Pontifícia Universidade Católica
do Rio de Janeiro

Manoel Francisco de Souza Pereira

Nonparametric Estimation of Risk-Neutral Distribution via the Empirical Esscher Transform

TESE DE DOUTORADO

Thesis presented to the Programa de Pós-Graduação em Engenharia Elétrica of the Departamento de Engenharia Elétrica, PUC-Rio as partial fulfillment of the requirements for the degree of Doutor em Engenharia Elétrica.

Advisor: Prof. Alvaro de Lima Veiga Filho

Rio de Janeiro
December 2016

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO

Manoel Francisco de Souza Pereira

Nonparametric Estimation of Risk-Neutral Distribution via the Empirical Esscher Transform

Thesis presented to the Programa de Pós-Graduação em Engenharia Elétrica of the Departamento de Engenharia Elétrica do Centro Técnico Científico da PUC-Rio, as partial fulfillment of the requirements for the degree of Doutor.

Prof. Alvaro de Lima Veiga Filho
Advisor
Departamento de Engenharia Elétrica – PUC-Rio

Prof. Cristiano Augusto Coelho Fernandes
Departamento de Engenharia Elétrica – PUC-Rio

Prof. Caio Ibsen Rodrigues de Almeida
FGV

Gyorgy Varga
FCE

Jose Renato Haas Ornelas
Banco Central do Brasil

Prof. Márcio da Silveira Carvalho
Coordinator of the Centro Técnico
Científico da PUC-Rio

Rio de Janeiro, December 12th, 2016

Manoel Francisco de Souza Pereira
Graduated in Economic Sciences at Universidade Federal Rural do Rio de Janeiro. Master in electrical engineering at Pontifical Catholic University of Rio de Janeiro.

Bibliographic data

Pereira, Manoel Francisco de Souza

Nonparametric Estimation of Risk-Neutral Distribution via the Empirical Esscher Transform / Manoel Francisco de Souza Pereira; advisor: Prof. Álvaro de Lima Veiga Filho – 2016.

127 f.; 30 cm

Tese (Doutorado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Elétrica, 2016.

Inclui bibliografia.

1. Engenharia Elétrica – Teses. 2. Apreçamento de opções. 3. Medida Neutra ao Risco. 4. Estimação Não Paramétrica. 5. Preço de Estado. 6. Transformada de Esscher Empírica. I. Veiga Filho, Álvaro de Lima. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Elétrica. III. Título.

CDD: 621.3

Acknowledgments

To God.

To CNPQ for the financial support and PUC-RJ for opening the doors of knowledge in my life.

To my advisor, Alvaro Veiga, who since my master's degree shared with me all his experience and knowledge. Moreover, in the most difficult moments of my life, you helped me with words of courage. I do not have words to thank you enough.

To Professor Cristiano Fernandes for the corrections and for providing the information needed to complete this work.

To Professor Caio Almeida for all the information about bibliography that was essential to conclude this work.

To the secretarial staff for their assistance and kindness, especially to Alcina Portes and Adriane.

To Mrs. Ana Maria for offering me all the necessary infrastructure during my periods of studies. Without her help the burden would be heavier.

To my brothers and sisters for all the incentive.

To Flávia for the support and love.

To my daughter Laura who brought joy in difficult times.

Abstract

Pereira, Manoel Francisco de Souza; Veiga Filho, Álvaro de Lima (Advisor). **Nonparametric Estimation of Risk-Neutral Distribution via the Empirical Esscher Transform.** Rio de Janeiro, 2016. 127p. PhD Thesis – Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

This thesis is comprised of three studies concerning the use of an empirical version of the Esscher Transform for nonparametric option pricing. The first one introduces the empirical Esscher transform and compares its performance against some well-known parametric option pricing approaches. In our proposal, we make only mild assumptions on the pricing kernel and there is no need for a risk-neutral model. In the second study, we propose a method for nonparametric option pricing under a GARCH framework with nongaussian innovations. Several papers have extended nonparametric option pricing and provided evidence that this methodology performs adequately in the presence of realistic financial time series. To represent a realistic time series, we use a new class of observation driven model, called dynamic conditional score model, proposed by Harvey (2013), for modeling the volatility (and heavy tails) of the asset's price. Finally, in the third study, we introduce a new approach for indirect estimation of state-prices implicit in financial asset prices from empirical Esscher transform. First, we generalize the discrete version of the Breeden and Litzenberger (1978) method for the case where states are not equally spaced. Second, we use the historical distribution of the underlying asset's price and the observed option prices to estimate the implicit Esscher parameter.

Keywords

Nonparametric estimation; indirect estimation; state-price; risk-neutral probability; empirical Esscher transform.

Resumo

Pereira, Manoel Francisco de Souza; Veiga Filho, Álvaro de Lima (Orientador). **Estimação não Paramétrica da Distribuição Neutra ao Risco através da Transformada de Esscher Empírica**. Rio de Janeiro, 2016. 127p. Tese de Doutorado - Departamento de Engenharia Elétrica, Pontifícia Universidade Católica do Rio de Janeiro.

Esta tese é composta de três estudos referentes ao uso de uma versão empírica da Transformada de Esscher para o apreçamento não paramétrico de opções. O primeiro introduz a transformada Esscher empírica e compara seu desempenho contra algumas bem conhecidas abordagens de apreçamento de opções paramétricas. Em nossa proposta, fazemos apenas suposições simples sobre o *pricing kernel* e não há necessidade de um modelo neutro ao risco. No segundo estudo, propomos um método de apreçamento de opções não paramétrico sob uma estrutura GARCH com inovações não Gaussianas. Vários artigos estenderam o apreçamento de opções não paramétrico e fornecendo evidências que esta metodologia funciona adequadamente na presença de séries temporais financeiras realistas. Para representar uma série temporal realista, usamos uma nova classe de modelo conduzido pela observação, denominado modelo de *score* condicional dinâmico, proposto por Harvey (2013), para modelar a volatilidade (e a cauda pesada) do preço do ativo. Finalmente, no terceiro estudo, introduzimos uma nova abordagem para a estimação indireta dos *state-prices* implícitos nos preços dos ativos financeiros a partir da transformada Esscher empírica. Primeiro, generalizamos a versão discreta do método de Breeden e Litzenberger (1978) para o caso em que os estados não são igualmente espaçados. Em segundo lugar, utilizamos a distribuição histórica do preço do ativo subjacente e os preços das opções observadas para estimar o parâmetro Esscher implícito.

Palavras-chave

Estimação não paramétrica; estimação indireta; preço de estado; probabilidade neutra ao risco; transformada de Esscher empírica.

Dedico esta tese a minha querida mãe Laura Inês de Souza *in memorian*.

Contents

List of figures

List of Tables

1
Introduction

1.1
Risk-Neutral Pricing

In most option pricing models, the fair price is determined from the expected value of its cash flow, under a risk-neutral probability (measure Q), and discounted by a risk-free rate. Under the assumption that the market is dynamically complete, it could be shown that every derivative security can be hedged and the measure Q is unique (Bingham and Kiesel, 2004). However, incomplete markets exist for many reasons and, according to the second fundamental theorem of asset pricing, we have an infinite number of measures Q under which one can get prices of derivatives. Then, one of the central questions in quantitative finance is how to get a measure of the risk-neutral probability that provides theoretical prices closest to those observed in the market.

According to Danthine and Donaldson (2015), the literature highlights two approaches to this problem: models based on the general equilibrium (Arrow, 1964, Debreu, 1959, Lucas, 1978, Rubinstein, 1976) and the models based on absence of arbitrage (Black-Scholes, 1973, Cox and Ross, 1976, Harrison and Kreps, 1979, Harrison and Pliska, 1981). In the general equilibrium, the supply and demand interacts in various markets affecting the prices of many goods simultaneously. The valuation of assets occurs when the markets are balanced, that is, when supply equals the demand. Thus, from a theoretical connection between macroeconomics (aggregate consumption) and financial markets, the marginal rate of substitution is used to determine a measure Q by solving a utility maximization problem.

In absence of arbitrage, we are appealing to the law of one price. This states that the equilibrium prices of two separate units of what is essentially the same good should be identical. If this was not the case, a riskless and costless arbitrage opportunity would open up: buy extremely large amounts at the low price and sell them at the high price, forcing the two prices to converge. The first fundamental theory of asset pricing says that, if a market model has a risk-neutral measure Q, then it does not admit arbitrage. The conditions that the risk-neutral probability structure must satisfy are that the discounted price process has zero drift and it must also be equivalent to the original structure (i.e., the same set of price paths must have positive probability under both structures). Then, a class of pricing kernels, or Radon-Nikodym derivatives, can be specified and impose restrictions that ensure the existence of a measure Q. In this case, the measure Q can be obtained without completely characterizing equilibrium in the economy (Christoffersen, Elkamhi, Feunou, and Jacobs, 2010, Christoffersen, Jacobs and Ornthanalai, 2013).

In both cases, these approaches require the formulation of an explicit risk-neutral model and are restricted to a few probability distributions for modeling the economy's uncertainty. However, empirical observations of asset returns showed several stylized facts, which highlight the parametric misspecification risk for the used stochastic process. Hence, due to the poor empirical performance of parametric methods, the nonparametric option pricing techniques have expanded rapidly in recent years, because they offer an alternative by avoiding possibly biased parametric restrictions (Haley and Walker, 2010).

This thesis presents two methods to determine the measure Q and to price European call option from nonparametric methods.

1.2
Objectives

The main objective of this thesis is to verify if simple assumptions on empirical pricing kernel are able to generate a measure Q that produces theoretical prices closer to those observed in the market.

From our investigation, we are able to derive three articles. The first article introduces the empirical Esscher transform and studies the nonparametric option pricing. In the second, we demonstrate that our proposal is flexible and performs very well in the presence of realistic financial time series and, in the third article, we use the empirical Esscher transform to include assets' and derivatives' data to get a risk-neutral probability.

1.3 Main Contributions

In general, in nonparametric option pricing methods, the historical distribution of the asset's prices is used to predict the distribution of future prices and the maximum entropy principle is employed to transform the empirical distribution into its risk-neutral counterpart, by minimizing some information criterion (Stutzer, 1996, Haley and Walker, 2010, Almeida and Azevedo, 2014). As the change of measure does not involve the distribution of the model's innovations, this method of risk-neutralization is applicable even when there are restrictions under the innovations' probability distribution.

Roughly speaking, on nonparametric option pricing there are two main objectives, not necessarily excludent. In the first type, the main objective is the adoption of other functions (for example, the Cressie-Read family) as alternative ways of measuring distance in the space of probabilities. While the second one is more focused on providing evidence that the methodology is flexible and that it performs adequately in the presence of realistic financial time series. Our first and second contributions are in this direction.

In our first contribution (Chapter 2), we assume that the empirical pricing kernel is known and given by an empirical version of the Esscher transform (1932). This assumption is reasonable, because it is well known in information theory that a problem of entropy maximization has its solution in the form of the Esscher transform (Buchen and Kelly, 1996, Stutzer, 1996, Duan, 2002).

In our second contribution (chapter 3), we use a recent proposal of dynamic model, proposed by Harvey (2013), which offers an alternative to model the volatility (and heavy tails) of observed underlying asset price using GARCH models and analyzes the nonparametric option pricing method.

In our third contribution (Chapter 4), we introduce a new approach for indirect estimation of the implicit state-price in financial asset prices using the empirical Esscher transform. First, we generalize the discrete version of the Breeden and Litzenberger (1978) method for the case where states are not equally spaced. Second, we use the empirical Esscher transform to include underlying assets' and derivatives' data. We use the historical distribution of the underlying asset prices and the observed option prices to estimate the implicit empirical

Esscher parameter. Then, we fit polynomials between the implied Esscher parameter and the strike price, as in Shimko (1993), to obtain our measure Q.

In order to evaluate the flexibility of the proposed method to adapt to different levels of maturity and moneyness, experiments (with synthetic data and real data) are performed and compared to main benchmarking.

1.4
Outline

The papers in this thesis share the common theme of asset pricing under a nonparametric structure. Each of these papers examines a distinct, well defined research problem related to the main topic and occupies a separate chapter. Moreover, the thesis contains sections on literature review, methodology and case study.

The remainder of this thesis is organized as follows. Chapter 2 presents an empirical version of the Esscher transform (1932). Chapter 3 uses a more realistic data generator process to analyze the nonparametric option pricing method. Chapter 4 introduces a new approach for indirect estimation of implicit state-price in financial asset price using empirical Esscher transform. In chapter 5 we conclude by summarizing the main results of this thesis in more detail. All the works cited in the thesis and other relevant documents are presented in the Bibliography. The Appendix is included at the end of this document.

2
Nonparametric Option Pricing

Chapter Abstract:

This paper introduces an empirical version of the Esscher transform for nonparametric option pricing. Traditional parametric methods require the formulation of an explicit risk-neutral model and are operational only for a few probability distributions for the returns of the underlying asset. In our proposal, we make only mild assumptions on the price kernel and there is no need for the formulation of the risk-neutral model. First, we simulate sample paths for the returns under the physical measure *P*. Then, based on the empirical Esscher transform, the sample is reweighted, giving rise to a risk-neutralized sample from which derivative prices can be obtained by a weighted sum of the options' payoffs in each path. We analyze our proposal in experiments with artificial and real data.

Keywords: nonparametric estimation, risk-neutral density, option pricing, empirical Esscher transform.

2.1
Introduction

In most option pricing models, the fair price is determined from the expected value of its cash flow, under a risk-neutral probability (measure Q), and discounted by a risk-free rate. Under the assumption that the market is dynamically complete, it could be shown that every derivative security can be hedged and the measure Q is unique (Bingham and Kiesel, 2004). However, incomplete markets exist for many reasons and, according to the second fundamental theorem of asset pricing, we have an infinite number of measures Q under which one can get prices of derivatives. Then, how to choose a measure Q from an infinite set of possible measures?

