

UN ANÁLISIS NO PARAMÉTRICO DE LOS TIPOS DE INTERÉS A CORTO PLAZO EN ESPAÑA

José Manuel Gavilán Ruiz José María Alba Riesco
Universidad de Sevilla Universidad de Sevilla
e-mail: gavi@us.es e-mail: jmalbar@us.es

28 de febrero de 2003

Resumen

En este trabajo se modeliza la evolución temporal de los tipos de interés a corto plazo en España mediante una difusión homogénea unifactorial. Las estimaciones son llevadas a cabo en el contexto no paramétrico propuesto en Stanton (1997)[29] y usando datos diarios de los tipos de interés de los depósitos interbancarios desde enero de 1988 hasta septiembre de 2002. La deriva estimada exhibe un comportamiento no lineal como el encontrado en Stanton (1997)[29] para la economía estadounidense y en Corzo y Gómez (2001)[11] para la economía española, concluyendo estos autores con una evidencia de no linealidad en la deriva. Sin embargo a la luz del artículo de Chapman y Pearson (2000)[9] esta conclusión parece algo precipitada. Con objeto de profundizar en este tema, se contrasta la linealidad de la deriva usando es test de Aït-Sahalia, Bickel, y Stoker (2001)[2] y se encuentra que no es posible rechazar la linealidad de la deriva. En cuanto a la volatilidad estimada, parece estar en consonancia con la propuesta en Chan *et al.* (1992)[8].

Palabras clave: *Estructura temporal de los tipos de interés, modelos unifactoriales, estimación no paramétrica, regresión kernel, test de bondad de ajuste.*

Clasificación JEL: C14, C52, E43

1. Introducción

El disponer de buenas estimaciones sobre la evolución temporal de los tipos de interés es de una enorme importancia en Economía, siendo de gran interés en temas relacionados con la inflación, la política monetaria, la formación de expectativas, las relaciones entre los tipos de interés a corto y largo plazo,... En el campo de la Economía Financiera permite la valoración correcta de activos financieros dependientes del tipo de interés y sus derivados, lo que lleva a un conveniente diseño de estrategias de inversión y cobertura.

En la moderna teoría de la valoración de activos es usual representar la evolución temporal de las variables subyacentes como un proceso estocástico que evoluciona en tiempo continuo. En este contexto, uno de los temas que ha recibido una mayor atención por parte de los investigadores ha sido la modelización de la dinámica de la Estructura Temporal de los Tipos de Interés (ETTI en adelante), donde el objetivo fundamental es el estudio del proceso de los tipos de contado $\{r_t\}_{t \geq 0}$, los cuales representan el tipo instantáneo de retorno libre de riesgo en cualquier instante.

Se han propuesto una gran cantidad de modelos diferentes para la modelización de tal proceso estocástico. Sin embargo, es poco el acuerdo existente sobre cual de ellos es el más natural utilizar. La mayoría de tales modelos se han propuesto en un contexto paramétrico atendiendo quizás más a razones de tratabilidad analítica y computacional que a razones de tipo económico. Algunos trabajos recientes usan técnicas no paramétricas con el objetivo de evitar una posible mala especificación del modelo impuesto. Entre ellos, el trabajo de Stanton (1997)[29] propone una modelización completamente no paramétrica de la difusión unifactorial homogénea que gobierna los movimientos del tipo de interés a corto plazo libre de riesgo. Las diferentes aproximaciones obtenidas a la deriva y la volatilidad de la difusión se comprueba que son muy buenas cuando se estiman, a modo de comprobación, los modelos de Cox, Ingersoll y Ross (1985)[13] y una versión de Black, Derman y Toy (1990)[3], los cuales son muy usados en la práctica.

El resto de este trabajo se organiza de la siguiente manera: En la Sección 2 se plantea el

contexto no paramétrico desarrollado en Stanton (1997)[29], obteniendo las aproximaciones a la deriva y la volatilidad de los tipos a corto plazo. En la Sección 3 se describen y se hace un primer análisis de los datos que se van a usar para llevar a cabo la estimación de la difusión propuesta para modelizar los movimientos de los tipos de interés a corto plazo. En la Sección 4 se lleva a cabo la estimación no paramétrica de tal difusión. La Sección 5 se dedica a contrastar la posible no linealidad mostrada en las estimaciones de la deriva, y finalmente en la Sección 6 se exponen las conclusiones obtenidas.

2. Estimación no paramétrica de modelos de difusión

En el marco de los modelos unifactoriales para la modelización de la dinámica de los tipos de interés es usual proponer la siguiente difusión homogénea para representar la dinámica de los tipos de interés a corto plazo libre de riesgo,

$$dr_t = \mu(r_t) dt + \sigma(r_t) dW_t \quad (1)$$

donde $\{W_t\}$ es un movimiento Browniano estándar, y la deriva $\mu(\cdot)$ y la volatilidad $\sigma(\cdot)$ son funciones que dependen únicamente del valor actual del tipo de interés. Algunos casos particulares incluidos en esta formulación general aparecen en el cuadro 1. Una breve descripción de estos modelos y de su comportamiento empírico se puede encontrar en Moreno (2000)[26].

Sin embargo, otros modelos similarmente planteados son no homogéneos en el tiempo y no caerían dentro de la clase de modelos planteados en la ecuación (1), algunos de ellos se muestran en el cuadro 2. Una breve descripción de estos modelos y de su comportamiento empírico se puede encontrar en Moreno (2000)[26].

La principal ventaja de estos modelos no homogéneos en el tiempo frente a los anteriores homogéneos es que al permitir que su deriva y/o volatilidad dependan del tiempo se pueden ajustar perfectamente a la estructura actual observada en los tipos de interés. Sin embargo,

Merton(1973)[25]	$dr_t = \alpha_0 dt + \beta_0 dW_t$
Black y Scholes(1973)[5]	$dr_t = \alpha_1 r_t dt + \beta_1 r_t dW_t$
Cox(1975)[12]	$dr_t = \alpha_1 r_t dt + \beta_1 r_t^\tau dW_t$
Vasicek(1977)[30]	$dr_t = (\alpha_0 + \alpha_1 r_t) dt + \beta dW_t$
Dothan(1978)[15]	$dr_t = \beta_1 r_t dW_t$
Brennan y Schwartz(1980)[6]	$dr_t = (\alpha_0 + \alpha_1 r_t) dt + \beta_1 r_t dW_t$
Cox, Ingersoll y Ross(1985)[13, 14]	$dr_t = (\alpha_0 + \alpha_1 r_t) dt + \beta_1 \sqrt{r_t} dW_t$
Constantinides(1992)[10]	$dr_t = (\alpha_0 + \alpha_1 r_t + \alpha_2 \sqrt{r_t - \alpha_3}) dt + \beta_0 (r_t - \beta_1) dW_t$
Chan <i>et al.</i> (1992)[8]	$dr_t = (\alpha_0 + \alpha_1 r_t) dt + \beta_2 r_t^\tau dW_t$
Duffie y Kan(1996)[16]	$dr_t = (\alpha_0 + \alpha_1 r_t) dt + \sqrt{\beta_0 + \beta_1 r_t} dW_t$
Ait-Sahalia(1996)[1]	$dr_t = (\alpha_0 + \alpha_1 r_t + \alpha_2 r_t^2 + \alpha_3 r_t^{-1}) dt + (\beta_0 + \beta_1 r_t + \beta_2 r_t^\tau) dW_t$

Cuadro 1: Algunos modelos homogéneos unifactoriales paramétricos.

