Volumen VII, No. 14, Enero-Junio, 2012, pp. 7-27

NONPARAMETRIC SPECIFICATION TESTING FOR CONTINUOUS TIME MODELS FOR INTEREST RATES IN MEXICO

José A. Núñez Mora*
Carlos A. Martínez Reyes**

(Recibido: Julio 2011 / Aprobado: Diciembre 2011)

Resumen

En este artículo se propone una prueba estadística para la especificación de los modelos paramétricos de dos factores. Se presentan tres pruebas diferentes. Las dos primeras se basan en una comparación de la estimación de la densidad de núcleo de la función de densidad desconocida y la estimación de la función de densidad marginal mediante el método Delta. La última prueba se basa en la idea de la comparación entre la estimación de la densidad de núcleo y el modelo paramétrico de la densidad de núcleo suavizado para evitar los efectos de sesgo. En particular, esta prueba se aplicó para determinar si la dinámica de la estructura temporal de tasa de interés de Cetes en México para el período 2002-2009 puede ser modelada a partir de los supuestos de los dos modelos, el de Brennan-Schwartz y el de Schaefer y Schwartz; los resultados de la prueba muestran que ambos modelos continuos son rechazados y por lo tanto no son capaces de describir los datos de los Cetes en México.
Palabras clave: modelo de tiempo continuo, función de densidad marginal, método Delta, estimación no paramétrica, proceso de difusión
Clasificación JEL: C14, C44, C51

---

* Instituto Tecnológico y de Estudios Superiores de Monterrey, Campus Ciudad de México. Electronic mail: <janm@itesm.mx>.

** Instituto Tecnológico y de Estudios Superiores de Monterrey, Campus Estado de México. Electronic mail: <creyes@itesm.mx>.

Abstract

In this paper we propose a statistical test for the specification of parametric models of two factors. We present three different tests. The first two are based on a comparison between the estimate of kernel density of the unknown density function and the estimate of marginal density function by the Delta method. The last test is based on the idea of comparison between the estimate of kernel density and the parametric model of the smoothed kernel density to avoid skew effects. Particularly, this test was applied to determine if the dynamic of the term structure of the Mexican Cetes interest rate in the period 2002-2009 can be modeled from the assumptions of two models, that of Brennan-Schwartz and that of Schaefer-Schwart; the test results show that both continuous models are rejected and therefore are unable to describe the data of the Mexican Cetes.
Keywords: continuous-time model, marginal density function, Delta method, nonparameter estimation, diffusion process
JEL Classification: C14, C44, C51

1. Introduction

A common approach to model the term structure of interest rates is such that express the interest rate in terms of one or more stochastic factors, which in turn follow continuous time stochastic processes. Several studies by Dybvig (1989) and Steeley (1991) have concluded that the variability of rates with different maturity dates can be best explained more than one stochastic factor. This has lead researchers to develop time structure models that use two or more stochastic factors. Multi factor models are proposed by Brennan and Schwartz (1979), Schaefer and Schwartz (l984), Longstaff and Schwartz (l992), Hull and White (1990), Fong and Vasicek (1991), Heath, Jarrow and Morton (1992), Duffie and Kan (1995). The model of Brennan y Schwartz (1979) assumes that the *short and long* rates are the driving forces for the time structure, while the model advanced by Fong y Vasicek (1991) include the *short rate* and volatility as the main factors.

In the absence of a theoretical framework that defines in detail the term structure of interest rates, can be testing different models with real data, without using any observation that coming from a set of prices derived from interest rates. Aït-Sahalia (1996) proposes an approach in this direction, it is assumed that the properties of a model that gives structure to terms is determined completely by a difusion process. This process is characterized by its two first moments of time continuum, tendency and diffusion. Each parametric model of the term structure has a certain density function characterized by the tendency and diffusion functions.

The statistical test is based in a comparison between the density function obtained from a parametric model of the time structure and a non-parametric estimate of the density function, which in turn is derived form data. The density function is valid even if the parametric model of the time structure is not well specified.

Nevertheless it should be pointed out that there are limits to this statistical test: first, it is only applied to one-factor models for term structure. Second, the most important assumption for this test is that data are smoothed when the non-parametric estimate of the density function is constructed, that means, it possess a high variance so the estimate introduces too much “noise” expressed by many “illegitimate” modes (relative maximums) which in turn do not appear in the desired density calculation, in this case, the non-parametric estimate is less than optimal. Therefore is not clear what happens to this test when the data are over smoothed. Thirdly, the test do no consider the skew effects inherent to the non parametric estimate of the density function. Next, the test of Yacine Aït-Sahalia (1996) will be explored in the aforementioned directions.

2. The model and null hypothesis

For a complete probability space (Ω, F, P) and augmented filtration $\{F_t : t \geq 0\}$ generated by a standard Brownian movement W in R^d, a time continuum model tipically depends on a stationary diffusion process X that takes values from some open subset D of R^d, with a dynamics represented by the Ito stochastic differential equation,

$$dX_t = \mu(X_t, \beta)dt + \sigma(X_t, \beta)dW_t \tag{1}$$

So that for any β in a bounded subset $\Theta \subset R^d$, $\mu(\cdot, \beta) \in R^d$ and $\sigma(\cdot, \beta) \in R^{d \times d}$ are the tendency and difussion functions respectively. The distribuition of the process is completely characterized by its tendency and its diffusion. For example, for a one dimension stationary diffusion process, the marginal density function can be written as

$$\pi(x,\beta) = \frac{\eta(\beta)}{\sigma^2(x,\beta)} \int_{x_0}^{x} exp\{\frac{2\mu(u,\beta)}{\sigma^2(u,\beta)}\} du \tag{2}$$

Where the process is distributed over R and $\eta(\beta)$ is a standarization constant that ensures that density integrates to one.

