

DOCUMENTS DE TREBALL
DE LA DIVISIÓ DE CIÈNCIES JURÍDIQUES
ECONÒMIQUES I SOCIALS

Col·lecció d'Economia

**UNA EXTENSIÓN DE LA REGRESIÓN PROPUESTA POR
GEWEKE Y PORTER-HUDAK PARA LA ESTIMACIÓN DEL
ORDEN DE DIFERENCIACIÓN EN MODELOS ARFIMA**

Ernest Pons Fanals

Jordi Suriñach i Caralt

Adreça correspondència:

Departament d'Econometria, Estadística i Economia Espanyola
Grup d'Anàlisi Quantitativa Regional
Facultat d'Econòmiques - Universitat de Barcelona
Avda. Diagonal, 690 - 08034 Barcelona
Email: epons@eco.ub.es

Recepció document: Setembre 1999

ABSTRACT

Over the last years there has been considerable interest in the application of long memory time series models in economics using ARFIMA models. Nowadays, the most popular estimator of the difference parameter in economic applications is that proposed by Geweke and Porter-Hudak (GPH) although has been shown that, in particular circumstances, this estimator can be badly biased. So we propose a biased-corrected extension of that estimator using the exponential model proposed by Bloomfield.

The performance of both estimators are investigated for moderate-size samples and it is concluded that the extension proposed has smaller mean squared error than the GPH estimator.

KEY WORDS: Fractional integration, long memory, ARFIMA model.

JEL classification: C22, E3

RESUM

Recentment, ha augmentat molt l'interès per l'aplicació dels models de memòria llarga a variables econòmiques, sobretot els models ARFIMA. Sens dubte, el mètode més usat per l'estimació d'aquests models a l'àmbit de l'anàlisi econòmica és el proposat per Geweke i Porter-Hudak (GPH) tot i que en treballs recents s'ha demostrat que, en certs casos, aquest estimador presenta un biaix molt important. Per aquest motiu, es proposa una extensió d'aquest estimador a partir del model exponencial proposat per Bloomfield, i que permet corregir aquest biaix.

A continuació, s'analitza i compara el comportament d'ambdós estimadors en mostres no gaire grans i es comprova com l'estimador proposat presenta un error quadràtic mitjà menor que l'estimador GPH.

PARAULES CLAU: Integració fraccional, memòria llarga, models ARFIMA.

CODIS JEL: C22, E3

1. Introducción.

Durante los últimos años se ha generado un gran interés por la aplicación de modelos de series temporales con memoria larga en gran número de disciplinas, muy especialmente en economía, debido a la observación de que gran número de magnitudes económicas presentan grados de persistencia o memoria poco coherentes con las características de los modelos ARMA o ARIMA.

Buena parte de este reciente interés se debe al desarrollo de los modelos autorregresivos, media móvil y fraccionalmente integrados (ARFIMA) que permiten, de manera relativamente simple, la modelización de situaciones intermedias entre los modelos ARMA (estacionarios y con poca persistencia) y los modelos ARIMA (con raíces unitarias y, por tanto, con una persistencia infinita de los posibles shocks).

Concretamente, se dice que un proceso estocástico X_t sigue un proceso autorregresivo, media móvil y fraccionalmente integrado (ARFIMA)¹ si:

$$\mathbf{f}(L)(1-L)^d X_t = \mathbf{q}(L)\mathbf{e}_t \quad (1)$$

donde los polinomios de (1) están definidos a partir de:

$$\mathbf{f}(L) = 1 - \mathbf{f}_1 L - \dots - \mathbf{f}_p L^p \quad (2)$$

$$\mathbf{q}(L) = 1 - \mathbf{q}_1 L - \dots - \mathbf{q}_q L^q \quad (3)$$

$$(1-L)^d = 1 - dL - \frac{d}{2!}(1-d)L^2 \dots \quad (4)$$

y \mathbf{e}_t es un proceso ruido blanco². Si los polinomios (2) y (3) que describen el comportamiento a corto plazo tienen todas sus raíces fuera del círculo unidad y el parámetro d se encuentra en el intervalo $(-1/2, 1/2)$ el proceso es estacionario e

¹ Véase Granger y Joyeux (1980) y Hosking (1981) para una introducción a la integración fraccional y a las propiedades de los modelos ARFIMA.

invertible. En cambio, si $d \neq 1/2$ entonces se trata de un proceso no estacionario que siempre es posible convertir en estacionario utilizando el operador $(1-L)$.

Entre todos los métodos propuestos en la literatura para la estimación del parámetro de integración fraccional d , uno de los más populares y sin duda, el más utilizado en las aplicaciones prácticas en el ámbito económico³, es el propuesto por Geweke y Porter-Hudak (1983) (GPH). No obstante, en Agiakloglou *et al.* (1993) se muestra como, en ciertas circunstancias, dicho estimador puede estar seriamente sesgado.

A continuación se propone una extensión del estimador GPH basada en el modelo exponencial de Bloomfield (1973) que permite eliminar dicho sesgo y mejorar las propiedades de dicho estimador. La eliminación de este sesgo es especialmente importante cuando se pretende realizar algún tipo de inferencia acerca del parámetro de integración.

El trabajo presentado está organizado de la siguiente manera. En el apartado 2 definimos una extensión del estimador GPH mediante la estimación de un modelo de regresión ampliado en el dominio de las frecuencias. En el apartado 3, se presentan los resultados de un experimento de Monte-Carlo sobre las propiedades del estimador propuesto en muestras finitas. A continuación, se ilustra la aplicación práctica del método propuesto con el Índice de Precios al Consumo (IPC) español en el apartado 4. Finalmente, se concluye el trabajo con un resumen de los principales resultados obtenidos.

² También se dice, en este caso, que X_t es fraccionalmente integrado de orden d , o $X_t \sim \text{FI}(d)$.

