ABSTRACT

The present work proposes an application of a non-parametric methodology to extract the risk-neutral probability density function (RND) to USD/BRL options. This methodology consists in complementing the RND extracted from the implied volatility smile with tails drawn from a GEV (generalized extreme value) distribution. These non-parametrics risk-neutral densities are compared to parametric distributions that are frequently mentioned in the literature - specifically, the mixture of log-normal densities (MLN) and the generalized beta of second kind (GB2) - through a study of their moments. The present work proposes a straightforward methodology to apply the estimated moments for trading of USD/BRL futures, finding strategies that produce greater returns relative to a simple buy and hold strategy.

Keywords: Risk-neutral density. Implied volatility smile. USD/BRL options. Generalized extreme value. Mixture of log-normal densities. Generalized beta of second kind. Futures trading.

RESUMO

O presente trabalho propõe a aplicação de uma metodologia não paramétrica para extração da medida de probabilidade neutra ao risco para opções sobre taxa de câmbio de reais por dólar dos Estados Unidos USD/BRL. Esta metodologia consiste em complementar a função de densidade de probabilidade extraída a partir do *smile* de volatilidade implícita utilizando as caudas de uma distribuição GVE (generalizada de valores extremos). Estas densidades neutras ao risco não parametricas são comparadas com distribuições parametricas frequentemente discutidas na literatura - especificamente, a mistura de densidades log-normais (MLN) e a beta generalizada de segundo tipo (GB2) - através do estudo de seus momentos. O presente trabalho propõe uma metodologia simples para aplicação desses momentos para fins de *trading* de contratos futuros de taxa de câmbio USD/BRL, encontrando retornos superiores a uma estratégia *buy and hold* simples.

Palavras-chave: Densidade neutra ao risco. Volatilidade implícita. Opções de USD/BRL. Distribuição generalizada de valores extremos. Mistura de densidades log-normais. Distribuição generalizada beta de segundo tipo. Trading de futuros.

List of Figures

List of Tables

List of abbreviations and acronyms

USD/BRL	Amount of Brazilian reais per one unit of US dollars
RND	Risk-neutral probability density function
GEV	Generalized extreme value
MLN	Mixture of log-normal densities
GB2	Generalized beta of second kind
FX	Foreign exchange
ATM	At-the-money
OTM	Out of the money
ITM	In the money

Contents

1 Introduction

Over the past decades, the understanding of financial markets' behaviour and expectations became evermore important, as the interconnectedness between markets' dynamics and domestic/international economics increased substantially. Investors, scholars and policy makers have increasingly been looking at ways to correctly assess the market's relative probabilities attributed to different future scenarios for a given asset, in order to improve decision taking in several areas.

In this sense, the concept of the risk-neutral probability density function of an asset arises as an appealing approach to obtaining such understanding of the markets' sentiment. Under a number of assumptions, the risk-neutral density maps the probabilities associated with different prices of an asset at a given maturity. With the complete distribution of probabilities in hands for different maturities, it is possible to measure its evolution through time and the impact of information shocks on its shape. The applications of the risk-neutral density are numerous, ranging from pricing of derivatives of that underlying asset, to assessment of market's expectations (for instance, the VIX index) in the context of central bank policy making and even for trading.

There is no single framework for extracting an asset's risk-neutral density function, and the existing ones can be divided into two main categories: one that assumes that the underlying density function has a known functional form and revolves around finding the parameters that best fit a set of theoretical options with respect to observable options; and other that does not assume any particular shape for the bulk of density function, but rather make assumptions on the shape of its tails in order to retrieve a complete distribution, as it fills the lack of (reliable) data from away-from-the-money options. These frameworks are commonly referred as parametric and non-parametric approaches to extracting a risk-neutral density, with both having theoretical and practical advantages.

In the case of Brazilian assets, the discussions regarding the extraction of the risk-neutral density have been focused predominantly on the USD/BRL - mostly due to the market's importance and the more prominent liquidity of its options. These works have focused on parametric techniques. The first goal of this work is to present a non-parametric approach to extracting the risk-neutral density that draws the tails of the density from a generalized extreme value (GEV) distribution. Afterwards we compare the results obtained with two of the main parametric distributions used for modelling the USD/BRL risk-neutral density, namely, the generalized beta of second kind and the mixture of log-normal distributions.

After presenting and discussing the three techniques mentioned, the second goal of this work is to test whether it is possible to use the informational content of the risk-neutral densities (captured via the estimation of its variance, skewness and kurtosis) for trading. In other words, we look for using the moments of the risk-neutral densities to build a trading strategy for USD/BRL itself.

The work is divided in the following way: in Chapter 2, the literature regarding the extraction of the risk-neutral density is reviewed. Chapter 3 presents in-depth the theoretical framework behind the techniques used to extract the risk-neutral density functions. Chapter 4 discusses the methodology used to apply these frameworks for USD/BRL and the estimation of its moments. Chapter 5 presents the trading model adopted and the returns obtained by applying it to the moments previously estimated; and Chapter 6 presents the conclusions and final remarks.

2 Literature Overview

Back in 1973, Black and Scholes published a seminal work where they presented to the world an explicit formula to calculate the price of European options based on a number of parameters and a continuous-time economy (BLACK; SCHOLES, 1973). While the volatility of the underlying asset until maturity was an unobservable input of the model, the option prices were directly observable in the market; when these were input in the Black-Scholes framework, the 'implied volatility' could be extracted. However, although the original Black-Scholes model assumed that the asset's volatility until maturity was constant for all strikes, different implied volatilities could be observed for different observable strikes with the same maturity - revealing an inconsistency in the original model proposed.

The observation of such phenomenon led to the development of the concept of the implied volatility smile - the collection of implied volatilities extracted from options with different strikes and the same maturity - and the implied volatility surface - the collection of smiles for different maturities. Irrespective of the inconsistencies found, traders continued to resort to the Black-Scholes framework as a way to guarantee standardization and pricing consistency across different products, which would be latter called the 'practitioner's Black and Scholes' approach.

As the studies regarding the options market advanced, it became clear that a collection of option prices for the same maturity revealed a lot more about the market's assumptions and preferences than just the implied volatility smile. Should this collection of prices respect the principle of convexity [1], when the future expected pay-offs are taken at present using the risk-free rate, the risk-neutral probability density function can be computed straightforwardly. As noted by Breeden and Litzenberger (1978), the extraction of the options' underlying probability distribution does not rely on any assumptions on the stochastic process governing the asset nor on the pricing model adopted. Rather, the authors show that a risk-neutral probability function is simply the second partial derivative of (a continuum of) European call option prices with respect to their strikes. As the difference between strikes tend to zero, these second derivatives become the Arrow-Debreu 'elementary' state contingent-claim (BAHRA, 1997).

The possibility to extract the market's risk-neutral probability density function (commonly referred as the risk-neutral density, or henceforth simply as 'RND') is a powerful feature, as it captures the market's expectations and risk preferences for a given time horizon. Given the different applications for such tool, a considerable number of research

[1] This property that was proven by Merton (1973).

works have delved onto how to extract, fit and apply risk-neutral densities.

In this regard, Jackwerth (2004) and Figlewski (2018) provide a comprehensive summary of the main publications on this topic and their conclusions. They both note that the extraction of the RNDs using option prices can be organized into two different categories: a parametric approach and a non-parametric one. The former chooses a probability distribution with a closed functional form to price options for different strikes, with its parameters being optimized to minimize the difference between the theoretical and actual prices. The latter makes no assumption on the distribution governing the underlying stochastic process. The most relevant parametric and non-parametric approaches to modelling the underlying risk-neutral probability distributions will be discussed as follows.

2.1 Parametric approaches

Several works have tried to overcome the drawbacks of the original Black-Scholes model while still remaining under the Gaussian distributions framework, whether by making changes in the underlying process or generalizing the shape of the RND. Following the line of work of previous publications that highlight pricing inconsistencies in the original Black-Scholes framework, Corrado and Su (1996) propose a Gram-Charlie series expansion of the Normal distribution of log-prices in order to arrive in a new pricing equation.

Bahra (1997), on the other hand, notes that drawing assumptions for the shape of the RND is more general than assuming which stochastic process drives prices of the underlying asset: a given terminal distribution can be generated by several different stochastic processes, while a single process can only generate one distribution. As the author stresses out that financial assets' price distributions are in the neighbourhood of the log-normal one, he chooses to model the final distribution as a mixture of two log-normal distributions, deriving closed formulas for the prices of European calls and put options. With this, it is possible to minimize the sum of squared errors between the theoretical and actual option prices by altering the five parameters that describe the distributions and by setting the interest rate to the observable market rate. The combination of parameters that optimize this minimization problem can therefore completely describe the risk-neutral density.

Bliss and Panigirtzoglou (2002) also choose the mixture of two log-normal distributions as the parametric framework to fit FTSE option data and compare with a non-parametric approach presented by them. They find strong evidence of the superior stability of the latter when shocks are applied to the prices and the RNDs are re-estimated.

Liu et al. (2007) compare the results obtained from the mixture of two log-normal densities (MLN) with the generalized beta of second kind (GB2) distribution. Their focus is to analyse the transformation from risk-neutral density functions into 'real-world' densities by assuming a power utility function for the representative agent (similarly to Bliss and Panigirtzoglou (2002)). They note that the GB2 density allows for a convenient transformation between probability density functions.

Fabozzi, Tunaru and Albota (2009) compare several different parametric distributions when looking at interest rate options. The authors derive closed formulas for call and put prices under the Weibull, generalized beta of second kind and generalized gamma distributions, and find the distribution parameters that minimize an error function that accounts for the difference between model and actual option prices. They argue that the generalized gamma distribution provides a stable distribution with less parameters (3 rather than 4) than the more commonly used GB2 distribution.

2.2 Non-parametric approaches

The Breeden and Litzenberger result can be considered the cornerstone of any non-parametric approach to extracting the risk-neutral density function. However, in order to obtain a continuum of option prices that allow to calculate the price of Arrow-Debreu securities[2], some interpolation technique has to be applied in order to fill the gaps between the (discrete) observable option prices. Bates (1991) applies a constrained cubic spline in the ratio of S&P 500 option prices to future prices in order to obtain enough interpolated points to proceed with the extraction of a RND.

Shimko (1993), on the other hand, states that smoother results can be achieved by interpolating implied volatility smiles, rather than the collection of available call option prices, as the volatility smile provides a more well-behaved space. He fits the implied volatilities with a least-squares quadratic equation that does not require that the estimated function necessarily goes through the data points, and inputs the estimated implied volatilities into the Black-Scholes model to obtain the equivalent prices and the RND. The option to use the Black-Scholes formula to retrieve prices does not require it to be true, as it works simply as a transformation device that allows the interpolation to be conducted in implied volatility measurement space (BAHRA, 1997); (FIGLEWSKI, 2008). After the risk-neutral distribution is obtained from prices, Shimko assumes the tails of the distribution to be log-normal in points beyond the range of traded strikes.

Malz (1997) also proposes to fill the discontinuity between traded options by interpolating the implied volatility smile – in his case, looking at over-the-counter currency options market. He stresses out that differently from other assets such as equities, the convention in the foreign exchange (FX) market is to refer to implied volatilities in terms

[2] Arrow-Debreu securities are securities whose pay-off is equal to one unit if a specific state materialize, and zero otherwise.

of the option delta (the first derivative of the option price with respect to the underlying asset's price), rather than strikes. Given this, he opts to interpolate the implied volatility smile in the volatility/option delta space by combining option structures such as strangles and risk reversals along with at-the-money (ATM) implied volatility. After a continuous function in the implied volatility smile is produced, the volatilities are translated to volatility/option strike space. The standard Black-Scholes formula is used to obtain the prices that will be used to apply the Breeden-Litzenberger result.

This quadratic function fit technique on the implied volatility/option delta space is compared to Shimko's approach to apply the same framework in the implied volatility/option strike space. While the estimated implied volatilities coincide around the at-the-money, Malz points that the framework of Shimko tends to overestimate volatilities for very low and high deltas, leading to price estimates that have substantially higher estimating errors when compared to actual market prices.

When conducting a non-parametric estimation of the RND, there is a trade-off between smoothness of the interpolated volatility smile and the goodness of fit of estimates in points where market prices are observable. A standard cubic spline, for instance, would require the interpolating function to be continuous at all data points, thus incorporating market micro-structure's noise into the estimation of the density function (FIGLEWSKI, 2008). Bliss and Panigirtzoglou (2002) propose to use a natural cubic spline with a smoothness penalty parameter that impacts the curvature of the piece-wise function – with the goal of balancing the smoothness and goodness of fit of the interpolated curve.

According to Jackwerth (2004), while frameworks based on curve-fitting of the implied volatility smile tend to be the most stable ones to extract risk-neutral densities, their results will yield very similar estimates for the center of the distributions and the differences will occur mostly in the tails. He proposes to minimize the difference between the estimated and observed implied volatilities through a function with a specific functional form that takes into consideration the second derivatives of these implied volatilities and includes an *ad-hoc* smoothness parameter (or trade-off parameter). The first and last estimated implied volatilities are extrapolated outside the estimation range. He then calculates the Black-Scholes price associated with each implied volatility, and 're-weights' each price so that the probabilities sum to one for the complete RND.

Figlewski (2008) also argues for the use of a smoothing spline to interpolate the volatility smile. In order to allow greater flexibility in the estimated function, he opts to conduct the interpolation with a fourth degree spline where the only point where the function necessarily goes through the data (referred as a 'knot') is placed ATM. Additionally, the author also takes advantage of the availability of information on bid-offer spreads in S&P 500 options to further treat the data: weighting parameters are added in the interpolated function so as to give more importance to estimates that fall within the bid-offer spread than those that don't. Moreover, given that the lack of liquidity of deep out of the money (OTM) options might affect the estimation of the RND, the bid-offer spreads are also used to trim the volatility smile according an *ad-hoc* maximum spread value. Finally, the tails of the estimated distribution are drawn according to the GEV distribution, as opposed to the log-normal choice adopted in previous works.

2.3 The case of Brazilian assets

The issue of estimating the RND of Brazilian financial assets has been mostly tackled using parametric frameworks.

Tabak and Chang (2002) extract the risk-neutral density function from USD/BRL options non-parametrically: they map out implied volatilities in the implied volatility/option strike space from actual prices and estimate additional points from the available information using a quadratic regression. The authors process the interpolation with as few as three points, which may carry a low information content about the underlying asset. The option prices are then calculated by inputting the estimated implied volatilities into the Garman-Kohalgen pricing model and applying the Breeden-Litzenberger result (GARMAN; KOHLHAGEN, 1983).

Abe, Chang and Tabak (2007) choose to use the GB2 distribution to parametrically estimate the risk-neutral density function for USD/BRL options between 2000 and 2005, as this particular distribution allows for a straightforward estimation of the 'real-world' probability density function and the risk-aversion parameter. They extract the closed-formulas for call and put options and then estimate these parameters by minimizing the sum of squared differences between theoretical and actual option prices. Then, after producing monthly estimates of the RND for one-month maturity options, the authors test whether the estimated density functions produce accurate interval forecasts for USD/BRL, with mixed conclusions.