Generally, for any $\beta \in \Theta$ we use $\pi(x,\beta)$ to express the marginal density function which is implicit in the parametric model $\pi(x)$ to express the true marginal density function. The null hypothesis and alternative hypothesis are:

$$H_0 : exists\ \beta_0 \in \Theta\ such\ that\ \pi(x,\beta_0) = \pi(x)$$

$$H_1 : \pi(\cdot,\beta) \neq \pi(\cdot)\ for\ any\ \beta \in \Theta$$

As in Aït-Sahalia (1996), our statistical test is based in a pondered integral of the squared difference between $\pi(x)$ and $\pi(x,\beta_0)$,

$$I = \int (\pi(x) - \pi(x,\beta))^2 \pi(x) dx \tag{3}$$

We can use the measure I as a pointer of the incorrect specification of the model, because $I \geq 0$ and $I = 0$ if and only if the marginal density function implied by the model is correctly specified. In order to obtain a consistent of parameter β_0, a non parametric method is use the *Delta method of nonparametric kernel functionals* of Yacine Aït-Sahalia (1992) as follows,

$$\hat{\beta}_n = \min_{\beta \in \Theta} L_n(\beta) \tag{4}$$

where

$$L_n(\beta) = \frac{1}{n} \sum_{i=1}^{n} (\pi(x_i,\beta) - \hat{\pi}(x_i))^2 \tag{5}$$

and $\hat{\pi}(x)$ represents the kernel estimate of the marginal density function $\pi(x)$, *i.e.*,

$$\hat{\pi}(x) = \frac{1}{nh^d}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) \qquad (6)$$

where $K(\cdot)$ is a kernel function and $h = h_n$ is a smoothing parameter, bandwidth or window that does not depend on the data. Therefore, under H_0 the estimate of the Delta method of $\pi(x,\beta)$ is $\pi(x,\hat{\beta}_n)$ and the true and unknown marginal density function of x, $\pi(x)$, can be estimated consistently with the kernel estimate $\hat{\pi}(x)$ regardless to the correct specification of the parametric model.

If the estimates of $\pi(x)$ and $\pi(x,\beta)$ are $\hat{\pi}(x)$ and $\pi(x,\hat{\beta}_n)$ respectively, they can be substituted in the definition of I producing the following estimate of I,

$$I_n = \int (\hat{\pi}(x) - \pi(x,\hat{\beta}_n))^2\, \hat{\pi}(x)dx \qquad (7)$$

3. The limit distribution of I_n under H_0

The following assumptions are used to obtain the limit distribution of I_n.

Premise 1. *For any* $\beta \in \Theta$, $\sigma(\cdot,\beta)$ *is locally bounded and is Borel measurable.*

Premise 2. *Let be q a measure of induced probability over* R^d *by* X_0 *and* $\int_{R^d} \lambda\varphi dq = 0$ *for any bounded and continous function* $\phi(x)$ *in* R^d, *where* λ *is the infinitesimal generator created by the diffusion process* $\{X_t, t \geq 0\}$.

Premise 3. *The kernel function* $K(\cdot)$ *is a symmetric and bounded function in* R^d *that satisfies:*

$$\int |K(u)|\, du < \infty, \| u \|^d |K(u)| \to 0 \textit{ when } \| u \| \to \infty,$$

$$\int u_i K(u)du = 0, \int u_i u_j K(u)du = 2k\delta_{ij}, \textit{ for } 1 \leq i, j \leq d, k \in R^+$$

Premise 4. *The density function* $\pi(\cdot)$ *and its partial derivative of second order are bounded and uniformly continuous in BBR*^{d}.

Premise 5. *The smoothing parameter* $h = h_n$ *satisfie* $h \to 0$, $nh^d \to \infty$ *when* $n \to \infty$.

Premise 6. *The function of parametric density* $\pi(u,\beta)$ *and its partial derivative of second order respect* β *are uniformely bounded and uniformly continuous.* $\pi(u,\beta)$ *and its partial derivative of second degree respect x are bounded and uniformly continuous in R*^{d}*. Also,* $\int |D'\pi(u,\beta)|\,dx < \infty$, *where* $D'\pi(u,\beta)$ *is a px1 partial derivative vector of first degree of the function* $\pi(u,\beta)$ *respect* β.

Premise 7. *The succesion of observed data* $\{X_i, 1 \le i \le n\}$ *is strictly stationary.*

Premise 8. *Exist* $\beta_* \in \Theta$ *such that* $\hat{\beta}_n \to \beta_*$ *a.s and*

$$\hat{\beta}_n = \beta_* + \int_{-\infty}^{+\infty} \varphi_F(u)\,d(F_n(u) - F(u)) + o_p\left(n^{-\frac{1}{2}}\right)$$

$$= \beta_* + \frac{1}{n}\sum_{i=1}^{n}\left(\varphi_F(X_i) - E\varphi_F(X_i)\right) + o_p\left(n^{-\frac{1}{2}}\right) \qquad (8)$$

where $F(\cdot)$ and $\varphi_F(\cdot)$ are the function of cumulative density and the derivative of $F(\cdot)$ respectively, associated with the unknown function of marginal density $\pi(\cdot)$.