³ Véase, por ejemplo, Diebold y Rudebusch (1989) o Porter-Hudak (1990).

2. Una extensión del estimador GPH

La aproximación propuesta por Geweke y Porter-Hudak (1983) se basa en que la densidad espectral del modelo (1) cumple⁴:

$$f_X(\mathbf{w}) = \left\{4\sin^2(\mathbf{w}/2)\right\}^{-d} f_Y(\mathbf{w}) \quad (5)$$

donde $f_Y(\mathbf{w})$ es la densidad espectral de $Y_t=(I-L)^d X_t$. Al tomar logaritmos en la expresión anterior y reordenando de manera adecuada se obtiene:

$$\ln f_X(\mathbf{w}) = \ln f_Y(0) - d \ln\left\{4\sin^2(\mathbf{w}/2)\right\} + \ln \frac{f_Y(\mathbf{w})}{f_Y(0)} \quad (6)$$

Si se dispone de observaciones X_t , $t=1,\dots,T$, puede utilizarse el periodograma para evaluar la expresión (6) en las frecuencias armónicas $\mathbf{w}_j=2\mathbf{j}/T$, obteniéndose la siguiente igualdad:

$$\ln I(\mathbf{w}_j) = \ln f_Y(0) - d \ln\left\{4\sin^2(\mathbf{w}_j/2)\right\} + \ln \frac{f_Y(\mathbf{w}_j)}{f_Y(0)} + \ln \frac{I(\mathbf{w}_j)}{f_X(\mathbf{w}_j)} \quad (7)$$

Dicha igualdad puede aprovecharse para obtener una estimación del parámetro d si se tiene en cuenta que, en frecuencias suficientemente próximas a la frecuencia cero, el término $\ln\left\{f_Y(\mathbf{w}_j)/f_Y(0)\right\}$ es muy próximo a cero y puede ser despreciado en comparación con los otros, de manera que, los autores proponen obtener una estimación de d a través del siguiente modelo de regresión:

$$\log I(\mathbf{w}_j) = \mathbf{a} + \mathbf{b}R_j + e_j \quad (8)$$

⁴ A pesar de que a lo largo del trabajo se propone el método propuesto para la estimación del orden de diferenciación fraccional en modelos ARFIMA, las condiciones impuestas són muy flexibles, lo que permite la utilización de dicho estimador en un entorno semiparamétrico.

donde $j=m_1, m_1+1, \dots, m$, el regresando $\ln I(\mathbf{w}_j)$ es el logaritmo del periodograma en la frecuencia $\mathbf{w}_j=2\mathbf{j}/T$ con T el numero de observaciones, la constante \mathbf{a} es el logaritmo del espectro en cero de $(1-L)^d X_t = u_t$, el regresor R_j está definido por $R_j = \ln\{4\sin^2(\mathbf{w}_j/2)\}$ y el error del modelo es $e_j = \ln\{I(\mathbf{w}_j)/f_X(\mathbf{w}_j)\}$. Una vez se dispone de una estimación del parámetro d , el cálculo de $Y_t = (1-L)^d X_t$ permite obtener estimaciones de los parámetros autorregresivos y media móvil.

Naturalmente, las propiedades de este estimador del parámetro d dependen de las características estocásticas del término de error, e_j , aspecto que ha generado especial interés en la literatura reciente. Pero además, dependiendo de las características del proceso $Y_t = (1-L)^d X_t$, la no constancia del término $\ln\{f_Y(\mathbf{w}_j) / f_Y(0)\}$ puede generar un sesgo importante. En particular, cuando el modelo (1) incluye parámetros autorregresivos o media móvil de valor elevado, el estimador puede presentar un sesgo muy importante, como se muestra en Agiaklogou *et al.* (1993).

De todas maneras, a nuestro entender, es posible reducir dicho sesgo utilizando una extensión del modelo de regresión (8). Más concretamente, supóngase que se pretende estimar el parámetro d de un modelo ARFIMA(p, d, q): $\mathbf{f}(L)(1-L)^d X_t = \mathbf{q}(L)\mathbf{e}_t$. Si se pudiera añadir al modelo de regresión (8) algún regresor adicional que permita recoger la influencia de los parámetros autorregresivos y media móvil sobre el logaritmo del periodograma, es de esperar que la estimación del parámetro d tenga un sesgo menor. Pero la densidad espectral de X_t es, en logaritmos:

$$\begin{aligned} \ln f_X(\mathbf{w}) = & \ln\{\mathbf{s}_e^2\} - \ln(2\mathbf{p}) - d \ln\{4\sin^2(\mathbf{w} / 2)\} + \\ & + 2\ln\{\mathbf{q}(\mathbf{w})\} - 2\ln\{\mathbf{f}(\mathbf{w})\} \end{aligned} \quad (10)$$

donde $\mathbf{q}(\mathbf{w}) = 1 - \mathbf{q}_1 e^{-i\mathbf{w}} - \dots - \mathbf{q}_q e^{-iq\mathbf{w}}$ y $\mathbf{f}(\mathbf{w}) = 1 - \mathbf{f}_1 e^{-i\mathbf{w}} - \dots - \mathbf{f}_q e^{-ip\mathbf{w}}$. Por tanto, $\ln f_X(\mathbf{w})$ depende de los parámetros autorregresivos y media móvil de manera no lineal.