Ornelas, Fajardo and Farias (2012) estimate the RND implied in USD/BRL options between 1999 and 2010 using the framework of the mixture of two log-normal distributions. Then, they assume that the representative agent has power utility function and constant relative risk aversion, and proceed to estimate the real-world probability distribution function based on these assumptions. The authors conclude that the real-world density produces better density forecasts than the RND.

The work of Santos and Guerra (2015) provides an interesting comparison between different RND estimation methods and results when applied to USD/BRL. First, the authors apply the Heston stochastic volatility pricing model in order to obtain what is referred as the 'true' risk-neutral density function (HESTON, 1993). Then, other RNDs are estimated with parametric frameworks such as the mixture of two log-normal densities, the density function based on confluent hypergeometric expansion, the Edgeworth expansion, and the non-parametric framework of the interpolated implied volatility smile proposed by Bliss and Panigirtzoglou (2002). Santos and Guerra use the period of June 2006 to January 2010 to conduct estimations.

In order to compare the different methodologies, the authors input into the Heston model a set of parameters according to scenarios for different regimes of volatility and skewness of USD/BRL. Then, the call options are priced according to the closed-formula proposed by the Heston framework under each scenario, and a random noise perturbation is added to these prices. The RNDs are generated by applying the parametric and non-parametric methodologies listed above to this set of shocked option prices, and these resulting densities are compared to the 'true' ones (generated by Heston model) to measure accuracy and stability. The comparisons are made by measuring the summary statistics (mean, variance, skewness and kurtosis) of the estimated and 'true' RND, as well as the root mean integrated squared error. They conclude that the mixture of two log-normal and the density functional based on confluent hypergeometric expansion provide the most accurate estimates, while the smoothed implied volatility smile proved to be the most stable method and the one that produced most accurate variance estimates.

3 Theoretical Framework

The goal of the present work is to propose a framework to extract the risk-neutral probability distribution function non-parametrically – that is, without previously assuming any functional form for probability distribution – and to complement such estimation with distribution tails assuming the shape of the GEV distribution. The resulting distribution will then be compared to more-often used parametric frameworks such as the MLN and the GB2, which will be more profoundly discussed in Section 3.4.

This chapter will discuss in depth the theoretical framework behind the extraction of the risk-neutral probability densities used on this work. Also will be highlighted the methodological choices that were taken in order to make such process viable. Then, the practical details behind the implementation of such model will be laid out in Chapter 4.

3.1 The underlying risk-neutral probability density

Options are derivatives that allow the buyer the possibility to buy or to sell a quantity of the underlying asset for a pre-determined price if it desires to do so, in exchange of a price (or premium) for this optionality. The characteristics of option contracts give market agents the possibility to build flexible future pay-offs according to the desires to increase or mitigate exposure in a certain asset under specific conditions. The fact that the option prices are determined by supply and demand in market places means that the agents' perceived likelihood that a scenario materializes will vary along with the price paid for options that reflect such outcome.

In this sense, the price of an option at time t for the underlying asset S, strike K and maturity T can be written as:

$$C(t,T,K) = e^{-r(T-t)} \int_{K}^{\infty} f(S_T)(S_T - K)\, dS_T \,, \tag{3.1}$$

$$P(t,T,K) = e^{-r(T-t)} \int_{-\infty}^{K} f(S_T)(K - S_T)\, dS_T \,, \tag{3.2}$$

where $C(t,T,K)$ and $P(t,T,K)$ are, respectively, the prices of European calls and puts, r is the risk-free interest rate, S_T is the price observed for S at time T, and $f(S_T)$ is the risk-neutral probability density function for S_T.

As shown in Breeden and Litzenberger (1978), if we twice differentiate the pricing formula (3.1) with respect to K, we arrive at the following relationship:

$$\frac{\partial^2 C}{\partial K^2} = e^{-r(T-t)} f(K)\,, \tag{3.3}$$

where we are able to extract the risk-neutral density function directly from a continuum of option prices. As the authors note, no assumptions regarding the underlying price's dynamics or the density function were made: it is only assumed that markets are perfect (that is, transaction costs are nil, short sales are allowed to unlimited quantities and borrowing can be made with the riskless interest rate) and that in the continuous case, the risk-neutral density function is twice differentiable.

In practice, observable option prices will not be available for a continuum of strikes, but rather only for a discrete number of observations that reflect liquidity aspects and market-specific conventions. In this case, the link between the available option prices and the risk-neutral density function will be presented by state-contingent prices or Arrow-Debreu prices – that is, prices that are equal to one unit if a specific state materialize, and zero otherwise.

It is possible to generate an Arrow-Debreu type of pay-off by building a butterfly spread between options with the same maturity and subsequent strikes. A pay-off of a butterfly structure centred around theoretical strike $K = 2$ is graphically represented in the Figure 1. With a simple combination of call options, it was possible to generate a structure whose pay-off is equal to 1 when $K = 2$, and 0 otherwise.

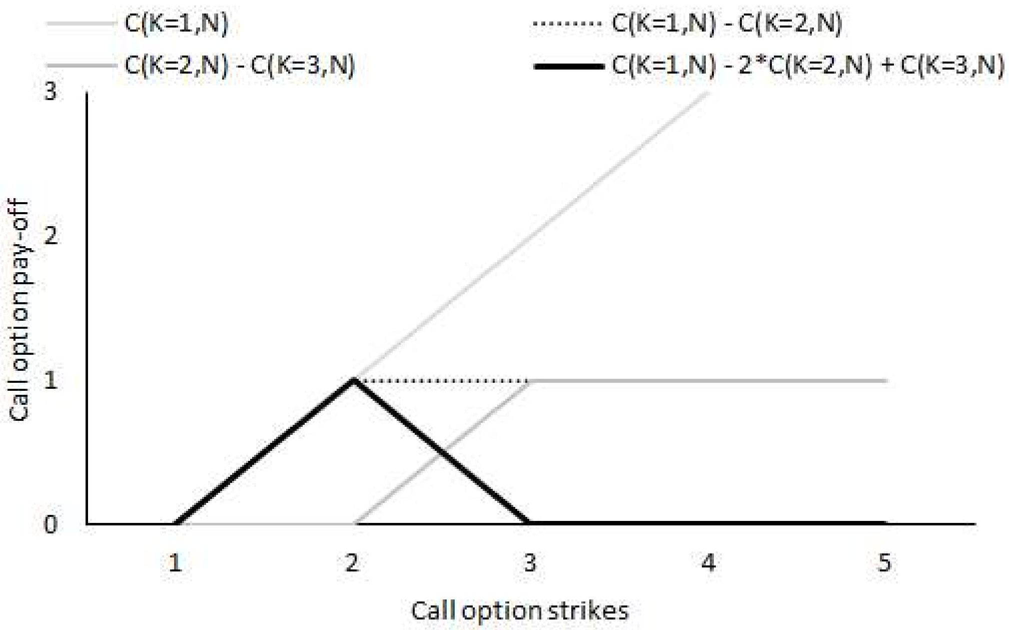

Figure 1 – Pay-off of the combined call option structures.

More generally, if we assume a step size of ΔM, a portfolio that has a payoff of 1 at K and 0 otherwise can be written as:

$$\frac{P(K,T,\Delta M)}{\Delta M} = \frac{[c(K-\Delta M,T) - 2c(K,T) + c(K+\Delta M,T)]}{(\Delta M)^2}.$$

By dividing the portfolio price by step size and taking the limit where the step size ΔM goes to zero, Breeden and Litzenberger show that:

$$\lim_{\Delta M \to 0} \frac{P(K,T,\Delta M)}{\Delta M} = \left.\frac{\partial^2 C}{\partial K^2}\right|_{K=M}. \tag{3.4}$$

That is, by taking infinitely small steps and pricing butterflies for a continuum of states, we have the complete state pricing function.

3.2 Filling the gaps between observable option prices

The result obtained in equation (3.3) is powerful, in the sense that one may obtain the market's complete distribution of probabilities for a given maturity by constructing simple portfolios of options without making assumptions on the underlying process governing the asset prices. This is an interesting feature of a non-parametric framework, as there is no *ex-ante* reason to believe that a risk-neutral probability distribution would follow any known parametric distribution.

However, while non-parametrically estimating the state pricing function allows for additional degrees of freedom in shape the distribution, two main hurdles in the estimation immediately arise. First, as noted before, only a handful of strikes are traded in the market for a given maturity, thus requiring some type of interpolating function in order to approximate prices in between observable data points.

Second, the estimation of the probability density function will only be possible up to the last strike observed, and the absence of deep in-the-money (ITM) and OTM data points will make the estimation of the tails of the distribution impossible without an extrapolation technique or additional assumptions regarding its shapes.

Regarding the interpolation of the observable option prices, the choice of technique will vary depending on the data available: by taking into account market microstructure's 'noises' such as bid-offer spreads (rather than mid prices) and/or price quotes from different time periods within a day, for instance, it may be required that the data is smoothed out prior to the estimation of the density function.

The interpolated points also have to be controlled for respecting the no-arbitrage conditions as put out originally by Merton (1973) – that is, respecting the principle of monotonicity (two European calls with same maturity and increasing strikes should be necessarily decreasing in price) and convexity (all the butterflies for a given maturity should be convex with respect to strikes). The convexity arbitrage condition is particularly important in the context of risk-neutral density functions, as applying the result of equation (3.4) in a non-convex case would translate into negative probabilities in the RND. Therefore, there is an underlying trade-off between the accuracy and smoothness of the estimated option prices relative to the original data.

Several works have presented a number of different approaches to interpolate the existing data points and thus obtain more observations for the estimation of the RND. While the work of Bates (1991) focused on obtaining an interpolating function between observable prices in the price/option strike space, Shimko (1993) indicated that this estimation should rather be conducted in the implied volatility/option strike space.

This is a naturally smoother function, as in practice market participants see the prices as almost as a 'numerical transformation' of the more-informational implied volatility surface. Since this work was published, the focus of the non-parametric estimations of the RND has almost entirely shifted into interpolating the implied volatility smile. After the work of Malz (1997), some papers opted to work with implied volatility/option delta. For the purposes of the methodology of the present work, a vector of implied volatilities calculated from observable option prices will be the object of interpolation and treatment in order to extract the RND.

In this regard, Bliss and Panigirtzoglou (2002) argue for the use of a natural cubic spline with a penalty parameter λ that gives more or less weight to the smoothness of the estimated piecewise function, depending on the value it assumes. The minimization also takes into consideration a weighting parameter, which they proposed to use the options' vega (the first derivative of the option price with respect to the volatility) in order to give more importance for observations that are closer to being ATM. The concept of a smoothing parameter is also adopted by Jackwerth (2004). While Figlewski (2008) opts for the use of a fourth degree polynomial to obtain smoothness, he imposes the condition that the estimated function necessarily go through one point (knot) placed at-the-money.

In this work, we chose to follow the methodology proposed by Bliss and Panigirtzoglou (2002) - also adopted in Santos and Guerra (2015) –, of using the natural smoothing spline while accounting for the use of *ad-hoc* knots as proposed by Figlewski (2008). The k-th order spline function is first described as:

$$\hat{\sigma}_i^k(\theta) = \sum_{n=1}^{k} a_n(\theta_i - \theta_{i-1})^n + a_{k+1}\,, \tag{3.5}$$

where θ can be a sequence of strikes or deltas, depending on how the implied volatility smile is represented. For a given vector of parameters a that determine the cubic spline and a spline function $f(\theta, a)$, the smoothing spline results from the following function minimization:

$$\min_{a} \quad w_i \sum_{n=1}^{N} (\sigma_i - \hat{\sigma}_i)^2 + \lambda \int_{-\infty}^{\infty} (f''(\theta, a))^2 d\theta \,, \tag{3.6}$$

where w_i is the relative weight attributed to each observation and λ is the smoothing penalty parameter. By altering the weight vector w, the knots where the function goes through the data can be promptly determined and adjusted. Conversely, λ controls the smoothness of the function by multiplying a measure of the degree of the curvature of the function. The combination of both parameters allow for a greater flexibility in determining the shape of the interpolating function. The choice of parameters for the present work is related to the type of data used in this work, and will be discussed in detail in Chapter 4.

After the smoothing spline function is estimated from the observable implied volatility data points, it is trivial to estimate other implied volatility points from the function and populate the implied volatility smile with infinitely small changes in θ, whether it is strike or delta. The new vector of implied volatilities may now be inputted into the Black-Scholes model to obtain the equivalent prices and the RND via relationship (3.4). This does not require the Black-Scholes formula to be true, as it works simply as a transformation device from the implied volatility measurement space back into prices (BAHRA, 1997);(FIGLEWSKI, 2008).

3.3 Treating the tails of the risk-neutral density function

As previously mentioned, one of the shortfalls of the non-parametric extraction of the risk-neutral density function is that the mapping of the function near its tails is dependent on information of deep ITM and OTM options. Given that these tend to be less liquid than the ones close-to-the-money, the estimation of the RND will be incomplete without making additional assumptions. A natural candidate to complete the tails of the RND is to assume log-normal tails by extrapolating the implied volatilities outside of the range of available information (SHIMKO, 1993); (JACKWERTH, 2004).

However, as noted in Figlewski (2008), the abundant evidence of fat tails in empirical return distributions make this a suboptimal choice, particularly if the end-goal of the estimation of the RND is to assess extreme events. Thus, the alternative proposed – which will be adopted in this work – is to draw the tails from the GEV distribution.

The author's choice for this specific distribution is based on the Fisher-Tippet theorem, stating that "under weak regularity conditions the largest value in a sample drawn from an unknown distribution will converge in distribution to one of three types of probability laws, all of which belong to the GEV family". This feature of the GEV distribution is therefore ideal to model the tails of a distribution.

The cumulative density function of the GEV has the following form:

$$F(z,\xi) = \begin{cases} \exp[-(1+z\xi)^{-\frac{1}{\xi}}] & \text{if } \xi \neq 0 \\ \exp[-\exp(-z)] & \text{if } \xi = 0 \end{cases}, \qquad z = \frac{S_T - \mu}{\sigma}, \tag{3.7}$$

with μ and σ serving as location and scale parameter, respectively. The parameter ξ is related to the shape of the distribution, assuming the form either of the Frechet distribution with heavy tails ($\xi > 0$), the Gumbel distribution ($\xi = 0$) or the Weibull distribution ($\xi < 0$). The probability density function of the GEV has the following form:

$$f(z,\xi) = \begin{cases} \frac{1}{\sigma}\,[(1+z\xi)^{-1-\frac{1}{\xi}}]\exp[-(1+z\xi)^{-\frac{1}{\xi}}] & \text{if } \xi \neq 0 \\ \frac{1}{\sigma}\exp(-z)\exp[-\exp(-z)] & \text{if } \xi = 0 \end{cases} \tag{3.8}$$

Following the notation adopted in Figlewski (2008), let us denote $F_{emp}(.)$ and $f_{emp}(.)$ as the empirical cumulative density function (CDF) and probability distribution function (PDF) estimated via the interpolation of the implied volatility smile. If we consider α as a quantile of the empirical cumulative density function, that is,

$$F_{emp}(K(\alpha)) = \alpha\,,$$

we should choose two significance levels α_1 and α_2 that are simultaneously close to the tails of the empirical distribution and existent for both the right and left tails - that is,

$$\alpha_{1L} = 1 - \alpha_{1R} \qquad \text{and} \qquad \alpha_{2L} = 1 - \alpha_{2R}.$$

These will serve as connection points where both the empirical non-parametric RND and the estimated GEV functions will be evaluated.