Premise 9. *The parameters space* $\Theta \subset R^d$ *is compact and* $E\left[(\partial\pi(u,\beta_0)/\partial\beta)(\partial\pi(u,\beta_0)/\partial\beta')\right]$ *has a full range.*

- Premise 1 guarantees the existence and uniqueness of a solution for the stochastic differential equation. Given that our model is a stochastical homogeneus differential equation in time and of Markovian type, also this premise is condition enough to ensure a non explosive solution.
- Premise 2 ensures that the solution process is stationary.
- Premises 3, 4 and 5 are used to obtain a limit distribution of the integral of the squared error for the estimate of the kernel density.

- Premises 6 and 8 are used to examine the effect of the estimate of $\pi(u,\beta_0)$ by $\tilde{\pi}(x)$ over the limit distribution I_n. In particular, the last term, $o_P(1)$, from the right side of the equation (8), can be guaranteed assuming that the derivative $\varphi_F(\cdot)$ is a cadlag bounded function.
- Premise 7 serves the purpose to limit the dependence of the discrete observations so that asymptotic theory can be used.
- Premise 9 ensures that the linear term in the Taylor expansion of a functional is not degenerate. If the linear term degenerates then the asyntotic distribution might be given by a term of greater degree present in the Taylor expansion.

From equation (7) we can write the squared integrable difference between $\hat{\pi}(x)$ and $\tilde{\pi}(x)$ as follows,

$$
\begin{aligned}
I_n &= \int(\hat{\pi}(x)-\pi(x))^2\,\hat{\pi}(x)dx + \int(\tilde{\pi}(x)-\pi(x))^2\,\hat{\pi}(x)dx \\
&\quad -2\int(\hat{\pi}(x)-\pi(x))(\tilde{\pi}(x)-\pi(x))\hat{\pi}(x)dx \\
&= I_{1n}+I_{2n}-2I_{3n} \qquad (9)
\end{aligned}
$$

Theorem 1. *Let be* $c(n)=(nh^d)^{-1}\int K^2(x)dx\int\pi^2(x)dx+\int(E\hat{\pi}(x)-\pi(x))^2)\pi(x)dx$, $\nabla\pi(x)=\sum_{i=1}^{d}\partial^2\pi(x)/\partial x_i^2$ *and we define*

$$
d(n)=\begin{cases} n^{1/2}h^{-2} & \text{si } nh^{d+4}\to\infty \\ nh^{d/2} & \text{si } nh^{d+4}\to 0 \end{cases}
$$

then, under the conditions established over $K(\cdot)$ and $\pi(\cdot)$ and assuming that $nh^{2d}\to\infty$, when $n\to\infty$, we have

$$
\begin{aligned}
d(n)(I_{1n}-c(n)) &= 2d(n)(n^2h^{2d})^{-1}\sum_{1\le i\le j\le n}H_n(X_i,X_j) \\
&\quad +2d(n)\int(\hat{\pi}(x)-E\hat{\pi}(x))(E\hat{\pi}(x)-\pi(x))\pi(x)dx+O_p(d(n)n^{-3/2}h^{-3d/2})
\end{aligned}
$$

where $H_n(X_i,X_j)=\int\left[K\left(\frac{x-X_i}{h}\right)-EK\left(\frac{x-X_i}{h}\right)\right]\left[K\left(\frac{x-X_j}{h}\right)-EK\left(\frac{x-X_j}{h}\right)\right]\pi(x)dx$

$$nh^{d/2}\left[\left(n^2h^{2d}\right)^{-1}\sum_{1\le i\le j\le n}H_n(X_i,X_j)\right]\xrightarrow{d}\sqrt{2}\sigma_2 N$$

$$n^{1/2}h^{-2}\int[\hat{\pi}(x)]^2\pi^2(x)dx-\{\int[\nabla^2\pi(x)]\pi^2(x)dx\}^2$$

where

$$\sigma_1^2=\int[\nabla^2\pi(x)]^2\pi^2(x)dx-\left\{\int[\nabla^2\pi(x)]\pi^2(x)dx\right\}^2$$

$$\sigma_2^2=\int\pi^4(x)dx\left\{\int\left[\int K(u)K(u+v)du\right]^2dv\right\}$$

$$d(n)(I_{1n}-c(n))\xrightarrow{d}\begin{cases}2k\sigma_1 N & \text{si } nh^{d+4}\to\infty\\ \sqrt{2}\sigma_2 N & \text{si } nh^{d+4}\to 0\end{cases}$$

Theorem 1 implies that the limit distribution of the integrable squared error of $\hat{\pi}(x)$, I_{1n} depends upon the amount of smoothing that is applied to the data. The limit distributions form Theorem 1 can be used to construct of proof for the null hypothesis $\pi(x)=\pi_0(x)$ against the alternative $\pi(x)\neq\pi_0(x)$, where $\pi_0(x)$ is a density function which is completely unknown. Furthermore, in order to Theorem 1 to be observed, the smoothing parameter h must satisfy $nh^{2d}\to\infty$ or $nh^{d+4}\to 0$ (under smoothing data) or $nh^{d+4}\to\infty$ (over smoothing data). So d must satisfy $d\le 3$ for under smoothing data. Nevertheless, when the null hypothesis is composite, that is, there are a finite number of unknown parameters, the last two terms of the right side of the equation (9) must be taken under consideration. For this purpose, the Taylor expansion and premise 8, under H_0, we have

$$\tilde{\pi}(x)-\pi(x,\beta_0)=D'\pi(x,\beta_0)\frac{1}{n}\sum(\varphi_F(X_i)-E\varphi_F(X_i))+o_p(n^{-1/2})$$

then we express I_{3n} as follows,

$$I_{3n}\equiv J_{1n}(\beta_0)+J_{2n}(\beta_0)$$

where

$$J_{1n}(\beta_0)=\int[\hat{\pi}(x)-E\hat{\pi}(x)][\tilde{\pi}(x)-\pi(x,\beta_0)]\hat{\pi}(x)dx \qquad (10)$$

and

$$J_{2n}(\beta_0)=\int[E\hat{\pi}(x)-\pi(x)][\tilde{\pi}(x)-\pi(x,\beta_0)]\hat{\pi}(x)dx \qquad (11)$$

The following lemma summarizes the behavior of $J_{1n}(\beta_0)$ and $J_{2n}(\beta_0)$, and therfore that of I_{3n}.

