.4	0.891	1.291	1.629	1.821	1.831	1.867	1.765	1.688	1.661	1.485
	0.8	0.496	0.875	1.340	1.852	2.287	2.630	2.873	3.040	2.522	2.510
0.4	0.0	1.154	1.453	1.547	1.724	1.756	1.838	1.698	1.648	1.541	1.460
	0.4	0.751	1.101	1.416	1.629	1.711	1.854	1.862	1.703	1.641	1.524
	0.8	0.447	0.555	0.998	1.213	1.488	2.068	2.578	2.769	2.801	2.695

Quadre 8.28. EQM del predictor d' $X_{T+k}$  si  $(1-L)^d(1-\phi L)X_t = \varepsilon_t$ . Mètode D.

d	$\phi$	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
0.0	0.0	1.216	1.097	1.117	1.107	1.112	1.051	1.061	1.089	1.117	1.061
	0.4	0.970	1.170	1.159	1.071	1.098	1.110	1.160	1.167	1.060	1.174
	0.8	0.471	0.654	0.833	0.967	1.051	1.014	1.030	1.045	1.101	1.116
0.1	0.0	0.911	1.013	1.025	0.982	0.887	0.990	0.972	0.914	0.968	0.918
	0.4	0.751	1.016	0.928	1.002	0.968	1.004	1.001	0.951	0.934	1.005
	0.8	0.338	0.570	0.787	0.857	0.937	0.949	1.001	1.001	1.032	1.012
0.2	0.0	0.950	0.953	1.045	0.949	0.934	0.877	0.949	1.020	1.042	0.903
	0.4	0.621	0.877	0.911	0.950	0.947	1.020	0.975	0.951	0.941	0.825
	0.8	0.252	0.474	0.665	0.822	0.962	1.027	1.066	1.123	1.108	1.024
0.3	0.0	0.908	0.991	1.073	1.115	1.029	1.085	1.058	0.967	0.980	0.930
	0.4	0.591	0.858	1.003	1.050	1.058	1.092	0.998	1.003	0.971	0.887
	0.8	0.185	0.388	0.578	0.776	0.917	1.034	1.101	1.150	1.119	1.047
0.4	0.0	0.800	0.936	0.939	0.993	0.998	1.046	0.975	0.929	0.853	0.761
	0.4	0.464	0.746	0.945	0.967	1.006	1.021	1.029	0.929	0.904	0.795
	0.8	0.126	0.278	0.483	0.681	0.815	0.892	0.933	1.001	1.019	0.970

Quadre 8.29. EQM del predictor d' $X_{T+k}$  si  $(1-L)^d(1-\phi L)X_t = \varepsilon_t$ . Mètode E.

d	$\phi$	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8	k=9	k=10
0.0	0.0	1.046	0.958	0.951	0.979	0.990	0.931	0.959	0.970	0.999	0.995
	0.4	0.997	1.036	1.008	0.954	0.960	0.987	1.020	1.048	0.982	1.092
	0.8	0.694	0.722	0.782	0.871	0.944	0.912	0.937	0.942	0.985	0.994
0.1	0.0	0.857	0.944	0.985	0.955	0.885	0.948	0.939	0.939	0.978	0.926
	0.4	0.843	1.062	0.892	0.970	0.924	0.952	0.988	0.937	0.929	0.973
	0.8	0.465	0.675	0.761	0.780	0.851	0.862	0.910	0.939	0.980	0.957
0.2	0.0	1.073	1.073	1.059	1.027	1.068	0.938	1.023	1.086	1.138	0.995
	0.4	0.796	1.019	0.955	1.005	1.032	1.077	1.028	1.018	0.998	0.915
	0.8	0.379	0.659	0.781	0.898	1.006	1.040	1.084	1.094	1.074	1.020
0.3	0.0	1.035	1.168	1.230	1.268	1.149	1.290	1.189	1.137	1.148	1.091
	0.4	0.808	1.109	1.167	1.187	1.201	1.187	1.156	1.166	1.116	1.053
	0.8	0.269	0.567	0.706	0.852	0.963	1.047	1.080	1.109	1.112	1.090
0.4	0.0	0.990	1.314	1.254	1.213	1.219	1.324	1.266	1.214	1.107	1.020
	0.4	0.652	1.091	1.191	1.176	1.184	1.182	1.253	1.232	1.194	1.083
	0.8	0.216	0.476	0.639	0.815	0.911	0.964	1.002	1.109	1.145	1.087

De totes maneres, és notable el fet que amb un model ARMA(1,1) s'obtenen prediccions força bones per  $k$  no gaire gran ( $k < 5$ ) quan les dades s'han generat a partir del model ARFIMA(0, $d$ ,0). Cal destacar també que els problemes per  $k > 5$  són especialment greus quan  $d = 0.4$ .

