Ensayos–Volumen XXVII, núm.1, mayo 2008, pp. 33-52

Parametric vs. non-parametric methods for estimating option implied risk-neutral densities: the case of the exchange rate Mexican peso – US dollar[1][2][3]

Guillermo Benavides Perales[*]
Israel Felipe Mora Cuevas[**]

Abstract

This research paper presents statistical comparisons between two methods that are commonly used to estimate option implied Risk-Neutral Densities (RND). These are: 1) mixture of lognormals (MXL); and, 2) volatility function technique (VFT). The former is a parametric method whilst the latter is a non-parametric approach. The RNDs are extracted from over-the-counter European-style options on the Mexican Peso–US Dollar exchange rate. The non-parametric method was the superior one for out-of-sample evaluations. The implied mean, median and mode were, in general, statistically different between the competing approaches. It is recommended to apply the VFT instead of the MXL given that the former has superior accuracy and it can be estimated when there is a relatively short cross-section of option exercise price range. The results have implications for financial investors and policy makers given that they could use the information content in options to analyze market's perceptions about the future expected variability of the financial asset under study.

***Key words**:* currency option implied volatility, exchange rate, parametric methods, non-parametric methods, risk-neutral densities.

---

[1] The title in spanish is 'Métodos paramétricos vs. no-paramétricos en la estimación de densidades de riesgo-neutral Implícitas en precios de opciones: Un análisis empírico para el tipo de cambio peso-dólar'.

[2] This paper's findings, interpretations, and conclusions are entirely those of the authors and do not necessarily represent the views of Banco de México (Central Bank of Mexico), its Board or its Staff Officials. The authors acknowledge financial support from CONACYT (especially Sistema Nacional de Investigadores).

[3] Acknowledgements: The authors want to thank Alberto Torres, Alejandro Gaytán, Alfonso Guerra and Alejandro Díaz de León for valuable comments of an earlier version of this paper. Thanks are also given to seminar participants at Tecnológico de Monterrey, Campus Ciudad de México, conference participants at the VII International Conference in Finance 2007 held at the Tecnológico de Monterrey, Campus Monterrey and participants at LACEA 2007 Conference held in Bogota, Colombia. We thank two anonymous referees for very helpful suggestions. Any remaining errors are responsibility of the authors.

[*] Banco de México, Dirección General de Investigación Económica and Tecnológico de Monterrey, Campus Ciudad de México. Correo electrónico: gbenavid@banxico.org.mx

[**] University of Essex

Resumen

En este trabajo de investigación, se realiza un análisis comparativo del poder predictivo de dos modelos que comúnmente se utilizan para estimar Densidades de Riesgo-Neutral (RNDs), obtenidas de los precios de opciones. Los modelos son: 1) mezcla de lognormales (MXL); y, 2) función de volatilidad implícita (VFT). El primero corresponde a un método paramétrico y el segundo, a un método no-paramétrico. Las RNDs son obtenidas de información diaria de opciones extrabúrsatiles de estilo-europeo del tipo de cambio Peso Mexicano – Dólar Estadounidense. En términos generales, los momentos de las RNDs fueron estadísticamente diferentes, entre ambas técnicas aplicadas. Se recomienda el uso del modelo VFT sobre el MXL, debido a que el primero obtuvo mejor desempeño en evaluación fuera de la muestra; mientras que el segundo no se puede estimar, si no se cuenta con un rango de precios de ejercicio suficientemente amplio. Los resultados pueden ser de interés tanto para inversionistas financieros, como para quienes se encuentran en situación de toma de decisiones. Esto es así, debido a que estos métodos pueden extraer las expectativas que tiene el mercado, estimadas a través de las RNDs, sobre la futura distribución de precios del activo que aquí se analiza.

Palabras clave*:* volatilidad implícita en opciones de divisa, tipo de cambio, métodos paramétricos, métodos no paramétricos, densidades de riesgo neutral.
Clasificación JEL: C14, C52, F31, G13.

Introduction

The present paper analyzes statistical differences between the two most commonly used methods to extract Rish Neutral Densities (RNDs) from option prices. These are: mixture of lognormals (MXL) and the volatility function technique (VFT). The former is a parametric method whilst the latter uses a non-parametric approach. The main objective is to identify statistical differences between the moments of the distributions as well as to analyze which method is the superior one in terms of goodness-of-fit and forecast accuracy. The forecast accuracy uses the first moments of the distributions and statistically compares them with spot prices relevant to the expiration date of the options (to analyze forecast power). The metric for the selection of the best method is done in terms of mean squared errors (MSE). These are obtained with the differences between implied first moments and spot prices at the maturity day of the options. Up to day, there is no general consensus in favor of one individual method statistically proven to be the most accurate. However, the results presented here are consistent with a significant part of the literature that has found that VFT is more accurate for estimating exchange rate (Bliss and Panigirtzoglou: 2000, Castrén: 2005).

To add relevant findings to the literature of RND estimation the above mentioned methods are analyzed in order to test the following null hypothesis:

> H_0: The Mixture of Lognormals method and the Volatility Function Technique do not provide different statistical estimates of Risk-Neutral Densities.

