

18/  
49  
(19408)



## Documento de Trabajo

Una nota sobre la estimación eficiente de modelos  
con parámetros cambiantes

Sonia Sotoca

No. 9408

Junio 1994



Instituto Complutense de Análisis Económico

UNIVERSIDAD COMPLUTENSE

FACULTAD DE ECONOMICAS

Campus de Somosaguas

28223 MADRID

Teléfono 3942811 - FAX 3942613



Instituto Complutense de Análisis Económico

UNIVERSIDAD COMPLUTENSE



W  
49  
(19408)

**UNA NOTA SOBRE LA ESTIMACION EFICIENTE DE MODELOS  
CON PARAMETROS CAMBIANTES**

**Sonia Sotoca\***  
Instituto Complutense de Análisis Económico  
Universidad Complutense  
Campus de Somosaguas  
28223 Madrid

**ABSTRACT**

---

Standard estimation procedures for the time-varying parameters model suppose that the variances of the noises in the model are known. Obviously, this assumption is not realistic in most econometric applications. Besides, the results of these methods are sensitive to the initial conditions of the algorithm, a fact that is often overlooked by the literature. In this paper, we propose an extension of the recursive algorithm proposed by Cooley, Rosenberg y Wall (1977), which is independent of initial conditions and includes on-line estimation of all the relevant variances. The results obtained with this method compare favourably with those obtained by standard procedures.

---

**RESUMEN**

---

Los procedimientos estándar para estimar modelos de parámetros cambiantes suponen conocidas las varianzas de los términos de error presentes en el modelo. Obviamente, éste no es un supuesto realista en la mayor parte de las aplicaciones econométricas. Por otra parte, los resultados que proporcionan estos métodos son sensibles a las condiciones iniciales, hecho que habitualmente es ignorado por la literatura. En este trabajo se propone una extensión del algoritmo recursivo debido a Cooley, Rosenberg y Wall (1977), que es independiente de condiciones iniciales e incorpora la estimación *on-line* de todas las varianzas relevantes. Los resultados obtenidos con este procedimiento se comparan favorablemente con los obtenidos usando los métodos habituales.

---

\* Quiero agradecer los comentarios y sugerencias recibidos de Rafael Flores, Miguel Jerez y Alfonso Novales.

## Introducción

En ocasiones, una estructura lineal de parámetros fijos no permite modelizar adecuadamente determinadas relaciones causales. Esto puede deberse a diversos factores como cambios en la estructura de la relación, existencia de no linealidades, así como otras formas en que la incorrecta especificación puede llevar a variación en los parámetros. En este contexto, diversos autores como Cooley y Prescott (1973, 1976), Pagan (1980) y Young (1984), proponen usar modelos de parámetros cambiantes, que pueden reducirse al siguiente modelo en forma de espacio de los estados:

$$y_t = x_t^T \beta_t + \epsilon_t$$

$$\beta_t = \Phi \beta_{t-1} + u_t$$

donde la ley de evolución del vector  $\beta_t$  es la ecuación de estado del sistema,  $\epsilon_t$  y  $u_t$  son variables aleatorias independiente e idénticamente distribuidas con esperanza nula y varianzas  $\sigma^2$  y  $Q$ , respectivamente. Estas varianzas suelen ser desconocidas y el tamaño relativo de las mismas indica cuán cerca estamos de un modelo de parámetros fijos o de un modelo de parámetros cambiantes en el tiempo. Haciendo diferentes supuestos sobre la matriz  $\Phi$  y modificando el vector de estado como sea necesario, pueden formularse distintos modelos, como es el modelo de constante adaptativa de Cooley y Prescott (1973), en el que la matriz  $\Phi$  es la identidad.

La aplicación de este tipo de modelos está cada vez más extendida en áreas como Economía financiera, Macroeconomía, así como en otros contextos donde el supuesto de constancia de los parámetros no es creíble. Trabajos en esta línea son, entre otros, los de Fabozzi y Francis (1978), Alexander y Benson (1982), Young (1984), Margaritis (1990) y Wells (1994).

Sin embargo, este trabajo se centra sólo en la estimación de una clase de modelos de parámetros cambiantes. Para estimar éstos, suele utilizarse un algoritmo en dos etapas en el que se aplican, sucesivamente, el filtro de Kalman y el smoother óptimo de intervalo fijo [ver Harvey (1989), cap.3]. Los resultados de este algoritmo (que, en adelante denominaremos FK-SIF) dependen crucialmente de disponer de estimaciones "adecuadas" del vector inicial de coeficientes y de las varianzas de los términos de error presentes en el modelo o, al menos, de la proporción entre ellas.

Nº E: 5306520295

Nº C: X-53-230684-0

El primer problema<sup>1</sup> fue abordado por Cooley, Rosenberg y Wall (1977), quienes propusieron un algoritmo (a partir de ahora, algoritmo CRW) independiente de condiciones iniciales, pero que requiere conocer las varianzas de las perturbaciones que intervienen en el modelo. Una posible solución a este problema es estimar el modelo por máxima verosimilitud mediante el filtro de Kalman [ver Cooley y Prescott (1976)]. Sin embargo, la experiencia práctica sugiere que: a) evaluar la función de verosimilitud de este tipo de modelos resulta costoso computacionalmente y b) el perfil de dicha función suele ser bastante plano en un entorno del máximo [ver García-Ferrer et al. (1993)]. Estos problemas han dado lugar a que algunos autores [ver Young (1984), cap. 5] sugieran la utilización de un ratio de varianzas arbitrario o "manual". Sin embargo, nuestra experiencia indica que dicha arbitrariedad afecta tanto a la eficiencia de los estimadores del resto de parámetros, como a la secuencia de estimaciones.

En este trabajo se propone una extensión del algoritmo CRW (a partir de ahora CRW1) que incorpora una estimación recursiva de dichas varianzas, además de la trayectoria óptima de los parámetros propios del modelo. Esta nueva versión también es independiente de condiciones iniciales, al estar basado en un filtro de información en lugar del filtro de Kalman.

La estructura del trabajo es la siguiente. En el apartado 1 se plantea el modelo objeto de estudio y el procedimiento de estimación CRW, poniendo especial énfasis en las ventajas de este algoritmo con respecto a otros alternativos.