- Usant un model AR(10), s'obtenen prediccions encara millors que al usar un model ARMA(1,1). De fet, tot i que presenten un EQM superior a les que s'obtenen al usar el model ARFIMA(1, $d$ ,0) correcte, les diferències són petites excepte quan  $d = 0.4$  i  $\phi = 0.8$  en que el model no és prou flexible per recollir una persistència tan elevada. De totes maneres, en aquest cas les prediccions que s'obtenen amb un AR(10) són molt millors que les obtingudes amb un ARMA(1,1).
- Finalment, quan s'utilitza un model ARFIMA(1, $d$ ,0) per calcular les prediccions però amb paràmetres estimats enlloc dels paràmetres correctes, deixa de ser cert que s'obtinguin sempre millors prediccions que a partir dels models ARMA(1,1) o AR( $p$ ). Si hom compara els resultats amb els que s'obtenen al usar un model AR( $p$ ), s'obté que en tots els casos l'EQM de les prediccions és millor usar un AR( $p$ ) que un ARFIMA(1, $d$ ,0), tot i ser aquest

últim model el correcte. En altres resultats no recollits als quadres 8.25 a 8.29, es pot comprovar com sols per  $k > 25$  (quan  $T=200$ ), s'obtenen prediccions clarament millors al usar un ARFIMA(1, $d$ ,0) estimat que al usar un AR(10) estimat.

- Al comparar els resultats obtinguts amb el model ARMA(1,1) amb els obtinguts amb el model ARFIMA(1, $d$ ,0) quan cal estimar els paràmetres d'aquest últim, les diferències entre ambdós mètodes es redueixen molt i sols quan  $d > 0.1$  i  $\phi = 0.8$  és clarament millor la utilització del model ARFIMA.
- Curiosament, la necessitat d'estimar per MV els paràmetres del model ARFIMA, tot i que aquest estigui ben especificat, afecten de manera especial a la qualitat de les prediccions a curt termini.

Aquests resultats són aparentment contradictoris amb els resultats obtinguts als quadres 8.22, 8.23 i 8.24 que mostren com l'error derivat d'aproximar un model AR( $\infty$ ) a partir d'un número finit de coeficients autoregressius pot ser molt important. Al nostre entendre, en l'aplicació pràctica d'un model AR( $p$ ) per la predicció de variables amb memòria llarga no es presenta aquest problema perquè els  $p$  coeficients autoregressius estimats no coincideixen amb els primers coeficients de la representació autoregressiva infinita.

En definitiva, sempre que no sigui necessari calcular prediccions a llarg termini d'una sèrie amb memòria llarga i no es disposi d'un nombre gaire gran d'observacions, una molt bona opció és la d'estimar un model AR( $p$ ) amb  $p$  gran, tot i que els coeficients no siguin tots ells significatius perquè el benefici d'especificar i estimar un model ARFIMA és poc important.

Però si l'objectiu és el càlcul de prediccions a llarg termini d'una sèrie amb memòria llarga, és millor usar un model ARFIMA i, en aquest cas, tot i que és millor usar l'algorisme proposat per Peiris i Perea, és més senzill i s'obtenen resultats gairebé iguals filtrant les dades amb l'operador diferència fraccional i calculant prediccions per les dades filtrades a partir d'un model ARMA .

## 8.5. Conseqüències per l'extracció de senyals

Sovint, l'evolució temporal de certes variables econòmiques presenta excessives oscil·lacions que fan difícil la seva anàlisi. Així, per exemple, en l'anàlisi econòmica conjuntural, és habitual usar estimacions de la sèrie desestacionalitzada o bé estimacions d'un senyal cicle-tendència de la variable.

En canvi, altres tipus d'anàlisi econòmica s'interessen pel comportament de certes variables econòmiques lliures de comportaments tendencials o de llarg termini. Un bon exemple n'és l'anàlisi del cicles econòmics que es basa en una estimació de l'evolució de la component cíclica de la sèrie.

Totes aquestes aplicacions es basen en la consideració que una sèrie temporal està formada, conceptualment, per diverses components no observables. Sota aquest supòsit, l'estimació d'una o diverses d'aquestes components és un cas particular del problema estadístic de l'extracció de senyals, motiu pel qual les tècniques d'extracció de senyals han generat un interès creixent en l'anàlisi econòmica.

El concepte d'extracció de senyals es pot definir com un conjunt de tècniques estadístiques dissenyades amb la finalitat de seleccionar una part de tota la informació que conté un conjunt de dades de tall temporal. La traducció econòmica d'aquestes tècniques és la consideració que tota sèrie temporal conté diferents tipus d'informació i l'objectiu és separar aquella informació rellevant continguda en cada una de les sèries considerades.

Així, donada una successió temporal d'observacions  $X_t$ , es considera que aquesta està formada per un conjunt de  $n$  components no observables  $X_t^1, X_t^2, \dots, X_t^n$  de manera que:

$$X_t = X_t^1 + X_t^2 + \dots + X_t^n \quad (8.83)$$

Si el fet d'agregar les diferents components de manera additiva no sembla raonable es pot optar per una representació multiplicativa:

$$X_t = X_t^1 X_t^2 \dots X_t^n \quad (8.84)$$

i, aleshores, prenent logaritmes de les dades es pot recuperar la notació additiva.