Versión continua de Ho y Lee (1986)[19]	$dr_t = \alpha(t) dt + \beta dW_t$
Black, Derman y Toy (1990)[3]	$d \ln r_t = \left[\alpha(t) - \frac{\beta'(t)}{\beta(t)} \right] dt + \beta(t) dW_t$
Black y Karasinski (1991)[4]	$d \ln r_t = \alpha_2(t) [\ln \alpha_1(t) - \ln r_t] dt + \beta(t) dW_t$
Hull y White (1990)[21]	$dr_t = \alpha_2(t) [\alpha_1(t) - r_t] dt + \beta(t) \sqrt{r_t} dW_t$

Cuadro 2: Algunos modelos no homogéneos unifactoriales paramétricos.

para mantener este ajuste perfecto, estos modelos deberían ser replanteados y reestimados en cada instante de tiempo, lo que llevaría a un enorme esfuerzo computacional.

Un problema potencialmente grave de cualquier modelo paramétrico es que el imponer una determinada estructura puede llevar a **errores de especificación** y en consecuencia se tendrían errores graves al usar el modelo con propósitos de valoración y cobertura de activos de tipos de interés (ver Canabarro (1995)[7] para mayores detalles sobre este punto). En un intento por evitar tales posibles errores de especificación, Stanton (1997)[29] propone un método de estimación completamente no paramétrico, el cual se puede estimar también de forma paramétrica con, por ejemplo, fines comparativos.

La metodología propuesta se basa en desarrollar la esperanza condicionada $E_t [f(r_{t+\Delta}, t + \Delta)]$ al modo de un desarrollo en serie de Taylor, siendo $f(\cdot, \cdot)$ una función arbitraria y Δ el intervalo de tiempo entre observaciones sucesivas. Bajo determinadas condiciones sobre $\mu(\cdot)$, $\sigma(\cdot)$ y $f(\cdot, \cdot)$, ver por ejemplo Hille y Phillips (1957)[18], se tiene que

$$E_t [f(r_{t+\Delta}, t + \Delta)] = f(r_t, t) + \mathcal{L}f(r_t, t) \cdot \Delta + \frac{1}{2} \mathcal{L}^2 f(r_t, t) \cdot \Delta^2 + \dots + \frac{1}{n!} \mathcal{L}^n f(r_t, t) \cdot \Delta^n + O(\Delta^{n+1}), \quad (2)$$

siendo \mathcal{L} el generador infinitesimal del proceso $\{r_t\}$, el cual viene dado por

$$\begin{aligned} \mathcal{L}f(r_t, t) &= \lim_{\tau \downarrow t} \frac{E[f(r_\tau, \tau)/r_t=r] - f(r, t)}{\tau - t} = \\ &= \frac{\partial f(r, t)}{\partial t} + \frac{\partial f(r, t)}{\partial r} \mu(r) + \frac{1}{2} \frac{\partial^2 f(r, t)}{\partial r^2} \sigma^2(r). \end{aligned} \quad (3)$$

Nótese que cuando $\sigma(\cdot) \equiv 0$ este desarrollo coincide con el desarrollo en serie de Taylor estándar no estocástico. El objetivo es ahora hacer elecciones convenientes de la función $f(\cdot, \cdot)$ y construir aproximaciones en base a una cantidad suficiente de datos de la esperanza condicionada en la ecuación (2) para obtener aproximaciones de $\mu(\cdot)$ y $\sigma(\cdot)$.

En concreto, de la expresión (2) se deduce que

$$\mathcal{L}f(r_t, t) = \frac{1}{\Delta} E_t [f(r_{t+\Delta}, t + \Delta) - f(r_t, t)] + O(\Delta), \quad (4)$$

expresión que se usará para obtener aproximaciones de primer orden para las funciones $\mu(\cdot)$ y $\sigma(\cdot)$.

Para obtener las aproximaciones de segundo orden, en primer lugar se añade un término más en la expresión (4) y se toman tiempos entre observaciones de Δ y 2Δ , obteniendo respectivamente:

$$\mathcal{L}f(r_t, t) = \frac{1}{\Delta} \mathbb{E}_t [f(r_{t+\Delta}, t + \Delta) - f(r_t, t)] - \frac{1}{2} \mathcal{L}^2 f(r_t, t) \Delta + O(\Delta^2), \quad (5)$$

$$\mathcal{L}f(r_t, t) = \frac{1}{2\Delta} \mathbb{E}_t [f(r_{t+2\Delta}, t + 2\Delta) - f(r_t, t)] - \frac{1}{2} \mathcal{L}^2 f(r_t, t) 2\Delta + O(\Delta^2). \quad (6)$$

Para que no aparezca el término $\mathcal{L}^2 f(r_t, t)$, el cual involucra a las funciones desconocidas $\mu(\cdot)$ y $\sigma(\cdot)$ y a sus derivadas, se multiplica la ecuación (5) por 2 y se le resta la ecuación (6) obteniendo la expresión que dará lugar a las aproximaciones de segundo orden de $\mu(\cdot)$ y $\sigma(\cdot)$. En concreto, la expresión que se obtiene es la siguiente:

$$\begin{aligned} \mathcal{L}f(r_t, t) = & \frac{1}{2\Delta} \{4\mathbb{E}_t [f(r_{t+\Delta}, t + \Delta) - f(r_t, t)] - \\ & \mathbb{E}_t [f(r_{t+2\Delta}, t + 2\Delta) - f(r_t, t)] + O(\Delta^2)\}. \end{aligned} \quad (7)$$

De manera análoga, para obtener las aproximaciones de tercer orden se añade un término adicional en la expresión (5) y se escribe con intervalos de tiempo Δ , 2Δ y 3Δ , obteniendo:

$$\mathcal{L}f(r_t, t) = \frac{1}{\Delta} \mathbb{E}_t [f(r_{t+\Delta}, t + \Delta) - f(r_t, t)] - \frac{1}{2} \mathcal{L}^2 f(r_t, t) \Delta - \frac{1}{6} \mathcal{L}^3 f(r_t, t) \Delta^2 + O(\Delta^3) \quad (8)$$

$$\mathcal{L}f(r_t, t) = \frac{1}{2\Delta} \mathbb{E}_t [f(r_{t+2\Delta}, t + 2\Delta) - f(r_t, t)] - \frac{1}{2} \mathcal{L}^2 f(r_t, t) (2\Delta) - \frac{1}{6} \mathcal{L}^3 f(r_t, t) (2\Delta)^2 + O(\Delta^3) \quad (9)$$

$$\mathcal{L}f(r_t, t) = \frac{1}{3\Delta} \mathbb{E}_t [f(r_{t+3\Delta}, t + 3\Delta) - f(r_t, t)] - \frac{1}{2} \mathcal{L}^2 f(r_t, t) (3\Delta) - \frac{1}{6} \mathcal{L}^3 f(r_t, t) (3\Delta)^2 + O(\Delta^3) \quad (10)$$

Finalmente, se busca la combinación lineal de las tres ecuaciones que elimina las cantidades $\mathcal{L}^2 f(r_t, t)$ y $\mathcal{L}^3 f(r_t, t)$. En concreto, si a la ecuación (10) le restamos la ecuación (9) multiplicada por 3 y le sumamos la ecuación (8) obtenemos la expresión que proporcionará las aproximaciones de tercer orden para $\mu(\cdot)$ y $\sigma(\cdot)$. En concreto, tal expresión resulta ser:

$$\begin{aligned} \mathcal{L}f(r_t, t) = & \frac{1}{6\Delta} \{18\mathbb{E}_t [f(r_{t+\Delta}, t + \Delta) - f(r_t, t)] - 9\mathbb{E}_t [f(r_{t+2\Delta}, t + 2\Delta) - f(r_t, t)] + \\ & 2\mathbb{E}_t [f(r_{t+3\Delta}, t + 3\Delta) - f(r_t, t)] + O(\Delta^3)\}. \end{aligned} \quad (11)$$