Different to most works in the literature, this paper not only includes a statistical comparison between MXL and VFT, but also presents an evaluation of each model's out-of-sample forecast accuracy based upon statistical comparisons of first moments versus ex-post exchange rate data i.e. comparisons between the implied mean, median, mode and ex-post exchange spot prices. The main motivation is to extend the current literature about statistically comparing parametric and non-parametric methods in the estimation of RNDs. One of the main objectives is to compare the predictive accuracy of these widely used methodologies. Statistically significance tests for equal forecast accuracy have been rarely reported in the literature. Finally, it is worth mentioning that up to now these types of models have not been compared using the Mexican peso–US Dollar exchange rate, which makes an additional contribution to do the proposed research.[4]

The findings of this work could be of interest for agents involved in making risk management decisions related to exchange rates, particularly, the Mexican peso–US Dollar. Such agents could be academic researchers, bankers, derivatives traders, investors, policy makers, central banks, among others.

The layout of this paper is as follows. The literature review of the MXL and VFT approaches are presented in Section I. The models are explained in detail in Section II. Information about the data used for the empirical research is shown in Section III. The results are presented in Section IV. Finally, the conclusions are presented.

[4] Díaz de León and Casanova (2004) estimated RNDs for the Mexican peso – USD exchange rate and for oil prices. However, they do not present a rigorous statistical comparison of the MXL vis-à-vis VFT mainly because they have different research questions.

I. Academic Literature of Risk-Neutral Densities

I.1. Risk-Neutral Densities Definition

The idea to estimate RNDs implied by option prices was first postulated by Breeden and Litzenberger (1978). The main reason to do this was the belief that derivative markets provided a rich source of forward-looking financial information embedded in them. A way to extract this information is by estimating an implicit probability distribution from option prices, which were traded in financial markets for a specific underlying asset. That is, the underlying assets' distribution implied by the observed market prices of those options. Given that the models used to estimate these probabilities have the assumption that the agents are risk-neutral, the resulting probability density is called risk-neutral density.[5]

The underlying assumptions of an option valuation model like the one postulated by Garman and Kohlhagen (1983: GK) are the following: 1) interest rates are non-stochastic, 2) there are no arbitrage profits, 3) all options are European-style, 4) agents are risk-neutral, 5) there are no transaction costs or taxes; and, 6) the price for the underlying asset follows a Geometric Brownian Motion.[6] The GK is presented in Equation 1 below.

$$c(X, T) = Se^{-rfT}N(d_1) - Xe^{-rT}N(d_2) \qquad (1)$$

$$p(X, T) = Xe^{-rT}N(-d_2) - Se^{-rfT}N(-d_1)$$

$$d_1 = \frac{\ln\left(\frac{S}{X}\right) + \left(r - r_f + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}$$

$$d_2 = d_1 - \sigma\sqrt{T}$$

Where c is the value of the European-style call option, p is the value of the European-style put option, T represents the time-to-maturity of the option, $N(x)$ is the cumulative probability distribution function, which is normally distributed. In other words, the probability that a variable with a standard normal distribution, $\psi(0, 1)$ will be less than x. The exercise price is

---

[5] RNDs have some important theoretical concepts that are consider in their interpretations. However, it should be kept in mind that 'real-world' densities may be significantly different to the 'theoretical' RNDs extracted from options.

[6] An European-style option can only be exercised at the maturity date of the option. On the other hand, an American-style option can be exercised at any time during the life of the option (including its maturity date).

represented by X, S refers to the spot price, ln is the natural logarithm, σ is the asset's annualised volatility measured as it annualised price-return standard deviation, r is the domestic risk-free interest rate, and r_f is the foreign risk-free interest rate.[7]

If observed option prices in the market are used instead of theoretical ones, it is possible to implicitly extract the probability distribution that the agents had when they traded the options. Supposing that the c, p, S, X, r, r_f, T in Equation 1 above are observed. After making an assumption about σ the RND can be implicitly estimated by finding the risk-neutral probability function $q(S_T)$ in terms of spot prices at the maturity of the option (S_T), relevant to that specific traded option premium. So, instead of assuming a standard $N(x)$ as it is shown in Equation 1 above, the RND is implicitly extracted from the model using the observed variables.

In a detailed option-pricing model derivation Breeden and Litzenberger (1978) proved that the RND that it is contained in option prices can be extracted by calculating the second partial derivative of the call price function $C(X, T)$, with respect to the different exercise prices (X):

$$\left(\frac{\partial^2 C(X,T)}{\partial X^2}\right) = e^{-rT} q(S_T) \; , \tag{2}$$

Rearranging terms it is possible to have the following definition,

$$q(S_T) = e^{rT}\left(\frac{\partial^2 C(X,T)}{\partial X^2}\right) \tag{3}$$

where $q(S_T)$ represents the risk-neutral probability function (RND) of the underlying asset (spot prices at maturity of the option) and the other variables are as defined previously. The problem with this definition is the assumption that the call price function is continuous for the range of exercise prices. As it is known this is not realistic given that in practice only some prices in discrete time are available or observed. Considering this limitation Shimko (1993) proposed an interpolation method using the exercise prices available. In subsequent research, Malz (1997b) proposed to interpolate

---

7 Price-returns are defined as the natural logarithmic difference between the contemporaneous price and the price one period before i.e. ln P_t – ln P_{t-1} (where P represents price and ln is the natural logarithm).

across implied volatilities (obtained with a GK model) and the delta.[8] In this case, the delta has to pass through at least three points on the volatility smile as it will be explained in more detail below in Subsection *II.2* (the non-parametric model subsection).[9]

RNDs estimations do not only give a point estimate forecast about a specific underlying asset but they give the whole distribution expected by the market around a point estimate forecast. Extracting a RND provides information about market sentiment. For example, if an exchange rate shows RNDs with skewness that are systematically positive through time, the interpretation is that the market is expecting one of the currencies to depreciate (or keep depreciating) in the near future.