El problema de l'extracció de senyals es pot plantejar de la següent manera. Un cop escollides *a priori* un conjunt d'aquestes components (una o vàries):

$$X_t^* = X_t^{i_1} + X_t^{i_2} + \dots + X_t^{i_k} \quad i_1, \dots, i_k \in \{1, \dots, n\} \quad (8.85)$$

es vol estimar aquesta nova component  $X_t^*$  a partir de les dades observades  $X_t$ .

Normalment, però, aquesta formulació tan general es redueix a considerar, en el cas de sèries econòmiques, només quatre components prou conegudes: tendència, cicle, estacionalitat i soroll.

Per tant:

$$X_t = T_t + C_t + S_t + U_t \quad (8.86)$$

En general, si es suposa que la sèrie observada ( $X_t$ ) és la combinació d'un senyal interès ( $Y_t$ ) i d'un terme residual ( $N_t$ ):

$$X_t = Y_t + N_t \quad (8.87)$$

la recerca d'una estimació del senyal es basa habitualment en una combinació lineal dels valors de la sèrie observada:

$$\tilde{Y}_t = \sum_{j=-\infty}^{\infty} h_j X_{t-j} \quad (8.88)$$

Sota el supòsit que ambdues components són estacionàries, a més de certes condicions de regularitat, es pot comprovar que el filtre òptim en el sentit que minimitza l'error quadràtic mig de l'estimació compleix la següent condició:

$$H(\omega) = \frac{f_{Y^*X}(\omega)}{f_X(\omega)} \quad (8.89)$$

on  $H(\omega)$  és la funció de transferència del filtre<sup>41</sup>,  $f_X(\omega)$  és la densitat espectral del procés  $X_t$  i  $f_{YX}(\omega)$  és la densitat espectral creuada dels processos  $Y_t$  i  $X_t$ . Si a més es suposa (com és habitual) que el senyal  $Y_t$  i el soroll  $N_t$  són independents, la condició (8.88) és equivalent a:

$$H(\omega) = \frac{f_Y(\omega)}{f_X(\omega)} \tag{8.90}$$

Aquest resultat tan important es coneix a la literatura estadística sota el nom de filtre de Wiener-Kolmogorov ja que foren aquests dos autors els que en dos treballs paral·lels ho comproven (Wiener, 1949 i Kolmogorov, 1940)

A l'anàlisi econòmica s'ha desenvolupat àmpliament aquesta metodologia a partir de la hipòtesi addicional de que ambdues components es poden representar mitjançant models ARMA<sup>42</sup>. En aquest cas, si es compleix:

$$\phi_Y(L)Y_t = \theta_Y(L)a_t \tag{8.91}$$

i:

$$\phi_N(L)N_t = \theta_N(L)b_t \tag{8.92}$$

llavors la sèrie observada també es pot representar mitjançant un model ARMA<sup>43</sup>.

$$\phi(L)X_t = \theta(L)\varepsilon_t \tag{8.93}$$

En aquest cas, es pot comprovar com, en notació temporal, el filtre òptim és:

$$\tilde{Y}_t = H(L, F)X_t = \frac{\sigma_a^2 \theta_Y(L)\theta_Y(F)\phi_N(L)\phi_N(F)}{\sigma_\varepsilon^2 \theta(L)\theta(F)} X_t \tag{8.94}$$

<sup>41</sup> La funció de transferència d'un filtre  $\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q$  es defineix com:

$$H_\theta(\omega) = \theta(e^{i\omega}) = 1 + \theta_1 e^{i\omega} + \theta_2 e^{2i\omega} + \dots + \theta_q e^{qi\omega}$$

Per cada freqüència  $\omega$ , el valor  $|H_\theta(\omega)|^2$  o funció de potència del filtre és el quocient de les densitats espectrals de la sèrie filtrada i la sèrie original; per tant, aquesta funció recull l'efecte que té sobre cada freqüència l'aplicació del filtre. Sovint, s'utilitza també la funció de guany d'un filtre, que es defineix en aquest cas com  $|H_\theta(\omega)|$ .

<sup>42</sup> Vegi's per exemple Maravall (1996).

<sup>43</sup> Per una demostració d'aquesta propietat es pot consultar Fuller (1976).

on  $F$  és un operador d'avançaments, és a dir  $F=L^{-1}$ .

A més, en els treballs de Cleveland i Tiao (1976) i Bell (1984) s'estén aquest resultat al cas de sèries que contenen arrels unitàries permetent l'aplicació del filtre de Wiener-Kolmogorov no sols a models ARMA sinó també a models ARIMA.