Pero utilizando una parametrización diferente de Y_t es posible escribir $\ln f_X(\mathbf{w})$ como función lineal del parámetro d y de los parámetros asociados con la memoria corta de la serie. Para ello, es suficiente utilizar una modelización propuesta por Bloomfield (1973), teniendo en cuenta que Y_t es estacionario:

$$f_Y(\mathbf{w}) = \frac{t^2}{2P} \exp \left\{ 2 \sum_{k=1}^l \mathbf{b}_k \cos(k\mathbf{w}) \right\} \quad (11)$$

Para garantizar que tiene sentido la utilización del modelo (11) no es necesario recurrir a ninguna hipótesis adicional, gracias al siguiente Lema:

Lema 1 (*Condiciones de Dirichlet*). Sea $g(x)$ una función que cumple las siguientes condiciones:

- (i) $g(x)$ está definida y tiene un valor único con excepción de un número finito de puntos en el intervalo $(-L, L)$
- (ii) $g(x)$ es periódica con período $2L$
- (iii) $g(x)$ y $g'(x)$ son continuas por intervalos en $(-L, L)$

y se definen los coeficientes de Fourier asociados a esta función como:

$$a_n = \frac{1}{L} \int_{-L}^L g(x) \cos \frac{n\mathbf{P}x}{L} dx \quad b_n = \frac{1}{L} \int_{-L}^L g(x) \sin \frac{n\mathbf{P}x}{L} dx \quad (12)$$

entonces, la siguiente serie converge a $g(x)$ siempre que x sea un punto de continuidad de dicha función:

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\mathbf{P}x}{L} + b_n \sin \frac{n\mathbf{P}x}{L} \right) \quad (13)$$

Demostración: Véase, por ejemplo, Körner (1988)

Teorema 1: Sea $f(\mathbf{w})$ la densidad espectral de un proceso ARMA Entonces, para cualquier $\mathbf{e} > 0$, existe una sucesión de coeficientes $\mathbf{b}_k, k=1, \dots, l_e$ tales que:

$$\left| f(\mathbf{w}) - \frac{\mathbf{t}^2}{2\mathbf{P}} \exp \left\{ 2 \sum_{k=1}^{l_e} \mathbf{b}_k \cos(k\mathbf{w}) \right\} \right| < \mathbf{e} \quad (14)$$

Demostración. Para demostrar el Teorema es suficiente aplicar el Lema anterior a la función $g(\mathbf{w}) = \ln f(\mathbf{w})$ en el intervalo $(-L, L) = (-\mathbf{P}\mathbf{P}, \mathbf{P}\mathbf{P})$ ya que se cumplen las condiciones necesarias.

Por tanto, todo modelo ARMA puede aproximarse a través de un modelo con densidad espectral de la forma (11), de manera que, sustituyendo la expresión (11) en (1) se obtiene que la densidad espectral de X_t (en logaritmos) es:

$$f_X(\mathbf{w}) = \left\{ 4 \sin^2(\mathbf{w}/2) \right\}^{-d} \frac{\mathbf{t}^2}{2\mathbf{P}} \exp \left\{ 2 \sum_{k=1}^l \mathbf{b}_k \cos(k\mathbf{w}) \right\} \quad (15)$$

de manera que, si se introduce el periodograma evaluado en las frecuencias armónicas, se obtiene la siguiente igualdad:

$$\ln I(\mathbf{w}_j) = \ln \frac{\mathbf{t}^2}{2\mathbf{P}} - d \ln \left\{ 4 \sin^2 \frac{\mathbf{w}_j}{2} \right\} + 2 \sum_{k=1}^l \mathbf{b}_k \cos k\mathbf{w} + \ln \frac{I(\mathbf{w}_j)}{f_X(\mathbf{w}_j)} \quad (16)$$

hecho que motiva la utilización del siguiente modelo de regresión:

$$\ln I(\mathbf{w}_j) = \mathbf{a} + \mathbf{b}R_j + \sum_{k=1}^l \mathbf{b}_k \cos(k\mathbf{w}) + e_j \quad (17)$$

Nótese como este modelo es una extensión del modelo (8) donde se han añadido regresores adicionales. La ventaja de utilizar este modelo ampliado consiste en que los regresores incorporados permiten discriminar las componentes de memoria larga de las componentes de memoria corta. De esta manera, en comparación con el desarrollo anterior, no es necesario suponer la constancia del término $\ln\{f_Y(\mathbf{w}_j) / f_Y(0)\}$ ya que éste ha sido substituido por una parametrización explícita a través del modelo de Bloomfield. Para distinguir este estimador del estimador GPH lo notaremos como estimador GPH ampliado (GPHA).

Finalmente, cabe destacar que, aunque en el desarrollo del método de estimación se ha hecho referencia al modelo de Bloomfield debido a sus especiales características, entendemos que es aplicable a modelos ARFIMA ya que los parámetros autorregresivos y media móvil deben estimarse en una etapa posterior.

Pero además, las condiciones impuestas sobre la densidad espectral son tan flexibles, que también puede utilizarse dicho estimador de manera semiparamétrica. Así, sea X_t un proceso estocástico estacionario tal que existe un parámetro $d \in (-1/2, 1/2)$ de manera que $Y_t = (1-L)^d X_t$ tiene una densidad espectral acotada y estrictamente positiva, definición habitual de proceso fraccionalmente integrado de orden d , o $X_t \sim FI(d)$.

No parece demasiado restrictivo, en este caso, suponer que además la densidad espectral del proceso cumple las condiciones del **Lema 1**, de manera que puede utilizarse el método de estimación propuesto sin necesidad de limitar el análisis al supuesto de que se trata de un modelo ARFIMA. Naturalmente, en este caso es posible obtener un estimador del parámetro d sin disponer de una modelización paramétrica plena del proceso generador de datos.

De todas maneras, debido al gran interés generado por los modelos ARFIMA y a su especial simplicidad y flexibilidad, en el apartado siguiente se

analizan las propiedades de este estimador con muestras pequeñas en el contexto de modelos ARFIMA.

3. Validación mediante un ejercicio de simulación

Para analizar las características del estimador GPHA del parámetro de diferencia fraccional y compararlas con las del estimador GPH, se ha realizado un ejercicio de simulación. Concretamente, se han generado grupos de 10.000 replicaciones de series generadas a partir de un modelo ARFIMA(1,d,1):

$$(1 - fL)(1 - L)^d X_t = (1 - qL)e_t \quad (18)$$

con diferentes parámetros y diferentes tamaños muestrales. A continuaciónPara cada una de las series generadas, se han utilizado tanto el estimador GPH como el GPHA para obtener estimaciones del parámetro d utilizando m ordenadas del periodograma donde $m = \lceil T^\alpha \rceil$ con distintos valores de α .