According to Danthine and Donaldson (2015), the literature highlights two approaches to this problem: models based on the general equilibrium (Arrow, 1964, Debreu, 1959, Lucas, 1978, Rubinstein, 1976, Brennan, 1979) and the models based on absence of arbitrage (Black-Scholes, 1973, Cox and Ross, 1976, Harrison and Kreps, 1979, Harrison and Pliska, 1981). In the general equilibrium, the supply and demand interacts in various markets affecting the prices of many goods simultaneously. The valuation of assets occurs when the markets are balanced, that is, when the supply equals the demand. Thus, from a theoretical connection between macroeconomics (aggregate consumption) and financial markets, the marginal rate of substitution is used to determine a measure Q by solving a utility maximization problem.

In absence of arbitrage, we are appealing to the law of one price. This states that the equilibrium prices of two separate units of what is essentially the same good should be identical. If this was not the case, a riskless and costless arbitrage opportunity would open up: buy extremely large amounts at the low price and sell them at the high price, forcing the two prices to converge. The first fundamental theory of asset pricing says that, if a market model has a measure Q, then it does not admit arbitrage. The conditions that the risk-neutral probability structure must satisfy are that the discounted price process has zero drift and it must also be equivalent to the original structure. Then, a class of pricing kernels, or Radon-Nikodym derivatives, can be specified and impose restrictions that ensure the existence of a risk-neutral measure. In this case, the measure Q can be obtained without completely characterizing equilibrium in the economy (Christoffersen, Elkamhi, Feunou, and Jacobs, 2010, Christoffersen, Jacobs and Ornthanalai, 2013).

In both cases, these approaches require the formulation of an explicit risk-neutral model and are restricted to a few probability distributions for the measure Q. First, because it is difficult to characterize the general equilibrium setup underlying a Risk-Neutral Valuation Relationship (RNVR), see for example Duan (1995, 1999). Second, it is possible to investigate option valuation for a large class of conditionally heteroskedastic processes (Gaussian or non-Gaussian), provided that the conditional moment generating function exists. Christoffersen, Jacobs and Wang (2004), cite that they help explain some stylized facts (smile effect, volatility variability over time and presence of clusters in certain periods) in a qualitative sense, but the magnitude of the effects is insufficient to completely solve the biases. The resulting pricing errors have the same sign as the Black-Scholes (1973) pricing errors, but are smaller in magnitude.

Due to the poor empirical performance of parametric methods, according to Haley and Walker (2010), the nonparametric option pricing techniques have expanded rapidly in recent years. In these methods, the historical distribution of prices is used to predict the distribution of future asset prices. According to Stutzer (1996), by using past data to estimate the payoff distribution at expiration, it permits more accurate assessment of the likely pricing impact caused by investors' data-based beliefs about the future value distribution. Moreover, it offers an alternative by avoiding possibly biased parametric restrictions and reducing the misspecification risk.

There are two ways to nonparametric estimate risk-neutral probabilities implicit in financial instruments: the methods that seek to infer the empirical risk-neutral probability from option market[1] (kernel, maximum entropy, and curve fitting) and the methods that seek to infer the empirical risk-neutral probability from asset price (with or without option price), as canonical valuation developed by Stutzer (1996). In the case of canonical valuation, the maximum entropy principle is employed to transform the empirical distribution into its risk-neutral counterpart, by minimizing the Kullback–Leibler information criterion (KLIC).

Several papers have extended Stutzer's original work and demonstrated that the methodology is flexible and performs very well in the presence of realistic financial time series, see Gray and Newman (2005), Gray, Edwards, and Kalotay (2007), Alcock and Carmichael (2008), Haley et al (2010) and Almeida and Azevedo (2014). Other researchers, as Haley et al (2010) and Almeida et al (2014), suggested the adoption of members of the Cressie-Read family of discrepancy functions as alternative ways of measuring distance in the space of probabilities.

This paper introduces an empirical version of the Esscher transform (1932) for nonparametric option pricing. We assume that the empirical pricing kernel is known and given by an empirical version of the Esscher transform (1932). This assumption is reasonable, because it is well-known in the information theory that a problem of maximum entropy has its solution in the form of the Esscher transform (Buchen and Kelly, 1996, Stutzer, 1996, Duan, 2002).

Duan (2002) also develops a nonparametric option pricing theory based on Esscher transform (1932). He uses a binary search to find the Esscher parameter and the measure Q is evaluated using the standard polynomial approximation formula. In our case, we use a consistent estimator for the moment generation function and we avoid the use of intensive computational methods. As the change of measure does not involve the distribution of the model's innovations, this method of risk-neutralization is applicable even when the moment generating

[1] There are also parametric methods such as Expansion, Distributions and Generalized Mixture. Parametric methods use a known probability distribution adjusted to observed option prices.

function of the innovations' probability distribution does not exist. Hence, we obtain a method that does not require a set of restrictive assumptions for the formulation of a specific model; that provides a clear and easy way to obtain a risk-neutral distribution; is adaptable and flexible to respond to changes in the data generating process; and explores the whole cross-section information contained in the underlying asset's price.

In many applications, the empirical pricing kernel is the object of interest because it describes risk preferences of a representative agent in an economy, and the risk aversion function estimates the investors' expectations about future return probabilities (Hansen and Jagannathan, 1991, Aït-Sahalia and Lo, 2000, and Jackwerth, 2000). Our objective is to verify if mild assumptions on the empirical pricing kernel are able to obtain option prices closer to the observed in the market.

In order to evaluate the flexibility of the proposed method to adapt to different contexts, three experiments (two with synthetic data and one with real data) are performed. We compare the proposed pricing method to the Black-Scholes (1973) model and the Heston (1993) model.

The paper is organized as follows. Section 2.2 discusses the Esscher transform. In Section 2.3, we introduce the empirical Esscher transform. Section 2.4 presents the methodology we use to compare the different pricing methods, and the results are discussed in Section 2.5. Finally, Section 2.6 concludes.

2.2
The Esscher transform

Let X be a random variable with probability density function $f(x)$ and let θ be a real number. Then, the Esscher transform (ET) of $f(x)$ with Esscher parameter θ is given by $f(x;\theta)$, defined as:

$$f(x;\theta) = \frac{e^{\theta x}}{\int_{-\infty}^{+\infty} e^{\theta x} f(x)dx} f(x). \tag{2.1}$$

Note that $f(x;\theta)$ is also a probability density function since it integrates one. Furthermore, the ET can be interpreted as a reweighted version of $f(x)$, with reweighting function given by:

$$m(x;\theta) = \frac{e^{\theta x}}{\int_{-\infty}^{+\infty} e^{\theta x} f(x)dx}. \quad (2.2)$$

The denominator of this expression represents the moment generating function (mgf) of $f(x)$, denoted by:

$$M(\theta) = E\left[e^{\theta X}\right] = \int_{-\infty}^{+\infty} e^{\theta x} f(x)dx. \quad (2.3)$$

In this case, for the Esscher transform to exist, the mgf of X must exist, which precludes some well-known density functions, like the t-student. Hence, the ET of $f(x)$ can be expressed as:

$$f(x;\theta) \quad = \quad m(x;\theta)f(x) \quad = \quad \frac{e^{\theta x}}{M(\theta)} f(x). \quad (2.4)$$

Consider now the ET of the density $f(x_T)$ of X_T, the log-return of an asset for a period T, given by:

$$f(x_T;\theta) = m(x_T;\theta)f(x_T) = \frac{e^{\theta x_T}}{M(\theta)} f(x_T). \quad (2.5)$$

Gerber and Shiu (1994) proposed to use the ET of X_T as the risk-neutral distribution (RND) for the log-return of this asset. They call it Risk-Neutral Esscher Transform (RNET). In this context, $f(x_T)$ is referred to as the physical probability measure P and $f(x_T;\theta)$, the ET of $f(x_T)$, is identified as the risk-neutral measure Q or, still, the equivalent martingale measure.

Let S_t be the price of an asset at time t. According to the fundamental theorem of asset pricing (Bingham et al, 2004), the risk-neutral value

$v\left(g(S_T)/S_0\right)$ of a derivative $g(S_T)$ with maturity T is given by the expected value of the payoff under the measure Q, discounted by the risk-free rate of return for period T, r_T:

$$v\left(g(S_T)/S_0\right) = e^{-r_T}E^Q\left[g(S_T)/S_0\right]. \tag{2.6}$$

This is also true if the derivative is the asset itself so that $v\left(g(S_T)/S_0\right) = S_0$ and $g(S_T) = S_T$ with $S_T = S_0e^{X_T}$. This imposes the non-arbitrage constraint:

$$S_0 = e^{-r_T}E^Q\left[S_0e^{X_T}/S_0\right] \rightarrow e^{r_T} = E^Q[e^{X_T}]. \tag{2.7}$$

Now, defining the measure Q as the ET of $f(x_T)$, we obtain the following condition for the value of θ:

$$\begin{gathered} e^{r_T} = \int_{-\infty}^{+\infty} e^{X_T}f(x_T;\theta)dx_T \\ e^{r_T} = \int_{-\infty}^{+\infty} \frac{e^{(\theta+1)x_T}}{M(\theta)}f(x_T)dx_T = \frac{M(\theta+1)}{M(\theta)}. \end{gathered} \tag{2.8}$$

Hence, the measure Q is given by $f(x_T;\theta^*)$, with $\theta^* = arg_\theta\left\{e^{r_T} = \frac{M(\theta+1)}{M(\theta)}\right\}$.

Gerber et al (1994) explores several different distributional assumptions to X_T, price dynamics and log-returns of an asset. They show that the RNET encompasses the classical option pricing formula of Black-Scholes (1973)[2] for Wiener processes, and the Binomial Model (Cox, Ross and Rubinstein, 1979).[3]

Duan (2002) explored empirical distribution of X_T and develops a nonparametric option pricing theory based on the first equation in (2.8). In this case, he used a binary search to find θ^*, the integral was evaluated numerically and $f(x_T;\theta)$ was evaluated using the standard polynomial approximation formula.

The RNET can also be applied to incomplete markets, which admit infinite measure Q. It provides an economic justification for selecting this particular transform, since it emerges as the solution for the problem of pricing a derivative under a power utility function (see Gerber et al, 1996).

Moreover, using the relative entropy principle, the risk-neutral density can be obtained from the following problem:

$$\min_{g(x)} \int_{-\infty}^{\infty} g(x) ln \frac{g(x)}{f(x)} dx \qquad (2.9)$$

where $g(x)$ is the model known and the discrepancy between it and another model $f(x)$ can be obtained by minimization of an information criterion. It is well-known in the information theory that the programming problem in (2.9) has its solution in the form of the ET (Buchen and Kelly, 1996, Stutzer, 1996, Duan, 2002).

2.3
The empirical Esscher transform

Consider a random sample of size n from X_T, denoted by $\{X_{T,i}\}_{i=1,n}$. Then, we define the Empirical Esscher Transform (EET) as:

[2] See appendix 6.1.
[3] The Esscher transform can be used in both cases, discrete time (Bühlmann, 1996 and Siu, Tong and Yang, 2004) and continuous time (Gerber et al, 1994). Chan (1999) showed the relationship between minimal entropy equivalent measure and the Esscher transform when asset prices follow the Lévy process (see Chorro, Guégan and Ielpo, 2008 and Ornthanalai, 2011). For more studies see Monfort and Pegoraro (2011), Li and Badescu (2012) and Guégan, Ielpo and Lalaharison (2013).

$$q_{i,\theta} = \frac{e^{\theta X_{T,i}}}{\sum_{j=1}^{n} e^{\theta X_{T,j}}}. \tag{2.10}$$

Note that $\{q_{i,\theta}\}_{i=1,n}$ constitutes a probability mass function since $\sum_{i=1}^{n} q_{i,\theta} = 1$ and $q_{i,\theta} > 0\ \forall i$. Furthermore, in analogy to the ET, it can be interpreted as a reweighted version of the original sample $\{X_{T,i}\}_{i=1,n}$. Then, one can write:

$$q_{i,\theta} = m_q(X_{T,i};\theta)p_i \tag{2.11}$$

with $p_i = 1/n$ being the original weight and $m_q(X_{T,i};\theta)$ the reweighting function, given by:

$$m_q(X_{T,i};\theta) = \frac{e^{\theta X_{T,i}}}{\frac{1}{n}\sum_{j=1}^{n} e^{\theta X_{T,j}}}. \tag{2.12}$$

The denominator of this expression represents an estimator of the moment generating function (mgf) of $f_{X_T}(x)$:

$$\widehat{M}(\theta) = \frac{1}{n}\sum_{j=1}^{n} e^{\theta X_{T,j}}. \tag{2.13}$$

Then, if $\{X_{T,i}\}_{i=1,n}$ is a i.i.d sample, the weak law of large numbers assures that, if $E[e^{\theta X}]$ exists for all $\theta \in \mathfrak{R}$, then $\widehat{M}(\theta)$ is a consistent estimator of $M(\theta)$, i.e.:

$$\widehat{M}(\theta) \xrightarrow{P} M(\theta). \tag{2.14}$$

Now, take the sample version of the fundamental theorem of asset pricing stated in the preceding section. An estimate of the risk-neutral value

$\hat{v}\left({g(S_T)}/{S_0} \right)$ is then given by the estimated expected value, denoted by $\hat{E}^Q[\cdot]$ of the payoff under the measure Q, discounted by the risk-free rate of return for period T, r_T:

$$\hat{v}\left({g(S_T)}/{S_0} \right) = e^{-r_T}\hat{E}^Q\left[{g(S_T)}/{S_0} \right] \tag{2.15}$$

with $S_{T,i} = S_0 e^{X_{T,i}}$ and $\hat{E}^Q\left[{g(S_T)}/{S_0} \right] = e^{-r_T} \sum_{i=1}^{n} g(S_0 e^{X_{T,i}}) q_{i,\theta}$.

Using the above expressions, it is easy to check that the sample version of the no-arbitrage condition is given by:

$$e^{r_T} = \frac{\widehat{M}(\theta + 1)}{\widehat{M}(\theta)}. \tag{2.16}$$

Then, the empirical risk-neutral measure Q is given by $\{q_{i,\hat{\theta}^*}\}_{i=1,n}$, with $\hat{\theta}^* = arg_\theta \left\{ e^{r_T} = \frac{\widehat{M}(\theta+1)}{\widehat{M}(\theta)} \right\}$. Again, if $E\left[e^{\theta X}\right]$ exists for all $\theta \in \Re$ then $\widehat{M}(\theta + 1)$ and $\widehat{M}(\theta)$ will converge in probability to their respective population values and, by consequence, the solution of the non-arbitrage constraint will also converge, i.e., $\hat{\theta}^* \overset{P}{\rightarrow} \theta^*$.