Para obtener las aproximaciones de primer, segundo y tercer orden a la deriva propuesta para los movimientos del tipo de interés a corto libre de riesgo, bastaría tomar $f(r, t) = r$, y entonces $\mathcal{L}f(r_t, t) = \mu(r_t)$. Las aproximaciones (4), (7) y (11) ahora particularizadas para esta función son las siguientes:

$$\mu(r_t) = \frac{1}{\Delta} \mathbb{E}_t[r_{t+\Delta} - r_t] + O(\Delta) \quad (12)$$

$$\mu(r_t) = \frac{1}{2\Delta} \{4\mathbb{E}_t[r_{t+\Delta} - r_t] - \mathbb{E}_t[r_{t+2\Delta} - r_t]\} + O(\Delta^2) \quad (13)$$

$$\mu(r_t) = \frac{1}{6\Delta} \{18\mathbb{E}_t[r_{t+\Delta} - r_t] - 9\mathbb{E}_t[r_{t+2\Delta} - r_t] + 2\mathbb{E}_t[r_{t+3\Delta} - r_t]\} + O(\Delta^3) \quad (14)$$

Seguidamente, para obtener las aproximaciones a $\sigma^2(r_t)$ basta tomar ahora $f(r, t) = (r - r_t)^2$, con lo que $\mathcal{L}f(r_t, t) = \sigma^2(r_t)$. En esta situación, las aproximaciones (4), (7) y (11) quedan de la siguiente forma:

$$\sigma^2(r_t) = \frac{1}{\Delta} \mathbb{E}_t[(r_{t+\Delta} - r_t)^2] + O(\Delta) \quad (15)$$

$$\sigma^2(r_t) = \frac{1}{2\Delta} \{4\mathbb{E}_t[(r_{t+\Delta} - r_t)^2] - \mathbb{E}_t[(r_{t+2\Delta} - r_t)^2]\} + O(\Delta^2) \quad (16)$$

$$\sigma^2(r_t) = \frac{1}{6\Delta} \{18\mathbb{E}_t[(r_{t+\Delta} - r_t)^2] - 9\mathbb{E}_t[(r_{t+2\Delta} - r_t)^2] + 2\mathbb{E}_t[(r_{t+3\Delta} - r_t)^2]\} + O(\Delta^3) \quad (17)$$

Tomando raíces cuadradas a ambos lados de cada una de las tres ecuaciones anteriores obtenemos las siguientes aproximaciones a la volatilidad $\sigma(r_t)$:

$$\sigma(r_t) = \sqrt{\frac{1}{\Delta} \mathbb{E}_t[(r_{t+\Delta} - r_t)^2]} + O(\Delta) \quad (18)$$

$$\sigma(r_t) = \sqrt{\frac{1}{2\Delta} \{4\mathbb{E}_t[(r_{t+\Delta} - r_t)^2] - \mathbb{E}_t[(r_{t+2\Delta} - r_t)^2]\}} + O(\Delta^2) \quad (19)$$

$$\sigma(r_t) = \sqrt{\frac{1}{6\Delta} \{18\mathbb{E}_t[(r_{t+\Delta} - r_t)^2] - 9\mathbb{E}_t[(r_{t+2\Delta} - r_t)^2] + 2\mathbb{E}_t[(r_{t+3\Delta} - r_t)^2]\}} + O(\Delta^3) \quad (20)$$

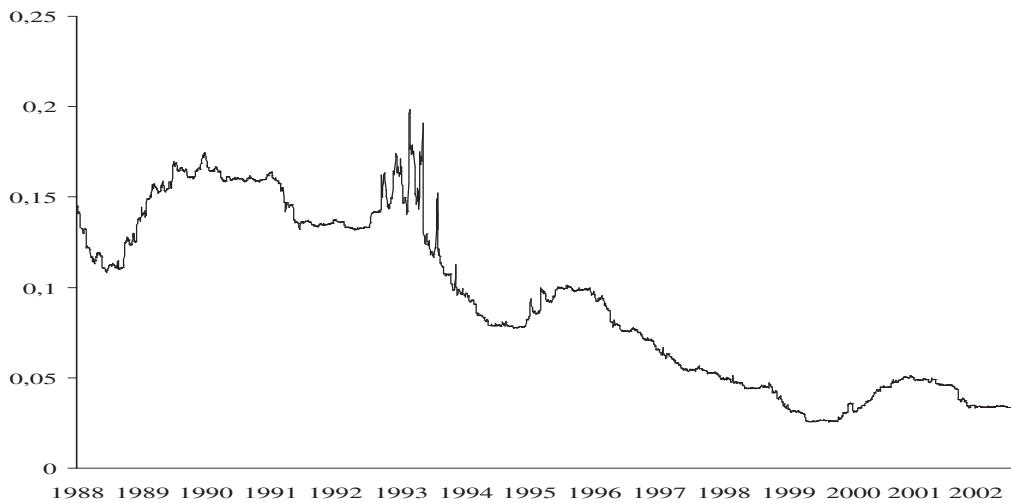


Figura 1: *Serie de datos diarios de los tipos de interés interbancarios a un mes.*

3. Los datos, un primer análisis

Para estimar la estructura temporal de los tipos a corto plazo para la economía española usando la metodología expuesta en la Sección anterior se va a usar la serie de datos diarios del tipo de interés interbancario a un mes desde enero de 1988 hasta septiembre de 2002.

Esta serie, con frecuencia mensual, ya ha sido usada en García (1998)[17] para estudiar el modelo paramétrico propuesto en Chan *et al.*(1992)[8], el cual es un caso particular del modelo de evolución propuesto en la ecuación (1). En el artículo de García (1998)[17] se comenta cómo es común usar los tipos de interés a un mes para describir la dinámica de los tipos de interés a corto plazo. En este sentido, Nowman (1977)[27] usa los tipos de interés interbancarios en Gran Bretaña y Chan *et al.*(1992)[8], Longstaff y Schwartz (1992)[24] y Vetzal (1997)[31] usan los tipos de interés de las letras del tesoro americanas a un mes.

En la figura 1 se muestra la evolución de la serie diaria de los tipos de interés interbancarios a un mes anualizados en España, y en la figura 2 un histograma de frecuencia.

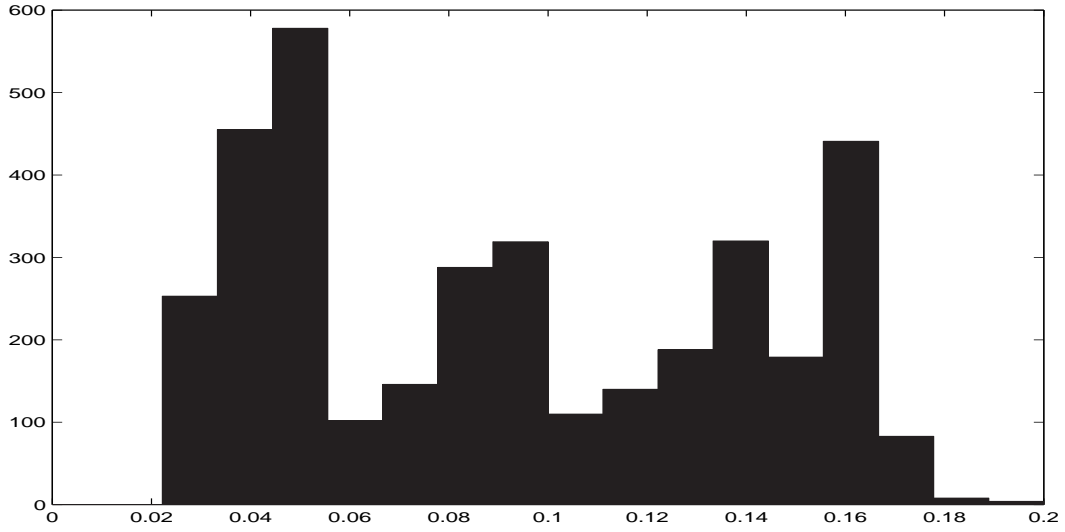


Figura 2: *Histograma de frecuencias de la serie diaria de tipos de interés interbancarios a un mes.*