En el apartado 2 se describe la forma de incorporar en el algoritmo CRW la estimación recursiva de las varianzas de los términos de error del modelo, dando lugar al criterio llamado CRW1.

En el apartado 3 se presentan los resultados obtenidos con este procedimiento usando datos simulados. Estos resultados se comparan favorablemente con los obtenidos mediante los algoritmos FK-SIF y CRW, siendo éste último el punto de referencia para comparar, dado que en el mismo se fijan los verdaderos valores de todas las varianzas relevantes.

Finalmente, en el apartado 4 se resumen las principales conclusiones del trabajo, así como sus posibles extensiones.

<sup>1</sup> Para una discusión detallada, ver Sotoca (1993).

## 1. El algoritmo CRW.

Sea el siguiente modelo de regresión con parámetros cambiantes en el tiempo:

$$y_t = x_t^T \beta_t + \epsilon_t \quad (1)$$

$$\beta_{t+1} = \beta_t + u_t \quad (2)$$

donde  $y_t$  y el vector  $x_t^T$  representan las variables observables del sistema,  $\beta_t$  es un vector de  $k$  parámetros desconocidos y las perturbaciones  $\epsilon_t$  y  $u_t$  son variables aleatorias independiente e idénticamente distribuidas con esperanza nula y varianzas  $\sigma^2$  y  $Q$ , respectivamente. El objetivo es estimar eficientemente el vector  $\beta_t$  a partir de las observaciones  $\{y_1, y_2, \dots, y_N\}$ .

El algoritmo CRW proporciona una solución al problema de estimación óptima del modelo dado por (1) y (2), conocidas las varianzas  $\sigma^2$  y  $Q$ . Este algoritmo se basa en la combinación de dos filtros de información [ver Anderson y Moore (1979), cap. 6], lo que da lugar a un procedimiento recursivo de estimación independiente del vector inicial de parámetros<sup>2</sup>.

Denotando las matrices de información en las etapas de predicción y actualización por  $H_{t|t-1}$  y  $H_{tt}$ , respectivamente, las variables auxiliares que juegan el papel de los parámetros en un filtro de información se definen como:

$$f_{t|t-1} = H_{t|t-1} \hat{\beta}_{t|t-1} \quad (3)$$

$$f_{tt} = H_{tt} \hat{\beta}_{tt}$$

y las correspondientes matrices de covarianzas del vector  $\hat{\beta}_t$  son:

$$P_{t|t-1} = H_{t|t-1}^{-1} \quad (4)$$

$$P_{tt} = H_{tt}^{-1}$$

El primer paso del algoritmo CRW consiste en propagar un filtro de información que recorre la muestra hacia adelante a partir de las condiciones iniciales:

<sup>2</sup> En un filtro de información se propagan las inversas de las matrices de covarianzas en lugar de las propias matrices. Esto hace que el algoritmo disponga de condiciones iniciales exactas y que las variables del filtro se propagen de forma estable a partir de su inicialización [ver Anderson y Moore (1979), cap. 6].

$$H_{1/0} = 0 \quad ; \quad f_{1/0} = 0 \quad (5)$$

Con esta inicialización, la fase de predicción de un filtro de información aplicado al modelo (1)-(2) viene dada por:

$$K_{t-1} = [I + H_{t-1/t-1}Q]^{-1} \quad (6)$$

$$H_{t/t-1} = K_{t-1}H_{t-1/t-1} \quad (7)$$

$$f_{t/t-1} = K_{t-1}f_{t-1/t-1} \quad (8)$$

donde  $K_t$  es la ganancia del filtro. Cuando un nuevo dato está disponible, las ecuaciones de actualización de este filtro son:

$$f_{t/t} = f_{t/t-1} + \frac{x_t y_t}{\sigma^2} \quad (9)$$

$$H_{t/t} = H_{t/t-1} + \frac{x_t x_t^T}{\sigma^2} \quad (10)$$

Evidentemente, los resultados de las ecuaciones (6)-(8) pueden sustituirse en (9)-(10), dando lugar a una única fase de estimación, lo que resulta ventajoso desde un punto de vista computacional.

El siguiente paso consiste en propagar un segundo filtro de información que recorre la muestra hacia atrás, es decir, en el orden  $t = N, N-1, \dots, 1$ . Denotando por  $G_{t/t+1}$  y  $G_{t/t}$  las matrices de información en las etapas de predicción y actualización de este nuevo filtro y las correspondientes variables auxiliares como  $r_{t/t+1}$  y  $r_{t/t}$ , respectivamente, la inicialización del filtro es de nuevo:

$$G_{N/N+1} = 0 \quad ; \quad r_{N/N+1} = 0 \quad (11)$$

Las fórmulas de predicción de este segundo filtro son:

$$J_{t+1} = [I + G_{t+1/t+1}Q]^{-1} \quad (12)$$

$$G_{t/t+1} = J_{t+1}G_{t+1/t+1} \quad (13)$$

$$r_{t/t+1} = J_{t+1}r_{t+1/t+1} \quad (14)$$

y las ecuaciones de actualización:

$$r_{t/t} = r_{t/t+1} + \frac{x_t y_t}{\sigma^2} \quad (15)$$

$$G_{t/t} = G_{t/t+1} + \frac{x_t x_t^T}{\sigma^2} \quad (16)$$

Comparando las ecuaciones (5)-(10) con (11)-(16), se observa que:

- 1) La estructura general de ambos filtros es idéntica.
- 2) Ambos algoritmos son independientes; esto es, para aplicar uno de ellos no es necesario conocer los resultados del otro<sup>3</sup>.
- 3) El algoritmo no requiere invertir la matriz de covarianzas  $Q$ , como ocurre con algunas versiones de este tipo de filtros [ver Anderson y Moore (1979), cap. 6]. Esto es importante ya que, en la práctica, es frecuente que dicha matriz sea singular<sup>4</sup>.
- 4) A la vista de las ecuaciones (5)-(16), es fácil demostrar que el algoritmo puede escribirse en función de ratios de varianzas, sin más que redefinir las variables auxiliares  $f$  y  $r$  como  $\sigma^2 f$  y  $\sigma^2 r$ , respectivamente.