The price of a European call on a non-dividend-paying stock is obtained under the risk-neutral distribution $q(S_T)$ and the payoff is discounted at the deterministic risk-free rate r:

$$C(K,T) = e^{-rT} \int_{-\infty}^{\infty} (S_T - K)^+ q(S_T) d\, S_T \tag{2.17}$$

or, alternatively,

$$C(K,T) = e^{-rT}\int_{-\infty}^{\infty}(S_T - K)^{+}m(S_T)f(S_T)d\,S_T \tag{2.18}$$

where T is the time to maturity, S_T is the underlying asset price, K is the strike price, $f(S_T)$ is the physical distribution of the asset price at the option's expiration and $m(S_T) = q(S_T)/f(S_T)$ is the pricing kernel, characterizing the change of measure $f(S_T)$ to $q(S_T)$.

Consider the discretization of integral in (2.18):

$$C(K,T) = e^{-rT}\left[\sum_{j=1}^{n}\left(S_{T,j} - K\right)^{+}m(S_T)p(S_T)\right] \tag{2.19}$$

where $q(S_T) = m(S_T)p(S_T)$ is the risk-neutral probability mass function.

2.4
Methodology

This section presents the methodology that is used to compare the proposed method to artificial and real data. To investigate its applicability in some settings, the empirical Esscher transform is applied to price European call options across a range of moneyness and maturities. The algorithm for our method is:

1. Simulate the physical distribution for $S_{T,i}, i = 1, \dots, n$;
2. Compute the empirical Esscher parameter, $\hat{\theta}^{*}$, using the equation (2.16);
3. Compute the option price with the equation (2.19).

In experiment 1, the asset price (step 1) follows a Geometric Brownian Model (GBM), under the physical measure:

$$dS_t = \mu S_t dt + \sigma S_t dz_t \tag{2.20}$$

where S_t is the underlying asset's price at time t, μ is the expected rate of return, σ is the volatility and dz_t follows a Wiener process. We compare the empirical Esscher transform simulated option prices to the true price, Black-Scholes model.

Using the assumptions that the future asset price follows a lognormal distribution, and the returns follow a normal distribution, we obtain the Black-Scholes (1973) formula for the price, at time 0, of a European call option on a non-dividend-paying stock,

$$C = S_0 N(d_1) - Ke^{-rT} N(d_2) \tag{2.21}$$

where,

$$d_1 = \frac{ln(S_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}} \tag{2.22}$$

$$d_2 = \frac{ln(S_0/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}} \tag{2.23}$$

The functions $N(d_1)$ and $N(d_2)$ are the cumulative probability distribution function for a normal variable with zero mean and variance equal to 1, C is the price of the call option, S_0 is the stock price at time 0, r is the risk-free interest rate continuously compounded, and σ is the asset volatility. Based on the works of Hutchinson and Poggio (1994), Stutzer (1996) and Gray et al (2005), we use an annualized volatility of 20%, a drift of 10% and the riskless rate of interest is assumed to be a constant of 5%.

In experiment 2, the asset price (step 1) follows the Heston (1993) model, which assumes a diffusion process for the asset price and another stochastic process for the volatility. The asset price S_t follows the diffusion, under the physical measure:

$$dS_t = \mu S_t dt + \sqrt{v_t} S_t dz_{1,t} \tag{2.24}$$

where $z_{1,t}$ is a Wiener process. The volatility $\sqrt{v_t}$ follows the diffusion:

$$dv_t = \kappa(\theta - v_t)dt + \sigma\sqrt{v_t} dz_{2,t} \tag{2.25}$$

$$dz_{1,t} dz_{2,t} = p dt \tag{2.26}$$

where $z_{2,t}$ is a Wiener process that has correlation ρ with $z_{1,t}$, θ is the long-run mean of the variance, κ is a mean reversion parameter, σ is the volatility of variance. We compare the empirical Esscher transform simulated option prices to the true price, the Heston models.

The price of a European call option with time to maturity $(T-t)$, is given by:

$$C = S_t P_1 - e^{-r(T-t)} K P_2. \quad (2.27)$$

The quantities P_1 and P_2 are the probabilities that the call option expires in-the-money, conditional on the log of the asset price, $x_t = ln(S_t)$, and on the volatility v_t, each at time t. The probabilities P_j can be obtained by inverting the characteristic functions f_j defined below. Thus:

$$P_j = \frac{1}{2} + \frac{1}{\pi}\int_0^{\infty} Re\left[\frac{e^{-i\phi ln(K)} f_j}{i\phi}\right] d\phi \quad (2.28)$$

$$f_j = exp\left(C_j + D_j v_t + i\phi x\right) \quad (2.29)$$

$$C_j = r\phi i\tau + \frac{\kappa\theta}{\sigma^2}\left\{\left(b_j - \rho\sigma\phi i + d_j\right)\tau - 2ln\left[\frac{1 - g_j e^{d_j\tau}}{1 - g_j}\right]\right\} \quad (2.30)$$

$$D_j = \frac{b_j - \rho\sigma\phi i + d_j}{\sigma^2}\left[\frac{1 - e^{d_j\tau}}{1 - g_j e^{d_j\tau}}\right] \quad (2.31)$$

$$g_j = \frac{b_j - \rho\sigma\phi i + d_j}{b_j - \rho\sigma\phi i - d_j} \quad (2.32)$$

$$d_j = \sqrt{\left(\rho\sigma\phi i - b_j\right)^2 - \sigma^2\left(2u_j\phi i - \phi^2\right)}. \quad (2.33)$$

In these expressions $\tau = T - t$ is the time to maturity, $i = \sqrt{-1}$ is the imaginary unit, $u_1 = 1/2$, $u_2 = -1/2$, $b_1 = \kappa + \lambda - \rho\sigma$, and $b_2 = \kappa + \lambda$. The parameter λ represents the price of volatility risk as a function of the asset price, volatility, and time.

We use the Euler discretizations for the stochastic process of price and volatility. Based on the works of Lin, Strong and Xu (2001), Zhang and Shu (2003) and Gray et al (2005), we use the following values: $\kappa = 3.00$, θ =0.04, σ =0.40 and $\rho = -0.50$. The initial value of the volatility equals to its long-term average. For consistency with the Black-Scholes world simulations, the drift of 10% and the riskless rate of interest is assumed to be a constant 5% continuously compounded.

In these artificial experiments, the moneyness (S/K) is equal to 0.90, 0.97, 1.00, 1.03, 1.125 and the maturities are equal to $1/12$, $1/4$, $1/2$ and 1 years. For each time to maturity T, 200 returns are drawn to generate the distribution of T-year forward. We obtain the risk-neutral measure (step 2) and we calculate the option price (step 3). This procedure is repeated 10.000 times and we calculate the Mean Absolute Percentage Error (MAPE). We repeat the artificial experiments with 5×10^4 returns, and 200 repetitions, to analyze if the accuracy improves with an increase in the sample size.

In experiment 3, we evaluate nonparametric opting pricing with real data. We compare the prices of the proposed method (EET) to Stutzer prices (STZ) and Black-Scholes prices (BSM). For each time to maturity T, we perform bootstrap with replacement on historical returns of the underlying asset. We follow the sequence: (a) we construct a single trajectory for the asset price by drawing a certain quantity of historical log returns. For example, if the option has 17 days to maturity, then we draw the same quantity; (b) we accumulate the log returns of this trajectory and we obtain one price; (c) we repeat the process (a) and (b) 252 times to construct the physical distribution for the price at maturity (step 1). We obtain the risk-neutral measure (step 2) and we calculate the option price (step 3). We repeat this procedure 15.000 times and we calculate the MAPE.

The Stutzer (1996) method begins with the asset's historical distribution of T-year gross returns R_i, $i = 1, \ldots, n$, which are expressed as price relatives. By the maximum entropy principle, it shows that the risk-neutral probabilities are:

$$\hat{\pi}_i^* = \frac{exp\left(\gamma^* \frac{R_i}{(1+r)^T}\right)}{\sum_{i=1}^{n} exp\left(\gamma^* \frac{R_i}{(1+r)^T}\right)} \tag{2.34}$$

where γ^* is the Lagrange multiplier, given by the following minimization problem:

$$\gamma^* = arg \min_{\gamma} \sum_{i=1}^{n} exp\left[\gamma\left(\frac{R_i}{(1+r)^T} - 1\right)\right]. \tag{2.35}$$

Note that the equation (2.34) is similar to risk-neutral probabilities of the empirical version of the Esscher transform in the equation (2.10).[4] Canonical option prices follow from the equation:

$$C(K,T) = \frac{1}{(1+r)^T}\left[\sum_{i=1}^{n}\left(P_{T,i} - K\right)^+ \hat{\pi}_i^*\right], \qquad P_{T,i} = P_0 R_i, \quad i = 1, \dots, n. \tag{2.36}$$

where $P_{T,i}$ are the prices of underlying asset.

We consider the closing values of two daily databases of Vale's and Petrobras' prices from January 17, 2011 to January 17, 2012, containing 251 observations for each database.[5] We set time 0 to January 17, 2012 (the end point of the data sample period). The closing value of Vale on that day was $S_0 = 41.13$ and of Petrobras was $S_0 = 24.37$. We use the corresponding true market price on the valuation date as a benchmark.[6] The true market prices of the options with the strikes and maturities under consideration are shown in Table 2.2. The maturities are equal to $17/252$, $40/252$, $59/252$ and $121/252$ years for Petrobras (PETR4) and only the three first for Vale (VALE5). The interest risk-free rate was 10.3499% (17/252), 10.2485% (40/252), 10.1721% (59/252) and 10.032%

[4] Stutzer (1996) reports that the performance of canonical valuation improves when a small amount of option data is used.
[5] These data are specifics to the Brazilian market.
[6] All required data are obtained from Bovespa (http://www.bmfbovespa.com.br).

(121/252) obtained by linear interpolations.[7] Table 2.1 presents the main descriptive statistics of the log returns.

Table 2.1: Descriptive Statistics of the Log Returns.

	Mean	Standard Deviation	Skewness	Kurtosis	Maximum	Minimum
Petrobras	-0.0007	0.0180	0.5621	5.2172	0.0480	-0.0788
Vale	-0.0011	0.0174	-0.6538	7.1367	0.0577	-0.0962

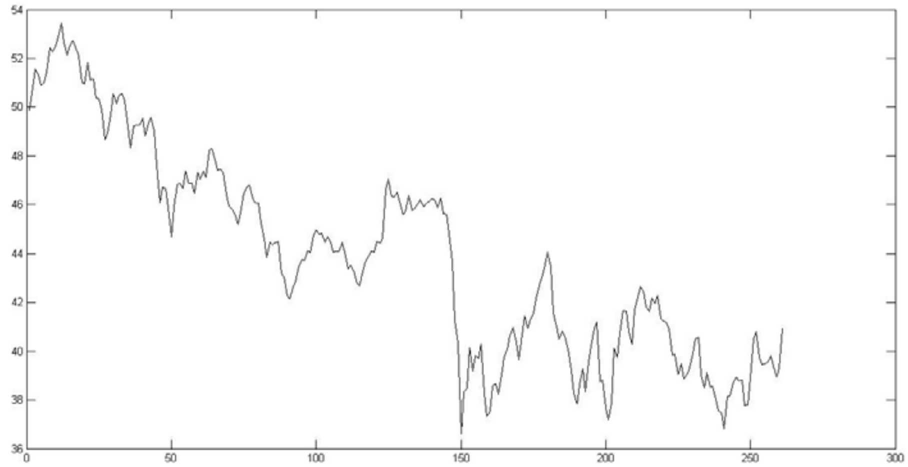


Figure 2.1: Vale prices from January 17, 2011 to January 17, 2012.

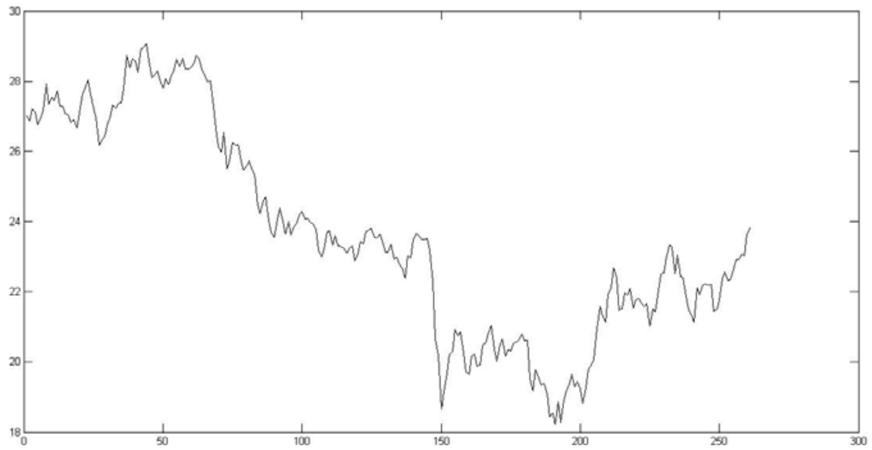


Figure 2.2: Petrobras prices from January 17, 2011 to January 17, 2012.

7 http://www.bcb.gov.br.

Table 2.2: Market prices of options from Vale and Petrobras.