In order for the GEV density function to serve as the right tail of the non-parametric RND, it has to meet the following conditions:

$$\begin{cases} F_{GEV(R)}(K(\alpha_{1R})) = \alpha_{1R} \\ f_{GEV(R)}(K(\alpha_{1R})) = f_{emp(R)}(K(\alpha_{1R})) \\ f_{GEV(R)}(K(\alpha_{2R})) = f_{emp(R)}(K(\alpha_{2R})) \end{cases} \tag{3.9}$$

That is, the GEV CDF should have its α_{1R}-th quantile at point $K(\alpha_{1R})$ and the GEV PDF has to be equal to the empirical probability density function when evaluated at points α_{1R} and α_{2R}. Similar conditions have to be met for the left tail:

$$\begin{cases} F_{GEV(L)}(-K(\alpha_{1L})) = 1 - \alpha_{1L} \\ f_{GEV(L)}(-K(\alpha_{1L})) = f_{emp(L)}(K(\alpha_{1L})) \\ f_{GEV(L)}(-K(\alpha_{2L})) = f_{emp(L)}(K(\alpha_{2L})) \end{cases} \tag{3.10}$$

The parameter vector $\Theta = [\mu, \sigma, \xi]$ that describes the GEV distribution is the result of the following minimization problem:

$$\min_{\Theta} \quad (F_{GEV}(K(\alpha_1)) - \alpha_1)^2 \tag{3.11}$$

Subject to constraints (3.9) and (3.10), depending on which tail is being evaluated.

Figure 2 below shows a generic non-parametric risk-neutral density function extracted from a collection of option prices. By solving the minimization (3.11) and constraints (3.9) and (3.10), it is possible to estimate the GEV distributions that best fit the right and left tails of the distribution. The result is shown in Figure 3.

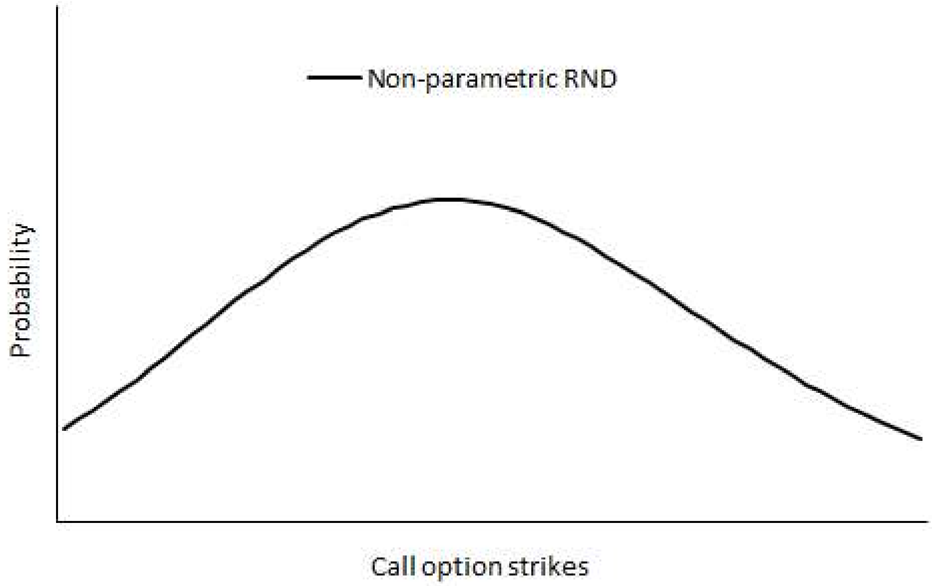

Figure 2 – Generic non-parametric RND estimated from a collection of option prices.

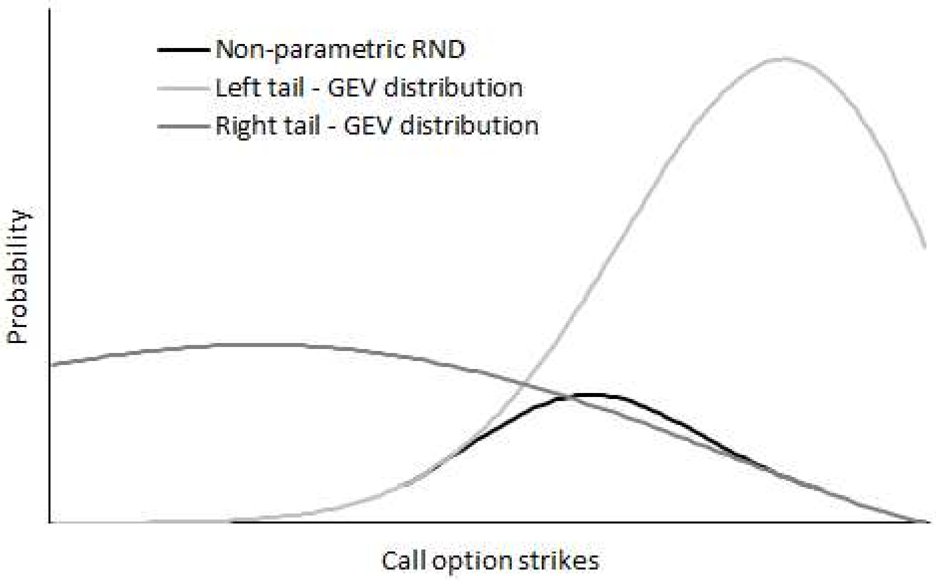

Figure 3 – GEV tails estimated from the truncated non-parametric RND.

Once the optimal parameters $\Theta^* = [\mu^*, \sigma^*, \xi^*]$ are used to build the GEV distributions, the non-parametric RND will be truncated at points α_{1R} and α_{1L} and replaced by the GEV for $K \geq K(\alpha_{1R})$ and $K \leq K(\alpha_{1L})$.

The newly estimated risk-neutral probability density function should be standardized by dividing each observation by the sum of all observations, so that all probabilities integrate to one (JACKWERTH, 2004). This yields a complete distribution that is estimated both non-parametrically by interpolating the implied volatility smile and parametrically via the adoption of tails that come from a GEV distribution.

3.4 The parametric estimations of the RND

So far, the extraction of the risk-neutral probability density function was discussed under non-parametric techniques, that is, without making assumptions regarding the form of the distribution. While this type of framework allows for greater flexibility in the shape of the RND, it also requires that the proper treatment for the distribution tails is applied, and creates additional hurdles to estimate the moments of the distribution.

In that regard, assuming that the risk-neutral density function has a parametric functional form has its advantages, as this type of framework will always yield a complete distribution (rather than a truncated one) whose moments can be calculated based on the estimated parameters. Thus, the goal of any parametric extraction of the RND revolves around finding a distribution (or combination of distributions) that allows for enough flexibility in its shape, while avoiding to be too cumbersome for the numerical estimation.

In this work, we will follow two choices of distribution that are widely used in the literature: the mixture of log-normal distributions (MLN) and the generalized Beta of second kind (GB2) distribution.

3.4.1 Generalized Beta of Second Kind

The GB2 distribution is discussed and implemented in Liu et al. (2007), Jondeau, Poon and Rockinger (2007) and, in the case of the Brazilian Real, in Abe, Chang and Tabak (2007) and Santos and Guerra (2015). The main advantage of using this type of distribution to model the RNDs is that only a few modifications are required in order to obtain a real-world density function, allowing a great number of applications in the results obtained.

The GB2 distribution is fully described by four parameters: $\theta = (a, b, p, q)$ and its probability density function has the form:

$$f_{GB2}(K, \theta) = \frac{a}{b^{ap} B(p,q)} K^{ap-1} \left[1 + \left(\frac{K}{b} \right)^{a} \right]^{-(p+q)}, \ s > 0, \tag{3.12}$$

where $B(p, q) = \Gamma(p)\Gamma(q)/\Gamma(p + q)$. The cumulative density function of the GB2 is related to the density function of the Beta distribution and has the form:

$$F_{GB2}(K, \theta) = F_{\beta}(y(K, a, b), \theta), \qquad y(K, a, b) = \frac{(K/b)^a}{1 + (K/b)^a}. \tag{3.13}$$

The martingale condition is imposed in order to guarantee risk neutrality and to facilitate the optimization of parameters:

$$S_t = e^{-r(T-t)} \frac{bB(p + \frac{1}{a}, q - \frac{1}{a})}{B(p,q)}. \tag{3.14}$$

The formulas for the price of European call option prices under the GB2 distribution are described in Liu et al. (2007) and Jondeau, Poon and Rockinger (2007). However, given that this work is focused on the estimation of FX risk-neutral densities, the domestic and foreign interest rates r_d and r_f have to be incorporated as follows:

$$\begin{aligned} C(K, \theta) &= e^{-r(T-t)} \int_{K}^{\infty} (S_T - K)^{+} f(K, \theta) dS_T \\ &= F e^{-r_f(T-t)} \left[1 - F_{GB2} \left(K, a, b, p + \frac{1}{a}, q - \frac{1}{a} \right) \right] \\ &\quad - K^{-r_d(T-t)} [1 - F_{GB2}(K, a, b, p, q)]. \end{aligned} \tag{3.15}$$

After a closed-formula for the price of the option is obtained, a vector of parameters $\theta^{*} = (a^{*}, b^{*}, p^{*}, q^{*})$ that solve the following minimization problem have to be found:

$$\min_{\theta} \quad \sum_{i=1}^{N_c}(C_{emp,i,T} - C_{GB2}(K(C_{emp,i,T}),\theta,T))^2 + \sum_{i=1}^{N_p}(P_{emp,i,T} - P_{GB2}(K(P_{emp,i,T}),\theta,T))^2 \,. \tag{3.16}$$

Subject to the restriction (3.14). The N_c and N_p are, respectively, the number of observable call and put prices with same maturity T. By minimizing the difference between the estimated and empirical option prices, the parametric risk-neutral density function can be described with the parameters θ^*.

3.4.2 Mixture of Log-normal Distributions

A very popular choice of parametric distribution to describe the risk-neutral density function is the mixture of log-normal densities (MLN). It allows for more flexibility in the estimation, as the number of mixed densities may be freely altered (although the majority of the literature use a combination of two log-normal densities), and the use of the log-normal density is a natural candidate as an extension of the original Black-Scholes model.

For the purposes of this work, we will adopt and describe a mixture of two log-normals. The PDF of the MLN will be simply the weighted combination of two log-normal densities:

$$g_{MLN}(K,\theta) = \alpha\, g_{LN}(K,\mu_1,\sigma_1,T) + (1-\alpha)\, g_{LN}(K,\mu_2,\sigma_2,T)\,, \tag{3.17}$$

with $0 < \alpha < 1$, where α distributes the total weight between the two log-normal distributions, μ_i are the scale parameters describing each of the i distributions, and σ_i are the variance parameters. Conversely, it can be written that:

$$F_i = \exp\left(log(S) + (\mu_i - 0.5\times\sigma_i^2)T + 0.5\times(\sigma_i\times\sqrt{T})^2\right), \tag{3.18}$$

where F_i will be the ATM forward related to each log-normal density function, and $(\mu_i - 0.5\times\sigma_i^2)T$ is the mean associated with the log-normal process.

Each PDF is described as:

$$g_{LN}(K,F,\sigma,T) = \frac{1}{K\sigma\sqrt{2\pi T}}\exp\left(\frac{1}{2}\left[\frac{log(K) - log(F) + 0.5\sigma^2 T}{\sigma\sqrt{T})}\right]^2\right). \tag{3.19}$$

Conversely, the price of an European call option under mixture of log-normals framework can be described as a weighted average of two options priced under the original equation proposed by Black (1976) (LIU et al., 2007);(JONDEAU; POON; ROCKINGER, 2007):

$$\begin{aligned} C_{MLN}(K, F_1, F_2, \sigma_1, \sigma_2, T) = \quad & \alpha \times C_1^{LN}(K, F_1, \sigma_1, T) \\ & +(1-\alpha) \times C_2^{LN}(K, F_2, \sigma_2, T) \end{aligned} \tag{3.20}$$

where C_{MLN} represents the price of an European call under the MLN distribution and C_i^{LN} represent the price of each European call under the log-normal distribution.

Similar to the procedure to find the optimal set of parameters under the GB2 distribution, the is to find $\theta^* = (F_1^*, \sigma_1^*, F_2^*, \sigma_2^*, \alpha^*)$ such that:

$$\begin{aligned} \min_{\delta} \quad & \sum_{i=1}^{N_c} (C_{emp,i,T} - C_{MLN}(K(C_{emp,i,T}), \theta, T))^2 \\ & + \sum_{i=1}^{N_p} (P_{emp,i,T} - P_{MLN}(K(P_{emp,i,T}), \theta, T))^2 \end{aligned} \tag{3.21}$$

subject to

$$F = \alpha \times F_1 + (1-\alpha) \times F2 \tag{3.22}$$

in order to ensure the martingale condition that mean of the risk-neutral density equals the ATM forward of the underlying asset.

4 Methodology

In this chapter, the methodology adopted in the estimation of the risk-neutral densities and its moments will be described and discussed in details.

4.1 Data treatment and implied volatility smile smoothing

The source of information for this work is the database provided by B3 (Brazilian stock exchange) between January 2014 and February 2019. In their website, it is possible to extract the information necessary to calculate the RNDs of the USD/BRL, such as: spot currency values, foreign and domestic interest rates, options' reference prices (that is, option prices published by B3 on a daily basis for a range of strikes and maturities for mark-to-market purposes), and most importantly, implied volatility structures. An extract of the options' reference prices and implied volatility structure per strike published by B3 on 12 April 2017 is depicted in the Table 1 below. The option strikes are quoted in USD/BRL future points.

Table 1 – Extract of the implied volatility and reference prices per strike and maturity published by B3 for USD/BRL European call options.

Maturity	Strike	Implied volatility (%)	Option price
02/May/2017	2975	12.86	191.13
02/May/2017	3000	12.65	166.57
02/May/2017	3025	12.44	142.34
02/May/2017	3050	12.23	118.67
02/May/2017	3075	12.11	96.07
02/May/2017	3100	12.16	75.32

Regarding the implied volatilities, the B3 follows an elaborate methodology in order to collect contributions from brokers/dealers and publish final IV estimates for different products daily. The methodology behind the calculation of implied volatilities for FX will be described in brief[1].

First, each broker/dealer listed as contributors is required to provide their particular set of implied volatilities for different maturities. These are measured in an implied volatility/option delta space, and the deltas requested to be filled in are currently: 99%, 95%, 90%, 75%, 63%, 50%, 37%, 25%, 10%, 5%, and 1%.