Lemma 1. *Under premises 1-9, for any* $0<\delta<1$ *we have*

a) $J_{1n}(\beta_0)=O_p(n^{-\delta})$ *and also* $n^{1/2}h^{-2}J_{2n}(\beta_0)\xrightarrow{d}k\sigma_3 N$ *where*

$$\sigma_3^2=\left\{\int D'_{\beta_0}\pi(x,\beta_0)\nabla^2\pi(x)dx\right\}V\left\{\int D_{\beta_0}\pi(x,\beta_0)\nabla^2\pi(x)dx\right\}$$

$$V=Var(\varphi_F(x_1))+\sum_{k=1}^{\infty}Cov(\varphi_F(X_1),\varphi_F(X_{1+k}))$$

consequently $I_{3n}=J_{2n}(\beta_0)+O_p(n^{\delta})$ *y* $n^{1/2}h^{-2}I_{3n}\xrightarrow{d}k\sigma_3 N$

b) If the null hypothesis is satisfied, then $I_{2n}=O_p(n^{-1})$ and also $n^{1/2}h^{-2}I_{3n}\xrightarrow{d}k\sigma_3 N$

If the null hypothesis is satisfied from equation (11) and Lemma 1 we obtain the following expression,

$$I_n=I_{1n}-2J_{2n}(\beta_0)+O_p(n^{-\delta}) \qquad (12)$$

this equation lays down that the parametric estimate of $\pi(x,\beta_0)$ under the effects of the null hypothesis is the limit distribution of I_n only through $J_{2n}(\beta_0)$. Therefore if the first term of the right side of equation (12) dominates the second term asimptotically, which is the case for under smoothed data, the parametric estimate of $\pi(x,\beta_0)$ under the null hypothesis does not affect the limit distribuition of I_n as Yacine Aït-Sahalia (1996) demonstrates for $d=1$. Nevertheless, if data are over smoothed, the limit distribution of A will be affected by non parametric estimates as the following theorem shows.

Theorem 2. *Under the assumptions of Lemma 1, if the null hypothesis is satisfied and* $\frac{d+\delta}{2(d+4)}\leq\delta<1$, *then we have*

$$d(n)(I_n - c(n)) \xrightarrow{d} \begin{cases} 2k\sigma_4 N & \text{si } nh^{d+4} \to \infty \\ \sqrt{2}\sigma_2 N & \text{si } nh^{d+4} \to 0 \end{cases}$$

where $c(n), d(n), \sigma_1^2, \sigma_2^2$ are the same as in Lemma 1, $\sigma_4^2 = V_1 + 2V_2$ and:

$$V_1 = Var(\nabla^2\pi(X_1)) - 2\int D'\pi(y,\beta)\nabla^2\pi(y)dy$$

$$+\int D'\pi(y,\beta)\nabla^2\pi(y)dy \times Var(\varphi_F(X_1)) \times \int D\pi(y,\beta)\nabla^2\pi(y)dy \qquad (13)$$

$$V_2 = \sum_{k=2}^{\infty} Cov(\nabla^2\pi(X_1), \nabla^2\pi(X_k)) + (\int D'\pi(y,\beta)\nabla^2\pi(x)dx)$$

$$\times \sum_{k=2}^{\infty} Cov(\varphi_F(X_1), \varphi_F(X_k)) \times (\int D\pi(y,\beta)\nabla^2\pi(y)dy) \qquad (14)$$

$$-\int D'\pi(y,\beta)\nabla^2\pi(y)dy \times \sum_{k=2}^{\infty} Cov(\nabla^2\pi(X_1), \varphi_F(X_k))$$

$$-\int D'\pi(y,\beta)\nabla^2\pi(y)dy \times \sum_{k=2}^{\infty} Cov(\nabla^2\pi(X_k), \varphi_F(X_1)) \qquad (15)$$

Corollary 1. *Under the assumptions of Lemma 1, and if the null hypothesis is satisfied, we have*

a) *If* $nh^{d+4} \to \infty$, *then* $n^{1/2}h^{-2}[I_n - \int [E\hat{\pi}(x) - \pi(x,\beta_0)]^2 \hat{\pi}(x)dx] \xrightarrow{d} 2k\sigma_4 N$

b) *If* $nh^{d/2+4} \to 0$, *then* $nh^{d/2}[I_n - \frac{1}{nh^d}\int K^2(u)du \int \pi^2(u)du] \xrightarrow{d} \sqrt{2}\sigma_2 N$