	Vale (VALE5)		Petrobras (PETR4)	
Maturity (days)	**Strike**	**Option price**	**Strike**	**Option price**
17	44.00	0.11	25.66	0.18
	43.14	0.21	25.16	0.27
	42.00	0.54	24.83	0.43
	41.14	0.99	23.66	1.12
	41.00	1.03	22.83	1.80
	40.14	1.72	21.66	2.88
	39.07	2.50	20.83	3.63
	38.57	2.75	19.66	4.94
	37.14	4.24	18.66	5.78
	37.00	4.41	17.66	6.78
	36.14	5.32	15.16	9.19
	36.00	5.34		
	35.00	6.05		
	34.00	7.31		
	30.14	11.41		
	30.00	11.65		
	28.00	13.59		
40	46.07	0.15	27.83	0.09
	46.00	0.20	27.00	0.17
	45.57	0.19	25.83	0.40
	44.07	0.42	25.33	0.61
	43.07	0.86	25.00	0.77
	42.07	1.25	23.83	1.38
	42.00	1.32	22.83	2.11
	41.00	1.83	21.66	3.08
	40.57	2.01	21.00	3.70
	40.00	2.51	19.66	4.95
	38.00	4.00	18.66	5.79
	37.00	4.75	17.83	6.60
	36.57	5.21		
	35.07	6.50		
	32.00	9.30		
59	48.00	0.21	26.00	0.58
	44.14	0.87	24.00	1.50
	44.00	0.90	21.83	3.21
	41.07	2.30		
	40.00	3.00		
121			25.50	1.85

Figures 2.1 and 2.2 present the prices' behavior during the studied period. Since the 2008 crisis, the prices of these companies have presented a downward trend. The main factors that have contributed to the prices' fall were the Arab spring, the downgrade in the US credit rating from Standard and Poor's and the crisis in the Eurozone (Greece, Italy, Ireland and Portugal declared inability to pay their debts).

2.5
Results

Experiment 1 compares the Empirical Esscher transform (EET) method performance when the conditions are the same as in the Black-Scholes model through the MAPE (Mean Absolute Percentual Error). Results in the table 2.3 show that the EET prices reproduced BSM prices when the sample size increases. Moreover, results also show that the EET presents the highest pricing errors in situations Deep-Out-of-the-money.

Experiment 2 compares the EET method performance when the conditions are the same as in the Heston model through the MAPE. Results in the table 2.4, as in the table 2.3, show that the EET prices improves with sample size.

In experiment 3, we compare the EET, the Stutzer method (STZ) and the Black-Scholes model (BSM) in real data (from Petrobras and Vale). Tables 2.5 and 2.6 show that the lowest pricing errors are between the nonparametric methods. Table 2.5, for the maturity equal to 17, the nonparametric methods present similar results. For others maturities, the proposed method presents the lowest MAPE for moneyness equal to deep-out-of-the-money, out-of-the-money and at-the-money. Results in the table 2.6 are similar to the ones in table 2.5.

Table 2.3: MAPE of Empirical Esscher Transform Estimates in a Black-Scholes World.

This table contains the prices for a European call option, where the underlying prices are simulated by a geometric Brownian motion with $\mu = 0.1$ and $\sigma = 0.20$, and they are compared to the true Black-Scholes call prices. The top and bottom numbers reported for each combination are the mean absolute percentage error (MAPE) of the EET with 200 returns and with 5×10^4 estimated returns respectively over 10.000 and 200 simulations.

	Time to expiration (years)			
Moneyness (S/K)	1/12	1/4	1/2	1
Deep-out-of-the-money	25.1459	5.9172	2.9991	2.0394
(0.90)	1.6812	0.3472	0.1720	0.1325
Out-of-the-money	2.8328	1.8461	1.6741	1.6500
(0.97)	0.1870	0.1166	0.1041	0.1054
At-the-money	1.4156	1.4219	1.4480	1.5203
(1.00)	0.0932	0.0817	0.0898	0.0964
In-the-money	0.9585	1.1688	1.2757	1.3936
(1.03)	0.0612	0.0710	0.0787	0.0862
Deep-in-the-money	0.1512	0.5227	0.7681	0.9774
(1.125)	0.0092	0.0320	0.0477	0.0628

Table 2.4: MAPE of Empirical Esscher Transform Estimates in a Heston World.

This table contains the prices for a European call option, where the underlying prices are simulated by stochastic volatility with $\mu = 0.1$, $\kappa = 3.00$, $\theta = 0.04$, $\sigma = 0.40$ and $\rho = -0.50$, and they are compared to the true Heston call prices. The top and bottom numbers reported for each combination are the mean absolute percentage error (MAPE) of the EET with 200 returns and with 5×10^4 estimated returns respectively over 10.000 and 200 simulations.

	Time to expiration (years)			
Moneyness (S/K)	1/12	1/4	1/2	1
Deep-out-of-the-money	36.7566	13.4996	9.5056	7.6861
(0.90)	4.8636	4.9842	5.8228	6.0545
Out-of-the-money	7.2624	5.6063	5.1907	5.0316
(0.97)	1.2073	2.3118	3.2218	3.9084
At-the-money	3.8561	3.9084	4.0431	4.2104
(1.00)	0.6422	1.5878	2.4599	3.2255
In-the-money	2.0482	2.7513	3.1526	3.5286
(1.03)	0.3030	1.0644	1.8609	2.6532
Deep-in-the-money	0.2332	0.8956	1.4536	2.0503
(1.125)	0.0171	0.2615	0.7414	1.4129

Table 2.5: MAPE of Empirical Esscher Transform Estimates in Petrobras.

This table contains the prices for a European call option, where the underlying prices are simulated from bootstrap with replacement on historical returns, and they are compared to the true market price. The numbers reported for each combination are the MAPE of the EET (proposed method), STZ (Stutzer method) and BSM (Black-Scholes model) with 252 returns and the simulation is repeated 15,000 times.

Maturity	Moneyness (spot/strike)		EET	STZ	BSM
T = 17/252	Deep-out-of-the-money	0.95	60.2770	61.5768	71.3672
	Out-of-the-money	0.97	63.0951	63.9341	70.4573
		0.98	32.2144	32.6802	36.5306
	In-the-money	1.03	8.6391	8.6142	9.3748
		1.07	2.5588	2.4217	2.5840
	Deep-in-the-money	1.13	0.4307	0.3273	0.2894
		1.17	1.6701	1.5079	1.6076
		1.24	1.9466	2.0699	1.9588
		1.31	0.9689	0.8672	0.9671
		1.38	0.7246	0.6422	0.7244
		1.61	1.3303	1.2781	1.3303
T =40/252	Deep-out-of-the-money	0.88	131.3580	140.6030	157.7360
		0.9	102.7306	108.7454	118.8495
		0.94	62.9430	65.8111	70.0874
		0.96	36.6367	38.4690	41.1048
	Out-of-the-money	0.97	26.2260	27.6094	29.5751
	At-the-money	1.02	15.0387	15.5256	16.3215
	In-the-money	1.07	7.3354	7.4032	7.7823
	Deep-in-the-money	1.13	4.0906	3.9246	4.1547
		1.16	2.4239	2.2025	2.4157
		1.24	1.9286	1.6844	1.8991
		1.31	3.8815	3.6555	3.8650
		1.37	3.4446	3.2459	3.4375
T = 59/252	Deep-out-of-the-money	0.94	52.6996	56.6112	58.6471
	At-the-money	1.02	20.8055	21.8995	22.4323
	Deep-in-the-money	1.12	3.5929	3.5943	3.8422
T = 121/252	Deep-out-of-the-money	0.96	6.4109	8.8424	6.3078

Table 2.6: MAPE of Empirical Esscher Transform Estimates in Vale.

This table contains the prices for a European call option, where the underlying prices are simulated from bootstrap with replacement on historical returns, and they are compared to the true market price. The numbers reported for each combination are the MAPE of the EET (proposed method), STZ (Stutzer method) and BSM (Black-Scholes model) with 252 returns and the simulation is repeated 15,000 times.

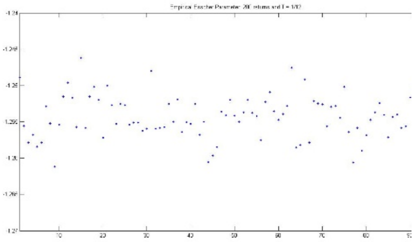
Maturity	Moneyness (spot/strike)		EET	STZ	BSM
T = 17/252	Deep-out-of-the-money	0.93	175.9523	182.4160	207.2992
		0.95	134.9038	138.7734	154.2885
	Out-of-the-money	0.98	61.8209	63.3503	69.8371
	At-the-money	1	28.7498	29.5111	32.7514
		1	31.0346	31.7410	34.7763
		1.02	8.7360	9.0209	10.4410
	In-the-money	1.05	5.9744	6.0083	6.5923
		1.07	11.1145	11.0728	11.4601
		1.11	1.9831	1.8404	1.9505
		1.11	1.0093	0.8644	0.9658
	Deep-In-the-money	1.14	0.9143	1.0715	0.9810
		1.14	1.2425	1.0812	1.1752
		1.18	5.4708	5.3079	5.4227
		1.21	0.7746	0.6326	0.7515
		1.36	1.8743	1.9600	1.8745
		1.37	2.7023	2.7859	2.7025
		1.47	1.9758	2.0428	1.9759
T = 40/252	Deep-out-of-the-money	0.89	193.5641	211.4168	228.3421
		0.89	126.1923	139.6921	152.3182
		0.9	178.4586	193.7011	207.7265
		0.93	111.0552	119.0766	125.9295
		0.95	41.1680	45.1967	48.5118
	Out-of-the-money	0.98	29.8445	32.5372	34.7016
		0.98	25.3795	27.9143	29.9489
	At-the-money	1	17.9030	19.5386	20.8631
		1.01	19.4561	20.8401	21.9774
	In-the-money	1.03	9.5245	10.4966	11.3238
		1.08	4.4847	4.7102	5.0866
		1.11	5.2491	5.2654	5.5572
	Deep-In-the-money	1.12	3.0515	3.0024	3.2671
		1.17	3.3883	3.2020	3.4340
		1.29	3.8424	3.6109	3.8307
T = 59/252	Deep-out-of-the-money	0.86	116.3855	135.1830	144.3093
		0.93	51.9814	58.8867	61.0783
		0.93	52.1273	58.8390	60.9299
	At-the-money	1	15.6564	18.0059	18.5215
	In-the-money	1.03	9.4821	11.0098	11.3472

We also analyze the behavior of the empirical Esscher parameter. The results are presented in Figure 2.3 and tables 2.7, 2.8, 2.9 and 2.10. In the Figure 2.3, the panels (a) and (d) present the empirical Esscher parameters obtained for 200 returns (and 10,000 repetitions) in the Black-Scholes and Heston worlds, respectively. We note a cloud of points when the size of the sample is small. Histograms in panels (b) and (e) show that the Esscher parameter is symmetric in the Black-Scholes world and negatively skewed in the Heston world, what can suggest that the probability distribution of the empirical Esscher parameter follows the behavior of data generating process. The same behavior was observed in others maturities. The panels (c) and (f) present the empirical parameters for 5x10 returns (and 200 repetitions) in the Black-Scholes and Heston worlds, respectively. These figures show that the empirical parameter converges for one specific value when the sample increases.

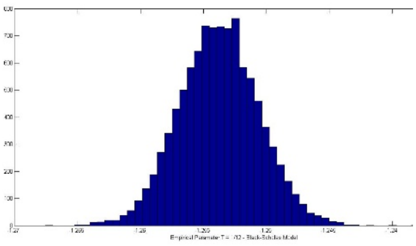
Tables 2.7 and 2.8 present the main descriptive statistics of the empirical parameter in the Black-Scholes and Heston world. We can highlight that the standard deviation decreases along with the maturity and with the increase in the sample size, and the statistics' values begin to converge to a constant value in larger samples.

Tables 2.9 and 2.10 present the main descriptive statistics of the empirical parameters of both methods: Esscher (θ^*) and Stutzer (γ^*). Note that the values are close. When we compare only the empirical Esscher parameter obtained for synthetic data with the one obtained for real data, the more important change is the signal. That is, the Esscher parameters obtained with synthetic data are simulated with a drift $(\mu = 10.00\%)$ greater than the risk-free rate $(r = 5.00\%)$. Thus, the negative parameter shifts the risk-neutral distribution to the left, what eliminates the risk premium and assures the average yield equal to risk-free rate. In real data, the opposite happens. The positive parameter shifts the risk-neutral distribution to the right. This is contrary to financial theory. However, this does not constitute an arbitrage opportunity, because the daily risk-free rate is between the worst and the best daily return (see Cox et al, 1979). As we can see in figures 2.1 and 2.2, and in Table 2.2, the price time series are in fall, and in this case, applications in risk-free interest rates are paying more than these stocks.

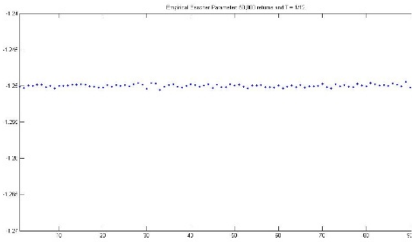
Black-Scholes World



(a) 200 returns

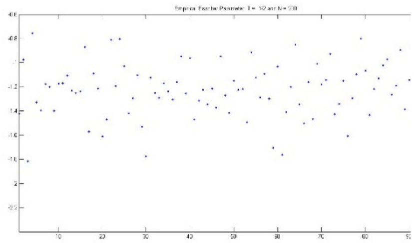


(b) Histogram for 200 returns

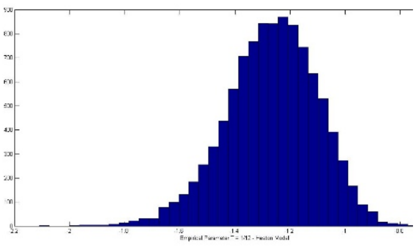


(c) $5x10^4$ returns

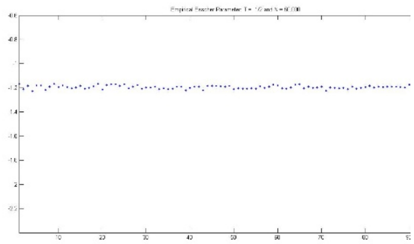
Heston World



(d) 200 returns



(e) Histogram for 200 returns



(f) $5x10^4$ returns

Figure 2.3: Empirical Esscher Parameter T=1/12.

Table 2.7: Empirical Escher Parameter in Black-Scholes World.

	200 returns and 10.000 repetitions				5x10^4 returns and 200 repetitions			
	$T = 1/12$	$T = 1/4$	$T = 1/2$	$T = 1$	$T = 1/12$	$T = 1/4$	$T = 1/2$	$T = 1$
Mean	-1.2538	-1.2539	-1.2539	-1.2543	-1.2500	-1.2500	-1.2500	-1.2500
Std Deviation	0.0032	0.0056	0.0079	0.0110	0.0002	0.0003	0.0005	0.0008
Maximum	-1.2379	-1.2255	-1.2194	-1.2020	-1.2495	-1.2492	-1.2487	-1.2482
Minimum	-1.2676	-1.2792	-1.2836	-1.3007	-1.2506	-1.2509	-1.2518	-1.2520

Table 2.8: Empirical Escher Parameter in Heston World.

