Continuous Time Models of Interest Rate with Jumps:
Testing the Mexican Data (1998-2006)

José Antonio Núñez Mora *
José Luis de la Cruz**
Elizabeth Ortega***

Resumen
Como una extensión del artículo de Núñez, De la Cruz y Ortega (2007), se prueban diferentes modelos paramétricos con saltos, usando la metodología desarrollada por Ait-Sahalia y Peng (2006), basada en la función de transición. Los datos analizados son la tasa de interés Mexicana en el período 1998-2006. Los resultados confirman que la mayoría de los modelos no tienen suficiente precisión para describir los datos de México.

Abstract
As an extension of the article by Núñez, De la Cruz and Ortega (2007), different parametric models with jumps are tested with the methodology developed by Ait-Sahalia and Peng (2006), based on the transition function. Data analyzed are the Mexican interest rates in the period 1998-2006. The results confirm that most of interest rate models do not have enough precision in order to describe the Mexican data.

PALABRAS CLAVE: Nivel de remuneraciones, distribución contrafactual, selección ocupacional, sector formal e informal
CLASIFICACIÓN JEL: J24, J31, J44, O15, O17

Recibido el 13 de febrero de 2009. Revisado y aceptado el 25 de junio de 2009
**Introduction**

In this paper we applied the methodology developed in Ait-Sahalia, Fan and Peng (2006) in order to evaluate different models for the Mexican interest rate data (CETES). Núñez, De la Cruz and Ortega (2007) analyzed the models proposed by Ait-Sahalia (1996), which have tried to describe the empirical evolution of the interest rate. Their results allow to affirm that the models of interest rate shown in that paper were unable to describe the data of the Mexican CETES. Nevertheless, in empirical applications could be some misspecification problems of parametric models. To avoid pricing or hedging strategies mistakes, Ait-Sahalia, Fan and Peng (2006) considered directly the divergence measures among the transition density function under the null hypothesis, and that under the alternative model.

Basically, Ait-Sahalia, Fan and Peng (2006), proposed an alternative specification test for the transition density of the process. In their proposal the authors compare directly the parametrically and nonparametrically estimated transition densities.

The typical dynamics specified is the jump diffusion process \( X_t \), defined on a probability space \( (\Omega, \mathcal{F}, P) \) with filtration \( \{\mathcal{F}_t\} \) (Protter, 2005)

\[
dX_t = \mu(X_t, \theta)dt + \sigma(X_t, \theta)dW_t + J_t dN_t
\]  

(1)

Where \( X_t \) is the state vector and \( W_t \) is a standard Brownian motion. \( \theta_t \in \mathbb{R}^n \) is a finite dimensional parameter to be estimated. The functions \( \mu(., \theta), \sigma(., \theta) \), and \( N \), are respectively, the drift, diffusion and pure jump processes. \( N \) has stochastic intensity \( \lambda(X_t, \theta) \) with jumps of size 1. \( J_t \) is
independent of $F_t$ and has probability density $\nu(., \theta)$ with non empty support interior $C$.

1.1 About jumps
Many studies (see for example Das 1998, 1999, Andersen and Lund, 1997) have demonstrated that diffusions cannot generate nonnormalities of the interest rate data. In a classical diffusion model,

$$dX_t = \mu(X_t, \theta)dt + \sigma(X_t, \theta)dW_t$$

the information structure generated is $F_t = \sigma(B_s, s \leq t)$, with $B_s$ a Brownian motion and the increments of the random variable are approximately normal if the time interval of observation is small. No surprising events are possible in these stochastic environment, (see Huang, 1985; Johannes, 2004).

In particular, large changes in the Mexican data were observed for long time. Nowadays, the existence of significant movements in the interest rates can be found, but not with the same intensity as in past decades.

Figure 1 shows the changes in CETES 28 days, and it can be observed the presence (frequently) of spikes, here interpreted as jumps. So, it is possible that the inclusion of jumps in the stochastic differential equations can serve to describe in a better way the Mexican data.
Interest rate is a key variable in economics, is subject to macroeconomic shocks, like inflation, default of economic agents, political an economics government decisions, variation in the exchange rate, etc. At the same time, agents take decisions about projects, investments, credit, etc, based on a macroeconomic scenario of stability of this variable.

From statistical point of view, the presence of jumps is very useful capturing the kurtosis excess in the data. This excess of kurtosis cannot be captured by the classical model of diffusion. It is known that many of the typical models used in the financial industry to explain interest rates behavior are constructed on a Gaussian component only. The new generation of models (Jarrow and Turnbull, 1995; Das, 1998; Segundo, 2002; Schonbucher, 2003) have added jumps. A very common way to specify these jumps is with a Poisson random variable.