De entre todos los resultados obtenidos, se comentan únicamente los más significativos. En primer lugar, tiene especial interés tiene analizar las propiedades de dichos estimadores cuando se obtienen los datos a partir de un modelo ARMA(1,1).

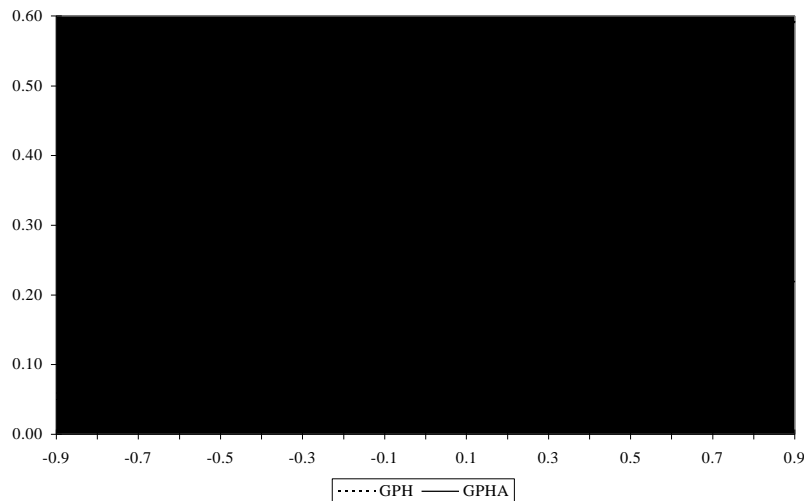
Así, los resultados de la **Tabla 1** permiten comprobar como la utilización del estimador GPH a series generadas según un modelo AR(1) con $f > 0.6$ (con $T=100$ observacions) conduce a estimaciones del parámetro d fuertemente sesgadas. En el **Gráfico 1** puede comprobarse como este sesgo es tan importante como para que en más del 70% de los casos, la utilización del estimador GPH conduzca al rechazo de la hipótesis nula de que $d=0$ aún siendo cierta.

Tabla 1. Resultados de la estimación de d a partir de simulación con series de 100 observaciones de procesos AR(1) (10.000 replicaciones).

f	GPH						GPHA ($l=1$)					
	a=1/2		a=2/3		a=3/4		a=1/2		a=2/3		a=3/4	
	med.	d.e.	med.	d.e.	med.	d.e.	med.	d.e.	med.	d.e.	med.	d.e.
-0.90	-0.01	0.30	-0.04	0.18	-0.09	0.15	0.00	0.59	0.01	0.33	0.01	0.26
-0.80	-0.02	0.30	-0.04	0.19	-0.09	0.15	-0.01	0.58	0.00	0.33	0.01	0.26
-0.70	-0.01	0.31	-0.04	0.19	-0.09	0.15	0.00	0.60	0.00	0.33	0.01	0.26
-0.60	-0.01	0.30	-0.04	0.19	-0.09	0.15	0.00	0.60	0.00	0.33	0.01	0.26
-0.50	-0.01	0.30	-0.04	0.19	-0.08	0.15	-0.01	0.60	0.00	0.33	0.01	0.26
-0.40	-0.01	0.31	-0.03	0.19	-0.07	0.15	-0.01	0.60	0.01	0.33	0.01	0.26
-0.30	-0.02	0.30	-0.03	0.19	-0.06	0.15	-0.01	0.60	0.00	0.34	0.00	0.26
-0.20	-0.01	0.31	-0.03	0.19	-0.05	0.15	-0.01	0.59	0.00	0.33	0.00	0.26
-0.10	-0.01	0.30	-0.01	0.19	-0.03	0.15	0.01	0.60	0.00	0.33	0.00	0.26
0.00	0.00	0.31	0.00	0.19	0.00	0.15	0.00	0.59	-0.01	0.34	-0.01	0.26
0.10	0.00	0.31	0.02	0.19	0.03	0.15	-0.01	0.60	0.00	0.34	0.00	0.26
0.20	0.01	0.30	0.04	0.19	0.07	0.15	0.00	0.59	0.00	0.33	0.00	0.26
0.30	0.03	0.30	0.08	0.19	0.13	0.15	-0.01	0.60	0.01	0.33	0.02	0.26
0.40	0.04	0.30	0.13	0.19	0.20	0.15	0.00	0.60	0.01	0.33	0.02	0.26
0.50	0.08	0.30	0.19	0.19	0.27	0.15	0.01	0.59	0.02	0.33	0.01	0.26
0.60	0.13	0.31	0.28	0.19	0.37	0.15	0.01	0.59	0.03	0.34	0.01	0.26
0.70	0.21	0.31	0.40	0.19	0.49	0.15	0.05	0.60	0.05	0.33	0.05	0.26
0.80	0.37	0.31	0.55	0.19	0.64	0.15	0.15	0.59	0.16	0.33	0.17	0.26
0.90	0.71	0.31	0.77	0.19	0.82	0.15	0.43	0.60	0.41	0.33	0.38	0.26

Med.: Media de \tilde{d} . D.e.: Desviación estándar de \tilde{d}

Gráfico 1. Porcentaje de rechazos de la hipótesis $d=0$ en series generadas según un AR(1) en función del parámetro f (nivel de significación del 5%).



En dicha tabla puede comprobarse como dicho porcentaje se reduce de manera significativa si se utiliza el estimador GPHA mediante la incorporación de un regresor adicional ($l=1$) en el modelo de regresión aplicado.