El *smoothing* óptimo en el algoritmo CRW, se obtiene mediante una combinación lineal de los resultados de ambos filtros. Concretamente, la trayectoria óptima del vector  $\beta_t$  y sus correspondientes matrices de covarianzas pueden calcularse de la forma [ver Liebelt (1967)]:

$$P_{t/N} = [H_{t/t} + G_{t/t+1}]^{-1} \quad (17)$$

$$b_{t/N} = P_{t/N} [f_{t/t} + r_{t/t+1}] \quad (18)$$

<sup>3</sup> Esta propiedad es de gran interés computacional, ya que permite aplicar procesos de cálculo en paralelo.

<sup>4</sup> Concretamente, cuando alguno de los parámetros del modelo se suponga constante en el tiempo. Ver, por ejemplo, Cooley y Prescott (1973).

## 2. Estimación recursiva de las varianzas del modelo: el algoritmo CRW1.

El principal problema del algoritmo CRW es que supone conocidas las varianzas  $\sigma^2$  y  $Q$ , lo que no es habitual en la práctica econométrica con datos reales. Por otra parte, la elección arbitraria del valor de estas varianzas influye en todo el proceso de estimación. En concreto, cuanto mayor sea  $\sigma^2$  comparada con las varianzas asociadas a la ley de variación de los parámetros, más cerca estaremos de un modelo de parámetros fijos. En el caso contrario, permitimos mucha más variabilidad al parámetro(s) cambiante(s), pudiendo llegar a que las fluctuaciones de la variable a explicar coincidan prácticamente con las fluctuaciones de los parámetros.

Seguindo a Margaritis (1990), una forma de estimar recursivamente dichas varianzas es:

$$\sigma_{it}^2 = \sigma_{it-1}^2 + \frac{\bar{y}_{it-1}^2 - \sigma_{it-1}^2}{t} \quad (19)$$

$$Q_{it} = Q_{it-1} + \frac{\bar{\beta}_i \bar{\beta}_i^T - Q_{it-1}}{t} \quad (20)$$

siendo

$$\bar{y}_{it-1} = y_t - x_t^T \hat{\beta}_{it-1} \quad (21)$$

$$\bar{\beta}_i = \hat{\beta}_{it} - \hat{\beta}_{it-1} \quad (22)$$

Es decir, (21) define el error de predicción a horizonte un período de la variable  $y$ , y (22) la diferencia entre el vector de estimaciones de  $\beta$  en la etapa de actualización y predicción del filtro. Las ecuaciones (19) y (20) surgen de la discretización de las siguientes ecuaciones diferenciales [ver Söderström et al. (1978)]:

$$\frac{d\sigma_t^2}{dt} = M(\beta_t) - \sigma_t^2 \quad (23)$$

$$\frac{dQ_t}{dt} = V(\beta_t) - Q_t \quad (24)$$

donde

$$M(\beta_t) = E(\bar{y}_{it-1}^2) \quad ; \quad V(\beta_t) = E(\bar{\beta}_i \bar{\beta}_i^T) \quad (25)$$

Una simple observación de (19) y (20) muestra que ambas ecuaciones son independientes de la inicialización de la que se parta. De hecho, dado cualquier valor inicial  $\sigma_{0,t}^2$  y  $Q_{0,t}$ , cuando es procesado el primer dato, la estimación de la varianza residual del modelo es:

$$\sigma_{1/1}^2 = \bar{y}_{1/0}^2 = (y_1 - x_1^T \hat{\beta}_{1/0})^2$$

y una vez procesadas las  $N$  observaciones disponibles:

$$\sigma_{N/N}^2 = \frac{\sum_{t=1}^N \bar{y}_{it-1}^2}{N} \quad (26)$$

Es decir, (19) indica que un estimador de la varianza residual del modelo es la varianza muestral de los errores de predicción a horizonte un período de  $y_t$ . Una interpretación similar tiene la ecuación (20), que estima recursivamente la matriz de covarianzas ( $Q$ ) del vector de perturbaciones ( $u_t$ ) asociadas a la ley de evolución de los parámetros  $\beta$ . La expresión (26) es similar al estimador máximo-verosímil de la varianza residual en un modelo de parámetros fijos. Por tanto, dicho estimador será sesgado aunque consistente.

La idea fundamental del trabajo es incorporar la estimación recursiva de las varianzas  $\sigma^2$  y  $Q$  en el algoritmo de CRW, con objeto de combinar de forma óptima la secuencia de estimaciones de las varianzas obtenidas en los dos filtros. Con esta idea, el nuevo filtro (que llamamos CRW1) necesita también condiciones iniciales de las varianzas<sup>5</sup>:

$$\begin{aligned} H_{1/0} &= 0 & ; & & f_{1/0} &= 0 \\ \sigma_{1/0}^2 &= 1 & ; & & Q_{1/0} &= 0 \end{aligned} \quad (27)$$

La etapa de predicción del filtro que recorre la muestra hacia adelante es:

$$K_{t-1} = [I + H_{t-1/t-1} Q_{t-1/t-1}]^{-1} \quad (28)$$

<sup>5</sup> La condición inicial de  $\sigma^2$  puede ser arbitraria, pero distinta de cero si queremos propagar el filtro en función de las variables  $f$  y  $r$  y no en función de  $\sigma^2 f$  y  $\sigma^2 r$  [ver ecuaciones (31) y (32)].

$$H_{t-1} = K_{t-1} H_{t-1|t-1} \quad (29)$$

$$f_{t-1} = K_{t-1} f_{t-1|t-1} \quad (30)$$

y la etapa de actualización:

$$f_{t|t} = f_{t|t-1} + \frac{x_t y_t}{\sigma_{t|t-1}^2} \quad (31)$$

$$H_{t|t} = H_{t|t-1} + \frac{x_t x_t^T}{\sigma_{t|t-1}^2} \quad (32)$$

$$\sigma_{t|t}^2 = \sigma_{t|t-1}^2 + \frac{(y_t - x_t^T H_{t|t-1}^{-1} f_{t|t-1})^2 - \sigma_{t|t-1}^2}{t} \quad (33)$$

$$Q_{t|t} = Q_{t|t-1} + \frac{[(H_{t|t}^{-1} f_{t|t} - H_{t|t-1}^{-1} f_{t|t-1})(H_{t|t}^{-1} f_{t|t} - H_{t|t-1}^{-1} f_{t|t-1})^T] - Q_{t|t-1}}{t} \quad (34)$$