As explained in Núñez, De la Cruz and Ortega (2007), there is no theoretical rationale for the election of the parametric drift and diffusion, and so, in this paper some of the most relevant interest rate models with jumps from
literature have been tested to prove if they can describe the Mexican CETES dynamics.

In the financial environment, several derivatives are based or strongly related to the interest rate, and therefore is an important factor in pricing of such financial instruments.

The paper is organized as follows. Section 2 presents an overview of the Methods, developed in Ait-Sahalia, Fan and Peng (2006) and its assumptions. Section 3 presents the application and results from the empirical research. Section 4 presents conclusions.

II. Methods and Assumptions

Definition 1. The transition probability density \( p(\Delta, y|x, \theta) \), when it exists, is the conditional density of \( X_{t+\Delta} = y \) given \( X_t = x \).

Assumption 1. The variance matrix \( V(x) \) is positive definite for all \( x \) in the domain of the process \( X \).

Assumption 2. The stochastic differential equation (1) has a unique solution. The transition density \( p(\Delta, y|x, \theta) \) is continuously differentiable with respect to \( \Delta \), twice differentiable with respect to \( x \) and \( y \).

Assumption 3. The boundary of the process \( X \) is unattainable.

Assumption 4. \( \nu(\cdot), \mu(\cdot), \sigma(\cdot) \) and \( \lambda(\cdot) \) are infinitely differentiable almost everywhere in the domain of \( X \).
Proposition 1. Under assumption 2, the transition density satisfies the backward and forward Kolmogorov equations given by

$$
\frac{\partial}{\partial \Delta} p(\Delta, y|x) = A^B p(\Delta, y|x)
$$

(2)

$$
\frac{\partial}{\partial \Delta} p(\Delta, y|x) = A^F p(\Delta, y|x)
$$

(3)

where the infinitesimal generators $A^B$ and $A^F$ are defined as in Protter(2005).

Using the backward and forward equations, it can be demonstrated that the transition density has the form,

$$
p(\Delta, y|x) = \Delta^{-\frac{n}{2}} \exp \left[ - \frac{C^{(-1)}(x, y)}{\Delta} \sum_{k=0}^{m} C^{(k)}(x, y) \Delta^k + \sum_{k=1}^{m} D^{(k)}(x, y) \Delta^k \right]
$$

(4)

In (4) functions $C^{(k)}(x, y)$ and $D^{(k)}(x, y)$ must be determined.

As showed in Ait-Sahalia (2006) an approximation of order $m > 0$ is obtained

$$
p^{(m)}(\Delta, y|x) = \Delta^{-\frac{n}{2}} \exp \left[ - \frac{C^{(-1)}(x, y)}{\Delta} \sum_{k=0}^{m-1} C^{(k)}(x, y) \Delta^k + \sum_{k=1}^{m} D^{(k)}(x, y) \Delta^k \right]
$$

(5)

The term $\Delta^{-\frac{n}{2}} \exp \left[ - \frac{C^{(-1)}(x, y)}{\Delta} \sum_{k=0}^{m-1} C^{(k)}(x, y) \Delta^k$ captures the behavior of $p(\Delta, y|x)$ at $y$ near $x$, and the term $\sum_{k=1}^{m} D^{(k)}(x, y) \Delta^k$ captures the tail behavior of $p(\Delta, y|x)$.
From theorem 1 below, the coefficients $C^{(k)}$ and $D^{(k)}$ can be founded (Ait-Sahalia, 2006).

**Theorem 1.** The backward equation imposes the following restrictions,

i.  
$$C^{(-1)}(x, y) = \frac{1}{2} \left[ \int_{s} \sigma(s)^{-1} ds \right]^2$$

ii.  
$$C^{(0)}(x, y) = \frac{1}{\sqrt{2\pi} \sigma(y)} \exp \left[ - \int_{s} \frac{\mu(s)}{\sigma^2(s)} ds \right]$$

iii.  
$$C^{(k+1)}(x, y) = \int_{s} \sigma^{-1}(s) ds \times \int_{s} \exp \left( - \int_{s} \frac{\sigma'(u)}{2\sigma(u)} - \frac{\mu(u)}{\sigma^2(u)} \right) du \cdot \sigma^{-1}(s) \left( \int_{s} \sigma^{-1}(u) du \right)^k \cdot \left[ \lambda(s) - L^k \right] C^{(k)}(s, y) ds$$

