Implied probability density functions: Estimation using hypergeometric, spline and lognormal functions

A thesis presented

by

André Duarte dos Santos

to

The Department of Finance

in partial fulfillment of the requirements

for the degree of

Master of Science in Finance

in the subject of

Finance

UNIVERSIDADE TÉCNICA DE LISBOA

Supervisor: Prof. Doutor João Guerra

Dissertation Committee:

Prof. Doutora Teresa Garcia, Chairman

Prof. Doutor Jorge Barros Luís

Prof. Doutor João guerra

Lisbon, Portugal

May 2011

Abstract

This thesis examines the stability and accuracy of three different methods to estimate Risk-Neutral Density functions (RNDs) using European options. These methods are the Double-Lognormal Function (DLN), the Smoothed Implied Volatility Smile (SML) and the Density Functional Based on Confluent Hypergeometric function (DFCH).

These methodologies were used to obtain the RNDs from the option prices with the underlying USDBRL (price of US dollars in terms of Brazilian reals) for different maturities (1, 3 and 6 months), and then tested in order to analyze which method best fits a simulated "true" world as estimated through the Heston model (accuracy measure) and which model has a better performance in terms of stability.

We observed that in the majority of the cases the SML outperformed the DLN and DFCH in capturing the "true" implied skewness. The DFCH and DLN methods were better than the SML model at estimating the "true" Kurtosis. However, due to the higher sensitivity of the skewness and kurtosis measures to the tails of the distribution (all the information outside the available strike prices is extrapolated and the probability masses outside this range can have infinite forms) we also compared the tested models using the root mean integrated squared error (RMISE) which is less sensitive to the tails of the distribution. We observed that using the RMISE criteria, the DFCH outperformed the other methods as a better estimator of the "true" RND.

Besides testing which model best captured the "true" world's expectations, we analyzed the historical summary statistics of the RNDs obtained from the FX options on the USDBRL for the period between June 2006 (before the start of the subprime crisis) and February 2010 (seven months before the Brazilian general election).

Acknowledgement

First of all, I would like to thank my supervisor, Professor João Guerra, for all the support and interesting discussions during the preparation of this thesis.

I would like also to acknowledge to my dear friends João Pedro, Pedro Gonçalves and Tiago Neves for their useful advice and help in the thesis preparation.

All my friends for their friendship and encouraging support.

At last I would like to thank all my family and my love for their unconditional love, their patient in the hard time when I did not have much time for them and because they had always believed in my work.

Contents

List of Figures

Chapter 1

Introduction

It is accepted by market participants that the prices of financial derivatives provide information about future expectations of the underlying asset prices, especially forwards, futures and options. Forwards and futures only give us the expected value for the underlying asset under the assumptions of risk neutrality, which makes using cross-sections of observed option prices more attractive because they allow estimation of an implied probability density function.

For market agents, the attractiveness of using an implied probability density function relies on being able to attribute probabilities to a range of future events, using market perceptions at a certain time. Several decision makers and analysts use this information source when analyzing market sentiment, uncertainty and extreme event scenarios, especially for interest rates and exchange rates.

It is known that the Black and Scholes model has several limitations, because it assumes that the price of the underlying asset evolves according to the geometric Brownian Motion (GBM) with a constant expected return and a constant volatility. The volatility is constant until maturity and also across all quoted strikes, which ignores phenomena like volatility smile and as such distorts probabilities for extreme scenarios. To tackle these problems, various methods have been suggested to extract Risk-Neutral Density Functions (RNDs) from option prices and several studies have been carried out to examine

the robustness of these estimates and their information power.

In this thesis we compare three methods of extracting RNDs from USDBRL European type exchange rate options. These methods are the Double-Lognormal Function, the Smoothed Implied Volatility Smile and the Density Functional Based on Confluent Hypergeometric function. We test the stability of the estimated RNDs and their robustness as regards small errors by randomly perturbing option prices by half of the quotation of the tick size as in Bliss and Panigirtzoglou (2002) before re-estimating the RNDs and their accuracy by experimenting their capacity to recover the "true" RNDs. The "true" probability density function (pdf) was estimated using the method developed in Cooper (1999), who generated pseudo prices from Heston's stochastic volatility model, and then compared the performance of the different methods using Monte Carlo simulations in order to obtain RNDs, whereby the input was the option prices calculated by these pseudo prices.

The remainder of this thesis is organized into seven chapters. Chapter Two gives a brief explanation of option pricing and a presentation of the Black and Scholes model and its theoretical background. We also describe the limitations of this model and its failure to capture the volatility smile contributions, due to the difference between the lognormal distribution mapped by the model and the real distribution of the underlying asset prices of the market (the difference between the theoretical B&S prices and the market prices). In this chapter, we also describe how option prices can provide information about implied probabilities given by market participants to future events and its use as an instrument to extract probability density functions of future prices using the formula proposed in Breeden and Litzenberger (1978).

Chapter Three describes some alternative option pricing methods that try to mitigate the limitations and restrictions of the B&S model, including the four models used in this thesis (DLN, SML, DFCH and Heston). Jondeau et al. (2006) divide the alternative methods into two categories: structural and non-structural. A structural model assumes a specific dynamic for the price or volatility process. A non-structural method allows the

estimation of a RND without describing any evolving process for the price or volatility of the underlying asset. The non-structural approaches can be divided into three subcategories: parametric (propose a form for the RND without assuming any price dynamics for the underlying asset), semi-parametric (suggest an approximation of the true RND) and non-parametric models (do not propose an explicit form for the RND).

Chapter Four explains the technical details of the strategies used in this thesis in order to estimate the RNDs and describe the measures used to evaluate the performance of the three models tested (MLN, SML and DFCH) in terms of accuracy and stability.

The results of the Monte Carlo simulation experiments and the comparisons of the models tested are presented and discussed in Chapter Five and Six. In Chapter Five we analyze the accuracy and stability performance using the "true" RNDs generated by the Heston parameters proposed in Cooper (1999). In Chapter Six, a similar analysis was carried out. However, the "true" RNDs were obtained through the previously calibrated Heston parameters. The Heston parameters were calibrated taking into account the observed quotes for the USDBRL European options between June 2006 and February 2010. The historical RND summary statistics obtained for the USDBRL in the time period described above are discussed in Chapter Seven. Finally, Chapter Eight presents the conclusions and discusses some research perspectives.

Chapter 2

Standard option pricing and extraction of RND

2.1 Option pricing and Black & Scholes model

Let us begin by introducing two elementary types of options. A European call option gives the buyer the right to buy the underlying asset for a certain price (strike price) at a certain date (maturity), whereas a European put option gives the buyer the right to sell the underlying asset for a certain price at a certain date. American options can be exercised at any time until expiration. In this thesis we will focus on European options. At maturity, the holder of the option only exercises it if he has a positive payoff (if the price of the underlying asset is above the exercise price for the call option or if the price of the underlying asset is below the exercise price for the put option).

Assuming that there are no transaction costs, we can represent the payoff of an European option at maturity through the following formulas (call option and put option), where X is the exercise price of the option, S_T is the price of the underlying asset at expiration date and T is the expiration date:

$$C(S_T, T, X) = \max(S_T - X; 0) \tag{2.1}$$

$$P(S_T, T, X) = \max(X - S_T; 0) \tag{2.2}$$

Intuitively, it can be inferred that the price of a call option reflects the ability to exercise the option when it brings a profit. This depends on the probability of the price of the underlying asset being greater than the strike price.

The widely used Black and Scholes model [Black and Scholes (1973)] for option pricing assumes that the underlying asset price has a lognormal distribution and evolves until reaching maturity in line with a geometric Brownian motion (GBM) stochastic process, with a constant expected return and a constant volatility:

$$dS_t = S_t \mu dt + S_t \sigma dW_t \tag{2.3}$$

where S_t is the price of the underlying asset at time t, dS_t denotes instantaneous price change, μ is the expected return, σ is the standard deviation of the price process and dW are increments from a Brownian motion process. The parameters μ and σ are assumed to be constant.

Besides constant volatility during the term of the option, the B&S model also assumes the same volatility across the whole range of strike prices.

Itô's Lemma states that an asset whose value depends on S_t and t has dynamics defined by the following stochastic differential equation:

$$df(S_t, t) = \left[\frac{1}{2}\frac{d^2 f}{dS_t^2}\sigma_t^2 + \frac{df}{dS_t}\mu_t + \frac{df}{dt}\right] dt + \frac{df}{dS_t}\sigma_t dW_t \tag{2.4}$$

Considering Itô's Lemma (see appendix A) and applying it to equation (2.3) results in S_t having a lognormal distribution and $\log(S_t) \sim N(\alpha, \beta)$ where $\alpha = \log(S_0) + (\mu - \frac{1}{2}\sigma^2)t$ and $\beta = \sigma^2 t$, which means that the underlying asset price has a lognormal distribution and the underlying returns are normally distributed.

If we consider a portfolio comprising one unit of a derivative asset and a short position

of Δ units ($\frac{df}{dS}$) of the underlying asset, we can apply the partial differential equation (2.4) to this portfolio getting the Black and Scholes partial differential equation (see Jondeau et al. (2006)):

$$\frac{1}{2}\frac{d^2f}{dS_t^2}S_t^2\sigma^2 + \frac{df}{dS_t}S_t r + \frac{df}{dt} - rf = 0 \tag{2.5}$$

The value of the option depends on r (risk free rate), σ and the boundary condition of the option contract in equations (2.1) and (2.2), respectively for calls and puts.

Solving the PDE in equation (2.5), in accordance with the boundary conditions, results in the Black and Scholes Pricing formula (call and put price):

$$C(S;t) = SN(d_1)\text{-}Xe^{-r(T-t)}N(d_2)\ ;\ S > 0\ ;\ t \in [0;T] \tag{2.6}$$

$$P(S;t) = Xe^{-r(T-t)}N(-d_2) - SN(-d_1)\ ;\ S > 0\ ;\ t \in [0;T] \tag{2.7}$$

with

$$d_1 = \frac{\ln(\frac{S}{X}) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{(T-t)}} \tag{2.8}$$

and

$$d_2 = \frac{\ln(\frac{S}{X}) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{(T-t)}} \tag{2.9}$$

We can observe that the parameter μ is not in equation (2.5), which means that the expected return does not appear in the B&S formula and consequently the value of the option does not depend on the investors' risk preferences (the solution of the equation is the same regardless of the risk premium required by each investor). In fact, instead of μ, equation (2.5) has r, which is the risk free rate (assumption that investors are risk neutral). In a world in which prices are lognormally distributed with constant volatility and expected returns, this theory allows option pricing and the creation of a risk free portfolio using delta hedging. The return of this hedged portfolio becomes certain and

does not depend on the change of the stock price.

2.2 Implied Volatility and limitations of the Black & Scholes model

The Black & Scholes model assumes that the price of the underlying asset follows a stochastic model with constant expected return and constant volatility. The final assumptions made by Black and Scholes' argument rely on the fact that if the future prices of the underlying asset are lognormally distributed, an option can be dynamically hedged using the underlying asset in order to build a portfolio that depends exclusively on the risk free rate.

However, in the real world we do not know the distribution of the prices in the future (traders do not have full knowledge of probabilities for future events) and dynamic hedging implies continuous trading (transaction cost problem, liquidity restrictions and not possible in practice).

The parameter regarding the instantaneous volatility in the underlying asset's return (σ) is not known. However, it can be estimated inverting Black and Scholes' formula in terms of σ (implied volatility) and then using market prices of options as inputs. The investors observe that the implied volatility calculated for each strike price is different, and that the implied volatilities are different across maturities (a volatility curve changes with maturity), which is not consistent with the Black and Scholes lognormal assumptions that define volatility as being constant across the whole range of strike prices and maturities. Implied volatilities observed in the market are a convex function of strike prices (usually out-of-the money and in-the-money options have higher volatility compared with at-the-money options), which creates the well known phenomenon called volatility smile.

The volatility smile indicates that traders make more complex assumptions about the path of the underlying asset price until maturity than the ones assumed by the GBM,

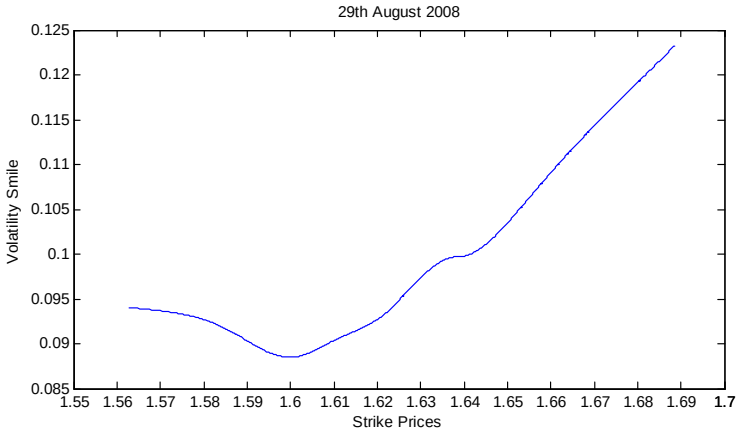


Figure 2-1: Volatility Smile curve at 29/08/2008 calculated using USDBRL options prices that expire in one month

which results in fatter tails of the true probability density function (pdf) when compared with a lognormal pdf. This indicates that the investors attribute higher probabilities to extreme events and that there is a gap between the true market RND and the Black and Scholes lognormal RND. In fact, higher volatilities for strike prices deep out-of-the-money make it more likely that future prices will be very different from current market values. This in turn increases the probability of these option prices being in-the-money in the future and leads to more expensive prices for deep out-of-the-money options, when compared to prices calculated through the B&S model.

2.3 Relation between option prices and the extraction of RNDs

It is possible to combine call options that have the same time to maturity but different exercise prices, in order to obtain a payoff at expiration that is dependent on the state of the economy at a particular time. The price of these combined securities (state-contingent securities) also reflects the probabilities that investors attribute to those particular states in the future.

This relation between probabilities and the price of a contingent claim[1] was initially proposed in Arrow (1964) [2] who applied a contingent claim model to the securities market. It was shown that the prices of an elementary claim (Arrow-Debreu security)[3] are proportional to the risk-neutral probabilities attached to each of the states.

This Arrow-Debreu security has an important information value and can be replicated with a combination of European call options, called *butterfly spread*, which consists of a long position in two calls with strikes $(X - \Delta M)$ and $(X + \Delta M)$ and a short position in two calls with strike (X) , where $\Delta M > 0$.

Breeden and Litzenberger (1978) applied the developments by Arrow and Debreu and used a state contingent claim in the form of a *butterfly spread* to show that the second partial derivative of a call option pricing function with respect to the strike prices yields the discounted RND $(f(S_T) \times e^{-rT})$.

In fact, a *butterfly spread* centered on X implies a payoff of ΔM if the price of the underlying asset at maturity T is equal to X (see Example 1).

Example 1 *(Breeden and Litzenberger (1978))*

Portfolio composed by $[c(0,T) - c(1,T)] - [c(1,T) - c(2,T)]$ with $T = Maturity$ will pay 1 unit if the state $M(T) = 1$ (butterfly spread centered in 1)

[1] a claim that can be made when a specific outcome occurs.

[2] who introduced uncertainty into the notion of competitive equilibrium and Pareto Optimality (Pareto equilibrium refers to a situation in economy where it's impossible to benefit an economic agent without harming another agent).

[3] a security paying one unit if a state s occurs and zero otherwise.

Aggregate Wealth	$c(0,T)$	$c(1,T)$	$c(2,T)$
$M(T)=0$	0	0	0
$M(T)=1$	1	0	0
$M(T)=2$	2	1	0
$M(T)=3$	3	2	1
.	.	.	.
$M(T)=N$	N	$N-1$	$N-2$

Aggregate Wealth	$c(X-\Delta M,T)$	$c(X,T)$	$c(X+\Delta M,T)$	Payoff of Butterfly Spread
$M(T)=X-\Delta M$	0	0	0	0
$M(T)=X$	ΔM	0	0	ΔM
$M(T)=X+\Delta M$	$2\Delta M$	Δ	0	0
$M(T)=X+2\Delta M$	$3\Delta M$	2Δ	Δ	0
.	.	.	.	0
$M(T)=X+N\Delta M$	$(N+1)\Delta M$	$N\Delta M$	$(N-1)\Delta M$	0

These authors also show that this relation can be generalized in the following formula (portfolio that pays 1 if scenario $M(T)=X$ occurs in T periods):

$$P(M,T;\Delta M)=\frac{[c(M-\Delta,T)-c(M,T)]-[c(M,T)-c(M+\Delta,T)]}{\Delta M} \tag{2.10}$$

with the $P(M,T;\Delta M)$ being the price of the elementary claim security in the discrete case (to have a payoff of 1 in the state $M(T)=X$ we have to buy $1/\Delta M$ units of the butterfly spread). For continuous M (step size Δ tends to zero) the price of the butterfly spread at state $M=X$ is the second partial derivative of the portfolio of calls with respect to X (Strike Price):

$$\lim_{\Delta M\to 0}\frac{P(M,T;\Delta M)}{\Delta M}=\frac{d^2C(X,T)}{dX^2}|_{X=M} \tag{2.11}$$

The price of an Arrow-Debreu security is equal to its expected payoff, which is calculated by multiplying the present value of 1 by the risk-neutral probability corresponding to its state (discounted using risk free rate). Applying this relation to a range of continuum possible future values for the underlying asset, leads to the estimation of the discounted Risk-Neutral Density:

$$\frac{d^2C(X,T)}{dX^2} = e^{-rT} f(S_T) \tag{2.12}$$

This condition only holds if $C(X,T)$ is monotonic decreasing and convex in the exercise price, otherwise there are arbitrage opportunities and the RND could be negative [Bahra (1997)].

Chapter 3

RND estimation - Alternative methods

Despite being widely used, the B&S model has several limitations because the log normal assumption does not hold in practice and calculates prices that are different from market values, which creates the need to analyze and study different methods in order to find a model that is more efficient at capturing market expectations and prices.

There are many alternative models to estimate the Risk-Neutral Density (RND). According to Jondeau et al. (2006), the models can be divided into two categories: structural and non-structural. A model is structural if it takes on a specific price dynamic and proposes a certain volatility process. A non-structural model yields a RND without describing the dynamics for the price or volatility.

In this chapter we give an overview of some methods developed in order to obtain estimates which closely reflect the expectations of the option market.

3.1 Structural Models

3.1.1 Jump Diffusion Model

In the structural category we can find stochastic models like the one developed in Malz (1996) which consists of assuming a stochastic process for the underlying asset, where S_t (log normal jump diffusion) corresponds to the sum of a Geometric Brownian Motion (GBM) and a Poisson jump process. The price process is

$$dSt = \mu S_t dt + \sigma S_t dW_t + k S_t dq_t \tag{3.1}$$

where q_t represents a variable with a Poisson distribution with the parameters k being the jump dimension and λ the average rate of jump occurrence.

For simplicity, Malz assumed that until the maturity of the option it will be at most one jump of constant size (referred to as Bernoulli version of jump diffusion), which results in the following prices for calls and puts:

$$\begin{aligned} C =& (1 - \lambda T) C_{BS}(S_t, T, K, \sigma, r, r^* + \lambda k) \\ &+ (\lambda T) C_{BS}(S_t(1+k), T, K, \sigma, r, r^* + \lambda k) \end{aligned} \tag{3.2}$$

$$\begin{aligned} P =& (1 - \lambda T) P_{BS}(S_t, T, K, \sigma, r, r^* + \lambda k) \\ &+ (\lambda T) C_{BS}(S_t(1+k), T, K, \sigma, r, r^* + \lambda k), \end{aligned} \tag{3.3}$$

where $(1 - \lambda T)$ represents the probability of no jump before maturity, C_{BS} and P_{BS} are the Black and Scholes pricing formulas for call and put options. After estimating the model's parameters, they are inserted into a pdf function in order to obtain the RND.

This kind of methods can be used when analyzing markets without option prices or with scarce liquidity of this kind of derivatives.

3.1.2 RND estimation using a model based on stochastic volatility - Heston Model

The Heston Model was developed in Heston (1993) and represents the classical stochastic volatility pricing model. It is used in this thesis to estimate the density corresponding to the 'true' world. This method adds a second Wiener Process to the price dynamics (volatility modeling), which leads to the dynamics of the underlying asset price (S_t) based on the geometric Brownian Motion with time varying volatility,

$$\begin{aligned} dS_t &= \mu S_t dt + S_t \sqrt{v_t} dZ_{1,t} \\ dv_t &= \kappa(\theta - v_t)dt + \sigma_v \sqrt{v_t} dZ_{2,t} \end{aligned} \tag{3.4}$$

where $\sqrt{v_t}$ denotes current volatility of the underlying asset price, $Z_{1,t}$ and $Z_{2,t}$ represents the correlated Brownian motion processes with correlation parameter ρ, v_t is the volatility of the underlying asset, θ is the long run volatility, σ_v is the volatility of the volatility process and κ is the speed at which the volatility returns to its long run average.

These parameters guide the trajectory of the square root process, which means that along its path, v_t goes around θ, crossing the long run volatility more frequently when k is higher and the trajectory of v_t is more volatile when σ is higher.

The parameter ρ defines the correlation between returns and volatility and can change the form of the RND, generating skewness in asset returns. For example, if $\rho > 0$ the volatility of the asset price increases when the asset price increases, and in this way the weight of the right tail of RND will increase. In contrast, when $\rho < 0$ the decrease in price leads to an increase in volatility and the weight of the left tail of RND will increase.

The derivation of the Heston option pricing formula also uses Itô's lemma. Like in the Black & Scholes model, in order to obtain a risk-neutral portfolio, Heston's model also considers a portfolio of assets. Nevertheless, the volatility needs to be hedged due to its stochastic nature, so a second derivative is added. For example, a short position of

one unit of a call option is covered by a long position of δ units of the underlying asset and γ units of a second derivative on the same underlying asset:

$$d\Pi_t = r(C - \delta S_t - \gamma C_1)dt, \tag{3.5}$$

where Π_t is the portfolio at time t, C is the covered call option, S_t is the underlying asset and C_1 is the second option on the same underlying asset.

Heston introduced the following closed formula for the European call option price:

$$\begin{aligned}
C(S_t, v_t, X, T) &= S_t e^{-r^*(T-t)} P_1 - X e^{-r(T-t)} P_2 \ ; \ S > 0 \ ; \ t \in [0; T], \\
P_j &= \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathrm{Re}[\frac{e^{-i\phi \ln(k)} f_j(\ln(S_t), v_0, T-t, \phi)}{i\phi}] d\phi, \\
f_j(\ln(S_t), v_0, T-t, \phi) &= e^{C(T-t,\phi) + D(T-t,\phi) v_t + i\phi \ln(S_t)}, \\
C(T-t, \phi) &= (r - r^*)\phi i (T-t) + \frac{a}{\sigma_v^2}\{(b_j - \rho\sigma_v \phi i + d)(T-t) \\
&\quad - 2\ln[\frac{1 - g e^{d(T-t)}}{1-g}]\}, \\
D(T-t, \phi) &= \frac{b_j - \rho\sigma_v\phi i + d}{\sigma_v^2}[\frac{1 - e^{d(T-t)}}{1 - g e^{d(T-t)}}], \\
g &= \frac{b_j - \rho\sigma_v\phi i + d}{b_j - \rho\sigma_v\phi i - d}, \\
d &= \sqrt{(\rho\sigma_v\phi i - b_j)^2 - \sigma_v^2(2u_j\phi i - \phi^2)}, \\
u_1 &= \frac{1}{2}, u_2 = -\frac{1}{2}, a = \kappa\theta, b_1 = \kappa + \lambda - \rho\sigma_v, b_2 = \kappa + \lambda, i = \sqrt{-1}
\end{aligned} \tag{3.6}$$

3.2 Non-Structural Models

3.2.1 Parametric models

A model is considered parametric if it proposes a RND without assuming a specific price or volatility dynamic and proposes a form for the RND without assuming any price

dynamics for the underlying asset.

Mixture of lognormal distribution

The mixture of lognormals (MLN) was proposed by Bahra (1997) and Melick and Thomas (1997) and assumes a functional form for the risk-neutral density (RND) that accomodates various stochastic processes for the underlying asset price. Instead of specifying a dynamic for the underlying asset price (which leads to a unique terminal value), it is possible to make assumptions about the functional form of the RND function itself and then obtain the parameters of the distribution by minimizing the distance between the observed option prices and those that are generated by the assumed parametric form. According to the authors, this makes this model more flexible than the Black and Scholes model and increases its ability to capture the main contributions to the smile curve, namely the skewness and the kurtosis of the underlying distribution.

It is known that the prices of European call and put options can be expressed as the discounted sum of all expected future payoffs:

$$\begin{aligned} C_0(X,T) &= e^{-rT}\int_X^\infty q(S_t)(S_t - X)dS_t \\ P_0(X,T) &= e^{-rT}\int_X^\infty q(S_t)(X - S_t)dS_t \end{aligned} \tag{3.7}$$

According to Bahra (1997), any functional form for the RND $q(S_t)$ can be assumed because the parameters would be estimated through optimization (minimizing the difference between the prices obtained through the MLN model and market prices). Nevertheless, the author assumed that the asset price distributions are closer to the lognormal distribution and consequently it would be plausible to use a weighted sum of lognormal density functions,

$$q(S_t, \theta) = \sum_{i=1}^{k} [w_i L(\alpha_i, \beta_i, S_t)] \tag{3.8}$$

where $L(\alpha_i, \beta_i, S_t)$ is the ith lognormal distribution with parameters α_i and β_i . It has the following expression:

$$\begin{aligned} L(\alpha_i, \beta_i, S_t) &= \frac{1}{S_t \beta_i \sqrt{2\pi}} e^{[-(\ln(S_t)-\alpha_i)^2/2\beta_i^2]}, \\ \alpha_i &= \ln(S_t) + (\mu_i - \frac{1}{2}\sigma_i^2)(T-t), \\ \beta_i &= \sigma_i \sqrt{(T-t)}. \end{aligned} \tag{3.9}$$

The term θ represents the vector of unknown parameters α_i, μ_i, σ_i for $i = 1, ..., k$, and k defines the number of mixtures describing the RND. In order to guarantee that q is a probability density, $w_i \geqslant 0$ for $i = 1, ...k$, and $\sum_{i=1}^{k} w_i = 1$. In this way q will be a combination of the lognormal densities.

While Melick and Thomas applied this method on the extraction of RNDs from the prices of American options on crude oil futures using a mixture of three independent lognormals, Bahra obtained the RNDs using European options on LIFFE equity index, LIFFE interest rate options and Philadelphia Stock Exchange currency options using a mixture of two independent lognormals. The choice of a mixture of two lognormals is based on the lower number of parameters to be estimated (5 parameters). In fact, options are traded across a relatively small range of exercise prices, hence there are limits on the number of parameters that can be estimated from the data.

Extending the mixture of lognormals to the European call option prices given by equation (3.7) we have the following option prices for each strike price (X_i) and with time to maturity $\tau = (T - t)$:

$$c(X_i,\tau) = e^{-r\tau}\int_X^\infty (S_t - X)\sum_{i=1}^k w_i L(\alpha_i,\beta_i,S_T)dS_t, \tag{3.10}$$
$$c(X_i,\tau) = e^{-r\tau}\sum_{i=1}^k w_i \int_X^\infty (S_t - X)L(\alpha_i,\beta_i,S_T)dS_t.$$

The integral in equation (3.10) can be rewritten as (see Jondeau et al. (2006)):

$$\begin{aligned} c(X_i,\tau) &= e^{-r\tau}\sum_{i=1}^k w_i e^{\alpha_i+\frac{1}{2}\beta_i^2} N(\frac{-\ln(X)+\alpha_i+\beta_i^2}{\beta_i}) \\ &\quad -e^{-r\tau}X\sum_{i=1}^k N(\frac{-\ln(X)-\alpha_i}{\beta_i}]. \end{aligned} \tag{3.11}$$

Applying the mixture of two lognormals used by Bahra, we get the following closed formula for a European call option,

$$\begin{aligned} c(X,\tau) &= e^{-r\tau}\{w[e^{\alpha_1+\frac{1}{2}\beta_1^2}N(d_1) - XN(d_2)] \\ &\quad +(1-w)[e^{\alpha_2+\frac{1}{2}\beta_2^2}N(d_3) - XN(d_4)]\} \end{aligned} \tag{3.12}$$

where

$$\begin{aligned} d_1 &= \frac{-\ln(X)+\alpha_1+\beta_1^2}{\beta_1}, \\ d_2 &= d_1 - \beta_1, \\ d3 &= \frac{-\ln(X)+\alpha_2+\beta_2^2}{\beta_2}, \\ d_4 &= d_3 - \beta_2. \end{aligned} \tag{3.13}$$

For the European put option, Bahra presented the following pricing formula,

$$\begin{aligned} p(X,\tau) &= e^{-r\tau}\{w[-e^{\alpha_1+\frac{1}{2}\beta_1^2}N(-d_1) - XN(-d_2)] \\ &\quad +(1-w)[-e^{\alpha_2+\frac{1}{2}\beta_2^2}N(-d_3) - XN(-d_4)]\}. \end{aligned} \tag{3.14}$$

In order to find the parameters of the implied RND (vector θ) we have to solve the minimization problem,

$$\begin{aligned} &\min_{\alpha_1,\alpha_2,\beta_1,\beta_2,w_i} \sum_{i=1}^{n}[c(X,\tau) - \widehat{c}]^2 + \sum_{i=1}^{n}[p(X,\tau) - \widehat{p}]^2 \\ &+[we^{\alpha_1+\frac{1}{2}\beta_1^2} + (1-w)e^{\alpha_2+\frac{1}{2}\beta_2^2} - e^{r\tau}S] \end{aligned} \tag{3.15}$$

where the first two terms refer to the sum of the squared deviation between option prices estimated through MLN and the observed market prices. Call and put prices can be considered in equation (3.15) because both refer to the same underlying distribution. The third term of the equation states that the expected value of the RND must be equal to the forward price of the underlying asset in order to avoid the violation of the arbitrage condition (martingale condition). After estimating parameters $\alpha_1, \alpha_2, \beta_1, \beta_2, w$, we insert them into equation (3.8) and then the implied RND is obtained.

The optimization problem (3.15) can be affected by a problem related to the symmetry between the densities because in an optimization program, various parameter vectors can be associated to the same density, which in turn can result in numerically unstable programs where the optimizer goes round in an infinite loop. In Jondeau et al. (2006), the authors recommended the imposition of $\beta_1 > \beta_2$ (first density will have a larger standard deviation than the second one) in order to avoid this symmetry problem.

This model was tested in this thesis for the extraction of the RND from the currency option USDBRL. The details of the method applied are explained in section 4.4.1.

Mixture of hypergeometric functions

This method allows the estimation of a probability density function (pdf) without assuming a specific functional form for it. It consists of the use of a formula that encompasses various densities, such as normal, gamma, inverse gamma, weibull, pareto and mixtures of these probability densities.

In Abadir and Rockinger (2003), the authors developed a function based on the confluent hypergeometric function (${}_1F_1$), also known as the function for the case of double integrals of densities. These authors believe the usefulness of ${}_1F_1$ relies on the fact that it includes special cases of incomplete gamma, normal distributions and mixtures of the two. In fact, this function has the advantage of being more efficient than fully nonparametric estimation for small samples and more flexible than parametric methods because it does not restrict functional forms.

The Kummer function ${}_1F_1$ can be defined by:

$$
\begin{aligned}
{}_1F_1 &\equiv \sum_{j=0}^{\infty} \frac{(\alpha)_j}{\beta_j}\frac{z^j}{j!} \equiv 1 + \frac{\alpha}{\beta}z + \frac{\alpha(\alpha+1)}{\beta(\beta+1)}\frac{z^2}{2} + \frac{\alpha(\alpha+1)}{\beta(\beta+1)}\frac{z^2}{2} + ..., \qquad (3.16)\\
(\alpha)_j &\equiv (\alpha)(\alpha+1)...(\alpha+j-1) \equiv \frac{\Gamma(a+j)}{\Gamma(a)}
\end{aligned}
$$

with $\Gamma(v)$, for $v \in \mathbb{R}$ being the gamma function and $\beta \in \mathbb{N}$.

The function considered for option pricing is called DFCH (density function based on confluent hypergeometric functions) and specifies the European call price as a mixture of two confluent hypergeometric functions:

$$
\begin{aligned}
C(X) = c_1 + c_2X + l_{X>m_1}a_1((X-m_1)^{b_1}){}_1F_1(a_2;a_3;b_2(X-m_1)^{b_3}) \qquad (3.17)\\
+ (a_4){}_1F_1(a_5;a_6;b_4(X-m_2)^2),
\end{aligned}
$$

where $a_{3,}a_6 \in \mathbb{N}$, $b_2, b_4 \in \mathbb{R}^-$ and $a_{,}, a_2, a_4, a_5, b_1, b_3 \in \mathbb{R}$. The indicator function l

represents a component of the density with bounded support.

The first ${}_1F_1$ function can represent a gamma or other asymmetric generalizations, whereas the second ${}_1F_1$ covers symmetric quadratic exponentials, such as the normal.