Este histograma, se puede contemplar como una estimación no paramétrica de la densidad estacionaria de la serie temporal. De forma genérica, dada una anchura h y un origen r_0 , los intervalos que constituyen el histograma vienen dados por $I_j = [r_0 + jh, r_0 + (j + 1)h]$, para $j \in \mathbb{Z}$, y el histograma vendría entonces dado por la función

$$\hat{f}(r) = \frac{\sum_{j=1}^n I_{[r_0+jh, r_0+(j+1)h]}(r_t)}{nh}, \quad r \in I_j. \quad (21)$$

Nótese que la elección del parámetro h controla la cantidad de suavidad que se obtendrá en la densidad estimada resultante. Un importante inconveniente de este estimador es que todas las observaciones que caen fuera de I_j no tienen ninguna influencia en $\hat{f}(r)$, lo cual es particularmente negativo en los extremos del intervalo I_j . Un remedio para este inconveniente es la construcción de un histograma adaptado, en el cual para estimar la densidad en un punto r se tienen en cuenta todas las observaciones en un intervalo centrado en el punto r y de longitud h . Así, su expresión analítica viene dada por la siguiente función:

$$\hat{f}(r) = \frac{\sum_{t=1}^n I_{[r-\frac{h}{2}, r-\frac{h}{2}]}(r_t)}{nh}, \quad (22)$$

la cual se puede escribir también bajo la siguiente forma:

$$\hat{f}(r) = \frac{1}{nh} \sum_{t=1}^n K_0\left(\frac{r - r_t}{h}\right), \quad x \in \mathbb{R}, \quad (23)$$

siendo $K_0(u) = I_{[-\frac{1}{2}, \frac{1}{2}]}(u)$ el llamado kernel intuitivo. Para obtener estimaciones más suaves, se suelen usar otros kernel.¹ Dado un kernel $K(\cdot)$, el estimador kernel de la densidad viene dado por

$$\hat{f}(r) = \frac{1}{nh} \sum_{t=1}^n K\left(\frac{r - r_t}{h}\right), \quad x \in \mathbb{R}, \quad (24)$$

donde el parámetro h es la llamada anchura de la ventana. En esta situación, $\hat{f}(r)$ es una mixtura de n funciones de densidad y por tanto es también una función de densidad.

En consecuencia, para poder llevar a cabo esta estimación de la densidad se debe elegir un núcleo $K(\cdot)$ y una anchura de ventana h . Algunos resultados teóricos prueban que la elección del kernel no es importante desde el punto de vista del error asintótico de estimación y se suele elegir atendiendo a otras consideraciones como diferenciabilidad o esfuerzo computacional asociado a su cálculo². Muy usado en la práctica, y el que se usará siempre en este trabajo, es el kernel Gaussiano, en el cual $K(\cdot)$ es la densidad de la distribución normal estándar

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \quad u \in \mathbb{R}. \quad (25)$$

En este caso $\hat{f}(r)$ es la media de n densidades normales centradas en r_t y de varianza h . Por tanto, cuanto más lejos esté un punto muestral del punto de estimación menos contribuirá a la densidad estimada, así la densidad estimada será mayor cerca de las grandes concentraciones de datos muestrales y menor donde las observaciones sean escasas. Por otra parte, al ser en este caso el kernel infinitamente diferenciable la densidad estimada también lo será.

En cuanto a la anchura de la ventana h , su elección es de gran importancia para la precisión del estimador, ya que $\hat{f}(r)$ es muy sensible a la elección de h . En el caso en que la verdadera densidad sea Gaussiana y se use un kernel Gaussiano, la elección óptima desde el punto de vista de obtener un menor error cuadrático integrado medio viene dada por

$$h = 1,06\hat{\sigma}n^{-\frac{1}{5}}, \quad (26)$$

siendo $\hat{\sigma}$ una estimación de la desviación estándar de la distribución. Sin embargo, tal elección sobre-suaviza las distribuciones multimodales y asimétricas, el cual parece ser el caso de la

¹Un kernel es una función de densidad simétrica acotada tal que $|u|K(u) \rightarrow 0$ cuando $|u| \rightarrow 0$ y con $\int u^2 K(u) du < \infty$.

²Ver Silverman (1986)[28] para más detalles.

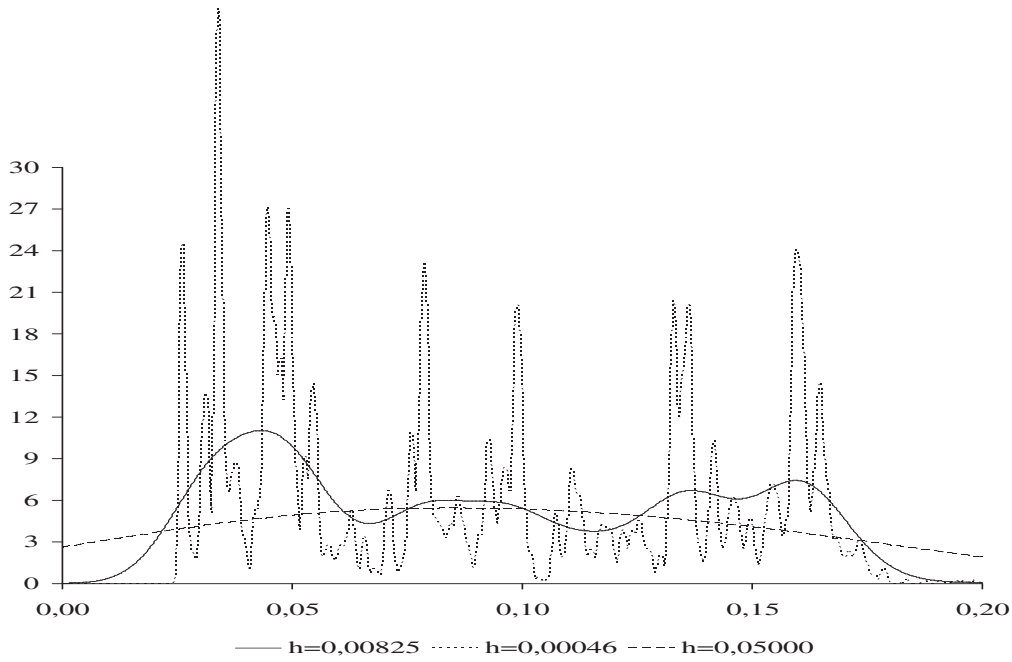


Figura 3: *Densidad incondicional de los tipos de interés interbancarios a un mes estimada para diferentes anchuras de ventana.*

distribución que pretendemos estimar a la vista del histograma de la figura 2. Mejores resultados en una gran variedad de distribuciones produce la siguiente elección:

$$h = 0,9 \min \left\{ \hat{\sigma}, \frac{R_Q}{1,34} \right\} n^{-\frac{1}{5}}, \quad (27)$$

siendo R_Q el recorrido intercuartílico de la distribución subyacente.