Las ecuaciones de predicción del filtro que recorre la muestra hacia atrás son ahora:

$$G_{NN+1} = 0 \quad ; \quad r_{NN+1} = 0 \quad (35)$$

$$\sigma_{NN+1}^2 = 1 \quad ; \quad Q_{NN+1} = 0$$

$$J_{t+1} = [I + G_{t+1|t+1} Q_{t+1|t+1}]^{-1} \quad (36)$$

$$G_{t+1} = J_{t+1} G_{t+1|t+1} \quad (37)$$

$$r_{t+1} = J_{t+1} r_{t+1|t+1} \quad (38)$$

y la fase de actualización:

$$r_{t|t} = r_{t|t+1} + \frac{x_t y_t}{\sigma_{t|t+1}^2} \quad (39)$$

$$G_{t|t} = G_{t|t+1} + \frac{x_t x_t^T}{\sigma_{t|t+1}^2} \quad (40)$$

$$\sigma_{t|t}^2 = \sigma_{t|t+1}^2 + \frac{(y_t - x_t^T G_{t|t+1}^{-1} r_{t|t+1})^2 - \sigma_{t|t+1}^2}{t} \quad (41)$$

$$Q_{t|t} = Q_{t|t+1} + \frac{[(G_{t|t}^{-1} r_{t|t} - G_{t|t+1}^{-1} r_{t|t+1})(G_{t|t}^{-1} r_{t|t} - G_{t|t+1}^{-1} r_{t|t+1})^T] - Q_{t|t+1}}{t} \quad (42)$$

El smoother óptimo se obtiene ponderando las secuencias de estimaciones tanto de los parámetros  $\beta$ :

$$P_{t|N} = [H_{t|t} + G_{t|t+1}]^{-1} \quad (43)$$

$$b_{t|N} = P_{t|N} [f_{t|t} + r_{t|t+1}] \quad (44)$$

como de las varianzas  $\sigma^2$  y  $Q$ :

$$\sigma_{t|N}^2 = \frac{(x_t^T H_{t|t}^{-1} x_t) \sigma_{t|t}^2 + (x_t^T G_{t|t+1}^{-1} x_t) \sigma_{t|t+1}^2}{x_t^T H_{t|t}^{-1} x_t + x_t^T G_{t|t+1}^{-1} x_t} \quad (45)$$

$$Q_{t|N} = \frac{H_{t|t}^{-1} Q_{t|t} + G_{t|t+1}^{-1} Q_{t|t+1}}{\text{tr}(H_{t|t}^{-1} + G_{t|t+1}^{-1})} \quad (46)$$

Las ecuaciones (45) y (46) muestran que la ponderación dada a la estimación de las varianzas  $\sigma^2$  y  $Q$  viene dada por la incertidumbre asociada a dicha estimación en cada uno de los dos filtros.

### 3. Resultados con datos simulados.

Para validar el algoritmo descrito en el apartado anterior, se ha aplicado a la estimación de un modelo de regresión con un parámetro cambiante usando datos simulados. El experimento ha consistido en estimar para distintos tamaños muestrales la siguiente especificación:

$$y_t = \alpha_t + \beta x_t + \epsilon_t \quad (47)$$

$$\alpha_t = \phi \alpha_{t-1} + u_t \quad (48)$$

donde  $\beta = 0.5$ , el parámetro autorregresivo  $\phi$  toma los valores 1, 0.95 y 0.5, el valor inicial de  $\alpha_t$  es cero y las perturbaciones  $\epsilon_t$  y  $u_t$  se han generado como variables aleatorias independientes e idénticamente distribuidas como normales con los siguientes momentos:

$$\begin{aligned} \epsilon_t &\sim N(0, 9) \\ u_t &\sim N(0, 1) \end{aligned} \quad (49)$$

La variable explicativa  $x_t$  también ha sido generada como una variable aleatoria independiente de  $\epsilon_t$  y con distribución  $N(0, 25)$ .

Por tanto, se han simulado tres modelos de acuerdo con la especificación dada por (47) y (48). El modelo I ( $\phi = 1$ ) es el modelo con constante adaptativa de Cooley y Prescott (1973). El modelo II supone que la ley de variación del parámetro está cerca de la no estacionariedad ( $\phi = 0.95$ ) y en el modelo III, la constante sigue un proceso AR(1) bien condicionado al fijar  $\phi = 0.5$ . La razón de generar modelos donde el parámetro cambiante sigue una ley de evolución estacionaria y no estacionaria, es comprobar si este supuesto afecta a la estimación puntual del parámetro constante  $\beta$ , de la varianza residual del modelo, de la varianza de la perturbación asociada al parámetro cambiante, y sobre todo, a la precisión con la que son estimados dichos parámetros.

Los modelos I, II y III han sido estimados utilizando tres procedimientos: (1) el algoritmo CRW, fijando el valor teórico de las varianzas  $\sigma_\epsilon^2$  y  $\sigma_u^2$ ; (2) el algoritmo CRW1 y (3) el algoritmo FK-SIF eligiendo las siguientes condiciones iniciales para el vector de parámetros y su correspondiente matriz de covarianzas:  $\alpha_{1|0} = 0$ ,  $\beta_{1|0} = 0$  y  $P_{1|0} = \tau I$  con  $\tau = 10^6$ . La razón de utilizar este último algoritmo es mostrar la importancia de las condiciones iniciales de los parámetros  $\alpha$  y  $\beta$  en la estimación recursiva de las varianzas  $\sigma_\epsilon^2$  y  $\sigma_u^2$ . De hecho, a medida que  $\tau$  tiende a ser un número "muy" grande, la estimación

final de las varianzas se acerca más al verdadero valor, pero el filtro se degrada numéricamente<sup>6</sup>. Este hecho muestra que es necesario utilizar un *smoother* de tipo información para hacer independientes los resultados de la inicialización.

Los resultados de este primer experimento se muestran en las Tablas 1, 2 y 3. Las estimaciones que se muestran en las tablas son la media de los resultados obtenidos con 100 realizaciones. Los errores estándar de las estimaciones del parámetro  $\beta$ , que se presentan entre paréntesis, son el valor medio de las desviaciones típicas obtenidas con 100 realizaciones. Los errores estándar de las varianzas  $\sigma_\epsilon^2$  y  $\sigma_u^2$  se han calculado a partir de las varianzas muestrales.