To get the implied probability density function, the formula stated in equation (2.12) is applied to $C(X)$:

$$
\begin{aligned}
e^{-rT} f(X) = \frac{d^2 C(X)}{dX^2} &= l_{X>m1} a_1 (X - m_1)^{b1-2} [b_1 (b_1 - 1)_1 F_1(a_2; a_3; b_2 (X - m_1)^{b_3}) \\
&+ \frac{a_2}{a_3} b_2 b_3 (2b_1 + b_3 - 1)(X - m_1)^{b_3} \\
&\times_1 F_1(a_2 + 1; a_3 + 1; b_2 (X - m_1)^{b_3}) + \frac{a_2 (a_2 + 1)}{a_3 (a_3 + 1)} b_2^2 b_3^2 (X - m_1)^{2b_3} \\
&\times_1 F_1(a_2 + 2; a_3 + 2; b_2 (X - m_1)^{b_3})] \\
&+ 2a_4 \frac{a_5}{a_6} b_4 [{}_1F_1(a_5 + 1; a_6 + 1; b_4 (X - m_2)^2) \\
&+ 2\frac{a_5 + 1}{a_6 + 1} b_4 (X - m_2)^2 {}_1F_1(a_5 + 2; a_6 + 2; b_4 (X - m_2)^2)].
\end{aligned}
\tag{3.18}
$$

The integral of the density is given by:

$$
\begin{aligned}
\frac{dC(X)}{dX} &= c_2 + l_{X>m_1} a_1 (X - m_1)^{b_1 - 1} [(b_1)_1 F_1(a_2; a_3; b_2 (X - m_1)^{b_3}) \\
&+ \frac{a_2}{a_3} b_2 b_3 ((X - m_1)^{b_3})_1 F_1(a_2 + 1; a_3 + 1; b_2 (X - m_1)^{b_3})] \\
&+ 2a_4 \frac{a_5}{a_6} b_4 (X - m_2)_1 F_1(a_5 + 1; a_6 + 1; b_4 (X - m_2)^2).
\end{aligned}
\tag{3.19}
$$

As stated above, some restrictions must be set in order to guarantee that $f(X)$ integrates to 1 between X_l and X_u,

$$
\int_{X_l}^{X_u} f(X) dX = 1. \tag{3.20}
$$

Through these restrictions, we obtained the following expressions for c_2 and a_4 (the details are presented in the Appendix).

$$c_2 = -1 + a_4\sqrt{-b_4\pi}, \tag{3.21}$$
$$a_4 = \frac{1}{2\sqrt{-b_4\pi}}\left[1 - a_1(-b_2)^{-a_2}\frac{\Gamma(a_3)}{\Gamma(a_3 - a_2)}\right].$$

As X tends to ∞, the value of the call option price will be approximately 0 ($C(\infty) = 0$), which is the obvious conclusion for call options very nearly out of the money (the probability to become in the money is near 0). The option value in equation (3.17) will pay a minimum of c_1, which leads to the following simplification:

$$c_1 = -c_2 m_2$$

Using assumptions $b_1 = 1 + a_2 b_3, a_5 = -\frac{1}{2}, a_6 = \frac{1}{2}$, formula (3.17) can be further simplified (see Abadir and Rockinger (2003)).

The final reduction was based on the no-arbitrage condition $S_t = \exp^{-r(T-t)} E(S_T)$, with r being the risk free rate and $E(X)$ the expected value of the underlying price at maturity,

$$E(X) = a_1\frac{\Gamma(a_3)}{\Gamma(a_3 - a_2)}(-b_2)^{-a_2}(m_1 - m_2) + m_2. \tag{3.22}$$

With the restrictions defined above, the number of parameters to estimate in the calculation of the theoretical price in equation (3.17) is reduced to seven.

In order to obtain the implicit RND we have to proceed with the following minimization problem:

$$\min_{a_2,\alpha_3,b_2,b_3,b_4,m_1,m_2} \sum_{i=1}^{n} [c(X_i, \tau) - \widehat{c_i}]^2 \tag{3.23a}$$

where a_2, a_3, b_2, b_3, b_4, m_1 and m_2 are the parameters to be estimated. Given the

restrictions above, $c(X, \tau)$ is the theoretical price given in equation (3.17), $\widehat{c}$ are the option prices observed in the market and n is the number of strike prices. The RND is obtained by inserting the parameters into equation (3.18).

The details about the extraction process of the implied RNDs using this method are described in section 4.4.2.

3.2.2 Non-parametric models

A model is considered non-parametric if it does not propose an explicit form for the RND.

Spline methods

This method consists of the derivation of the RND using the results of Breeden and Litzenberger (1978), but with a preliminary process of smoothing the volatility smile. The first approach using this method was made by Shimko (1993), who proposed smoothing the volatility smile via a low order polynomial (using a quadratic polynomial) that fitted the implied volatilities (on the y-axis) and the associated strike prices (on the x-axis),

$$\sigma_i = a_0 + a_1 K_i + a_2 K_i^2, \text{ for } i = 1, ..., N, \tag{3.24}$$

with N as the number of observed strike prices. The continuous implied volatility function obtained (on strike prices space) is then inserted back into Black and Scholes formula (2.6) and the probability density function is obtained through $\frac{dC^2}{dS^2}$. The option currency markets are quoted in terms of implied volatility for a specific delta ($\Delta = \frac{dC}{dS}$), which makes it necessary to convert the deltas into strike prices via the Black and Scholes model.

Malz (1996) applied smoothing of the volatility smile using the delta as the x-axis instead of the strike price. Using delta rather than strike, away-from-the-money groups implied volatilities closer than near-the-money implied volatilities, which gives more weight

to the centre of the distribution where the data is more reliable (more frequently traded).

Campa et al. (1997) used the spline method instead of the quadratic polynomial to smooth the smile curve. A natural cubic spline was applied using the strike prices as the X-axis. This method allows the smoothness of the fitted curve to be controlled and is less restrictive about the shape of the fitted function.

Bliss and Panigirtzoglou (2002) applied a natural cubic spline in the volatility/delta space.

The cubic spline interpolation consists of connecting the adjacent points (Δ_i, σ_i), $(\Delta_{i+1}, \sigma_{i+1})$, using the cubic functions $\hat{\sigma}_i; i = 0, ..., n-1$ in order to piece together a curve with continuous first and second order derivatives.

$$\hat{\sigma}_i = \begin{cases} \hat{\sigma}_0(\Delta) & \text{if } \Delta < \Delta_1 \\ \hat{\sigma}_1(\Delta) & \text{if } \Delta_1 < \Delta < \Delta_2 \\ \vdots & \\ \hat{\sigma}_{n-1}(\Delta) & \text{if } \Delta_{n-1} < \Delta < \Delta_n \\ \hat{\sigma}_n(\Delta) & \text{if } \Delta > \Delta_n \end{cases} \tag{3.25}$$

where $\hat{\sigma}_i$ is a third order polynomial defined by:

$$\hat{\sigma}_i(\Delta) = d_i + c_i(\Delta - \Delta_i) + b_i(\Delta - \Delta_i)^2 + a_i(\Delta - \Delta_i)^3 \tag{3.26}$$

with Δ being in the interval $[\Delta_i, \Delta_{i+1}]$. At Δ_i the value of the function is d_i.

The first and second derivatives of equation (3.26) are:

$$\hat{\sigma}_i'(\Delta) = c_i + 2b_i(\Delta - \Delta_i) + 3a_i(\Delta - \Delta_i)^2, \tag{3.27}$$

$$\hat{\sigma}_i''(\Delta) = 2b_i + 6a_i(\Delta - \Delta_i), \tag{3.28}$$

which means that the second-order derivative ($\hat{\sigma}_i''$) is given as a linear interpolation between knot points.

The condition that the cubic functions $\hat{\sigma}_i$ and $\hat{\sigma}_{i-1}$ must meet at the point (Δ_i, y_i) is expressed as:

$$\hat{\sigma}_{i-1}(\Delta_i) = \hat{\sigma}_i(\Delta_i) = y_i \tag{3.29}$$
$$y_i = d_i = a_{i-1}(\Delta_i - \Delta_{i-1})^3 + b_{i-1}(\Delta_i - \Delta_{i-1})^2 + c_{i-1}(\Delta_i - \Delta_{i-1}) + d_{i-1}$$

The conditions regarding the continuous nature of the first and second derivatives in the knot points are:

$$\hat{\sigma}'_{i-1}(\Delta_i) = \hat{\sigma}'_i(\Delta_i) \tag{3.30}$$
$$3a_{i-1}(\Delta_i - \Delta_{i-1})^2 + 2b_{i-1}(\Delta_i - \Delta_{i-1}) + c_{i-1} = c_i$$

$$\hat{\sigma}''_{i-1}(\Delta_i) = \hat{\sigma}''_i(\Delta_i) \tag{3.31}$$
$$6a_{i-1}(\Delta_i - \Delta_{i-1})^2 + 2b_{i-1}(\Delta_i - \Delta_{i-1}) = 2b_i$$

In Bliss and Panigirtzoglou (2002) the authors used a natural smoothing spline, whereby the second order derivatives in the extreme knot points were 0, $S''(x_0) = 0$ and $S''(x_n) = 0$ (leading to a spline function that is linear outside the range of avail-

able data). This condition can result in negative values when extrapolating outside the extreme points, which can yield a negative fitted fdp in the extrapolated points (in this thesis we did not have this problem). In a natural spline, the smoothness of the interpolating polynomial is controlled by a smoothness parameter λ, which weights the degree of curvature of the spline function. According to Bliss and Panigirtzoglou (2002), the cubic interpolating spline has the disadvantage of following the same random fluctuations as the data points, which distorts the nature of the underlying function, which explains why they used a cubic smoothing spline.

The natural spline minimizes the following objective function:

$$\min_{\theta}(1-\lambda)\sum_{i=1}^{N} w_j(\sigma_i - \hat{\sigma}_i(\Delta_i,\theta))^2 + \lambda \int_{-\infty}^{\infty} (\sigma''(\Delta;\theta))^2 d\Delta, \tag{3.32}$$

where N is the number of quoted deltas ($\Delta = \frac{dC}{dS}$), $\hat{\sigma}_i(\Delta_i,\theta)$ is the implied volatility corresponding to the spline parameters represented by vector θ and w_i represents the weight attributed to each observation. The first term measure the goodness of fit and the second term measures the smoothness of the spline. If $\lambda = 0$ the cubic spline has an exact fit to the data (the closeness of the spline to the data is the only concern). If $\lambda = 1$ the interpolating function will be a straight line (smoothness is all that matters).

Chapter 4

Accuracy and Stability analysis of the tested PDF estimation methods

4.1 Data

The RNDs analyzed in this thesis were extracted from currency OTC options with the underlying USDBRL (price of US dollars in terms of Brazilian reals).

The quotes used as inputs were taken from the daily settlement bid prices in Bloomberg for Offshore USDBRL FX Options [1] . The data collected covers the period from June 2006 (half a year before the problems regarding the subprime crisis started to worsen) to February 2010 (seven months after the Brazilian general election) and comprises the monthly quotes (end of month prices).

This four-year period witnessed economic growth in Brazil, despite the financial crisis. In fact, the worst global recession since the 1930s left Brazil relatively unscathed (it was one of the last countries that experienced a downturn and one of the first to recover; the economy shrank for only two quarters as can be seen in figure (11-10). The Brazilian

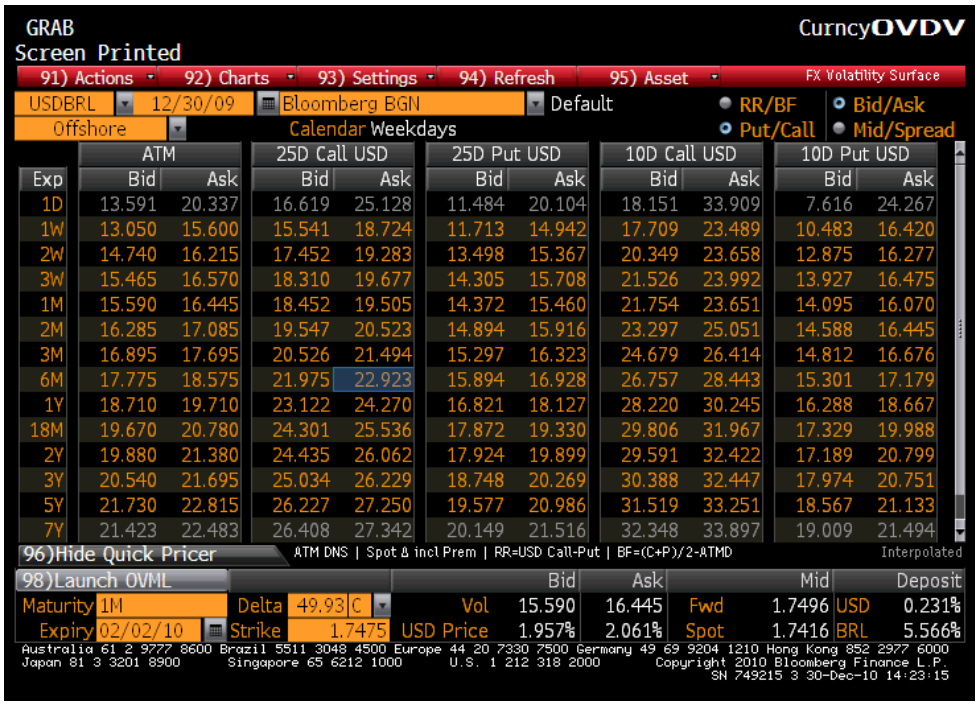
[1]Information provided by Bloomberg for the OTC Market. The USDBRL is quoted in volatility in terms of delta according to international conventions (does not use the specific maturity of BM&F calendar and a day count of business days/252 just like other financial instruments traded in BM&F)

Real was introduced in December 1993 and succeeded the Cruzeiro Real as the Brazilian currency in order to solve chronic problems at that time like hyperinflation [2] and unstable exchange rates (these two problems were caused mainly by inflationary expectations). Within a year, the Real plan had managed to control price rises and after 1999 the exchange-rate peg was abandoned and the currency was allowed to float. As such, the data included was obtained in an environment of free-floating currency market.

The calls and puts used are of the European type and are priced in volatility as a function of delta. As shown in the screen below, the grid of quoted deltas is 0.05, 0.1, 0.15, 0.25, 0.35 and 0.5 deltas. This means that we only considered out-of-the money options (calls and puts) and at-the-money options[3], which confirms the general understanding that out-the-money options tend to be more liquid than in-the-money options. In this thesis, we estimate the RNDs using 1, 3 and 6 months to maturity options.

[2] for example, according to the official numbers of Instituto Brasileiro de Geografia e Estatística, the Brazilian CPI (Consumer Price Index) was always above 25% from January 1993 to June 1994.

[3] The delta value varies from 0 for very out-the-money options to 1 for deeply in-the-money options. At the money options have a delta close to 0.5.



OTC USDBRL European options quotes

4.2 Testing PDF estimation techniques using Monte Carlo approach

This section describes the method used to test the performance of the three estimation techniques applied in this work and explained in Chapter 3: the Double-Lognormal Function (DLN), the Smoothed Implied Volatility Smile (SML) and the Density Functional Based on Confluent Hypergeometric Functions (DFCH).

To test the accuracy of these methods at capturing the risk-neutral density functions, we have to see how closely they fit the true risk-neutral density. Unfortunately, the true RND is unobservable, so we use the method proposed in Cooper (1999).

In the absence of the true RND, Cooper suggested the use of simulated option prices data that correspond to a given risk-neutral density function, and then, using these

simulated prices as input, test what methods produce a better performance in recovering the given RND.

To generate the "true" risk-neutral density functions, Cooper applied the Heston stochastic volatility model because it is an interesting technique able to generate a wide range of different shapes reflecting different market conditions: high or low volatility, positive or negative skewness and it is also able to generate data for a full range of maturities.

As explained previously, under the Heston model the underlying asset price dynamics is described by equation (4.1):

$$\begin{aligned} dS_t &= \mu S_t dt + S_t \sqrt{v_t} dZ_{1,t} \\ dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t} dZ_{2,t} \ , \end{aligned} \tag{4.1}$$

where σ is the volatility of volatility and θ is the long run volatility. The correlation between Z_1 and Z_2 is measured by ρ (correlation between returns and volatility) and can change the form of the RND generating skewness in asset returns. For example, if ρ is negative, there is a negative correlation between shocks to asset price and volatility, which means that a negative shock to the price will increase the volatility and consequently increase the likelihood of getting further big downward movements. A positive correlation between asset price and volatility has the opposite effect. The figure 4-1 shows the effect of changing ρ on the RND.

Heston (1993) shows that under stochastic volatility assumptions, the European call options have the closed form given in equation (3.6).

In order to obtain the true density and its associated summary statistics, we apply the second partial derivative of equation $C(S_t, v_t, X, T)$ with respect to the strike price (d^2C/dX^2) Breeden and Litzenberger (1978).

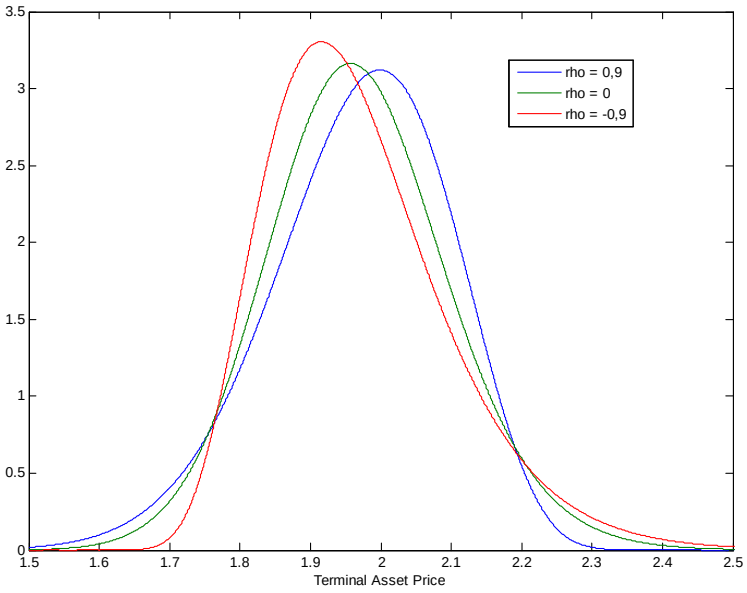


Figure 4-1: Implied RND under aternative values for the correlation parameter

In order to test the ability of the estimation methods tested to capture a wide range of possible shapes of the "true" RNDs, we establish a set of six scenarios divided into low and high volatility and which have three levels of skewness (strong negative skewness, weak positive skewness and strong positive skewness) as in Cooper (1999).

Table 4-1: Parameters used in Heston model under each scenario

	Strong negative Skew	Weak positive skew	Strong positive skew
	Scenario 1	Scenario 2	Scenario 3
Low volatility	$\kappa = 2, \sqrt{\theta} = 0.1$	$\kappa = 2, \sqrt{\theta} = 0.1$	$\kappa = 2, \sqrt{\theta} = 0.1$
	$\sigma = 0.1, \rho = -0.9$	$\sigma = 0.1, \rho = 0$	$\sigma = 0.1, \rho = 0.9$
	Scenario 4	Scenario 5	Scenario 6
High volatility	$\kappa = 2, \sqrt{\theta} = 0.3$	$\kappa = 2, \sqrt{\theta} = 0.3$	$\kappa = 2, \sqrt{\theta} = 0.3$
	$\sigma = 0.4, \rho = -0.9$	$\sigma = 0.4, \rho = 0$	$\sigma = 0.4, \rho = 0.9$

In generating these scenarios we try to replicate the environment from USDBRL FX Options. Therefore, as input we considered a grid of strike prices which results from the

average of historical strike prices between January 1996 and February 2010 (end of month prices) for each delta, in order to obtain the average interval between strike prices for this period. Because the quotes are given in volatility in terms of delta, at each considered date, we converted the deltas into strike prices using the formulas

$$\begin{aligned} X_{call} &= S_t e^{-N^{-1}(\Delta_{call} e^{r_{usd}T})\sigma\sqrt{T}+(r_{brl}-r_{usd}+\sigma^2/2)T} \\ X_{put} &= S_t e^{N^{-1}(-\Delta_{put} e^{r_{usd}T})\sigma\sqrt{T}+(r_{brl}-r_{usd}+\sigma^2/2)T} , \end{aligned} \tag{4.2}$$

where S_t is the USDBRL exchange rate (the price of one unit of the US dollar, which is the foreign currency, expressed in BRL real, the domestic currency), r_{brl} is the domestic risk-free interest rate (Brazilian interest rate) and r_{usd} the foreign interest rate (US interest rate) Espen (2007). As with strike prices, in the Heston model we also used the average and the volatility of the spot USDBRL FX rate for the period starting on June 2006 and finishing on February 2010 for S_0 (USDBRL price at $t = 0$) and v_0 (volatility of the USDBRL price at $t = 0$). The interest rates r_{brl} and r_{usd} are also an average from the money market rates (US Libor and SICOR for Brazil) for the same period and have a maturity of 1, 3 or 6 months, depending on the maturity of the "true" RND.

In total, we generate six scenarios for each maturity which results in eighteen different RNDs. The other parameters used for producing the different scenarios are the same as in Bu and Hadri (2007) and Cooper (1999). The authors chose the long-run volatility based on the levels of implied volatility typically observed within equity markets and for the low volatility scenarios chose the long run volatility typically observed in stock index, currency and interest rate markets.

Our goal using this method was to produce risk-neutral densities that incorporate the different shapes and scenarios discussed above (different levels of skewness and kurtosis) in order to test the capacity of the MLN, SML and DFCH methods to recover these RNDs. Doing this does not assume that equation (3.4) explains the asset price dynamics

in the real world.

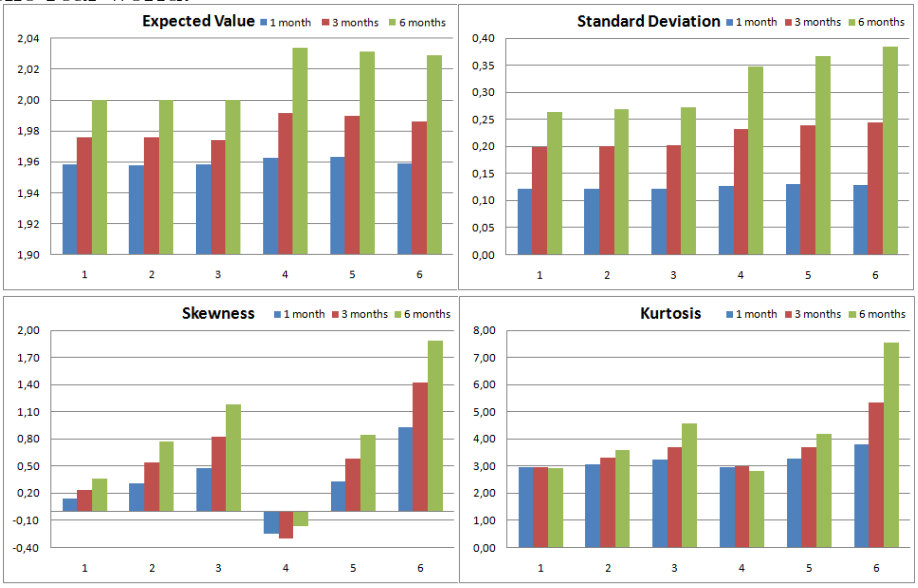


Figure 4-2: Summary statistics of the "true" RND obtained through the Heston model

The summary statistics for the eighteen RNDs obtained through the Heston model are presented in figure 4-2. The wide range of different shapes that the different RNDs can assume can be seen. For example, the skewness range between -0.1651 and 1.8839 and the kurtosis between 2.8316 (thin tails) and 7.5411 (fat tails) in the high volatility scenario for the 6-month horizon. We can also see that the variance increases with the maturity, as it should be expected.

To test the robustness of the MLN, SML and DFCH models in recovering the "true" RNDs, we first derive the call option prices using equation (3.6) in section 3.1.2 for each combination of scenario and maturity. We then add a uniformly distributed random noise perturbation in prices of size between minus half and half of the tick size (according to BM&FBOVESPA, the minimum tick size is 0.001) as in Bliss and Panigirtzoglou (2002). Given these shocked option prices, we use the MLN, SML and DFCH methods (the details of the optimization process are described in section 4.4) to estimate the RNDs. This process of first shocking prices and then fitting the RND is repeated 500 times for the eighteen combinations of maturities and scenarios (Monte Carlo Simulation).

In order to approximate the methodology described above to the characteristics of

the USDBRL option market, we proceed with the calibration of the Heston model for the end of month USDBRL option quotes between June 2006 and February 2010 (the results are presented in figure 11-9 in Appendix B) and we also produced the tests described above for 12 dates (6 low volatility dates and 6 high volatility dates). For the low volatility dates we select the period between October 2006 and March 2007 (before the increased problems regarding the subprime crisis). For the high volatility dates we select the period between September 2008 and February 2009 (peak of the subprime crisis). For these periods, we generate the "true" RNDs using the calibration parameters and the strike prices obtained for each tested date.

The different methods are then compared using some statistical measures that will be described below.

4.3 Statistics used in comparison of different techniques

In this thesis the different methods were compared using different approaches adopted by different authors.

In Cooper (1999) and Bliss and Panigirtzoglou (2002) the mean, standard deviation, skewness and kurtosis of the estimated RNDs were analyzed. However, Bliss and Panigirtzolou focused on stability analysis.

In Cooper the robustness of the MLN and SML methods was studied by comparing the mean of the summary statistics obtained from the Monte Carlo simulations with the summary statistics of the "true" RNDs. The process of shocking the prices [4] and then fitting the RNDs was repeated 100 times. The author also tested the stability of these models by analyzing the standard deviation of the summary statistics, arguing that the model with the best performance in terms of stability has a lower standard deviation for

[4] each price was shocked by a random number uniformly distributed from -1/2 to +1/2 a tick size

the different descriptive statistics. He concluded that the SML method performed better than the MLN method in terms of accuracy and stability.

Bliss and Panigirtzolou tested the stability of the MLN and SML methods, but instead of shocking the fitted prices obtained from the Heston model, they introduced a noise in market prices. The authors believed a good estimation method would have better behavior in the convergence results of the processed simulations. These authors did not adopt the methods followed by Cooper, arguing that goodness-of-fit results outside the range of available strike prices (tails of the distribution) can be unreliable, in the sense that there is an infinite variety of probability masses in the tails of the RNDs obtained. In fact, the summary statistics with higher moments like skewness and kurtosis are very sensitive to the tails of the distribution, and the data outside the set examined is heavily dependent on the estimation method used. For example, when the assumed PFD has a double-lognormal functional form, the MLN estimation method may do better than the other methods. We agree with these arguments and hence we give more importance to the RMISE analysis (root mean integrated analysis) as in Bu and Hadri (2007).

Bu and Hadri (2007) tested the accuracy and stability of the DFCH and SML methods using the root mean integrated squared error (RMISE), which has the advantage of being less sensitive to the tails of the distribution. Another advantage of RMISE is that it can be broken down into RISB (root integrate square bias) that measures the accuracy and RIV (root integrated variance) which indicates the stability of the distribution. As in Cooper (1999), Bu and Hadri also compared the performance of the methods in terms of a "true" PDF produced from an assumed Heston stochastic volatility price and using the pseudo-prices generated from the PDFs as input. For each combination of maturity and scenario, the authors carry out 500 simulations and find that in the majority of the cases the DFCH method performs better than the SML method in terms of accuracy (RISB) and stability (RIV).

In this thesis we tested both accuracy and stability of the DFCH, MLN and SML methods using the RMISE as in Bu and Hadri (2007), in Bondarenko (2003) and in

Lee (2008), and using the mean, standard deviation, skewness and kurtosis summary statistics as in Cooper (1999) and Bliss and Panigirtzoglou (2002).

A definition of these statistics is provided below:

i. *mean*: expected value of the implied PDF.

ii. *standard deviation:* square root of the variance of the implied PDF.

iii. *skewness:* the third central moment of the implied pdf standardized by the third power of the standard deviation.

$$Skewness = \frac{E[X - \bar{X}]^3}{\sigma^3} \tag{4.3}$$

It provides a measure of asymmetry, measuring the relative probabilities above and below the mean. By weighting the relative probabilities through the cubic distances, the weighting of the relative probabilities above the mean becomes positive and the weighting of the relative probabilities below the mean becomes negative.

iv. *kurtosis:* the fourth central moment of the implied pdf standardized by the fourth power of the standard deviation. Provides a measure of the degree of "fatness" of the tails of the implied pdf. The kurtosis of the normal distribution is equal to three. A higher kurtosis usually implies a greater probability for extreme changes. This means that a distribution with a higher kurtosis when compared with the normal distribution has fatter tails than the normal distribution (normally associated with a greater degree of "peakedness" in the centre of the PDF).

$$Kurtosis = \frac{E[X - \bar{X}]^4}{\sigma^4} \tag{4.4}$$

v. *RMISE*: the root mean integrated squared error. By considering $\hat{f}(S_t)$ as the estimator of the true RND, then the RMISE is defined as

$$RMISE(\hat{f}) = \sqrt{E[\int_{-\infty}^{\infty} (\hat{f}(S_t) - f(S_t))^2 dS_t]} \tag{4.5}$$

representing a measure of the average integral of the squared error over the support of

the RND. It is a measure of the quality of the estimator and is not as sensitive to the tails of the distribution as the skewness and kurtosis.

The squared of RMISE can also be broken down into the sum of the squared RISB (root integrated squared bias) and squared RIV (root integrated variance):

$$RMISE^2(\hat{f}) = RISB^2(\hat{f}) + RIV^2(\hat{f}) \tag{4.6}$$

$$RISB(\hat{f}) = \sqrt{\int_{-\infty}^{\infty} (E[\hat{f}(S_t)] - f(S_t))^2 dS_t]} \tag{4.7}$$

$$RIV(\hat{f}) = \sqrt{\int_{-\infty}^{\infty} E[(\hat{f}(S_t) - E[\hat{f}(S_t)])^2] dS_t} \tag{4.8}$$

In the thesis we tested all the statistics explained above. However, because of the limitations of skewness and kurtosis in evaluating PDFs, we give more importance to RMISE as a measure of the overall quality of the estimator, whereby RISB is the measure of the accuracy and RIV is the measure of the stability.

4.4 Numerical aspects of estimating option prices using MLN, SML and DFCH

The optimizations we have performed for the calculus of the theoretical option prices and estimation of the risk-neutral densities using Double-Lognormal Function (DLN), the Smoothed Implied Volatility Smile (SML) and the Density Functional Based on Confluent Hypergeometric Function (DFCH) were produced using the MATLAB software.

4.4.1 Double-Lognormal Function

As explained in section 3.2.1, the mixture of lognormals (MLN) was proposed in Bahra (1997) and Melick and Thomas (1997) and assumes a functional form for the risk-neutral

density (RND) that is consistent with various stochastic processes for the underlying asset (instead of specifying underlying asset price dynamics as in Black and Scholes' model, which leads to a unique terminal RND). Using the MLN method, the RND is a weighted sum of lognormal density functions because according to Bahra the asset price distributions are closer to the lognormal distribution. For our purposes, we follow Bahra and adopt a Mixture of two lognormals in the estimation of the risk-neutral densities from the pseudo-option prices calculated through the Heston model (as described in section 4.2). The five parameters $(\alpha_1, \alpha_2, \beta_1, \beta_2, w)$ are estimated through the minimization problem defined in equation (3.15). The part of the minimization problem that corresponds to the non-arbitrage condition, restricting the expected value of the RND to be equal to the forward price of the underlying asset, is defined in our algorithm as the price of the underlying asset minus the theoretical price of a call option (using MLN model) with a strike price of 0, which has the same meaning as equation (3.15) but in a different form. In fact, this martingale condition means that for a strike price of 0, the option will always be exercised and at maturity it will be worth the value of the underlying asset. Therefore, we must solve the minimization problem:

$$\min_{\alpha_1,\alpha_2,\beta_1,\beta_2,w} \sum_{i=1}^{n} [c(X,\tau) - \widehat{c}]^2 + [S - c(0,\tau)] \tag{4.9a}$$

Due to the symmetry problems discussed in section 3.2.1, we impose $\beta_1 > \beta_2$ (the first density will have a larger standard deviation than the second one). The optimization was carried out using MATLAB with a non-linear least squares optimization algorithm and we follow the optimization steps proposed in Jondeau et al. (2006). We start by defining a vector of values for the weight parameter w in the interval $[0, 1]$. The points in this vector are equally spaced at intervals of 0.1. We then proceed to the optimization along the grid of w values and obtain the values for $\alpha_1, \alpha_2, \beta_1, \beta_2, w$ that minimize our problem. These parameters are inserted into the MLN's RND equation (3.8) in order to obtain the risk-neutral density. This procedure was repeated 500 times for each combination of maturity and scenario as described in section 4.2.

4.4.2 Density Functional Based on Confluent Hypergeometric Function

This method, described in section 3.2.1, was developed in Abadir and Rockinger (2003) and consists of the use of a formula that encompasses various densities, like normal, gamma, inverse gamma, weibull, pareto and mixtures of these statistical densities.