En todo caso, aunque la reducción es mucho mayor cuando se utiliza $\mathbf{a}=1/2$ (es decir, $m=10$) que cuando $\mathbf{a}=2/3$ (es decir $m=21$), debe tenerse en cuenta que ello es debido no a las mejores propiedades del estimador para un menor número de frecuencias en el modelo de regresión, sino a que con un menor número de observaciones, se obtiene un estimador con mayor varianza y, por tanto, un intervalo de confianza mayor.

Tabla 2. Resultados de la estimación de d a partir de simulación con series de 100 observaciones de procesos MA(1) (10.000 replicaciones).

q	GPH						GPHA ($l=1$)					
	$\mathbf{a}=1/2$		$\mathbf{a}=2/3$		$\mathbf{a}=3/4$		$\mathbf{a}=1/2$		$\mathbf{a}=2/3$		$\mathbf{a}=3/4$	
	med.	<i>d.e.</i>	med.	<i>d.e.</i>	med.	<i>d.e.</i>	med.	<i>d.e.</i>	med.	<i>d.e.</i>	med.	<i>d.e.</i>
-0.90	0.01	0.30	0.04	0.19	0.09	0.15	0.00	0.59	0.00	0.33	-0.01	0.26
-0.80	0.01	0.30	0.05	0.19	0.09	0.15	0.00	0.59	-0.01	0.34	-0.01	0.26
-0.70	0.01	0.31	0.04	0.19	0.09	0.15	-0.01	0.60	0.00	0.34	-0.01	0.26
-0.60	0.01	0.30	0.04	0.19	0.08	0.15	0.00	0.60	0.00	0.33	-0.01	0.26
-0.50	0.01	0.30	0.04	0.19	0.08	0.15	-0.01	0.60	-0.01	0.33	-0.01	0.26
-0.40	0.01	0.31	0.04	0.19	0.07	0.15	-0.01	0.60	0.00	0.33	-0.01	0.26
-0.30	0.00	0.30	0.03	0.19	0.06	0.15	-0.01	0.60	-0.01	0.34	-0.01	0.26
-0.20	0.00	0.31	0.02	0.19	0.05	0.15	-0.01	0.59	-0.01	0.33	0.00	0.26
-0.10	0.00	0.30	0.01	0.19	0.02	0.15	0.01	0.60	0.00	0.33	0.00	0.26
0.00	0.00	0.31	0.00	0.19	0.00	0.15	0.00	0.59	-0.01	0.34	-0.01	0.26
0.10	-0.01	0.31	-0.02	0.19	-0.04	0.15	-0.01	0.60	0.00	0.34	0.00	0.26
0.20	-0.02	0.30	-0.05	0.19	-0.08	0.15	0.00	0.59	0.00	0.33	-0.01	0.26
0.30	-0.03	0.30	-0.08	0.19	-0.13	0.15	-0.01	0.59	-0.01	0.33	-0.02	0.26
0.40	-0.05	0.30	-0.13	0.19	-0.19	0.15	0.00	0.60	-0.02	0.33	-0.04	0.26
0.50	-0.08	0.30	-0.19	0.19	-0.28	0.15	-0.01	0.59	-0.04	0.33	-0.09	0.26
0.60	-0.13	0.30	-0.28	0.19	-0.37	0.15	-0.03	0.59	-0.08	0.33	-0.15	0.26
0.70	-0.25	0.30	-0.34	0.19	-0.48	0.15	-0.06	0.59	-0.10	0.33	-0.18	0.26
0.80	-0.42	0.30	-0.49	0.19	-0.52	0.15	-0.11	0.59	-0.14	0.33	-0.19	0.26
0.90	-0.61	0.30	-0.68	0.19	-0.71	0.15	-0.16	0.59	-0.17	0.33	-0.20	0.26

Med.: Media de \tilde{d} . D.e.: Desviación estándar de \tilde{d}

Para realizar de manera correcta la comparación entre ambos métodos de estimación debe utilizarse el estimador GPHA con $\mathbf{a}=2/3$ ya que, en este caso, la varianza del estimador es parecida al varianza del estimador GPH con $\mathbf{a}=1/2$. En estas condiciones, puede comprobarse en la **Tabla 1** como el sesgo del estimador GPH es mucho menor que el sesgo del estimador GPH. Obsérvese

además, que a medida que se utiliza un mayor número de frecuencias en la estimación del parámetro d , el sesgo del estimador GPH aumenta mientras que el sesgo del estimador GPHA se reduce.

Cuando el proceso generador de datos contiene un parámetro media móvil en lugar de un parámetro autorregresivo las conclusiones son similares. Tal como se muestra en Agiakloglou *et al.* (1993), un valor positivo y elevado del parámetro produce un importante sesgo en el estimador GPH (aunque menor que bajo la presencia de parámetros autorregresivos) y por tanto, un tamaño efectivo del contraste muy superior al tamaño nominal. En la **Tabla 2** puede comprobarse como, nuevamente, el estimador GPHA presenta mejores propiedades que el estimador GPH en este caso.

Hasta el momento, se han analizado las propiedades del estimador GPHA sólo para procesos $FI(0)$. En la **Tabla 3** se comparan los estimadores GPH y GPHA cuando $d \neq 0$. Estos resultados permiten apreciar como la ampliación del modelo de regresión utilizado no provoca la aparición de ningún sesgo en la estimación del parámetro de integración fraccional.

En todo caso, el aumento de la varianza del estimador derivada de la ampliación del número de regresores utilizados puede compensarse con la utilización de un mayor número de frecuencias para reducir la varianza del estimador (en este caso, $m=21$ en lugar de $m=10$). Recuérdese que al utilizar el estimador GPH, el aumento del número de frecuencias utilizadas puede provocar la aparición de un mayor sesgo del estimador. En cambio, este aumento del número de frecuencias es posible con el estimador GPH sin incurrir en el peligro de aumentar el sesgo gracias, precisamente, a que la presencia de los regresores adicionales permiten eliminar el sesgo derivado de la posible presencia de parámetros de memoria corta.