Aquesta metodologia es pot generalitzar a la hipòtesi de l'existència de més de dos components. Així, si es suposa que una sèrie econòmica està formada per tres components, tendència-cicle ( $T_t$ ), estacionalitat ( $S_t$ ) i soroll ( $U_t$ ) és habitual suposar que cada una d'elles es pot representar mitjançant un model ARIMA:

$$\phi_T(L)T_t = \theta_T(L)a_t \quad (8.95)$$

$$\phi_S(L)S_t = \theta_S(L)b_t \quad (8.96)$$

$$\phi_U(L)U_t = \theta_U(L)c_t \quad (8.97)$$

A més, en aquestes condicions, es pot comprovar com els polinomis autoregressius i la mitjana mòbil del model de la sèrie  $X_t$  i els polinomis autorregressius i la mitjana mòbil de les components compleixen algunes relacions. Concretament es compleix que<sup>44</sup>:

$$\phi_X(L) = \phi_T(L)\phi_S(L)\phi_U(L) \quad (8.98)$$

i:

$$\begin{aligned} \theta_X(L)\sigma_\varepsilon^2 = & [\phi_S(L) + \phi_U(L)] \theta_T(L)\sigma_a^2 + [\phi_T(L) + \phi_U(L)] \theta_S(L)\sigma_b^2 + \\ & + [\phi_T(L) + \phi_S(L)] \theta_U(L)\sigma_c^2 \end{aligned} \quad (8.99)$$

La idea fonamental d'aquest enfocament consisteix en obtenir primer una estimació del model ARIMA de la sèrie agregada  $X_t$ , i, a partir d'aquest model, esbrinar els models ARIMA de cada una de les components.

Concretament els estimadors amb error quadràtic mitjà mínim per cada un dels components són, seguint la notació anterior de cada un dels models ARIMA:

---

<sup>44</sup> Vegi's Maravall (1996).

$$\tilde{T}_t = \frac{\theta_T(L) \theta_T(F) \phi_S(L) \phi_S(F) \phi_U(L) \phi_U(F) \sigma_a^2}{\theta_X(L) \theta_X(F) \sigma_e^2} X_t \quad (8.100)$$

$$\tilde{S}_t = \frac{\theta_S(L) \theta_S(F) \phi_T(L) \phi_T(F) \phi_U(L) \phi_U(F) \sigma_b^2}{\theta_X(L) \theta_X(F) \sigma_e^2} X_t \quad (8.101)$$

$$\tilde{U}_t = \frac{\theta_U(L) \theta_U(F) \phi_S(L) \phi_S(F) \phi_T(L) \phi_T(F) \sigma_c^2}{\theta_X(L) \theta_X(F) \sigma_e^2} X_t \quad (8.102)$$

Per tant, si es coneixen els models ARIMA de cada una de les components no observables i de la sèrie agregada, el problema d'estimar aquestes components és relativament senzill.

Una de les conseqüències fonamentals d'aquest desenvolupament formal és que el filtre òptim per estimar una component concreta depèn de les característiques dels models de la sèrie observada i de la component a estimar. En particular, si les dades contenen memòria llarga i això no es té en compte, es poden produir errors greus en l'estimació de les components.

En aquest sentit, i valgui com exemple, s'ha comentat que moltes de les variables econòmiques d'alta freqüència es poden aproximar bastant bé usant el model línies aèries, però que amb les eines estadístiques habituals no és possible saber si aquest model recull les característiques de la variable pel que fa a la seva memòria o bé si s'acaba seleccionant aquest model degut a un problema de sobrediferenciació d'una sèrie integrada fraccionalment.

Cal preguntar-se doncs quines conseqüències pot tenir l'estimació d'un senyal no observable quan les dades contenen memòria llarga però s'ha especificat de manera errònia un model ARIMA.

Per valorar aquestes conseqüències és necessari conèixer quin és l'estimador òptim del senyal quan les dades observades segueixen per exemple un model ARFIMA. Pel que sabem aquest és un aspecte de l'extracció de senyals sobre el que no es disposa fins al moment de cap resultat.



Suposi's que la sèrie observada segueix un model ARFIMA:

$$\phi(L)(1-L)^d X_t = \theta(L)\varepsilon_t \quad (8.103)$$

i que aquesta sèrie s'obté com agregació de les components tendència-cicle, component estacional i soroll:  $X_t = T_t + S_t + U_t$ . Sembla raonable, de manera similar al cas dels models ARMA, suposar que cada una de les components també es pot representar com un model ARFIMA.

En aquest context, l'objectiu és disposar d'algun resultat similar a (8.94) que permeti trobar a partir dels models de la sèrie i de les components quin és el filtre òptim per l'extracció de senyals en l'entorn de models ARFIMA.

Però per a que sigui possible aquesta hipòtesi cal garantir en primer lloc que la suma de models ARFIMA és també un model ARFIMA. El Teorema següent demostra aquesta propietat i és una primera passa cara a l'objectiu plantejat.