4. Statistical test and its asymptotical distributions

The following statistical tests are based upon the estimates of σ_2^2 and $\sigma_2^4 \cdot \sigma_2^2$ can be consistently estimated by $\frac{1}{n}\sum_{i=1}^{n}\hat{\pi}^3(X_i)\int[\int K(u)K(u+v)du]^2 dv$. To obtain consistent estimates of σ_4^2, we define

$$\widehat{Var}(\nabla^2\pi(X_1)) = \int(\nabla^2\hat{\pi}(x))^2\hat{\pi}(x)dx - \{\int(\nabla^2\hat{\pi}(x))\hat{\pi}(x)dx\};$$

$$\widehat{\sum_{k=1}^{\infty}} Cov(\varphi_F(X_1), \varphi_F(X_{k+1})) =$$

$$\sum_{k=1}^{G_n}\frac{2}{1-k+G_n}\left\{\frac{1}{n}\sum_{i=1}^{n-k}\varphi_{\hat{F}_n}(X_{i+k}) - \left(\frac{1}{n}\sum_{i=1}^{n}\varphi_{\hat{F}_n}(X_i)\right)^2\right\};$$

$$\widehat{\sum_{k=1}^{\infty}} Cov(\nabla^2\pi(X_1), \nabla^2\pi(X_{k+1})) =$$

$$\sum_{k=1}^{G_n}\frac{2}{1-k+G_n}\left\{\frac{1}{n}\sum_{i=1}^{n-k}\nabla^2\hat{\pi}(X_i)\nabla^2\hat{\pi}(X_{i+k}) - \left(\frac{1}{n}\sum_{i=1}^{n}\nabla^2\hat{\pi}(X_i)\right)^2\right\};$$

$$\widehat{\sum_{k=1}^{\infty}} Cov(\nabla^2\pi(X_1), \varphi_F(X_{k+1})) =$$

$$\sum_{k=1}^{G_n}\frac{2}{1-k+G_n}\left\{\frac{1}{n}\sum_{i=1}^{n-k}\nabla^2\hat{\pi}(X_i)\varphi_{\hat{F}_n}(X_{i+k}) - \frac{1}{n^2}\sum_{i=1}^{n}\nabla^2\hat{\pi}(X_i)\sum_{i=1}^{n}\varphi_{\hat{F}_n}(X_{i+k})\right\};$$

$$\widehat{\sum_{k=1}^{\infty}} Cov(\nabla^2\pi(X_{1+k}), \varphi_F(X_1)) =$$

$$\sum_{k=1}^{G_n}\frac{2}{1-k+G_n}\left\{\frac{1}{n}\sum_{i=1}^{n-k}\nabla^2\hat{\pi}(X_{i+k})\varphi_{\hat{F}_n}(X_i) - \frac{1}{n^2}\sum_{i=1}^{n}\nabla^2\hat{\pi}(X_{i+k})\sum_{i=1}^{n}\varphi_{\hat{F}_n}(X_i)\right\},$$

where G_n is a delay of a selected truncate such that $\underline{\lim} G_n = +\infty$ y $G_n = O(n^{1/3})$. This is the estimate of the spectral density at zero see (Yacine Aït-Sahalia, 1992). It is possible to find consistent estimates of V_1 and V_2 substituing the previous estimates in equations (13) and (14). Using the results obtained tests values for $H_0 : \pi(x) = \pi(x, \beta_0)$ against the general alternative $H_1 : \pi(x) \neq \pi(x, \beta)$ for any $\beta \in \Theta$.

5. Extension of the statistical test of Yacine Aït-Sahalia

Now we define the statistical tests as follows

$$T_{1n} = \frac{n^{1/2}h^{-2}\{I_n - \int [K_h * \tilde{\pi}(x) - \tilde{\pi}(x)]^2 \hat{\pi}(x)dx\}}{2k\hat{\sigma}_4} \quad (16)$$

$$T_{2n} = \frac{h^{-d/2}[nh^d I_n - \int K^2(z)dz \int \pi^2(z)dz]}{\sqrt{2}\hat{\sigma}_2} \quad (17)$$

where $K_h * \tilde{\pi}(x) = h^{-d} \int K\left(\frac{x-u}{h}\right)\tilde{\pi}(u)du$

Theorem 3. *Under premises 1 to 9, if H_0 is satisfied and* $\frac{d+\delta}{2(d+4)} \leq \delta < 1$ *then*

a) *If* $nh^{d/2+4} \to \infty$ *then* $T_{1n} \xrightarrow{d} N$

b) *If* $nh^{d/2+4} \to 0$ *then* $T_{2n} \xrightarrow{d} N$

Theorem 3 stablishes that a statistical test can be constructed to a significant level α to prove H_0 against H_1 that correspond to a diverse amount of smoothing applied to data. Also it is pointed that the statistical test T_{2n} is an extension Yacine Aït-Sahalia (1996) test that can be applied to two factors model, by the small value of the smoothed parameter.

6. Statistical test for skew correction

As established before, the skew introduced by the kernel estimate of the density function has a significant influence in the test of H_0 based upon the estimate of the squared error integral. The skew effects are reflected upon the restriction of the smoothing parameter, which can decrease to zero and cannot decrease too fast or too slow when the size of the sample tends to infinite. Based upon different restrictions in the smoothing parameter, two different tests can be obtained. Now, we present a statistical test that applies to any diffusion process of finite dimension, but also does not depend on the amount of smoothing applied to data. Under H_0, the expected value of the non parametric estimate $\hat{\pi}(x)$ is $E\hat{\pi}(x) = K_h * \pi(x, \beta_0) = \int h^{-d} K(\frac{x-u}{h}) \pi(u, \beta_0) du$. Therefore, to eliminate the skew effects, we can construct an statistical test for adjusted skew based upon the weighted integral of the square difference divided by the non parametric estimate of the marginal density function implied in the data and the estimate of smoothed kernel of the function of marginal density implied by the parametric model,

$$J_n = \int [\hat{\pi}(x) - K_h * \tilde{\pi}(x)]^2 dx$$

We can decompose J_n as follows

$$J_n = \int [\hat{\pi}(x) - E\tilde{\pi}(x)]^2 \hat{\pi}(x) dx + \int [E\tilde{\pi}(x) - K_h * \tilde{\pi}(x)]^2 \hat{\pi}(x) dx$$