En definitiva, estos resultados permiten validar de manera clara, las ventajas de utilizar un modelo de regresión ampliado en el dominio de las

frecuencias para minimizar los problemas de sesgo que presenta el estimador GPH en ciertas situaciones.

Tabla 3. Resultados de la estimación de d a partir de simulación con series de 100 observaciones de procesos FI(d) (10.000 replicaciones).

d	GPH						GPHA ($l=1$)					
	a=1/2		a=2/3		a=3/4		a=1/2		a=2/3		a=3/4	
	med.	d.e.	med.	d.e.	med.	d.e.	med.	d.e.	med.	d.e.	med.	d.e.
-0.45	-0.43	0.30	-0.43	0.19	-0.44	0.15	-0.44	0.59	-0.44	0.33	-0.44	0.26
-0.40	-0.36	0.30	-0.38	0.19	-0.38	0.15	-0.37	0.59	-0.36	0.33	-0.37	0.26
-0.35	-0.34	0.30	-0.33	0.19	-0.34	0.15	-0.33	0.59	-0.33	0.33	-0.33	0.26
-0.30	-0.27	0.31	-0.29	0.18	-0.29	0.15	-0.29	0.59	-0.28	0.33	-0.28	0.26
-0.25	0.24	0.30	-0.25	0.19	-0.23	0.15	-0.25	0.59	-0.24	0.33	-0.25	0.26
-0.20	-0.19	0.30	-0.20	0.19	-0.20	0.15	-0.20	0.60	-0.19	0.33	-0.20	0.26
-0.15	-0.15	0.30	-0.15	0.19	-0.15	0.15	-0.15	0.59	-0.15	0.33	-0.16	0.26
-0.10	-0.10	0.31	-0.10	0.19	-0.10	0.15	-0.10	0.60	-0.10	0.33	-0.10	0.26
-0.05	-0.05	0.30	-0.05	0.19	-0.05	0.15	-0.05	0.60	-0.05	0.33	-0.05	0.26
0.00	-0.01	0.30	0.00	0.19	0.00	0.15	-0.02	0.59	0.00	0.33	0.00	0.26
0.05	0.05	0.30	0.05	0.19	0.05	0.15	0.05	0.59	0.05	0.33	0.05	0.26
0.10	0.10	0.30	0.10	0.19	0.10	0.15	0.10	0.60	0.10	0.33	0.10	0.26
0.15	0.16	0.30	0.15	0.19	0.15	0.15	0.15	0.59	0.15	0.33	0.15	0.26
0.20	0.21	0.30	0.21	0.18	0.21	0.15	0.22	0.59	0.21	0.33	0.21	0.26
0.25	0.25	0.30	0.25	0.19	0.25	0.15	0.25	0.59	0.25	0.33	0.25	0.26
0.30	0.32	0.30	0.31	0.18	0.29	0.15	0.30	0.59	0.31	0.32	0.30	0.26
0.35	0.36	0.30	0.34	0.18	0.33	0.15	0.34	0.59	0.35	0.32	0.35	0.26
0.40	0.40	0.30	0.42	0.19	0.43	0.15	0.41	0.58	0.40	0.33	0.41	0.26
0.45	0.45	0.30	0.46	0.19	0.46	0.15	0.45	0.58	0.45	0.33	0.45	0.26

Med.: Media de \tilde{d} . D.e.: Desviación estándar de \tilde{d}

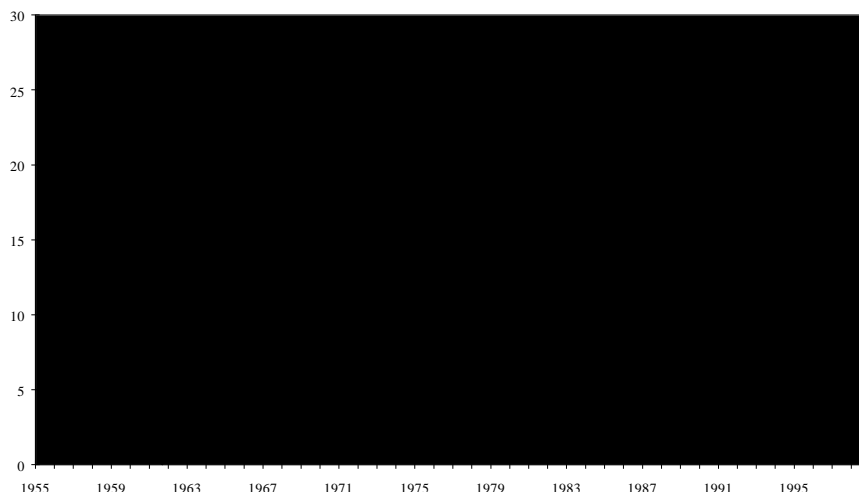
4. Una aplicación al IPC español

Para ilustrar la utilidad de la extensión propuesta del estimador GPH, se ha aplicado éste a la estimación del orden de integración de la serie de inflación en España. Se trata de un ejemplo de especial interés, teniendo en cuenta que es posible encontrar en la literatura econométrica una importante controversia en relación a la posibilidad de que las tasas de crecimiento del Índice de Precios al Consumo (IPC) español contenga una raíz unitaria.

Nótese que la aceptación de dicha hipótesis tiene importantes implicaciones de política económica ya que la presencia de una raíz unitaria en el crecimiento de los precios implica que, en tal caso, los *shocks* que afecten a la tasa de inflación en el presente deberían tener un efecto permanente en las tasas de inflación futuras.

Para analizar la verosimilitud de la presencia de una raíz unitaria en las tasas de inflación en España, se han utilizado los datos mensuales del IPC español desde marzo de 1954 hasta diciembre de 1998 (un total de 538 observaciones).

Gráfico 2. Tasas de inflación en España (1955.3 - 1998.12).