**Teorema 8.1.** Si una certa sèrie temporal  $X_t$  és combinació de dos components independents  $X_t = Y_t + N_t$  on  $Y_t \sim \text{ARFIMA}(p_Y, d_Y, q_Y)$  i  $N_t \sim \text{ARFIMA}(p_N, d_N, q_N)$  estacionaris amb  $d_Y \geq d_N$  llavors la composició de les dos components també es pot representar mitjançant un model ARFIMA( $p, d, q$ ).

#### **Demostració:**

Les dos components es poden representar com:

$$\phi_Y(L)(1-L)^{d_Y} Y_t = \theta_Y(L)a_t$$

$$\phi_N(L)(1-L)^{d_N} N_t = \theta_N(L)b_t$$

on  $a_t$  i  $b_t$  són processos soroll blanc. De manera que cada una de les components és estacionària amb densitats espectrals:

$$f_Y(\omega) = \frac{\sigma_a^2}{2\pi[2\sin(\omega/2)]^{2d_Y}} \frac{|\theta_Y(\omega)|^2}{|\phi_Y(\omega)|^2}$$

i:

$$f_N(\omega) = \frac{\sigma_b^2}{2\pi[2\sin(\omega/2)]^{2d_N}} \frac{|\theta_N(\omega)|^2}{|\phi_N(\omega)|^2}$$

Com a més són independents, la sèrie observada té per espectre la suma d'aquestes dos densitats espectrals. Així, sota el supòsit que  $d_Y \geq d_N$  aquesta densitat es pot escriure com:

$$\begin{aligned} f(\omega) &= \frac{\sigma_a^2}{2\pi[2\sin(\omega/2)]^{2d_Y}} \frac{|\theta_Y(\omega)|^2}{|\phi_Y(\omega)|^2} + \frac{\sigma_b^2}{2\pi[2\sin(\omega/2)]^{2d_N}} \frac{|\theta_N(\omega)|^2}{|\phi_N(\omega)|^2} = \\ &= \frac{\sigma_a^2 |\theta_Y(\omega)|^2 |\phi_N(\omega)|^2 + \sigma_b^2 |\theta_N(\omega)|^2 |\phi_Y(\omega)|^2 [2\sin(\omega/2)]^{2(d_Y - d_N)}}{2\pi[2\sin(\omega/2)]^{2d_Y} |\phi_Y(\omega)|^2 |\phi_N(\omega)|^2} = \\ &= \frac{g(\omega)}{2\pi[2\sin(\omega/2)]^{2d_Y} |\phi_Y(\omega)|^2 |\phi_N(\omega)|^2} \end{aligned}$$

on la funció  $g(\omega)$  compleix la següent condició<sup>45</sup>:

$$\int_{-\pi}^{\pi} \log\{g(\omega)\} d\omega > -\infty \tag{8.104}$$

Aquesta condició permet usar el següent Lema (Doob, 1953, pàg. 160):

**Lema 8.1.** Si una funció  $g(\omega)$  compleix (8.103), existeix una única successió de coeficients  $\theta_0, \theta_1, \dots$  amb  $\theta_0$  positiu i

$$\sum_{j=0}^{\infty} |\theta_j|^2 < \infty$$

tal que la funció  $\theta(z) = \sum_{jj=0}^{\infty} \theta_j z^j$  no té zeros dintre del cercle unitat, i permet factoritzar la

funció  $g(\omega)$  com:

$$g(\omega) = |\theta(\omega)|^2$$

<sup>45</sup> Vegi's el Teorema 8.3 de l'Annex 8.2 per a una demostració que la condició (8.102) és certa.

Normalitzant per a que  $\theta(0)=1$ , es poden trobar una constant que notarem com  $\sigma_\varepsilon^2$  i un polinomi:

$$\theta(\omega) = 1 + \theta_1 e^{-i\omega} + \dots + \theta_m e^{-im\omega}$$

de manera que la funció  $g(\omega)$  compleix:

$$g(\omega) = \sigma_\varepsilon^2 |\theta(\omega)|^2$$

Per tant, la sèrie observada  $X_t$ , es pot representar com:

$$\phi(L)(1-L)^d X_t = \theta(L)\varepsilon_t$$

on  $d=d_Y$  i  $\phi(L) = \phi_Y(L) \phi_N(L)$ .

A continuació, un cop s'ha garantit que la suma de models ARFIMA és també un model ARFIMA, cal trobar l'expressió del filtre òptim en aquest context. La via més senzilla per trobar aquest filtre consisteix a aprofitar el resultat de Wiener i Kolmogorov.