	200 returns and 10.000 repetitions				5x10^4 returns and 200 repetitions			
	T = 21	T = 63	T = 126	T = 252	T = 21	T = 63	T = 126	T = 252
Mean	-1.2665	-1.2545	-1.2434	-1.2318	-1.2484	-1.2366	-1.2264	-1.2139
Std Deviation	0.1665	0.1548	0.1406	0.1239	0.0116	0.0098	0.0088	0.0078
Maximum	-0.6649	-0.6557	-0.7943	-0.8156	-1.2189	-1.2126	-1.2073	-1.1982
Minimum	-2.1098	-2.0004	-1.8774	-1.7747	-1.2758	-1.2568	-1.2509	-1.2378

Table 2.9: Empirical Parameters (Esscher and Stutzer) – Petrobras.

	Esscher (θ^*)				Stutzer (γ^*)			
	T = 17	T = 40	T = 59	T = 121	T = 17	T = 40	T = 59	T = 121
Mean	2.8557	2.8891	2.9031	2.9309	2.8002	2.8253	2.8327	2.8460
Std Deviation	0.9265	0.6626	0.5826	0.4915	0.9276	0.6652	0.5881	0.5127
Maximum	7.0618	6.4443	5.7783	5.4746	6.9985	6.3773	5.6778	5.5471
Minimum	-0.5597	0.6241	0.7970	1.3742	-0.6260	0.5585	0.7238	1.1808

Table 2.10: Empirical Parameters (Esscher and Stutzer) – Vale.

	Esscher (θ^*)			Stutzer (γ^*)		
	T = 17	T = 40	T = 59	T = 17	T = 40	T = 59
Mean	4.8156	4.8404	4.8830	4.7550	4.7724	4.8129
Std Deviation	1.0412	0.8152	0.7700	1.0451	0.8298	0.7974
Maximum	9.7780	8.4302	8.8304	9.7615	8.4394	8.8084
Minimum	0.9284	2.2304	2.7088	0.8686	2.1183	2.3691

2.6 Conclusions

In this paper, we propose an empirical version of the Esscher transform for nonparametric option pricing. We conduct artificial experiments in Black-Scholes and Heston worlds and real experiments to explore the potential usefulness of the proposed method.

Artificial results show that the EET prices improve alongside sample size. EET also provides higher prices for all maturities, and the MAPE decreases as moneyness does.

Real data results show that, when the stochastic process of underlying asset is unknown, the lowest pricing errors are between the nonparametric methods. For a maturity equal to 17, the nonparametric methods present similar results. For others maturities, the proposed method presents the lowest MAPE for moneyness equal to deep-out-of-the-money, out-of-the-money and at-the-money.

We also analyze the behavior of the empirical Esscher parameter. We can highlight that the standard deviation decreases with the maturity and with the increase in the sample size, and the values of the descriptive statistics begin to converge to a constant value in larger samples. When we compare only the empirical Esscher parameter obtained for synthetic and real data, the more important change is the signal. That is, the Esscher parameters obtained with synthetic data are simulated with a drift $(\mu = 10.00\%)$ greater than the risk-free rate $(r = 5.00\%)$. Thus, the negative parameter shifts the risk-neutral distribution to the left, what eliminates the risk premium and assures the average yield equal to risk-free rate. With real data, the opposite happens. The positive parameter shifts the risk-neutral distribution to the right. This is contrary to financial theory. However, this does not constitute an arbitrage opportunity, because the daily risk-free rate is between the worst and the best daily return. Price time series are have been falling and in this case, applications in risk-free interest rates are paying more than in these stocks.

Further research can be done comparing the proposed method to other nonparametric pricing methodologies, verifying Monte Carlo simulation techniques and obtaining the Greeks.

3
Dynamic Conditional Score Model

Chapter Abstract:

We propose a method to nonparametric option pricing under a GARCH framework with non-Gaussian innovations. Several papers have extended nonparametric option pricing and provided evidence that this methodology performs adequately in the presence of realistic financial time series. To represent a realistic time series, we use a new class observation driven model, called dynamic conditional score model, proposed by Harvey (2013), for modeling the volatility (and heavy tails) of the asset price. These models use the score of the conditional distribution instead of square observations, what mitigates the effect of outliers. We identify our risk-neutral measure from empirical Esscher transform. We compare our proposal with the Black-Scholes (1973) and Heston and Nandi (2000) models in experiments with real data.

Keywords: nonparametric option pricing, Beta-t-GARCH, Empirical Esscher transform.

3.1
Introduction

Volatility is perhaps the most used risk measure in finance and it is a key factor in option pricing and asset allocation. Although asset return volatility is well defined, it is not directly observable. What we observe are prices of an asset and of its derivatives, leading to several important implications in studying and modeling the volatility. Empirical evidences also show that asset return volatility is stochastic and mean reverting, it responds asymmetrically to positive and negative returns, and the systematic patterns in implied volatility indicate that only an asymmetric probability distribution with heavier tails would be able to produce market prices (Tsay, 2013).

In option pricing, there are two directions in which we model volatility: continuous time and discrete time models. Christoffersen, Heston and Jacobs (2011) argue that continuous time models have become the workhorse of modern option pricing theory, given that they offer closed-form solutions for European option and have the flexibility of incorporating stochastic volatility, stochastic jumps, leverage effects and various types of risk premia (Heston, 1993, Bakshi, Cao and Chen, 1997 and Broadie, Chernov and Johannes, 2007).

In a discrete time setting, the stochastic volatility is often modeled using extensions of the autoregressive conditional heteroscedasticity (ARCH) model proposed by Engle (1982) and generalized by Bollerslev (1986) through

(GARCH) framework. When regard to option pricing, they offer, at least, five advantages with regard to continuous time pricing models. First, it may be considered an accurate numerical approximation of continuous time models (Nelson and Cao, 1992 and Nelson, 1996) avoiding any discretization bias. Second, their predictions are exactly compatible with the filter to extract the variance. Third, estimation is computationally fast. Fourth, the volatility is observable in each time point. Fifth, they can incorporate multiple factors (Engle and Lee, 1999), long memory (Bollerslev and Mikkelsen, 1996), and non-Gaussian innovations (Bollerslev, 1987, and Nelson, 1991).

The theoretical price of an option is evaluated as the discounted expected value of the payoff function under a martingale measure. In general, to change the physical measure to an artificial one, in continuous time models, used the Girsanov theorem.[8] In discrete time models, according to Christoffersen, Elkamhi, Feunou and Jacobs (2010), a class of pricing kernels, or Radon-Nikodym, can be specified and impose restrictions that ensure the existence of an artificial measure.

Obviously, option pricing models with Gaussian innovations cannot capture the skewness, the kurtosis and the leptokurtosis of the financial data. The use of stochastic jumps is the preferred approach to deal with the shortcoming of models with Gaussian innovations in continuous time (Bates, 2000, Eraker, Johannes and Polson, 2003, Chernov, Gallant, Ghysels and Tauchen, 2003). In discrete time, the moment generating function is only able to risk-neutralize a few probabilities distributions, beyond the Gaussian case. For example, gamma distribution (Siu, Tong and Yang, 2004), inverse gaussian innovations (Christoffersen, Jacobs and Heston, 2006), smoothly-truncated stable distribution (Menn and Rachev, 2005), generalized error distributions innovations (Christoffersen, Jacobs and Minouni, 2006) and generalized hyperbolic innovations (Chorro, Guégan and Ielpo, 2008 and Badescu, Elliott, Kulperger, Miettinen and Siu, 2011).

In situations where we need to model innovations with heavy-tailed distributions such as Student't, we can use three methods, according to Liu, Li and Ng (2015). The first method used is a Markov-switching GARCH model with Student's t innovations. To avoid changing the measure, Satoyoshi and Mitsui

[8] See appendix 6.3.

(2011) replaced the drift term in the conditional mean equation with the risk-free interest rate. They justified the drift term's replacement by using the assumption that investors in the real world are risk-neutral, requiring no compensation for risk.

The second method does not require any distributional assumption for the innovations. The asymmetric GJR-GARCH model[9], implemented by Barone-Adesi, Engle, and Mancini (2008), consists of two steps: first, they estimate the GARCH parameters using historical returns; in the second step, they calibrate the model to the observed market prices. The drift term in the conditional mean equation is determined in such a way that the expected asset return implied by the model was the risk-free interest rate.

The third method, propose by Badescu and Kulperger (2007), is similar to that of Barone-Adesi et al (2008), but without resorting to option prices. The authors estimate the GARCH parameters by quasi-maximum likelihood and then, approximate the unknown innovation distribution function using a kernel density estimator based on the standardized residuals. They compute option prices by simulating stock prices under the physical measure and by evaluating Radon-Nikodym derivatives.

Liu et al (2015) proposes one more method for heavy-tailed distributions under GARCH models with Hansen's skewed-t distributed innovations. They use the canonical valuation method, developed by Stutzer (1996), to identify a risk-neutral measure. That is, the risk-neutralization is applied to the empirical distribution of the sample paths generated from the assumed model. The change of measure does not involve the distribution of the model's innovations, this method of risk-neutralization is applicable even when the moment generating function of the innovations' probability distribution does not exist. Maximum entropy principle is then employed to transform the empirical distribution of the sample of future asset returns distribution into its risk-neutral counterpart, by minimizing the Kullback–Leibler information criterion (KLIC).

In a spirit similar to the canonical valuation method of Stutzer (1996), Duan (2002) develops a nonparametric option pricing theory without resorting to option prices. He formalizes the risk-neutralization process so that one can infer

[9] This model is used to handle leverage effects. The model uses zero as its threshold to separate the impacts of past shocks though a variable indicator for negative values.

directly from the price dynamics of the underlying asset to establish the risk-neutral pricing dynamic under GARCH framework.[10] Transforming one-period asset return empirical distribution to normality is a key step in constructing this nonparametric option pricing theory. Applying the relative entropy principle with the condition that the expected asset returns are equal to the risk-free rate, one can derive the risk-neutral distribution for the normalized asset return.

In this work, we propose a method to obtain European option prices under a GARCH framework with non-Gaussian innovations. Several papers have extended nonparametric option pricing and demonstrated that this methodology performs very well in the presence of realistic financial time series. To represent a realistic financial time series, we use a new class of time series models, dynamic conditional score, proposed by Harvey (2013), for modeling the volatility (and heavy tails) of the observed underlying asset price. These models replace the observations, or their squares, by the score of the conditional distribution. They are more robust in extreme events, when compared to standard GARCH, allowing the modeling of leverage effect, adding components of short and long-term volatility.[11]

To avoid the formulation of a restricted model, risk-neutralization is applied to the empirical distribution of the sample paths generated from the assumed model, as Liu et al (2015) and Duan (2002). To identify a risk-neutral measure, we use the empirical Esscher transform. Hence, the sample paths are reweighted, giving rise to a risk-neutralized sample from which option prices can be obtained by a weighted sum of the options' payoffs in each path. We will compare our approach empirically to two competing benchmarks: Black-Scholes (1973) and Heston and Nandi (2000).

The paper is organized as follows. In section 3.2 we introduce the proposed method. Section 3.3 presents the methodology used to compare the different pricing methods, and the results are discussed in section 3.4. Finally, section 3.5 concludes.

---

[10] It is in this regard that his nonparametric option pricing theory differs from the canonical evaluation method of Stutzer (1996).

[11] In this case it is used an exponential function as link, as the asymptotic distribution of the maximum likelihood estimators can be derived from it.

3.2 Proposed Method

In this section, we develop a data generator process dynamic that captures the most important features of the return process.

3.2.1 Asset Price Dynamic

Under physical measure the GARCH model assumed is:

$$y_t = ln(S_t/S_{t-1}) = \mu + a_t, \qquad a_t = \sqrt{h_t}z_t \tag{3.1}$$

where S_t is the asset price at time t, μ is a constant, h_t is the conditional variance, with its evolution captured by some GARCH-type model, and $\{z_t\}$ is the sequence of standardized innovations, which are independent and identically distributed random variables with zero mean and unit variance. It follows the conditional mean and variance of y_t are μ and h_t, respectively.

According to Tsay (2013), for most asset return series, the serial correlations are weak, if there is any. Thus, building a mean equation, μ, to remove the sample mean from data if the sample mean is significantly different from zero. In some cases, a simple auto-regressive model might be needed. In other cases, the equation mean may employ some explanatory variables, such as indicator variables to capture possible daily effects.

In option pricing, the form of the equation (3.1) depends on an underlying risk-neutral model. For example, in Duan (1995), μ was replaced by $r + \lambda\sqrt{h_t} - h_t/2$, so that the expected value in risk-neutral measure is equal to the risk-free rate.

3.2.2
Return volatility

Time series models in which a parameter of conditional distribution is a function of past observations are widely used in econometrics. Such models are termed ‘observation driven’ as opposed to ‘parameter driven’. Examples of observation driven models are the class of GARCH models. The stochastic volatility models are parameter driven in which the volatility is determined by an unobserved stochastic process, such as the Heston (1993) model. The works of Harvey (2013) and Creal, Koopman and Lucas (2013) propose a new approach to the formulation of observation driven models where time-varying parameters are driven by the score.[12]

In both cases, when the disturbance is Gaussian, the models are equivalent to the standard GARCH (1,1). If we assume that the disturbance follows a Student t distribution, there is an important difference between the standard t-GARCH (1,1) model of Bollerslev (1987) and the Harvey (2013) one: the updating equation for the conditional volatility is not the same in both models. According to Harvey (2013), the standard GARCH model responds too much to extreme observations, and this effect is slow to dissipate. Letting the dynamic equation for volatility depend on the conditional score of the t distribution, as it is the case of Harvey’s set up, mitigates the effect of outliers. Creal et al (2013) cite that, if the errors are modeled by a fat-tailed distribution, a large observation causes a more moderate increase in the variance when the dynamic equation depends on the conditional score.

Harvey (2013) presents the following models for volatility: Beta-t-(E)GARCH and Gamma-GED-(E)GARCH. In the first, the conditional score is a linear function of a variable that has a beta distribution, while in the latter the conditional score is a linear function of a variable that has a gamma distribution.