Even though substantial amount of research has been done about this topic, there still is a current debate in the literature about how different are the risk-neutral probability distributions compared with the market's 'true' probability distributions. For the case of exchange rates Christoffersen and Mazzotta (2004) found that RNDs provide reliable estimates of true density functions. The evidence was corroborated for stock indices (Liu, et. al. 2004).

Considering the actual evidence it could be assumed that from a theoretical viewpoint, RNDs are a parsimonious and reliable method for capturing the market's belief about a future asset price distribution.[10] The present research paper emphasizes on comparing two of the most popular methods (parametric and non-parametric) to extract RNDs for exchange rates. However, the research question about the differences between RNDs and 'true' probability distributions (real-world densities) is not analyzed here.

---

[8] Delta is defined in Hull (2003) as the rate of change of the price of a derivative with respect to the price of the underlying asset.

[9] The volatility smile shows the relationship in a two dimensional space between the option implied volatilities and their relevant exercise prices or deltas. For exchange rates these curves commonly have a u-shaped form i.e. a 'smile'. This is why they are named volatility smiles. Smile curves are also observed for stock options but, in that case, the curve is normally downward sloping (also named volatility skew).

[10] The exact date for this implied asset price distribution is the maturity date of the options.

I.2. Mixture of Lognormals

The mixture of Lognormals method is nowadays one of the most used methods to extract RNDs (Clews et. al. 2000). It was postulated around the mid 1990's by Rubinstein (1994), Bahra (1997) and Melick and Thomas (1997). The first two authors analyzed interest rate markets whilst Melick and Thomas did it on oil prices; the MXL is based on a weighted mixture of two or more lognormal distributions. According to Bahra (1997) the intuition of this approach is that it is possible to make assumptions about the functional form of the RND with the objective of recovering the parameters by minimizing the squared distance between the observed and the theoretical option prices; the latter generated by an assumed parametric form. So, instead of starting with an assumption about the stochastic process for the underlying asset (as other methods proposed), MXL initially assumes a lognormal terminal RND function.

I.3. The Volatility Function Technique

The volatility function technique was originally postulated by Malz (1997a). He extended an idea proposed by Shimko (1993) in which an interpolation of exercise prices was derived in order to recover a RND from option prices. Shimko's method proposed a parabolic function to estimate a curve for the implied volatility function vis-à-vis exercise prices i.e. the smile curve. The idea behind this method is to estimate a 'smoothed' smile implied volatility function, out of relatively few exercise prices (five or less) with a parabolic function and then generate smooth call option prices using the Black-Scholes equation (BS). With the estimated call prices the RND can be extracted by applying the Breeden and Litzenberger (1978) approach (explained in Subsection I.1 above). The main difference with Malz is that the latter does not use a parabolic function to estimate the smile curve but instead he applies implied volatilities from option pricing strategies (risk reversals and strangles). The objective is to estimate a curve matching implied volatility vis-à-vis the delta and then calculate the call option prices from it by using BS. Malz argued that his method is more accurate for modeling financial data given that option strategies' implied volatilities, like risk reversals (*rr*) and strangles (*str*), capture the market's expectations for the relative likelihood of exchange rate depreciations (implied skewness) and extreme events (excess implied kurtosis).[11]

---

[11] A risk reversal in an option trading strategy where an investor simultaneously takes a long position in an out-of-the-money (OTM) call option and a short position in an OTM put option, both on the same underlying asset and time-to-maturity. A strangle is a similar type of strategy but in this case an investor simultaneously takes a long position in both an OTM call and put options (Hull: 2003, Micu: 2004). These strategies are usually

Several studies have statistically compared MXL and VFT. The comparisons are generally in terms of goodness-of-fit and stability of the parameters. Among those studies Bliss and Panigirtzoglou (2000) extracted RNDs for the FTSE-100 stock index and short sterling futures. They concluded in favor of the VFT arguing that it had both higher goodness-of-fit and stability of the parameters. On the other hand, Mc Manus (1999) found that the MXL was the most accurate method by showing higher goodness-of-fit for the case of Eurodollar options. Micu (2004) extracted RNDs for twelve emerging markets currencies vis-à-vis the US Dollar. He concluded that there is a trade-off between goodness-of-fit accuracy and stability of the parameters.

II. The Models

II.1. Mixture of Lognormals

One of the most used methods to extract RNDs from option prices is the Mixture of Lognormals (Rubinstein: 1994, Melick and Thomas: 1997, Bahra: 1997, Bliss and Panigirtzoglou: 2000). One of the assumptions for the underlying model is that the prices for European-style options at time *t* can be defined as the present value of the sum of expected payoffs (methodology and notation follow those presented in Bahra: 1997),

$$c(X,T) = e^{-rT} \int_X^{\infty} q(S_T)(S_T - X)dS_T \tag{4}$$

$$p(X,T) = e^{-rT} \int_0^{X} q(S_T)(X - S_T)dS_T \tag{5}$$

where notation is the same as defined previously. In theory any functional specification for the density $q(S_T)$ can be used for Equations (4) and (5), and its parameters can be recovered (implicitly) through a numerical optimization method. According to Ritchey (1990) it is possible to consider the following definitions:

$$q(S_T) = \sum_{i=1}^{k} \left[\theta_i L(\alpha_i, \beta_i; S_T)\right] \tag{6}$$

where $L(\alpha_i, \beta_i; S_T)$ represents the lognormal density function *i* in the *k*-component of the mixture. The *k*-component indicates the total number of lognormals used in the mixture considering the parameters α_i, β_i (for example, for the case of the mixture of two lognormals the value of *K* is

undertaken by investors that believe that the underlying spot price will end up far away from the current (at-the-money) spot price.