Para ilustrar la influencia de h sobre la densidad estimada se presenta en la figura 3 estimaciones correspondientes a diferentes valores de h , en concreto se ha tomado $h = 0,00825$, dada por la expresión (27), $h = 0,00046$, un valor sensiblemente menor, y $h = 0,05$, un valor sensiblemente mayor. Se observa como en efecto la densidad estimada es muy sensible al valor de h , pasando de una estimación con $h = 0,00046$ poco suave y mostrando excesivos detalles³ a la obtenida con $h = 0,05$ que es excesivamente suave y oculta detalles importantes. Parece que la situación intermedia con $h = 0,008253$ obtenida a partir de la ecuación (27) es la más adecuada.

En la figura 4 acompañamos a la densidad estimada de bandas puntuales de confianza. Estas

³Esta elección corresponde al método de validación cruzada máximo verosímil. A pesar de ser un método más elaborado en este caso lleva claramente a una peor estimación que la elección automática dada en (27). Ver Silverman (1986)[28] para más detalles.

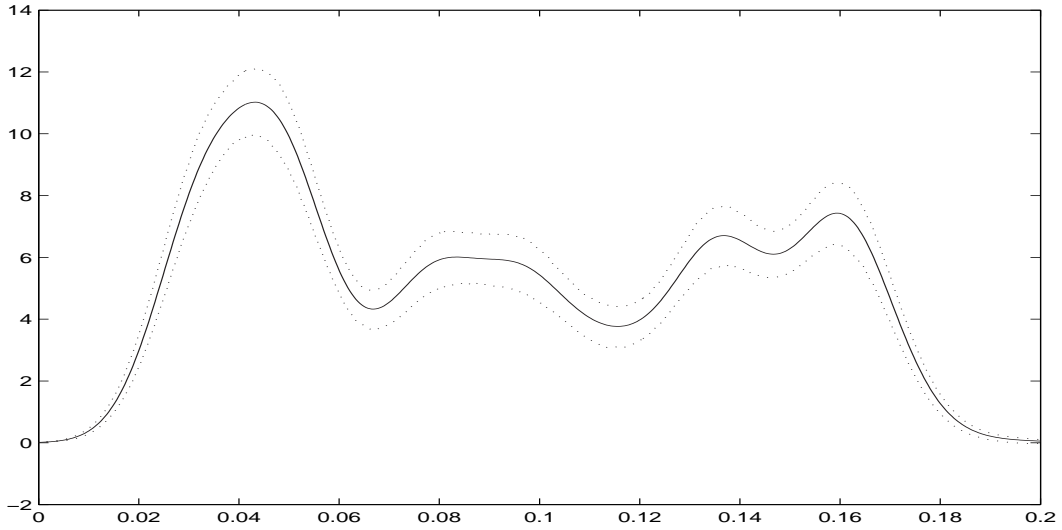


Figura 4: *Densidad incondicional estimada de los tipos de interés interbancarios a un mes y bandas puntuales al 95 % de confianza construidas con el método bootstrap de los bloques móviles.*

bandas de confianza se han obtenido mediante el método bootstrap de bloques móviles desarrollado en Künsch (1989)[22] y Liu y Singh (1992)[23], con el cual, al captar la dependencia débil entre las observaciones, se logran mejores resultados que con la clásica normalidad asintótica del estimador. Esta gráfica no invita a pensar que la evolución temporal de los tipos de interés en España obedezca a modelos tan simples como los modelo Gaussianos.

Con este método de remuestreo se obtiene la distribución en el muestreo del estimador $\hat{f}(r)$ de $f(r)$ remuestreando de los bloques móviles $B_1, B_2, \dots, B_{n-b+1}$, donde b es el tamaño de cada bloque y $B_j = \{r_j, r_{j+1}, \dots, r_{j+b-1}\}$ con $j = 1, 2, \dots, n - b + 1$. El método consiste en obtener una muestra de $K(\cdot)$ bloques⁴ de manera independiente e idénticamente distribuida que se denota por $(\xi_1, \xi_2, \dots, \xi_k)$ y que es la muestra bootstrap con la cual se obtiene el estimador bootstrap $f^*(r)$. Repitiendo el procedimiento para una gran cantidad de muestras bootstrap, se puede aproximar la distribución muestral de $\hat{f}(r) - f(r)$ por la de $f^*(r) - \hat{f}(r)$ y obtener un intervalo de confianza para $f(r)$.

⁴Aquí k se debe tomar para que kb sea del mismo orden que n .

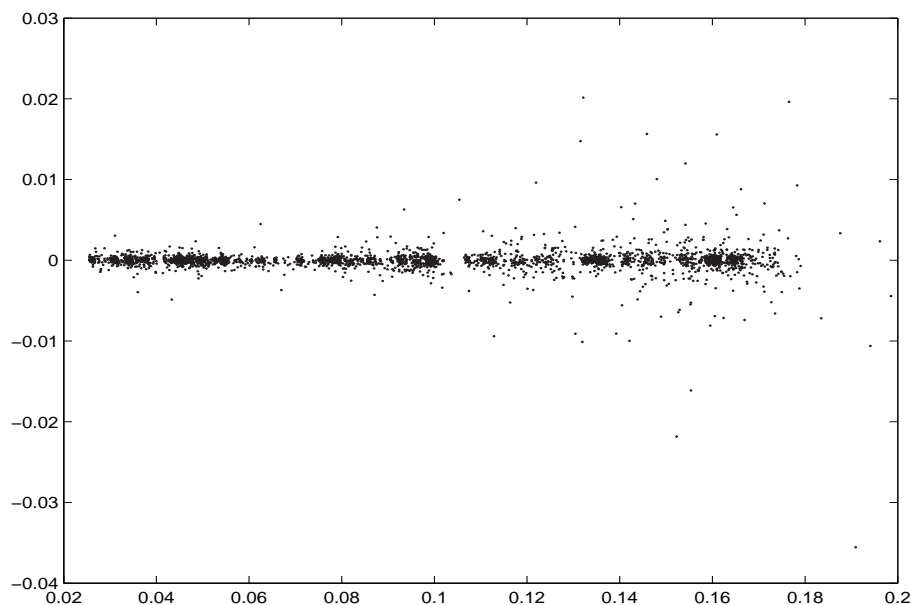


Figura 5: Variaciones diarias de los tipos de interés interbancarios a un mes frente a su valor el primero de los días.

4. Estimación del modelo

Esta sección está dedicada a la estimación de la difusión homogénea propuesta en la ecuación (1) como modelo para la evolución temporal de los tipos de interés a corto plazo en España.

Es importante observar que para llevar a cabo las aproximaciones a la deriva $\mu(r)$ dadas en las ecuaciones (12), (13) y (14), basta con estimar a partir de los datos las siguientes esperanzas condicionadas:

$$E[r_{t+j\Delta} - r_t / r_t = r] \quad j = 1, 2, 3. \quad (28)$$

Nótese que estimar la esperanza condicionada $E[r_{t+\Delta} - r_t / r_t = r]$, con $\Delta = \frac{1}{250}$ correspondiente a observaciones diarias, equivale a determinar la línea de regresión en la gráfica 5, en la cual se muestran los cambios diarios del tipo de interés en función del tipo de interés del primer día. Esta gráfica muestra cierta evidencia de heteroscedasticidad, ya que se observan mayores variaciones cuando los tipos son más altos.