(Insertar Tablas 1, 2 y 3)

A la vista de los resultados, puede concluirse que, en general, para las tres especificaciones consideradas, el valor medio del parámetro  $\beta$  es muy próximo a su verdadero valor con los tres procedimientos utilizados. Sin embargo, la precisión con la que se estima dicho parámetro es mayor con el criterio CRW1 que con FK-SIF, comparada con la desviación típica del parámetro obtenida a partir de CRW, que supone conocidas las varianzas  $\sigma_\epsilon^2$  y  $\sigma_u^2$ .

Además, con el algoritmo CRW1 y en cualquiera de los modelos considerados, el valor medio de las varianzas residuales estimadas se acerca más a su correspondiente valor teórico que cuando se utiliza el algoritmo FK-SIF. Por otro lado, hay que señalar que el procedimiento FK-SIF tiende a sobrestimar la varianza residual asociada al modelo en las tres especificaciones, aumentando el sesgo a medida que decrece el tamaño muestral. Sin embargo, cuando la estructura estocástica del parámetro cambiante es más estacionaria (ver Tabla 3), parece que este último criterio aproxima mejor la varianza residual del modelo.

Con respecto a la varianza asociada a la perturbación  $u_t$ , el algoritmo CRW1 tiende a sobrestimar su valor teórico, aunque dicho sesgo disminuye al aumentar el tamaño muestral y al hacerse más estacionario el proceso AR(1) que sigue el parámetro cambiante. Por el contrario, el criterio FK-SIF tiende a infraestimar dicho parámetro, aumentando el sesgo a medida que crece el tamaño muestral y que el parámetro autorregresivo se aleja de la no estacionariedad.

<sup>6</sup> Se ha elegido un valor de  $\tau = 10^6$ , porque es el habitualmente usado en los algoritmos de estimación recursiva de modelos con parámetros fijos [ver Young (1984)].



En la Tabla 4 se presenta el valor medio de los ratios estimados  $\hat{\sigma}_t^2 / \sigma_t^2$  al usar los procedimientos CRW1 y FK-SIF, en los tres modelos considerados. La conclusión es que el algoritmo CRW1 estima un ratio de varianzas más próximo al valor teórico (1/9) que el criterio FK-SIF. Esto asegura que si reformulamos los dos algoritmos en términos de los ratios de varianzas, también se compararían favorablemente los resultados obtenidos con el criterio CRW1 con respecto a FK-SIF.

(Insertar Tabla 4)

En las Figuras 1 y 2 se presenta la evolución media de la estimación del parámetro  $\alpha_t$  del modelo I, utilizando los algoritmos FK-SIF, CRW y CRW1, para un tamaño muestral de 100 y 200 observaciones, respectivamente<sup>7</sup>. Se observa que la evolución del parámetro cambiante cuando se estima por el procedimiento CRW1 se ajusta mejor a su evolución teórica obtenida mediante el algoritmo CRW. En general, siempre se observa que el perfil del parámetro cuando se estima con FK-SIF es bastante más suave que el obtenido con los otros dos procedimientos. Este resultado es lógico dado que el criterio FK-SIF tiende a sobrestimar la varianza residual del modelo y a la vez, infraestima la varianza correspondiente al parámetro cambiante.

(Insertar Figuras 1 y 2)

Tabla 1: Resultados de la simulación del modelo I:  $y_t = \alpha_t + \beta x_t + \epsilon_t$  con  $\alpha_t = \alpha_{t-1} + u_t$ . Valores teóricos de los parámetros  $\beta = 0.5$ ,  $\sigma_\epsilon^2 = 9$ ,  $\sigma_u^2 = 1$ .

N	Algoritmo CRW		Algoritmo CRW1		Algoritmo FK-SIF con $\tau = 10^6$		
	$\hat{\beta}$	$\hat{\sigma}_\epsilon^2$	$\hat{\beta}$	$\hat{\sigma}_\epsilon^2$	$\hat{\beta}$	$\hat{\sigma}_\epsilon^2$	$\hat{\sigma}_u^2$
100	0.4949 (0.0414)	8.7205 (2.0090)*	0.4993 (0.0517)	1.4245 (0.7685)*	0.4947 (0.0568)	24.0735 (13.5127)*	1.5836 (1.4675)*
200	0.5042 (0.0314)	8.1708 (1.3923)*	0.5067 (0.0384)	1.2309 (0.5057)*	0.5049 (0.0499)	19.2893 (8.8741)*	1.0734 (0.7519)*
1000	0.5089 (0.0138)	9.3823 (1.2079)*	0.4988 (0.0154)	1.2040 (0.6098)*	0.5029 (0.0235)	14.7103 (3.6885)*	0.5879 (0.2792)*

Notas: N: Tamaño muestral utilizado.

(\*) Las cifras entre paréntesis con un \* representan las desviaciones típicas muestrales.  
 Las cifras entre paréntesis sin un \* representan el valor medio de las desviaciones típicas estimadas.

<sup>7</sup> No se ofrecen los gráficos de evolución temporal de dicho parámetro resultante de otras especificaciones, ya que las conclusiones no cambian con respecto a las obtenidas de los gráficos presentados.

Tabla 2: Resultados de la simulación del modelo II:  $y_t = \alpha_t + \beta x_t + \epsilon_t$  con  $\alpha_t = 0.95\alpha_{t-1} + u_t$ .  
Valores teóricos de los parámetros  $\beta = 0.5$ ,  $\sigma_\epsilon^2 = 9$ ,  $\sigma_u^2 = 1$ .

N	Algoritmo CRW	Algoritmo CRW1		Algoritmo FK-SIF con $\tau = 10^6$			
	$\hat{\beta}$	$\hat{\beta}$	$\hat{\sigma}_\epsilon^2$	$\hat{\sigma}_u^2$	$\hat{\beta}$	$\hat{\sigma}_\epsilon^2$	$\hat{\sigma}_u^2$
100	0.4905 (0.0468)	0.5007 (0.0605)	14.6771 (6.6303)*	1.1097 (2.4454)*	0.4901 (0.0641)	17.0793 (4.5690)*	0.2318 (0.3338)*
200	0.4991 (0.0330)	0.5067 (0.0419)	14.3433 (3.5061)*	1.2865 (2.2947)*	0.4980 (0.0459)	17.8434 (4.4287)*	0.1351 (0.1618)*
1000	0.5047 (0.0138)	0.5027 (0.0178)	13.1920 (1.3134)*	0.8630 (1.5133)*	0.5045 (0.0206)	19.6167 (3.1481)*	0.0153 (0.0327)*

Notas: N: Tamaño muestral utilizado.