equal to two). The definitions for α_i, and β_i are taken from a lognormal distribution and are shown next:

$$\alpha_i = \ln S + \left(\mu_i - \frac{1}{2}\sigma_i^2\right)T \qquad (7)$$

$$\beta_i = \sigma_i \sqrt{T} \quad \text{for each } i \qquad (8)$$

where μ_i represents the return expectation of the series (i) and σ_i represents its standard deviation.[12] The weights of the probabilities θ_i, must satisfy the following condition,

$$\sum_{i=1}^{k} \theta_i = 1, \quad \theta_i > 0 \ \text{ for each } i \qquad (9)$$

It is well known that in many derivative exchanges, option contracts are traded within a relatively narrow range of exercise prices. Because of this, there are limitations about the number of parameters of the distributions that can be estimated from the data. Even though there is a narrow range, a mixture of two lognormals can be validly applied to estimate the optimal parameters, α_1, α_2, β_1, β_2 and θ as long as there are at least five different exercise prices. For that it is important to define call and put options prices in the following way:

$$c(X,T) = e^{-rT}\int_X^{\infty}\left[\theta L(\alpha_1,\beta_1;S_T) + (1-\theta)L(\alpha_2,\beta_2;S_T)\right](S_T - X)dS_T \qquad (10)$$

$$p(X,T) = e^{-rT}\int_0^{X}\left[\theta L(\alpha_1,\beta_1;S_T) + (1-\theta)L(\alpha_2,\beta_2;S_T)\right](X - S_T)dS_T \qquad (11)$$

For an easier estimation Equations 10 and 11 must be transformed to closed-form.[13] Once this is done the following can be expressed,

$$c(X,T) = e^{-rT}\left\{\theta\left[e^{\alpha_1+\frac{1}{2}\beta_1^2}N(d_1) - XN(d_2)\right] + (1-\theta)\left[e^{\alpha_2+\frac{1}{2}\beta_2^2}N(d_3) - XN(d_4)\right]\right\} \qquad (12)$$

$$p(X,T) = e^{-rT}\left\{\theta\left[XN(-d_2) - e^{\alpha_1+\frac{1}{2}\beta_1^2}N(-d_1)\right] + (1-\theta)\left[XN(-d_4) - e^{\alpha_2+\frac{1}{2}\beta_2^2}N(-d_3)\right]\right\} \qquad (13)$$

$$d_1 = \frac{-\ln(X) + \alpha_1 + \beta_1^2}{\beta_1} \qquad\qquad d_3 = \frac{-\ln(X) + \alpha_2 + \beta_2^2}{\beta_2}$$

---

[12] For the case of an exchange rate option μ represents the differential between the domestic and the foreign risk-free interest rate.

[13] To see more details about the derivation to closed-form solutions the interested reader can refer to Bahra (1997, pg. 50).

$d_2 = d_1 - \beta_1$ $\qquad\qquad$ $d_4 = d_2 - \beta_2$

Both call and put options are related to the same underlying asset. For this reason, it is consistent if both option premium series are included for minimizing one objective function. The idea is to find the optimal parameters α_1, α_2, β_1, β_2 and θ, which will minimize the difference between the theoretical and observed prices.[14] The minimization problem is,

$$Min_{\alpha_1,\alpha_2,\beta_1,\beta_2,\theta} \sum_{i=1}^{n} [c(X_i,T)-\hat{c}_i]^2 + \sum_{i=1}^{n} [p(X_i,T)-\hat{p}_i]^2 + \left[\theta e^{\alpha_1+\frac{1}{2}\beta_1^2} + (1-\theta)e^{\alpha_2+\frac{1}{2}\beta_2^2} - e^{rT}S_T\right]^2 \quad (14)$$

with the following restrictions, β_1 and $\beta_2 > 0$ and $0 \le \theta \le 1$, for the observed range of exercise prices X_1, X_2, X_3,…,X_n. $c(X_i, T)$ and $p(X_i, T)$ are the theoretical prices for the call and put options respectively, whilst $\hat{c}$ and $\hat{p}$ represent the observed call and put prices, respectively. Once the optimal parameters are obtained the RND is estimated using the closed-form Equations 12 and 13.

II.2. The Volatility Function Technique

Malz (1997a) shows that a RND can be estimated by interpolating the smile curve. Interpolation can be carried out in terms of the implied volatilities determined from market expectations. The implied volatilities considered are: at-the-money where the forward price (F) is equals to the exercise price (*atm*); risk reversal; and, strangle. For exchange rates these were taken from market traders.

The implied volatilities from the above mentioned option strategies for a 25-delta call and put option can be theoretically obtained as follows.

The *rr* is defined as,

$$rr_t^{25\Delta} = \sigma_t^{(\Delta_{0.25}^c)} - \sigma_t^{(\Delta_{0.25}^p)} \quad (15)$$

and the *str* can be expressed as

$$str_t^{25\Delta} = 0.5\left[\sigma_t^{(\Delta_{0.25}^c)} + \sigma_t^{(\Delta_{0.25}^p)}\right] - \sigma_t^{ATM} \quad (16)$$

[14] For details about the derivation to obtain the objective function the interested reader can refer to Bahra (1997, pg. 50).