De la misma manera, para obtener las aproximaciones a la volatilidad propuestas en la ecuaciones (18), (19) y (20), se deben estimar, a partir de los datos observados, las cantidades

$$\mathbb{E} \left[(r_{t+j\Delta} - r_t)^2 / r_t = r \right], \quad j = 1, 2, 3. \quad (29)$$

En definitiva, todo se reduce a obtener estimaciones no paramétricas de la regresión de una variable Y sobre una variable X , es decir, en estimar

$$m(x) = \mathbb{E}[Y/X = x]$$

En el contexto de la regresión no lineal de una variable Y sobre una variable X , la estimación análoga al histograma es el regresograma, el cual viene dado por

$$\hat{m}(x) = \frac{\sum_{i=1}^n y_i I_j(x_i)}{\sum_{i=1}^n I_j(x_i)} = \bar{Y}/X \in I_j, \quad (30)$$

donde los intervalos I_j son los descritos en la Sección 3 para el histograma. Esta estimación no paramétrica de la línea de regresión adolece de los mismos inconvenientes que el histograma y se le puede dar la misma solución. Por tanto, el estimador kernel propuesto para $m(x)$, llamado estimador de Nadaraya-Watson, viene dado por:

$$\hat{m}(x) = \frac{\sum_{i=1}^n y_i K\left(\frac{x_i - x}{h}\right)}{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right)} = \sum_{i=1}^n y_i w_i(x). \quad (31)$$

De esta manera, el estimador de $m(x)$ es una media aritmética ponderada de las respuestas observadas y_i , donde cada ponderación $w_i(x)$ depende de todo de todo el vector de observaciones X y es mayor cuanto más cerca esté x_i de x .

Nuevamente, para poder llevar a cabo esta estimación kernel es necesario especificar el kernel $K(\cdot)$ y la anchura de la ventana h , la cual controla el grado de suavidad de la función estimada $\hat{m}(x)$. Al igual que en el caso de la estimación de la densidad, la elección del kernel no es una cuestión crítica para la precisión del estimador y continuaremos usando el kernel Gaussiano.

Sin embargo, es de gran importancia para la precisión del estimador la elección de una anchura de ventana adecuada, la cual se suele buscar minimizando el error cuadrático medio

$$\text{ECM}(h) = \frac{\sum_{i=1}^n [m(x_i) - \hat{m}(x_i)]^2}{n}. \quad (32)$$

Un estimador intuitivo este error cuadrático medio se obtendría reemplazando el desconocido $m(x_i)$ por la observación y_i , llevando a la media de los cuadrados de los residuos de la regresión, dada por

$$\text{MCR}(h) = \frac{\sum_{i=1}^n [y_i - \hat{m}(x_i)]^2}{n}. \quad (33)$$

Sin embargo, este estimador sesgado del error cuadrático medio llevaría siempre a que la mejor elección es la menor anchura de ventana. La razón intuitiva de esto es que para predecir y_i se usa la propia observación y_i en $\hat{m}(x_i)$. Por tanto, una opción más adecuada sería excluir la observación (x_i, y_i) al estimar $m(x_i)$. Llamando a este estimador $\hat{m}_{-i}(x_i)$, se elegiría la anchura de la ventana mediante el llamado método de validación cruzada minimizando la expresión siguiente:

$$\text{CV}(h) = \sum_{i=1}^n [y_i - \hat{m}_{-i}(x_i)]^2. \quad (34)$$

Una sencilla manipulación de la expresión anterior prueba que se puede escribir también en la forma

$$\text{CV}(h) = \sum_{i=1}^n [y_i - \hat{m}(x_i)]^2 \Xi(w_i(x_i)), \quad (35)$$

donde $\Xi(u) = (1 - u)^{-2}$ y $w_i(x_i) = \frac{K(0)}{\sum_{j=1}^n K(\frac{x_i - x_j}{h})}$.

En consecuencia, el método de validación cruzada, el cual lleva a una anchura de ventana asintóticamente óptima, se puede interpretar también como la minimización de una suma de cuadrados de residuos ponderada, donde los valores pequeños de h se penalizan mediante la inclusión de la función $\Xi(\cdot)$. Para más detalles sobre este punto en particular y sobre la regresión no paramétrica en general se puede consultar el excelente texto de Härdle (1997)[20].

A modo de ilustración del método de validación cruzada, se presenta en la figura 6 la función

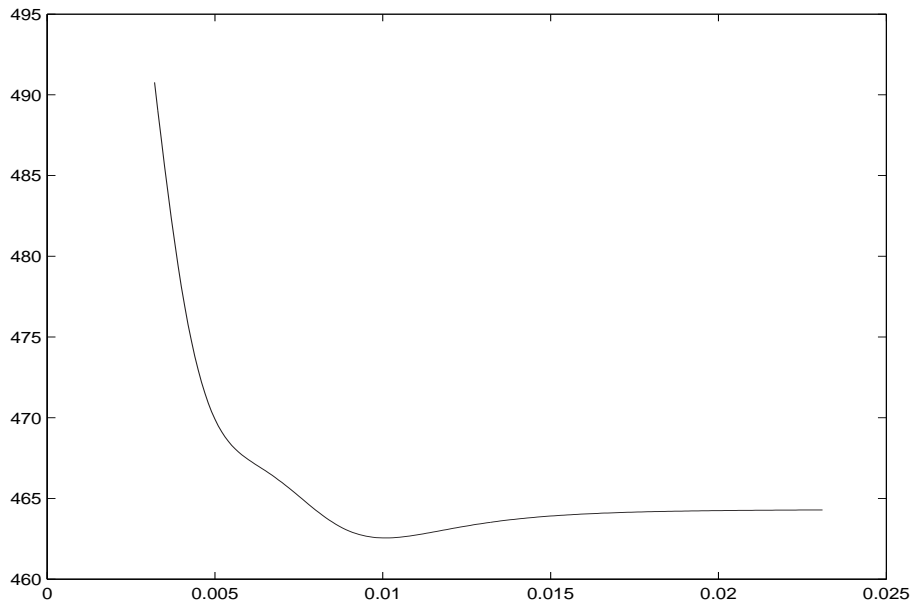


Figura 6: *Determinación de la anchura de la ventana mediante validación cruzada para la regresión no paramétrica conducente a la aproximación de orden uno de la deriva.*

$CV(h)$ correspondiente a la aproximación de orden uno de la deriva dada por la ecuación (12). El valor óptimo de h obtenido es 0,0101.

En las figuras 7 y 8 se representan las aproximaciones de orden uno, dos y tres a la deriva y a la volatilidad de los tipos a corto plazo dadas por las estimaciones kernel de las esperanzas condicionadas en las ecuaciones (12), (13),(14), (18), (19) y (20) y usando una anchura de ventana obtenida por validación cruzada. Se muestran también bandas de confianza puntuales para las aproximaciones de primer orden construidas mediante el método bootstrap de bloques móviles descrito en la Sección 3.

Lo primero a destacar es que las tres aproximaciones, tanto para la deriva como para la volatilidad, son muy similares, que es precisamente lo que cabe esperar que ocurra cuando los errores de aproximación son pequeños.