The Beta-t-GARCH model depends on ensuring a positive variance and the asymptotic distribution of the maximum likelihood estimators cannot be

[12]This idea was suggested independently in papers by Creal et al (2013) and Harvey and Chakravarty (2009). Creal et al (2013) went on to develop a whole class of score driven models, while Harvey and Chakravarty (2009) concentrated on GARCH type models. Creal et al (2013) named their work as Generalized Autoregressive Score (GAS), while Harvey (2013) refers to them as dynamic conditional score. However, only in Harvey’s paper the model asymptotic theory was addressed.

obtained. In Beta-t-EGARCH model, it works with a scale[13] which employs an exponential link function, and it is possible to derive closed form expressions for multistep predictions, moments the autocorrelations of absolute values together with the mean square error of these predictions. Moreover, it has all the advantages of the standard EGARCH model and the asymptotic distribution of the maximum likelihood estimators can be derived analytically.

When the conditional distribution of the innovations has a Generalized Error Distribution (GED), the Dynamic Conditional Score (DCS) approach leads to a complementary class of models in which score is a linear function of observations' absolute values raised to a positive power. These variables can be transformed in order to have a gamma distribution, so the properties of the model, denoted as Gamma-GED-EGARCH, can again be obtained. The normal distribution is a special case of the GED, as is the Laplace distribution. In his empirical studies, Harvey (2013) recognizes that Beta-t-(E)GARCH models overcome gamma-GED-(E)GARCH models. Moreover, Beta-t-GARCH model can be considered as approximating a Beta-t-EGARCH model. This work is concerned with the Beta-t-GARCH model for modeling the volatility.

3.2.2.1 Model specification

Let y_t be the dependent variable of interest and h_t the time-varying parameter, all at time t. The available information set at a time t consists of $Y_{t-1} = \{y_1, y_2, \dots, y_{t-1}\}$ and $h_{t-1} = \{h_1, h_2, \dots, h_{t-1}\}$. An observation-driven model is set up in terms of a conditional distribution for the *t-th* observation. The model assumes that y_t is generated by the observation probability density function (pdf):[14]

$$f(y_t | Y_{t-1}, h_{t-1}),\ t = 1, \dots, T. \tag{3.2}$$

The mechanism for updating the time-varying parameter h_t is given by:

---

[13] The scale is defined as $\varphi_t *= (\upsilon - 2)^{1/2} h_t$. The dynamic equation is then set up for the logarithm of scale $\lambda_t = ln\varphi_t$.

[14] The appendix 6.2 presents the link between equations (3.1) and (3.2).

$$\eta_{t-1} = \psi \frac{\partial lnf(y_t|Y_{t-1}, h_{t-1})}{\partial h_{t-1}} \tag{3.3}$$

where ψ is a finite constant which may be the information matrix, or some other constant, including unity. Let the information matrix be:

$$\psi = \left\{Var\left[\frac{\partial lnf(y_t|Y_{t-1}, h_{t-1})}{\partial h_{t-1}}/Y_{t-1}\right] = E\left[\left(\frac{\partial lnf(y_t|Y_{t-1}, h_{t-1})}{\partial h_{t-1}}\right)^2 \Big/ Y_{t-1}\right] = -E\left[\frac{\partial^2 lnf(y_t|Y_{t-1}, h_{t-1})}{\partial h_{t-1}^2}/Y_{t-1}\right]\right\}^{-1}. \tag{3.4}$$

The derivative $\partial lnf(y_t|Y_{t-1}, h_{t-1})/\partial h_{t-1}$, or score of the conditional distribution, is a random variable which has zero mean at the true parameter value. This variable is a martingale difference by construction,[15] as it is shown below:

$$E[\eta_{t-1}/Y_{t-1}] = \psi E\left[\frac{\partial lnf(y_t|Y_{t-1}, h_{t-1})}{\partial h_{t-1}}/Y_{t-1}\right] = 0 \tag{3.5}$$

$$Var[\eta_{t-1}/Y_{t-1}] = E\left[\left(\psi \frac{\partial lnf(y_t|Y_{t-1}, h_{t-1})}{\partial h_{t-1}}\right)^2 \Big/ Y_{t-1}\right] = \psi^2 E\left[\left(\frac{\partial lnf(y_t|Y_{t-1}, h_{t-1})}{\partial h_{t-1}}\right)^2 \Big/ Y_{t-1}\right] \tag{3.6}$$

$$E\left[\left(\frac{\partial lnf(y_t|Y_{t-1}, h_{t-1})}{\partial h_{t-1}}\right)^2 \Big/ Y_{t-1}\right] = \frac{Var[\eta_{t-1}/Y_{t-1}]}{\psi^2} = \frac{\sigma_\eta^2}{\psi^2}. \tag{3.7}$$

In the dynamic conditional score models, the idea is to replace the observation, or their squares, in the dynamic equation for the volatility, by the

[15] When markets are working efficiently, returns are martingale differences. In other words, they should not be predictable on the basis of past information. However, returns are not usually independent, and so features of the conditional distribution apart from the mean may be predictable.

score of the conditional distribution. For example, let the first-order Gaussian GARCH model be, according to equation (3.1) with $z_t \sim NID(0,1)$, then:

$$h_t = \delta + \beta h_{t-1} + \alpha a_{t-1}^2, \qquad \delta > 0, \alpha \geq 0, \beta \geq 0. \tag{3.8}$$

The conditions on δ, α and β ensure that the variance remains positive. The sum of α and β is typically close to one. The logarithm of the pdf is:

$$lnf(y_t|Y_{t-1}, \mu, h_{t-1}) = ln\left[\frac{1}{\sqrt{2\pi}}\right] - \frac{1}{2} lnh_{t-1} - \frac{(y_{t-1} - \mu)^2}{2h_{t-1}} \tag{3.9}$$

and the first derivative:

$$\frac{\partial lnf(y_t|Y_{t-1}, \mu, h_{t-1})}{\partial h_{t-1}} = \frac{1}{2h_{t-1}^2}(a_{t-1}^2 - h_{t-1}). \tag{3.10}$$

The information matrix is:

$$\psi = -\frac{1}{E\left[\frac{\partial^2 lnf(y_t|Y_{t-1}, h_{t-1})}{\partial h_{t-1}^2} / Y_{t-1}\right]} = 2h_{t-1}^2 \tag{3.11}$$

and by replacing (3.10) and (3.11) in (3.3), we have:

$$\eta_{t-1} = a_{t-1}^2 - h_{t-1}. \tag{3.12}$$

Rewriting it (3.12) as:

$$a_{t-1}^2 = \eta_{t-1} + h_{t-1} \tag{3.13}$$

and finally replacing (3.13) in (3.8), we have:

$$h_t = \delta + \phi h_{t-1} + \alpha \eta_{t-1}, \qquad \delta > 0, \phi \geq \alpha, \alpha \geq 0 \tag{3.14}$$

where $\phi = \alpha + \beta$ and η_{t-1} is given by (3.12). In the normal case, the mechanism for updating the h_t, given by score of the conditional distribution is similar to standard Gaussian GARCH(1,1).

3.2.2.2
Beta-t-GARCH Model

When the observations have a conditional t distribution, with υ degrees of freedom, we can write z_t in (3.1) as:

$$z_t = \left(\frac{\upsilon - 2}{\upsilon}\right)^{\frac{1}{2}} \varepsilon_t, \qquad \upsilon > 2, \tag{3.15}$$

where the serially independent zero mean variable ε_t has a standard t_υ distribution:

$$\varepsilon_t \sim t_\upsilon \left(0, \frac{\upsilon}{\upsilon - 2}\right). \tag{3.16}$$

Therefore, z_t has a t_υ distribution and standardized to have unit variance. Let the conditional pdf be given by a Student t distribution:

$$f(y_t|Y_{t-1}, \mu, h_{t-1}, \upsilon) = \frac{\Gamma((\upsilon+1)/2)}{\Gamma(\upsilon/2)\sqrt{\pi(\upsilon-2)}} \left(\frac{1}{h_{t-1}}\right)^{\frac{1}{2}} \left(1 + \frac{(y_{t-1} - \mu)^2}{(\nu - 2)h_{t-1}}\right)^{-\frac{\upsilon+1}{2}}. \tag{3.17}$$

Its logarithm is given by:

$$lnf(y_t|Y_{t-1}, \mu, h_{t-1}, \upsilon) = ln\Gamma\left(\frac{\upsilon+1}{2}\right) - ln\Gamma\left(\frac{\upsilon}{2}\right) - \frac{1}{2}ln\pi - \frac{1}{2}ln(\upsilon - 2)$$
$$- \frac{1}{2}lnh_{t-1} - \frac{\upsilon+1}{2}ln\left(1 + \frac{(y_{t-1} - \mu)^2}{(\upsilon - 2)h_{t-1}}\right). \tag{3.18}$$

Consider u_{t-1} another mechanism for updating the time varying parameter for the Student t distribution, similar to (3.3). Then, the score of u_{t-1} is:

$$\frac{\partial lnf(y_t|Y_{t-1},\mu,h_{t-1},\upsilon)}{\partial h_{t-1}} = -\frac{1}{2h_{t-1}}\left\{1-\left[\frac{(\upsilon+1)(y_{t-1}-\mu)^2}{(\upsilon-2)h_{t-1}+(y_{t-1}-\mu)^2}\right]\right\}. \tag{3.19}$$

Making $m = (y_t - \mu)$ and $\omega = (\upsilon - 2)h_{t-1}$, and dividing (3.19) by ω, we have:

$$\frac{\partial lnf(y_t|Y_{t-1},\mu,h_{t-1},\upsilon)}{\partial h_{t-1}} = -\frac{1}{2h_{t-1}}\left\{1-(\upsilon+1)\left[\frac{m^2/\omega}{1+m^2/\omega}\right]\right\}. \tag{3.20}$$

Lemma 1 If g is gamma (θ,α) and s is gamma (θ,β), then $x = g/(s+g)$ is beta (α,β).

Corollary 1 The variable $(m^2/\omega)/(1+m^2/\omega)$ has a beta $(1/2\,,\upsilon/2)$ distribution, whereas $1/(1+m^2/\omega)$ has a beta $(\upsilon/2\,,1/2)$ distribution.

Proof: Since m^2/ω is the ratio of a squared standard normal to a χ^2_υ, it follows from Lemma 1 that $(m^2/\omega)/(1+m^2/\omega) = (g/s)/(1+g/s) = g/(s+g)$. Similarly, $1/(1+m^2/\omega) = s/(s+g)$ ■

If $b \sim$ beta $(1/2\,,\upsilon/2)$, then its mean and variances will be given by:

$$E[b] = \frac{1}{(\upsilon+1)} \tag{3.21}$$

$$Var[b] = \frac{2\upsilon}{(\upsilon+3)(\upsilon+1)^2}. \tag{3.22}$$

The score in (3.20), according to (3.21) and (3.22), has mean zero and variance given by:

$$Var\left[\frac{\partial lnf(y_t|Y_{t-1},\mu,h_{t-1},\upsilon)}{\partial h_{t-1}}/Y_{t-1}\right]=\frac{(\upsilon+1)^2}{4h_{t-1}^2}\frac{2\upsilon}{(\upsilon+3)(\upsilon+1)^2}$$
$$=\frac{\upsilon}{2h_{t-1}^2(\upsilon+3)}. \quad (3.23)$$

Rewriting (3.7), but with u_{t-1}, we have:

$$\psi=\sqrt{\frac{\sigma_u^2}{Var\left[\frac{\partial lnf(y_t;\mu,\varphi,\upsilon)}{\partial h_t}\right]}}=\sqrt{\frac{\frac{2\upsilon}{(\upsilon+3)}}{\frac{\upsilon}{2h_{t-1}^2(\upsilon+3)}}}=2h_{t-1}. \quad (3.24)$$

This model is named Beta-t-GARCH because u_{t-1} is a linear function which its variable has a beta distribution. The principal feature of the Beta-t-GARCH class is that a linear combination of past values of the martingale difference is given by the conditional score:

$$u_{t-1}=\frac{(\upsilon+1)(y_{t-1}-\mu)^2}{(\upsilon-2)h_{t-1}+(y_{t-1}-\mu)^2}-1, \qquad -1\leq u_{t-1}\leq\upsilon, \qquad \upsilon>2. \quad (3.25)$$

The variable u_{t-1} may be expressed as:

$$u_{t-1}=(\upsilon+1)b_{t-1}-1 \quad (3.26)$$

where

$$b_{t-1}=\frac{(y_{t-1}-\mu)^2/(\upsilon-2)h_{t-1}}{1+(y_{t-1}-\mu)^2/(\upsilon-2)h_{t-1}}, \qquad 0\leq b_{t-1}\leq 1,$$
$$0<\upsilon<\infty. \quad (3.27)$$

When $\upsilon=\infty$, $u_{t-1}=a_{t-1}^2/h_{t-1}-1$ and the standard GARCH model in (3.14), is obtained by setting $\eta_{t-1}=h_{t-1}u_{t-1}$. The Beta-t-GARCH (p,q) model, using (3.26), is given by:

$$h_t = \delta + \beta_1 h_{t-1} + \cdots + \beta_q h_{t-q} + \alpha_1 h_{t-1}(\upsilon + 1) b_{t-1} + \cdots + \alpha_q h_{t-q}(\upsilon + 1) b_{t-1-q} \quad (3.28)$$

where $\alpha_i = \theta_i$ and $\beta_i = \phi_i - \alpha_i$, $i = 1, \ldots, q$. In the limit as $\upsilon \to \infty, (\upsilon + 1) b_{t-1} = a_{t-1}^2$, it leads to the standard GARCH specification. A sufficient condition for the conditional variance to remain positive is $\delta > 0$, $\beta_i \geq 0$, and $\alpha_i \geq 0$, $i = 1, \ldots, q$. The beta-t-GARCH (1,1) model is:

$$h_t = \delta + \phi h_{t-1} + \alpha h_{t-1} u_{t-1}. \quad (3.29)$$

The Beta-t-GARCH (1,1) model may be extended to include leverage effects by adding the indicator variable $I(y_{t-1} < 0)$:

$$h_t = \delta + \phi h_{t-1} + \alpha h_{t-1} u_{t-1} + I(y_{t-1} < 0) \alpha^* h_{t-1} u_{t-1}. \quad (3.30)$$

Forecasts for more than one step ahead of the conditional variance can be made for Beta-t-GARCH models, as in standard GARCH case, by using the law of iterated expectations.[16]

3.2.3 Risk-Neutral Measure

Consider a random sample of size n of an asset's log-return for a period T, denoted by $\{y_{T,i}\}_{i=1,n}$. Then, the empirical risk-neutral Q measure is given by $\{q_{i,\hat{\theta}^*}\}_{i=1,n}$, with $\hat{\theta}^* = arg_\theta \left\{ e^{r_T} = \frac{\hat{M}(\theta+1)}{\hat{M}(\theta)} \right\}$. See section 2.3.