Incorporating the volatilities to a quadratic function, it is then possible to obtain the following smile curve (Malz: 1997a),

$$\sigma(\delta) = atm - 2rr(\delta - 0.5) + 16str(\delta - 0.5)^2, \tag{17}$$

where σ(.) represents the market's implied volatility as a function of delta (δ). Once this curve is obtained a transformation is performed in which the implied volatility can be expressed in terms of exercise price (X) and not in terms of the delta. Thus, the definition of the delta function is now as follows:

$$\delta = e^{-rfT} * N\left(\frac{\ln\left(\frac{F_t}{X}\right) + \left(\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right), \tag{18}$$

where the notation is as defined previously. Equation 18 is substituted in Equation 17 and then Equation 19 below is obtained,

$$atm - 2rr\left[\left(e^{-rfT} * N\left(\frac{\ln\left(\frac{F}{X}\right) + \left(\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right)\right) - 0.5\right] + 16str\left[\left(e^{-rfT} * N\left(\frac{\ln\left(\frac{F}{X}\right) + \left(\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right)\right) - 0.5\right]^2 - \sigma = 0$$

To estimate the density function for the underlying asset the Breeden and Litzenberger (1978) result is applied here (Equation 3 above). Substituting in Equation 19 it is then possible to estimate the probability function for the underlying asset, which is expressed as follows,

$$q(S_T) = e^{rT}\left[\left[F\left(\left(\frac{n(d_1)}{X^2 v\sqrt{T}}\right) + n(d_1)d_1\left(\frac{1}{Xv\sqrt{v}}\right)\right)^2\right] - \left[X\left(\left(\frac{n(d_2)}{X^2 v\sqrt{T}}\right) + n(d_2)d_2\left(\frac{1}{Xv\sqrt{T}}\right)\right)\right]^2\right] \tag{20}$$

where,

$$d_1 = \frac{\ln\left(\frac{F}{X}\right) + \left(\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}, \qquad d_2 = \frac{\ln\left(\frac{F}{X}\right) - \left(\frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}},$$

and v represents the option implied volatility (σ), which makes Equation 19 equals to zero. Finally, by using different values of X it is then possible to extract the RND through the prices of options.

III. Data

The data for the exchange rate consists of daily spot and forward quotes obtained from a Banco de México's financial database ('Central de Información Financiera (CIF)').[15] The forward and option prices are daily over-the-counter (OTC) quotes with one-month-to-expiration traded by major Mexican banks and other financial intermediaries.[16] The exchange rate *atm*, *rr*, and *str* implied volatilities were obtained from Switzerland-based UBS[17] Financial Institution data base.[18] The domestic discount rate consists of daily 28-day secondary market interest rates of Mexican Certificates of Deposit (CDs) obtained from the same source.[19] United States CDs were obtained from the Federal Reserve (FED) web page with the same maturity in order to include the equivalent foreign risk-free discount rate in the estimations.[20] The sample period under analysis is more than four years, from 28/05/2002 to 04/07/2006. The final selected sample size consists of 39 daily observations for the exchange rates.

IV. Results

IV.1. Exchange Rate

This subsection presents the estimated RNDs implied moments for the exchange rate through time. The implied moments are the mean, annualized volatility (implied standard deviation), skewness and kurtosis. The mean gives the market's expected value for the underlying asset at the maturity date of the options. The annualized volatility provides an indication of the dispersion of expectations around the mean exchange rate. The probability of a large upward movement compared to a large downward movement is explained with the implied skewness. This also measures the asymmetry of the distribution. Kurtosis indicates the possibility of large changes in

---

[15] Banco de México's 'Central de Información Financiera' has exchange OTC forward and option prices. This database is not available to the general public, however, daily spot exchange rates can be downloaded from the Web page, which is http://www.banxico.org.mx

[16] For exchange rate options OTC markets have significantly higher volume of trade compared with exchange-traded options. This difference can be more than fourth fold (Castrén: 2005). European-style option data is used in this document.

[17] UBS is one of the world's leading financial firms and operate in two locations. The web page is http://www.ubs.com/

[18] The quotes from UBS were from financial traders that quoted volatility and not option prices. This is normally done for exchange rate options (see Cooper and Talbot (1999) for details).

[19] In Spanish these are 'Certificados de la Tesorería de la Federación (CETES)'.

[20] The FED web page is http://www.federalreserve.gov/.

exchange rates prior the maturity date of the options (Mc Manus, 1999). It is important to point out that a skewed left (negative skewness) distribution gives greater weight to the possibilities that the exchange rate will be far below instead of far above (with respect to the mean value) at the maturity date of the options. In this situation, it is expected that the implied mean will be lower than the implied median and mode. The opposite holds true i.e. for the case of a positive skewness.

Implied moments through time for the exchange rate are presented in Figures 1-6.[21] It can be observed that both methods have relatively different implied measures between them. For example, the MXL has significantly more variation compared to VFT in all the implied moments. These differences show that the implied measures tend to be model dependent. Because of this, it makes it difficult to rely on them as indicators of market sentiment (Mc Manus, 1999). However, these results show that the statistical differences between the estimation models are relevant depending on the chosen model. This is in line with that part of the academic literature that considers the importance of choosing a correct RND model. The type of the model creates a statistical difference in terms of gauging market sentiment. Table 1 shows results from tests about statistical differences between the implied mean, median and variance for both methods.[22] In the majority of the cases the null hypothesis of equality of the implied mean, median and variance is clearly rejected (*p*-values reported).