La estimación de la deriva parece ser no lineal, mostrando una ligera reversión a la media para tipos de interés medios y bajos y cayendo fuertemente cuando los tipos de interés se

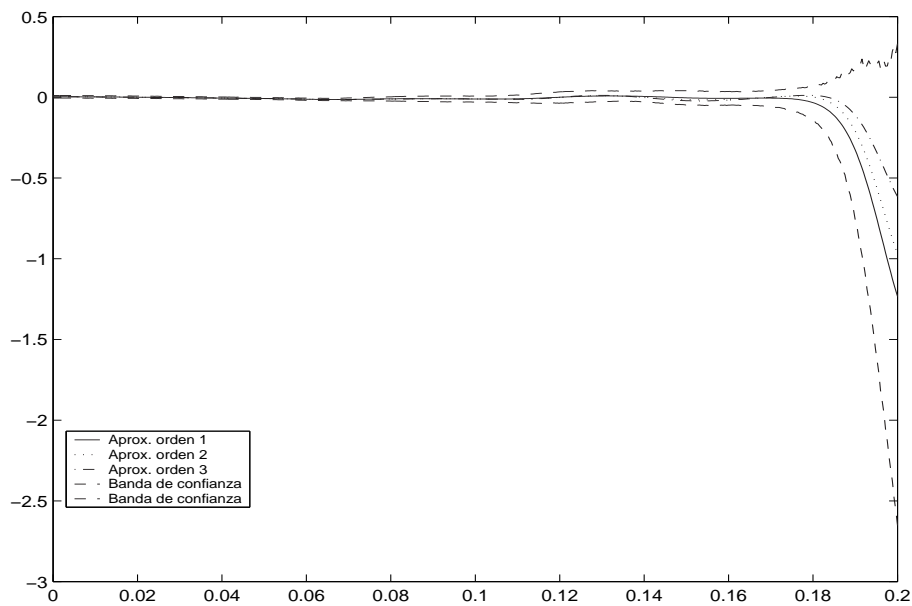


Figura 7: Aproximaciones de primer, segundo y tercer orden a la deriva de la difusión propuesta como modelo de evolución de los tipos de interés a corto plazo y banda de confianza puntual al 95 % para la aproximación de primer orden, construida mediante el método bootstrap de los bloques móviles.

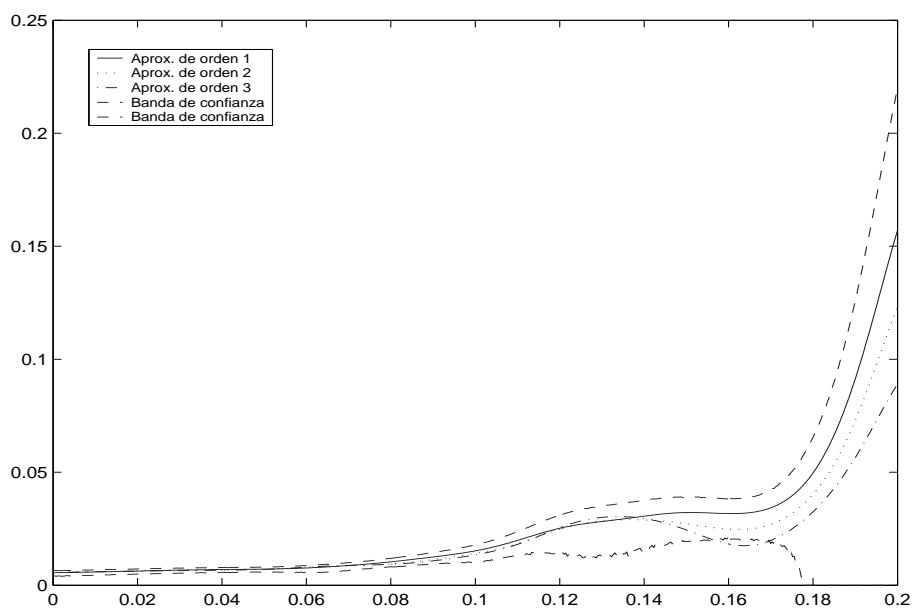


Figura 8: Aproximaciones de primer, segundo y tercer orden a la volatilidad de la difusión propuesta como modelo de evolución de los tipos de interés a corto plazo y banda de confianza puntual al 95 % para la aproximación de primer orden, construida mediante el método bootstrap de los bloques móviles.

sitúan por encima del 17%. Este comportamiento aparentemente no lineal en la deriva es el mismo que encuentran Stanton (1997)[29], usando datos diarios de los tipos de interés de la economía estadounidense, y Corzo y Gómez (2001)[11], usando datos semanales de la economía española. Estos autores concluyen, quizás algo precipitadamente, a partir de las estimaciones que realizan que existe evidencia de no linealidad en la deriva. Sin embargo, Chapman y Pearson (2000)[9] analizan el comportamiento muestral del estimador de la deriva y prueban, usando como referencia el modelo de Cox, Ingersoll y Ross, que aunque la verdadera deriva sea lineal el estimador mostrará una no linealidad precisamente como la exhibida en estos trabajos. En lugar de afirmar la no linealidad de la deriva únicamente a partir de la estimaciones, vamos a usar el test propuesto por Aït-Sahalia, Bickel, y Stoker (2001)[2] para contrastarla. Este será el objetivo de la Sección siguiente.

En cuanto a la estimación de la volatilidad de la difusión, se observa que en general aumenta conforme los tipos de interés son mayores. Por tanto, no parece estar en consonancia con la volatilidad constante propuesta en Vasicek (1977)[30], sino más bien con la propuesta en Chan *et al.* (1992)[8].

Las bandas de confianza obtenidas para la deriva y la volatilidad son bastantes estrechas para tipos de interés bajos y medios, donde hay muchas observaciones. Sin embargo, para tipos de interés altos, donde las observaciones son escasas, las bandas de confianza se hacen mucho más anchas reflejando una menor confianza en las estimaciones.

5. Contrastando la no linealidad de la deriva

El objetivo de esta Sección es contrastar la aparente no linealidad de la deriva exhibida en las estimaciones presentadas en la figura 7. Para ello usaremos el test propuesto en Aït-Sahalia, Bickel y Stoker (2001) [2], el cual se plantea en el contexto no paramétrico de la regresión kernel. Aquí se usará una versión del test general, en concreto, cuando la hipótesis alternativa

es paramétrica.

El contraste que vamos a realizar viene especificado por las siguientes hipótesis nula y alternativa:

$$H_0 : \Pr [m(r) = \alpha_0 + \alpha_1 r] = 1$$

$$H_1 : \Pr [m(r) = \alpha_0 + \alpha_1 r] < 1$$

siendo $m(r) = \frac{1}{\Delta} E [r_{t+\Delta} - r_t / r_t = r]$ la aproximación de primer orden de la deriva.

La forma del estadístico considerada en principio para el contraste es

$$\hat{\Gamma} = \frac{1}{n} \sum_{i=1}^n [\hat{m}_h(r_i) - (\hat{\alpha}_0 + \hat{\alpha}_1 r_i)]^2 a(r_i), \quad (36)$$

que es la suma de los cuadrados de las diferencias de los residuos bajo las hipótesis nula y alternativa, ponderados por una función no negativa $a(\cdot)$. Esta función peso, que suele ser la función indicador en un soporte compacto, permite restringirse a áreas de datos con suficiente densidad y dejar fuera las áreas donde los datos son escasos.

El resultado principal en Aït-Sahalia, Bickel y Stoker (2001) [2] es que la distribución asintótica de $\hat{\Gamma}$ es normal sesgada. Más precisamente, se tiene que

$$\hat{\tau} = \sigma_{11}^{-1} \left[nh^{\frac{1}{2}} \hat{\Gamma} - h^{-\frac{1}{2}} \gamma_{12} \right] \rightarrow N(0, 1), \quad (37)$$

donde

$$\hat{\sigma}_{11}^2 = \frac{2C_{11}}{n} \sum_{i=1}^n \frac{\hat{\sigma}_h^4(r_i)}{\hat{f}_h(r_i)} a^2(r_i) \quad \text{y} \quad \hat{\gamma}_{12} = \frac{C_{12}}{n} \sum_{i=1}^n \frac{\hat{\sigma}_h^2(r_i)}{\hat{f}_h(r_i)} a^2(r_i).$$

En estas últimas expresiones $\hat{\sigma}_h(r)$ es una estimación de la volatilidad de la difusión, $\hat{f}_h(r)$ de la densidad incondicional y las constantes C_{11} y C_{12} vienen determinados por la función kernel elegida en la estimación no paramétrica. En concreto, para un kernel Gaussiano, que es el que se está usando a lo largo de todo este trabajo se tiene que $C_{11} = \frac{1}{2\sqrt{2\pi}}$ y $C_{12} = \frac{1}{2\sqrt{\pi}}$.