for $k \geq 0$

Where

$$L^k (\mathcal{G}(.)) = \sum_{i=1}^{n} \mu_i(x) \frac{\partial}{\partial x_i} \mathcal{G}(\cdot) + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} v_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} \mathcal{G}(\cdot)$$

iv.  
$$D^{(0)}(x, y) = \lambda(x) \nu(y - x)$$

v.  
$$D^{(k+1)}(x, y) = \frac{1}{1+k} \times \left[ A^k D^{(k)}(x, y) + \sqrt{2\pi} \lambda(x) \sum_{r=0}^{k} \left( \frac{M^r_x}{2^r (2r)!} \frac{\partial^{2r}}{\partial w^{2r}} g_{k-r}(x, y, w) \right) \right] \text{ para } k > 0.$$
where

\[ g_k(x, y, w) = C^{(k)}(w^{-1}(w), y) \cdot \nu(w^{-1}(w) - x) \cdot \sigma(w^{-1}(w)) \]

and

\[ M^{1}_{2r} = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\left(-\frac{s^2}{2}\right)s^{2r}ds \quad \text{and} \quad w_B(x, y) = \int_{y}^{x} \sigma(s)^{-1}ds \]

Explanation of Theorem 1:
The coefficients in the pure diffusive case, resulting in a system of equations, can be solved with the initial values \((k=0, 1)\), \textit{i.e.} (i) and (ii). The higher order terms are including the effect of the jumps are generated recursively from these initial values. The terms are obtained in the same recursive way as in Yu (2007).

### III. Results

The methodology is applied on a set of classical models with jumps, which are showed in Table 1. As we have explained, obtaining a jump model explaining the dynamic of interest rate would be very useful in the financial sector. From a theoretical point if view, description of jumps related to bonds and bond option prices can be found for example in Duffie and Kan (1996), Baz and Das (1996) and Chacko (1997). Das (1999, 1998) has tested models with jumps for interest rates in a very formal way, but with some problems about the possibility of negative rates, a nonsense financial result, but permitted in the statistical point of view. In fact, Das (2001) studied data with a Vasicek model adding jumps.
Among the most important models of interest rates we have:

Vasicek (1977). This model was developed under the assumptions

(i) The interest rate follows a diffusion process
(ii) The price of a zero coupon bond depends on the short rate only, and
(iii) There is transaction costs

One of the most important contributions of this paper is price of zero coupon bond in context of no arbitrage.

CIR. The model of Cox, Ingersoll and Ross (CIR, 1985) is a general equilibrium model, with one factor

\[ r(t) = \delta^* Y(t) \]

Where \( Y(t) \) is a stochastic state variable. Under this model, the price of a bond is a function of \( r \) and \( t \), and the price satisfy a partial differential equation, from which we can obtain the market price of risk, and the market risk parameter. The short rate follows a chi-square distribution.

BS . In the model of Brennan and Schwartz (1979) we have two factors. In this model they worked with the short interest rate and the long term interest rate. The dynamics is specified by a system of two stochastic differential equations. The equation which describes the short rate has a mean reversion dynamics, and the long term one is described as a classical diffusion process. Merton (1973).
The instantaneous short-term interest rate is described by a stochastic differential equation of the form

\[ dr(t) = \theta dt + \sigma dW(t) \]

where \(\theta\) and \(\sigma\) are constants and \(W(t)\) is the standard Brownian motion. As already noted by Merton, the normality permits negative values of the interest rate.

Chan (1992). Chan used the generalized method of moments to demonstrate that the dynamic of short term interest rate, permits a high sensitivity in the volatility. The conditional changes in the mean and variance of the interest rate depends on the level of this interest rate.

Dothan (1978). Takes the assumption of Vasicek, about the free of risk elements of the market. Uses microeconomics tools to maximize certain utility under specific preferences and free arbitrage context.

CEV. The dynamics is given by

\[ dX_t = \left( \alpha_0 + \alpha_1 X + \alpha_2 X^2 + \alpha_3 / X \right) dt + \left( \beta_0 + \beta_1 X + \beta_2 X^{\beta_3} \right) dW_t \]

A particular case from this model is case studied by Cox (1985):

\[ dX_t = \alpha_1 X dt + \beta_1 X^\gamma dW_t \]