As explained in section 3.2.1, the number of parameters to be estimated using this model was reduced to seven, due to the restrictions:

$$c_2 = -1 + a_4\sqrt{-b_4\pi}, \tag{4.10}$$

$$a_4 = \frac{1}{2\sqrt{-b_4\pi}}\left[1 - a_1(-b_2)^{-a_2}\frac{\Gamma(a_3)}{\Gamma(a_3 - a_2)}\right], \tag{4.11}$$

$$c_1 = -c_2 m_2, \tag{4.12}$$

$$b_1 = 1 + a_2 b_3, \tag{4.13}$$

$$a_5 = -\frac{1}{2}, \tag{4.14}$$

$$a_6 = \frac{1}{2}, \tag{4.15}$$

$$E(X) = a_1\frac{\Gamma(a_3)}{\Gamma(a_3 - a_2)}(-b_2)^{-a_2}(m_1 - m_2) + m_2. \tag{4.16}$$

The minimization defined in equation (3.23a) (regardless of the method used, the objective is to minimize some function of the squared distance between the observed option prices and the fitted prices derived from the estimated PDF) was performed in Matlab using non-linear least square optimization.

As described in section 4.2, the estimation of the risk-neutral densities used the pseudo-option prices calculated through the Heston model as input.

Given the high number of parameters to be estimated, the choice of initial points to be used in the optimization plays an important role. We opted for the values used in Abadir and Rockinger (2003) since these authors proved that they worked well.

These parameters coincide with the parameters of a Gaussian RND for the third term of equation (3.17):

$$a_5 = -\frac{1}{2}, \quad a_6 = \frac{1}{2}, \quad b_4 = -\frac{1}{2 * \sigma^2(K)}, \; m_2 = mean(K). \tag{4.17}$$

Moreover, for the second-term of equation (3.17) the starting parameters correspond to the parameters of a restricted gamma RND:

$$b_1 = 1 + a_2 b_3, \quad b_3 = 1, \quad a_3 = a_2 + 2, \; a_2 = 4, \quad m_1 = m_2 \tag{4.18}$$

Owing to the highly non-linear nature of the function, it was also important to state lower and upper bounds to the parameters of the function during the optimization processes in order to achieve better stability and fit for the results obtained.

4.4.3 Smoothed Implied Volatility Smile

In the estimation of the RNDs through the SML model we used the method proposed in Bliss and Panigirtzoglou (2002) which consists of an interpolation of volatility/delta space using a natural smoothing cubic spline, whereby the second-order derivatives in the extreme knot points were 0 (spline function is linear outside the range of available data). This method is explained in detail in section 3.2.1.

The variable regarding the weight parameter w in equation (3.32) is described by Bliss and Panigirtzoglou as a source of price error. It is known that in the context of the Black and Scholes formula, the only unobservable parameter is volatility (σ), which means that the uncertainty regarding the PDF lies in σ. The greek vega (v) measures the relationship between volatility σ and option price ($v = \frac{dC}{d\sigma}$) and reflects the uncertainty concerning the volatility. The value of v is approximately 0 for far deep-out-the-money options and reaches its maximum for at-the-money options [5]. The authors use this v weighting when

[5] The value of out-the-money and in-the-money options relies mainly in the intrinsic value. The part that depends of the time value, (which depends on σ) is smaller.

fitting the volatility smile because this weighting scheme places more weight on near-the-money options and less weight on away-from-the-money options. However, the authors admitted that it was difficult to choose a good weighting scheme that takes into account all the sources of price error. In this thesis we tested the SML model using both vega weighting ($w_i = v_i$) and equal weighting ($w_i = 1$) and observed that the performance is similar for both (with a slight improvement for the vega weighting).

The smoothed parameter in function (3.32) , λ, multiplies a measure of curvature in function (3.32) and allows the smoothness of the spline and its shape to be controlled. In this thesis we tested this method using the value that minimizes the RMISE (root mean integrated squared error) as the smoothed parameter. Nevertheless, because in the real world we don't know the "true" RND, we are unable to get the λ that minimizes RMISE. As such, we also performed the SML technique using a specific value for the smoothing parameter ($\lambda = 0.9$).

In conclusion, we tested this method using different schemes for the weighting parameter ($w_i = v_i$ and $w_i = 1$) and for the smoothness of the spline (λ that minimizes the RMISE and $\lambda = 0.9$). We observed that the performance is very similar for the different schemes (see figures 11-2 and 11-4 in Appendix B).

Chapter 5

Comparison of different methods using the Cooper scenarios

The different methods tested in this thesis, the Double-Lognormal Function (DLN), the Smoothed Implied Volatility Smile (SML) and the Density Functional Based on Confluent Hypergeometric Function (DFCH) were compared in terms of accuracy and stability. The performance of the three techniques was measured based on two different approaches: analysis using the summary statistics (mean, standard deviation, skewness and kurtosis) and analysis using the RMISE (root mean integrated squared error) as explained in section 4.3.

The approach that measured accuracy based on the summary statistics was analysed as in Cooper (1999). We obtained the mean, standard deviation, skewness and kurtosis for the 500 simulations performed for each combination of scenario and maturity, and then compared the mean of these summary statistics with the values obtained in the "true" RND estimated through the Heston stochastic volatility model. In this approach, the stability is measured as in Cooper (1999) and Bliss and Panigirtzoglou (2002), which takes into account the standard deviation of the higher moments of summary statistics (variance, skewness and kurtosis), in order to measure how much the estimates are likely to be affected by data imperfections or computational problems. In line with

this approach, the model with better accuracy would present mean values of summary statistics closer to the "true" RND and the model with more stability would have a lower standard deviation of summary statistics.

However, as explained in section 4.3, skewness and kurtosis are very sensitive to the tails of the distribution and the data outside the examined set is heavily dependent on the estimation method used. That is why we also follow the approach used in Bu and Hadri (2007), who tested the accuracy and stability of the DFCH and SML methods using the root mean integrated squared error (RMISE), which is a more reliable measure of the robustness of the RND estimators.

5.1 Analysis using mean, standard deviation, skewness and kurtosis

5.1.1 Accuracy

The accuracy using this approach was analyzed by comparing the average values of the mean, standard deviation, skewness and kurtosis estimated from the 500 Monte-Carlo simulations, which were applied to each combination of scenario and maturity (the scenarios are defined in table 4-1). The method with the best performance has an average value of the summary statistics that is close to the "true" ones (figures 5-1 and 5-2).

To understand the unbiasedness of the MLN, DFCH and SML models we have to focus on the difference between these mean statistics and the "true" ones, so we present figure 5-3, which calculates the difference between the true and the mean summary statistics as a percentage of the "true" summary statistics: (true-mean)/true.

	Scenarios	Expected Value	Volatility	Skewness	Kurtosis	RISB	RMISE
1 month	low volatility and negative skewness	SPLINE	SPLINE	SPLINE	HYPERGEOM	HYPERGEOM	HYPERGEOM
	low volatility	SPLINE	MLN	SPLINE	HYPERGEOM	SPLINE	MLN
	low volatility and positive skewness	SPLINE	SPLINE	SPLINE	SPLINE	HYPERGEOM	HYPERGEOM
	high volatility and negative skewness	MLN	HYPERGEOM	HYPERGEOM	HYPERGEOM	HYPERGEOM	HYPERGEOM
	high volatility	MLN	MLN	MLN	MLN	MLN	MLN
	high volatility and positive skewness	SPLINE	HYPERGEOM	SPLINE	MLN	HYPERGEOM	HYPERGEOM
3 months	low volatility and negative skewness	SPLINE	SPLINE	SPLINE	HYPERGEOM	HYPERGEOM	HYPERGEOM
	low volatility	SPLINE	MLN	SPLINE	SPLINE	SPLINE	SPLINE
	low volatility and positive skewness	SPLINE	SPLINE	SPLINE	HYPERGEOM	MLN	MLN
	high volatility and negative skewness	MLN	SPLINE	MLN	HYPERGEOM	HYPERGEOM	HYPERGEOM
	high volatility	MLN	MLN	MLN	MLN	MLN	MLN
	high volatility and positive skewness	MLN	SPLINE	SPLINE	MLN	HYPERGEOM	HYPERGEOM
6 months	low volatility and negative skewness	SPLINE	SPLINE	SPLINE	HYPERGEOM	HYPERGEOM	HYPERGEOM
	low volatility	SPLINE	SPLINE	SPLINE	SPLINE	SPLINE	SPLINE
	low volatility and positive skewness	SPLINE	SPLINE	SPLINE	HYPERGEOM	MLN	MLN
	high volatility and negative skewness	MLN	SPLINE	MLN	SPLINE	HYPERGEOM	HYPERGEOM
	high volatility	MLN	MLN	MLN	MLN	MLN	MLN
	high volatility and positive skewness	SPLINE	MLN	MLN	HYPERGEOM	MLN	MLN

Figure 5-1: Best method in terms of accuracy for each combination of scenario and maturity

Figure 5-2: Summary statistics obtained for Heston model (true density) and mean of summary statistics obtained for DFCH, MLN and SML methods. The results estimated for the SML method were processed with v weighting and with the smoothing parameter λ that minimizes RMISE.

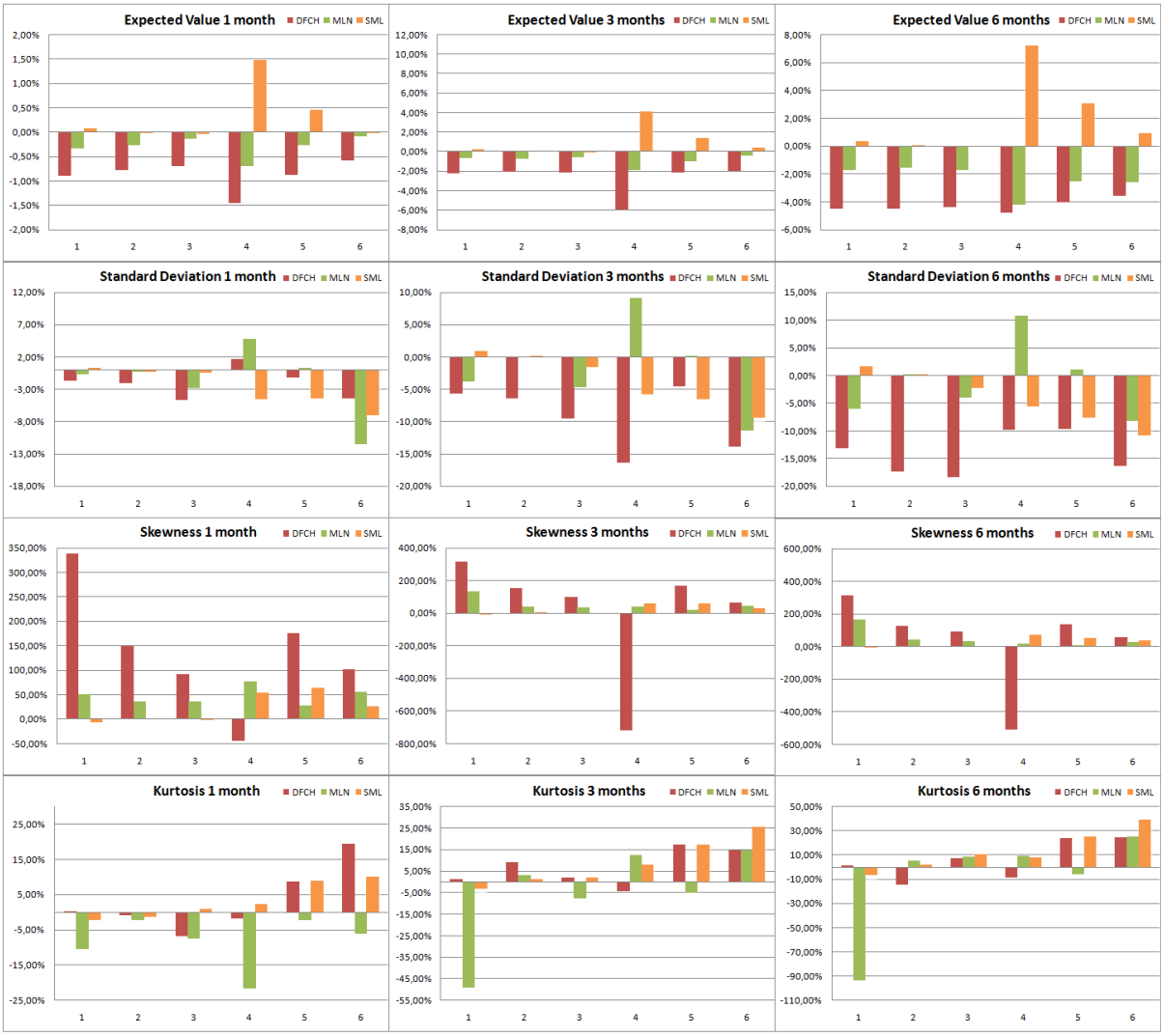


Figure 5-3: Difference between the "true" and the mean summary statistics in percentange of the "true" statistics.The results estimated for the SML method were processed with v weighting and with the smoothing parameter λ that minimizes RMISE.

Expected Value

If we look at the expected value, we see that the SML method has a better performance than the MLN method, with the exception of scenarios 4 and 5 (for all the maturities), where the MLN model is slightly closer to the "true" RND. The DFCH method has a biased expected value for almost all scenarios and maturities.

Standard Deviation

Analyzing the volatility, we see that for "one month to maturity" the SML outperforms the DFCH and the MLN methods in scenarios 1, 2 and 3. The DFCH method almost always has the worst performance, with the exception of scenarios 4 and 6, where it has the less biased implied volatility.

In the "three months to maturity" the SML technique has a better fit in scenarios 1, 3, 4 and 6. The DFCH has the least fitted implied volatility.

Considering the six-month term, we notice that the SML and MLN methods outperform the DFCH one, with the SML method showing better results for the low volatility scenarios and the MLN one the best in the high volatility scenarios.

In general terms, we see that the SML method has a better performance in capturing the volatility of the "true" density. The volatility of all tested RNDs increases in line with longer time to maturity, which confirms the higher uncertainty attached to longer maturities.

Skewness

Considering all the maturities, the SML and MLN methods have skewness values that have a close fit to the "true" skewness when compared to the DFCH model. We also observe a slightly better performance using the SML method (if we consider an equal weighting scheme, these results improve slightly). If we look carefully, we see that for 3 and 6-month terms the SML method usually has an unbiased skewness in lower volatility scenarios and the MLN has better results in higher volatility scenarios. All the models

tested were able to capture the different levels of skewness corresponding to each scenario, which demonstrates their ability to incorporate the changes in skewness observed in the real world.

Kurtosis

Analyzing the 1-month kurtosis, it can be noticed that the DFCH method has a close fit to the "true" RND in a bigger proportion of scenarios, having the best fit in scenarios 1, 2 and 4. The MLN underperforms in relation to the other methods in the estimation of the "true" kurtosis, except in scenarios 5 and 6.

For the “3 months to maturity”, the less biased estimator for kurtosis is obtained more times through the DFCH model (in scenarios 1, 3 and 4). The MLN and the SML methods have a similar performance, with the MLN having a closer fit to the "true" RND in scenarios 5 and 6 and the SML outperforming the other methods in scenario 2.

In the 6-month term the observed results were similar to the “3 months to maturity”, with the DFCH method outperforming the SML and MLN methods by a large proportion. The SML and the MLN methods have very similar behavior (MLN does better in higher volatility scenarios and the SML does better in lower volatility scenarios).

Analyzing all the maturities, it can be seen that the DFCH method has a close fit in the majority of the scenarios. We also observed that the MLN does better than the SML when the uncertainty is higher and worse than the SML in low volatility scenarios.

5.1.2 Stability

In this thesis, the stability is measured in line with the approach proposed in Cooper (1999) and Bliss and Panigirtzoglou (2002), which consists of slightly perturbing the option prices and then re-estimating the RNDs as explained in section 4.2, in order to measure how much estimates are likely to be affected by data imperfections or computational problems. The most stable method would have a lower standard deviation for the higher moments of summary statistics (variance, skewness and kurtosis).

The standard deviation values of the summary statistics are shown in figure 5-5.

	Scenarios	Volatility	Skewness	Kurtosis	RIV
1 month	low volatility and negative skewness	HYPERGEOM	SPLINE	SPLINE	SPLINE
	low volatility	MLN	SPLINE	SPLINE	MLN
	low volatility and positive skewness	SPLINE	SPLINE	SPLINE	MLN
	high volatility and negative skewness	SPLINE	SPLINE	SPLINE	SPLINE
	high volatility	HYPERGEOM	SPLINE	SPLINE	SPLINE
	high volatility and positive skewness	SPLINE	SPLINE	SPLINE	SPLINE
3 months	low volatility and negative skewness	SPLINE	SPLINE	SPLINE	SPLINE
	low volatility	MLN	SPLINE	SPLINE	MLN
	low volatility and positive skewness	SPLINE	SPLINE	SPLINE	SPLINE
	high volatility and negative skewness	SPLINE	SPLINE	SPLINE	SPLINE
	high volatility	MLN	SPLINE	SPLINE	SPLINE
	high volatility and positive skewness	SPLINE	SPLINE	SPLINE	SPLINE
6 months	low volatility and negative skewness	SPLINE	SPLINE	SPLINE	SPLINE
	low volatility	MLN	SPLINE	SPLINE	MLN
	low volatility and positive skewness	SPLINE	SPLINE	SPLINE	SPLINE
	high volatility and negative skewness	SPLINE	SPLINE	SPLINE	SPLINE
	high volatility	SPLINE	SPLINE	SPLINE	SPLINE
	high volatility and positive skewness	MLN	SPLINE	SPLINE	MLN

Figure 5-4: The most stable method for each combination of scenario and maturity

Figure 5-5: Standard Deviation of the summary statistics for the SML, MLN and DFCH methods

Variance

For all the maturities considered, the SML method returns the lowest standard deviation of the variance, which indicates that this method is the most stable one for the volatility estimates.

It should be mentioned that the stability of the SML method increases when the v weighting scheme is applied (see figure 11-6 in Appendix B).

Skewness

The SML method has the most stable skewness estimates across all the combinations of scenarios and maturities.

We observe a stability improvement for the SML method if we consider a v weighting scheme (see figure 11-6 in Appendix B).

Kurtosis

The SML method has the highest degree of stability for all the tested maturities. We also observed the same phenomenon as in the standard deviation of variance and skewness, with the performance of the SML method again improving upon the adoption of a v weighting scheme. The MLN method is the worst performer for 1-month and 3-month terms and the DFCH has the lowest stability for the 6-month term.

5.2 Analysis using RMISE

As explained in section 3.3, the summary statistics skewness and kurtosis are highly sensitive to the tails of the distributions, which can lead to unreliable results outside the range of available strike prices.

In view of these limitations, in the statistical analysis we test the accuracy and stability of the DFCH, MLN and SML methods using the RMISE which is a measure of the average of the integral of the squared deviation over the support of the distribution and is less sensitive to the tails of the distribution. The RMISE can be broken down into RISB (measure of accuracy) and RIV (measure of stability). The best model will have a lower RMISE (lower RISB if it is more accurate and lower RIV if it is more stable).

The values obtained for the eighteen combinations of scenarios and maturities are presented in figure 5-6.

Figure 5-6: Values for RMISE, RISB and RIV. The results shown for the SML method were processed with v weighting and the smoothing parameter λ that minimizes RMISE

5.2.1 SML with v weighting or with equal weighting

As in the analysis of the summary statistics, we examined the results considering the impact on the SML method of using both the v weighting scheme and the equal weighting scheme and the opti∇mal λ (minimizes RMISE) as the smoothing parameter and $\lambda = 0.9$. In terms of the overall quality of the estimator which is measured by RMISE, we observe a better performance of the SML method when it applies a v weighting approach. These results are in accordance with those observed in the previous section, where the skewness and kurtosis estimated by the SML model were closer to the "true" skewness and kurtosis when the v weighting was applied. The decrease of RMISE when using the v weighting is due mainly to the increase in stability, which is measured through RIV. The accuracy,

which is measured using RISB, is almost the same if we use equal weighting in the RND estimation. The impact of using an optimal λ was insignificant.

5.2.2 Best Performance of the DFCH and MLN as the estimators of the "true"RND

Examining the results for the different maturities, we observe that the RNDs estimated with the DFCH and MLN methods perform better than the distributions obtained with the SML in terms of the overall quality of the RND estimator. In fact, the lower RMISE of the DFCH and MLN methods is observed in the majority of the eighteen combinations of scenarios and maturities (the DFCH method has the best RND estimator 9 times and the MLN method 7 times). This higher quality of the DFCH and MLN estimators is due to the better accuracy of these methods, which translates into a lower RISB. The DFCH method has a higher flexibility and was superior in capturing the different shapes of the "true" distributions under the various scenarios, only showing fragilities in the estimation of the lower skewness scenarios and in terms of stability.

5.2.3 Comparing DFCH with MLN accuracy

The DFCH method performs less well in terms of accuracy in the central scenarios (lower skewness) for the different maturities. For the central scenarios, the SML method has a higher overall quality as an estimator (lower RMISE) and better accuracy in the lower volatility scenarios (except for one-month term) and the MLN has a lower RMISE and RISB in the higher volatility scenarios.

Comparing the accuracy of the DFCH and MLN models for the negative skewness scenarios, we observe for all the maturities that the DFCH has an higher quality and accuracy as an estimator of the "true" RND, with the accuracy of these methods almost the same for the 6-month term.

In positive skewness scenarios, the DFCH method does better in terms of accuracy in

"one month to maturity". However, for longer maturities the MLN has a lower RMISE and RISB in the majority of these positive expectations scenarios.

5.2.4 Stability

Comparing the stability of the DFCH, MLN and SML methods and considering a v weighting scheme for the SML model, because of the more stable performance using this technique, we conclude that the SML method outperforms the other models across all the maturities. In terms of stability, the DFCH method underperforms in relation to the MLN and SML methods in the majority of the cases. Nevertheless, the impact of its lower stability is insufficient to offset its superiority as the estimator of the "true" RND.

5.3 Comparison of our results with other studies

In Cooper (1999), the MLN model was compared with the SML method in terms of accuracy and stability using the summary statistics approach and in Bu and Hadri (2007) the DFCH method was compared with the SML method in line with RMISE criteria. In both studies, the accuracy was measured using the Cooper technique of generating the "true" world through the Heston model and the SML was estimated interpolating across the volatility smile in "delta-space" via a cubic smoothing spline (as in our study). In Cooper, the SML method had a better stability performance and in terms of accuracy neither technique outperformed the other in skewness and kurtosis estimates. In Bu and Hadri (2007) the DFCH had a higher accuracy (lower RISB) and the SML method was more stable in the majority of scenarios (lower RIV).

As in Cooper, it was difficult to define which method (MLN or SML) was better in capturing the "true" skewness and the "true" kurtosis. Nevertheless, we notice that the SML method marginally outperformed the MLN method in skewness and kurtosis accuracy and was the best model at capturing the "true" expected value and the "true" volatility. In terms of stability, we obtained the same results as in Cooper. In fact, the

summary statistics estimates calculated through the SML method were the most stable.

According to the RMISE criterion, the DFCH was the most accurate model in the majority of scenarios (lower RISB) and the SML model was the most stable (lower RIV), as in Bu and Hadri's study.

Chapter 6

Comparison of different methods using USDBRL Heston calibrated parameters

In order to approximate the method proposed in Cooper (1999) to the characteristics of the USDBRL option market, we calibrated the Heston model for the end of month USDBRL option quotes between June 2006 and February 2010 (the results are presented in figure 11-9 in Appendix B) and produced the Monte Carlo simulations in order to re-estimate the RNDs using the calibration parameters and the strike prices for 12 dates (6 low volatility dates and 6 high volatility dates). We selected the period between October 2006 and March 2007 (before the increase of the problems regarding the subprime crisis) as the low volatility dates. The period between September 2008 and February 2009 (peak of the 2007-2010 financial crisis) was selected as the high volatility dates.

6.1 Analysis using mean, standard deviation, skewness and kurtosis

6.1.1 Accuracy

	Scenarios	Expected Value	Volatility	Skewness	Kurtosis	RISB	RMISE
1 month	Outubro 06	MLN	HYPERGEOM	SPLINE	MLN	HYPERGEOM	HYPERGEOM
	Novembro 06	MLN	HYPERGEOM	SPLINE	MLN	HYPERGEOM	HYPERGEOM
	Dezembro 06	HYPERGEOM	MLN	SPLINE	HYPERGEOM	HYPERGEOM	HYPERGEOM
	Janeiro 07	MLN	MLN	MLN	MLN	HYPERGEOM	HYPERGEOM
	Fevereiro 07	MLN	SPLINE	SPLINE	HYPERGEOM	HYPERGEOM	HYPERGEOM
	Março 07	MLN	HYPERGEOM	SPLINE	MLN	HYPERGEOM	HYPERGEOM
3 months	Outubro 06	SPLINE	MLN	SPLINE	MLN	MLN	MLN
	Novembro 06	MLN	MLN	SPLINE	MLN	HYPERGEOM	HYPERGEOM
	Dezembro 06	SPLINE	HYPERGEOM	SPLINE	MLN	HYPERGEOM	HYPERGEOM
	Janeiro 07	MLN	MLN	MLN	MLN	HYPERGEOM	HYPERGEOM
	Fevereiro 07	SPLINE	MLN	SPLINE	MLN	HYPERGEOM	HYPERGEOM
	Março 07	SPLINE	MLN	SPLINE	MLN	MLN	MLN
6 months	Outubro 06	SPLINE	MLN	SPLINE	MLN	MLN	MLN
	Novembro 06	SPLINE	MLN	SPLINE	MLN	MLN	MLN
	Dezembro 06	SPLINE	MLN	SPLINE	MLN	HYPERGEOM	HYPERGEOM
	Janeiro 07	SPLINE	MLN	SPLINE	MLN	MLN	MLN
	Fevereiro 07	SPLINE	MLN	SPLINE	MLN	MLN	MLN
	Março 07	SPLINE	MLN	SPLINE	MLN	MLN	MLN

Figure 6-1: Best method in terms of accuracy for the low volatility dates

	Scenarios	Expected Value	Volatility	Skewness	Kurtosis	RISB	RMISE
1 month	Setembro 08	MLN	HYPERGEOM	SPLINE	MLN	HYPERGEOM	HYPERGEOM
	Outubro 08	HYPERGEOM	HYPERGEOM	SPLINE	MLN	HYPERGEOM	HYPERGEOM
	Novembro 08	HYPERGEOM	MLN	SPLINE	HYPERGEOM	HYPERGEOM	HYPERGEOM
	Dezembro 08	HYPERGEOM	HYPERGEOM	SPLINE	HYPERGEOM	HYPERGEOM	HYPERGEOM
	Janeiro 09	MLN	HYPERGEOM	SPLINE	MLN	HYPERGEOM	HYPERGEOM
	Fevereiro 09	MLN	HYPERGEOM	SPLINE	MLN	HYPERGEOM	HYPERGEOM
3 months	Setembro 08	SPLINE	MLN	SPLINE	MLN	MLN	MLN
	Outubro 08	HYPERGEOM	HYPERGEOM	SPLINE	HYPERGEOM	HYPERGEOM	HYPERGEOM
	Novembro 08	SPLINE	HYPERGEOM	SPLINE	MLN	HYPERGEOM	HYPERGEOM
	Dezembro 08	SPLINE	HYPERGEOM	SPLINE	MLN	HYPERGEOM	HYPERGEOM
	Janeiro 09	MLN	HYPERGEOM	SPLINE	MLN	HYPERGEOM	HYPERGEOM
	Fevereiro 09	MLN	HYPERGEOM	SPLINE	MLN	HYPERGEOM	HYPERGEOM
6 months	Setembro 08	SPLINE	MLN	SPLINE	MLN	MLN	MLN
	Outubro 08	MLN	HYPERGEOM	MLN	HYPERGEOM	MLN	MLN
	Novembro 08	MLN	HYPERGEOM	SPLINE	MLN	HYPERGEOM	HYPERGEOM
	Dezembro 08	MLN	HYPERGEOM	SPLINE	MLN	MLN	MLN
	Janeiro 09	SPLINE	HYPERGEOM	SPLINE	MLN	MLN	MLN
	Fevereiro 09	SPLINE	HYPERGEOM	SPLINE	MLN	HYPERGEOM	HYPERGEOM

Figure 6-2: Best method in terms of accuracy for the high volatility dates

Figure 6-3: Low Volatility Dates: Difference between the "true" and the mean summary statistics in percentange of the "true" statistics: (true-mean)/true. The results for the SML method were processed with v weighting and with the smoothing parameter λ that minimizes RMISE.

Figure 6-4: High Volatility Dates: Difference between the "true" and the mean summary statistics in percentange of the "true" statistics: (true-mean)/true. The results for the SML method were processed with v weighting and with the smoothing parameter λ that minimizes RMISE.

Expected Value

Low volatility Dates The mean of the expected values estimated using the SML method have a close fit to the "true" expected value in the 3 and 6-month terms. The DFCH method has the biased expected value for the majority of dates and maturities (see figure 6-3).

High volatility Dates The MLN method outperforms the other models for most dates and maturities. The DFCH model underperforms the other models for the 3 and 6-month terms. The SML method has the best fit for the 3-month term and the worst fit for the 1-month term (see figure 6-4).

Standard Deviation

Low volatility Dates The implied volatility estimated using the MLN method has the closest fit to the "true" standard deviation for the majority of dates and maturities. The SML method performs worse in terms of capturing the "true" volatility (see figure 6-3).

High volatility Dates For the high volatility dates, the best volatility fit was estimated using the DFCH method. The SML method has the worst performance for implied volatility in the 6-month term and the MLN has the worst performance in the 1 and 3-month terms (see figure 6-4).

Skewness

Low volatility Dates The SML method outperformed the other models in capturing the "true" skewness and the DFCH model has the worst performance for the majority of dates and maturities (see figure 6-3).

High volatility Dates For the high volatility dates, the best fit for skewness was estimated using the SML method. The MLN method has the worst performance for

implied volatility for the 1-month term and the DFCH method has the worst performance in the 3 and 6-month terms (see figure 6-4).

Kurtosis

Low volatility dates The implied kurtosis estimated using the MLN method was closer to the "true" kurtosis for most dates and maturities. The SML method has the highest biased implied volatility in the 1-month term and the DFCH has the worst performance in the 3 and 6-month terms (see figure 6-3).

High volatility dates For the high volatility dates, the MLN again outperformed the other models. The SML method has the worst performance at capturing the "true" kurtosis (see figure 6-4).

6.1.2 Stability

	Scenarios	Volatility	Skewness	Kurtosis	RIV
1 month	Outubro 06	SPLINE	SPLINE	SPLINE	MLN
	Novembro 06	SPLINE	SPLINE	SPLINE	SPLINE
	Dezembro 06	SPLINE	SPLINE	SPLINE	SPLINE
	Janeiro 07	HYPERGEOM	SPLINE	SPLINE	SPLINE
	Fevereiro 07	SPLINE	SPLINE	SPLINE	SPLINE
	Março 07	MLN	SPLINE	SPLINE	MLN
3 months	Outubro 06	HYPERGEOM	SPLINE	SPLINE	HYPERGEOM
	Novembro 06	SPLINE	SPLINE	SPLINE	SPLINE
	Dezembro 06	SPLINE	SPLINE	HYPERGEOM	HYPERGEOM
	Janeiro 07	SPLINE	SPLINE	SPLINE	HYPERGEOM
	Fevereiro 07	HYPERGEOM	SPLINE	SPLINE	HYPERGEOM
	Março 07	SPLINE	SPLINE	SPLINE	SPLINE
6 months	Outubro 06	HYPERGEOM	SPLINE	SPLINE	HYPERGEOM
	Novembro 06	SPLINE	SPLINE	SPLINE	MLN
	Dezembro 06	HYPERGEOM	SPLINE	SPLINE	HYPERGEOM
	Janeiro 07	HYPERGEOM	SPLINE	HYPERGEOM	HYPERGEOM
	Fevereiro 07	HYPERGEOM	SPLINE	SPLINE	HYPERGEOM
	Março 07	HYPERGEOM	SPLINE	SPLINE	MLN

Figure 6-5: The most stable method for the low volatility dates

	Scenarios	Volatility	Skewness	Kurtosis	RIV
1 month	Setembro 08	HYPERGEOM	SPLINE	SPLINE	SPLINE
	Outubro 08	HYPERGEOM	SPLINE	SPLINE	MLN
	Novembro 08	HYPERGEOM	SPLINE	MLN	MLN
	Dezembro 08	MLN	SPLINE	SPLINE	MLN
	Janeiro 09	SPLINE	SPLINE	SPLINE	SPLINE
	Fevereiro 09	SPLINE	SPLINE	SPLINE	MLN
3 months	Setembro 08	SPLINE	SPLINE	SPLINE	SPLINE
	Outubro 08	HYPERGEOM	MLN	MLN	HYPERGEOM
	Novembro 08	HYPERGEOM	SPLINE	SPLINE	MLN
	Dezembro 08	HYPERGEOM	MLN	SPLINE	HYPERGEOM
	Janeiro 09	HYPERGEOM	SPLINE	SPLINE	SPLINE
	Fevereiro 09	HYPERGEOM	SPLINE	SPLINE	MLN
6 months	Setembro 08	SPLINE	SPLINE	SPLINE	HYPERGEOM
	Outubro 08	HYPERGEOM	MLN	MLN	MLN
	Novembro 08	SPLINE	SPLINE	SPLINE	SPLINE
	Dezembro 08	MLN	MLN	SPLINE	MLN
	Janeiro 09	HYPERGEOM	SPLINE	SPLINE	MLN
	Fevereiro 09	HYPERGEOM	SPLINE	SPLINE	HYPERGEOM

Figure 6-6: The most stable method for the high volatility dates

Figure 6-7: Low Volatility Dates: Standard Deviation of the summary statistics for the SML, MLN and DFCH methods

Figure 6-8: High Volatility Dates: Standard Deviation of the summary statistics for the SML, MLN and DFCH methods

Standard Deviation

For the low volatility dates, the SML method was the most stable model and for the high volatility dates the DFCH method outperformed the other models. The MLN method was the most unstable model for the majority of the dates tested (see figures 6-7 and 6-8).