Presentamos en el cuadro 3 el resultado del contraste usando una ventana del tipo $h = h_0 n^{-\frac{1}{\delta}}$ con $\delta \in (2, 4.5)$ para que se cumplan las condiciones requeridas en el contraste⁵, el modelo lineal

⁵En concreto, se está tomando $h_0 = 0,40$ y $\delta = 2,25$.

Modelo restringido:	$0,001878 - 0,103621r$
Estadístico: $(\hat{\tau})$	0.275984
p-value:	0.391280

Cuadro 3: *Resultado del contraste de hipótesis de deriva no lineal.*

se ha estimado mediante mínimos cuadrados ordinarios y se ha tomado $a(r) \equiv 1$, es decir se usan todos los datos disponibles.

Teniendo en cuenta la distribución asintótica normal estándar de $\hat{\tau}$ y que el contraste es de carácter unilateral, se rechaza la hipótesis nula si los residuos son muy diferentes bajo las hipótesis del contraste, no hay evidencias para rechazar una deriva lineal.

Considerando además los comentarios de Chapman y Pearson (2000)[9] sobre el comportamiento del estimador kernel de la deriva, ya citados en la Sección anterior, y mencionando que el contraste realizado tiene cierta tendencia al sobre-rechazo, no está justificado abandonar la hipótesis de una deriva lineal en la modelización de los tipos de interés a corto plazo en España mediante una difusión unifactorial.

6. conclusiones

En este trabajo se modeliza la evolución temporal de los tipos de interés a corto plazo en España mediante la difusión homogénea unifactorial

$$dr_t = \mu(r_t) dt + \sigma(r_t) dW_t. \quad (38)$$

Las estimaciones de la deriva y de la volatilidad se llevan a cabo en el contexto no paramétrico propuesto en Stanton (1997)[29], usando la regresión kernel y datos diarios desde enero de 1988 hasta septiembre de 2002 de los tipos de interés interbancarios a un mes.

La deriva estimada exhibe un comportamiento aparentemente no lineal como el obtenido

por Stanton (1997)[29] y Corzo y Gómez (2001)[11], mostrando una ligera reversión a la media para tipos bajos y medios y cayendo rápidamente para tipos altos. La volatilidad estimada es creciente con el tipo de interés y parece estar en consonancia con la propuesta en Chan *et al.*(1992)[8].

Sin embargo, teniendo en cuenta que, como exponen Chapman y Pearson (2000)[9], el estimador kernel de la deriva puede exhibir este tipo de comportamiento no lineal aunque la verdadera deriva sea lineal, una conclusión de no linealidad en la deriva a partir de las estimaciones parece precipitada. Para ahondar un poco más en este asunto, contrastamos la linealidad de la deriva usando el test de Aït-Sahalia, Bickel, y Stoker (2001)[2] y obtenemos que no es posible rechazar la hipótesis de linealidad. Por tanto, a partir de estas consideraciones está justificado modelizar la evolución temporal de los tipos de interés en España mediante una difusión con deriva lineal.

Referencias

- [1] Aït-Sahalia, Y. (1996). Testing continuous-time models of the spot rate. *The Review of Financial Studies*, volumen 9, 385–426.
- [2] Aït-Sahalia, Y., Bickel, P. J., y Stoker, T. M. (2001). Goodness-of-fit test for kernel regression with an application to option implied volatilities. *Journal of Econometrics*, volumen 105, 363–412.
- [3] Black, F., Derman, E., y Toy, W. (1990). A one-factor model of interest rates and its application to treasury bond options. *Financial Analysts Journal*, volumen 46, 33–39.
- [4] Black, F. y Karasinski, P. (1991). Bond and option pricing when short rates are lognormal. *Financial Analysts Journal*, volumen 47, 52–59.
- [5] Black, F. y Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, volumen 81, 637–659.

- [6] Brennan, M. y Schwartz, E. (1980). Analyzing convertible bonds. *Journal of Financial and Quantitative Analysis*, volumen 15, 4, 907–929.
- [7] Canabarro, E. (1995). Where do one-factor interest rate models fail? *J. Fixed Income*, volumen 5, 31–52.
- [8] Chan, K. C., Karolyi, G. A., Longstaff, F. A., y Sanders, A. B. (1992). An empirical comparison of alternative models of the term structure of interest rates. *The Journal of Finance*, volumen 47, 1209–1228.
- [9] Chapman, D. A. y Pearson, N. D. (2000). Is the short rate drift actually nonlinear? *Journal of Finance*, volumen 55, 1, 355–388.
- [10] Constantinides, G. (1992). A theory of the nominal term structure of interest rates. *The Review of Financial Studies*, volumen 5, 531–552.
- [11] Corzo, T. y Gómez, J. (2001). Nonparametric estimation of interest rates processes in europe. *Revista de Economía Aplicada*, volumen 9, 27, 83–101.
- [12] Cox, J. (1975). Notes on option pricing I: Constant elasticity of variance diffusions. Working Paper, Stanford University (reimpreso en *Journal of Portfolio Management* 1996, 22, 15–17).
- [13] Cox, J. C., Ingersoll, Jr., J. E., y Ross, S. A. (1985). An intertemporal general equilibrium model of asset prices. *Econometrica*, volumen 53, 363–384.
- [14] Cox, J. C., Ingersoll, Jr., J. E., y Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica*, volumen 53, 385–407.
- [15] Dothan, L. U. (1978). On the term structure of interest rates. *Journal of Financial Economics*, volumen 6, 59–69.
- [16] Duffie, D. y Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance*, volumen 6, 379–406.
- [17] García, J. (1998). Tipos de interés a corto plazo en España. *Revista de Economía Aplicada*, volumen 18, 6, 5–26.

- [18] Hille, E. y Phillips, R. S. (1957). *Functional Analysis and Semigroups*. (American Mathematical Society, Providence, RI).
- [19] Ho, T. y Lee, S. (1986). Term structure movements and pricing interest rate contingent claims. *Journal of Finance*, volumen 41, 1011–1029.
- [20] Härdle, W. (1997). *Applied Nonparametric Regression*. Cambridge University Press.
- [21] Hull, J. y White, A. (1990). Pricing interest rate derivative securities. *Review of Financial Studies*, volumen 3, 573–592.
- [22] Künsch, H. (1989). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, volumen 17, 3, 1217–1241.
- [23] Liu, R. Y. y Singh, K. (1992). Moving blocks jackknife and bootstrap capture weak dependence. En: *Exploring the Limits of Bootstrap*. Editado por LePage, R, y Billard, L.
- [24] Longstaff, F. A. y Schwartz, E. S. (1992). Interest rate volatility and the term structure: A two-factor general equilibrium model. *The Journal of Finance*, volumen 47, 1259–1282.
- [25] Merton, R. C. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science*, volumen 4, 141–183.
- [26] Moreno, M. (2000). Modelización de la estructura temporal de los tipos de interés: valoración de activos derivados y comportamiento empírico. *Revista Española de Financiación y Contabilidad*, volumen 29, 104, 345–376.
- [27] Nowman, K. (1997). Gaussian estimation of single-factor continuous time models of the term structure of interest rates. *Journal of Finance*, volumen 52, 4, 1695–1706.
- [28] Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall.
- [29] Stanton, R. (1997). A nonparametric model of term structure dynamics and the market price of interest rate risk. *Journal of Finance*, volumen 52, 5, 1973–2002.
- [30] Vasicek, O. A. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, volumen 5, 2, 177–188.

- [31] Vetzal, K. R. (1997). Stochastic volatility, movements in short-term interest rates and bond option values. *Journal of Banking and Finance*, volumen 21, 169–196.