Si es suposa que les tres components  $X_t$ ,  $Y_t$  i  $N_t$  són estacionàries, n' existeixen representacions de Wold amb successió de coeficients convergent a zero:

$$X_t = \psi(L)\varepsilon_t, \quad Y_t = \psi_Y(L)a_t, \quad N_t = \psi_N(L)b_t$$

Si a més, es suposa que el resultat de Wiener i Kolmogorov és vàlid, s'obté que el filtre òptim per l'estimació de  $Y_t$  compleix:

$$H(\omega) = \frac{f_Y(\omega)}{f_X(\omega)} \quad (8.105)$$

i, per tant:

$$\begin{aligned} \tilde{Y}_t &= H(L, F)X_t = \frac{\sigma_a^2 \psi_Y(L)\psi_Y(F)}{\sigma_\varepsilon^2 \psi(L)\psi(F)} = \\ &= \frac{\sigma_a^2 \theta_Y(L)\theta_Y(F)\phi_N(L)\phi_N(F)}{\sigma_\varepsilon^2 \theta(L)\theta(F)} X_t \end{aligned} \quad (8.106)$$

Però per a que aquest raonament sigui correcte, cal demostrar que en el context de models ARFIMA el resultat de Wiener-Kolmogorov és cert. El problema és que en la demostració

d'aquests resultats sovint s'utilitzen supòsits que els models ARFIMA no compleixen. Per exemple, les demostracions més generals d'aquest resultat basades en el domini de les freqüències com poden ser les presentades a Fuller (1976) i Priestley (1981) es basen en la hipòtesi que les densitats espectrals de la sèrie observada i de la component a estimar són contínues, hipòtesi que no compleixen els models ARFIMA.

En tot cas, al nostre entendre, la demostració de Maravall (1987) pel cas estacionari (pàg. 54-57) és vàlida ja que es basa sols en l'existència de la representació de Wold, condició que compleixen els models ARFIMA estacionaris.

Pel cas no estacionari, la cosa és més delicada ja que la demostració de Cleveland i Tiao (1976) es basa directament en el supòsit de que es tracta de models ARIMA i Bell (1984) parteix del supòsit que les autocovariàncies són absolutament sumables, hipòtesi que tampoc es compleix sota memòria llarga.

Queda pendent la demostració del cas no estacionari però creiem que l'argument usat a Maravall (1987) per incloure la possibilitat d'arrels unitàries en la representació autoregressiva és suficient per estendre el resultat a models ARFIMA no estacionaris.

En tot cas, calen alguns comentaris addicionals:

1) En el filtre que s'obté no apareix cap terme de la forma  $(1-L)^\delta$  degut al supòsit  $d_Y \geq d_N$ . Si la component que es vol estimar fos  $N_b$ , és a dir, la que presenta una menor memòria, llavors es pot comprovar com el filtre òptim és:

$$\tilde{N}_t = H(L, F)X_t = \frac{\sigma_b^2}{\sigma_\varepsilon^2} \frac{\psi_N(L)\psi_N(F)}{\psi(L)\psi(F)} X_t \quad (8.107)$$

2) Tal com passa si s'usen models ARIMA, la component de major memòria té el mateix ordre d'integrabilitat que la sèrie observada<sup>46</sup>.

<sup>46</sup> De fet, l'ordre d'integració fraccional també és una propietat dominant en el sentit de Granger (1990).

3) Si en el cas d'usar models ARIMA hi ha problemes amb la convergència dels filtres, aquests també es presenten al usar models ARFIMA. En tot cas, una possible solució consisteix en la utilització de l'algorisme de Wilson. Aquest algorisme permet descomposar un filtre bidireccional  $H(L,F)$  del tipus:

$$H(L, F) = \frac{C(L, F)}{\theta(L)\theta(F)}$$

com la suma de dos filtres unidireccionals:

$$H(L, F) = \frac{D(L)}{\theta(L)} + \frac{D(F)}{\theta(F)}$$

De totes maneres, per valorar les possibles conseqüències de no considerar la possibilitat de memòria llarga a l'hora d'estimar components no observables cal fixar-se en els efectes pràctics més que en els efectes a nivell teòric. De fet, podria succeir que tot i l'error en l'especificació del model per la variable observada, la component final estimada sigui molt similar a la que s'obtidria especificant el model correcte.

Un dels models ARIMA més habituals en la modelització de sèries econòmiques és el model de les línies aèries i, atès que a l'apartat 8.2 s'ha comprovat que és molt probable especificar aquest model quan les dades s'han generat segons un model ARFIMA, es pot usar aquest cas com exemple per valorar les conseqüències dels errors d'especificació del tipus de memòria en l'estimació de components no observables.

Per simplicitat, s'analitza a continuació les conseqüències que té per l'extracció de senyals la sobrediferenciació a la freqüència zero. Així, consideri's una sèrie generada segons el procés:

$$(1 - L)^d X_t = \varepsilon_t \quad (8.108)$$

i suposi's que es pretèn obtenir una estimació de la tendència de la sèrie per treure el màxim soroll possible, és a dir, es pretèn descomposar  $X_t$  com:

$$X_t = Y_t + N_t \quad (8.109)$$

on  $N_t$  sigui soroll blanc.

A partir dels resultats ja comentats, per trobar el filtre òptim que permeti l'estimació d' $Y_t$ , a partir de les observacions d' $X_t$ , sols cal conèixer quina és la densitat espectral d' $Y_t$ .