(\*) Las cifras entre paréntesis con un \* representan las desviaciones típicas muestrales.

Las cifras entre paréntesis sin un \* son el valor medio de las desviaciones típicas estimadas.

Tabla 3: Resultados de la simulación del modelo III:  $y_t = \alpha_t + \beta x_t + \epsilon_t$  con  $\alpha_t = 0.50\alpha_{t-1} + u_t$ .  
Valores teóricos de los parámetros  $\beta = 0.5$ ,  $\sigma_\epsilon^2 = 9$ ,  $\sigma_u^2 = 1$ .

N	Algoritmo CRW	Algoritmo CRW1		Algoritmo FK-SIF con $\tau = 10^6$			
	$\hat{\beta}$	$\hat{\beta}$	$\hat{\sigma}_\epsilon^2$	$\hat{\sigma}_u^2$	$\hat{\beta}$	$\hat{\sigma}_\epsilon^2$	$\hat{\sigma}_u^2$
100	0.4907 (0.0469)	0.4985 (0.0469)	11.2499 (2.4502)*	2.2184 (5.1327)*	0.4907 (0.0539)	12.0256 (2.4302)*	0.0256 (0.0321)*
200	0.4991 (0.0329)	0.5067 (0.0339)	10.8407 (1.3882)*	1.3476 (1.2828)*	0.4993 (0.0369)	11.4698 (1.5142)*	0.0106 (0.0164)*
1000	0.5040 (0.0138)	0.5027 (0.0158)	10.6426 (0.7092)*	1.0002 (0.0951)*	0.5040 (0.0156)	11.0129 (0.7099)*	0.0005 (0.0007)*

Notas: N: Tamaño muestral utilizado

(\*) Las cifras entre paréntesis con un \* representan las desviaciones típicas muestrales.

Las cifras entre paréntesis sin un \* representan el valor medio de las desviaciones típicas estimadas.

Tabla 4: Valor medio de los ratios de varianzas  $\sigma_{\alpha}^2/\sigma_{\alpha}^2$  estimados con CRW1 y FK-SIF.

N	Modelo I $\alpha = 0.5$		Modelo II $\alpha = 0.95$		Modelo III $\alpha = 1.0$	
	CRW1	FK-SIF	CRW1	FK-SIF	CRW1	FK-SIF
100	0.2051	0.0020	0.0781	0.0151	0.1283	0.0399
200	0.1242	0.0008	0.0921	0.0084	0.1507	0.5565
1000	0.0944	$4.22 \times 10^{-5}$	0.0653	0.0009	0.1634	0.6578

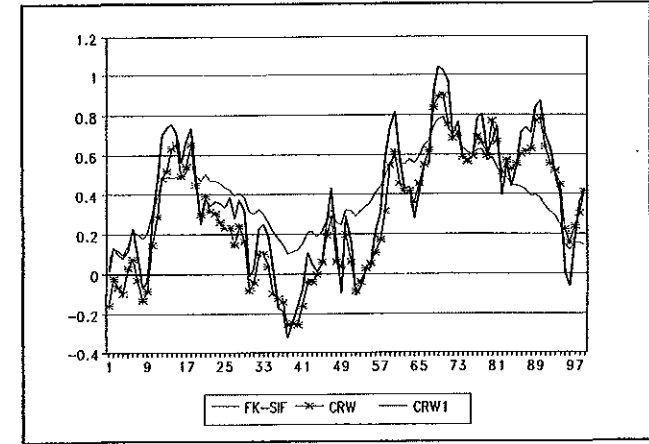


Figura 1: Evolución media de la estimación de  $\alpha$ , obtenida con los criterios FK-SIF, CRW y CRW1 para N = 100.

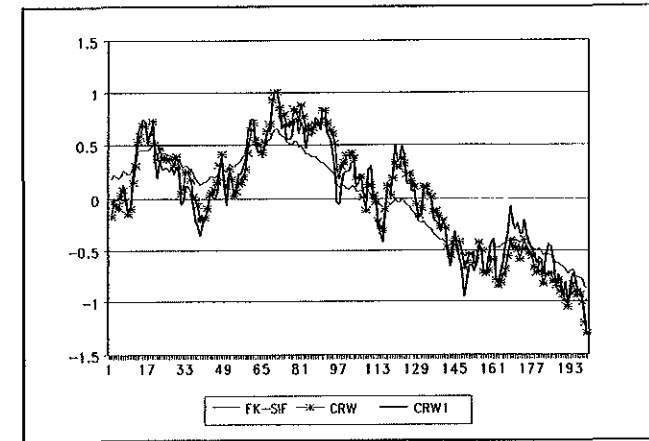


Figura 2: Evolución media de la estimación de  $\alpha$ , obtenida con los criterios FK-SIF, CRW y CRW1 para N = 200.

#### 4. Conclusiones.

La estimación recursiva de modelos de parámetros cambiantes mediante el criterio habitual en dos etapas, consistente en propagar un filtro de Kalman y después un *smoother* de intervalo fijo (denotado por FK-SIF), exige conocer o fijar las varianzas de todos los términos de error presentes en el modelo. Dado que los resultados de este criterio son sensibles a la inicialización del mismo, es preferible utilizar un algoritmo como el de Cooley, Rosenberg y Wall (1977) (denotado por CRW), independiente de condiciones iniciales. Sin embargo, este algoritmo también supone conocidas las varianzas de las perturbaciones existentes. Este supuesto no es habitual en la práctica econométrica con datos reales, y en este trabajo se muestra que la elección arbitraria de dichas varianzas, o de los ratios entre ellas, afecta tanto a la eficiencia de los estimadores del resto de parámetros, como a la verdadera secuencia de estimaciones.