$$+ 2\int [\hat{\pi}(x) - E\tilde{\pi}(x)][E\tilde{\pi}(x) - K_h * \tilde{\pi}(x)] \hat{\pi}(x) dx$$

$$= \int [\hat{\pi}(x) - E\tilde{\pi}(x)]^2 \hat{\pi}(x) dx + J_{11} + 2J_{12}$$

Theorem 4. *Under premises 1 to 9, if H_0 is observed and* $\frac{d+\delta}{2(d+\delta)} \leq \delta < 1$, *then*

$$T_n = \frac{nh^{d/2}[J_n - \frac{1}{nh^d} \int K^2(z) dz \int pi^2(z) dz]}{\sqrt{2}\hat{\sigma}_2} \xrightarrow{d} N \qquad (18)$$

where $\hat{\sigma}_2^2$ is defined as before.

The asymptotic distribution of the adjusted skew test is the same regardless of the nature of the data as under or over smoothing. It is also important to note that the adjusted skew test can be applied to de marginal density function of any diffusion process of finite dimensions.

7. Parametric models of two factors

There is a general acceptance that one factor models of the term structure cannot succesfully explain several characteristics of bond returns. The reason for this is that the factor of the term structure and the returns and outcomes of the bonds must be perfectly correlated. As we have seen in the Introduction, several authors accept that moved by two uncertainty factors at least. So actual research focus into models of term structure that uses two state variables or stochastic factors. These models are mutually exclusive and do create different values when used to predict the price derived from different rates. From the two model research we can find that of Brennan y Schwartz (1979) and that of Schaefer Schwartz (1984) that create their specifications based upon stochastic processes using two interest rates. Before we use the specification test of Brennan-Schwartz and the model of Schaefer-Schwartz, we will find the marginal density functions of such models.

7.1 Models of term structure of Brennan-Schwartz and Schaefer-Schwartz

7.1.1 Model of Term Structure of Brennan-Schwartz

Brennan and Schwartz (1979) developed a straddle model for term structure of interest rates under the assumption that all the term structure can be expressed at any time in terms of the outcomes of instruments of short and long terms. If r defines the instant interest rate and l is the long term interest rate its model can be expressed as,

$$dr_t = r[\alpha(\ln l_t - p - \ln r_t) + 1/2\sigma_r^2]dt + (\sigma_r \cdot r_t)dW_1 \qquad (19)$$

$$dl_t = [l_t(k(\theta - \ln l_t) + 1/2\sigma_l^2]dt + (\sigma_l \cdot l_t)dW_2 \qquad (20)$$

where p is the difference between the middle levels of 1nl and 1nr; α is the adjustment coefficient for the speed in which 1nr returns to $(\ln l - p)$; k is the adjustment coefficient for the speed in which 1nl returns to the mean level of θ; and $E(dW_1 dW_2) = \delta$ Using the Ito´s lemma, the previous equations can be written as,

$$d\ln r_t = \alpha((\ln l_t - p) - \ln r_t)dt + \sigma_r dW_1 \qquad (21)$$

$$d\ln l_t = k(\theta - \ln l_t)dt + \sigma_l dW_2 \qquad (22)$$

To put into practice our test, we have to estimate the unknown parameters assumed under the null hypothesis. The bivariate processes for 1nr and 1nl is stationary in a joint form if α and k are greater than zero. Also, the solution of equations (21) and (22) is a Gaussian process if and only if the initial value is constant or normaly distributed. As a matter of fact the solution of the linear stochastic differential equation (21) and (22) can be expressed as

$$X_t = \mu + \exp(A(t-t_0)) \cdot (X_{t_0} - \mu) + \int \exp(A(t-s))B^{1/2} dW \qquad (23)$$

where

$$X_t = \begin{pmatrix} \ln r_t \\ \ln l_t \end{pmatrix}, \quad \mu = \begin{pmatrix} \theta - p \\ \theta \end{pmatrix}, \quad A = \begin{pmatrix} -\alpha & \alpha \\ 0 & -k \end{pmatrix}, \quad B = \begin{pmatrix} \sigma_r^2 & 0 \\ 0 & \sigma_l^2 \end{pmatrix}, \quad dW = \begin{pmatrix} dW_1 \\ dW_2 \end{pmatrix}$$

Otherwise, if α and k are greater than zero, then the solution follows a Gaussian stationary process. For a stochastic stationary process, the marginal density function equals the density of the initial observation, therefore, we can obtain a stationary marginal density function of this process once we have the density function of the initial random variable. The marginal density can be obtained with equation (22) as follows,

$$f(X) = \frac{1}{2\pi |\Sigma|^{1/2}} \exp(-(X-\mu))' \Sigma^{-1} (X-\mu)) \qquad (24)$$

where

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \quad X = \begin{pmatrix} \ln\ r_t \\ \ln\ l_t \end{pmatrix}$$

$$\Sigma_{11} = \frac{\sigma_r^2}{2\alpha} + 2\sigma_r^2\sigma_l^2\delta(\frac{\alpha}{\alpha-k})(\frac{1}{(\alpha+k)-\frac{1}{2\alpha}}) + \sigma_l^2(\frac{\alpha}{(\alpha-k)})^2(\frac{1}{2k} - \frac{2}{(\alpha+k)} + \frac{1}{2\alpha})$$

$$\Sigma_{12} = 2\sigma_r\sigma_l\delta(\frac{1}{\alpha+k}) + \sigma_l^2(\frac{\alpha}{\alpha-k})(\frac{1}{2k} - \frac{1}{(\alpha+k)})$$