Table 1 we show the models studied and their respective transition density.
| Parametric Model | Transition density $p(\Delta, y|x)$ |
|------------------|-----------------------------------|
| **Vasicek (1977)** | $\frac{1}{\sqrt{2\pi \beta_0^2 \Delta}} e^{\left\{ \frac{y-x}{\beta_0^2} \frac{1}{2} (y-x) + \alpha_0 + \frac{1}{2} \alpha_1 (y+x) \right\} - \left( \frac{\alpha_0 + \alpha_1 y}{2 \beta_0^2} + \lambda \right) \Delta}$ 
$+ \frac{\lambda \Delta}{\sqrt{2\pi \sigma_S^2}} e^{-\frac{1}{2} \left( \frac{y-x-\mu_S}{\sigma_S} \right)^2}$ |
| **CIR-Cox, Ingersoll y Ross (1985)** | $\frac{1}{\sqrt{2\pi \Delta \beta_1}} \left( \frac{y}{x} \right)^{\alpha_0 / \beta_1} e^{\left\{ \frac{-2}{2 \beta_1^2 \Delta} (\sqrt{y-x})^2 + (y-x) \left( \frac{\alpha_1}{2 \beta_1^2} \frac{3}{4} \right) \left( \frac{\alpha_0 + \alpha_1 y}{2 \beta_1^2} + \lambda \right) \Delta \right\}}$ 
$+ \frac{\lambda \Delta}{\sqrt{2\pi \sigma_S^2}} e^{-\frac{1}{2} \left( \frac{y-x-\mu_S}{\sigma_S} \right)^2}$ |
| **CIR VR** | $\frac{1}{\sqrt{2\pi \Delta x^3 \beta_1}} \left( \frac{y}{x} \right)^{\frac{9}{4}} e^{\left\{ \frac{1}{2 \beta_1^2 \Delta} \left( \frac{1}{\sqrt{x}} - \frac{1}{\sqrt{y}} \right)^2 \right\} + \lambda \Delta}$ 
$+ \frac{\lambda \Delta}{\sqrt{2\pi \sigma_S^2}} e^{-\frac{1}{2} \left( \frac{y-x-\mu_S}{\sigma_S} \right)^2}$ |
| **Brennan & Schwartz (1982)** | $\frac{1}{\sqrt{2\pi \Delta \beta_2 x^2}} \left( \frac{y}{x} \right)^{3 / 2} e^{\left\{ \frac{a_0}{2 \beta_2^2 x^2} \frac{1}{y^2} + \frac{1}{y} \right\} + \frac{1}{2 \lambda \beta_2^2 \Delta} \left( \frac{\alpha_0 y}{\beta_2} \right)}$ 
$+ \frac{\lambda \Delta}{\sqrt{2\pi \sigma_S^2}} e^{-\frac{1}{2} \left( \frac{y-x-\mu_S}{\sigma_S} \right)^2}$ |
\[
\text{Chan (1992)} \\
\frac{1}{\sqrt{2\pi\Delta\beta_1^2 x^{\beta_1-1}}} \left( \frac{y}{x} \right)^{\frac{3}{2} \beta_1} e^{\frac{x^{\beta_1} - y^{\beta_1}}{2 \beta_1^2 (1-\beta_1)x} - \frac{\alpha_0 (y^{\beta_1} - 1^{\beta_1})}{\beta_1^2 (1-2\beta_1)} + \frac{\alpha_1^2 (y^{\beta_1} - 1^{\beta_1})}{2 \beta_1^2 \beta_1 (1-2\beta_1)} - \frac{(\alpha_0 + \alpha_1)^2}{2 \beta_1^2 y^{\beta_1} - \Delta}} \\
+ \frac{\lambda \Delta}{\sqrt{2\pi\sigma_S^2}} e^{-\frac{1}{2} \left( \frac{y-x-\mu_S}{\sigma_S} \right)^2}
\]

| Parametric Model | Transition density \( p(\Delta, y|x) \) |
|------------------|---------------------------------------|
| **Merton**       | \[
\frac{1}{\sqrt{2\pi\Delta\beta_0}} e^{\left( -\frac{1}{2\beta_0^2} (y-x)^2 + \frac{\alpha_0}{\beta_0^2} (y-x) - \left( \frac{\alpha_0^2}{2 \beta_0^2} + \lambda \right) \Delta \right)} \\
+ \frac{\lambda \Delta}{\sqrt{2\pi\sigma_S^2}} e^{-\frac{1}{2} \left( \frac{y-x-\mu_S}{\sigma_S} \right)^2}
\] |
| **Dothan**       | \[
\frac{1}{\sqrt{2\pi\Delta\beta_1 x^{\beta_1}}} \left( \frac{y}{x} \right)^{-\frac{3}{2}} e^{\left( -\frac{1}{2\beta_1^2} \left( \ln \frac{y}{x} \right)^2 - \lambda \Delta \right)} \\
+ \frac{\lambda \Delta}{\sqrt{2\pi\sigma_S^2}} e^{-\frac{1}{2} \left( \frac{y-x-\mu_S}{\sigma_S} \right)^2}
\] |
| **GBM**          | \[
\frac{1}{\sqrt{2\pi\Delta\beta_1 x^{\beta_1}}} \left( \frac{y}{x} \right)^{\frac{3}{2} \beta_1} e^{\left( -\frac{1}{2\beta_1^2} \left( \ln \frac{y}{x} \right)^2 - \left( \frac{\alpha_1^2}{\beta_1^2} + \lambda \right) \Delta \right)} \\
+ \frac{\lambda \Delta}{\sqrt{2\pi\sigma_S^2}} e^{-\frac{1}{2} \left( \frac{y-x-\mu_S}{\sigma_S} \right)^2}
\] |
III.1 About Nonparametric Estimation of the Transition Density