[1] The complete methodology adopted to price in options and treat the contributions can be found on B3's website: <http://www.b3.com.br/en_us/market-data-and-indices/data-services/market-data/reports/derivatives-market/methodology/bm-fbovespa-pricing-manual/>

With a set of implied volatility surfaces in hands, B3 assesses whether they respect arbitrage conditions - that is, if by inputting these implied volatilities in the Black-Scholes formula:

(i) the prices of the resulting European call options are strictly decreasing as the strikes increases for a given maturity;

(ii) the prices are strictly increasing as the time-to-expiry increases for a given strike;

(iii) the convexity of the prices with respect to the strike is strictly positive, preventing that butterfly structures with a positive payout (self-financing) are built.

The implied volatility estimates are also checked for arbitrage opportunities between participants, and the treated data points are fitted (in the case of USD/BRL, the SVI parametrization[2]) in order to produce a consolidated smile that meets the no-arbitrage conditions between implied volatility points.

Even though the methodology described treats the information in the implied volatility/option delta space, B3 only publishes implied volatility/option strike smiles. The transformation performed by B3 from the former to the latter is done according to the following formula:

$$\sigma_i = \sigma_p . \left(\frac{\sigma_n}{\sigma_p}\right)^{\left(\frac{K_i - K_p}{K_n - K_p}\right)}, \tag{4.1}$$

where:

- σ_i is the implied volatility for strike $K = i$;
- σ_p is the previous implied volatility measured in the implied volatility/option delta space (that is, the implied volatility contributed by the dealer/brokers whose strike is the closest lower to $K = i$);
- σ_n is the next implied volatility measured in the implied volatility/option delta space;
- K_i is the strike $K = i$;
- K_p is the strike equivalent of implied volatility σ_p;
- K_n is the strike equivalent of implied volatility σ_n.

[2] The stochastic volatility inspired (SVI) models are techniques to fit the market's implied volatility smile that guarantee the absence of statistic arbitrage, as developed in Gatheral (2004).

It should be highlighted that after this transformation is applied, the implied volatility smiles are no longer arbitrage-free – that is, self-financing butterflies are observed in the reference prices published. For the purpose of the estimation of risk-neutral densities, these self-financing butterflies translate into negative probabilities when the Breeden-Litzenberger result (3.4) is applied.

It is not possible to recover the (arbitrage-free) implied volatility/option delta smiles from the information published without making additional assumptions about unobserved variables, as the non-linear nature of the delta function often results in missing information when backtracking equation (4.1).

Nevertheless, irrespective of the existing arbitrage opportunities, the implied volatility information published by B3 still has the significant advantage of synchronicity, as it based on end-of-day contributions from market participants. That is, the fact that contributors are required to provide several complete implied volatility smiles regardless of deals having occurred at every tenor mitigates the issue that implied volatility estimates based on actual prices might be non-synchronous and reflect different market conditions - which is a frequently mentioned issue in the literature.

Given that we believe that the synchronicity feature out-weights the drawbacks from eventual price arbitrages, rather than relying on a different data source for implied volatility smiles, we chose to fix arbitrage issues by applying the smoothing interpolation methodology described in Section 3.2 when treating the implied volatility/option strike smiles.

4.1.1 Constant maturities

By taking the decision to work with data coming from B3, another aspect of the Brazilian market has to be considered. Differently from other countries, the Brazilian onshore products listed on B3 such as interest rate futures and options are traded with constant expiry dates, rather than constant maturities. That is, instead of trading derivatives with maturities clustered as 3 months, 6 months, and so on, contracts will have fixed maturities of 02 January 2020, 01 February 2020, and so on.

Given this, the time-to-expiration of every contract will change daily. Given that the goal of the work is to conduct an analysis of the moments of the RNDs, and that using constant expiry dates could lead to issues related to the optimal roll-over dates for such analysis, it was decided to work with constant-maturity tenors. More specifically, 1-month, 2-month and 3-month (21, 42 and 63 business days, respectively) constant maturity risk-neutral probability density functions were chosen as the end-goals of the estimation. Given that the liquidity in the Brazilian FX market tends to concentrate in short-term maturities, we chose not to further extend the analysis into longer-date tenors.

In order to obtain such constant maturity RNDs, it is necessary conduct an interpolation of the interest rates and the implied volatility smiles. These interpolations were conducted via straightforward cubic splines of the DI (domestic interest rate) and the cupom cambial (foreign interest rate) curves. In the case of the implied volatility smiles, they were only interpolated for the strikes that were common to the maturities immediately before and after the number of business days to be interpolated.

For example, when building a 42-business days (2-month) constant maturity implied volatility smile on 12 April 2017, the closest B3 dates available were 01 June 2017 (33 business days) and 03 July 2017 (54 business days). The common strikes between the two maturities ranged from 2825 to 3800, and implied volatilities above or below this interval were discarded. Each point was interpolated using a cubic spline, and a graphical representation of the result is shown in Figure 4.

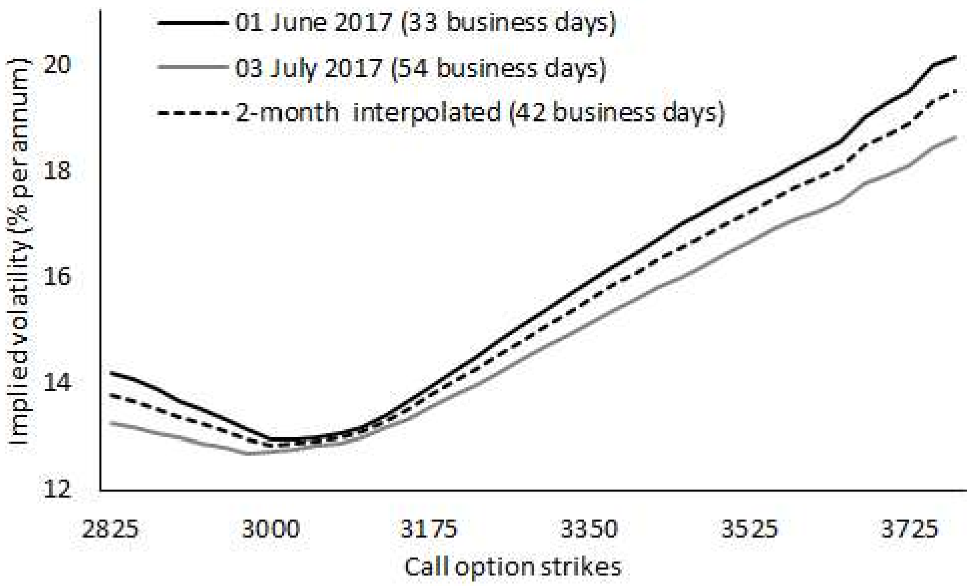

Figure 4 – Interpolation of a 2-month constant volatility smile from fixed-date maturities.

4.1.2 Smoothing spline

After the volatility smiles and interest rates are interpolated for the constant maturities, a smoothing spline interpolation is applied to the implied volatility smiles. The goal is to simultaneously generate more implied volatility estimates (in order to apply the result of Equation (3.4)), correct eventual convexity issues arising from the transformation discussed in Section 4.1, and reduce the noises in the smile that and may generate 'kinks' in the probability density function.

As discussed in Chapter 3, there is a trade-off between the informational content of the estimated RND and its smoothness: if the implied volatility smiles are excessively smoothed out, the analysis of the density functions might become unreliable simply because the "true" RND was shaped into a generic figure. Having this in mind, it was chosen to start off with a low level of smoothness that can be gradually increased up to a desired point.

The smoothing spline function applied is shown by Equation (3.5). The weight parameter w is used to assign relative importance to different points in the curve.

For instance, in Bliss and Panigirtzoglou (2002), the weights adopted are the options' vegas. For a given set of options, their vegas will be higher when closer to the ATM option and increasingly lower when closer to deep ITM and OTM options. This weighing structure aims to mitigate distortions caused by the low liquidity of away-from-the-money options relative to those closer to ATM.

However, given that the information provided by B3 is built from surveys with market participants - rather than actual negotiations -, applying a weighting structure that reduces the importance of away-from-the-money options might lead to loss of relevant information.

Thus, we opted to use the following structure: $w = 100$ for the ATM, deepest ITM, and deepest OTM options; and $w = 1$ otherwise. With this, it is possible to smooth out the implied volatility structure while preserving its shape around the point of highest liquidity (ATM) and at the edges of the implied volatility smile.

As per the smoothing parameter λ, an empirical observation of the shape of the estimated density functions suggests that an initial setting of $\lambda = 1 \times 10^{-4}$ when interpolating the volatility smile provides a reasonable balance between smoothness and preservation of the original shape of the curve.

After an initial smoothing interpolation, it is necessary to assess whether eventual convexity issues in the smile still persist. This is done by converting the implied volatilities into prices and checking whether self-financing butterfly structures can be built. If so, the λ parameter is recursively increased as: $\lambda = \lambda + 1 \times 10^{-5}$, until the smoothing interpolation is sufficient to guarantee that the convexity condition is met on all points.

Then, it is trivial to interpolate implied volatility points along the smile for a larger collection of strikes - in our case, an implied volatility is estimated every 0.001 BRL per USD or 1 future point. These will be input into the Black-Scholes formula and converted into European option prices.

In order to illustrate the process, the interpolated smile shown on Figure 4 was adjusted using the weighing structure described above and considering $\lambda = 1 \times 10^{-4}$. The obtained smile is shown on Figure 5 below.

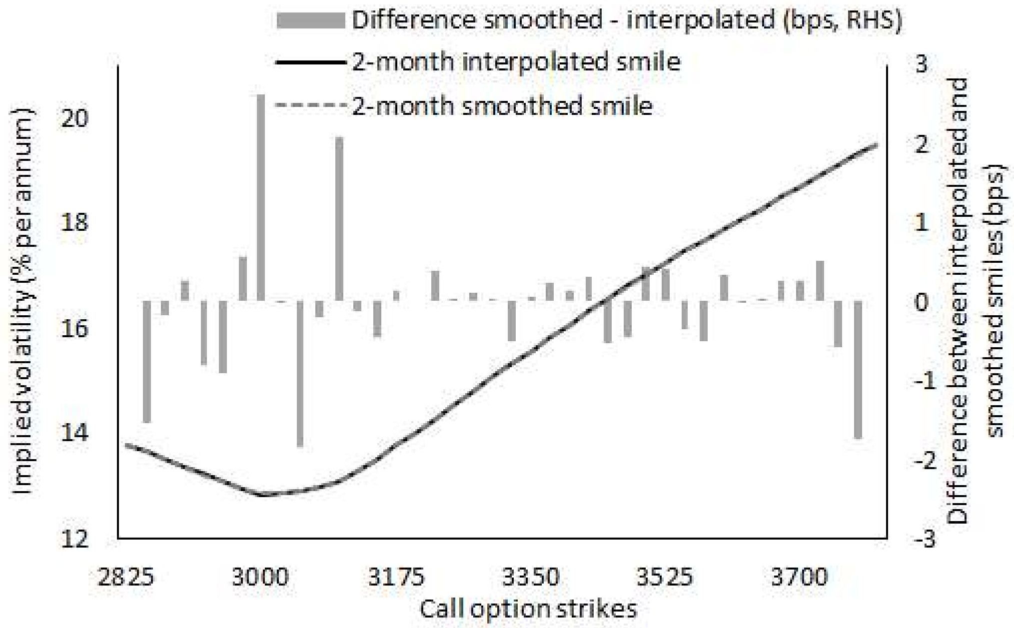

Figure 5 – Interpolated and smoothed implied volatility smiles plotted against changes in estimates due to smoothing (in bps, RHS).

Although the interpolated and smoothed implied volatility smiles appear to be the same when compared, their differences become much more apparent when the risk-neutral probability density function is estimated by applying the Breeden-Litzenberger result, as shown on Figure 6.

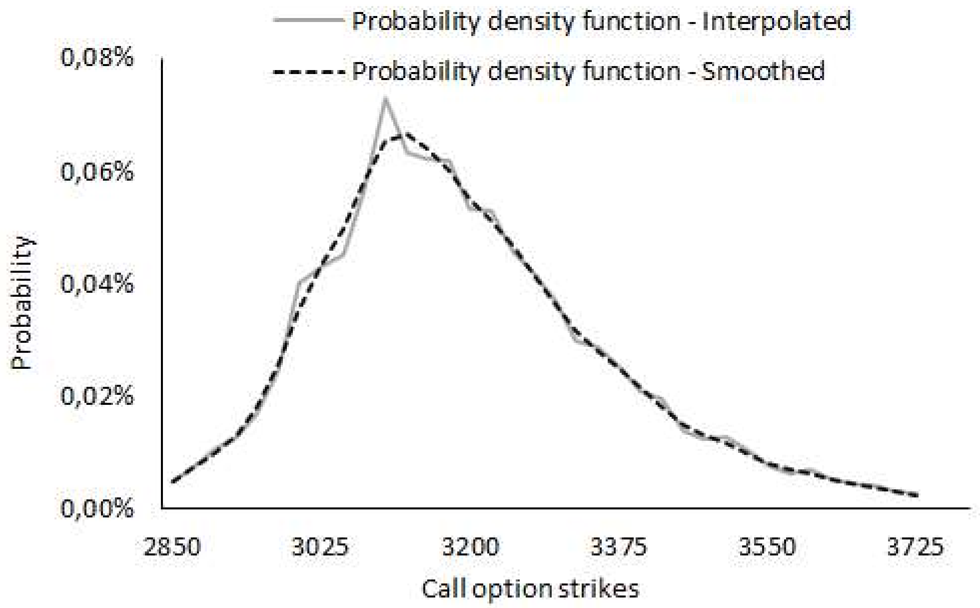

Figure 6 – Risk-neutral probability density functions estimated using the implied volatility smiles of Figure 5.

From this point onwards, the extraction of the RND will depend on the technique adopted.

4.2 Non-parametric estimation of RND with GEV tails

It is possible to extract both the risk-neutral PDF and CDF from the collection of implied volatilities calculated above. One characteristic of the non-parametric framework is that the RND will only be estimated up to the last strike with a corresponding implied volatility in the smile. Given that this approach utilizes the result shown on equation (3.4), the probability density function will be truncated for strikes that are deep away-from-the-money.

While it would be simpler to extrapolate the implied volatility, this would be a very strong assumption about the market's preferences in these extreme scenarios. Rather, we fitted the distribution tails according to the GEV distribution, as described in Section 3.3. A few particularities of the database provided by B3 will be discussed as follows.

As it is possible to see from Figure 6, the implied volatility smiles provided by B3 translate into relatively complete risk-neutral density functions. The fact that the bulk of the RNDs are already determined means that the inclusion of the tails will have a smaller impact on its final shape and the subsequent calculation of its moments.

In order to find parameters that best fit each GEV tail, four significance levels α_1 e α_2 that are sufficiently close to the edges of the CDF are chosen. They will be serve as equality restrictions in Equations (3.9) and (3.10) and as inputs of the objective function of the optimization routine (3.11). An excessive distance between α_1 and α_2 in each tail will have a negative impact on the performance of the optimization; therefore, a relative distance of 3% (that is, $\alpha_{2R} = \alpha_{1R} - 3\%$ and $\alpha_{2L} = \alpha_{1L} + 3\%$) was adopted. All the optimizations were conducted on Matlab using the fmincon (constrained optimization) function.