En primer lugar, y a pesar de que tanto el estimador GPH como el estimador GPHA tienen propiedades conocidas en la estimación del orden de

integración de procesos estacionarios, se han utilizado estos estimadores, a modo de análisis exploratorio, para obtener una estimación del orden de integración de la serie de tasas de inflación. Así, en la **Tabla 4** se recogen las estimaciones obtenidas con diferentes números de frecuencias entre $m=20$ y $m=40$.

Tabla 4. Resultados de la estimación de d a partir de la tasa de inflación en España (Marzo 1955 - Diciembre 1998).

m	GPH		GPHA ($l=1$)		GPHA ($l=2$)	
	\tilde{d}	e.s.(\tilde{d})	\tilde{d}	e.s.(\tilde{d})	\tilde{d}	e.s.(\tilde{d})
20	0.86	0.021	1.02	0.067	0.86	0.147
21	0.84	0.019	1.03	0.059	0.88	0.127
22	0.87	0.018	0.94	0.057	0.87	0.125
23	0.90	0.016	0.89	0.053	0.92	0.111
24	0.89	0.015	0.90	0.048	0.87	0.100
25	0.91	0.014	0.87	0.043	0.89	0.089
26	0.95	0.014	0.81	0.041	0.92	0.080
27	0.95	0.012	0.82	0.038	0.87	0.074
28	0.98	0.012	0.79	0.035	0.88	0.067
29	1.02	0.012	0.74	0.034	0.89	0.061
30	1.02	0.012	0.77	0.032	0.80	0.061
31	1.00	0.011	0.83	0.032	0.85	0.064
32	1.00	0.010	0.84	0.029	0.84	0.059
33	1.04	0.011	0.79	0.029	0.89	0.057
34	1.03	0.010	0.83	0.028	0.83	0.055
35	1.03	0.009	0.85	0.026	0.80	0.052
36	1.07	0.010	0.80	0.027	0.88	0.052
37	1.08	0.010	0.80	0.025	0.86	0.049
38	1.10	0.010	0.79	0.024	0.87	0.046
39	1.11	0.009	0.80	0.022	0.84	0.044
40	1.15	0.010	0.76	0.023	0.89	0.043

Fuente: Elaboración propia.

De estos resultados, destaca la proximidad a la unidad de las estimaciones basadas en el estimador GPH, incluso con estimaciones superiores a la unidad a medida que aumenta m . Téngase en cuenta, no obstante, que a medida que aumenta el número de frecuencias utilizadas aumenta el sesgo potencial de este estimador.

En cambio, los resultados obtenidos mediante el estimador propuesto en este trabajo (con $l=1$ y $l=2$), parecen sugerir la no presencia de una raíz unitaria

en las tasas de inflación en favor de un parámetro de diferenciación inferior a la unidad.

Para obtener conclusiones más fiables se han calculado estimaciones del parámetro de integración fraccional a partir de los incrementos de las tasas de crecimiento. Así, sea X_t la serie de tasas de inflación, se define una nueva serie $Y_t = X_t - X_{t-1}$. Si $X_t \sim \text{FI}(d)$ entonces $Y_t \sim \text{FI}(d^*)$ con $d = 1 + d^*$. En la **Tabla 5** se recogen las estimaciones del parámetro d^* obtenidas con diferentes números de frecuencias entre $m=20$ y $m=40$.

Tabla 5. Resultados de la estimación de d^* a partir del incremento de la tasa de inflación en España (Marzo 1955 - Diciembre 1998).

m	<i>GPH</i>		<i>GPHA (l=1)</i>		<i>GPHA (l=2)</i>	
	\tilde{d}	e.s.(\tilde{d})	\tilde{d}	e.s.(\tilde{d})	\tilde{d}	e.s.(\tilde{d})
20	-0.16	0.018 **	0.00	0.059	-0.12	0.130
21	-0.18	0.017 **	0.02	0.052	-0.11	0.112
22	-0.15	0.016 **	-0.07	0.051	0.02	0.110
23	-0.11	0.015 **	-0.14	0.049 **	0.09	0.099
24	-0.11	0.014 **	-0.14	0.044 **	0.04	0.090
25	-0.10	0.013 **	-0.16	0.040 **	0.03	0.080
26	-0.07	0.012 **	-0.20	0.037 **	0.05	0.072
27	-0.06	0.011 **	-0.20	0.034 **	0.01	0.066 *
28	-0.02	0.011	-0.24	0.032 **	0.02	0.060 *
29	0.01	0.012	-0.28	0.031 **	0.03	0.054 *
30	0.01	0.011	-0.24	0.030 **	-0.08	0.058 **
31	-0.01	0.010	-0.19	0.029 **	-0.17	0.059 **
32	0.00	0.010	-0.19	0.027 **	-0.18	0.054 **
33	0.04	0.010 **	-0.23	0.027 **	-0.12	0.052 **
34	0.03	0.010 **	-0.19	0.026 **	-0.19	0.052 **
35	0.03	0.009 **	-0.17	0.025 **	-0.21	0.049 **
36	0.07	0.010 **	-0.22	0.025 **	-0.15	0.047 **
37	0.08	0.009 **	-0.22	0.023 **	-0.16	0.045 **
38	0.10	0.009 **	-0.22	0.022 **	-0.16	0.042 **
39	0.11	0.009 **	-0.21	0.021 **	-0.18	0.040 **
40	0.15	0.009 **	-0.25	0.020 **	-0.14	0.039 **