It can be observed in Table 2 below that the VFT has higher accuracy predicting observed data for the exchange rate. In other words, the implied mean, median and mode from the VFT estimations were closer to the spot exchange rates at the options' maturity date i.e. the day the forecasts are estimated for. Following Diebold and Mariano (1995) the statistical difference between the MSEs from both methods is reported in the last column.[23] The MSEs are the lowest for the VFT method. The estimations are

---

[21] Figures 1-6, table 1 and table 2 are presented in the Appendix.

[22] The statistical tests used here are the standard *t*-test (for the mean), a Wilcoxon / Mann-Whitney (for the median) and a *F*-test (for the variance). Details about these standard tests are available to the reader upon request.

[23] This method requires generating a time series, which is the differential of the squared-forecast errors from two different forecast models i.e. $d_t = \left(\sigma_t^2 - \hat{\sigma}_{1,t-1}\right)^2 - \left(\sigma_t^2 - \hat{\sigma}_{2,t-1}\right)^2$, where d_t is the differential of the series and $\hat{\sigma}_i$ is the forecast of the *i* model. The *t*-statistic is obtained in the following way, $\dfrac{\bar{d}}{\sqrt{\dfrac{sd}{n}}}$ where $\bar{d}$ is the sample mean and *sd* is equal to the standard deviation of the series (*d*). *n* represents the total number of observations.

statistically different between the MSEs at the 5% level. The implied median had higher accuracy compared with the implied mean and mode.

IV.2. Analysis of the Results

There are advantages of using the MXL and VFT methods to estimate RNDs. The former gives the possibility that any return on a given asset could be drawn from two lognormal distributions. However, MXL requires a minimum number of cross-sections of option prices with maturity on the same date in order to be estimated (at least five per trading day in this case). On the other hand, the VFT method does not actually need a minimum number of observed option market prices. This is because the RNDs are extracted from the implied volatility function 'smile curve' through interpolations. An advantage of the VFT over the MXL is that the former can be estimated for almost every trading day as long as there is information about the implied volatilities, *atm*, *rr* and *str* (either traded or estimated). Also, VFT appears to be more stable (in the higher moments) through time. This can be observed in Figures 1 – 6.

It is concluded that the VFT model was the superior one for the out-of-sample evaluation, which considers the comparison between the implied first moments of the implied distribution and the observed spot prices at the maturity-date of the options. This is an out-of-sample test given that it compares the expectations at the day of trading (by the estimation of the implied mean) and the actual observed (spot) value for that specific maturity date. In terms of statistical significance of MSEs for the out-of-sample evaluation, it was shown that there are statistical significant differences between both methods, being the VFT the most accurate one. The findings also show clear discrepancies between the higher implied moments estimated with both methods. Unfortunately, these statistical differences show unstable higher moments, therefore it is difficult to conclude that RNDs can, in an unbiased way, gauge market sentiment. Nonetheless, the out-of-sample tests confirm the advantages of the VFT method over the MXL in terms of forecasting accuracy for the period under analysis.

Recent works in the literature by Abadir and Rockinger (2003) and Bu and Hadri (2007) have applied hypergeometric functions to estimate RNDs. Their results are interesting given that both research papers acknowledge the benefits of using these types of hypergeometric functions compared to other methods. The results we find in this paper are in line with those of Bu and Hadri (2007) in that VFT-type methods could provide higher accuracy and stability compared to MXL-type methods. However, they do not perform a more rigorous testing of these methods considering that there are several features a RND estimation should be considered (see Taylor: 2005).[24] These authors based their conclusions only on accuracy and stability tests. They compare a method similar to the one applied here (Malz) but not exactly the same method. So we believe that their conclusions add insights to the literature but at this moment there is not consensus yet about the superiority of one method over the rest. It is worth mentioning that Micu (2004) concluded that there is not a unique method that will outperform others given that there appears to be a trade off between accuracy and stability.

Finally, the work by Hördahl and Vestin (2003) shows the importance of a risk premium when interpreting RNDs. One of their conclusions is that this risk premium is not constant through time. They argue that the relationship between RNDs and RWDs is neither constant nor equivalent. The underlying assets they apply bond with one and ten year maturities. They obtain interest rates for one and ten years respectively. Given that the underlying asset they consider differs from the one applied here there is no evidence that their conclusions could apply to our case. It is important to mention that our null hypothesis is stated to find statistical differences between the methods used in this paper. Even though we also conduct accuracy tests our main research question is based on the statistical differences between the estimation methods presented.

Suggestions for further research

For the Mexican case, it is suggested that future research should analyze specific monetary events and its influence to market expectations for the underlying asset. The former can be identified with recorded monetary policy decisions and the latter can be estimated with RNDs. For example, RNDs can be estimated for their underlying asset around the date when there is a Monetary Policy Committee Meeting at the Mexican Central Bank. It could be that the differences on these two estimation procedures could be higher around an important monetary policy event (Mc Manus 1999, Castrén 2005). It is also important to test other methods like Hermite Polynomial Approximations, Maximum Entropy, piecewise cubic polynomial and other non-parametric ones in order to make a comparison to the results presented here. Finally, it could be important to analyze in more detail the possible application of a premium to compensate investors for a systematic

---

[24] According to Taylor (2005) RNDs should be 1) always positive, 2) permit skewness and kurtosis, 3) have fatter tails compared to lognormal distributions, 4) there are analytic formulae for the density and the call price formula, 5) estimates should be stable, 6) solutions to the estimation problem must be relatively simple, 7) no subjective decisions must be made, 8) RNDs should be easily transform to RWDs.