$$\Sigma_{21} = \Sigma_{12}$$

$$\Sigma_{22} = \frac{\sigma_l^2}{2k}$$

7.1.2 Model of term structure of Schaefer-Schwartz

Schaefer and Schwartz (1984) use the same information of the interest rate as Brennan and Schwartz (1979), but express its model in terms of the long term interest rate and the difference between it and the short term interest rate. This is a redefinition of the variables that make possible to obtain an analytical solution for the problem of valuation. Furthermore, Schaefer and Schwartz assume that the difference follows the process of Ornstein-Uhlenbeck (a process of the reversion of the mean and the constant diffusion function). In financial research, the Ornstein-Uhlenbeck process has been used to model the short term interest rate. Nevertheless, it is reasonable to assumme a better probability that it is the difference and not the short term rate what follows this kind of process due to the fact that the model allows negative values. The specific form of the stochastic process of Schaefer and Schwartz is,

$$dS_t = \beta(\alpha - S_t)dt + \gamma dW_1 \tag{25}$$

$$dl_t = \mu_2(S_t, l_t, t)dt + \sigma\sqrt{(l)}dW_2 \tag{26}$$

where dW_1 and dW_2 are standard Wiener processes with $E[dW_1] = E[dW_2] = 0, dW_1^2 = dW_2^2 = dt$. As in the model of Brennan-Schwartz, Schaefer and Schwartz assume that the difusion function of the console rate depends on its level. The tendency function of the console rate stays as in its general form. The process of Ornstein-Uhlenbeck, $dS_t = \beta(\alpha - S_t)dt + \gamma dW_1$, has a transition density function given by

$$f(S_t = s, t \mid S_{t_0} = s_0, t_0) = \frac{1}{\sqrt{2^2(t)}} \exp\{-\frac{(s - \alpha - (s_0 - \alpha)e^{-\beta(t-t_0)^2})}{2\nu^2(t)}\}$$

where

$$v^2(t) = \frac{\sigma^2}{2\beta}[1 - e^{-2\beta(t-t_0)}].$$

Let us suppose that the spread process shows the property of reversion to the mean ($\beta > 0$), then when $t_0 \to -\infty$ or $(t - t_0) \to \infty$, the marginal density of the stochastic process is invariant through time, *i.e.*, the spread process is stationary and the marginal density function can be expressed as follows:

$$\sqrt{\frac{\pi\sigma^2}{\beta}} \exp\left\{-\frac{\beta(s-\alpha)^2}{\sigma^2}\right\}$$

For stationary diffusion processes, the only stationary processes with explicit transition density functions are those who have the linear functional specification for the tendency function and the specification of the cuadratic function for the diffusion function.

7.2 Empirical tests and specification analysis of the term structure of the models of Brennan-Schwartz and Schaefer-Schwartz

The time series used in this analysis include daily observations of Cetes for 1 day to 10 years of maturity. The analyzed sample period covers from july 9th of 2002 up to december 7th of 2009, which comprises 1 867 observations for each term. Based upon the stationary density functions derived previously, we can estimate the unknown parameters, reducing to a minimal expression the cuadratic error between the estimate of the non parametric density and the density of the parametric model. The results of the parametric estimate of both models are shown in tables 1 and 2.

The estimates of $(\int K^2(z)dz)(\int \pi^2(x)dx)$ and $(\int(\int K(u)K(u+x)du)^2)(\int \pi^4(x)dx)$ are respectively:

$$\int K^2(z)dz)(\frac{1}{n}\sum_{i=1}^{n}\hat{\pi}_0(\log(r_i),\log(l_i))$$

$$(\int(\int K(u)K(u+x)du)^2 dx)(\frac{1}{n}\sum_{i=1}^{n}\hat{\pi}_0^3(\log(r_i),\log(l_i)))$$

Tabla 1
Parameter values estimated with the Brennan and Schwarts model

Parameter	α	k	θ	p	σ_r	σ_l	δ
Estimate	0.98725	0.23873	0.10191	- 0.01533	0.4956	0.21096	0.24627

Tabla 2
Values of the parameters estimate with the Schaefer and Schwartz model

Parameter	α	β	σ
Estimate	1.7612	1.50114	0.1796

Taking a critical value of 6.32 to obtain a test level of 0.05%. For the normal kernel, the two kernel constants are:

$$\left(\int K^2(z)dz\right)=\begin{cases}\frac{1}{2\sqrt{\pi}} & if\ d=1\\ \frac{1}{4\pi} & if\ d=2\end{cases},\ \int\left(\int K(u)K(u+x)du\right)^2 dx=\begin{cases}\frac{1}{2\sqrt{2\pi}} & if\ d=1\\ \frac{1}{8\pi} & if\ d=2\end{cases}$$

To estimate the marginal density, the smoothed parameter $h_n = h$ is selected such that $\lim_{n\to\infty} nh^2 = \infty$ and $\lim_{\to\infty} nh^5 = 0$. The quality of the estimate of the density depends mainly on the selection of the smoothing parameter rather than on the kernel. We select $h = c_n n^{-1/5}$, where $c_n = c(\ln(n)$. c is choosed such that it minimizes the mean cuadratic integrable error of the estimate. The results of this test are shown in Table 3. Both models Brennan-Schwartz and Schaefer-Schwartz are rejected with a significance level of 0.05%. The main reasons for rejection of both models are: first, the constant diffusions limit volatilities to be uniformly increasing. Second, if a model has a linear especification for the drift and also has a constant diffusion, as in the case of Brennan-Schwartz and Schwartz-Schaefer, the parameters founded in the tendency and difussion functions might not be homogeneous, that is, shifts in the economical regime implies the possible lack of stationarity of the parameters during the modeling process.