By choosing the significance levels empirically and enforcing conditions (3.9) and (3.10) in the maximization, the cumulative density of the tails estimated will necessarily complement the density of empirical distributions so the combination integrates to one. That is, if a level of $\alpha_{1R} = 98\%$ is chosen, the right GEV tail will have a cumulative density of 2%. This restriction prevents the optimization from finding parameters that lead to tails whose density is different from that established by market prices.

The CDF that is equivalent to the smoothed RND depicted in Figure 6 is shown in Figure 7 below. In this case, the significance levels used were $\alpha_{1L} = 2\%$, $\alpha_{2L} = 5\%$, $\alpha_{1R} = 98\%$, and $\alpha_{2R} = 95\%$.

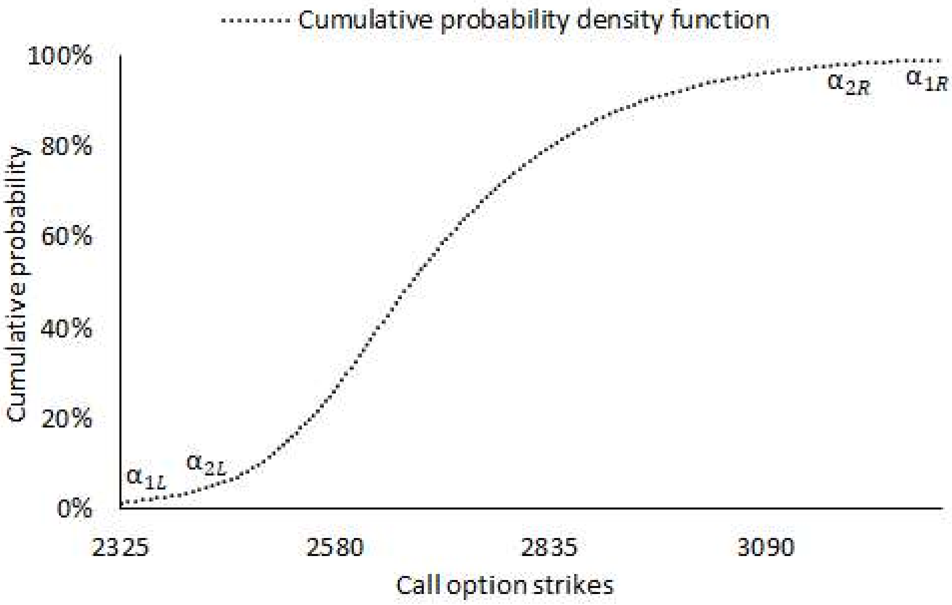

Figure 7 – Cumulative probability density function and significance levels.

The optimization routine is performed first for the right tail, resulting in the parameters $\Theta_R^* = [\mu_R^*, \sigma_R^*, \xi_R^*]$ that minimize the objective function. Then, the routine for the left tail is conducted with the additional restriction that $\xi_R^* < \xi_L < 0$. By adopting this additional restriction in the parameter that determines the behaviour of the GEV tail, the left tail of the RND will decrease towards zero more rapidly than the right tail.

In the case of the risk-neutral probability density function of USD/BRL, the right tail is associated with scenarios of a depreciation of the Real against the US Dollar. Therefore, the inclusion of the additional restriction when optimizing the left tail in practice will attribute higher importance to tail risks of a rise in USD/BRL - which is well supported by Emerging Markets currency performance and FX hedging patterns using options.

After estimating the parameters $\Theta_L^* = [\mu_L^*, \sigma_L^*, \xi_L^*]$, both probability density functions can be estimated according to Equation (3.7). The result of the optimizations for the smile of Figure 6 and the CDF of Figure 7 is shown in Figure 8 below.

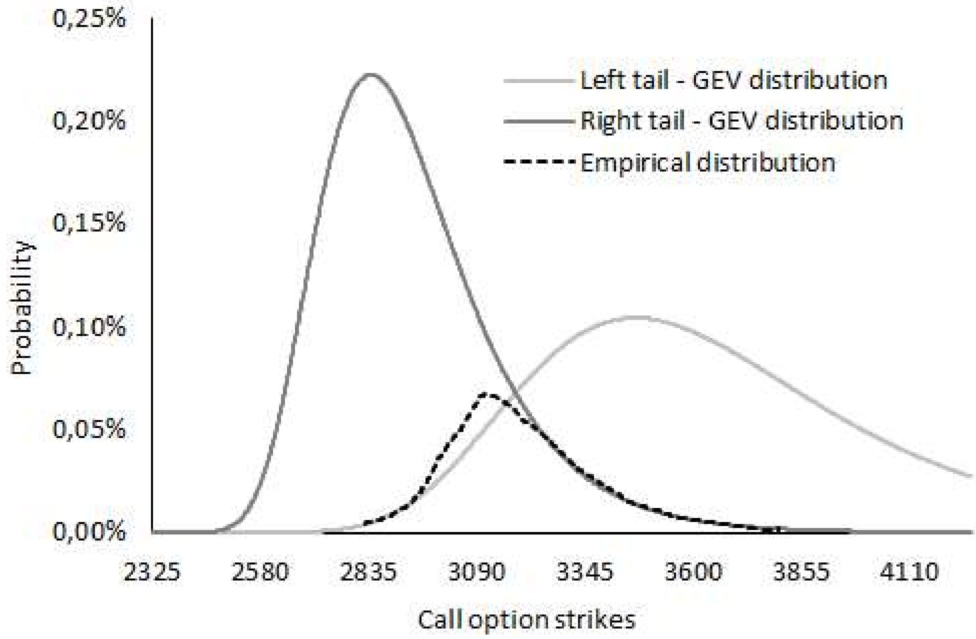

Figure 8 – Empirical RND plotted and GEV probability density functions estimated with optimal parameters.

The original (truncated) risk-neutral density is connected to the GEV tails at the significance level where the difference between the original PDF and the GEV PDF is the smallest. Then, the distribution is re-based so that the area below the RND is equal to one, creating the complete risk-neutral density function shown on Figure 9.

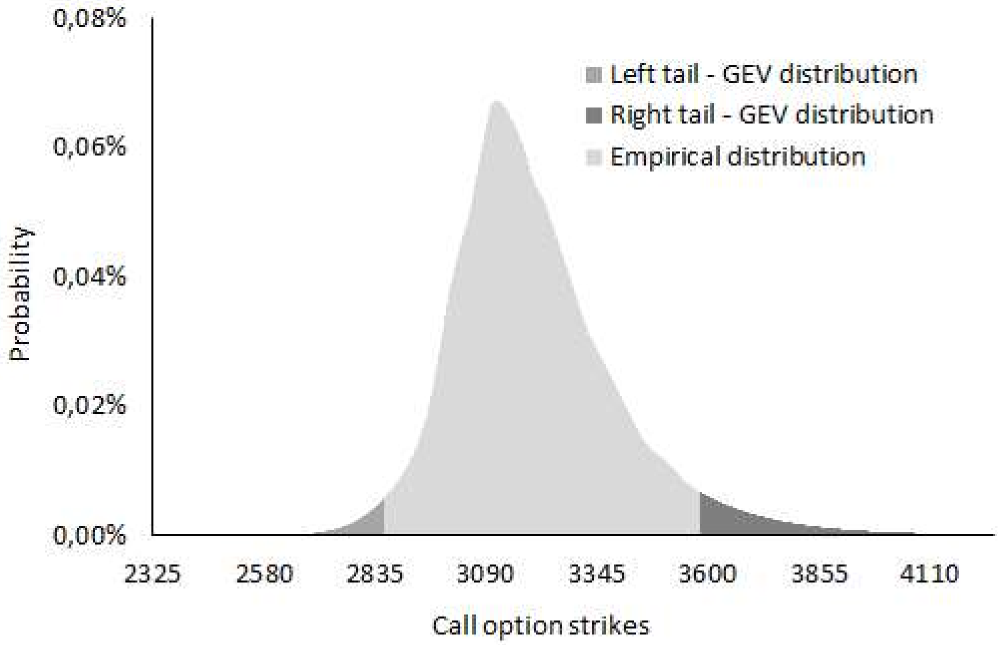

Figure 9 – Complete risk-neutral density function with GEV tails.

4.3 Parametric estimation of RND

The estimation of the RND under a parametric framework revolves around finding a set of optimal parameters that minimize the squared difference between observable option prices and the theoretical prices under a given distribution. In our case, we will focus on the GB2 and MLN distributions, and use the reference prices published by B3 as the observable option prices that the optimization will take as reference.

In the case of the GB2 distribution, the optimization routine will minimize the objective function described by Equation (3.16), subject to the martingale restriction described in Equation (3.14). The risk-neutral probability density function is obtained by inputting the optimal parameters $\theta^* = (a^*, b^*, p^*, q^*)$ into Equation (3.12), which describes the probability density function of the GB2 distribution. The RND estimated with the same dataset of its non-parametric equivalent is depicted on Figure 10.

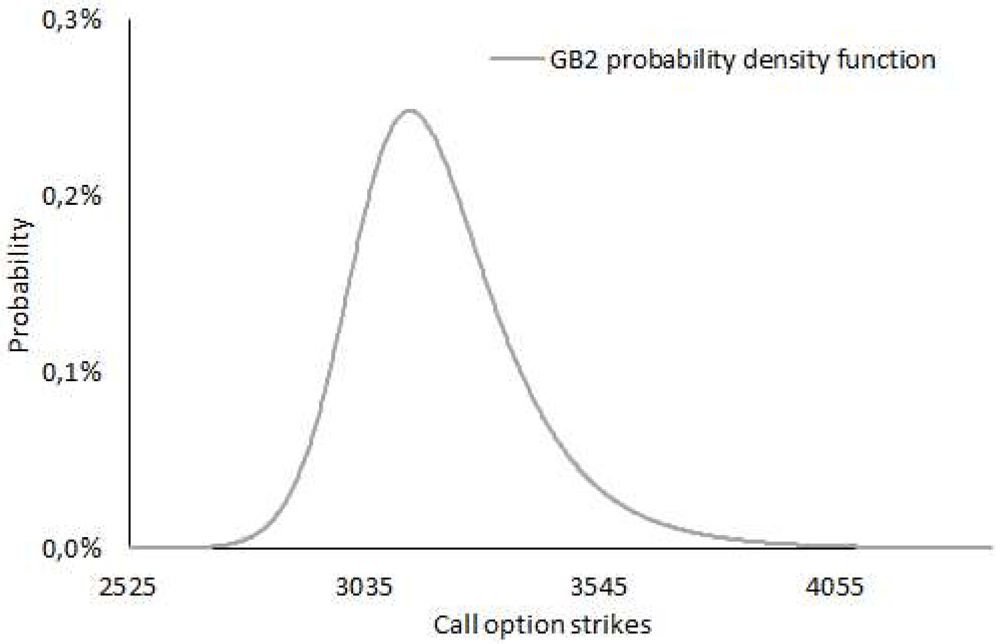

Figure 10 – Risk-neutral probability density function under the GB2 distribution.

The RND of Figure 10 can be compared to the non-parametric result shown in Figure 9 and the result is shown on Figure 11.

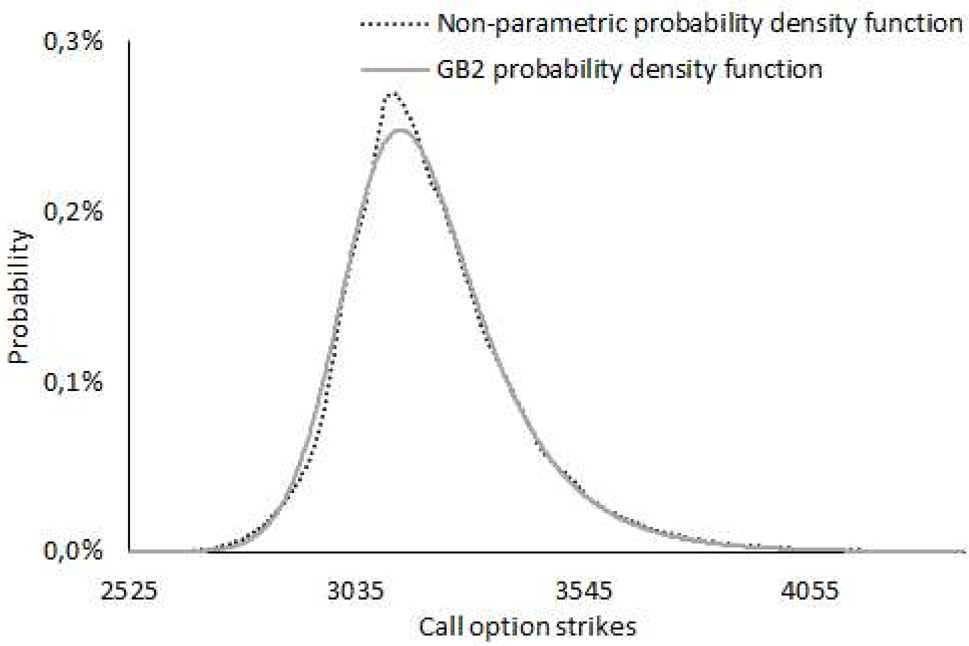

Figure 11 – Comparison between the GB2 and non-parametric RNDs.

The estimation of the RND under the MLN distribution is very similar to the procedure described for the GB2 distribution: the optimization routine will minimize objective function (3.21) subject to the martingale restriction (3.22). The risk-neutral probability density function is described by Equation (3.18) using the optimal parameters $\theta^* = (\mu_1^*, \sigma_1^*, \mu_2^*, \sigma_1^*, \alpha^*)$. Differently from the GB2, it is also possible to estimate each log-normal PDF. Each density function and the combined RND are shown in Figure 12.

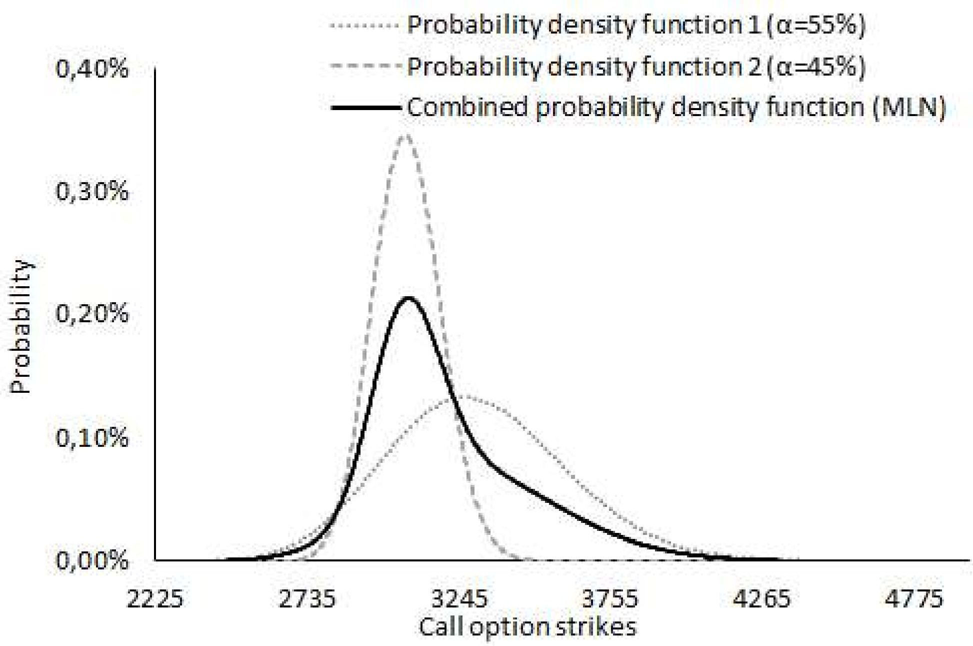

Figure 12 – Risk-neutral probability density function under the MLN distribution.