En tot cas, pels resultats de l'apartat 8.4.3, és molt probable que a partir dels estadístics habituals es decideixi la utilització de l'operador diferència regular, de manera que, a partir dels resultats obtinguts, és molt probable que el model finalment escollit sigui un model IMA(1,1):

$$(1 - L)X_t = (1 - \theta L)\varepsilon_t \quad (8.109)$$

A partir d'aquest model, la densitat espectral d' $X_t$  és:

$$f_X(\omega) = \frac{\sigma_\varepsilon^2(1 + \theta^2 - 2\theta \cos(\omega))}{4\pi(1 - \cos(\omega))} \quad (8.110)$$

Com  $Y_t$  i  $N_t$  han de ser ortogonals, la densitat espectral d' $X_t$  ha de ser la suma de les densitats espectrals d' $Y_t$  i  $N_t$ . Per simplificar la descomposició és millor fer un canvi de variable de manera que  $x = \cos(\omega)$ . Com a més, la densitat espectral d' $N_t$  és constant, s'obté:

$$f_X(x) = \frac{\sigma_\varepsilon^2(1 + \theta^2 - 2\theta x)}{4\pi(1 - x)} = f_Y(x) + C \quad (8.111)$$

De totes maneres, existeixen infinites descomposicions que compleixen la condició (8.111). Per trobar-ne una solució única és habitual imposar el que es coneix com condició canònica, és a dir, fixar com a constant  $C$  el valor màxim possible que compleixi aquesta condició. D'aquesta manera, s'assigna a  $N_t$  el màxim de variància possible per a que la tendència sigui el més estable possible.

Com per definició, una densitat espectral no pot prendre valors negatius, el valor de la constant  $C$  s'obté calculant el valor mínim de l'espectre d' $X_t$ , a l'interval  $\omega \in [0, \pi]$ . En aquest cas concret, com  $f_X(x)$  és una funció decreixent, el mínim s'assoleix en  $x = -1$  (és a dir,  $\omega = \pi$ ) i és igual a:

$$C = f_X(-1) = \frac{\sigma_\varepsilon^2(1 - \theta)^2}{8\pi} \quad (8.112)$$

Per tant, la densitat espectral de la tendència a estimar és:

$$f_Y(x) = f_X(x) - C = \frac{\sigma_\varepsilon^2(1+\theta^2-2\theta x)}{4\pi(1-x)} - \frac{\sigma_\varepsilon^2(1-\theta)^2}{8\pi} = \frac{\sigma_\varepsilon^2(1-\theta)^2(1+x)}{8\pi(1-x)} \quad (8.113)$$

de manera que el filtre de Wiener-Kolmogorov per l'estimació de la tendència a utilitzar si s'ha escollit per  $X_t$ , un model IMA(1,1) té per funció de guany:

$$G^*(\omega) = \frac{f_Y(\omega)}{f_X(\omega)} = \frac{(1-\theta)^2(1+\cos(\omega))}{2(1+\theta^2-2\theta\cos(\omega))} \quad (8.114)$$

Però aquest fora el filtre òptim si el procés generador de dades fos efectivament un IMA(1,1) però recordi's que s'ha suposat que el model correcte és un ARFIMA(0,d,0). Per tant, té especial interès comparar la funció de potència del filtre que s'aplica degut a l'error en l'especificació del model  $G^*(\omega)$ , amb la funció de potència del filtre òptim per valorar si l'error comès és gaire gran.

Per trobar el filtre òptim, cal tenir en compte que la densitat espectral d' $X_t$ , un cop s'utilitza el model correcte és:

$$f_X(x) = \frac{\sigma_\varepsilon^2}{2\pi[2(1-x)]^d} \quad (8.115)$$

Com es tracta d'una funció decreixent a l'interval d'interès, el mínim s'assoleix en  $x=-1$  (és a dir,  $\omega=\pi$ ) i és igual a:

$$C = f_X(-1) = \frac{\sigma_\varepsilon^2}{2\pi 4^d} \quad (8.116)$$

Per tant, el filtre òptim per l'estimació de la tendència té per funció de guany la següent:

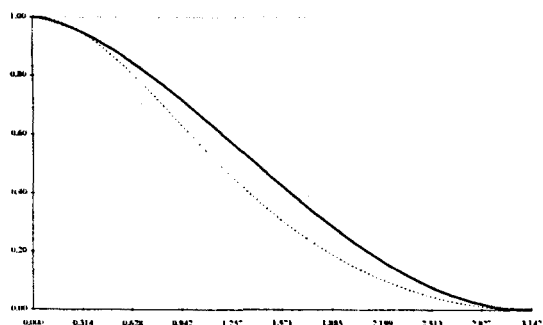
$$G(\omega) = \frac{f_Y(\omega)}{f_X(\omega)} = \frac{\frac{\sigma_\varepsilon^2}{2\pi[2(1-\cos(\omega))]^d} - \frac{\sigma_\varepsilon^2}{2\pi 4^d}}{\frac{\sigma_\varepsilon^2}{2\pi[2(1-\cos(\omega))]^d}} = 1 - \left[ \frac{1-\cos(\omega)}{2} \right]^d \quad (8.117)$$