Suppose that the observed process \( \{X_i\} \) is sampled at the regular time points \( \{i\Delta, i = 1, \ldots, n+1\} \). We make the dependence on the transition function and related quantities on \( \Delta \) implicit by redefining

\[
X_i = X_{i\Delta}, i = 1, \ldots, n+1
\]

which is assumed to be stationary and \( \beta - \text{mixing} \) process. Let \( p(y \mid x) \) be the transition density of the series \( \{X_i, i = 1, \ldots, n+1\} \). A possible estimate of the transition distribution \( P(y \mid x) = P(X_{i+1} < y \mid X_i = x) \) is given by

\[
\hat{P}(y \mid x) = \frac{1}{nh_1} \sum_{i=1}^{n} W_n \left( \frac{X_i - x}{h_1}; x \right) I(X_{i+1} < y)
\]

where \( h_1 \) is the bandwidth, \( W_n \) is the effective kernel induced by the local linear fit. We have obtained the following results which we show in the table 2. (Definitions of \( \hat{E}_M \) and \( \hat{V}_M \) can be found in Ait-Sahalia, Fan and Peng, 2006).
Table 2

Results for the non-parametric model

<table>
<thead>
<tr>
<th>Number of observation</th>
<th>2,196</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_M$</td>
<td>28.34</td>
</tr>
<tr>
<td>$V_M$</td>
<td>590,613.48</td>
</tr>
<tr>
<td>Confidence lever 95%</td>
<td>1.64485</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>0.0024</td>
</tr>
<tr>
<td>$c(\alpha)$</td>
<td>28.3397</td>
</tr>
</tbody>
</table>

With a confidence level of 95%, the critical value is 28.33970, i.e., if the statistics $M$ is below this value, we cannot reject the null hypothesis (the model can explain data). Table 3 shows the statistics for the parametric models. Minimum is referred to the minimization used to calculate the parameters of each model.

Table 3

<table>
<thead>
<tr>
<th>VASICEK</th>
<th>Brennan &amp; Schwartz</th>
<th>Chan</th>
<th>MERTON</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>3.98576</td>
<td>12,011.73826</td>
<td>12,014.40247</td>
</tr>
<tr>
<td>$M$ statistic</td>
<td>20.83483</td>
<td>62,789.21253</td>
<td>62,803.13917</td>
</tr>
<tr>
<td>Result</td>
<td>Non Rejected</td>
<td>Rejected</td>
<td>Rejected</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CIR VR</th>
<th>Dothan</th>
<th>GBM</th>
<th>CEV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum</td>
<td>5,495.03913</td>
<td>1,965.32643</td>
<td>9,052.06187</td>
</tr>
<tr>
<td>$M$ statistic</td>
<td>28,724.33381</td>
<td>10,273.39225</td>
<td>47,318.03374</td>
</tr>
<tr>
<td>Result</td>
<td>Rejected</td>
<td>Rejected</td>
<td>Rejected</td>
</tr>
</tbody>
</table>

As we can see in the table, the null has been rejected for each of the models, except for the Vasicek model, i.e., we can find parameters from this model, and therefore CETES data are represented in an acceptable way.
IV. Conclusions

In order to avoid disappointing results in the risk management, contemporaneous global financial uncertainty requires the application of accurate quantitative financial tools. In our previous paper we studied several classical models for interest rates without jumps, and all of them were rejected to describe Mexican CETES. However, in the current paper we have found that Vasicek model with jumps is good enough to represent Mexican data. In that way the Ait-Sahalia, Fan and Peng (2006) methodology allow us confirm our previous results (Núñez, de la Cruz and Ortega, 2007): CETES dynamics cannot be described by some interest rate models. Consequently, the empirical application of the inadequate model has a negative effect over the measuring of relevant financial variables in the Mexican case. The next step in our agenda of researching is the consequences of this finding when we use derivatives based on interest rate.

References