The RNDs of Figure 11 can be compared to the final (combined) result shown in Figure 12, and the result is shown on Figure 13.

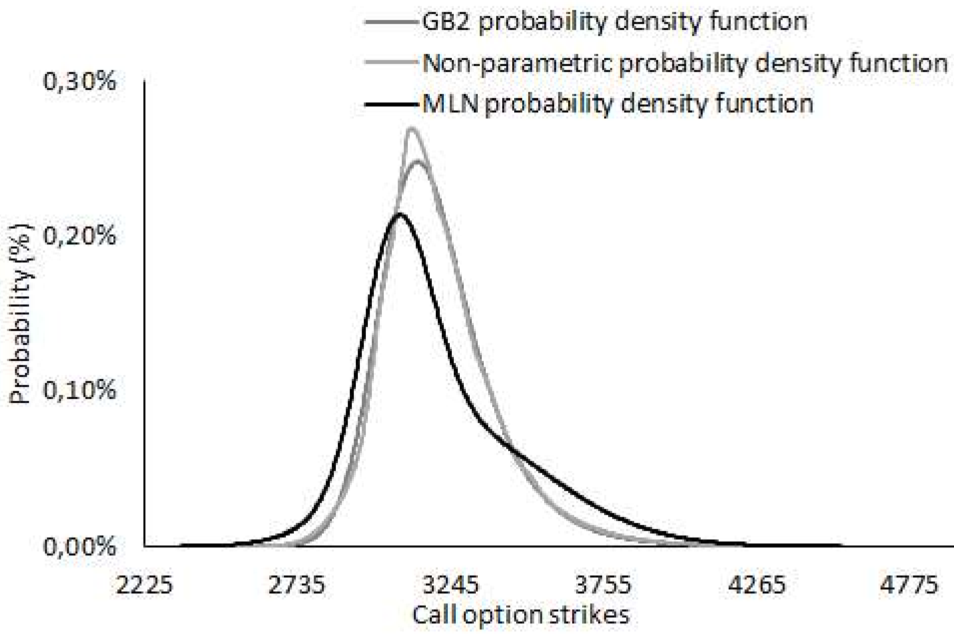

Figure 13 – Comparison between the GB2, non-parametric and MLN RNDs

Each framework can be more properly compared to each other by estimating the moments of the RNDs.

4.4 Estimating the moments of the RNDs

Given that the non-parametric framework does not rely on a specific function form for the probability density function, its moments are calculated via numerical integration on Matlab. The mean μ_{np}, variance σ^2_{np}, skewness S_{np} and kurtosis K_{np} are given by the following equations:

$$\mu_{np} = \int_{-\infty}^{\infty} K f(K) \mathrm{d}K \tag{4.2}$$

$$\sigma^2_{np} = \int_{-\infty}^{\infty} (K - \mu_{np})^2 f(K) \mathrm{d}K \tag{4.3}$$

$$S_{np} = \frac{\int_{-\infty}^{\infty} (K - \mu_{np})^3 f(K) \mathrm{d}K}{(\sigma_{np})^3} \tag{4.4}$$

$$K_{np} = \frac{\int_{-\infty}^{\infty} (K - \mu_{np})^4 f(K) \mathrm{d}K}{(\sigma_{np})^4} \tag{4.5}$$

One of the advantages of parametric estimations of risk-neutral densities is that they allow for the direct estimation of its moments through the parameters, without the need of using numerical integration. Also, given that the martingale condition was included as an equality condition in the optimization routines, the mean of the parametric RNDs necessarily corresponds to the ATM forward, requiring the estimation only of the variance, skewness and kurtosis of the distributions.

In the case of the GB2 distribution, Jondeau, Poon and Rockinger (2007) note that, for a given strike X:

$$E[X^n] = \frac{b^n B(p + \frac{n}{a}, q - \frac{n}{a})}{B(p,q)}, \tag{4.6}$$

for $n < aq$ and with $B(p,q) = \Gamma(p)\Gamma(q)/\Gamma(p+q)$.

As per the MLN estimation, Liu et al. (2007) show the equivalent result:

$$E[X^n] = \theta\alpha_1^n e^{\frac{1}{2}(n^2-n)\beta_1^2 T} + (1-\theta)\alpha_2^n e^{\frac{1}{2}(n^2-n)\beta_2^2 T}, \tag{4.7}$$

where:

$$\alpha_i = lnS + (\mu_i - \frac{1}{2}\sigma_i^2)T \,, \beta_i = \sigma_i\sqrt{T}, \tag{4.8}$$

as shown in Bahra (1997). Also, μ_i and σ_i are the parameters that characterize each log-normal distribution, S is the spot USD/BRL rate, T is the time-to-expiry and θ is the weighting parameter for the density mixture.

With this, is is possible to calculate the parametric variance σ_p^2, parametric skewness S_p and parametric kurtosis K_p by applying the results (4.6) and (4.7) in the following equations:

$$\mu_p = E[K], \tag{4.9}$$

$$\sigma_p^2 = E[K^2] - E[K]^2, \tag{4.10}$$

$$S_p = \frac{E[K^3] - 3(\mu_p\sigma_p^2) - \mu_p^3}{\sigma_p^3}\,, \tag{4.11}$$

$$K_p = \frac{E[K^4] - 4(\mu_p E[K^3]) + 6(\mu_p^2 E[K^2]) - 3(\mu_p^4)}{\sigma_p^4}\,. \tag{4.12}$$

The process of extraction of the RND can be summarized as the following steps:

1. Download from B3 the information that will be necessary to conduct the estimation of the RND for a given day: spot USD/BRL, local and foreign interest rate curve settlements (for all maturities), and implied volatility/option strike smiles[3].

[3] The daily information from January 2014 to February 2019 was downloaded directly from B3's FTP server using a script written in R language. The files come in a .txt file that took around 10 seconds/day to download.

2. Interpolate the interest rate curves and the implied volatility smiles in order to obtain 1-month, 2-month and 3-month constant maturity implied volatility smiles.
3. Apply the smoothing spline in order to correct for eventual convexity issues (that is, self-financing butterflies) and to smooth out excessively noisy parts of the risk-neutral probability density function.
4. Use this interpolation structure to evaluate the implied volatility smile in a more granular grid of strikes, thus obtaining a much more complete set of observations within the smile.
5. Apply the standard Black-Scholes formula using the interpolated implied volatilities, interpolated interest rates and other parameters in order to obtain a collection of options prices (expressed as BRL per option).

Now, the non-parametric estimation of the RND can be summarized as the following steps:

1. Estimate the risk-neutral PDF and CDF by applying the Breeden-Litzenberger result in collection of prices. Unless the implied volatility smile is extraordinarily complete, this estimation will lead to a truncated empirical RND (missing the tails).
2. Chose two significance levels for each of the left and right tails that are available in the empirical CDF. Assess the equivalent strikes and probabilities in the PDF associated with these significance levels. They will serve as 'connecting points' to the GEV tails.
3. Run an optimization routine for the right tail in order to find the three parameters for the GEV distribution that minimize the squared difference between the estimated RND's PDF and CDF in the two significance levels.
4. Run the same optimization routine for the left tail, adding the additional restriction that the ξ parameter has to be higher than the one of the right tail.
5. With the set of parameters that optimize the functions for both tails, estimate the PDF values for a larger set of strikes.
6. Determine which among the two significance levels has the lowest absolute difference between the truncated RND and the GEV tail – this will be the cut-off point between the two curves. Use the GEV distribution before/after these points, creating a complete distribution.
7. Re-weight the collection of points so that the area below the PDF is equal to one.

8. Estimate the mean, variance, skewness and kurtosis of the RNDs through numerical integration.
9. In order to improve the stability of the RNDs estimated, use the optimal parameters of optimization at time t as the initial guess of the optimization routine at time $t + 1$.

Finally, the process for the parametric estimation of the RND is summarized as the following steps:

1. With a collection of observed prices in hands, run an optimization routine in order to find the parameters that simultaneously: (a) minimize the squared difference between empirical prices and theoretical prices under the GB2/MLN distribution; (b) meet the martingale condition that the mean of the RND is equal to the at-the-money forward.
2. Estimate the mean, variance, skewness and kurtosis of the RNDs by using the specific analytic formulas of each distribution.
3. In order to improve the stability of the RNDs estimated, use the optimal parameters of optimization at time t as the initial guess of the optimization routine at time $t + 1$.

The results of the estimation of the moments will be discussed in depth in Chapter 5.

5 Results and trading with the moments of the RND

The goal of this chapter is to discuss the results of the estimation of the moments of the RNDs estimated using the methodology described in Chapter 4 and to provide a practical application for the moments.

5.1 Analysing the estimated moments of the RNDs

The procedures described in Section 4.4 were applied to daily data of USD/BRL market between January 2014 and February 2019, for a total of 1,264 business days. The Table 2 below summarize the median variance, skewness and kurtosis for 1-month, 2-month and 3-month constant maturities of the non-parametric risk-neutral density (the means of the RNDs are not depicted given that the martingale equality restriction included in the optimization routines required it to be equal to the ATM forward).

Table 2 – Median of the descriptive statistics for the non-parametric RND.

Statistic	**2015**	**2016**	**2017**	**2018**	**2019**
Variance (1-month)	34,978	34,341	14,648	26,633	24,769
Variance (2-months)	69,000	72,166	33,077	59,658	48,825
Variance (3-months)	108,555	106,713	53,740	94,044	71,595
Skewness (1-month)	0.75	0.71	0.95	0.87	0.96
Skewness (2-months)	0.98	0.94	1.23	1.07	1.02
Skewness (3-months)	1.18	1.11	1.39	1.21	0.96
Kurtosis (1-month)	4.47	4.36	5.05	4.93	5.35
Kurtosis (2-months)	5.35	5.18	6.30	5.85	5.90
Kurtosis (3-months)	6.19	5.86	7.10	6.22	5.50

The scale of the variance is related to the fact that B3 publishes the strikes of USD/BRL options in terms of future points (*i.e.*, a strike of USD/BRL at 4.00 is depicted as 4,000). In terms of the data, it is possible to see that there is an increase in variance, skewness and kurtosis when the time-to-expiry increases - reflecting higher uncertainty (expressed by the increase in variance and kurtosis) but also a higher perceived risk of a rise in USD/BRL (expressed by the increase in skewness). The increase in variance across all estimates also shows that there is a stochastic process that prevents the existence of calendar spreads. The only exception is the skewness in 2019, where the coefficients remained about constant for all maturities.

The Tables 3 and 4 below show the equivalent descriptive statistics for the GB2 and MLN frameworks.

Table 3 – Median of the descriptive statistics for the GB2 RND.

Statistic	**2015**	**2016**	**2017**	**2018**	**2019**
Variance (1-month)	33,557	31,806	14,362	26,508	22,832
Variance (2-months)	68,734	66,168	31,733	57,535	47,439
Variance (3-months)	100,433	100,710	50,695	87,911	73,509
Skewness (1-month)	0.74	0.79	1.01	0.90	0.80
Skewness (2-months)	1.02	0.97	1.15	1.11	0.96
Skewness (3-months)	1.28	1.18	1.34	1.32	1.16
Kurtosis (1-month)	4.81	4.90	5.99	5.59	4.98
Kurtosis (2-months)	5.72	5.40	6.34	6.12	5.43
Kurtosis (3-months)	6.94	6.28	7.26	7.09	6.13

Table 4 – Median of the descriptive statistics for the MLN RND.

Statistic	**2015**	**2016**	**2017**	**2018**	**2019**
Variance (1-month)	40,681	41,744	26,659	40,225	37,272
Variance (2-months)	79,550	77,843	49,178	76,260	68,530
Variance (3-months)	119,302	117,009	78,517	123,227	109,595
Skewness (1-month)	0.51	0.53	0.60	0.60	0.54
Skewness (2-months)	0.73	0.72	0.73	0.76	0.70
Skewness (3-months)	0.94	0.91	0.91	0.98	0.90
Kurtosis (1-month)	3.36	3.39	3.61	3.61	3.51
Kurtosis (2-months)	3.91	3.90	4.05	4.08	4.02
Kurtosis (3-months)	4.54	4.53	4.71	4.82	4.62

The GB2 estimation of the RND yielded very similar descriptive statistics to those observed for the non-parametric estimation: the variance and skewness coefficients stood around the same levels for all periods. However, the RND estimated with the GB2 distribution showed a higher level of kurtosis for the period of 2015-2018 relative to the non-parametric framework.

Regarding the MLN estimation, the level of variance stood higher than the other two frameworks for all periods and maturities analysed, while the skewness and kurtosis were lower than the non-parametric and GB2 estimations. The estimated 1-month MLN and GB2 kurtosis for the period of 2015-2017 is shown in Figure 14.

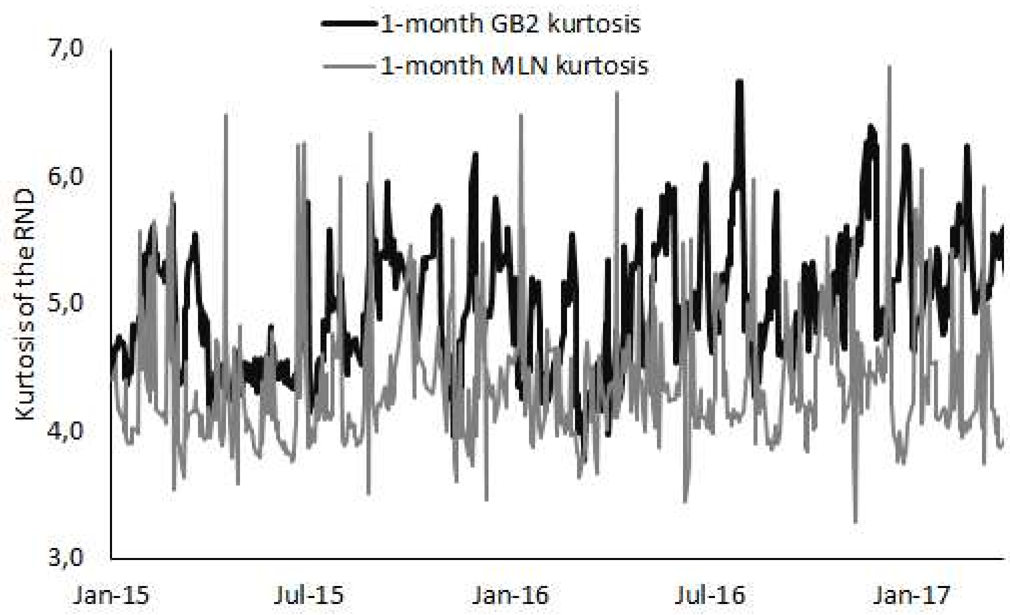

Figure 14 – Comparison between the estimated 1-month kurtosis coefficients under the MLN and GB2 frameworks.