En este trabajo, se ha derivado una versión del algoritmo CRW (1977) que permite obtener una estimación recursiva de las varianzas relevantes de un modelo de parámetros cambiantes, además de la secuencia óptima de estimaciones de los parámetros propios del modelo. Además, esta nueva versión del filtro CRW es independiente de las condiciones iniciales de todos los parámetros, considerando también como parámetros las varianzas de las distintas perturbaciones.

Al igual que el CRW, este nuevo algoritmo está basado en la utilización de dos filtros de información que se propagan de forma independiente entre sí, pudiendo correr en paralelo para después combinar de forma óptima los resultados de ambos. Este hecho supone una ventaja computacional con respecto a otros algoritmos de *smoothing* más clásicos, como es el criterio FK-SIF.

Se ha aplicado este procedimiento a la estimación de modelos de regresión con un único parámetro cambiante que sigue o un proceso autorregresivo o un paseo aleatorio. Los primeros resultados obtenidos con este nuevo criterio usando datos simulados, indican el buen funcionamiento del mismo para distintos tamaños muestrales. En concreto, se han comparado los resultados de los procedimientos FK-SIF y CRW1 con los obtenidos con CRW. En general, para todas las especificaciones consideradas, el criterio CRW1 aproxima mejor la varianza residual del modelo y la varianza asociada al parámetro cambiante que el algoritmo FK-SIF, generando una evolución temporal del parámetro cambiante prácticamente indistinguible de la verdadera (resultante del procedimiento CRW). El criterio FK-SIF, en general, sobrestima la varianza residual e infraestima la varianza asociada al coeficiente cambiante, generando, por tanto, una secuencia temporal

del mismo excesivamente suave con respecto a su evolución teórica. Por último, se observa que una estimación inadecuada de dichas varianzas hace aumentar la desviación típica estimada del parámetro constante, teniendo como punto de referencia el error estándar obtenido con el criterio CRW.

Dada la relación existente entre modelos de parámetros cambiantes y los modelos ARCH [ver Wolff (1988)], una posible extensión del trabajo es investigar qué tipo de modelos del primer tipo generan las estructuras ARCH más habituales en la práctica econométrica con datos económicos reales. Los contrastes desarrollados de heterocedasticidad condicional podrían ser un método para validar la utilización del modelo de parámetros aleatorios asociado. Por supuesto, el planteamiento básico es investigar cuál de estas dos formulaciones ofrece más información en la etapa de estimación del modelo y cuál es preferible si el objetivo es predicción.

## Referencias

- Alexander, G.J. y P.G. Benson (1982). "More on beta as a random coefficient". *Journal of Financial and Quantitative Analysis*, vol. XVII, n°1, pag. 27-36.
- Anderson, B.D.O. y J.B. Moore (1979). *Optimal filtering*. Prentice-Hall, Inc., New Jersey.
- Cooley, T.F. y E.C. Prescott (1973). "An adaptative regression model". *International Economic Review* 14, pag. 364-371.
- Cooley, T.F. y E.C. Prescott (1976). "Estimation in the presence of stochastic parameter variation". *Econometrica*, vol. 44, n°1, pag. 167-184.
- Cooley, T.F., B. Rosenberg y K.D. Wall (1977). "A Note on Optimal Smoothing for Time Varying Coefficient Problems". *Annals of Economic and Social Measurement* 6, n°4, pag. 453-456.
- Fabozzi, F.J. y J.C. Francis (1978). "Beta as a random coefficient". *Journal of Financial and Quantitative Analysis*, pag. 101-116.
- García-Ferrer, A. J. del Hoyo, A. Novales y P.C. Young (1993). "Further Evidence on Forecasting International GNP Growth Rates Using Unobserved Components Transfer Function Models". *Documento de Trabajo ICAE*, n° 9312.
- Harvey, A.C. (1989). *Forecasting structural time series models and the Kalman filter*. Cambridge University Press.
- Liebelt, P. (1967). *An Introduction to Optimal Estimation*. Reading: Mass.: Addison-Wesley.
- Margaritis D. (1990). "A Time-Varying Model of Rational Learning". *Economics Letters* 33, pag. 309-314.
- Pagan, A.R. (1980). "Some identification and estimation results for regression models with stochastically varying coefficients". *Journal of Econometrics* 13, pag. 341-363.
- Rosenberg, B.M. (1973a). "A Survey of Stochastic Parameter Regression". *Annals of Economic and Social Measurement* 3, n°4, pag. 381-397.
- Rosenberg, B.M. (1973b). "The Analysis of a Cross Section of Time Series by Stochastically Convergent Parameter Regression". *Annals of Economic and Social Measurement* 3, n°4, pag. 399-428.
- Söderström T., L. Ljung y I. Gustavsson (1978). "A theoretical analysis of recursive identification methods". *Automatica*, 14, pag. 321-344.
- Sotoca, S. (1992). "El problema de las condiciones iniciales en los algoritmos de estimación recursiva de modelos lineales". *Estadística Española*, vol. 35, n° 132, pag. 89-115.
- Wells, C. (1994). "Variable betas on the Stockholm exchange 1971-1989". *Applied Economics*, 4, pag. 75-92.
- Wolff, C. (1988). "Autoregressive conditional heteroscedasticity. A comparison of ARCH and Random Coefficient Models". *Economic Letters*, 27, pag. 141-143.
- Young, P. (1984). *Recursive estimation and time-series analysis. An introduction*. Springer-Verlag, Heidelberg.

## APENDICE:

El algoritmo FK-SIF consta de dos etapas. En la primera, se propaga el filtro de Kalman (FK) para el modelo (1)-(2) empezando en  $t = 1, 2, \dots, N$  y con condiciones iniciales arbitrarias:

$$\hat{\beta}_{1/0} = 0 \quad ; \quad P_{1/0} = \tau I \quad \tau \gg 0$$

## Etapa de predicción:

$$\hat{\beta}_{t/t-1} = \hat{\beta}_{t-1/t-1} \quad (A.1)$$

$$P_{t/t-1} = P_{t-1/t-1} + Q_{t-1/t-1} \quad (A.2)$$

## Etapa de actualización:

$$K_t = P_{t/t-1} x_t (x_t^T P_{t/t-1} x_t + \hat{\sigma}_{t-1}^2)^{-1} \quad (A.3)$$

$$\bar{z}_t = y_t - x_t^T \hat{\beta}_{t/t-1} \quad (A.4)$$

$$\hat{\beta}_{t/t} = \hat{\beta}_{t/t-1} + K_t \bar{z}_t \quad (A.5)$$

$$P_{t/t} = (I - K_t x_t^T) P_{t/t-1} \quad (A.6)$$

$$\hat{\sigma}_t^2 = \hat{\sigma}_{t-1}^2 + \frac{(\bar{z}_t^2 - \hat{\sigma}_{t-1}^2)}{t} \quad (A.7)$$

$$Q_{t/t} = Q_{t-1/t-1} + \frac{\{(K_t \bar{z}_t)(K_t \bar{z}_t)^T - Q_{t-1/t-1}\}}{t} \quad (A.8)$$