depreciation of the Mexican peso vis-à-vis United States dollar. The systematic depreciation of this exchange rate was observed because of its positive implied skewness (in pesos per USD) throughout the sample.[25]

Conclusion

In the present research project RNDs estimation models were compared to each other in order to find the forecasting model with higher predictive accuracy for the Mexican peso – US Dollar exchange rate. The models were Mixture of Lognormals and the Volatility Function Technique. The former is a parametric method whilst the latter is a non-parametric technique. Tests were performed for out-of-sample evaluations. These were conducted to compare the implied mean, median and mode of each RND vis-à-vis the expost asset price at maturity of the options (the forecast day). The metrics used were mean squared errors. The VFT contained most of the information content of the realized spot exchange rate at maturity of the options. According to the results, the null hypothesis that the Mixture of Lognormals method and the Volatility Function Technique do not provide different statistical estimates of Risk-Neutral Densities is rejected. Considering that the VFT can be estimated when there are relatively few cross-section of option prices whilst the MXL can not be estimated, it is recommended to use the former instead of the latter.[26] A word of caution is given to these conclusions considering that it was shown that the higher moments of the RNDs were unstable and model dependent.

Bibliography

Abadir, K. M., and Rockinger, M. (2003). "Density functionals, with an option-pricing application." *Econometric Theory*. Vol. 19, pp. 778-811.

Bahra, B. (1997). "Implied Risk Neutral Probability Density Functions from Options Prices: Theory and Application." *Bank of England. Workingpaper No.* 66.

[25] The reader can refer to Díaz de León and Casanova (2004) for more details about a premium to compensate investors for the systematic depreciation of the Mexican peso vs. USD.

[26] For Mexican exchange rate OTC options it is very common to have less than five cross-section observations of option prices that mature on the same day in any given trading day. This makes it very difficult to obtain accurate RND estimates for the MXL method (as explained above).

Bliss, R., and Panigirtzoglou, N. (2000). "Testing the Stability of Implied Probability Density Functions." *Bank of England*. ISSN 1368-5562.

Breeden, D., and Litzenberger, R. (1978). "Prices of State-Contingent Claims Implicit in Option Prices." *The Journal of Business*. Vol. 51, pp. 621-51.

Bu, R., and Hadri, K. (2007). "Estimating option implied risk-neutral densities using spline and hypergeometric functions." *The Econometrics Journal.* Vol. 10, pp. 216-244.

Castrén, O. (2005). "Estimating and Analyzing Currency Options Implied Risk-Neutral Density Functions for the Largest New EU Member Status." *European Central Bank. Workingpaper* No. 440. February.

Christoffersen, P., and Mazzotta, S. (2004). "The Informational Content of Over-the-Counter Currency Options." *European Central Bank.* Working paper No. 366.

Clews, R., Panigirtzoglou, N., and Proudman, J. (2000). "Recent Developments in Extracting Information from Option Markets." *Bank of England Quarterly Bulletin*. February.

Cooper, N., and Talbot, J. (1999). "The Yen/Dollar Exchange Rate in 1998: Views from Options Markets." *Bank of England.* Working paper.

Díaz de León, A., and Casanova, M. E. (2004). "Expectativas del Mercado Implícitas en los Precios de Instrumentos Derivados: Aplicaciones al Mercado Cambiario y Petrolero." *Banco de México.* Documento de investigación No. 2004-01.

Diebold, F. X., and Mariano, R. S. (1995). "Comparing Predictive Accuracy." *Journal of Business and Economic Statistics.* Vol. 13, pp. 253-263.

Garman, M. B., and Kohlhagen, S. W. (1983). "Foreign Currency Option Values." *Journal of International Money and Finance.* Vol. 2, pp. 231-37. May.

Hördahl, P., and Vestin, D. (2003). "Interpreting implied risk-neutral densities: the role of risk premia." *European Central Bank. Workingpaper* No. 274.

Hull, J. (2003). "Options, Futures and Other Derivatives." 5th. Edition. Prentice Hall.

Liu, X., Shackelton, M., Taylor, S., and Xinzhong, X. (2004). "Closed-form Transformations from Risk-Neutral to Real-world Distributions." *Lancaster University.* Working paper.

Malz, A. M. (1997a). "Option-implied Probability Distributions and Currency Excess Returns." *Federal Reserve Bank of New York.* Staff Reports, No. 32.

Malz, A. M. (1997b). "Estimating the Probability Distribution of the Future Exchange Rate from Option Prices." *Journal of Derivatives.* Vol. 5, pp. 18-36.

Mc Manus, D. J. (1999). "The Information Content of Interest Rate Futures Options." *Bank of Canada. Workingpaper* No. 99-15.

Melick, W. R., and Thomas C. P. (1997). "Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil during the Gulf Crisis." *Journal of Financial and Quantitative Analysis*. Vol. 32, pp. 91-115.

Micu, M. (2004). "Extracting Expectations from Currency Option Prices: A Comparison of Methods." *Bank of International Settlements.* Working paper. September.

Ritchey, R. J. (1990). "Call Option Valuation for Discrete Normal Mixtures." *Journal of Financial Research.* Vol. XIII, No. 4, pp. 285-96.

Rubinstein, M. (1994). "Implied Binomial Trees." *Journal of Finance.* Vol. 49, pp. 771-818. July.

Shimko, D. (1993). "Bounds of Probability." *Risk*. Vol. 6, pp. 33-7. April.