3.3 Methodology

This section presents the methodology used to compare the proposed method to artificial and real data. To investigate its applicability in some settings,

[16] See appendix 6.4 for properties of first-order model.

the empirical Esscher transform is applied to price European call option across a range of moneyness and maturities.

The real-world parameters are estimated using the log returns of the closing values of two daily databases of Vale's and Petrobras' prices from January 2, 2011 to January 17, 2012, containing 260 daily observations for each database.[17] We set time 0 to January 17, 2012 (the end point of the data sample period) and Vale's closing value on that day was $S_0 = R\$\,41.13$ and Petrobras' was $S_0 = R\$\,24.37$. The corresponding true market price on the valuation date was used as a benchmark.[18] The true market prices of the options with the strikes and maturities under consideration are shown in Table 2.2. The maturities are equal to $17/252$, $40/252$, $59/252$ and $121/252$ years. The risk-free interest rate was 10.3499% (17/252), 10.2485% (40/252), 10.1721% (59/252) and 10.032% (121/252) obtained by linear interpolations.[19] Table 3.1 presents main descriptive statistics for the time series returns.

Table 3.1: Descriptive Statistics of Log Returns.

	Log return	
Statistics	**Vale**	**Petrobras**
Mean	-0.00076	-0.00048
Standard deviation	0.01740	0.01760
Skewness	-0.64820	-0.67110
Kurtosis	7.05440	5.24140
Maximum	0.05750	0.04530
Minimum	-0.09610	-0.07800

Pricing procedures

The Black–Scholes model for the price, at time 0, of a European call option on a non-dividend-paying stock is given by the formulae: (2.21), (2.22) and (2.23). See section 2.4.

In the model of Heston and Nandi (2000), the logarithmic return $y_t = ln(S_t/S_{t-1})$ is assumed to follow GARCH (1,1) in the mean process driven by the following pair of equations, under physical measure:

[17] These data are specific to the Brazilian market.
[18] All required data are obtained from Bovespa (http://www.bmfbovespa.com.br).
[19] http://www.bcb.gov.br

$$y_t = r + \lambda\sigma_t^2 + \sigma_t z_t \tag{3.31}$$

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha(z_{t-1} - \gamma\sigma_{t-1})^2 \tag{3.32}$$

where r is the risk-free interest rate, λ represents the risk premium, σ_t^2 is the conditional variance, z_t is the error term distributed as a standard normal variable, $z_t \sim N(0,1)$. The α determines kurtosis, γ determines skewness and variance persistence is $\beta + \alpha\gamma^2$, the process will be mean-reverting if $\beta + \alpha\gamma^2 < 1$. The risk-neutral version of this model can be written as:

$$y_t = r - \frac{1}{2}\sigma_t^2 + \sigma_t z_t^* \tag{3.33}$$

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha(z_{t-1}^* - \gamma^*\sigma_{t-1})^2 \tag{3.34}$$

where λ is replaced by -1/2, z_t^* is equal to $z_t^* = z_t + \left(\lambda + \frac{1}{2}\right)\sigma_t$ and γ is replaced by $\gamma^* = \gamma + \lambda - 1/2$. The price at time *t* of a European call option with maturity at time *t*+*T* is given by:

$$C = e^{-rT}E_t^*[(S_{t+T} - K)^+] = S_t P_1 - Ke^{-rT}P_2 \tag{3.35}$$

where T is the time to maturity, $E_t^*[]$ is the expectation at time t under risk-neutral distribution, S_t is the price of the underlying asset at time *t* and P_1, P_2 are the risk-neutral probabilities.

The quantities P_1 and P_2 are the probabilities that can be obtained by inverting the characteristic functions $f^*(\phi)$:

$$P_1 = \frac{1}{2} + \frac{e^{-rT}}{\pi S_t}\int_0^\infty Re\left[\frac{K^{-i\phi}f^*(i\phi + 1)}{i\phi}\right]d\phi \tag{3.36}$$

$$P_2 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty Re\left[\frac{K^{-i\phi}f^*(i\phi)}{i\phi}\right]d\phi. \tag{3.37}$$

$$f(\phi) = S_t^\phi exp(A_t + B_t\sigma_t^2) \tag{3.38}$$

$$A_t = A_{t+1} + \phi r + B_{t+1}\omega - \tfrac{1}{2}log(1 - 2\alpha B_{t+1}) \tag{3.39}$$

$$B_t = \phi(\lambda + \gamma) - \frac{1}{2}\gamma^2 + \beta B_{t+1} + \frac{\frac{1}{2}(\phi - \gamma)^2}{1 - 2\alpha B_{t+1}}. \tag{3.40}$$

Note that A_t and B_t are defined recursively, by working backwards from the maturity time t+T of the option. Note also that A_t and B_t are functions of t, t+T, and ϕ, so that $A_t \equiv A(t; t + T, \phi)$ and $B_t \equiv B(t; t + T, \phi)$. Both these terms can be solved recursively from time t+T, working back through time and using the terminal conditions:

$$A_{t+T} = B_{t+T} = 0. \tag{3.41}$$

The next terms in the backward recursion would be $A_{t+T} = \phi r$ and:

$$B_{t+T} = \phi(\lambda + \gamma) - \frac{1}{2}[\gamma^2 + (\phi - \gamma)^2]. \tag{3.42}$$

As mentioned earlier, for pricing options the risk-neutral distribution must be used. To obtain the risk-neutral generating function $f^*(\phi)$, in all terms A_t and B_t, we replace λ with -1/2, and γ with γ^*.

For our proposal, the algorithm is:

1. Simulate the physical distribution for S_T;
2. Compute the Esscher parameter, $\hat{\theta}^*$, using the equation (2.16);
3. Compute the option price with the equation (2.19);

In the step 1, we simulate the returns from equation (3.1) and the volatility is obtained from equation (3.30), after that we obtain the parameter values through maximum likelihood estimator. For each time to maturity T, 252 returns are drawn to generate the distribution of T-year forward. We repeat the experiments with 5×10^4 returns to analyze if the accuracy improves with an increase on the sample size. We obtain the risk-neutral measure (step 2) and we calculate the option price (step 3). We repeat this procedure 15.000 times (or 200 times) and we calculate the Mean Absolute Percentage Error (MAPE).

For each time to maturity T, we performed bootstrap with replacement on the underlying asset's historical returns. We follow the sequence: (a) we construct a single trajectory for the asset price by drawing a certain quantity of historical log returns. For example, if the option has 17 days to maturity, then we draw the same quantity; (b) we accumulate the log returns of this trajectory and we obtain one price; (c) we repeat the process (a) and (b) 252 times (or 5×10^4 times) to construct the physical distribution for the price at maturity. We obtain the risk-neutral measure (step 2) and we calculate the option price (step 3). We repeat this procedure 15.000 times (or 200 times) and we calculate the Mean Absolute Percentage Error (MAPE).

3.4
Results

Testing for serial dependence in the data

The basic idea behind volatility study is that the series of log returns is serially uncorrelated or with minor lower order serial correlations, but is a dependent series. The Ljung-Box statistics[20] show that *Q(12) = 22.173* with *p* value *0.03563* for Vale, except lag 3, and *Q(12) = 13.275* with *p* value *0.3494* for Petrobras.

The Ljung-Box statistics for the squared log returns show that *Q(12)= 35.558* with *p* value *3.814e-04* for Vale and *Q(12) = 43.438* with *p* value *1.903e-05* for Petrobras. We can conclude that daily log returns of Vale's and Petrobras' stocks are serially uncorrelated, but non-linearly dependent.

If we rewrite (3.1) as $\tilde{a}_t = \tilde{y}_t - \tilde{\mu}$, where $\tilde{\mu}$ is the mean of y_t (observed market returns), then $\tilde{a}_t$ will be the centered return. The squared series $\tilde{a}_t^2$ is then used to check for conditional heteroscedasticity, which is also known as the ARCH effect. The Ljung-Box statistics of $\tilde{a}_t^2$ show strong ARCH effects with

[20] The null hypothesis of the test statistic is $H_0: \rho_1 = \cdots = \rho_m = 0$ and alternative hypothesis is $H_a: \rho_i \neq 0$ for some i between 1 and m. Under the null hypothesis, $Q(m) = T(T+2)\sum_{l=1}^{m} \frac{\hat{\rho}_l^2}{T-l}$ follows asymptotically a chi-squared random variable with *m* degrees of freedom. The decision rule is to reject H_0 if $Q(m) > \chi_\alpha^2$, where χ_α^2 denotes the $100(1-\alpha)$th percentile of a chi-squared distribution with *m* degrees of freedom.

Q(12) = 36.011, the *p* value of which is 3.227e-04 for Vale and *Q(12) = 43.849* with *p* value 1.62e-05 for Petrobras.

Parameter estimation

Regarding the optimization of parameters, we used the following heuristics. From the sets of generated initial conditions, we start the optimization using the Nelder-Mead method. The values of the resulting parameters will serve as initial conditions for optimization through BFGS method, and these are later passed to the Nelder-Mead method and so on, until the difference between the solutions is less than a tolerance value of 0.10.

All parameters were estimated by maximum likelihood method from the underlying asset's historical prices. In the Heston-Nandi (2000) model, the estimated parameter values to Vale data were: $\widehat{\omega} = 1.6835e - 05$, $\hat{\beta} = 0.9295$, $\hat{\alpha} = 2.9611e - 06$, $\hat{\gamma} = 5.1837e - 05$, $\hat{\lambda} = 2.2808e - 07$. And the estimated parameter values to Petrobras data were: $\widehat{\omega} = 2.51387e - 06$, $\hat{\beta} = 0.9866$, $\hat{\alpha} = 1.7055e - 06, \hat{\gamma} = 3.6264e - 06, \hat{\lambda} = 0.1407$.

In the proposed method, the estimated parameter values to Vale data were: $\delta = 2.7639\text{e} - 05$, $\phi = 0.9006$, $\alpha = 0.1627$, $\alpha^* = 0.0413$, $\mu = -6.4173\text{e} - 04$, $\upsilon = 6.3983$. And the estimated parameter values to Petrobras data were: $\delta = 2.0921\text{e} - 05$, $\phi = 0.9307$, $\alpha = 0.0577$, $\alpha^* = 0.0684$, $\mu = -1.3085\text{e} - 04, \upsilon = 6.9949$.

The sufficient condition for the conditional variance to remain positive was obtained ($\delta > 0$, $\phi \geq 0$, $\alpha \geq 0$ and $\alpha^* \geq 0$) and y_t is strictly stationary and ergodic because $\phi < 1$. The estimated degrees of freedom (υ) have values that are typical of distribution with heavy tails. The leverage effect parameter (α^*) also shows that a positive y_t contributes $\alpha u_{t-1} h_{t-1}$ to h_t, whereas a negative y_t has a larger impact $(\alpha + \alpha^*) u_{t-1} h_{t-1}$ with $\alpha^* \geq 0$. Figures 3.1 and 3.2 show the estimated volatility with the absolute value of returns for Vale and Petrobras, respectively. We observe the same robustness for conditional volatility in the presence of outliers, as seen in Harvey (2013).

Diagnostics

We can check the adequacy of a fitted model by examining its series of standardized residuals. In particular, the Ljung-Box statistics of $\hat{a}_t$.[21] can be used to check the adequacy of the mean equation and $\hat{a}_t^2$ can be used to test the validity of the volatility equation.

In Beta-t-GARCH (1,1), the Ljung-Box test on the standardized residuals gives *Q(12)* = 16.944 with *p* value 0.1517 and the *Q-statistics* of $\{\hat{a}_t^2\}$ give *Q(12)=5.3256* with *p* value *0.9462* for Vale. For Petrobras, the results are *Q(12) = 12.798* with *p* value *0.3839,* while the *Q-statistics* of $\{\hat{a}_t^2\}$ give *Q(12) = 13.435* with *p* value *0.3382.*

In the GARCH (1,1) of Heston-Nandi (2000), the Ljung-Box test on the standardized residuals gives *Q(12)* = 20.086 with *p* value 0.06547 and the *Q-statistics* of $\{\hat{a}_t^2\}$ give *Q(12)=26.679* with *p* value *0.08592* for Vale. For Petrobras, the results are *Q(12) = 13.992* with *p* value *0.3012,* while the *Q-statistics* of $\{\hat{a}_t^2\}$ give *Q(12) = 13.752* with *p* value *0.3382.*

The test results show no significant serial correlations in the squared standardized residuals, suggesting that the models are sufficient to explain the heteroscedasticity in the log returns series of both stocks.

We investigate if the standardized residuals have a normal distribution (null hypothesis) through the Jarque Bera test.[22] In the beta-t-GARCH (1,1), the values are: p-value = 1.991e-13 for Petrobras and p-value = 2.997e-06 for Vale. In the GARCH (1,1) of Heston-Nandi (2000), the values are equals to: p-value = 2.2e-16.

[21] Where $\hat{a}_t = (y_t - \hat{\mu})/\sqrt{h_t}$ are standardized residuals and $\hat{\mu}$ is estimated by maximum likelihood.
[22] This test statistic is asymptotically distributed as chi-square random variables with 2 degrees of freedom.