Skewness

The SML model was the most stable one for all dates. The MLN method was the most unstable model for the majority of the low volatility dates and the DFCH was the least stable model for most of the high volatility dates.

Kurtosis

As in the analysis using Cooper's Scenarios (section 5.1.2), the SML method was the most stable model for a bigger proportion of the low volatility and high volatility dates. The MLN method performs the worst in terms of stability in the low volatility dates and the DFCH was the least stable model for the high volatility dates.

6.2 Analysis using RMISE

Figure 6-9: Low Volatility Dates: Values for RMISE, RISB and RIV. The SML results were processed with v weighting and the smoothing parameter λ that minimizes RMISE

Figure 6-10: High Volatility Dates: Values for RMISE, RISB and RIV. The SML results were processed with v weighting and the smoothing parameter λ that minimizes RMISE

6.2.1 Best Performance of the DFCH and MLN model

The DFCH model was the best estimator of the "true" RND for all the dates tested with a maturity of 1 month and for almost all the dates with a maturity of 3 months. The MLN method was the best estimator of the "true" 6-month RND. Overall, the DFCH method returns the best performance at capturing the true RND (the DFCH method has a lower RMISE 24 times and the MLN method 12 times). The SML method performed worse than all the other methods in terms of accuracy (see figures 6-9 and 6-10).

6.2.2 Stability

In the stability analysis, we obtain different results than in the analysis of the Cooper scenarios (section 5.2.4), where the SML method outperforms all the other models. For the lower volatility dates, the DFCH method has the lower RIV for the majority of 3 and 6-month RNDs. The SML method has a lower RIV for most 1-month RNDs. For the high volatility dates there is no clear "winner" in terms of stability performance.

Chapter 7

Information contained in the option implied risk-neutral probability density function

Besides analyzing the accuracy and stability of the MLN, SML and DFCH methods, we also estimated the end of month RNDs extracted from the USDBRL option prices for the period between June 2006 and February 2010 in order to compare the measures obtained for the three models tested and to interpret the information provided by these implied distributions.

7.1 Analyzing changes of implied pdf summary statistics over time

7.1.1 Comparing MLN, SML and DFCH

Before analyzing the information provided by the statistical measures, we compare the summary statistics calculated for the MLN, SML and DFCH methods and see if the results (regarding the mean, the uncertainty, the skewness and the probability of extreme

moves) are similar for the methods considered, or if they show a similar trend.

The expected values of the estimated distributions evolve very closely to one another for the three models tested (figure 7-1), which shows that the reliability of the average value of all possible outcomes has low dependence on the method used to estimate the pdf. The advantage of using risk-neutral densities is that they provide information about a range of possible events in the future and for the estimation of the expected value there is no need to estimate implied distributions, because the prices of forward or future contracts already give us the expected value for the underlying asset.

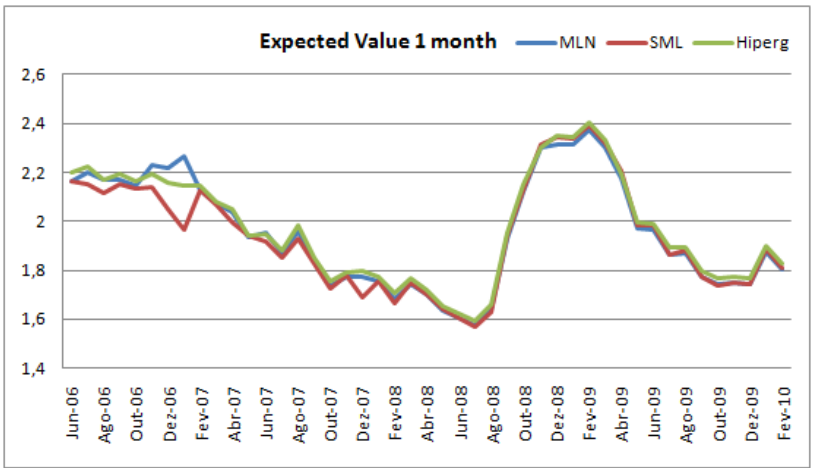


Figure 7-1: Evolution of one month to maturity expected value

The uncertainty around the mean, measured through the standard deviation of the estimated RNDs, has a strong correlation between the SML/DFCH pair. The MLN/DFCH and MLN/SML pairs have lower correlations for the standard deviation estimates.

1 month		
ρ	SML	DFCH
MLN	0,790	0,638
SML		0,944

3 months		
ρ	SML	DFCH
MLN	0,336	0,319
SML		0,981

6 month		
ρ	SML	DFCH
MLN	0,386	0,445
SML		0,987

Table 7-1: Correlations between the standard deviations calculated through MLN, DFCH and MLN for the period between June 2006 and February 2010

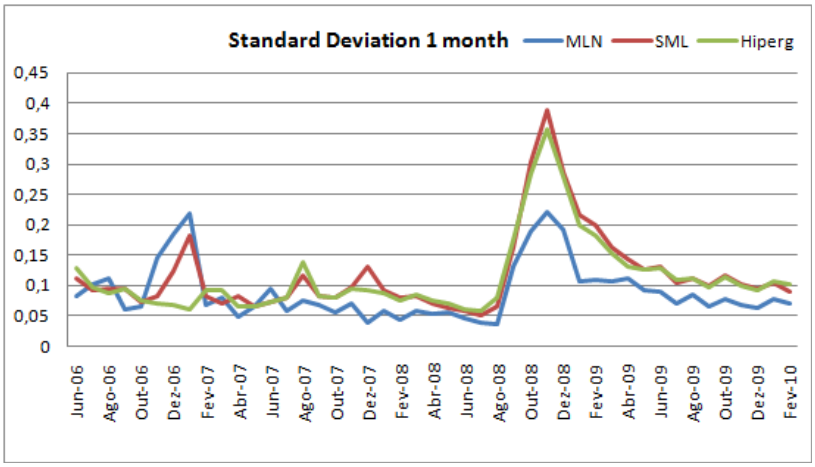


Figure 7-2: Evolution of one month to maturity standard deviation

Skewness is an indicator of the probability mass around the mean. If the implied distribution is positively skewed, the right tail is greater than the left tail and it suggests that market participants are positive about the future prices. However, a positively skewed distribution has an unweighted probability above the mean smaller than that below the mean (expected value is above the median and the mode), because the positive expectations lead to an upward revision of the expected price. Looking at figure 7-3, it is clear that for the period under consideration it is easier to find a trend for the implied skewness calculated for the DFCH and SML methods than for the MLN method (maintained a level close to 0.2 after December 2007 for "one month to maturity" term). The correlation level between the MLN method and the other methods is almost null and the correlation between the SML and DFCH methods is much lower when compared to the estimated values for the expected value and standard deviation (between 0.4 and 0.54).

1 month		
ρ	SML	DFCH
MLN	0,135	-0,044
SML		0,409

3 months		
ρ	SML	DFCH
MLN	0,159	-0,075
SML		0,537

6 month		
ρ	SML	DFCH
MLN	0,046	0,140
SML		0,436

Table 7-2: Correlations between the skewness calculated through MLN, DFCH and MLN for the period between June 2006 and February 2010

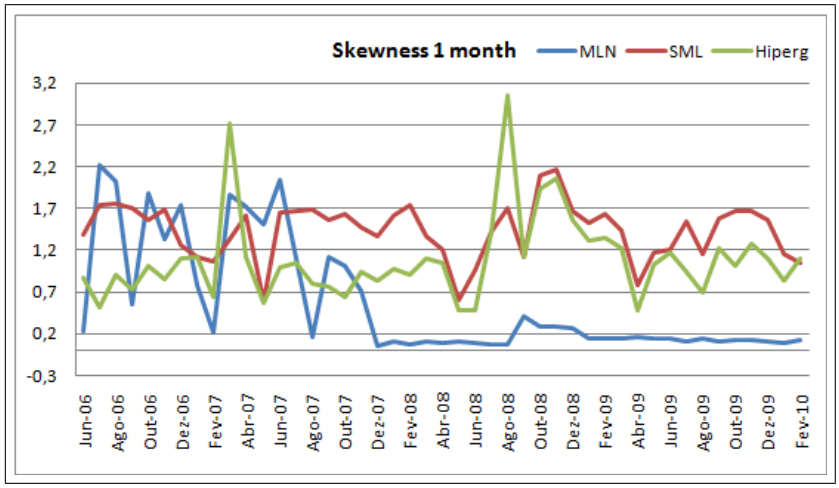


Figure 7-3: Evolution of one month to maturity skewness

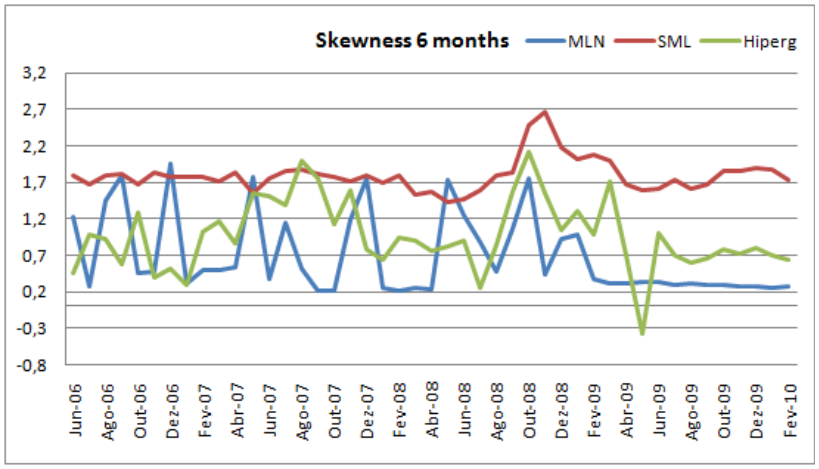


Figure 7-4: Evolution of six months to maturity skewness

As mentioned earlier in this thesis, the skewness is very sensitive to the tails of the distribution, which decreases the reliability of this measure. We therefore calculated the values for Pearson's skewness coefficients which are less sensitive to the tails of the distribution.

$$\text{Pearson median skewness} = \frac{E[X] - \text{median}}{\sigma} \tag{7.1}$$

$$\text{Pearson mode skewness} = \frac{E[X] - \text{mode}}{\sigma} \tag{7.2}$$

For both Pearson measures we saw almost no positive correlation between the observed values, which is shown in tables 6-3 and 6-4.

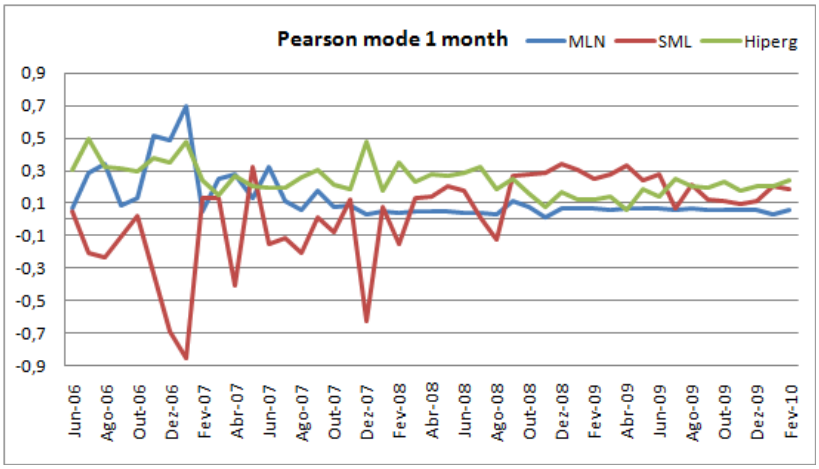


Figure 7-5: Evolution of one month to maturity Pearson mode

1 month		
ρ	SML	DFCH
MLN	-0,696	0,411
SML		-0,468

3 months		
ρ	SML	DFCH
MLN	-0,721	0,049
SML		-0,362

6 month		
ρ	SML	DFCH
MLN	-0,446	-0,105
SML		-0,325

Table 7-3: Correlations between the Pearson median skewness calculated through MLN, DFCH and MLN for the period between June 2006 and February 2010

1 month		
ρ	SML	DFCH
MLN	-0,707	0,507
SML		-0,758

3 months		
ρ	SML	DFCH
MLN	-0,641	0,295
SML		-0,726

6 month		
ρ	SML	DFCH
MLN	-0,397	0,344
SML		-0,526

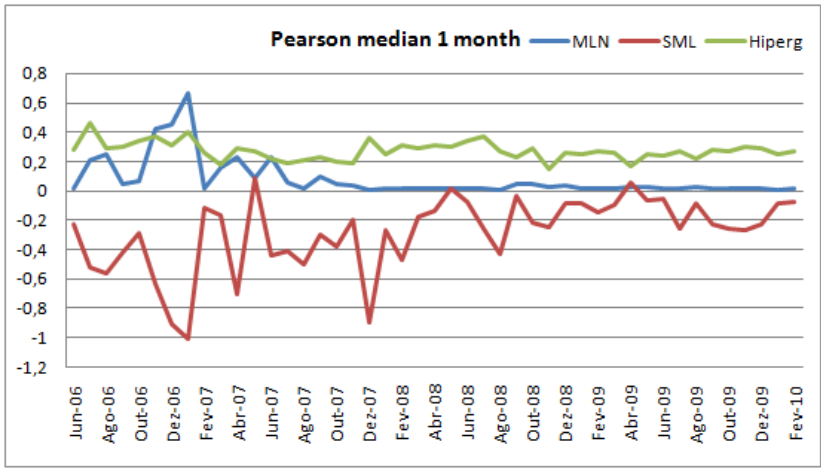


Figure 7-6: Evolution of one month to maturity Pearson median

Table 7-4: Correlations between the Pearson mode skewness calculated through MLN, DFCH and MLN for the period between June 2006 and February 2010

We applied the kurtosis as a measure of the probability for extreme changes (it also indicates how peaked a distribution is). However, as written earlier in this thesis, this measure is highly sensitive to the tails of the distribution, whose shape can have infinite forms and is heavily dependent on the method used to estimate the implied RNDs. As such, the reliability of the kurtosis measure is poor and should be interpreted with care. Like in the skewness analysis, the MLN method shows almost no changes after December 2007. In fact, the estimated MLN implied kurtosis for the one-month term was less able to capture the increase in kurtosis during the peak of the subprime crisis (August and September 2008). Once more, the correlation between the different methods was low. This correlation was higher between the DFCH and SML methods for the one month and three-month terms.

1 month		
ρ	SML	DFCH
MLN	-0,080	-0,062
SML		0,370

3 months		
ρ	SML	DFCH
MLN	-0,183	0,041
SML		0,469

6 month		
ρ	SML	DFCH
MLN	-0,010	0,232
SML		-0,023

Table 7-5: Correlations between the kurtosis calculated through MLN, DFCH and MLN for the period between June 2006 and February 2010

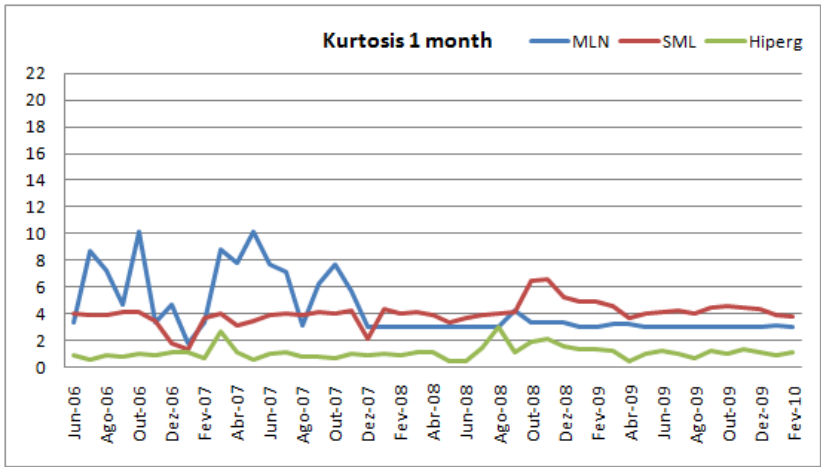


Figure 7-7: Evolution of one month to maturity Kurtosis

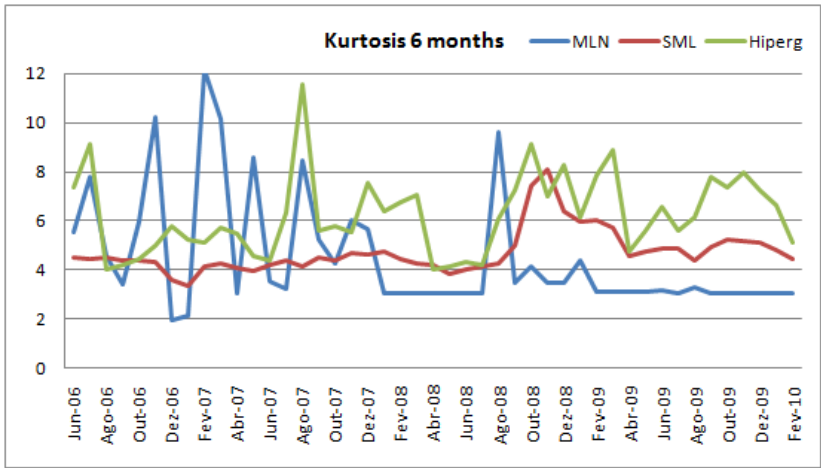


Figure 7-8: Evolution of 6 months to maturity Kurtosis

In conclusion, a higher correlation between the DFCH and SML methods was observed for the expected values and standard deviations of the implied RNDs. However, the statistical measures which correspond to the asymmetry and probability of extreme movements show different results depending on the method used. The correlation was higher between the SML and DFCH models. The historical values between June 2006 and February 2010 show the low reliability of the skewness and kurtosis measures that arises from the higher uncertainty of the estimated tails of the RNDs which are heavily dependent on the estimation method chosen. This increases the need to use a RND estimation method that is better able to capture the market expectations from the real world because the estimated statistical measures can be different depending on the model used. The RNDs estimated through the tested methods are shown in figure 7-9.

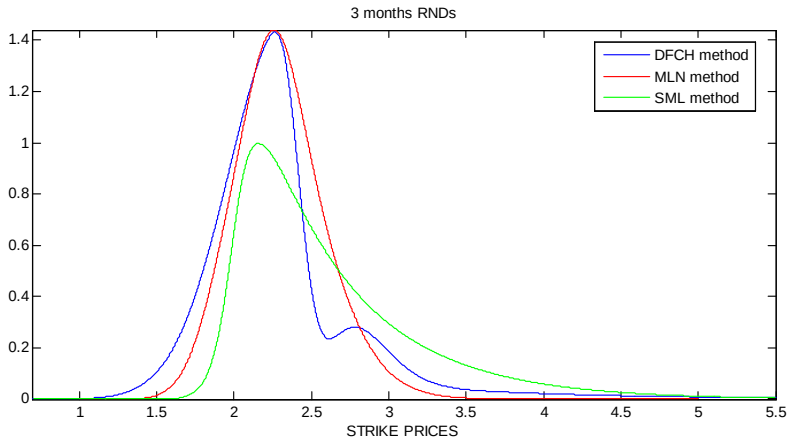


Figure 7-9: 3 months RNDs at 28th November 2008 estimated through DFCH, MLN and SML methods using USDBRL FX options

In Chapters 5 and 6 we concluded that the DFCH method was better at capturing the real world expectations when compared to the SML and MLN methods according to RMISE criterion. Therefore, we will analyze the information provided by the end of the month implied RNDs, between June 2006 and February 2010, using the implied distributions calculated through the DFCH method.

7.1.2 Historical behavior of implied summary statistics

The second half of 2006 and first half of 2007 was a period of BRL appreciation and a slight decrease in volatility (from 0.12 to 0.08 between June 2006 and July 2007 for "one month to maturity" RND). This decrease in volatility can be seen in figure 7-11 and in figures 7-15, 7-16 and 7-17 with the tightening of the gap between the 25th and 75th percentiles (Interquartile Range). During these two semesters, the expected values for the cross USDBRL were always above the spot USDBRL (see figure 7-10), which indicates the weakening expectations for the BRL. During this period the levels of skewness were positive (BRL depreciation) which goes in line with the higher expected values for USDBRL. This positive skewness can also be perceived if we compare the mean value, the mode and the median. Usually, if the mean value is above the mode and the median, there is a positive skewness (positive expectations lead to an upward revision of the expected price). However, by the end of the first semester of 2007 the distance between the mean and the mode narrowed a little, which could be a sign of downward revisions regarding an increase in USDBRL. During this period, there were consecutive SEDIC (overnight reference rate of the Brazilian inter-bank money market) rate cuts by COPOM (Brazil Monetary Policy Committee) which could be partially motivated by the need to force the BRL to depreciate (figure 11-13 in Appendix B). Nevertheless, these expectations of a depreciation in BRL did not materialize due to the healthier Brazilian macroeconomic conditions when compared with US data and the increasing credit and mortgage issues (the subprime crisis).

The trend described in the previous paragraph was temporarily interrupted between July and September 2007, with a peak in volatility, skewness and an increase in USDBRL. This increase in volatility and skewness along with a peak in Kurtosis (in figure 7-18 we see an increase in kurtosis in August 2007 for the 3 and 6-month terms) indicates that there was an increase of the probability of an extreme devaluation of BRL. This BRL weakening was related to heightened fears that the subprime and credit crisis in US would potentially reduce the global risk appetite for the emerging markets. It was

the first shock concerning the subprime crisis, with shortages and lack of liquidity in the money market. During these months, there were also rumors about some financial institutions experiencing liquidity difficulties, such as Northern Rock, a British Bank (at that time the biggest British mortgage lender) that was asking the Bank of England for emergency funding due to liquidity problems (in February 2008 Northern Rock Bank was nationalized).

Between September 2007 and July 2008 we continued to see a BRL appreciation, but this time, this movement was also supported by a COPOM tightening policy (in April 2008 it started a series of four consecutive rate rises) which increased the rate's differential between the Brazilian and US interest rates (at June 2006 the FED started a cycle of Federal Funds rate lowering) and augmented the pressure on the BRL strength (figures 11-13 and 11-12 in Appendix B). During this period the volatility was relatively constant with the implied standard deviation ranging between 0.21 and 0.26 for a maturity of six months (a higher range than in the first semester of 2007) and the central expectations measured by the interquartile range were concentrated in a lower range, which indicates a downward revision of the expectations concerning the USDBRL expected value (BRL appreciation). Despite increasing expectations for BRL appreciation, we noticed an increase in the asymmetry of the expectations in favor of a BRL depreciation, that could be related to the fact that the market attributed increasing likelihood for a correction of the BRL strengthening movement (the higher skewness can be seen through the increase of the Pearson mode and Person median, and through the increase of the difference between the mean and mode for the 3 and 6-month maturities, with the mean higher than the 75th percentile between March and August 2008 in figure 7-17).

In August 2008 the BRL appreciation came to an end after reaching a minimum of 1.558 (USDBRL). In August 2008 the markets pointed to two main reasons for the end of the Dollar depreciation: the end of the rises in oil prices (historically there is a negative correlation between oil prices and the dollar) and commodity prices (as a commodity exporter Brazil's trade surplus would be negatively affected), and the improvement in

the US Balance of Payments. There was also a huge increase in uncertainty, which could be seen in figure 7-11 and in the widening gap between the 25th and 75th percentiles (figures 7-15, 7-16 and 7-17). The growth in volatility could be due to the doubts in the financial markets about the extent of this dollar rally. There was also an increase in skewness and kurtosis caused by an increase in the probability of an extreme dollar appreciation. This upward movement in volatility (along with a rise in the expected value) reached its maximum in November 2008 after a sequence of negative events (in September 2008 Government-sponsored enterprises Fannie Mae and Freddie Mac which owned or guaranteed about half (56.8%) of the U.S mortgage market were being placed into conservatorship of the FHFA[1], Lehman Brothers filed for bankruptcy and the Bank of America purchased Merril Lynch, in October 2008 the US government bailed out Goldman Sachs and Morgan Stanley) that increased the risk aversion and the fears that the capital inflows for the emerging economies such as Brazil would be reduced, which would depreciate its exchange rate.

After December 2008 the USD stopped its rally and the volatility started to decrease, despite fears regarding the decrease of capital inflows into emerging markets. This new trend was partially caused by the increase in the US quantitative easing [2] and by the decrease of the Fed Reserve Target Rate to 0.25% in December 2008.

The decrease in volatility and BRL appreciation were more pronounced until May 2009, which can be seen through the decreasing of the USDBRL expected value and by the narrowing of the Interquartile Range. The skewness also dropped from the maximum values reached between August and December 2008 which could be provoked by the pressures as regards the dollar devaluation (increase in money supply due to quantitative

[1] The Federal Housing Finance Agency is an independent federal agency created on July 30, 2008, when the President George Bush signed into law the Housing and Economic Recovery Act of 2008. The Act objective was to create a world-class, empowered regulator with all of the authorities necessary to oversee vital components of US's secondary mortgage markets – Fannie Mae, Freddie Mac, and the Federal Home Loan Banks.

[2] Quantitative easing was used by the FED to increase the supply of money by increasing the excess reserves of the banking system, through buying not only government bonds, but also troubled assets in order to improve the liquidity of these assets.

easing).

After June 2009, the uncertainty came back to a range closer to the volatility levels prior to the turbulent period that started in August 2008 (nevertheless, until February 2010 it remained at higher levels than before the peak of the crisis), which could be related to the perception in the financial markets that the worst of the global recession was over.

In January and February 2010, we observed an increase in USDBRL (BRL depreciation) that was accompanied by an increase in the level of skewness (implied skewness with "one month to maturity" increased as well as the Pearson mode and Pearson median for all the considered maturities).

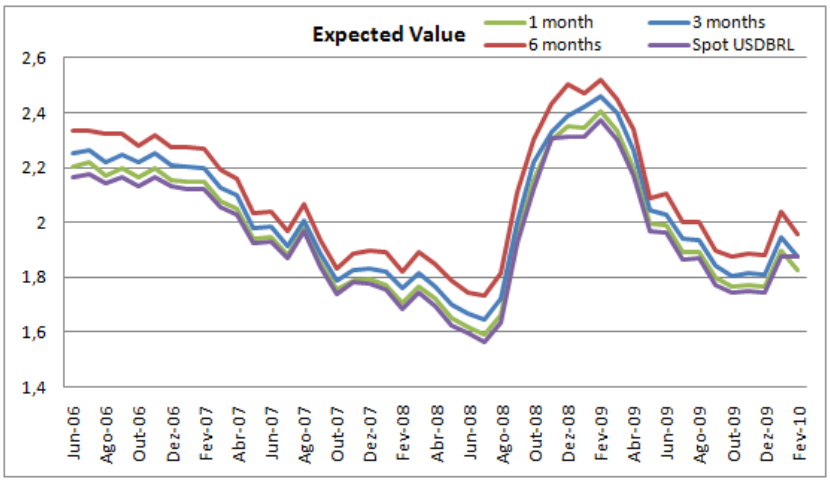


Figure 7-10: Evolution of implied expected value estimated through DFCH method

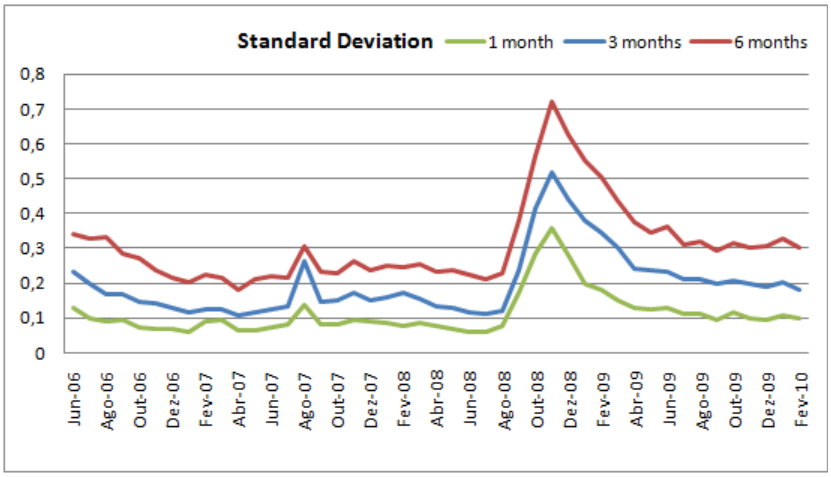


Figure 7-11: Evolution of implied standard deviation estimated through DFCH method