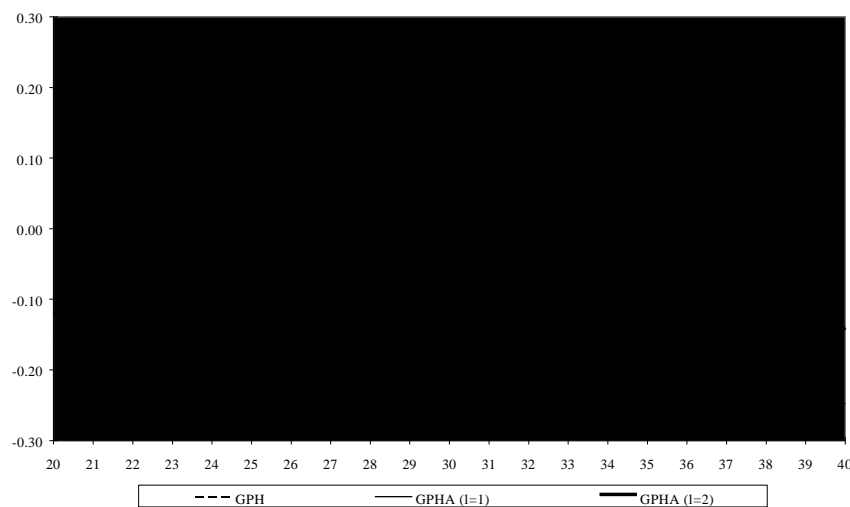
Notación: * indica valores significativos al 5% y ** valores significativos al 1%.

Nótese como estos resultados confirman la apreciación anterior, es decir, que con el estimador GPHA se confirma la no presencia de una raíz unitaria en las tasas de inflación en España ya que para $m > 25$ se obtienen estimaciones

negativas y significativas del parámetro de diferenciación fraccional. Esto contrasta con los resultados que se obtienen si se utiliza el estimador GPH que no permiten rechazar la presencia de una raíz unitaria para $m > 27$ y incluso, para $m > 33$ se obtienen valores positivos significativos.

A nuestro entender, las diferencias que aparecen entre las estimaciones obtenidas con GPH y con GPHA y que pueden apreciarse de manera muy clara en el **Gráfico 3** son debidas a los problemas de sesgo del estimador GPH, sesgo que aumenta a medida que se utiliza un mayor número de frecuencias en la estimación. Por tanto, este parece un claro ejemplo de como la utilización de dicho estimador puede conducir a conclusiones que, como mínimo, deberían ser dudosas.

Gráfico 3. Estimación del orden de integr. fraccional de la inflación en España.



Finalmente, destacar que los resultados obtenidos confirman la no estacionariedad de las tasas de inflación en España y su alta persistencia o memoria ya que se obtienen valores de d claramente superiores a 0.5 pero también parece claro que puede rechazarse la presencia de una raíz unitaria, de manera que los *shocks* que puedan alterar estas tasas de inflación en un cierto

momento del tiempo no pueden tener un efecto permanente en las tasas de inflación futuras.

5. Conclusiones

Como ya se ha comentado, se está popularizando durante los últimos años la utilización de los modelos ARFIMA para el análisis de la evolución temporal de diversas magnitudes económicas, modelos cuya principal característica es la de que permiten la presencia de órdenes de integración fraccional, sin restringir estos a valores naturales.

Entre los métodos propuestos para la estimación del parámetro de integración fraccional en estos modelos, el más utilizado es el propuesto por Geweke y Porter-Hudak (1983) a pesar de que en Agiakloglou *et al.* (1993) se muestra como, en ciertas circunstancias, dicho estimador puede estar seriamente sesgado.

En el trabajo presentado se ha propuesto una extensión de dicho estimador GPH basada en el modelo exponencial de Bloomfield (1973) que permite eliminar dicho sesgo y mejorar de manera clara sus propiedades. La eliminación de este sesgo es especialmente importante cuando se pretende realizar algún tipo de inferencia acerca del parámetro de integración.

Así, mientras que en el trabajo de Agiakloglou *et al.* (1993) se concluye que los problemas de sesgo asociados al estimador GPH demuestran la necesidad de estimar de manera conjunta todos los parámetros del modelo (1) siguiendo las propuestas de Fox y Taqqu (1986) o Sowell (1992), los resultados presentados en este trabajo permiten, a nuestro entender, matizar esta conclusión, en el sentido de que es posible mejorar las propiedades del estimador GPH mediante una ampliación del modelo de regresión utilizado en el dominio de las frecuencias.

Para ilustrar las ventajas del procedimiento propuesto, se ha aplicado éste a la estimación del orden de integración de las tasas de inflación en España. Esta aplicación ha permitido, además de mostrar las ventajas del método propuesto, encontrar evidencia clara en contra de la hipótesis de que estas tasas de inflación contienen una raíz unitaria a pesar de su no estacionariedad. Esta conclusión es importante en la medida que conduce a la consideración de que los *shocks* que puedan afectar a la inflación en España no son permanentes sino transitorios.

Referencias bibliográficas

Agiakloglou, C.; P. Newbold i M. Wohar (1992): "Bias in an estimator of the fractional difference parameter", *Journal of Time Series Analysis* **14**, 235-246.

Bloomfield, P. (1973): "An Exponential model for the spectrum of a scalar linear time series", *Biometrika* **60**, 217-226.

Diebold, F.X. i G.D. Rudebusch (1989): "Long memory and persistence in aggregate output", *Journal of Monetary Economics* **24**, 189-209.

Fox, R. i M.S. Taqqu (1986): "Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series", *The Annals of Statistics* **14**, 517-532.

Geweke, J. i S. Porter-Hudak (1983): "The estimation and application of long memory time series models", *Journal of Time Series Analysis* **4**, 221-238.

Hosking, J.R.M. (1981): "Fractional differencing", *Biometrika* **68**, 165-176.

Körner, T.W. (1988): *Fourier Analysis*, Cambridge University Press, Cambridge.

Porter-Hudak, S. (1990): "An application of the seasonal fractionally differenced model to the monetary aggregates", *Journal of the American Statistical Association* **85**, 338-344.

Sowell, F. (1992): "Maximum likelihood estimation of stationary univariate fractionally integrated time series models", *Journal of Econometrics* **53**, 165-188.