Even though some of the estimated moments appear to be similar for the non-parametric and GB2 distribution, the stability of the estimations will be very different. Figure 15 shows the daily skewness estimates extracted from the 2-month RNDs under GB2 and non-parametric framework.

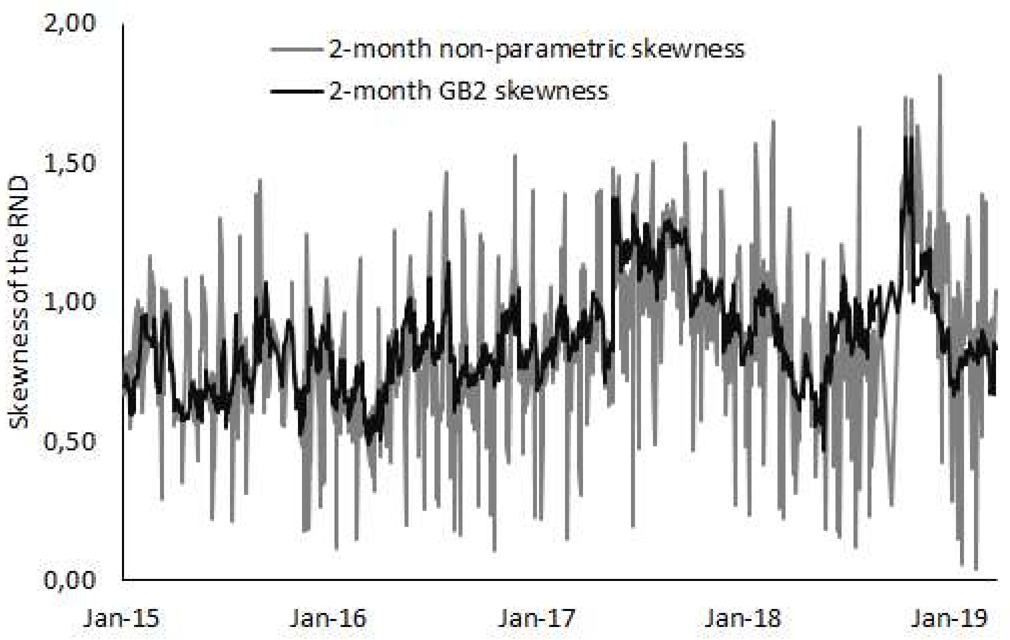

Figure 15 – Comparison between the estimated 2-month skewness coefficients under the non-parametric and GB2 frameworks.

Even though the overall pattern and median values across the period analysed appear to be similar, the moments of the risk-neutral density function estimated using a non-parametric framework are much more volatile and unstable than those estimated using a parametric approach. This conclusion can be extended to all the moments estimated and also when compared to the MLN framework.

This very important characteristic of the non-parametric framework has to be taken into consideration when choosing an approach to extract a risk-neutral density function. Should the application of the RND depend on the stability of the estimation - which is the case of the present work, as will be discussed later in this Chapter - the non-parametric framework may yield sub-optimal results when compared to parametric approaches.

5.2 Motivation of the application

The applications of an asset's risk-neutral density functions mentioned in the literature are numerous, ranging from better pricing of exotic derivatives to assessing the market's complete distribution of probabilities and applying that to policy making and/or investment decisions.

In the case of this study in particular, the goal is to assess whether the RND of USD/BRL can offer valuable information in order to make investment decisions for trading spot USD/BRL. The rationale behind this implementation will be broken down as follows.

The options market allows for market participants to express their views regarding an underlying asset's future performance - views that are reflected in the risk-neutral probability density function. It is reasonable to assume that information shocks that alter the asset's fundamentals will also affect (to different extents) the market's future expectations about that asset and therefore its risk-neutral densities.

The spot FX market and its options are interconnected many ways. Besides the relationship established by the asset's fundamentals and market's expectations mentioned above, there are more technical reasons as well. Given that some option positions and portfolios require some sort of interaction with the FX market (like delta hedging, for instance), part of the volume traded in the spot FX is due to investors and/or market dealers conducting this kind of operation.

At the same time, the impact of the options market in its underlying FX market cannot be exaggerated: it would be naive to assume that risk-neutral densities provide a roadmap to the future of the underlying asset, or that a currency's performance is solely driven by the behavior of its options (a case of 'the tail wagging the dog'). The hypothesis of this work's application does not rely on such overstatement, but rather builds on the fundamental relationship between the risk-neutral density functions and its underlying asset.

Part of the changes in a currency value are not purely explained by changes in its fundamentals, but are also consequence of aspects such as liquidity shifts, large investors' activity 'dragging' the market, or even due to intricate reasons more related to behavioral finance. If we assume that the majority of shocks of this type tend to be temporary and subside with time as the market retraces towards its fundamental value, the risk-neutral density may serve as a guide or a compass to detect if such type of market deviation is taking place.

In other words, given that the impact of an information shock in the fundamentals of a certain asset should be reflected both at present value (in our case, the spot FX) and in its expectations (the RND), the risk-neutral density may be used to detect 'unexplained' market deviations in spot USD/BRL that have a higher likelihood of being faded out in the short-term. If this hypothesis is correct, the appropriate detection of this type of behaviour prior the subsequent convergence towards fundamentals would generate a trading opportunity.

Building on this hypothesis, the goal is to propose a model that can successfully identify when such type of pattern occurs, and trade the spot USD/BRL using the signals.

The solution proposed in this work does not intent to be 'ultimate' FX trading tool or the best way to make use of the information contained in the RNDs. Rather, it aims to be a comprehensive and straightforward way to test whether the informational content of the RNDs can be used to improve the decision taking for trading.

5.3 The model

In what follows we are going to use the following notation:

- closing spot rate level, BRL per USD, on date t: S_t,
- percent return between closing levels at dates t and $t-1$:

$$R_t = \frac{S_t}{S_{t-1}} - 1\,,$$

- estimated moment[1] on date t: M_t.

We are going to prescribe the following model:

$$R_t = \beta_1\, M_t + \beta_2\, M_{t-1} + \varepsilon_t, \tag{5.1}$$

where ε is considered to be Gaussian with zero mean. The regression is computed on a daily basis on a 42 business days (2 months) rolling window. A smaller time window is

[1] This moment can be variance, skewness or kurtosis, estimated under non-parametric, GB2 or MLN frameworks.

preferred, as adding too much data points would incur the risk of mixing different market regimes, reducing the effectiveness of the estimation.

The model (5.1) was created as a way to establish a linear relationship between the distributions' moments and the returns of USD/BRL, as new information shocks are incorporated both in the spot market and in the RND. The β parameters accommodate the inter-temporal differences in these informational shocks. If the theoretical motivation of the exercise is correct, the unexplained returns (related to 'non-fundamental' activity) can be measured in a straightforward manner, also taking into account the mean-reverting nature of markets and the RND moments.

The means of the RNDs were not included among the moments studied given that the martingale condition was enforced as an equality restriction during the optimization routines. The fact that it will necessarily match the ATM forward at all times thus provide no information content to measure the type of activity that is the goal of the present work.

With the estimation of $\widehat{R}_t$ in hands, it is possible to retrieve what would be the modelled level of the spot rate at date t:

$$\widehat{S}_t = (1 + \widehat{R}_t) \times S_{t-1}\,. \tag{5.2}$$

It is also possible to create 'confidence bands' with the model's MSE (mean squared error):

Upper band:

$$\text{UB}_t = \alpha \times \sqrt{\text{MSE}_t} + \widehat{S}_t$$

Lower band:

$$\text{LB}_t = \alpha \times \sqrt{\text{MSE}_t} - \widehat{S}_t$$

where the α parameter allows for flexibility in the width of the bands.

The actual levels of USD/BRL is then compared against the model's central estimate and its bands. Given that all the information published by B3 to estimate the RNDs only becomes available after the domestic FX market is closed, any trading signals produced by the model at time t can only be put in place at time $t + 1$. The trading rule can be summarized as follows:

1. When all the information regarding the trading session t is available, estimate the RND, its moments and the USD/BRL return at time t, and input into Equation (5.1). The RND and its moments at time $t - 1$ will already be available.
2. Check whether the parameters of the model are statistically significant: any model signals will only be translated into actual positions if the β_1 and β_2 in Equation

(5.1) are statistically significant (at least) at 10%. This particular significance level was chosen as it is simultaneously a well-established threshold in the literature, and allows for greater flexibility in the regression. If the coefficients are not statistically significant, any signal will be ignored. This aims to reduce noises from signals that have a low informational content.

3. If the regression coefficients are statistically significant, the closing level of USD/BRL at t is compared with the model's upper and lower band. If it is above the upper band or below the lower band, a trading signal is triggered. The positions are built according to the hypothesis that a signal represent a currency's dislocation from fundamentals and that a convergence is likely:
   - (a) if USD/BRL has breached the upper band, the model sells USD (buys BRL);
   - (b) if USD/BRL has breached the lower band, the model buys USD (sells BRL).
4. Based on the signal triggered at t, a position is built at the beginning of the trading session of $t+1$ - more specifically, at 9:15 AM (15 minutes after the domestic market opens). This was chosen in order to allow for market's opening liquidity and volatility to settle (avoiding to use open prices that might be infeasible to build a position in practice). A bid/offer spread of 2 future points (0.002 USD/BRL) was considered as trading cost for each position to be opened and closed.
5. The extraction of the RNDs and updated model estimates/bands is performed daily. Assuming that a position is opened at $t + 1$ (from a signal triggered at t), it will be closed if:
   - (a) the USD/BRL retraces back inside the bands at $t + 1, t + 2, t + 3 \ldots$;
   - (b) the coefficients estimated in the Equation (5.1) are no longer statistically significant at $t + 1, t + 2, t + 3 \ldots$.
6. Following the previous example where a position is opened at $t + 1$ (market opening) and assuming that due to reasons (a) or (b) mentioned above there is no longer a trading signal at $t + 1$ (after market, when information is released by B3), the position is closed at 9:15 AM of the trading session of the following day (in this case, $t + 2$).
7. The performance of the strategy considered in this particular trade would be the trade level of the position at $t + 2$ against the trade level of the position at $t + 1$, minus trading costs.

The trading rule was applied to the RND moments estimated between January 2014 and February 2019 (a total of 1,264 trading days). The α levels used to build the confidence bands of the model were 0.2, 0.4, 0.6, 0.8 and 1.0. Finally, these confidence bands were applied to the moments estimated for the 3 time horizons estimated (1-month, 2-month, and 3-month constant maturity), resulting in a total of 135 strategies (9 moments using 5 different confidence bands applied for 3 different maturities).

The strategies were ranked and compared according to the following metrics:

- Total return: the total cumulative return of the strategy during the complete period of analysis;
- 'Signal consistency': the number of years with positive cumulative returns when the trading rule is applied;
- Gain/loss ratio: ratio between the median return of trades with positive return and the median return of trades with negative return. This was preferred over the 'hit-ratio' metric, as this ratio accounts for the sizes of gains and losses.

As a graphical example, the Figure 16 below shows the model's central estimate and confidence bands against the actual level of USD/BRL for the year of 2018 using the 2-month constant GB2 skewness and $\alpha = 0.6$. The shaded areas indicate periods where the trading signals were triggered. It is possible to see that the trading windows vary (from daily trades to longer periods).

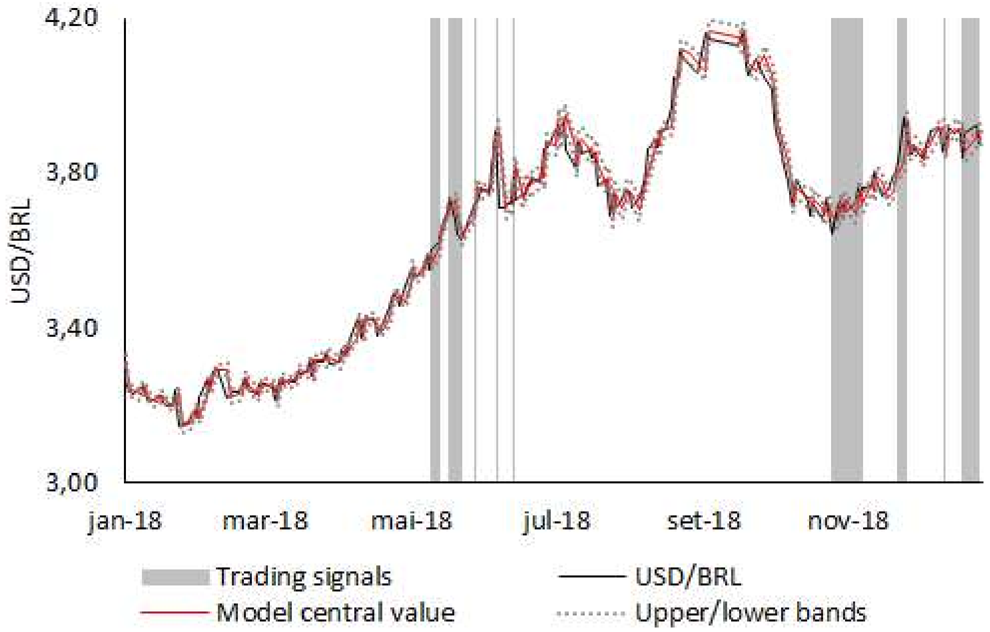

Figure 16 – Spot USD/BRL plotted against the model built using the GB2 skewness and confidence band of $\alpha = 0.6$ for the period of 2018.

The Figure 17 shows the model's result using the 3-month constant non-parametric variance and $\alpha = 0.2$ for a period between 2015 and 2016. It is possible to see that, as a consequence of the smaller value of α relative to the previous example, the model bands are narrower thus leading to an increase in the number of trading signals.