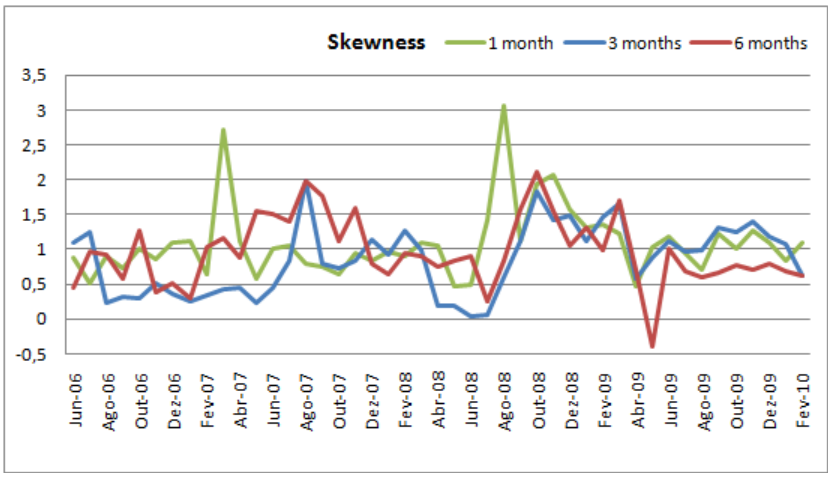


Figure 7-12: Evolution of implied skewness estimated through DFCH method

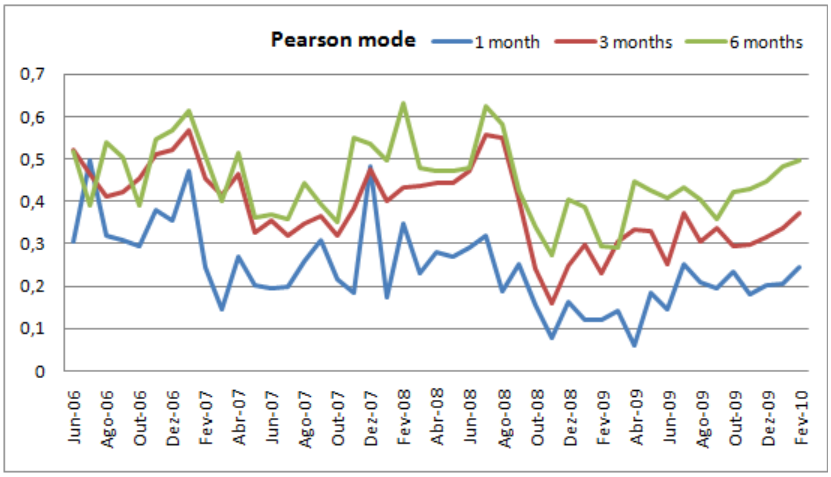


Figure 7-13: Evolution of implied Pearson mode estimated through DFCH method

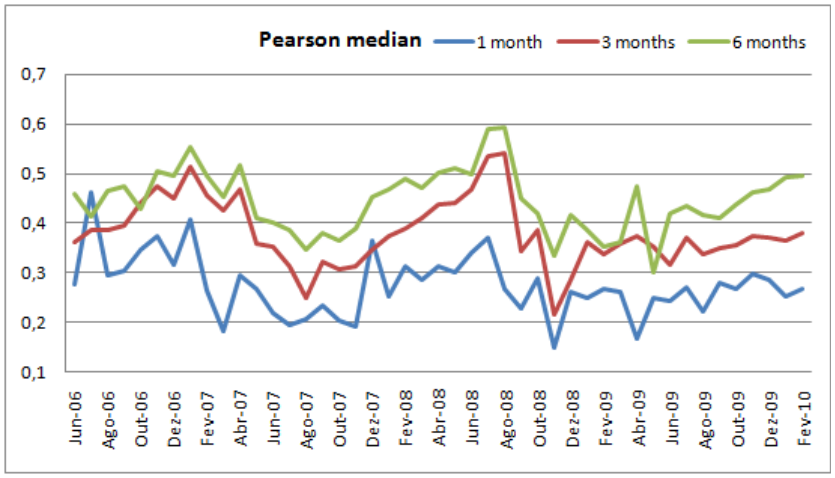


Figure 7-14: Evolution of implied Pearson median estimated through DFCH method

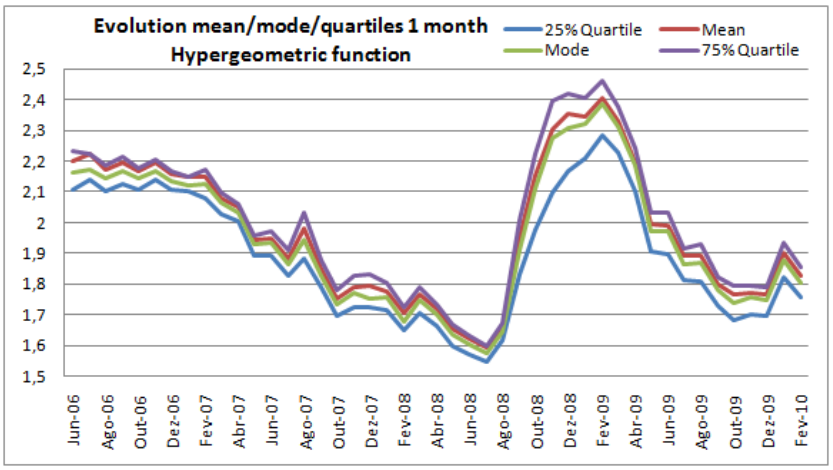


Figure 7-15: Evolution IQR 1 month

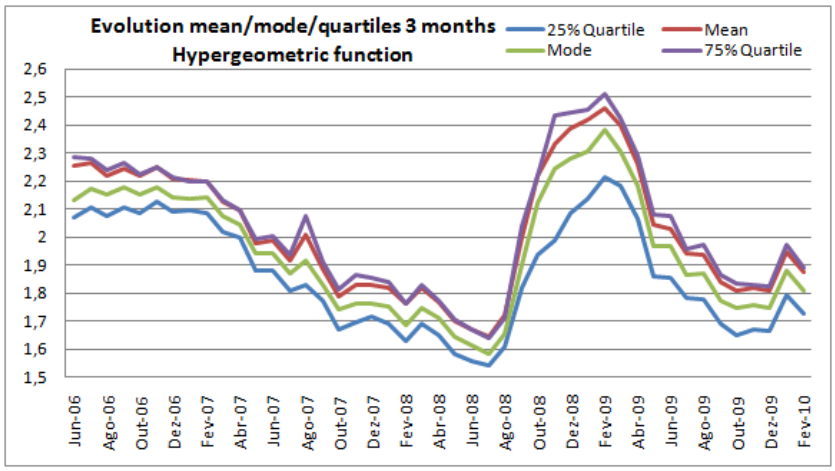


Figure 7-16: Evolution IQR 3 months

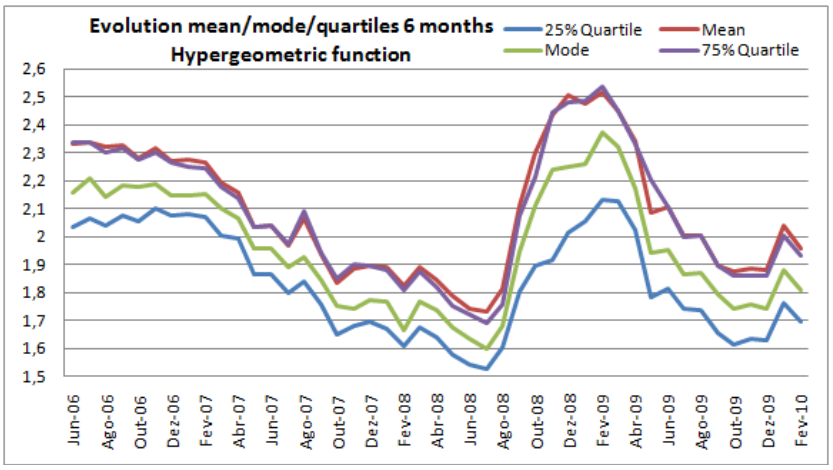


Figure 7-17: Evolution IQR 6 months

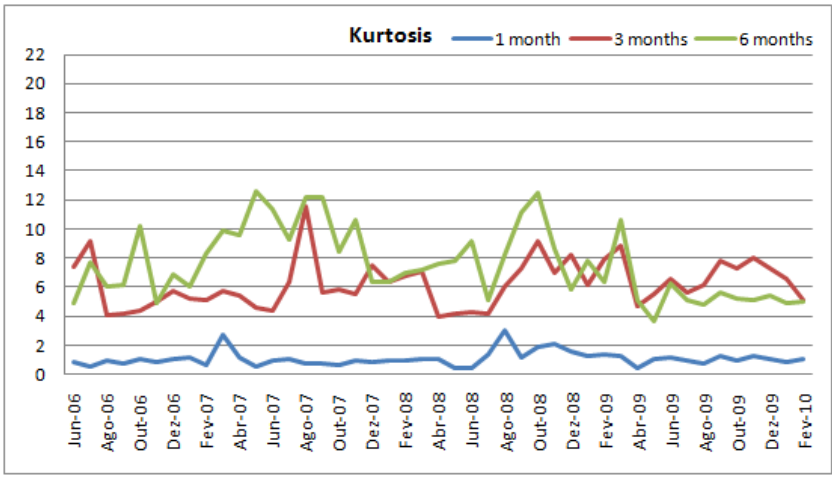


Figure 7-18: Evolution of implied kurtosis estimated through DFCH method

Chapter 8

Conclusion

This work compared the DFCH method with the widely known SML and MLN methods in the estimation of the Risk-Neutral Densities through option prices. The methodology adopted consisted of re-estimating the RNDs after adding a uniformly distributed random noise perturbation in theoretical option prices generated by Heston's stochastic volatility model for a set of different scenarios in order to test the ability of the different methods to recover the "true" RNDs under different market conditions. The "true" Heston model RNDs were produced using two approaches: in Chapter 5 we used the Heston parameters proposed in Cooper (1999) and in Chapter 6 we considered the Heston parameters that resulted from the calibration of this model for 6 low volatility dates (between October 2006 and March 2007) and 6 high volatility dates (between September 2008 and February 2009).

The three models tested were compared using two different approaches: analysis using the RMISE criteria which is a measure of the average distance between the "true" RND and the estimated ones and analysis using the summary statistics: mean, variance, skewness and kurtosis.

With the RMISE criteria we observed a higher performance of the DFCH method, especially for the low volatility dates (between October 2006 and March 2007) and high volatility dates (between September 2008 and February 2009). However, we noticed that

the MLN method was superior in capturing the "true" 6-month RNDs. In the stability analysis, we see the worst performance of the DFCH (higher RIV) in the Cooper scenarios, with the v weighting SML method showing the best results. For the high volatility dates and low volatility dates, the MLN model was the most unstable according to all statistical criteria. Despite its lower stability, the DFCH method showed a higher overall quality as the "true" RND estimator in accordance with its estimated implied RNDs which recovered the true RNDs more closely in the majority of the cases. We also found that the v weighting scheme applied to the SML method only generates improvements in terms of stability, with the overall quality of the SML being unaffected. For the SML model we also tested a theoretically optimal λ (minimizes RMISE) and λ equal to 0.9 (because in the real world we do not know the optimal λ) as the smoothing parameter (λ). We found that the comparative analysis of the methods tested was not sensitive to these two choices of the smoothing parameter.

The comparisons using the summary statistics were carried out in terms of accuracy (comparing the mean values of the summary statistics estimated from the Monte Carlo simulations and the "true" ones) and stability (standard deviation of the summary statistics). The results regarding the mean of the distributions were better for the SML method, with the DFCH method showing an expected value that is far from the "true" values in the majority of the cases. In terms of implied volatility, the SML method performed better in the majority of scenarios proposed in Cooper (1999). The results regarding the implied volatility for the period of low volatility were favorable to the MLN method and the results for the high volatility dates were better for the DFCH method. This indicates that no method clearly outperforms others in capturing the implied volatility. Concerning the skewness, the SML model was better than the MLN one and the DFCH method returned the worst results in the majority of scenarios (Cooper, low volatility and high volatility dates). The implied kurtosis obtained through the DFCH method was closer to the "true" kurtosis in the majority of the Cooper scenarios tested. The implied kurtosis for the periods of low and high volatility was favorable to the MLN method. We

also found for the skewness and kurtosis that the SML had a slight improvement when we adopted the v weighting scheme. In the stability analysis we conclude that the SML model significantly increases its stability when the v weighting is adopted. The SML method was the most stable for the variance, Skewness and kurtosis estimates.

Despite also analyzing the summary statistics, we focused our analysis on the RMISE criteria because of the higher sensitivity of skewness and kurtosis to the tails of the distribution (RNDs can have an infinite variety of probability masses outside the range of available strike prices and those shapes are very dependent on the estimation methods used).

To sum up, we conclude that the DFCH method is the best estimator of the "true" RND according to the RMISE criterion. It outperforms the widely used SML and MLN methods. It was also interesting to observe that the SML method did not outperform the MLN as an estimator of the "true" distribution according to the RMISE criterion (in Cooper (1999) the SML model was considered marginally better than the MLN model in terms of accuracy of summary statistics). In fact, despite being less stable than the SML method, the MLN method showed greater accuracy, having a lower RMISE than the SML model in most of the scenarios (Cooper, low volatility and high volatility dates). The SML was the most stable model, and its performance was enhanced when the v weighting was adopted.

In this thesis, we also obtained the USDBRL implied RNDs for the period between June 2006 and February 2010 in order to analyze the difference in the summary statistics estimated using DFCH, MLN and SML methods. We observed a higher correlation between the models tested for the expected value and volatility and found almost no relation between the methods for the skewness, kurtosis, Pearson mode and Pearson median values. From this low correlation arises the need to use a RND estimation method that has a higher capacity to capture the market expectations from the real world. The estimated RNDs and the alternative measures of uncertainty, asymmetry and extreme movement tendency were also used to analyze market expectations. We found that the

probability density functions estimated using the DFCH method were able to incorporate the changes that arise from the major USDBRL market events for this sample period.

Chapter 9

Further research

In the last section we concluded that the RNDs are a powerful tool for analyzing the effect of information in market expectations. Nevertheless, we noticed that the estimated RNDs failed to predict the Brazilian real appreciation between June 2006 and February 2010. In the future, in-depth investigations about the capacity of the estimated RNDs to predict the direction and volatility of future price movements can be made.

The accuracy and stability tests used in this thesis can be applied to other currencies from emerging markets and to currencies from markets with higher liquidity (EURUSD, GBPUSD, etc), stock index options, interest rate options and other markets in order to compare the returned results.

This study can be completed testing the accuracy and stability of semi-parametric models as Hermite Polynomials and Edgeworth expansions. The key idea of the Hermite Polynomials is that the RND can be obtained through a multiplicative perturbation to the normal distribution (reference density). These perturbations incorporate deviations to the normal densities. In the Edgeworth expansions the RND is approximated by an expansion around a lognormal distribution in order to generate more complicated functions that capture the higher moments with higher accuracy.

Further analysis can be made using the Lévy processes to generate the "true" RNDs and to price options due to their interesting theoretical architecture, which appears to

describe the observed reality of financial markets (asset price processes have jumps or spikes) in a more accurate way than models based on Brownian motion.

It would also be interesting to analyze the usefulness of the studied models in the estimation of risk measures used in financial risk management, namely value-at-risk (Var) calculations and for stress testing purposes. The empirical relevance of these models in this field could then be tested through backtesting methodology. The importance of these alternative option's pricing models as tools for hedging can be analyzed, comparing the efficiency and the cost of the hedging methods using the Black and Scholes model versus alternative methods as pricing tools.

Chapter 10

Appendix A

10.1 Geometric Brownian motion

Let us assume that the dynamics of the underlying asset is in the form of a stochastic differential equation (SDE) which evolves according to the following diffusion process:

$$dS_t = \mu dt + \sigma dW_t \tag{10.1}$$

where dS_t is the instantaneous price change, μ is the expected return, σ is the constant volatility of the price process and dW_t is an infinitesimal increment from a Wiener process with $dW_t \sim N(0, dt)$. The parameters μ and σ are assumed to be constant over time.

The Wiener process is a particular type of Markov stochastic process, with mean change of 0 and variance rate of 1 per year.

$$dW_t = \varepsilon\sqrt{dt}, \varepsilon \sim N(0,1) \tag{10.2}$$

$$dW_t \sim N(0, \sqrt{dt}) \tag{10.3}$$

A variance rate of 1 means that the variance in W_t in a time interval with length t equals t.

Nevertheless, these conditions do not guarantee the non-negativity of the stock price and implies that the expected return and volatility are constant, independently of the level of the stock price.

In order to face this problem, the expected return and the variability of the change should be proportional to the stock price, which gives:

$$dS_t = \mu S dt + \sigma S_t dW_t \tag{10.4}$$

The discrete version of this model is

$$\delta S_t = \mu S \delta t + \sigma S \delta W_t \tag{10.5}$$

$$\frac{\delta S_t}{S_t} = \mu \delta t + \sigma \delta W_t \tag{10.6}$$

$$\frac{dS_t}{S_t} \sim N(\mu \delta t, \sigma \sqrt{\delta t}) \tag{10.7}$$

Through this model, known as geometric Brownian motion, we can infer the dynamics of the underlying asset.

10.2 Itô's Lemma

After analyzing the dynamics of S_t we are interested in the dynamics of the price of the derivative asset, which we denote as $f(S_t, t)$.

If f admits a derivative we have the following discrete step for f:

$$\delta f(x) = f(x + \Delta) - f(x) = f'(x)\Delta + O(\Delta) \tag{10.8}$$

$$\lim_{\Delta \to 0} \frac{O(\Delta)}{\Delta} = 0 \tag{10.9}$$

Let us assume an SDE (stochastic differential equation) $dX_t = \mu_t dt + \sigma_t dW_t$, which has the discrete form $X_{t+\Delta} = X_t + \mu_t \Delta + \sigma_t (W_{t+\Delta} - W_t)$.

We are now interested in the dynamics of $f(X_t, t)$. For this purpose we get (using second order Taylor expansion):

$$\begin{aligned} f(X_{t+\Delta}, t+\Delta) &= f(X_t, t) + \frac{df}{dX}(X_{t+\Delta} - X_t) + \frac{df}{dt}(t+\Delta-t) \\ &+ \frac{1}{2}\left[\frac{d^2 f}{dX^2}(X_{t+\Delta} - X_t)^2 + 2\frac{df}{dX dt}(X_{t+\Delta} - X_t)(t+\Delta-t) + \frac{d^2 f}{dt^2}(t+\Delta-t)^2\right] \end{aligned} \tag{10.10}$$

If we keep the terms that are of the same order of magnitude as Δ and $(W_{t+\Delta} - W_t)$ and drop all the other terms that are smaller, we obtain (see the details in Jondeau et al. (2006)):

$$X_{t+\Delta} - X_t = \mu_t \Delta + \sigma_t (W_{t+\Delta} - W_t) \tag{10.12}$$

$$(X_{t+\Delta} - X_t)^2 = \mu_t^2 \Delta^2 + \sigma_t^2 (W_{t+\Delta} - W_t)^2 + 2\mu_t \sigma_t (W_{t+\Delta} - W_t)\Delta \approx \sigma^2 \Delta \tag{10.13}$$

$$(X_{t+\Delta} - X_t)\Delta = \mu_t \Delta^2 + \sigma_t (W_{t+\Delta} - W_t)\Delta \approx 0 \tag{10.14}$$

If we replace the terms in the Taylor expansion with the equations (10.12), (10.13) and (10.14) and taking the limit $\Delta \to 0$, we have the Itô's lemma (see the details in Jondeau et al. (2006)):

$$df = \left[\frac{1}{2}\frac{d^2 f}{dX^2}\sigma_t^2 + \frac{df}{dX}\mu_t + \frac{df}{dt}\right] dt + \frac{df}{dX}\sigma_t dW_t \tag{10.15}$$

If we define $S_t = f(X_t, t) = \exp(X_t)$, Itô's lemma gives:

$$dS_t = (\frac{1}{2} S_t \sigma^2 + S_t \mu) dt + \sigma S_t dW_t \tag{10.16a}$$

Defining $\mu_S = \mu + \frac{1}{2}\sigma^2$, we have the following dynamics for S_t:

$$dS_t = \mu_S S_t dt + \sigma S_t dW_t \tag{10.17a}$$

This is the geometric Brownian motion discussed previously.

We can see that S_t has some nice properties in describing the behavior of the asset price: S_t cannot be negative and the returns defined in this way have a constant variance, independently from the level:

$$\frac{S_t - S_{t-1}}{S_{t-1}} = \mu_S dt + \sigma dW_t \sim N(\mu_{_S}, \sigma^2) \tag{10.18a}$$

Applying $d\log(S_t) = \mu dt + \sigma dW_t$, with $\mu = (\mu_S - \frac{1}{2}\sigma^2)$ and then integrating, we have:

$$\log(S_t) - \log(S_0) = (\mu_S - \frac{1}{2}\sigma^2)t + \sigma(W_t - W_0) \tag{10.19}$$

If we consider $(W_t - W_0) \sim N(0, t)$, it follows that $\log(S_t) \sim N(\log(S_0) + (\mu_S - \frac{1}{2}\sigma^2)t, \sigma^2 t)$.

Thus, the price has a log-normal distribution and the returns are normally distributed.

Applying Ito's lemma in equation (10.17a) results in the pricing process:

$$df = \left[\frac{1}{2}\frac{d^2 f}{dS^2}S_t^2\sigma^2 + \frac{df}{dS}S_t\mu + \frac{df}{dt}\right]dt + \frac{df}{dS}S_t\sigma dW_t \tag{10.20}$$

If we create a portfolio composed by 1 unit of the derivative asset and a short position with a delta quantity ($\Delta = \frac{df}{dS}$) of the underlying asset, it can be shown that the dynamic of this portfolio does not have risk. In fact, the portfolio value is $V_t = f - \frac{df}{dS}S_t$, with the price evolving according to:

$$dV_t = df - \frac{df}{dS}dS_t \tag{10.21}$$

Substituting df and dS_t we obtain the formula:

$$
\begin{aligned}
dV_t &= \left[\frac{1}{2}\frac{d^2f}{dS^2}S_t^2\sigma^2 + \frac{df}{dS}S_t\mu + \frac{df}{dt}\right]dt + \frac{df}{dS}S_t\sigma dW_t - \frac{df}{dS}(S_t\mu dt + S_t\sigma dW_t) \qquad (10.22a)\\
&= \left[\frac{1}{2}\frac{d^2f}{dS^2}S_t^2\sigma^2 + \frac{df}{dS}S_t\mu + \frac{df}{dt} - \frac{df}{dS}S_t\mu\right]dt + \frac{df}{dS}S_t\sigma dW_t - \frac{df}{dS}S_t\sigma dW_t\\
&= \left[\frac{1}{2}\frac{d^2f}{dS^2}S_t^2\sigma^2 + \frac{df}{dS}S_t\mu + \frac{df}{dt} - \frac{df}{dS}S_t\mu\right]dt
\end{aligned}
$$

Since there is no (dW_t) term, the instantaneous return of this portfolio equal to the risk free rate (no arbitrage opportunities).

In fact, imposing the equality between equation (10.22a) and $r(f - \frac{df}{dS}S_t)dt$, results in the following equation:

$$
\frac{1}{2}\frac{d^2f}{dS^2}S_t^2\sigma^2 + \frac{df}{dS}S_tr + \frac{df}{dt} = rf \qquad (10.23)
$$

This is the Black and Scholes fundamental partial differential equation (PDE). It governs the prices of all derivatives, considering that S_t has the price dynamics defined by equation (10.17a). This equation establishes the conditions that must be satisfied by the price of a derivative written on S_t.

The solution for this PDE depends on the boundary conditions, which means that the options prices depend on the future price of the underlying asset and time to maturity:

$$
\begin{aligned}
C(S_T, T, X) &= \max(S_T - X; 0) \text{ if it is call} \qquad (10.24)\\
P(S_T, T, X) &= \max(X - S_T; 0) \text{ if it is put}
\end{aligned}
$$

10.3 Stochastic Volatility

In the Heston model we have the following two Wiener processes:

$$dS_t = \mu S_t dt + S_t \sqrt{v_t} dZ_{1,t} \tag{10.25}$$

$$dv_t = \kappa(\theta - v_t)dt + \sigma\sqrt{v_t} dZ_{2,t}, \tag{10.26}$$

where $Z_{1,t}$ and $Z_{2,t}$ are correlated Wiener processes ($Corr[dZ_{1,t}|dZ_{2,t}] = \rho dt$), v_t is the the volatility of the underlying asset, θ is the long run volatility, σ is the volatility of the volatility process and κ is the speed by which volatility returns to its long run average.

If we rewrite equations (10.25) and (10.26) in the shorter form:

$$dS_t = \mu_S dt + \sigma_S dZ_{1,t} \tag{10.27}$$

$$dv_t = \mu_v dt + \sigma_v dZ_{2,t} \tag{10.28}$$

and apply the Heston Model bivariate Itô's lemma, the dynamics for the option price is (see the details in Jondeau et al. (2006)):

$$dC = \left[\frac{1}{2}\frac{d^2C}{dS_t^2}\sigma_S^2 + \rho\sigma_S\sigma_v\frac{d^2C}{dS_t dv_t} + \frac{1}{2}\frac{d^2C}{dv_t^2}\sigma_v^2 + \mu_S\frac{dC}{dS_t} + \mu_v\frac{dC}{dv_t} + \frac{dC}{dt}\right]dt \tag{10.29}$$

$$+\sigma_S\frac{dC}{dS_t}dZ_{1,t} + \sigma_v\frac{dC}{dv_t}dZ_{2,t} \tag{10.30}$$

The risk free portfolio Π_t obtained by selling one unit of a call option (C), purchasing δ units of the underlying asset and γ units of a second derivative (C_2) on the same underlying, can be represented by:

$$
\begin{aligned}
d\Pi_t &= dC - \delta dS_t - \gamma dC_1 \qquad (10.31)\\
&= \left[\frac{1}{2}\frac{d^2C}{dS_t^2}\sigma_S^2 + \rho\sigma_S\sigma_v\frac{d^2C}{dS_t dv_t} + \frac{1}{2}\frac{d^2C}{dv_t^2}\sigma_v^2 + \mu_S\frac{dC}{dS_t} + \mu_v\frac{dC}{dv_t} + \frac{dC}{dt} - \delta\mu_S\right]dt\\
&- \left[\frac{1}{2}\frac{d^2C_1}{dS_t^2}\sigma_S^2 + \rho\sigma_S\sigma_v\frac{d^2C_1}{dS_t dv_t} + \frac{1}{2}\frac{d^2C_1}{dv_t^2}\sigma_v^2 + \mu_S\frac{dC_1}{dS_t} + \mu_v\frac{dC_1}{dv_t} + \frac{dC_1}{dt}\right]dt\\
&+ \left[\sigma_S\frac{dC}{dS_t} - \delta\sigma_S - \gamma\sigma_S\frac{dC_1}{dS_t}\right]dZ_{1,t} + \left[\sigma_v\frac{dC}{dv_t} - \gamma\sigma_v\frac{dC_1}{dv_t}\right]dZ_{2,t}
\end{aligned}
$$

The terms in $dZ_{1,t}$ and $dZ_{2,t}$ must be zero in order to obtain a portfolio without risk. This fact results in

$$\frac{dC}{dS_t} = \delta + \gamma\frac{dC_1}{dS_t} \qquad (10.32)$$

$$\frac{dC}{dv_t} = \gamma\frac{dC_1}{dv_t} \qquad (10.33)$$

The instantaneous return for this portfolio must be the risk-free rate to avoid arbitrage:

$$d\Pi_t = r(C - \delta S_t - \gamma C_1)dt \qquad (10.34a)$$

If equation (10.34a) is used on equation (10.31) and we replace δ and γ using the results in equations (10.32) and (10.33), then we get the same dynamics for derivative C and derivative $C1$:

$$
\begin{aligned}
&\left[\frac{1}{2}\frac{d^2C}{dS_t^2}\sigma_S^2 + \rho\sigma_S\sigma_v\frac{d^2C}{dS_t dv_t} + \frac{1}{2}\frac{d^2C}{dv_t^2}\sigma_v^2 + rS_t\frac{dC}{dS_t} + \mu_v\frac{dC}{dv_t} + \frac{dC}{dt} - rc\right]/\frac{dC}{dv_t} \qquad (10.35a)\\
&= \left[\frac{1}{2}\frac{d^2C_1}{dS_t^2}\sigma_S^2 + \rho\sigma_S\sigma_v\frac{d^2C_1}{dS_t dv_t} + \frac{1}{2}\frac{d^2C_1}{dv_t^2}\sigma_v^2 + rS_t\frac{dC_1}{dS_t} + \mu_v\frac{dC_1}{dv_t} + \frac{dC_1}{dt} - rC_1\right]/\frac{dC_1}{dv_t}
\end{aligned}
$$

We observe that both sides of the equation are equal and do not depend on the type

of the option. Both terms only depend on S_t, v_t and t and can be expressed as a function $\lambda(S_t, v_t, t)$ which is the volatility risk premium.

The Heston PDE is

$$\frac{1}{2}\frac{d^2C}{dS_t^2}S_t^2 v_t + \rho S_t \sigma v_t \frac{d^2C}{dS_t dv_t} + \frac{1}{2}\frac{d^2C}{dv_t^2}\sigma^2 v_t + [\kappa(\theta - v_t) - \lambda(S_t, v_t, t)]\frac{dC}{dv_t} + rS_t\frac{dC}{dS_t} + \frac{dC}{dt} = 0 \tag{10.36a}$$

Considering $x = \log(S_t)$ and $C(e^x, t)$, the PDE can be rewritten as:

$$\frac{1}{2}\frac{d^2C}{dx_t^2}v_t + \rho\sigma v_t \frac{d^2C}{dx dv_t} + \frac{1}{2}\frac{d^2C}{dv_t^2}\sigma^2 v_t + [\kappa(\theta - v_t) - \lambda(x_t, v_t, t)]\frac{dC}{dv_t} + r\frac{dC}{dx} + \frac{dC}{dt} = 0 \tag{10.37a}$$

For an European call option, the following boundary conditions must be satisfied:

$$C(S_T, v_t, r, X, T, t) = \max(S_T - X; 0) \tag{10.38}$$

$$C(0, v_t, r, X, T, t) = 0 \tag{10.39}$$

$$\frac{dC}{dS_t}(\infty, v_t, r, X, T, t) = 1 \tag{10.40}$$

10.4 Mixture of hypergeometric functions

The function DFCH (density function based on confluent hypergeometric functions), that specifies European call pricing as a mixture of two confluent hypergeometric functions, is given by (see the details in Abadir and Rockinger (2003)):

$$\begin{aligned} C(X) = {} & c_1 + c_2 X + l_{X>m_1} a_1((X - m_1)^{b_1})_1 F_1(a_2; a_3; b_2(X - m_1)^{b_3}) \\ & + (a_4)_1 F_1(a_5; a_6; b_4(X - m_2)^2), \end{aligned} \tag{10.41}$$

where $a_{3,}a_6 \in \mathbb{N}$ and $b_2, b_4 \in \mathbb{R}^-$. The indicator function l represent a component of the density with bounded support.

The Kummer's function ${}_1F_1$ was defined in equation (3.16):

$$ {}_1F_1(\alpha, \beta, z) \equiv \sum_{j=0}^{\infty} \frac{(\alpha)_j}{\beta_j} \frac{z^j}{j!} \equiv 1 + \frac{\alpha}{\beta} z + \frac{\alpha(\alpha+1)}{\beta(\beta+1)} \frac{z^2}{2!} + \frac{\alpha(\alpha+1)(\alpha+2)}{\beta(\beta+1)(\beta+2)} \frac{z^3}{3!} + ... \quad (10.42) $$

The first derivative of ${}_1F_1(\alpha, \beta, z)$ is

$$ {}_1F_1(\alpha, \beta, z)' \equiv \frac{\alpha}{\beta} + \frac{\alpha(\alpha+1)}{\beta(\beta+1)} z + \frac{\alpha(\alpha+1)(\alpha+2)}{\beta(\beta+1)(\beta+2)} \frac{z^2}{2!} + \frac{\alpha(\alpha+1)(\alpha+2)(\alpha+3)}{\beta(\beta+1)(\beta+2)(\beta+3)} \frac{z^3}{3!} + ... \quad (10.43) $$

$$ = \frac{\alpha}{\beta}[1 + \frac{(\alpha+1)}{(\beta+1)} z + \frac{(\alpha+1)(\alpha+2)}{(\beta+1)(\beta+2)} \frac{z^2}{2!} + ...] $$

$$ = \frac{\alpha}{\beta}[{}_1F_1(\alpha+1, \beta+1, z)]. $$

The Kummer's function has the following asymptotic representation for $X \in \mathbb{R}$ (see the details in Abadir (1999)),

$$ {}_1F_1(\alpha, \beta, z) = \begin{cases} \frac{\Gamma(\beta)}{\Gamma(\beta-a)} |z|^{-a} \left(1 + O\left(\left(\frac{1}{2}\right)\right), \right. & \text{as } z \longrightarrow -\infty \\ \frac{\Gamma(\beta)}{\Gamma(a)} |z|^{a-c} \exp^z \left(1 + O\left(\left(\frac{1}{2}\right)\right), \right. & \text{as } z \longrightarrow \infty \end{cases} \quad (10.44) $$

With the formula (10.43) we can obtain the implied probability density function which is given by the second derivative of $C(X)$ with respect to the strike price X.

$$\frac{d^2C(X)}{dX^2} = e^{-r\tau} f(X) = l_{X>m1} a_1 (X - m_1)^{b1-2} [b_1(b_1 - 1)_1F_1(a_2; a_3; b_2(X - m_1)^{b_3}) \tag{10.45}$$

$$\begin{aligned}
&+ \frac{a_2}{a_3} b_2 b_3 (2b_1 + b_3 - 1)(X - m_1)^{b_3} \\
&\times_1 F_1(a_2 + 1; a_3 + 1; b_2(X - m_1)^{b_3}) + \frac{a_2(a_2 + 1)}{a_3(a_3 + 1)} b_2^2 b_3^2 (X - m_1)^{2b_3} \\
&\times_1 F_1(a_2 + 2; a_3 + 2; b_2(X - m_1)^{b_3})] \\
&+ 2a_4 \frac{a_5}{a_6} b_4 [{}_1F_1(a_5 + 1; a_6 + 1; b_4(X - m_2)^2) \\
&+ 2\frac{a_5 + 1}{a_6 + 1} b_4 (X - m_2)^2 {}_1F_1(a_5 + 2; a_6 + 2; b_4(X - m_2)^2)]
\end{aligned}$$

The pdf (probability density function) derived from DFCH must be integrate to 1. To restrict the integral of $f(X)$ we derive

$$f(X) = \frac{dC(X)}{dX} = -\exp^{-r\tau}(1 - G(X)) = c_2 + a_1 b_1 (X - m_1)^{b_1 - 1} \; {}_1F_1(a_2, a_3, b_2(X - m_1)^{b_3}) \tag{10.46}$$

$$\begin{aligned}
&+ \frac{a_2}{a_3} \; {}_1F_1(a_2 + 1, a_3 + 1, b_2(X - m_1)^{b_3}) b_2 b_3 (X - m_1)^{b_3} \\
&\times a_1 (X - m_1)^{b_1 - 1} \\
&+ a_4 \; {}_1F_1(a_5 + 1, a_6 + 1, b_4(X - m_2)^2) 2b_4 (X - m_2) \frac{a_5}{a_6} \\
&= c_2 + l_{X>m_1} a_1 (X - m_1)^{b_1 - 1} [(b_1)_1F_1(a_2; a_3; b_2(X - m_1)^{b_3}) \qquad (10.47) \\
&+ \frac{a_2}{a_3} b_2 b_3 ((X - m_1)^{b_3})_1F_1(a_2 + 1; a_3 + 1; b_2(X - m_1)^{b_3})] \\
&+ 2a_4 \frac{a_5}{a_6} b_4 (X - m_2)_1F_1(a_5 + 1; a_6 + 1; b_4(X - m_2)^2)
\end{aligned}$$

In order to guarantee that $f(X)$ integrates to 1 between X_l and X_u we set,

$$\int_{X_l}^{X_u} f(X)dX = 1, \tag{10.48}$$

which is equivalent to

$$\frac{dC(X_l)}{dX} = G(X_l) - 1 = -1, \tag{10.49}$$

$$\frac{dC(X_u)}{dX} = G(X_u) - 1 = 0. \tag{10.50}$$

Solving the restrictions on equations (10.49) and (10.50), we obtain explicit formulas for the parameters c_2 and a_4. If we assume that $X_l < m_1$, from the constrain set in equation (10.49), we conclude that c_2 is defined as

$$c_2 = -1 - 2a_4\frac{a_5}{a_6}b_4(X_l - m_2){}_1F_1(a_5 + 1, a_6 + 1, b_4(X_l - m_2)^2) \tag{10.51}$$