TABLA 3
Tests of parametric specification of two factor models of time continuum

Brennan-Schwartz					
Function of Tendency	Function of Diffusion	Statistical Test I	Statistical Test II	Critical Value	Result
		T_{2n}	T_n		
$r[\alpha(\ln l - p - \ln r) + 1/2\sigma_r^2]$	$\sigma_r \cdot r$	46.2792		1.645	Reject
$[l(k(\theta - \ln l) + 1/2\sigma_l^2)]$	$\sigma_l \cdot l$		21.48973	1.645	Reject
Schaefer-Schwartz					
$\beta(\alpha - S_t)\mu(s,l,t)$	γ	75.07685		1.645	Reject
	$\sigma\sqrt{(l)}$		123.97681	1.645	Reject

8. Conclusions

It was proposed a statistical test for the specification of parametric models of two factors. We present three different tests. The first two are based upon a comparison between the estimate of the kernel density of the unknown density function and the estimate of $\pi(x,\beta)$ by the Delta method. The last test is based upon the idea of comparison between the estimate of the kernel density and the parametric model of the smoothed density kernel to avoid skew effects. The advantage of the last test respect the first two tests is its validity for under smoothed data and also over smoothed data. This test can be applied in many financial process which of research importance. Particularly, this test was applied to determine if the dynamic of the term structure of Cetes can be modeled from the assumptions of two models, that of Brennan-Schwartz (1979) and that of Schaefer-Schwartz (1984), the test results show that both continuous models are rejected at a 5% level as accurate describing the dynamic of Cetes. Finally the causes for rejection were analyzed for both models.

References

Aït-Sahalia, Y. (1996), "Testing Continuous-Time Models of The Spot Interest Rate". *Review of Financial Studies*, vol. 9, num. 2, pp. 385-426.

Aït-Sahalia, Y. (l992), "The Delta Method for Nonparametric Kernel Functionals", Working paper, MIT.

Bakshi G. S. and Z. Chen (1996), "Inflation, asset prices and the term structure of interest rates in monetary economies", *The Review of Financial Studies* 9, pp. 241-275.

Buraschi, A. and A. Jiltsov (2001), "Time-varying inflation risk premia and the expectations hypothesis: a monetary model of the treasury yield curve". *Journal of Financial Economics*.

Chen, L. (1996), "Stochastic mean and stochastic volatility-a three factor model of the term structure of interest rates and its application in derivatives pricing and risk management", Financial Markets, *Institutions and Instruments*, 5, pp. 1-87.

Chen, R. R. and L. Scott (1993), "Maximum Likelihood Estimation for a Multifactor Equilibrium Model of the Term Structure of Interest Rates", *Journal of Fixed Income*, 3, pp. 14-31.

Cortázar, G. and E. Schwartz (1994), "The valuation of commodity-contingent claims", *Journal of Derivatives*, vol. 1, num. 4, pp. 27-39 77.

Cox, J. C., J. E. Ingersoll and S. A. Ross (1985), "A Theory of the Term Structure of Interest Rates", *Econometrica*, 53, pp. 385-407.

Cultberson, J. (1957), "The term structure of interest rates", *Quarterly Journal of Economics*, vol. 71, pp. 485-517.

Dai, Q. and K. Singleton (2002), "Expectation puzzles, time-varying risk premia, and affine models of the term structure", *Journal of Financial Economics*, 63, 415-441.

D´Ecclesia, R. L. and S. A. Zenios (1994), "Risk Factor Analysis and Portfolio Immunization in the Italian Bond Market", *Journal of Fixed Income*, vol. 4, num. 2, pp. 51-58.

Duffie, D. and R. Kan (1996), "A Yield-Factor Model of Interest Rates", *Mathematical Finance*, vol. 6, num. 4, pp. 379-406.

Dybvig, P. H. (1989), "Bond and Option Pricing Based on The Current Term Structure", Working paper, Washington University.

Elton, E. J., M. Gruber and R. Michaely (1990), "The Structure of Spot Rates and Immunization", *The Journal of Finance*, vol. 45, num. 2, pp. 629-642.

Fisher, I. (1896), "Appreciation and interest", *AEA Publications*, vol. 3, num. 11, pp. 331-442.

Hicks, J. (1939), *Value and Capital*, Segunda Edición, Londres: Oxford University Press.

Knez, J., R. Litterman and J. Scheinkman (1996), "Exploration into Factors Explaining Money Markets Returns", *Journal of Finance*, vol. 49, num. 5, pp. 1861-1881.

Litterman, R. and J. A. Scheinkman (1991), "Common Factors Affecting Bond Returns", *Journal of Fixed Income*, 1, pp. 54-61.

Lund, J. (1997), *Non-Linear Kalman Filtering Techniques for Term Structure Models*, Aarhus School of Business.

Pearson, N. D. and T. S. Sun (1994), "Exploiting the Conditional Density in Estimating the Term Structure:An Application to the Cox, Ingersoll, and Ross Model", *Journal of Finance*, 49, pp. 1279-1304.

Steeley, J. M. (1990), "Modelling the dynamics of the term structure of interest rates", *Economic and Social Review*, vol. 21, num. 4, pp. 337-361.

Svensson, L. (1994), "Estimating and interpreting forward interest rates: Sweden 1992-94", CEPR Discussion Paper 1051.

Vasicek, O. A. (1997), "An equilibrium Characterization of the Term Structure", *Journal of Financial Economics*, 5, pp. 177-188.79.