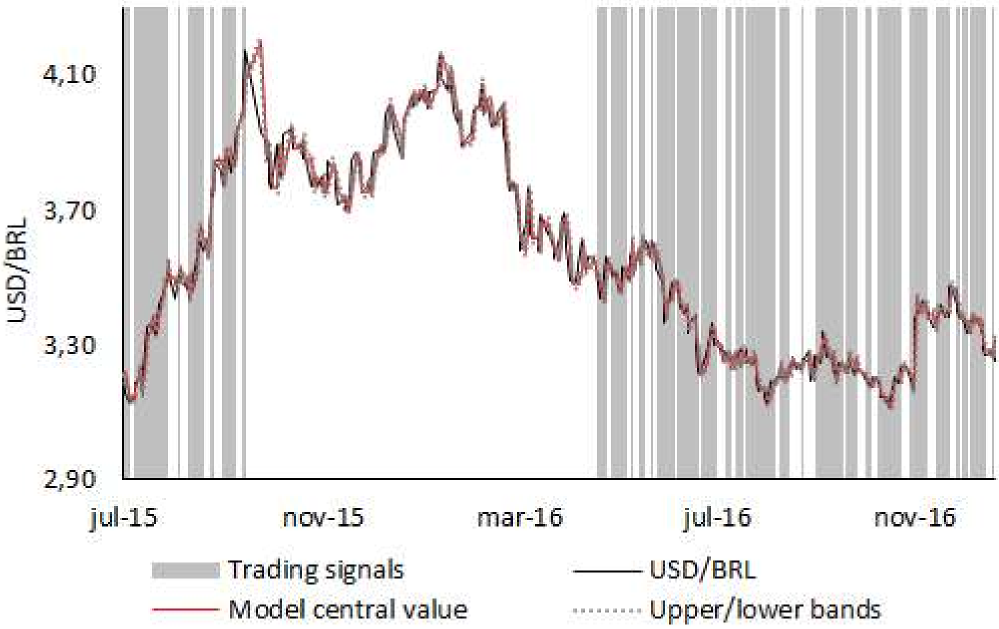

Figure 17 – Spot USD/BRL plotted against the model built using the non-parametric variance and confidence band of $\alpha = 0.2$ for the period of 2015-2016.

The Table 5 below breaks down the three best performing strategies under each of the three time horizon and also depicts the performance of the buy-and-hold strategy.

Table 5 – Comparative performance under different moments, confidence bands and time horizons.

Statistic	**Horizon**	α	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**	**Acum.**
MLN variance	1 month	0.8	-1.1%	21.2%	3.8%	5.7%	1.7%	-1.5%	31.6%
MLN variance	1 month	0.6	-1.1%	20.3%	1.7%	8.4%	1.4%	-3.6%	28.0%
GB2 variance	1 month	0.8	-2.7%	-2.4%	31.3%	-2.8%	14.0%	-1.4%	36.4%
GB2 skewness	2 month	0.6	4.7%	28.5%	6.5%	3.0%	7.6%	-0.1%	58.7%
Non-parametric variance	2 month	0.6	-1.1%	8.6%	38.8%	13.3%	-7.9%	0.5%	56.3%
GB2 variance	2 month	0.6	4.1%	-1.9%	51.7%	11.1%	-2.4%	-1.9%	64.7%
Non-parametric variance	3 month	0.2	12.8%	10.2%	6.9%	4.3%	-1.5%	4.5%	42.6%
GB2 variance	3 month	0.6	17.4%	-9.0%	51.7%	4.7%	-3.7%	-0.7%	62.5%
GB2 variance	3 month	0.4	18.5%	-10.4%	42.1%	11.6%	-10.7%	-2.8%	46.4%
Buy-and-hold			2.2%	31.9%	-27.5%	-7.2%	14.4%	-2.3%	1.6%

The cumulative performance of the strategies depicted in Figures 16 and 17 is shown in Figure 18. It is possible to see that the strategy based on the 2-month GB2 skewness generated a better and steadier performance with respect to the non-parametric one.

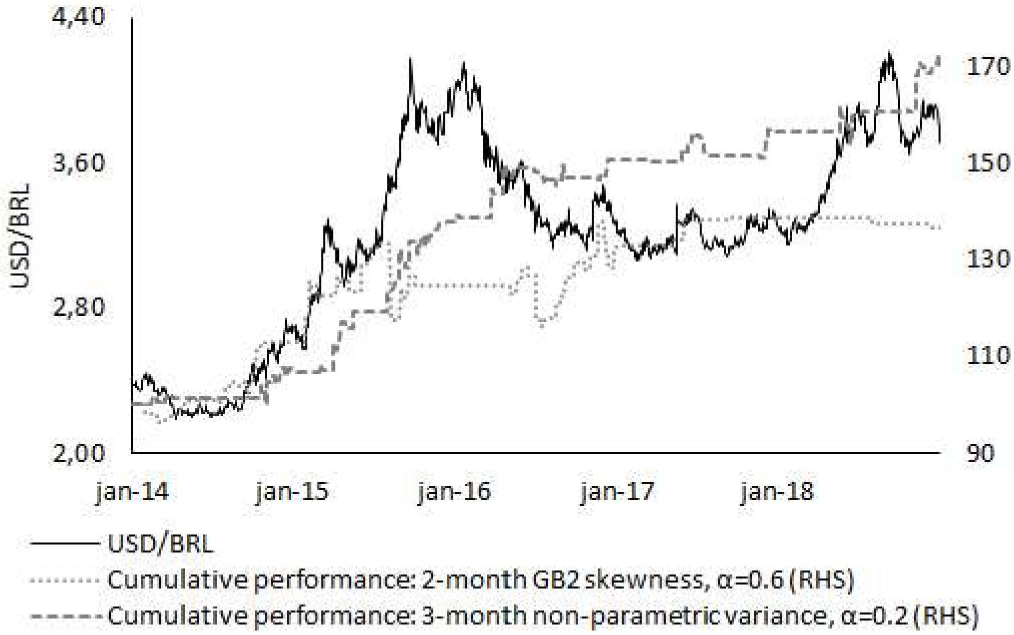

Figure 18 – Spot USD/BRL plotted against the cumulative performance of the two trading strategies shown in Figures 16 and 17.

The return of the buy-and-hold strategy was measured by comparing the closing price of the 1-year USD/BRL onshore future in the first trading session of the respective year against the closing price of the same future at the last trading session of that year. This was considered in order to account for the carry of the position. As per the trading strategies adopted, the carry was accounted by accruing the onshore local currency and USD interest rates throughout the period where the position remained open and incorporating that result into the overall gain/loss of the position.

In terms of the results, all the best performing strategies for the three maturities analysed outperformed the buy-and-hold strategy by a large margin in terms of total returns. Looking among the components, the strategies that used the moments extracted from the 2-month (42 business days) constant RND exhibited the best performance among the three maturities, with the highest average return. In contrast, the strategies using the 1-month (21 business days) RND moments had the lowest average return - which might indicate that the shortest maturity RND has a lower informational content, at least for the purpose of applying the risk-neutral density function for trading. This observation could also be related to the lower vegas of short-term options.

When comparing the different frameworks, the moments extracted from the GB2 distribution appear to yield the best results, as shown by its predominance among the best performing strategies depicted on Table 5. The non-parametric framework shows the worst performance among the three that were tested, when all nine moments and five confidence bands are taken into consideration. This is likely related to the fact that its moments are much more volatile than those of the parametric approaches, which could lead to an increase in false signals and/or positions that are prematurely closed. Given this, the parametric approaches to extracting the RND that were tested provide better final results and should be preferred if the end-goal of the application is trading under a similar framework to what was described in this chapter.

Finally, the use of the variances of the risk-neutral densities provide the highest returns among the moments calculated. This is true for all the maturities analysed. The skewness of the parametric frameworks (GB2 and MLN) also generated results that outperformed the buy-and-hold strategy.

It should be noted, however, that the predominance of the strategies that use the variance of the RND is related to the criteria chosen to rank the strategies. If the hit-ratio is chosen as one of the criteria, as opposed to the gain/loss ratio, the best performing strategies become predominantly those who use the skewness of the distributions - with a lower average cumulative return.

6 Conclusions

The goal of this work was to present a non-parametric technique for the estimation of the risk-neutral probability density function of the USD/BRL, compare it to other parametric frameworks that are frequently mentioned in the literature, and propose a trading application based on the moments of the RND.

A non-parametric estimation has the main advantage of not assuming any particular shape for the RND. While this is compelling, the estimated density function will also depend on the extent of the information available for that asset (in particular, the implied volatility smile) - that is, unless an extraordinarily complete set of option prices for a given asset is available, an additional assumption regarding the tails of the distribution has to be made. This work presented a methodology to complete the non-parametric RND using tails drawn from a GEV distribution.

The parametric techniques to extracting the risk-neutral density function aim to find a sufficiently flexible function (or combination of functions) to describe a distribution of asset prices in the future. Although the assumption that the market's distribution of outcomes for a given asset follows a known density function is less compelling than the underlying assumption of the non-parametric approach, the parametric framework yields a more 'well behaved' function that can be used in several different applications. This work applied the methodology of the GB2 and MLN densities, in order to draw a comparison with the non-parametric framework.

The non-parametric framework's feature of allowing its shape to be determined by market prices is also reflected on the stability of the RND, as it is more prone to have its shape altered by daily changes in prices. As a consequence, the non-parametric risk-neutral densities for a given asset will change more abruptly from one day to another. On the other hand, the parametric risk-neutral densities are less likely to incorporate the changes in prices, as the optimization routines required for retrieving the parameters do not require all option prices to match those observed in the market. This conclusion is supported by the analysis of the moments of the risk-neutral densities under different frameworks, which show that even though they show similarities in magnitude and direction, the moments of the non-parametric RND are much more volatile and unstable. The contribution of this work was to apply a non-parametric technique to estimate the RND of USD/BRL options and to compare the moments of the distributions under different frameworks.

Secondly, the chosen application was to test whether the USD/BRL risk-neutral densities' informational could improve the decision taking when trading the currency. The fundamental hypothesis of the application was: information shocks that affect the fundamentals of an asset are reflected (to different extents) in both spot, futures and options - and by consequence in its risk-neutral probability density function. However, the opposite is not necessarily always true, that is, not all shifts in asset prices are purely due to changes in fundamentals. Rather, they could also happen due to liquidity and other 'behavioural' reasons.

We assumed that these 'non-fundamental' shocks tend to have only a short-term impact, fading away with time. Then, we test whether the RNDs can be used to detect when this type of shock is taking place in spot/futures, and build trades that profit of the convergence towards its fundamental value. This was done by modelling the USD/BRL returns as a function of the moments of the estimated risk-neutral densities.

We were able to build simple strategies that outperformed the buy-and-hold strategy during the period analysed (2014-2019). The parametric frameworks (in particular, the GB2) had, on average, a higher total return than the non-parametric one. This is likely related to the instability of the non-parametric moments, which could lead to false signals and sub-optimal timing to open/close positions. In this sense, the contribution of this work is to provide evidence that the informational content of the moments of the RND can be used to improve trading performance of USD/BRL itself.

For the purpose of trading based on the model described in this work, the greater stability of parametric approaches more than compensated the drawbacks from assuming a particular functional form for the density function. In this sense, the RNDs estimated using the GB2 framework appeared to strike a balance between a fairly similar shape to their non-parametric equivalents, and a stability similar to the MLN; indeed, the GB2 approach showed the best performance among all frameworks tested.

Extensions of this work could include: improvements in the trading model such as optimization of the trading bands and inclusion of stop losses; development of alternative trading frameworks, including options trading; studies that apply the non-parametric framework to pricing and interval forecasting; extraction of RND from fixed-date information provided by B3.

References

ABE, M. M.; CHANG, E. J.; TABAK, B. M. Forecasting Exchange Rate Density Using Parametric Models: The Case of Brazil. Revista Brasileira de Financas, v. 5, n. 1, p. 29–39, 2007.

BAHRA, B. Implied Risk-Neutral Probability Density Functions from Option Prices: Theory and Application. Bank of England Working Paper No 66, 1997.

BATES, D. S. American Finance Association The Crash of ' 87 : Was It Expected ? The Evidence from Options Markets. The Journal of Finance, v. 46, n. 3, p. 1009–1044, 1991.

BLACK, F.; SCHOLES, M. The Pricing of Options and Corporate Liabilities. Journal of Political Economy, v. 81, n. 3, p. 637–654, 1973.

BLISS, R. R.; PANIGIRTZOGLOU, N. Testing the stability of implied probability density functions. Journal of Banking and Finance, v. 26, n. 2-3, p. 381–422, 2002.

BREEDEN, D. T.; LITZENBERGER, R. H. Prices of State-Contingent Claims Implicit in Option Prices. The Journal of Business, v. 51, n. 4, p. 621–651, 1978.

CORRADO, C. J.; SU, T. Skewness and kurtosis in S&P 500 index returns implied by option prices. The Journal of Financial Research, v. 19, p. 175–192, 1996.

FABOZZI, F. J.; TUNARU, R.; ALBOTA, G. Estimating risk-neutral density with parametric models in interest rate markets. Quantitative Finance, v. 9, n. 1, p. 55–70, 2009.

FIGLEWSKI, S. Estimating the Implied Risk Neutral Density for the U . S . Market Portfolio. Volatility and time series econometrics: Essays in honor of Robert F. Engle, 2008.

FIGLEWSKI, S. Risk Neutral Densities: A Review. Annual Review of Financial Economics, v. 51, p. 329–359, 2018.

GARMAN, M. B.; KOHLHAGEN, S. W. Foreign currency option values. Journal of international Money and Finance, v. 2, n. 3, p. 231–237, 1983.

GATHERAL, J. A parsimonious arbitrage-free implied volatility parameterization with application to the valuation of volatility derivatives. Presentation at Global Derivatives & Risk Management, Madrid, 2004.

HESTON, S. L. A closed-form solution for options with stochastic volatility with applications to bond and currency options. The review of financial studies, v. 6, n. 2, p. 327–343, 1993.

JACKWERTH, J. C. Option-Implied Risk-Neutral Distributions and Risk Aversion. [S.l.: s.n.], 2004.

JONDEAU, E.; POON, S.-H.; ROCKINGER, M. Financial Modeling Under Non-Gaussian Distributions. [S.l.: s.n.], 2007. 541 p.

LIU, X. et al. Closed-form Transformations from Risk-neutral to Real-world Distributions Closed-form Transformations from Risk-neutral to Real-world Distributions. Journal of Banking and Finance, v. 31, n. 5, p. 1501–1520, 2007.

MALZ, A. M. Estimating the probability distribution of the future exchange rate from option prices. Journal of Derivatives, v. 5, n. 2, p. 18–36, 1997.

MERTON, R. C. Theory of Rational Option Pricing. The Bell Journal of Economics and Management Science, v. 4, n. 1, p. 141–183, 1973.

ORNELAS, J. R. H.; FAJARDO, J.; FARIAS, A. Estimating Relative Risk Aversion, Risk-Neutral and Real-World Densities Using Brazilian Real Currency Options. Brazilian Central Bank working paper, n. 269, 2012.

SANTOS, A.; GUERRA, J. Implied Risk Neutral Densities From Option Prices: Hypergeometric, Spline, Lognormal, and Edgeworth Functions. Journal of Futures Markets, v. 35, n. 7, p. 655–678, 2015.

SHIMKO, D. Bounds of Probability. Risk, v. 6, n. 4, p. 33–37, 1993.

TABAK, B. M.; CHANG, E. J. Risk-Neutral Probability Densities. Financial Stability Report, Department of Research, Banco Central do Brasil, n. November, p. 132–141, 2002.