Applying the restriction on X_u defined in equation (10.50), we get c_2 plus the other terms of (10.46) which give the following explicit formula for c_2

$$\begin{aligned} c_2 = &-a_1(X_u - m_1)^{b_1-1}[(b_1){}_1F_1(a_2; a_3; b_2(X_u - m_1)^{b_3}) \\ &+ \frac{a_2}{a_3}b_2b_3((X_u - m_1)^{b_3}){}_1F_1(a_2 + 1; a_3 + 1; b_2(X_u - m_1)^{b_3})] \\ &- 2a_4\frac{a_5}{a_6}b_4(X_u - m_2){}_1F_1(a_5 + 1; a_6 + 1; b_4(X_u - m_2)^2). \end{aligned} \tag{10.52}$$

If we compare equations (10.51) and (10.52), we get an explicit formula for a_4,

$$a_4 = \left\{ \begin{array}{c} 1 - a_1(X_u - m_1)^{b_1 - 1}[(b_1)_1F_1(a_2; a_3; b_2(X_u - m_1)^{b_3}) \\ + \frac{a_2}{a_3} b_2 b_3((X_u - m_1)^{b_3})_1F_1(a_2 + 1; a_3 + 1; b_2(X_u - m_1)^{b_3})] \end{array} \right\} \tag{10.53}$$
$$\div \left\{ \begin{array}{c} 2\frac{a_5}{a_6} b_4[(X_u - m_2)_1F_1(a_5 + 1; a_6 + 1; b_4(X_u - m_2)^2) \\ -(X_l - m_2)_1F_1(a_5 + 1, a_6 + 1, b_4(X_l - m_2)^2)] \end{array} \right\}$$

In Abadir and Rockinger (2003), the assumptions $b_1 = 1 + a_2 b_3$; $a_5 = -\frac{1}{2}$ and $a_6 = \frac{1}{2}$ were applied in equations (10.51) and (10.53). Using the asymptotic representation of Kummer's function in equation (10.44), equation (10.51) simplifies to

$$\begin{aligned} c_2 &= -1 - 2a_4 \frac{a_5}{a_6} b_4(X_l - m_2) \frac{\Gamma(a_6 + 1)}{\Gamma(a_6 - a_5)} \left| b_4(X_l - m_2)^2 \right|^{-a_5 - 1} \\ &= -1 + a_4 \sqrt{-b_4 \times \pi} \end{aligned} \tag{10.54}$$

and equation (10.53) simplifies to

$$a_4 = \frac{1}{2\sqrt{-b_4 \pi}} \left[1 - a_1(-b_2)^{-a_2} \frac{\Gamma(a_3)}{\Gamma(a_3 - a_2)} \right]. \tag{10.55}$$

This formula was deduced by simplifying the two terms of equation (10.53) separately. The first term is

$$\begin{aligned} &1 - a_1(X_u - m_1)^{b_1 - 1}[(b_1)_1F_1(a_2; a_3; b_2(X_u - m_1)^{b_3}) \\ &+ \frac{a_2}{a_3} b_2 b_3((X_u - m_1)^{b_3})_1F_1(a_2 + 1; a_3 + 1; b_2(X_u - m_1)^{b_3})] \end{aligned} \tag{10.56}$$

applying the relation ${}_1F_1(\alpha, \beta, z) = \exp^z {}_1F_1(\beta - \alpha, \beta, -z)$ [1] we continue the simplification of the first term

[1] This transformation is shown by Karim Abadir (1999) in "An introduction to hypergeometric functions for economiste"

$$
\begin{aligned}
&1 - a_1(X_u - m_1)^{b_1-1}[b_1(\exp^{b_2(X_u-m_1)^{b_3}})_1F_1(a_3 - a_2; a_3; -b_2(X_u - m_1)^{b_3}) \qquad (10.57)\\
&+ \frac{a_2}{a_3}b_2b_3((X_u - m_1)^{b_3})(\exp^{b_2(X_u-m_1)^{b_3}})_1F_1(a_3 - a_2; a_3 + 1; -b_2(X_u - m_1)^{b_3})]\\
&= 1 - a_1(X_u - m_1)^{b_1-1}[b_1 \exp^{b_2(X_u-m_1)^{b_3}} \frac{\Gamma(a_3)}{\Gamma(a_3 - a_2)}(-b_2(X_u - m_1)^{b_3})^{-a_2} \exp^{-b_2(X_u-m_1)^{b_3}}\\
&+ \frac{a_2}{a_3}b_2b_3(X_u - m_1)^{b_3} \exp^{b_2(X_u-m_1)^{b_3}} \frac{\Gamma(a_3 + 1)}{\Gamma(a_3 - a_2)}(-b_2(X_u - m_1)^{b_3})^{-a_2-1} \exp^{-b_2(X_u-m_1)^{b_3}}]\\
&= 1 - a_1(X_u - m_1)^{b_1-1}b_1\frac{\Gamma(a_3)}{\Gamma(a_3 - a_2)}(-b_2(X_u - m_1)^{b_3})^{-a_2}\\
&- a_1(X_u - m_1)^{b_1-1}\frac{a_2}{a_3}b_2b_3(X_u - m_1)^{b_3}\frac{\Gamma(a_3 + 1)}{\Gamma(a_3 - a_2)}(-b_2(X_u - m_1)^{b_3})^{-a_2-1}.
\end{aligned}
$$

Taking into account that $\Gamma(a_3 + 1) = a_3!$ we transform the $\Gamma(a_3 + 1)$ into $\Gamma(a_3) \times a_3$, which gives

$$
\begin{aligned}
&1 - a_1(X_u - m_1)^{b_1-1}b_1\frac{\Gamma(a_3)}{\Gamma(a_3 - a_2)}(-b_2(X_u - m_1)^{b_3})^{-a_2} \qquad (10.58)\\
&- a_1(X_u - m_1)^{b_1-1}\frac{a_2}{a_3}b_2b_3(X_u - m_1)^{b_3}\frac{\Gamma(a_3)}{\Gamma(a_3 - a_2)}a_3(-b_2(X_u - m_1)^{b_3})^{-a_2-1}\\
&= 1 - a_1(X_u - m_1)^{b_1-a_2b_3-1}b_1\frac{\Gamma(a_3)}{\Gamma(a_3 - a_2)}(-b_2)^{-a_2}\\
&- a_1(X_u - m_1)^{b_1-a_2b_3-1}a_2b_3\frac{\Gamma(a_3)}{\Gamma(a_3 - a_2)}(-b_2)^{-a_2}b_2(-b_2)^{-1}\\
&= 1 - a_1(-b_2)^{-a_2}\frac{\Gamma(a_3)}{\Gamma(a_3 - a_2)}.
\end{aligned}
$$

Applying the same transformation set in the first term ${}_1F_1(\alpha, \beta, z) = e^z \, {}_1F_1(\beta-\alpha, \beta, -z)$, the second term of equation (10.53) is

$$
\begin{aligned}
&2\frac{a_5}{a_6}b_4[(X_u-m_2){}_1F_1(a_5+1;a_6+1;b_4(X_u-m_2)^2) \qquad (10.59)\\
&-(X_l-m_2){}_1F_1(a_5+1,a_6+1,b_4(X_l-m_2)^2)]\\
&=-2b_4[(X_u-m_2){}_1F_1(a_5+1;a_6+1;b_4(X_u-m_2)^2)\\
&-(X_l-m_2){}_1F_1(a_5+1,a_6+1,b_4(X_l-m_2)^2)]\\
&=-2b_4[(X_u-m_2)\exp^{b_4(X_u-m_2)^2}\ {}_1F_1(a_6-a_5;a_6+1;-b_4(X_u-m_2)^2)\\
&-(X_l-m_2){}_1F_1(a_5+1;a_6+1;b_4(X_l-m_2)^2]\\
&=-2b_4[(X_u-m_2)\exp^{b_4(X_u-m_2)^2}\ \frac{\Gamma(a_6+1)}{\Gamma(a_6-a_5)}(-(b_4(X_u-m_2)^2))^{-a_5-1}\exp^{-b_4(X_u-m_2)^2}\\
&-(X_l-m_2)\frac{\Gamma(a_6+1)}{\Gamma(a_6-a_5)}\left|b_4(X_l-m_2)^2\right|^{-a_5-1}]\\
&=-2b_4(X_u-m_2)(X_u-m_2)^{-2a_5-2}(-b_4)^{-a_5-1}\frac{\Gamma(a_6+1)}{\Gamma(a_6-a_5)}\\
&+2b_4(X_l-m_2)(-(X_l-m_2))^{-2a_5-2}(-b_4)^{-a_5-1}\frac{\Gamma(a_6+1)}{\Gamma(a_6-a_5)}\\
&=-2b_4(-b_4)^{-\frac{1}{2}}\frac{\Gamma(a_6+1)}{\Gamma(a_6-a_5)}+2b_4(-b_4)^{-\frac{1}{2}}\frac{(X_l-m_2)}{-(X_l-m_2)}\frac{\Gamma(a_6+1)}{\Gamma(a_6-a_5)}\\
&=-2b_4(-b_4)^{-\frac{1}{2}}\frac{\Gamma(a_6+1)}{\Gamma(a_6-a_5)}-2b_4(-b_4)^{-\frac{1}{2}}\frac{\Gamma(a_6+1)}{\Gamma(a_6-a_5)}\\
&=-4(-b_4)^{\frac{1}{2}}\frac{\Gamma(a_6+1)}{\Gamma(a_6-a_5)}=-4\sqrt{-b_4}\frac{\Gamma(\frac{3}{2})}{\Gamma(1)}=2\sqrt{b_4(\Gamma(\frac{3}{2})^2)2^2}=2\sqrt{b_4\pi}
\end{aligned}
$$

Taking into account that $(\Gamma(\frac{3}{2})^2)2^2=\pi$, we have that

$$
4\sqrt{-b_4}\Gamma(\frac{3}{2})=2\sqrt{-b_4(\Gamma(\frac{3}{2})^2)2^2}=2\sqrt{-b_4\pi}. \qquad (10.60)
$$

Chapter 11

Appendix B

DFCH

	Scenario	1 month	3 months	6 months
Expected Value	1	1,9761	2,0193	2,0899
	2	1,9733	2,0173	2,0900
	3	1,9721	2,0174	2,0878
	4	1,9914	2,1090	2,1306
	5	1,9806	2,0323	2,1135
	6	1,9703	2,0260	2,1014
Standard Deviation	1	0,1233	0,2105	0,2985
	2	0,1237	0,2142	0,3151
	3	0,1277	0,2216	0,3233
	4	0,1255	0,2700	0,3822
	5	0,1311	0,2508	0,4019
	6	0,1349	0,2775	0,4465
Skewness	1	-0,3423	-0,5169	-0,7718
	2	-0,1521	-0,2884	-0,1880
	3	0,0355	-0,0111	0,0882
	4	-0,3663	-2,4656	-1,0055
	5	-0,2466	-0,4104	-0,3136
	6	-0,0105	0,5078	0,7908
Kurtosis	1	2,9457	2,9032	2,8680
	2	3,0714	3,0209	4,1051
	3	3,4633	3,6068	4,2312
	4	3,0189	3,1169	3,0826
	5	2,9991	3,0450	3,1784
	6	3,0467	4,5411	5,6989

MLN

	Scenario	1 month	3 months	6 months
Expected Value	1	1,9651	1,9887	2,0346
	2	1,9630	1,9900	2,0312
	3	1,9610	1,9858	2,0345
	4	1,9766	2,0286	2,1194
	5	1,9686	2,0105	2,0823
	6	1,9608	1,9938	2,0810
Standard Deviation	1	0,1220	0,2068	0,2798
	2	0,1216	0,2012	0,2680
	3	0,1254	0,2117	0,2842
	4	0,1216	0,2109	0,3106
	5	0,1293	0,2393	0,3625
	6	0,1441	0,2714	0,4152
Skewness	1	0,0717	-0,0774	-0,2323
	2	0,1999	0,3246	0,4426
	3	0,3043	0,5148	0,7774
	4	-0,0595	-0,1786	-0,1362
	5	0,2369	0,4630	0,7684
	6	0,4149	0,7813	1,3350
Kurtosis	1	3,2627	4,3885	5,6363
	2	3,1160	3,2180	3,4002
	3	3,4862	3,9632	4,1651
	4	3,6043	2,6091	2,5644
	5	3,3568	3,8580	4,4296
	6	4,0123	4,5432	5,6260

Figure 11-1: Summary Statistics obtained for DFCH and MLN methods

With v weighting

SML (λ=0,9)

	Scenario	1 month	3 months	6 months
Expected Value	1	1,9569	1,9708	1,9932
	2	1,9586	1,9755	2,0001
	3	1,9591	1,9765	2,0005
	4	1,9341	1,9102	1,8892
	5	1,9547	1,9626	1,9712
	6	1,9596	1,9813	2,0118
Standard Deviation	1	0,1208	0,1976	0,2595
	2	0,1216	0,2011	0,2682
	3	0,1224	0,2054	0,2792
	4	0,1335	0,2456	0,3681
	5	0,1354	0,2559	0,3952
	6	0,1383	0,2678	0,4254
Skewness	1	0,1532	0,2677	0,3872
	2	0,3059	0,5355	0,7602
	3	0,4878	0,8409	1,1640
	4	-0,1182	-0,1199	-0,0540
	5	0,1147	0,2249	0,3744
	6	0,6767	0,9600	1,1544
Kurtosis	1	3,0172	3,0366	3,1005
	2	3,0841	3,2709	3,4996
	3	3,2101	3,6004	4,0741
	4	2,8935	2,7388	2,6008
	5	2,9884	3,0320	3,1177
	6	3,3975	3,9203	4,5497

SML (λ that minimizes RMISE)

	Scenario	1 month	3 months	6 months
Expected Value	1	1,9570	1,9708	1,9932
	2	1,9586	1,9755	2,0000
	3	1,9591	1,9765	2,0005
	4	1,9340	1,9099	1,8879
	5	1,9546	1,9621	1,9697
	6	1,9594	1,9786	2,0106
Standard Deviation	1	0,1208	0,1975	0,2595
	2	0,1216	0,2011	0,2683
	3	0,1225	0,2055	0,2793
	4	0,1335	0,2456	0,3677
	5	0,1353	0,2556	0,3946
	6	0,1383	0,2668	0,4251
Skewness	1	0,1533	0,2675	0,3873
	2	0,3061	0,5357	0,7609
	3	0,4882	0,8411	1,1640
	4	-0,1164	-0,1165	-0,0448
	5	0,1165	0,2305	0,3846
	6	0,6791	1,0243	1,1614
Kurtosis	1	3,0172	3,0367	3,1006
	2	3,0841	3,2708	3,5000
	3	3,2104	3,6007	4,0741
	4	2,8929	2,7378	2,5988
	5	2,9886	3,0329	3,1208
	6	3,3988	3,9628	4,5556

Without v weighting

SML (λ=0,9=0,9)

	Scenario	1 month	3 months	6 months
Expected Value	1	1,9559	1,9679	1,9773
	2	1,9586	1,9759	2,0011
	3	1,9596	1,9771	2,0015
	4	1,9316	1,9088	1,8957
	5	1,9610	1,9727	1,9836
	6	1,9645	1,9881	2,0190
Standard Deviation	1	0,1207	0,1964	0,2532
	2	0,1215	0,2013	0,2691
	3	0,1227	0,2060	0,2813
	4	0,1337	0,2465	0,3724
	5	0,1374	0,2609	0,4048
	6	0,1395	0,2714	0,4321
Skewness	1	0,1575	0,2573	0,3573
	2	0,3051	0,5312	0,7562
	3	0,4913	0,8541	1,2007
	4	-0,1237	-0,1445	-0,0814
	5	0,0150	0,1853	0,3752
	6	0,6038	0,9149	1,1291
Kurtosis	1	3,0130	3,0204	2,9966
	2	3,0831	3,2493	3,4834
	3	3,2162	3,6011	4,1656
	4	2,9011	2,7375	2,6161
	5	2,9969	3,0414	3,1673
	6	3,3753	3,9049	4,5538

SML (min RMISE)

	Scenario	1 month	3 months	6 months
Expected Value	1	1,9566	1,9695	1,9848
	2	1,9586	1,9759	1,9999
	3	1,9592	1,9765	2,0005
	4	1,9316	1,9072	1,8919
	5	1,9574	1,9648	1,9745
	6	1,9612	1,9820	2,0116
Standard Deviation	1	0,1206	0,1964	0,2537
	2	0,1216	0,2012	0,2690
	3	0,1226	0,2060	0,2812
	4	0,1337	0,2462	0,3713
	5	0,1367	0,2593	0,4017
	6	0,1395	0,2706	0,4300
Skewness	1	0,1368	0,2295	0,2766
	2	0,3050	0,5304	0,7686
	3	0,4992	0,8625	1,2080
	4	-0,1183	-0,1172	-0,0444
	5	0,0949	0,2808	0,4436
	6	0,6751	0,9788	1,1750
Kurtosis	1	3,0112	3,0161	3,0191
	2	3,0837	3,2486	3,4898
	3	3,2189	3,6059	4,1693
	4	2,9018	2,7308	2,6122
	5	3,0075	3,0697	3,1981
	6	3,4131	3,9554	4,5957

Figure 11-2: Summary Statistics obtained for SML method under 4 scenes: with or without v weighting and for each weighting approach using a smoothing parameter λ that minimizes RMISE or a smoothing parameter λ with a value of 0,9.

DFCH

	Scenario	1 month	3 months	6 months
Expected Value	1	-0,0090	-0,0221	-0,0448
	2	-0,0079	-0,0210	-0,0449
	3	-0,0070	-0,0218	-0,0439
	4	-0,0146	-0,0591	-0,0476
	5	-0,0088	-0,0213	-0,0405
	6	-0,0058	-0,0200	-0,0357
Standard Deviation	1	-0,0173	-0,0563	-0,1307
	2	-0,0202	-0,0646	-0,1728
	3	-0,0459	-0,0957	-0,1836
	4	0,0170	-0,1630	-0,0979
	5	-0,0118	-0,0459	-0,0962
	6	-0,0439	-0,1386	-0,1634
Skewness	1	3,3804	3,1564	3,1309
	2	1,4884	1,5347	1,2441
	3	0,9261	1,0135	0,9253
	4	-0,4491	-7,1751	-5,0902
	5	1,7555	1,7001	1,3709
	6	1,0113	0,6429	0,5802
Kurtosis	1	0,0030	0,0132	0,0139
	2	-0,0078	0,0896	-0,1441
	3	-0,0675	0,0213	0,0742
	4	-0,0185	-0,0448	-0,0886
	5	0,0869	0,1720	0,2377
	6	0,1947	0,1483	0,2443

MLN

	Scenario	1 month	3 months	6 months
Expected Value	1	-0,0034	-0,0066	-0,0172
	2	-0,0026	-0,0072	-0,0155
	3	-0,0013	-0,0058	-0,0173
	4	-0,0070	-0,0187	-0,0421
	5	-0,0026	-0,0103	-0,0251
	6	-0,0009	-0,0038	-0,0257
Standard Deviation	1	-0,0063	-0,0378	-0,0597
	2	-0,0031	0,0004	0,0026
	3	-0,0273	-0,0466	-0,0406
	4	0,0479	0,0915	0,1079
	5	0,0027	0,0019	0,0112
	6	-0,1149	-0,1134	-0,0818
Skewness	1	0,5014	1,3230	1,6414
	2	0,3580	0,3983	0,4254
	3	0,3663	0,3746	0,3410
	4	0,7648	0,4080	0,1749
	5	0,2743	0,2101	0,0912
	6	0,5517	0,4506	0,2914
Kurtosis	1	-0,1043	-0,4916	-0,9379
	2	-0,0225	0,0302	0,0524
	3	-0,0745	-0,0754	0,0887
	4	-0,2160	0,1254	0,0943
	5	-0,0220	-0,0491	-0,0624
	6	-0,0605	0,1479	0,2540

Figure 11-3: Difference between the "true" and mean summary statistics in percentage of the "true" statistics for the DFCH and MLN methods.

With v weighting

SML (λ=0,9)

	Scenario	1 month	3 months	6 months
Expected Value	1	0,0008	0,0025	0,0035
	2	-0,0003	0,0001	0,0001
	3	-0,0004	-0,0011	-0,0003
	4	0,0146	0,0407	0,0711
	5	0,0044	0,0138	0,0295
	6	-0,0003	0,0024	0,0084
Standard Deviation	1	0,0037	0,0084	0,0169
	2	-0,0028	0,0006	0,0017
	3	-0,0031	-0,0156	-0,0224
	4	-0,0453	-0,0578	-0,0573
	5	-0,0445	-0,0670	-0,0779
	6	-0,0698	-0,0987	-0,1084
Skewness	1	-0,0656	-0,1170	-0,0689
	2	0,0178	0,0072	0,0130
	3	-0,0156	-0,0215	0,0133
	4	0,5326	0,6024	0,6730
	5	0,6486	0,6164	0,5572
	6	0,2688	0,3249	0,3872
Kurtosis	1	-0,0212	-0,0321	-0,0660
	2	-0,0120	0,0142	0,0247
	3	0,0106	0,0230	0,1086
	4	0,0238	0,0819	0,0815
	5	0,0901	0,1755	0,2522
	6	0,1020	0,2647	0,3967

SML (λ that minimizes RMISE)

	Scenario	1 month	3 months	6 months
Expected Value	1	0,0008	0,0025	0,0035
	2	-0,0003	0,0002	0,0001
	3	-0,0004	-0,0011	-0,0003
	4	0,0147	0,0409	0,0717
	5	0,0045	0,0140	0,0303
	6	-0,0002	0,0038	0,0090
Standard Deviation	1	0,0034	0,0088	0,0172
	2	-0,0030	0,0010	0,0016
	3	-0,0039	-0,0159	-0,0226
	4	-0,0458	-0,0578	-0,0561
	5	-0,0442	-0,0659	-0,0763
	6	-0,0699	-0,0945	-0,1077
Skewness	1	-0,0660	-0,1162	-0,0693
	2	0,0172	0,0069	0,0121
	3	-0,0164	-0,0217	0,0132
	4	0,5396	0,6136	0,7289
	5	0,6431	0,6068	0,5451
	6	0,2662	0,2797	0,3835
Kurtosis	1	-0,0212	-0,0321	-0,0660
	2	-0,0120	0,0143	0,0246
	3	0,0104	0,0229	0,1086
	4	0,0240	0,0823	0,0822
	5	0,0901	0,1753	0,2515
	6	0,1017	0,2567	0,3959

Without v weighting

SML (λ=0,9)

	Scenario	1 month	3 months	6 months
Expected Value	1	0,0013	0,0039	0,0114
	2	-0,0003	-0,0001	-0,0005
	3	-0,0006	-0,0014	-0,0008
	4	0,0159	0,0414	0,0679
	5	0,0012	0,0087	0,0234
	6	-0,0028	-0,0010	0,0049
Standard Deviation	1	0,0044	0,0145	0,0409
	2	-0,0025	-0,0002	-0,0014
	3	-0,0052	-0,0188	-0,0298
	4	-0,0470	-0,0616	-0,0696
	5	-0,0598	-0,0882	-0,1043
	6	-0,0798	-0,1136	-0,1258
Skewness	1	-0,0952	-0,0733	0,0136
	2	0,0204	0,0152	0,0181
	3	-0,0229	-0,0375	-0,0179
	4	0,5108	0,5209	0,5070
	5	0,9541	0,6838	0,5562
	6	0,3476	0,3566	0,4007
Kurtosis	1	-0,0198	-0,0266	-0,0303
	2	-0,0117	0,0207	0,0292
	3	0,0086	0,0228	0,0886
	4	0,0213	0,0824	0,0761
	5	0,0875	0,1730	0,2404
	6	0,1079	0,2676	0,3961

SML (λ that minimizes RMISE)

	Scenario	1 month	3 months	6 months
Expected Value	1	0,0010	0,0031	0,0077
	2	-0,0004	0,0000	0,0001
	3	-0,0004	-0,0011	-0,0003
	4	0,0159	0,0422	0,0698
	5	0,0031	0,0126	0,0279
	6	-0,0011	0,0021	0,0085
Standard Deviation	1	0,0053	0,0143	0,0392
	2	-0,0030	0,0002	-0,0012
	3	-0,0043	-0,0188	-0,0297
	4	-0,0469	-0,0603	-0,0666
	5	-0,0550	-0,0814	-0,0957
	6	-0,0795	-0,1104	-0,1203
Skewness	1	0,0486	0,0427	0,2362
	2	0,0206	0,0166	0,0021
	3	-0,0393	-0,0477	-0,0241
	4	0,5320	0,6115	0,7313
	5	0,7094	0,5209	0,4754
	6	0,2705	0,3116	0,3763
Kurtosis	1	-0,0192	-0,0251	-0,0380
	2	-0,0119	0,0209	0,0274
	3	0,0078	0,0215	0,0878
	4	0,0210	0,0846	0,0775
	5	0,0843	0,1653	0,2330
	6	0,0979	0,2581	0,3906

Figure 11-4: Difference between the "true" and mean summary statistics in percentage of the "true" statistics for the SML method under 4 scenes: with or without v weighting and for each weighting approach using a smoothing parameter λ that minimizes RMISE or a smoothing parameter λ with a value of 0,9.

DFCH

	Scenario	1 month	3 months	6 months
Variance	1	0,0002	0,0024	0,0014
	2	0,0002	0,0010	0,0033
	3	0,0003	0,0007	0,0021
	4	0,0008	0,0038	0,0015
	5	0,0002	0,0014	0,0060
	6	0,0006	0,0012	0,0066
Skewness	1	0,0292	0,0197	0,0100
	2	0,0223	0,0557	0,1451
	3	0,0723	0,0242	0,0694
	4	0,1039	0,0436	0,0143
	5	0,0316	0,0275	0,0959
	6	0,1208	0,0471	0,1018
Kurtosis	1	0,0107	0,0126	0,0032
	2	0,0539	0,1020	0,3953
	3	0,1899	0,0458	0,1775
	4	0,2406	0,0332	0,0193
	5	0,0101	0,0548	0,1949
	6	0,1400	0,1532	0,3518

MLN

	Scenario	1 month	3 months	6 months
Variance	1	0,0005	0,0027	0,0080
	2	0,0001	0,0001	0,0000
	3	0,0003	0,0005	0,0007
	4	0,0014	0,0011	0,0009
	5	0,0004	0,0002	0,0015
	6	0,0005	0,0008	0,0008
Skewness	1	0,1091	0,3167	0,5376
	2	0,0246	0,0168	0,0098
	3	0,0242	0,0417	0,0333
	4	0,2673	0,0155	0,0067
	5	0,0204	0,0465	0,0375
	6	0,0867	0,0279	0,0064
Kurtosis	1	0,2327	0,8198	1,5266
	2	0,0394	0,0260	0,0174
	3	0,1476	0,1251	0,0917
	4	0,6022	0,0483	0,0253
	5	0,1590	0,2073	0,0938
	6	0,1693	0,0689	0,0292

Figure 11-5: Standard Deviation of the summary statistics for the DFCH and MLN

With v weighting

SML (λ=0,9)

	Scenario	1 month	3 months	6 months
Variance	1	0,0002	0,0004	0,0005
	2	0,0002	0,0004	0,0005
	3	0,0002	0,0004	0,0006
	4	0,0002	0,0004	0,0006
	5	0,0002	0,0005	0,0008
	6	0,0003	0,0005	0,0009
Skewness	1	0,0032	0,0011	0,0004
	2	0,0035	0,0013	0,0001
	3	0,0031	0,0006	0,0006
	4	0,0009	0,0008	0,0010
	5	0,0059	0,0035	0,0026
	6	0,0034	0,0014	0,0011
Kurtosis	1	0,0008	0,0003	0,0003
	2	0,0013	0,0020	0,0014
	3	0,0031	0,0026	0,0014
	4	0,0015	0,0010	0,0001
	5	0,0013	0,0026	0,0029
	6	0,0052	0,0036	0,0042

SML (λ that minimizes RMISE)

	Scenario	1 month	3 months	6 months
Variance	1	0,0002	0,0004	0,0005
	2	0,0002	0,0004	0,0005
	3	0,0002	0,0004	0,0006
	4	0,0002	0,0004	0,0006
	5	0,0002	0,0005	0,0007
	6	0,0003	0,0005	0,0008
Skewness	1	0,0039	0,0014	0,0003
	2	0,0040	0,0018	0,0005
	3	0,0033	0,0010	0,0001
	4	0,0011	0,0009	0,0011
	5	0,0062	0,0036	0,0027
	6	0,0036	0,0028	0,0012
Kurtosis	1	0,0009	0,0003	0,0003
	2	0,0014	0,0022	0,0017
	3	0,0032	0,0029	0,0017
	4	0,0016	0,0011	0,0000
	5	0,0013	0,0026	0,0030
	6	0,0052	0,0048	0,0042

Without v weighting

SML (smooth=0,9)

	Scenario	1 month	3 months	6 months
Variance	1	0,0004	0,0008	0,0048
	2	0,0003	0,0006	0,0011
	3	0,0003	0,0006	0,0009
	4	0,0004	0,0010	0,0015
	5	0,0003	0,0006	0,0009
	6	0,0003	0,0006	0,0010
Skewness	1	0,0198	0,0007	0,0218
	2	0,0144	0,0003	0,0038
	3	0,0123	0,0034	0,0026
	4	0,0603	0,0196	0,0006
	5	0,0017	0,0039	0,0031
	6	0,0023	0,0010	0,0013
Kurtosis	1	0,0047	0,0104	0,1591
	2	0,0021	0,0069	0,0134
	3	0,0033	0,0043	0,0070
	4	0,0308	0,0141	0,0084
	5	0,0018	0,0035	0,0038
	6	0,0044	0,0038	0,0046

SML (min RMISE)

	Scenario	1 month	3 months	6 months
Variance	1	0,0003	0,0007	0,0048
	2	0,0003	0,0006	0,0011
	3	0,0003	0,0006	0,0009
	4	0,0004	0,0008	0,0014
	5	0,0003	0,0005	0,0008
	6	0,0003	0,0006	0,0009
Skewness	1	0,0251	0,0284	0,1362
	2	0,0018	0,0020	0,0168
	3	0,0140	0,0073	0,0047
	4	0,0147	0,0018	0,0051
	5	0,0091	0,0063	0,0035
	6	0,0056	0,0025	0,0016
Kurtosis	1	0,0089	0,0143	0,1058
	2	0,0049	0,0060	0,0201
	3	0,0133	0,0108	0,0119
	4	0,0150	0,0080	0,0079
	5	0,0045	0,0048	0,0043
	6	0,0091	0,0052	0,0051

Figure 11-6: Standard deviation of the summary statistics for the SML method under 4 scenes: with or without v weighting and for each weighting approach using a smoothing parameter λ that minimizes RMISE or a smoothing parameter λ with a value of 0,9.

DFCH

	Scenario	1 month	3 months	6 months
RMISE	1	0,0275	0,0535	0,0839
	2	0,0416	0,0511	0,0918
	3	0,0437	0,0612	0,0961
	4	0,1524	0,0827	0,1040
	5	0,1102	0,0541	0,0724
	6	0,1762	0,1854	0,1801
RISB	1	0,0239	0,0466	0,0837
	2	0,0352	0,0489	0,0907
	3	0,0412	0,0601	0,0957
	4	0,1457	0,0813	0,1032
	5	0,0952	0,0506	0,0709
	6	0,1643	0,1852	0,1770
RIV	1	0,0135	0,0264	0,0057
	2	0,0222	0,0148	0,0139
	3	0,0148	0,0118	0,0088
	4	0,0449	0,0151	0,0129
	5	0,0555	0,0191	0,0149
	6	0,0637	0,0091	0,0330

MLN

	Scenario	1 month	3 months	6 months
RMISE	1	0,0468	0,0689	0,0854
	2	0,0079	0,0103	0,0182
	3	0,0438	0,0559	0,0397
	4	0,2090	0,1092	0,1077
	5	0,0156	0,0228	0,0420
	6	0,1860	0,1889	0,1270
RISB	1	0,0446	0,0672	0,0844
	2	0,0068	0,0102	0,0182
	3	0,0428	0,0553	0,0393
	4	0,2067	0,1035	0,1075
	5	0,0127	0,0202	0,0411
	6	0,1843	0,1881	0,1270
RIV	1	0,0142	0,0153	0,0127
	2	0,0040	0,0013	0,0005
	3	0,0089	0,0083	0,0054
	4	0,0310	0,0347	0,0069
	5	0,0090	0,0105	0,0089
	6	0,0252	0,0165	0,0010

Figure 11-7: RMISE, RISB and RIV for DFCH and MLN methods.