Per valorar si les diferències entre la funció  $G(\omega)$  i la funció  $G^*(\omega)$  són gaire importants, s'han comparat aquestes funcions quan  $\theta=1-d$  per a que siguin comparables, ja que a

l'apartat 8.4.3 de la tesi s'ha comprovat que en cas de sobrediferenciació, el valor esperat del paràmetre mitjana mòbil estimat és  $\theta=1-d$ .

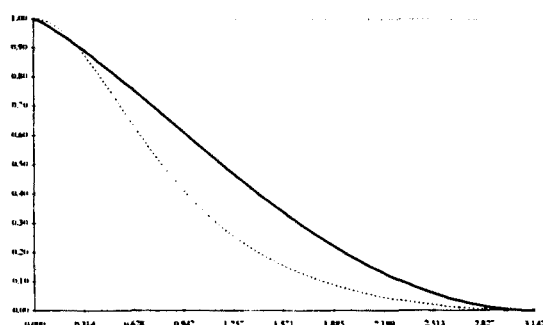
**Gràfic 8.15.** Comparació de les funcions de potència correcte i incorrecte suposant que

$$(1-L)^d X_t = \varepsilon_t \text{ amb } d=0.8.$$



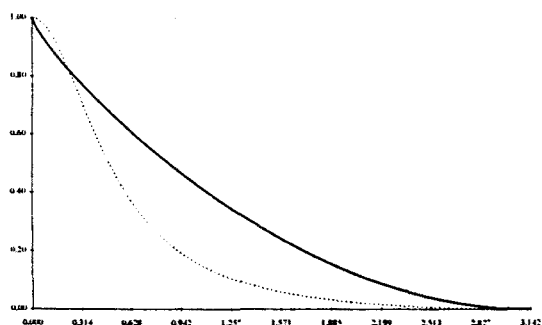
**Gràfic 8.16.** Comparació de les funcions de potència correcte i incorrecte suposant que

$$(1-L)^d X_t = \varepsilon_t \text{ amb } d=0.6.$$



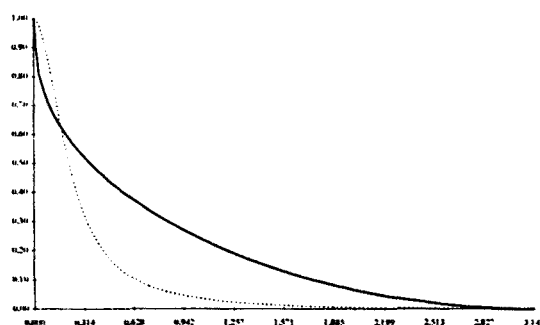
**Gràfic 8.17.** Comparació de les funcions de potència correcte i incorrecte suposant que

$$(1-L)^d X_t = \varepsilon_t \text{ amb } d=0.4.$$



**Gràfic 8.18.** Comparació de les funcions de potència correcte i incorrecte suposant que

$$(1-L)^d X_t = \varepsilon_t \text{ amb } d=0.2.$$



Als gràfics 8.15, a 8.18 es comparen ambdues funcions per diferents valors de  $d$ . Observi's com les diferències són molt importants de manera que errors en la identificació del tipus de memòria poden portar a errors molt importants en l'estimació de senyals no observables a partir de la teoria habitual del filtre de Wiener-Kolmogorov<sup>47</sup>.

<sup>47</sup> En els gràfics apareixen en línia contínua la funció de guany del filtre correcte,  $G(\omega)$ , per l'estimació de la tendència si s'ha especificat el model  $(1-L)^d X_t = \varepsilon_t$  mentre que en línia discontinua hi ha la funció de guany del filtre utilitzat si s'ha especificat el model  $(1-L)X_t = (1-\theta)\varepsilon_t$ ,  $G^*(\omega)$ .



De fet, els errors són més grans a mesura que el paràmetre  $d$  és inferior ja que l'error en la determinació de la necessitat de diferenciar és més greu. En tot cas, fins i tot per  $d=0.6$ , situació en que és probable que els tests d'arrels unitàries analitzats prèviament informin de la necessitat de diferenciar les dades, les diferències entre el filtre usat i el filtre òptim són molt grans.

Per tant, queda demostrada la necessitat de considerar la possibilitat de que les dades econòmiques presentin diferents tipus de memòria i d'especificar correctament el model usat si es pretèn usar aquestes dades per estimar algun tipus de component no obseables, ja sigui la tendència de la variable, com en l'exemple anterior, o un altre tipus de senyal com la sèrie desestacionalitzada o la component cíclica.