En la segunda etapa, se propaga un smoother de intervalo fijo empezando en el instante  $t = N-1, N-2, \dots, 1$ , con condiciones iniciales:

$$\hat{\beta}_{N/N} \quad ; \quad P_{N/N}$$

$$\hat{\beta}_{t/t} = \hat{\beta}_{t/t} + A_t (\hat{\beta}_{t+1/t} - \hat{\beta}_{t/t}) \quad (A.9)$$

$$P_{MN} = P_{Ml} + A_l(P_{l,MN} - P_{l,MH})A_l^T \quad (\text{A.10})$$

$$A_l = P_{Ml}P_{l,MH}^{-1} \quad (\text{A.11})$$

## SERIE DE DOCUMENTOS DE TRABAJO DEL ICAE

- 9301 "Análisis del Comportamiento de las Cotizaciones Reales en la Bolsa de Madrid bajo la Hipótesis de Eficiencia". Rafael Flores de Frutos. Diciembre 1992. (Versión final aceptada para publicación en Estadística Española)
- 9302 "Sobre la Estimación de Primas por Plazo dentro de la Estructura Temporal de Tipos de Interés". Rafael Flores de Frutos. Diciembre 1992.
- 9303 "Cambios de Estructuras de Gasto y de Consumo en el Cálculo del IPC". Antonio Abadía. Febrero 1993. (Versión final publicada en Revista de Economía Aplicada, Vol.1, N°1)
- 9304 "Tax Analysis in a Limit Pricing Model". Félix Marcos. Febrero 1993.
- 9305 "El Tipo de Cambio Propio: Reformulación del Concepto y Estimación para el Caso Español". José de Hevia Payá. Junio 1993. (Versión final aceptada para publicación en Revista Española de Economía)
- 9306 "Price Volatility Under Alternative Monetary Instruments". Alfonso Novales. Abril 1992.
- 9307 "Teorías del Tipo de Cambio: Una Panorámica". Oscar Bajo Rubio. Simón Sosvilla Rivero. Junio 1993. (Versión final publicada en Revista de Economía Aplicada, Vol.1, N°2).
- 9308 "Testing Theories of Economic Fluctuations and Growth in Early Development (the case of the Chesapeake tobacco economy)". Rafael Flores de Frutos. Alfredo M. Pereira. Diciembre 1992.
- 9309 "Maastricht Convergence Conditions: A Lower Bound for Inflation?". Jorge Blázquez. Miguel Sebastián. Marzo 1992.
- 9310 "Recursive Identification, Estimation and Forecasting of Nonstationary Economic Time Series with Applications to GNP International Data". A. García-Ferrer. J. del Hoyo. A. Novales. P.C. Young. Marzo 1993.
- 9311 "General Dynamics in Overlapping Generations Models". Carmen Carrera. Manuel Morán. Enero 1993. (Versión final aceptada para publicación en Journal of Economic Dynamics and Control)
- 9312 "Further Evidence on Forecasting International GNP Growth Rates Using Unobserved Components Transfer Function Models". A. García-Ferrer, J. del Hoyo, A. Novales, P.C. Young. Septiembre 1993. (De próxima aparición en un volumen de homenaje a A. Zellner)
- 9313 "Public Capital and Aggregate Growth in the United States: Is Public Capital Productive?". Rafael Flores de Frutos. Alfredo M. Pereira. Julio 1993.
- 9314 "Central Bank Structure and Monetary Policy Uncertainty". José I. García de Paso. Abril 1993.
- 9315 "Monetary Policy with Private Information: A Role for Monetary Targets". José I. García de Paso. Julio 1993.
- 9316 "Exact Maximum Likelihood Estimation of Stationary Vector ARMA Models". José Alberto Mauricio. Julio 1993. (Versión final aceptada para publicación en Journal of the American Statistical Association)
- 9317 "The Exact Likelihood Function of a Vector ARMA Model". José Alberto Mauricio. Julio 1993.
- 9318 "Business Telephone Traffic Demand in Spain: 1980-1991, An Econometric Approach". Teodosio Pérez Amaral. Francisco Alvarez González. Bernardo Moreno Jiménez. Septiembre 1993. (Versión final aceptada para publicación en Information Economics and Policy)

BIBLIOTECA UCM



5306520295

- 9401 "Contrastes de momentos y de la matriz de información". Teodosio Pérez Amaral. Junio 1994. (Versión final aceptada para publicación en Cuadernos Económicos del ICE)
- 9402 "A partisan explanation of political monetary cycles". José I. García de Paso. Junio 1994.
- 9403 "Estadísticos para la detección de observaciones anómalas en modelos de elección binaria: Una aplicación con datos reales". Gregorio R. Serrano. Junio 1994.
- 9404 "Effects of public investment in infrastructure on the spanish economy". Rafael Flores de Frutos. Mercedes Gracia Díez. Teodosio Pérez Amaral. Junio 1994.
- 9405 " Observaciones anómalas en modelos de elección binaria". Mercedes Gracia. Gregorio R. Serrano. Junio 1994. (Versión final aceptada para publicación en Estadística Española)
- 9406 " Permanent components in seasonal variables". Rafael Flores. Alfonso Novales. Junio 1994.
- 9407 " Money demand instability and the performance of the monetary model of exchange rates". Rodrigo Peruga. Junio 1994.
- 9408 " Una nota sobre la estimación eficiente de modelos con parámetros cambiantes". Sonia Sotoca. Junio 1994.
- 9409 " Distribución de la renta y redistribución a través del IRPF en España". Rafael Salas. Junio 1994.
- 9410 " Trade balances: Do exchange rates matter?". Rodrigo Peruga. Junio 1994.