With v weighting

SML (λ=0,9)

	Scenario	1 month	3 months	6 months
RMISE	1	0,0613	0,0801	0,0957
	2	0,0102	0,0074	0,0085
	3	0,0578	0,0787	0,0957
	4	0,2641	0,3000	0,2947
	5	0,1185	0,1561	0,1761
	6	0,2007	0,2507	0,2698
RISB	1	0,0605	0,0800	0,0957
	2	0,0032	0,0055	0,0079
	3	0,0570	0,0785	0,0957
	4	0,2640	0,3000	0,2947
	5	0,1182	0,1560	0,1761
	6	0,2005	0,2506	0,2698
RIV	1	0,0142	0,0153	0,0127
	2	0,0040	0,0013	0,0005
	3	0,0089	0,0083	0,0054
	4	0,0310	0,0347	0,0069
	5	0,0090	0,0105	0,0089
	6	0,0252	0,0165	0,0010

SML (λ that minimizes RMISE)

	Scenario	1 month	3 months	6 months
RMISE	1	0,0613	0,0802	0,0957
	2	0,0106	0,0072	0,0087
	3	0,0578	0,0787	0,0957
	4	0,2641	0,3001	0,2949
	5	0,1183	0,1556	0,1759
	6	0,2008	0,2506	0,2697
RISB	1	0,0605	0,0800	0,0957
	2	0,0033	0,0052	0,0080
	3	0,0571	0,0785	0,0957
	4	0,2640	0,3000	0,2949
	5	0,1180	0,1556	0,1759
	6	0,2006	0,2506	0,2697
RIV	1	0,0099	0,0048	0,0032
	2	0,0100	0,0049	0,0033
	3	0,0094	0,0048	0,0032
	4	0,0084	0,0036	0,0019
	5	0,0085	0,0032	0,0017
	6	0,0082	0,0030	0,0016

Without v weighting

SML (smooth=0,9)

	Scenario	1 month	3 months	6 months
RMISE	1	0,0631	0,0828	0,1255
	2	0,0137	0,0091	0,0112
	3	0,0587	0,0795	0,0967
	4	0,2696	0,3028	0,2932
	5	0,1240	0,1612	0,1813
	6	0,2035	0,2549	0,2750
RISB	1	0,0613	0,0822	0,1126
	2	0,0029	0,0061	0,0097
	3	0,0574	0,0793	0,0966
	4	0,2688	0,3025	0,2931
	5	0,1236	0,1612	0,1813
	6	0,2033	0,2548	0,2750
RIV	1	0,0149	0,0100	0,0554
	2	0,0134	0,0068	0,0056
	3	0,0121	0,0061	0,0044
	4	0,0204	0,0122	0,0057
	5	0,0101	0,0036	0,0018
	6	0,0090	0,0032	0,0017

SML (min RMISE)

	Scenario	1 month	3 months	6 months
RMISE	1	0,0634	0,0831	0,1243
	2	0,0134	0,0088	0,0115
	3	0,0587	0,0796	0,0968
	4	0,2691	0,3024	0,2930
	5	0,1229	0,1610	0,1805
	6	0,2036	0,2542	0,2741
RISB	1	0,0617	0,0824	0,1116
	2	0,0033	0,0058	0,0099
	3	0,0574	0,0793	0,0967
	4	0,2684	0,3021	0,2930
	5	0,1225	0,1610	0,1805
	6	0,2034	0,2541	0,2741
RIV	1	0,0149	0,0103	0,0548
	2	0,0130	0,0067	0,0058
	3	0,0124	0,0064	0,0046
	4	0,0204	0,0116	0,0056
	5	0,0102	0,0036	0,0017
	6	0,0090	0,0033	0,0017

Figure 11-8: RMISE, RISB and RIV for the SML method under 4 scenes: with or without v weighting and for each weighting approach using a smoothing parameter λ that minimizes RMISE or a smoothing parameter λ with a value of 0,9.

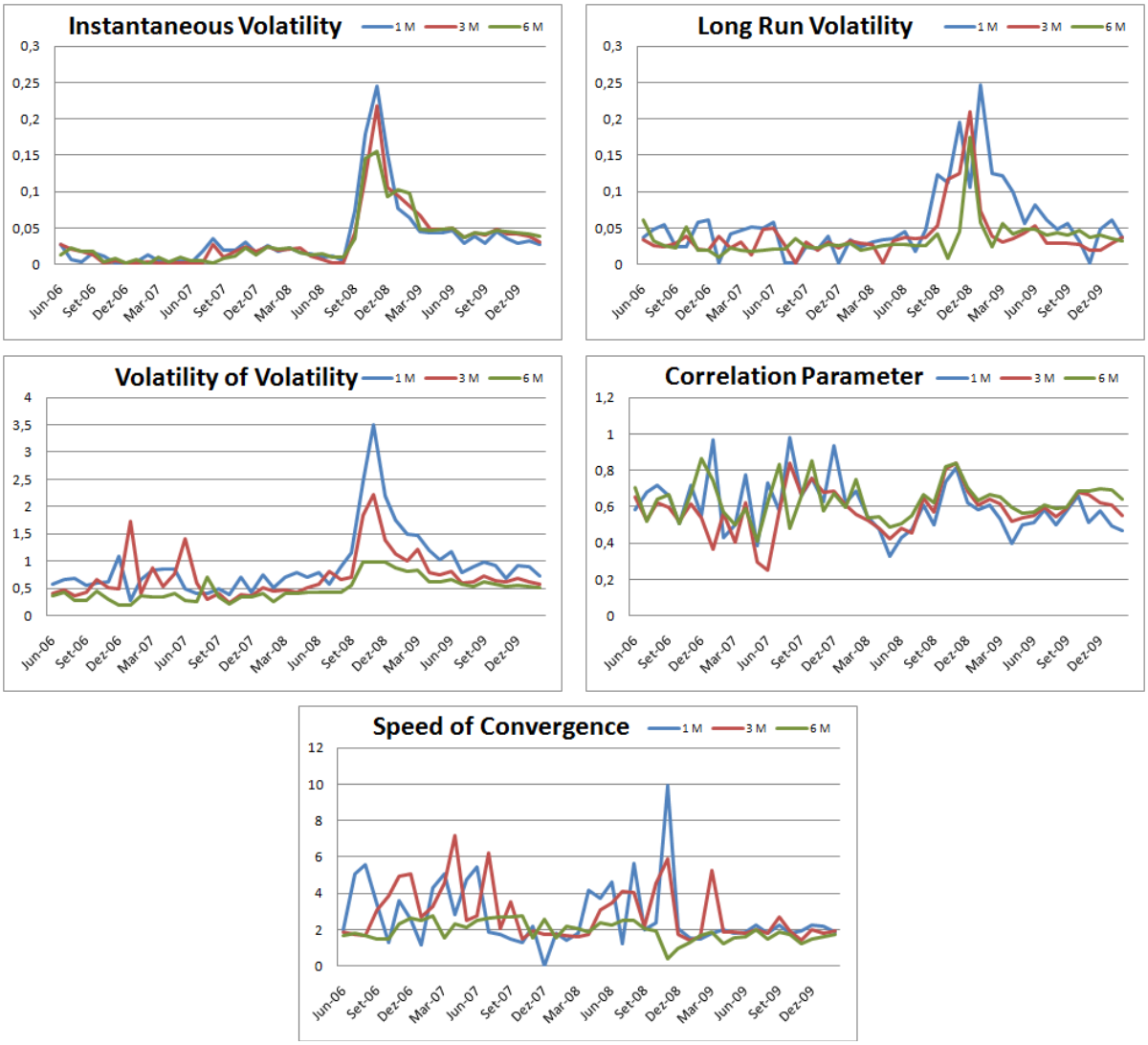


Figure 11-9: Heston model parameters obtained through calibration between June 2006 and February 2010

Figure 11-10: Brazil GDP

Figure 11-11: USD GDP

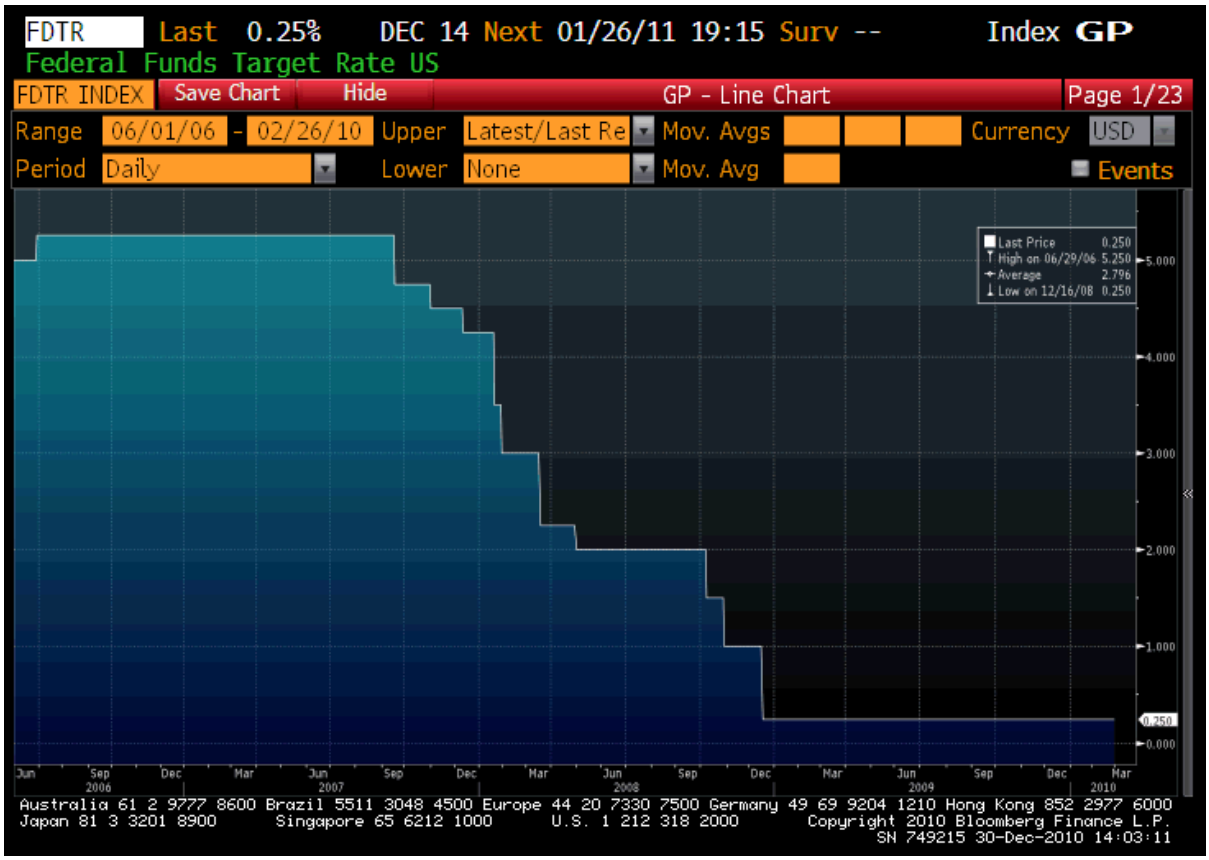


Figure 11-12: FED Funds target rate

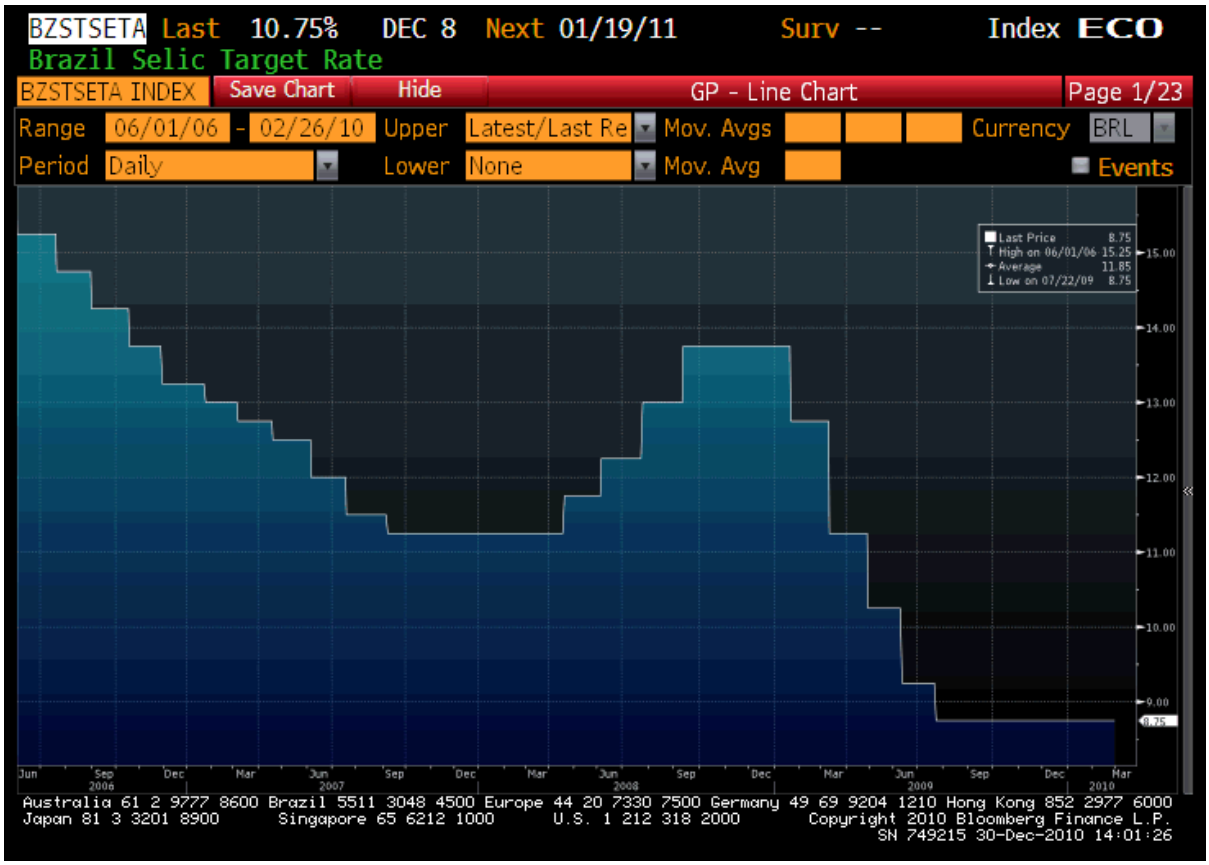


Figure 11-13: Brazil Selic Target Rate

Chapter 12

Matlab Codes

12.1 Heston model Codes

12.1.1 Generate Cooper Scenarios

To generate the "true" RNDs from the Cooper Scenarios and the Heston theoretical options pseudo-price we used as input the average strike prices, spot prices and interest rates for the period between June 1996 and February 2010.

```
function Test_HestonSPD()
%PURPOSE: runs Heston's formula yielding the risk neutral densities for
%the scenarios purpose in Cooper (1996) with the average strike prices for
%the period between June 2006 and February 2010
clc; clear;
%load the average strike prices, the average spot USDBRL, the average
%interest rates and the volatility for the period between June 2006 and
%February 2010
load (strcat('STRIKE_MEDIO','.','mat'),'mean_strike_1');
load (strcat('STRIKE_MEDIO','.','mat'),'mean_strike_2');
load (strcat('STRIKE_MEDIO','.','mat'),'mean_strike_3');
```

```
load (strcat('STRIKE_MEDIO','.','mat'),'mean_strike_4');
mean_strike = [mean_strike_1 mean_strike_2 mean_strike_3 mean_strike_4];
load (strcat('STRIKE_MEDIO','.','mat'),'mean_spot_1');
load (strcat('STRIKE_MEDIO','.','mat'),'mean_spot_2');
load (strcat('STRIKE_MEDIO','.','mat'),'mean_spot_3');
load (strcat('STRIKE_MEDIO','.','mat'),'mean_spot_4');
mean_spot = [mean_spot_1 mean_spot_2 mean_spot_3 mean_spot_4];
load (strcat('STRIKE_MEDIO','.','mat'),'var_spot_1');
load (strcat('STRIKE_MEDIO','.','mat'),'var_spot_2');
load (strcat('STRIKE_MEDIO','.','mat'),'var_spot_3');
load (strcat('STRIKE_MEDIO','.','mat'),'var_spot_4');
var_spot = [var_spot_1 var_spot_2 var_spot_3 var_spot_4];
load (strcat('STRIKE_MEDIO','.','mat'),'mean_rbrl_1');
load (strcat('STRIKE_MEDIO','.','mat'),'mean_rbrl_2');
load (strcat('STRIKE_MEDIO','.','mat'),'mean_rbrl_3');
load (strcat('STRIKE_MEDIO','.','mat'),'mean_rbrl_4');
mean_rbrl = [mean_rbrl_1 mean_rbrl_2 mean_rbrl_3 mean_rbrl_4];
load (strcat('STRIKE_MEDIO','.','mat'),'mean_r_1');
load (strcat('STRIKE_MEDIO','.','mat'),'mean_r_2');
load (strcat('STRIKE_MEDIO','.','mat'),'mean_r_3');
load (strcat('STRIKE_MEDIO','.','mat'),'mean_r_4');
load (strcat('STRIKE_MEDIO','.','mat'),'Tempo');
mean_r = [mean_r_1 mean_r_2 mean_r_3 mean_r_4];
for j=1:1 % estimate the price density for 1 month term
rbrl= mean_rbrl(j); %SICOR brazilian interest rate
r = mean_r(j); %Libor USD interest
tau=Tempo(j); %time to maturity
St=mean_spot(j); %spot USDBRL
```

```
vt=var_spot(j); % volatility at time 0
kap=2; %speed by which volatility returns to its long run average
SBTv= mean_strike(:,j); %grid with the selected strike prices
lda=0;
NSBTv=size(SBTv,1); %number of strike prices
for l=1:NSBTv
%Low volatility scenarios
 %estimate call prices for a scenario with negative skewness
SBT=SBTv(l);
th=0.01; %long run volatility
sig=0.1; %standard deviation of the volatility
rho=-0.9; %correlation parameter
dens1_low_neg(l) = (1/(2.*pi)).*quadgk(@p3,0,200);
P1(l) = 0.5 + (1/pi).*quadgk(@p1,0,200);
P2(l) = 0.5 + (1/pi).*quadgk(@p2,0,200);
Call_heston_low_neg (l) = St.*exp(-r.*tau).*P1(l) - SBT.*exp(-rbrl.*tau).*P2(l);

%estimate call prices for a scenario with no skewness
rho=0;
dens2_low(l) = (1/(2.*pi)).*quadgk(@p3,0,200);
P1(l) = 0.5 + (1/pi).*quadgk(@p1,0,200);
P2(l) = 0.5 + (1/pi).*quadgk(@p2,0,200);
Call_heston_low (l) = St.*exp(-r.*tau).*P1(l) - SBT.*exp(-rbrl.*tau).*P2(l);
%estimate call prices for a scenario with positive skewness
rho=0.9;
dens3_low_pos(l) = (1/(2.*pi)).*quadgk(@p3,0,200);
P1(l) = 0.5 + (1/pi).*quadgk(@p1,0,200);
P2(l) = 0.5 + (1/pi).*quadgk(@p2,0,200);
```

```
Call_heston_low_pos (l) = St.*exp(-r.*tau).*P1(l) - SBT.*exp(-rbrl.*tau).*P2(l);
%High volatility
%estimate call prices for a scenario with negative skewness
SBT=SBTv(l);
th=0.09; %long run volatility
sig=0.4; %devio padrao da volatilidade
rho=-0.9;
dens1_high_neg(l) = (1/(2.*pi)).*quadgk(@p3,0,200);
P1(l) = 0.5 + (1/pi).*quadgk(@p1,0,200);
P2(l) = 0.5 + (1/pi).*quadgk(@p2,0,200);
Call_heston_high_neg (l) = St.*P1(l) - SBT.*exp(-rbrl.*tau).*P2(l);
%estimate call prices for a scenario with no skewness
rho=0;
dens2_high(l) = (1/(2.*pi)).*quadgk(@p3,0,200);
P1(l) = 0.5 + (1/pi).*quadgk(@p1,0,200);
P2(l) = 0.5 + (1/pi).*quadgk(@p2,0,200);
Call_heston_high (l) = St.*P1(l) - SBT.*exp(-rbrl.*tau).*P2(l);
%estimate call prices for a scenario with positive skewness
rho=0.9;
dens3_high_pos(l) = (1/(2.*pi)).*quadgk(@p3,0,200);
P1(l) = 0.5 + (1/pi).*quadgk(@p1,0,200);
P2(l) = 0.5 + (1/pi).*quadgk(@p2,0,200);
Call_heston_high_pos (l) = St.*P1(l) - SBT.*exp(-rbrl.*tau).*P2(l);
end
%estimate a RND for each scenario (creation of a grid with 5000 points)
SBTv= linspace(1.5,2.4,5000)';
lda=0;
NSBTv=size(SBTv,1);
```

```
for l=1:NSBTv
%Low volatility
%estimate RND for a scenario with negative skewness
SBT=SBTv(l);
th=0.01;
sig=0.1;
rho=-0.9;
P1(l) = 0.5 + (1/pi).*quadgk(@p1,0,200);
P2(l) = 0.5 + (1/pi).*quadgk(@p2,0,200);
Call_heston_low_neg1 (l) = St.*exp(-r.*tau).*P1(l) - SBT.*exp(-rbrl.*tau).*P2(l);
 if l>=3
second_low_neg(l) = ((Call_heston_low_neg1 (l)- 2*Call_heston_low_neg1 (l-1)
+ Call_heston_low_neg1 (l-2))./((SBTv(l)-SBTv(l-1))^2))*exp((rbrl-r)*tau);
end
%estimate call prices for a scenario with no skewness
rho=0;
P1(l) = 0.5 + (1/pi).*quadgk(@p1,0,200);
P2(l) = 0.5 + (1/pi).*quadgk(@p2,0,200);
Call_heston_low1 (l) = St.*exp(-r.*tau).*P1(l) - SBT.*exp(-rbrl.*tau).*P2(l);

if l>=3
second_low(l) = ((Call_heston_low1 (l)- 2*Call_heston_low1 (l-1)
+ Call_heston_low1 (l-2))./((SBTv(l)-SBTv(l-1))^2))*exp((rbrl-r)*tau);
end
%estimate call prices for a scenario with positive skewness
rho=0.9;
P1(l) = 0.5 + (1/pi).*quadgk(@p1,0,200);
```

```
P2(l) = 0.5 + (1/pi).*quadgk(@p2,0,200);
Call_heston_low_pos1 (l) = St.*exp(-r.*tau).*P1(l) - SBT.*exp(-rbrl.*tau).*P2(l);
if l>=3
second_low_pos(l) = ((Call_heston_low_pos1 (l)- 2*Call_heston_low_pos1 (l-1)
+ Call_heston_low_pos1 (l-2))./((SBTv(l)-SBTv(l-1))^2))*exp((rbrl-r)*tau);
end
%High volatility
%estimate call prices for a scenario with negative skewness
SBT=SBTv(l);
th=0.09;
sig=0.4;
rho=-0.9;
P1(l) = 0.5 + (1/pi).*quadgk(@p1,0,200);
P2(l) = 0.5 + (1/pi).*quadgk(@p2,0,200);
Call_heston_high_neg1 (l) = St.*P1(l) - SBT.*exp(-rbrl.*tau).*P2(l);
if l>=3
second_high_neg(l) = ((Call_heston_high_neg1 (l)- 2*Call_heston_high_neg1 (l-1)
+ Call_heston_high_neg1 (l-2))./((SBTv(l)-SBTv(l-1))^2))*exp((rbrl-r)*tau);
end
%estimate call prices for a scenario with no skewness
rho=0;
P1(l) = 0.5 + (1/pi).*quadgk(@p1,0,200);
P2(l) = 0.5 + (1/pi).*quadgk(@p2,0,200);
Call_heston_high1 (l) = St.*P1(l) - SBT.*exp(-rbrl.*tau).*P2(l);
if l>=3
second_high(l) = ((Call_heston_high1 (l)- 2*Call_heston_high1 (l-1)
+ Call_heston_high1 (l-2))./((SBTv(l)-SBTv(l-1))^2))*exp((rbrl-r)*tau);
end
```

```
%estimate call prices for a scenario with positive skewness
rho=0.9;
P1(l) = 0.5 + (1/pi).*quadgk(@p1,0,200);
P2(l) = 0.5 + (1/pi).*quadgk(@p2,0,200);
Call_heston_high_pos1 (l) = St.*P1(l) - SBT.*exp(-rbrl.*tau).*P2(l);
if l>=3
second_high_pos(l) = ((Call_heston_high_pos1 (l)- 2*Call_heston_high_pos1 (l-1)
+ Call_heston_high_pos1 (l-2))./((SBTv(l)-SBTv(l-1))^2))*exp((rbrl-r)*tau);
end
end
end

figure(2)
plot(SBTv',second_low_neg);

differencial = diff(SBTv);
%estimation of Expected Value
Em_heston_low_neg = sum(second_low_neg.*differencial(1).*SBTv');
Em_heston_low = sum(second_low.*differencial(1).*SBTv');
Em_heston_low_pos = sum(second_low_pos.*differencial(1).*SBTv');
Em_heston_high_neg=sum(second_high_neg.*differencial(1).*SBTv');
Em_heston_high =sum(second_high.*differencial(1).*SBTv');
Em_heston_high_pos=sum(second_high_pos.*differencial(1).*SBTv');
%estimation of implied variance
Var_heston_low_neg = real(sum(second_low_neg.*((SBTv'-Em_heston_low_neg).^2)
.*differencial(1));
Var_heston_low = sum(second_low.*((SBTv'-Em_heston_low).^2)
.*differencial(1));
```

```
Var_heston_low_pos = sum(second_low_pos.*((SBTv'-Em_heston_low_pos).^2)
.*differencial(1));
Var_heston_high_neg=sum(second_high_neg.*((SBTv'-Em_heston_high_neg).^2)
.*differencial(1));
Var_heston_high =sum(second_high.*((SBTv'-Em_heston_high).^2)
.*differencial(1));
Var_heston_high_pos=sum(second_high_pos.*((SBTv'-Em_heston_high_pos).^2)
.*differencial(1));
%estimation of implied Skewness
Skew_heston_low_neg=(sum(second_low_neg.*((SBTv'-Em_heston_low_neg).^3)
.*differencial(1))./(Var_heston_low_neg.^(3/2));
Skew_heston_low =(sum(second_low.*((SBTv'-Em_heston_low).^3)
.*differencial(1))./(Var_heston_low.^(3/2);
Skew_heston_low_pos=(sum(second_low_pos.*((SBTv'-Em_heston_low_pos).^3)
.*differencial(1))./(Var_heston_low_pos.^(3/2));
Skew_heston_high_neg=(sum(second_high_neg.*((SBTv'-Em_heston_high_neg).^3)
.*differencial(1))./(Var_heston_high_neg.^(3/2));
Skew_heston_high =(sum(second_high.*((SBTv'-Em_heston_high).^3)
.*differencial(1))./(Var_heston_high.^(3/2));
Skew_heston_high_pos=(sum(second_high_pos.*((SBTv'-Em_heston_high_pos).^3)
.*differencial(1))./(Var_heston_high_pos.^(3/2));
%estimation of implied Kurtosis
Kurtosis_heston_low_neg=(sum(second_low_neg.*((SBTv'-Em_heston_low_neg).^4)
.*differencial(1))./(Var_heston_low_neg.^2);
Kurtosis_heston_low =(sum(second_low.*((SBTv'-Em_heston_low).^4)
.*differencial(1))./(Var_heston_low.^2;
Kurtosis_heston_low_pos=(sum(second_low_pos.*((SBTv'-Em_heston_low_pos).^4)
.*differencial(1))./(Var_heston_low_pos.^2);
```

```
Kurtosis_heston_high_neg=(sum(second_high_neg.*((SBTv'-Em_heston_high_neg).^4)
.*differencial(1))./(Var_heston_high_neg.^2);
Kurtosis_heston_high =(sum(second_high.*((SBTv'-Em_heston_high).^4)
.*differencial(1))./(Var_heston_high.^2);
Kurtosis_heston_high_pos=(sum(second_high_pos.*((SBTv'-Em_heston_high_pos).^4)
.*differencial(1))./(Var_heston_high_pos.^2);


function y= p3(phi)
 y= CF_SVj(log(St),vt,tau,(rbrl-r),kap*th,-0.5,kap+lda,rho,sig,phi,SBT);
end

function y= p1(phi)
 y= CF_SVj_forcallprice(log(St),vt,tau,(rbrl-r),kap*th,0.5,kap+lda-rho.*sig,rho,sig,phi,SBT);
end

function y= p2(phi)
y= CF_SVj_forcallprice(log(St),vt,tau,(rbrl-r),kap*th,-0.5,kap+lda,rho,sig,phi,SBT);
end

save (strcat('Heston_mat1','.mat'));
end
```

12.1.2 USDBRL Heston parameters

The Heston parameters estimated for the USDBRL options market for the end of month dates between June 2006 and February 2010 were calculated through the code bellow:

```
function heston
% Calibrate the Heston model parameters for the end of the month call
% prices between June 2006 and February 2010
```

```
clear; clc;
%open excel file with the selected dates in the first column (in this case 45 dates)
D=xlsread('data.xls');
Dia = D(1:45,1);
Mes = D(1:45,2);
Ano = D(1:45,3);
% Open tables with market information for 45 dates
for w=1:45;
load (strcat('USDBRL_',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.','mat')
,'AllInfo');
load (strcat('USDBRL_',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.','mat'),'N');
load (strcat('USDBRL_',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.','mat'),'M');
load (strcat('USDBRL_',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.','mat'),
'NbMat');
load (strcat('USDBRL_',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.','mat')
,'NbStrik');
load (strcat('USDBRL_',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.','mat')
,'NbCall');
load (strcat('USDBRL_',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.','mat')
,'NbPut');
% Variables
NbStrik=N/2; %number of stikes
NbMat =M; %number of considered maturities
NbCall =NbMat*NbStrik; %number of calls with different strikes
NbPut =NbMat*NbStrik; %number of puts with different strikes
NbOpt = NbCall+NbPut; % number of calls and puts out-the money
Res=[]; %here store the densities
ParamM=[]; % parameters to store
```

```
Skewness=[]; % skewness to store
Kurtosis =[]; % kurtosis to store
Mean =[]; % expected value to store
Std =[]; % standard deviation to store
for i=1:1 %RNDs with 1 month term

S0 = AllInfo(1,1); % spot USDBRL
KC = AllInfo(1+NbStrik*(i-1):NbStrik*i,9); %out the money strikes for call options
KP = AllInfo(NbCall+1+NbStrik*(i-1):NbCall+NbStrik*i,9); %out the money strikes for
put options
K = [KC; KP];
CPi= AllInfo(1+NbStrik*(i-1):NbStrik*2*i,3);
rbrl = AllInfo(1+NbStrik*(i-1),8); % brazilian interest rate (domestic)
T = AllInfo(1+NbStrik*(i-1),5); %time to maturity
C = AllInfo(1+NbStrik*(i-1):NbStrik*i,10); %call option
P = AllInfo(NbCall+1+NbStrik*(i-1):NbCall+NbStrik*i,10); %put option
r = AllInfo(1+NbStrik*(i-1),4); % Libor interest rate
Cs = P + S0*exp(-r*T) - KP.*exp(-rbrl*T); %convert put prices into call prices through
put-call parity
Call = [C; Cs];

b0=[0.04 0.6 0.5 2 0.03]; b0=b0';

lb=[ 0.001; 0.001; -1; 0; 0.001]; % lower bounds for parameters
ub=[ 5; 10; 1; 10; 5]; %upper bounds for parameters

%nonlinear least-squares optimization
```

```
options=optimset('Algorithm','trust-region-reflective','Diagnostics','on','Display','iter','TolX',1e-
6,'TolFun',
1e-6,'MaxFunEvals',150,'MaxIter',5000);
[beta,resnorm,residual,exitflag,output] = lsqnonlin(@(b)MD_Obj1(b,S0,K,rbrl,T,Call,r),b0,
lb,ub,options);
ParamM=[ ParamM; beta' ];

th = beta(1);
sig = beta(2);
rho = beta(3);
kap = beta(4);
vt = beta(5);
lda = 0;
tau= T;

z=linspace(1.6,2.8,5000); %grid definition for the RND
diferencial = diff(z);
NSBTv = size(z,2);

for l=1:NSBTv

SBT=z(l);
P1(l) = 0.5 + (1/pi).*quadgk(@p1,0,500);
P2(l) = 0.5 + (1/pi).*quadgk(@p2,0,500);
Call(l) = S0.*exp(-r.*tau).*P1(l) - SBT.*exp(-rbrl.*tau).*P2(l);

if l>=3
f(l) = ((Call(l)- 2*Call(l-1) + Call(l-2))./((z(l)-z(l-1))^2))*exp((rbrl-r)*tau);
```

```
end
end

Em = sum(z.*f.*diferencial(1));
Var = sum((z-Em).^2.*f.*diferencial(1));
Skewness = sum(((z-Em).^3).*f.*diferencial(1))./(Var.^(3/2));
Kurtosis = sum(((z-Em).^4).*f.*diferencial(1))./(Var.^2);

SBTv = size(K,1);

for l=1:SBTv
SBT=K(l);
P1_K(l) = 0.5 + (1/pi).*quadgk(@p1,0,500);
P2_K(l) = 0.5 + (1/pi).*quadgk(@p2,0,500);
Call_K(l) = S0.*exp(-r.*tau).*P1_K(l) - SBT.*exp(-rbrl.*tau).*P2_K(l);
end

end
figure(w)
plot(z,f)
title('RND')
xlabel('STRIKE PRICES')
disp('The parameters are ');
ParamM
save (strcat('Heston_mat1_',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.mat'));
end
%*********************************************************
%Objective function
```

```
function y=MD_Obj1(b,S0,K,rbrl,T,Call,r)
th = b(1);
sig = b(2);
rho = b(3);
kap = b(4);
vt = b(5);
lda = 0;
NbStrik=12;
tau=T;
Cemp=Call(1:NbStrik);
K=K(1:NbStrik);
S0=S0;
for l=1:NbStrik
SBT = K(l);
P1(l) = 0.5 + (1/pi).*quadgk(@p1,0,500);
P2(l) = 0.5 + (1/pi).*quadgk(@p2,0,500);
Cth(l) = S0.*exp(-r.*tau).*P1(l) - SBT.*exp(-rbrl.*tau).*P2(l);

end
y = Cemp - Cth';
end
%************************************************************
function y= p3(phi)
y= CF_SVj(log(S0),vt,tau,(rbrl-r),kap*th,-0.5,kap+lda,rho,sig,phi,SBT);
end
function y= p1(phi)
y= CF_SVj_forcallprice(log(S0),vt,tau,(rbrl-r),kap*th,0.5,kap+lda-rho.*sig,rho,sig,phi,SBT);
end
```

```
function y= p2(phi)
y= CF_SVj_forcallprice(log(S0),vt,tau,(rbrl-r),kap*th,-0.5,kap+lda,rho,sig,phi,SBT);
end
end
```