Taylor, S. J. (2005). "Asset Price Dynamics, Volatility, and Prediction." *Princeton University Press.*

Appendix

Figures 1 - 6. Mexican Peso vis-à-vis USD RNDs Implied Moments Through Time

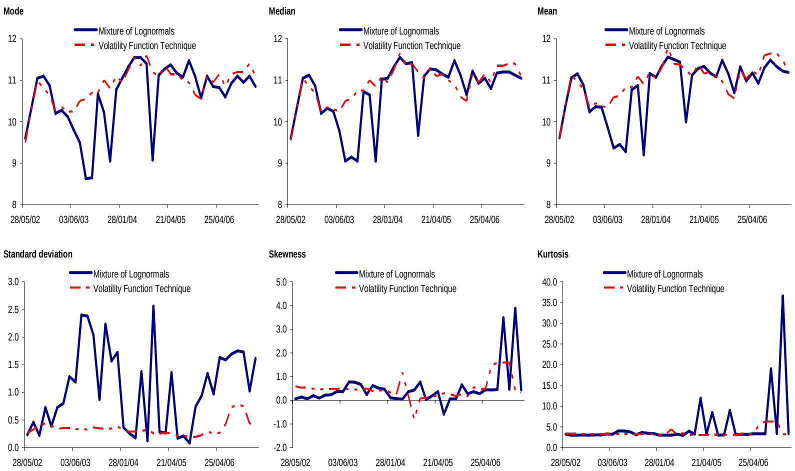

Table 1. Tests of Statistical Equality of the Implied Mean, Median and Variance for the Mixture Lognormal and the Volatility Function Technique: Exchange Rates

Date	Mean	Median	Variance
28/05/2002	0.8896	0.0206	0.0000
12/11/2002	0.0947	0.4327	0.0000
10/02/2003	0.0000	0.1782	0.0000
11/03/2003	0.0000	0.0000	0.0000
10/04/2003	0.0000	0.0000	0.7231
20/05/2003	0.0000	0.0000	0.0000
28/05/2003	0.0021	0.0016	0.0000
29/05/2003	0.3224	0.0455	0.0000
03/06/2003	0.0000	0.0000	0.0000
01/07/2003	0.0000	0.0000	0.0000
29/07/2003	0.0000	0.0000	0.0000
05/08/2003	0.0000	0.0000	0.0000
12/08/2003	0.0253	0.0124	0.0000
08/09/2003	0.0000	0.0000	0.0000
24/09/2003	0.0000	0.0000	0.0000
21/10/2003	0.2572	0.0099	0.0000
28/01/2004	0.0090	0.0005	0.0000
26/04/2004	0.0027	0.0000	0.0000
25/10/2004	0.0000	0.0000	0.0000
28/12/2004	0.0168	0.2080	0.0000
04/01/2005	0.0000	0.0654	0.0000
10/02/2005	0.0000	0.0000	0.0000
02/03/2005	0.4314	0.0560	0.0002
17/03/2005	0.0000	0.0000	0.0003
21/04/2005	0.0001	0.0055	0.0000
27/04/2005	0.0001	0.0990	0.0000
03/05/2005	0.0022	0.0005	0.0000
27/06/2005	0.0000	0.0000	0.0000
16/08/2005	0.0000	0.0000	0.0000
24/01/2006	0.0000	0.0003	0.0000
01/04/2006	0.0467	0.4291	0.0000
18/04/2006	0.1150	0.0376	0.0000
25/04/2006	0.1449	0.0080	0.0000
09/05/2006	0.0001	0.0000	0.0000
30/05/2006	0.0000	0.0000	0.0000
06/06/2006	0.0000	0.0000	0.0000
07/06/2006	0.0000	0.0000	0.0000
28/06/2006	0.0000	0.0000	0.0000
04/07/2006	0.9855	0.2184	0.0000

This table reports *p*-values for test of equality of the implied mean, median and variance for the methods analyzed: mixture lognormal and volatility function technique. The underlying variable is the Mexican peso -

USD exchange rate. Mean applies a *t*-test, median applies a Wilcoxon/ Mann-Whitney and variance applies an *F*-test. Option data was supplied by Banco de México. The number of observations chosen is 39 from days when both methods were able to estimate risk-neutral densities in one specific day. The sample size is from May 28th, 2002 to July 7th, 2006.

Table 2. Out-of-Sample Forecast Accuracy. MSEs for Mixture Lognormal and Volatility Function Technique: Exchange Rates.

MSE	Mixture Lognormal	Volatility Function Technique	Statistical Significance
Implied Mean	3.9520×10^{-1}	**7.1943×10^{-2}**	2.3419** (0.0220)
Implied Median	5.0716×10^{-1}	**5.6912×10^{-2}**	2.5285** (0.0136)
Implied Mode	6.7254×10^{-1}	**6.0169×10^{-2}**	2.5550** (0.0127)

This table reports the Mean-Square-Error (MSE) for the comparison between the implied mean, median, mode from the RNDs and the ex-post spot exchange rate. The implied values are for both methods analyzed: mixture lognormal and volatility function technique. The underlying variable is the Mexican peso - USD exchange rate. Option and exchange rate data was supplied by Banco de México. Statistical significance represents the Diebold and Mariano (1995) MSE equality test. The null hypothesis is MSE differences between the models are equal to zero. The statistic reported is a *t*-statistic and the *p*-value is expressed in parentheses. The sample size considers 37 observations of RNDs estimated applying both methods. Exchange rate call and put options were used in the estimations. The period under study is from May 28th, 2002 to July 7th, 2006. Bold indicates the smallest value i.e. best accuracy. ***/ Statistically significant at 1% confidence level, **/ Statistically significant at 5% confidence level.