12.2 Hypergeometric model codes

12.2.1 DFCH Monte Carlo simulations for USDBRL Heston Scenarios

To test the DFCH robustness in capturing the "true" RNDs generated by Heston model representing USDBRL Low Volatility Dates (between October 2006 and March 2007) and High Volatility Dates (between September 2008 and February 2009) we proceed the Monte Carlo simulations through the following matlab code (To test the Cooper Scenarios we change the input parameters):

```
function DFCH_CAL()
clc;clear;
%open excel file with the selected dates in the first column (in this case
%45 dates)
D=xlsread('data.xls');
Dia = D(1:45,1);
Mes = D(1:45,2);
Ano = D(1:45,3);
% Open tables with market information for 45 dates
for w=16:16; %simulations for the DFCH model using the Heston PARAMETERS for 28-
11-2008 date
load (strcat('USDBRL_',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.','mat'));
```

```
load (strcat('Heston_mat1_',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.mat'),'f');
load (strcat('Heston_mat1_',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.mat'),
'Call_K');
load (strcat('Heston_mat1_',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.mat'),'K');
% Variables:
NbStrik=N/2; %number of stikes
NbMat =M; %number of considered maturities
NbCall =NbMat*NbStrik; %number of calls with different strikes
NbPut =NbMat*NbStrik; %number of puts with different strikes
NbOpt = NbCall+NbPut; % number of calls and puts out-the money
Res=[]; %here store the DFCH densities
REAL=[]; %here store the true density
ParamM=[]; % parameters to store
Skewness=[]; %skewness to store
Kurtosis =[];% kurtosis to store
Mean =[]; % expected value to store
Std =[]; % standard deviation to store
Ruido = [];
realdens = f';
ruido = -0.001*0.5 + (0.001*0.5-(-0.001*0.5)).*rand(500,1);
Ruido = [Ruido; ruido];
for l=1:500
for j=1:1 %%RNDs with 1 month term

S0 = AllInfo(1,1); % spot USDBRL
CPi= AllInfo(1+NbStrik*(j-1):NbStrik*2*j,3);
rbrl = AllInfo(1+NbStrik*(j-1),8); % brazilian interest rate (domestic)
r = AllInfo(1+NbStrik*(j-1),4); %libor interest rate
```

```
T = AllInfo(1+NbStrik*(j-1),5); %time to maturity
Call = Call_K';
Call = [Call+Ruido(l), zeros(12,1)];
Call = max(Call,[],2);
av = linspace(-38,0,38); %optimization along a grid of possible values for b2
%as initial points
m= linspace(0.001,0.1,100); %optimization along a grid of possible values for m1
%as initial points

GridRes=[];
for j=1:size(av,2)
for g=1:size(m,2)
b= [4; 6; (-1/(2*var(K)))-av(j); 1; -1/(2*var(K)); mean(K)-m(g);mean(K)];

lb=[ 0; 0; -100; 0; -100; 0; 0]; % lower bounds for parameters
ub=[ 10; 10; -0.1; 5; -0.1; 3; 3]; %upper bounds for parameters

options=optimset('Algorithm','trust-region-reflective','Diagnostics','on','Display','iter','TolX',1e-
6,'TolFun',
1e-6,'MaxFunEvals',150,'MaxIter',5000);

[beta,resnorm,residual,exitflag,output] = lsqnonlin(@(b)MD_Obj(b,S0,K,rbrl,T,Call,r),b,lb,ub,options)

GridRes=[GridRes; [av(j) m(g) beta' resnorm]];
end
end
[mi,miidx]=min(GridRes(:,end));
beta=GridRes(miidx,3:9);
```

```
ParamM=[ ParamM; beta' ];
Residual=[ Residual; residual ];
Resnorm = [Resnorm; resnorm];
Output = [Output; output];

z=linspace(1.5,3.3,5000); z=z'; % support for the RND (5000 points)
diferencial = diff(z);

a2 = beta(1);
a3 = beta(2);
b2 = beta(3);
b3 = beta(4);
b4 = beta(5);
m1 = beta(6);
m2 = beta(7);
a5 = -1/2;
a6 = 1/2;
b1 = 1+a2*b3;
a1 = ((S0/exp(-(rbrl-r)*T))-m2)./(((gamma(a3))/(gamma(a3-a2))).*((-b2)^(-a2)).*(m1-m2));
a4 = (1/(2*sqrt(-b4*pi)))*(1-a1*((-b2)^(-a2))*((gamma(a3))/(gamma(a3-a2))));
c2 = -1 + a4 * sqrt(-b4*pi);
c1 = -c2*m2;

f = [];
for n=1:size(z,1)

l1 = z(n)>m1;
```

```
f(n) = ((l1.*a1.*((z(n)-m1).^(b1-2)).*(b1*(b1-1)*chgm(a2,a3,b2*((z(n)-m1).^b3))+...
(a2/a3)*b2*b3*(2*b1+b3-1)*((z(n)-m1).^b3).*chgm(a2+1,a3+1,b2*((z(n)-m1).^b3))+...
+((a2*(a2+1))/((a3)*(a3+1)))*(b2^2)*(b3^2)*((z(n)-m1).^(2*b3)).*chgm(a2+2,a3+2,b2*((z(n)-
m1).^b3)))+...
+ 2*a4*(a5/a6)*b4.*(chgm(a5+1,a6+1,b4.*((z(n)-m2).^2))+2*((a5+1)/(a6+1))*b4.*((z(n)-
m2).^2)
.*chgm(a5+2,a6+2,b4*(((z(n)-m2).^2))))))*exp((rbrl-r)*T);
end

f = f';

Res = [Res f];
REAL= [REAL realdens];
M1 = [M1 m1];
l= size(Res,2);

Em(l) = sum(z.*f.*diferencial(1));%expected value
Var(l) = sum((z-Em(l)).^2.*f.*diferencial(1)); %implied standard deviation
Skewness(l) = sum(((z-Em(l)).^3).*f.*diferencial(1))./(Var(l).^(3/2)); %implied skewness
Kurtosis(l) = sum(((z-Em(l)).^4).*f.*diferencial(1))./(Var(l).^2); %implied kurtosis

figure(1)
plot(z,f)
hold on
end
end
end
```

```
RIV = sqrt(sum((mean(((Res-kron(mean(Res,2),ones(1,size(Res,2)))).^2),2))
.*diferencial(1))); %implied RIV
RISB = sqrt(sum(((mean(Res,2)-realdens).^2).*diferencial(1))); %implied RISB
min_RMISE = sqrt(sum((mean(((Res-REAL).^2),2)).*diferencial(1))); %implied RMISE
mean_mean = mean(Em);
var_mean = mean(Var);
skew_mean = mean(Skewness);
kurt_mean = mean(Kurtosis);
var_stat_var = var(Var);
skew_stat_var = var(Skewness);
kurt_stat_var = var(Kurtosis);
figure(2)
plot(z,Res)
title('chgmETRIC')
xlabel('STRIKE PRICES')
axis tight
MixRND=Res;
save (strcat('HYPER_Nov08_mat1','.mat'));
%*********************************************************
function y=MD_Obj(b,S0,K,rbrl,T,Call,r)
%
a2 = b(1);
a3 = b(2);
b2 = b(3);
b3 = b(4);
b4 = b(5);
m1 = b(6);
m2 = b(7);
```

```
a5 = -1/2;
a6 = 1/2;
a1 = ((S0/exp(-(rbrl-r)*T))-m2)./(((gamma(a3))/(gamma(a3-a2))).*((-b2)^(-a2)).*(m1-m2));
b1 = 1+a2*b3;
a4 = (1/(2*sqrt(-b4*pi)))*(1-a1*((-b2)^(-a2))*((gamma(a3))/(gamma(a3-a2))));
c2 = -1 + a4 * sqrt(-b4*pi);
c1 = -c2*m2;
NbStrik=6;
Cemp=Call(1:2*NbStrik);
K=K(1:2*NbStrik);
for x=1:size(K,1)
l1 = K(x)>m1;

Cth(x)= c1 + c2.*K(x) + l1.*a1.*((K(x)-m1).^b1).*chgm(a2,a3,b2*((K(x)-m1).^b3))
+ a4.*chgm(a5,a6,b4.*((K(x)-m2).^2));
end
Cth = Cth';
y =Cemp - Cth ; %calls
```

12.2.2 DFCH USDBRL parameters

The Heston parameters estimated for the USDBRL options market for the end of month dates between June 2006 and February 2010 were calculated through the code bellow:

```
function hypergeometric()
% Calibrate the DFCH model parameters for the end of the month call
% prices between June 2006 and February 2010
clc;clear;
%abrir tabela com as datas da amostra disponiveis (final do mês)
```

```
D=xlsread('data.xls');
Dia = D(1:45,1);
Mes = D(1:45,2);
Ano = D(1:45,3);
for w=1:45;
% Open tables with market information for 45 dates
load (strcat('USDBRL_',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.','mat'));
% Variables
NbStrik=N/2; %number of stikes
NbMat =M; %number of considered maturities
NbCall =NbMat*NbStrik; %number of calls with different strikes
NbPut =NbMat*NbStrik; %number of puts with different strikes
NbOpt = NbCall+NbPut; % number of calls and puts out-the money
Res=[]; %here store the DFCH densities
ParamM=[]; % parameters to store
Skewness=[]; %skewness to store
Kurtosis =[]; %kurtosis to store
Mean =[]; %expected value to store
Std =[]; %standard deviation to store
Residual =[];
Resnorm =[];
Output =[];
for i=1:1 %RNDs with 1 month term

S0 = AllInfo(1,1); % spot USDBRL
KC = AllInfo(1+NbStrik*(i-1):NbStrik*i,9); %Call Strike prices
KP = AllInfo(NbCall+1+NbStrik*(i-1):NbCall+NbStrik*i,9); %Put Strike Prices
K = [KC; KP]; %Strike prices
```

```
CPi= AllInfo(1+NbStrik*(i-1):NbStrik*2*i,3);
rbrl = AllInfo(1+NbStrik*(i-1),8); % brazilian interest rate (domestic)
r = AllInfo(1+NbStrik*(i-1),4); %libor interest rate
T = AllInfo(1+NbStrik*(i-1),5); %time to maturity
C = AllInfo(1+NbStrik*(i-1):NbStrik*i,10); %call option
P = AllInfo(NbCall+1+NbStrik*(i-1):NbCall+NbStrik*i,10); %put option
Cs = P + S0*exp(-r*T) - KP.*exp(-rbrl*T); %convert puts prices (out-the money) into call prices (in-the-money) through put-call parity
Call = [C; Cs]; % calls out-the-money and calls in the money

av = linspace(-38,0,38); %optimization along a grid of possible values for b2 as initial points
m= linspace(0.001,0.1,100); %optimization along a grid of possible values for m1 as initial points
GridRes=[];
for j=1:size(av,2)
for g=1:size(m,2)

b= [4; 6; (-1/(2*var(K)))-av(j); 1; -1/(2*var(K)); mean(K)-m(g);mean(K)];

lb=[ 0; 0; -inf; 0; -inf; 0; 0]; %lower bounds for parameters
ub=[ inf; inf; -0.00001; inf; -0.00001; inf; inf]; %upper bounds for parameters
options=optimset('Algorithm','trust-region-reflective','Diagnostics','on','Display','iter','TolX',1e-6,'TolFun',1e-6,'MaxFunEvals',150,'MaxIter',5000);

[beta,resnorm,residual,exitflag,output] = lsqnonlin(@(b)MD_Obj(b,S0,K,rbrl,T,Call,r),b,lb,ub,options)
GridRes=[GridRes; [av(j) m(g) beta' resnorm]];
end
end
```

```
[mi,miidx]=min(GridRes(:,end));
beta=GridRes(miidx,3:9);

ParamM=[ ParamM; beta' ];
Residual=[ Residual; residual ];
Resnorm = [Resnorm; resnorm];
Output = [Output; output];


z=linspace(0.7,5.5,5000); z=z'; % support for the RND (5000 points)
diferencial = diff(z);

a2 = beta(1);
a3 = beta(2);
b2 = beta(3);
b3 = beta(4);
b4 = beta(5);
m1 = beta(6);
m2 = beta(7);
a5 = -1/2;
a6 = 1/2;
b1 = 1+a2*b3;
a1 = ((S0/exp(-(rbrl-r)*T))-m2)./(((gamma(a3))/(gamma(a3-a2))).*((-b2)^(-a2)).*(m1-m2));
a4 = (1/(2*sqrt(-b4*pi)))*(1-a1*((-b2)^(-a2))*((gamma(a3))/(gamma(a3-a2))));
c2 = -1 + a4 * sqrt(-b4*pi);
c1 = -c2*m2;

f = [];
```

```
for n=1:size(z,1)

l1 = z(n)>m1;

f(n) = ((l1.*a1.*((z(n)-m1).^(b1-2)).*(b1*(b1-1)*chgm(a2,a3,b2*((z(n)-m1).^b3))+...
(a2/a3)*b2*b3*(2*b1+b3-1)*((z(n)-m1).^b3).*chgm(a2+1,a3+1,b2*((z(n)-m1).^b3))+...
+((a2*(a2+1))/((a3)*(a3+1)))*(b2^2)*(b3^2)*((z(n)-m1).^(2*b3)).*chgm(a2+2,a3+2,b2*((z(n)-
m1).^b3)))+...
+ 2*a4*(a5/a6)*b4.*(chgm(a5+1,a6+1,b4.*((z(n)-m2).^2))+2*((a5+1)/(a6+1))*b4.*((z(n)-
m2).^2).*chgm(a5+2,a6+2,b4*(((z(n)-m2).^2))))))*exp((rbrl-r)*T);
end

f = f';

Res = [ Res f ];
Em_f = sum(z.*f.*diferencial(1));
Var_f = sum((z-Em_f).^2.*f.*diferencial(1));
Skewness_f = sum(((z-Em_f).^3).*f.*diferencial(1))./(Var_f.^(3/2));
Kurtosis_f = sum(((z-Em_f).^4).*f.*diferencial(1))./(Var_f.^2);

Skewness = [Skewness_f Skewness ];
Kurtosis =[Kurtosis_f Kurtosis ];
Mean = [Em_f Mean];
Std = [sqrt(Var_f) Std];
end

figure(w)
plot(z,Res)
```

```
title('hypergeomETRIC')
xlabel('STRIKE PRICES')
axis tight
MixRND=Res;
disp('The parameters are ');
ParamM
print ('-djpeg',strcat('HYPER',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w))))
Skew(w,:) = [Skewness];
Kurt(w,:) = [Kurtosis];
MEAN(w,:) = [Mean];
STD (w,:) = [Std];
save (strcat('HYPER_mat3',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.mat'));
end
figure(100)
plot(1:45,Skew)
title('SKEWNESS')
xlabel('Data');
print ('-djpeg','SKEWNESS')
figure(101)
plot(1:45,Kurt)
title('Kurtosis')
xlabel('Data');
print ('-djpeg','Kurtosis')
figure(102)
plot(1:45,MEAN)
title('Mean')
xlabel('Data');
print ('-djpeg','MEAN')
```

```
figure(103)
plot(1:45,STD)
title('Stdv')
xlabel('Data');
print ('-djpeg','STDV')
%*********************************************************
function y=MD_Obj(b,S0,K,rbrl,T,Call,r)
%
a2 = b(1);
a3 = b(2);
b2 = b(3);
b3 = b(4);
b4 = b(5);
m1 = b(6);
m2 = b(7);
a5 = -1/2;
a6 = 1/2;
a1 = ((S0/exp(-(rbrl-r)*T))-m2)./(((gamma(a3))/(gamma(a3-a2))).*((-b2)^(-a2)).*(m1-m2));
b1 = 1+a2*b3;
a4 = (1/(2*sqrt(-b4*pi)))*(1-a1*((-b2)^(-a2))*((gamma(a3))/(gamma(a3-a2))));
c2 = -1 + a4 * sqrt(-b4*pi);
c1 = -c2*m2;
NbStrik=6;
Cemp=Call(1:2*NbStrik);
K=K(1:2*NbStrik);
for x=1:size(K,1)
l1 = K(x)>m1;
```

```
Cth(x)= c1 + c2.*K(x) + l1.*a1.*((K(x)-m1).^b1).*chgm(a2,a3,b2*((K(x)-m1).^b3)) +
a4.*chgm(a5,a6,b4.*((K(x)-m2).^2));
end
Cth = Cth';
y =Cemp - Cth ; %calls
```

12.3 Spline model codes

12.3.1 SML USDBRL parameters

The USDBRL RNDs estimated through SML method for the end of month dates between June 2006 and February 2010 were calculated through the code bellow:

```
function Spline()
clc;clear;
%open an excel file with the selected dates in the first column (in this case 45
%dates)
D=xlsread('data.xls');
Dia = D(1:45,1);
Mes = D(1:45,2);
Ano = D(1:45,3);
for w=1:45; % Open tables with market information for 45 dates
load (strcat('USDBRL_',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.','mat'));
% Variables
NbStrik=N/2; %number of stikes
NbMat =M; %number of considered maturities
NbCall =NbMat*NbStrik; %number of calls with different strikes
NbPut =NbMat*NbStrik; %number of puts with different strikes
NbOpt = NbCall+NbPut; % number of calls and puts out-the money
```

```
RND=[]; %store de densities
Skewness=[];
Kurtosis =[];
Mean =[];
Std =[];
for i=1:1 %RNDs with 1 month term
S0 = AllInfo(1,1); %SPOT USBRL
KC = AllInfo(1+NbStrik*(i-1):NbStrik*i,9); %Call Strike prices
KP = AllInfo(NbCall+1+NbStrik*(i-1):NbCall+NbStrik*i,9); %Put Strike Prices
CPi= AllInfo(1+NbStrik*(i-1):NbStrik*i,3);
rbrl = AllInfo(1+NbStrik*(i-1),8); % brazilian interest rate (domestic)
T = AllInfo(1+NbStrik*(i-1),5); %time to maturity
C = AllInfo(1+NbStrik*(i-1):NbStrik*i,10); %call option
P = AllInfo(NbCall+1+NbStrik*(i-1):NbCall+NbStrik*i,10); %put option
id = AllInfo(1+NbStrik*(i-1),7);
ivc=AllInfo(1+NbStrik*(i-1):NbStrik*i,6)/100; % implied volatilities Calls
ivp=AllInfo(NbCall+1+NbStrik*(i-1):NbCall+NbStrik*i,6)/100; % implied volatilities Puts
r = AllInfo(1+NbStrik*(i-1),4); %libor interest rate
DC = AllInfo(1+NbStrik*(i-1):NbStrik*i,2); %Call deltas
DP = -AllInfo(NbCall+1+NbStrik*(i-1):NbCall+NbStrik*i,2).*exp(-r.*T)+exp(-r.*T); %con-
vert out-the-money put deltas into out-the-money call deltas
dc=linspace(0,0.5,500); dc=dc';
dp=linspace(0.5,1,500); dp=dp';
d =[dc;dp];
y = [ivc; ivp]; %implied volatilities
Dl= [DC; DP]; %put deltas and call deltas
x = Dl'; y = y(:).';
p=0.9; %smoothing parameter for the natural spline
```

spline1 = fnxtr(csaps(x,y,p),2); %natural spline curve giving the implied volatilities in terms of deltas with vega weighting

sighC = fnval(spline1,d); %implied volatilies given by the natural spline curve

z = S0.*exp(-norminv(d.*exp((rbrl-r).*T)).*(sighC).*sqrt(T)+(r+((sighC).^2)./2).*T); %convert the grid of deltas into strike prices

NaN=1-isnan(z); %detect NaN

f = find(NaN==0); %position of NaNs

z(f) = []; %delete NaN in strike vector

d(f) = []; %delete NaN in delta vector

Inf=1-isinf(z); %detect Inf

f = find(Inf==0); %position of Inf

z(f) = []; %delete Inf in strike vector

d(f) = []; %delete Inf in delta vector

sighC = fnval(spline1,d); %implied volatilies given by the natural spline curve after cleaning the delta vector

%interpolation between delta and strike prices in order to obtain

%equidistant strike prices and this way use the Riemman Sum to estimate the

%summary statistics

q=1;

spline2 = fnxtr(csaps(z,d,q),2); %natural spline curve giving the deltas in terms of strike prices

z = linspace(1.4,3.8,5000)';

nz = size(z,1);

d = max(real(fnval(spline2,z)),0); %deltas from the equidistant strike prices

sighC = real(fnval(spline1,d)); %implied volatilies given by the natural spline curve after obtaining deltas corresponding to equidistant deltas

BSCallDCUR(S0,z,rbrl,sighC,T,r); %insertion of the smile volatility curve into Black-Scholes

stau=sqrt(T);

```
d1 = ((log(S0./z) + ( rbrl - r + 0.5.*(sighC.^2)) .* T)./(sighC.*stau));
d2 = (d1 - sighC.*stau);
dif = diff(z);
f = ((BSCallDCUR(S0,z(3:nz),rbrl,sighC(2:(nz-1)),T,r)
- 2*BSCallDCUR(S0,z(2:(nz-1)),rbrl,sighC(2:(nz-1)),T,r)
+ BSCallDCUR(S0,z(1:(nz-2)),rbrl,sighC(2:(nz-1)),T,r))./(dif(1).^2))*exp((rbrl-r)*T);
diferencial = diff(z);
Em = sum(z(2:(nz-1)).*f.*diferencial(1)); %Expected Value
Var = sum((z(2:(nz-1))-Em).^2.*f.*diferencial(1)); %implied variance
Skew = sum(((z(2:(nz-1))-Em).^3).*f.*diferencial(1))./(Var.^(3/2)); %implied skewness
Kurt = sum(((z(2:(nz-1))-Em).^4).*f.*diferencial(1))./(Var.^2); %implied kurtosis
RND = [f RND];
figure(w)
plot(z(2:(nz-1)),RND)
title('RND')
xlabel('STRIKE PRICES')
axis tight
hold on
end
print ('-djpeg',strcat('SPLINE',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w))))
Skewness = [Skew Skewness];
Kurtosis = [Kurt Kurtosis];
Mean = [Em Mean];
Std = [sqrt(Var) Std];
save (strcat('SPLINE_mat1_',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.mat'));
end
```

12.4 MLN model codes

12.4.1 MLN USDBRL parameters

The USDBRL RNDs estimated through MLN method for the end of month dates between June 2006 and February 2010 were calculated through the code bellow:

```
function MLN()


clc; clear;
%open an excel file with the selected dates in the first column (in this case 45
%dates)
D=xlsread('data.xls');
Dia = D(1:45,1);
Mes = D(1:45,2);
Ano = D(1:45,3);
for w=1:45; %Open tables with market information for 45 dates
load (strcat('USDBRL_',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.','mat'));
% Variables
NbStrik=N/2; %number of stikes
NbMat =M; %number of considered maturities
NbCall =NbMat*NbStrik; %number of calls with different strikes
NbPut =NbMat*NbStrik; %number of puts with different strikes
NbOpt = NbCall+NbPut; % number of calls and puts out-the money
z=linspace(1,4,5000); z=z'; %RND support
Res=[]; %here store the densities
ParamM=[]; % parameters to store
Skewness=[];
Kurtosis =[];
```

```
Mean =[];
Std =[];
for i=1:1 %RNDs with 1 month term

S0 = AllInfo(1,1); %SPOT USBRL
KC = AllInfo(1+NbStrik*(i-1):NbStrik*i,9); %Call Strike prices
KP = AllInfo(NbCall+1+NbStrik*(i-1):NbCall+NbStrik*i,9); %Put Strike Prices
CPi= AllInfo(1+NbStrik*(i-1):NbStrik*2*i,3);
rbrl = AllInfo(1+NbStrik*(i-1),8); %brazilian interest rate (domestic)
T = AllInfo(1+NbStrik*(i-1),5); %time to maturity
C = AllInfo(1+NbStrik*(i-1):NbStrik*i,10); %call option
P = AllInfo(NbCall+1+NbStrik*(i-1):NbCall+NbStrik*i,10); %put option
r = AllInfo(1+NbStrik*(i-1),4); %libor interest rate
lb=[ -4; -4; 0.0001; 0.0001 ]; % lower bounds for parameters
ub=[ 4; 4; 0.8; 0.8]; %upper bounds for parameters

b0=[ 0.1; 0.1; 0.4; 0.01];

A=[0 0 -1 1];
b=0;
av=0.01:0.1:0.99; av=av';
GridRes=[];
for j=1:size(av,1)

options=optimset('Algorithm','trust-region-reflective','Diagnostics','on','Display','iter','TolX',1e-
6,'TolFun',1e-6,'MaxFunEvals',500,'MaxIter',5000);
```

```
[beta,Qmin,residual,exitflag,output] = lsqnonlin(@(b)MD_Obj1(b,S0,[KC;KP],CPi,rbrl,T,[C;
P],r,av(j)),b0,lb,ub,options);

GridRes=[GridRes; [av(j) beta' Qmin]];
end

[mi,miidx]=min(GridRes(:,end));
 b0=GridRes(miidx,1:5); b0=b0';

lb=[ 0.0001; -3; -3; 0.0001; 0.0001 ]; % lower bounds for parameters
ub=[ 0.9999; 3; 3; 0.9; 0.9]; %upper bounds for parameters

options=optimset('Algorithm','trust-region-reflective','Diagnostics','on','Display','iter','TolX',1e-
6,'TolFun',1e-6,'MaxFunEvals',150,'MaxIter',5000);
[beta,Qmin,residual,exitflag,output] = lsqnonlin(@(b)MD_Obj(b,S0,[KC;KP],CPi,rbrl,T,[C;
P],r),b0,lb,ub,options);

ParamM=[ ParamM; beta' ];
a = beta(1);
mu1 = beta(2);
mu2 = beta(3);
sig1 = beta(4);
sig2 = beta(5);
f = a * get_LN_RND(z,S0,mu1,T,r,sig1) +...
(1-a) * get_LN_RND(z,S0,mu2,T,r,sig2);

diferencial = diff(z);
```

```
Em = sum(z.*f.*diferencial(1));
Var = sum((z-Em).^2.*f.*diferencial(1));
Skew = sum(((z-Em).^3).*f.*diferencial(1))./(Var.^(3/2));
Kurt = sum(((z-Em).^4).*f.*diferencial(1))./(Var.^2);


end
figure(w)
plot(z,f)
title('RND')
xlabel('STRIKE PRICES')
MixRND=Res;
save MixRND;
fdp = figure(w);
filename = strcat('Final_',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.mat');
print(fdp, '-dpsc', filename)
Skewness = [Skew Skewness];
Kurtosis = [Kurt Kurtosis];
Mean = [Em Mean];
Std = [sqrt(Var) Std];
disp('The parameters are ');
ParamM
save (strcat('MLN_mat3_',num2str(Dia(w)),'-',num2str(Mes(w)),'-',num2str(Ano(w)),'.mat'));
end
%*********************************************************
function y=MD_Obj(b,S0,K,CPi,rbrl,T,CP,r)
%
a = b(1);
piv = [a; 1-a];
```

```
muv = b(2:3); muv=muv';
sigv = b(4:5); sigv=sigv';
NbStrik=6;
Cemp=CP(1:NbStrik);
Pemp=CP(1+NbStrik:2*NbStrik);
KC=K(1:NbStrik);
KP=K(1+NbStrik:2*NbStrik);
Cth=a*BSCallDCUR(S0,KC,muv(1),sigv(1),T,r)+...
(1-a)*BSCallDCUR(S0,KC,muv(2),sigv(2),T,r);
Pth=a*BSPutDCUR(S0,KP,muv(1),sigv(1),T,r)+...
(1-a)*BSPutDCUR(S0,KP,muv(2),sigv(2),T,r);
y1 = Cemp - Cth; %calls
y2 = Pemp - Pth; %puts
e = S0>=K(1:NbStrik);
y = y1.*e + y2.*(1-e);
dist1=y;
mart= S0 - a * BSCallDCUR(S0,0.0001,sigv(1),muv(1),T,r)-...
 (1-a) * BSCallDCUR(S0,0.0001,sigv(2),muv(2),T,r);
y = dist1 + mart^2;
%************************************************************
function y=MD_Obj1(b,S0,K,CPi,rbrl,T,CP,r,a)
y=MD_Obj([a;b],S0,K,CPi,rbrl,T,CP,r);
%************************************************************

function y=get_LN_RND(z,S0,rbrl,T,r,sig);
% the benchmark log-normal density
m = log(S0) + (rbrl-r-0.5*sig^2)*T;
s = sig*sqrt(T);
```

```
x = ( log(z)-m )/s;
y=1/s*pdfn(x) ./ z;
function y=pdfn(x);
y=1/sqrt(2*pi)*exp(-0.5*x.^2);
```

Bibliography

Abadir, K. M. (1999). *An introduction to hypergeometric functions for economists.* Econometric Reviews 18 (2003).

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

Arrow, K. J. and Debreu, G. (1954). *Existence of an equilibrium for a competitive economy.* Econometrica 22:265-290.

Arrow, K. J. (1964). *The Role of Securities in the Optimal Allocation of Risk-Bearing.*Review of Economic Studies, 31, No. 2 (April), pages 91-96.

Bahra, B. (1997). *Implied risk-neutral probability density functions from option prices: Theory and application.* Working Paper, Bank of England.

Black, F. and Sholes, M. (1973), *Pricing of Options and Corporate Liabilities*, Journal of Political Economy, 81, pages 637-659.

Bliss, R. and Panigirtzoglou, N. (2002). *Testing the stability of implied probability density functions.* Journal of Banking and Finance 26, 381–422.

Bondarenko, O. (2003). *Estimation of risk-neutral densities using positive convolution approximation.* Journal of Econometrics 116, 85–112.

Breeden, D. T. and Litzenberger, R. H. (1978). *Prices of state-contingent claims implicit in option prices.* Journal of Business 51, 621–51.

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

Campa, J.C., Chang, P.H.K., and Reider, R.L. (1997). *ERM bandwidths for EMU and after: evidence from foreign exchange options.* Economic Policy, 55–87.

Cooper, N. (1999).*Testing techniques for estimating implied RNDs from the prices of European and American options.* Working Paper, Bank of England.

Cox, J. and Ross, S. (1976), *The Valuation of Options for Alternative Stochastic Processes*, Journal of Financial Economics, 3, pages 145-66.

Espen, G. H. (2007). *Option Pricing Formulas, second edition.*

Heston, S. (1993). *A Closed-Form Solution for Options with Stochastic Volatility With Applications to Bond and Currency Options*, The Review of Financial Studies, Vol 6, No 2, pp 327-343.

Jondeau, E. and Rockinger, M. (2000). *Reading the smile: The message conveyed by methods which infer risk neutral densities.* Journal of International Money and Finance 19, 885–915.

Jondeau, E., Poon, S.H and Rockinger, M. (2006). *Financial Modeling Under Non-Gaussian Distributions.*

Lee, S. H. (2008). *Three essays on estimation of risk neutral measures using option pricing models.* Dissertation presented for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

Malz, A.M., (1996). Options-based estimates of the probability distribution of exchange rates and currency excess returns, mimeo. Federal Reserve Bank of New York.

Malz, A. M. (1997). *Estimating the probability distribution of the future exchange rate from options prices.* Journal of Derivatives, Winter, pages 18-36.

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, pages 91-115.

Shimko, D. C. (1993), *Bounds of probability*, Risk 6, pages 33-37.