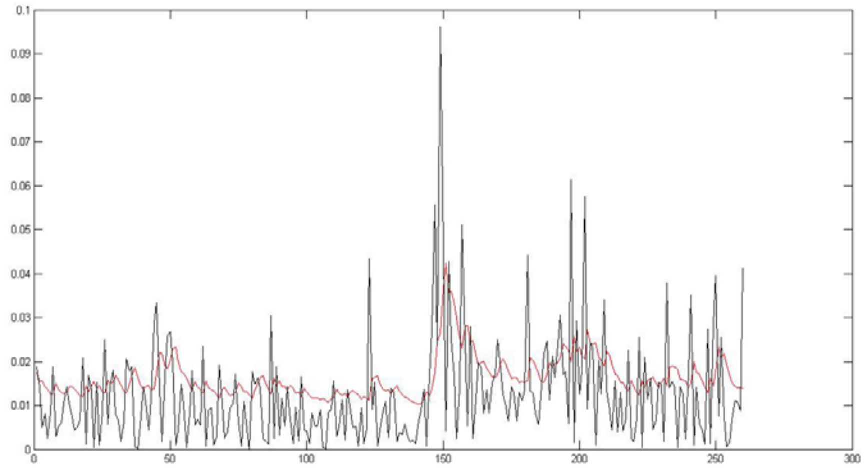


Figure 3.1: Estimated Volatility and the Absolute log returns from Vale.

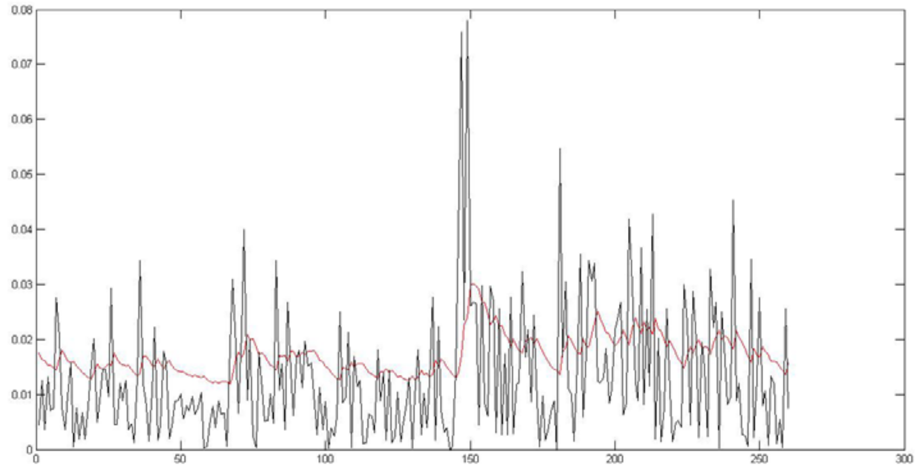


Figure 3.2: Estimated Volatility and the Absolute log returns from Petrobras.

Monte Carlo simulation analysis

Table 3.2 presents the results of the mean absolute percentage error (MAPE) obtained between pricing methods and the Petrobras database. The proposed method, with an assumed model to describe the empirical distribution of prices, in most cases, presents the lowest MAPE. Table 3.3 presents the results of the MAPE obtained between pricing methods and the Vale database. This table showed similar results to Table 3.2.

Table 3.4 shows the results of the MAPE obtained between the proposed method and the Petrobras database. We increased the size of the empirical distribution to analyze the impact on prices calculated by proposed method. For lower maturity (17/252), the increase of the sample reduced the MAPE. For other periods, a sample of 252 observations tends to have smaller MAPE. Table 3.5 shows the results of the MAPE obtained between the proposed method and the Vale database. This table showed similar results to Table 3.4.

Table 3.2: MAPE of Empirical Esscher Transform Estimates for Petrobras.

This table contains the prices for a European call option from EET-AM (empirical Esscher transform - assumed model), EET-B (bootstrap with replacement on historical returns), BS (Black-Scholes) and HN (Heston-Nandi) methods for different moneyness, maturities and they are compared to the true market price of Petrobras data. The numbers reported for each combination are the mean absolute percentage error (MAPE). In the proposed method, we use 252 returns and the simulation is repeated 15,000 times.

Maturity	Moneyness (spot/strike)		EET-AM	EET-B	BS	HN
T = 17/252	Deep-out-of-the-money	0.95	30.0697	51.6393	63.7784	77.4953
	Out-of-the-money	0.97	39.0752	56.8726	64.7611	78.4734
		0.98	16.1753	28.2949	32.8115	43.1060
	In-the-money	1.03	1.6857	7.5260	8.1260	13.4044
		1.07	0.8618	2.1482	2.0521	5.2937
	Deep-in-the-money	1.13	0.3616	0.3637	0.1661	1.8891
		1.17	1.4873	1.6451	1.5740	2.7965
		1.24	1.9716	1.9494	1.9618	1.1296
		1.31	0.9664	0.9686	0.9669	1.1198
		1.38	0.7244	0.7245	0.7244	0.7847
		1.61	1.3303	1.3303	1.3303	1.3654
T =40/252	Deep-out-of-the-money	0.88	97.0938	112.7758	140.7773	91.3502
		0.90	81.8346	90.5321	107.5664	91.6974
		0.94	56.2463	56.8326	64.2107	64.7549
		0.96	29.6578	32.5719	37.0929	39.6310
	Out-of-the-money	0.97	17.9689	23.0286	26.3649	29.5165
	At-the-money	1.02	8.0226	13.4987	14.6670	18.5379
	In-the-money	1.07	2.9488	6.5800	6.9186	10.3816
	Deep-in-the-money	1.13	2.2976	3.7873	3.7871	6.6295
		1.16	1.4679	2.2564	2.2069	4.6907
		1.24	1.7016	1.8856	1.8454	3.7674
		1.31	3.8063	3.8681	3.8496	5.4948
		1.37	3.4178	3.4406	3.4331	3.9858
T = 59/252	Deep-out-of-the-money	0.94	40.0601	47.5063	53.6051	54.6439
	At-the-money	1.02	19.8938	19.0197	20.5533	24.8240
	Deep-in-the-money	1.12	4.1748	3.1370	3.3037	6.7046
T = 121/252	Deep-out-of-the-money	0.96	0.6364	5.1689	3.9998	7.5215

Table 3.3: MAPE of Empirical Esscher Transform Estimates for Vale.

This table contains the prices for a European call option from EET-AM (empirical Esscher transform - assumed model), EET-B (bootstrap with replacement on historical returns), BS (Black-Scholes) and HN (Heston-Nandi) methods for different moneyness, maturities and they are compared to the true market price of Vale data. The numbers reported for each combination are the mean absolute percentage error (MAPE). In the proposed method, we use 252 returns and the simulation is repeated 15,000 times.

Maturity	Moneyness (Spot/strike)		EET-AM	EET-B	BS	HN
T = 17/252	Deep-out-of-the-money	0.93	132.3809	177.7454	207.6516	149.4234
		0.95	102.9132	136.3859	154.5106	124.6248
	Out-of-the-money	0.98	42.8648	62.7852	69.9354	59.8857
	At-the-money	1.00	15.7799	29.4206	32.8052	28.5659
		1.00	18.2929	31.6931	34.8277	30.9905
		1.02	0.6636	9.1424	10.4689	9.1643
	In-the-money	1.05	1.5949	6.2268	6.6073	6.6472
		1.07	7.6055	11.3211	11.4716	11.8876
		1.11	0.7302	2.0628	1.9543	2.1187
		1.11	0.0997	1.0810	0.9691	1.8023
	Deep-In-the-money	1.14	1.4029	0.8775	0.9795	0.1384
		1.14	0.8132	1.2765	1.1766	2.0321
		1.18	5.3185	5.4862	5.4232	6.2524
		1.21	0.7264	0.7801	0.7516	1.4623
		1.36	1.8745	1.8742	1.8745	1.8421
		1.37	2.7025	2.7022	2.7025	2.6711
		1.47	1.9758	1.9758	1.9759	1.9513
T = 40/252	Deep-out-of-the-money	0.89	97.1126	198.1214	228.7324	140.4928
		0.89	52.9349	129.5083	152.6143	86.4836
		0.90	98.3808	182.5218	208.0594	138.9450
		0.93	72.0110	113.4649	126.1089	96.9817
		0.95	21.8471	42.4646	48.6054	35.7837
	Out-of-the-money	0.98	17.2539	30.8073	34.7675	72.3618
		0.98	13.4745	26.2884	30.0114	23.0940
	At-the-money	1.00	8.7974	18.5782	20.9070	17.1773
		1.01	11.0570	20.0682	22.0163	19.1798
	In-the-money	1.03	2.7025	9.9956	11.3534	9.7122
		1.08	0.5307	4.7083	5.1001	5.5257
		1.11	2.7652	5.3974	5.5660	6.4849
	Deep-In-the-money	1.12	1.0646	3.1720	3.2741	4.3086
		1.17	2.4545	3.4474	3.4372	4.7063
		1.29	3.7298	3.8515	3.8311	4.9972
T = 59/252	Deep-out-of-the-money	0.86	22.6917	119.5118	144.6253	64.5129
		0.93	22.9524	53.6051	61.1890	46.6326
		0.93	23.9124	53.7180	61.0375	47.2443
	At-the-money	1.00	5.3325	16.2658	18.5636	16.1516
	In-the-money	1.03	3.1441	9.9069	11.3773	10.5606

Table 3.4: Comparison Between the MAPEs of the Proposed Method for Different Sample Sizes for Petrobras.

This table contains the prices for a European call option from EET-AM (empirical Esscher transform - assumed model) and the EET-B (bootstrap with replacement on historical returns) method for different moneyness, maturities and they are compared to the true market price of Petrobras data. The numbers reported for each combination are the mean absolute percentage error (MAPE). In the proposed method, we use 252 returns and the simulation is repeated 15,000 times and we repeat the experiment in EET-AM*, EET-B* with 50,000 returns, and the simulation is repeated 200 times.

Maturity	Moneyness (spot/strike)		EET-AM	EET-B	EET-AM*	EET-B*
T = 17/252	Deep-out-of-the-money	0.95	30.0697	51.6393	28.8035	52.4537
	Out-of-the-money	0.97	39.0752	56.8726	35.4213	57.5193
		0.98	16.1753	28.2949	12.6480	28.7008
	In-the-money	1.03	1.6857	7.5260	1.2814	7.6249
		1.07	0.8618	2.1482	0.3616	2.1653
	Deep-in-the-money	1.13	0.3616	0.3637	0.1626	0.3199
		1.17	1.4873	1.6451	1.5369	1.6473
		1.24	1.9716	1.9494	1.9543	1.9492
		1.31	0.9664	0.9686	0.9693	0.9687
		1.38	0.7244	0.7245	0.7247	0.7245
		1.61	1.3303	1.3303	1.3303	1.3303
T =40/252	Deep-out-of-the-money	0.88	97.0938	112.7758	118.6863	116.2486
		0.90	81.8346	90.5321	85.3186	92.6886
		0.94	56.2463	56.8326	48.1147	57.8006
		0.96	29.6578	32.5719	25.2225	33.1587
	Out-of-the-money	0.97	17.9689	23.0286	16.4807	23.4608
	At-the-money	1.02	8.0226	13.4987	9.2738	13.6603
	In-the-money	1.07	2.9488	6.5800	4.1976	6.6383
	Deep-in-the-money	1.13	2.2976	3.7873	2.7985	3.8054
		1.16	1.4679	2.2564	1.7425	2.2653
		1.24	1.7016	1.8856	1.7912	1.8884
		1.31	3.8063	3.8681	3.8544	3.8692
		1.37	3.4178	3.4406	3.4420	3.4408
T = 59/252	Deep-out-of-the-money	0.94	40.0601	47.5063	43.4797	48.4873
	At-the-money	1.02	19.8938	19.0197	15.9649	19.2947
	Deep-in-the-money	1.12	4.1748	3.1370	2.1643	3.1809
T = 121/252	Deep-out-of-the-money	0.96	0.6364	5.1689	1.2779	2.1500

Table 3.5: Comparison Between the MAPEs of the Proposed Method for Different Sample Sizes for Vale.

This table contains the prices for a European call option from EET-AM (empirical Esscher transform - assumed model) and the EET-B (bootstrap with replacement on historical returns) method for different moneyness, maturities and they are compared to the true market price of Vale data. The numbers reported for each combination are the mean absolute percentage error (MAPE). We use 252 returns and the simulation is repeated 15,000 times and we repeat the experiment in EET-AM*, EET-B* with 50,000 returns, and the simulation is repeated 200 times.

Maturity	Moneyness (Spot/strike)		EET-AM	EET-B	EET-AM*	EET-B*
T = 17/252	Deep-out-of-the-money	0.93	132.3809	177.7454	129.1951	180.9486
		0.95	102.9132	136.3859	92.3902	138.1864
	Out-of-the-money	0.98	42.8648	62.7852	36.1791	63.3947
	At-the-money	1.00	15.7799	29.4206	12.6889	29.6875
		1.00	18.2929	31.6931	15.4971	31.9417
		1.02	0.6636	9.1424	0.1448	9.2528
	In-the-money	1.05	1.5949	6.2268	1.3282	6.2675
		1.07	7.6055	11.3211	7.6805	11.3457
		1.11	0.7302	2.0628	1.0069	2.0680
		1.11	0.0997	1.0810	0.1570	1.0852
	Deep-In-the-money	1.14	1.4029	0.8775	1.2766	0.8763
		1.14	0.8132	1.2765	0.9220	1.2770
		1.18	5.3185	5.4862	5.3597	5.4856
		1.21	0.7264	0.7801	0.7432	0.7797
		1.36	1.8745	1.8742	1.8740	1.8742
		1.37	2.7025	2.7022	2.7021	2.7022
		1.47	1.9758	1.9758	1.9759	1.9758
T = 40/252	Deep-out-of-the-money	0.89	97.1126	198.1214	214.8283	203.1444
		0.89	52.9349	129.5083	141.1287	133.2440
		0.90	98.3808	182.5218	188.8369	186.5764
		0.93	72.0110	113.4649	104.8281	115.2330
		0.95	21.8471	42.4646	34.2552	43.2622
	Out-of-the-money	0.98	17.2539	30.8073	22.8009	31.2986
		0.98	13.4745	26.2884	18.5726	26.7483
	At-the-money	1.00	8.7974	18.5782	11.9412	18.8487
		1.01	11.0570	20.0682	13.7751	20.2916
	In-the-money	1.03	2.7025	9.9956	4.8394	10.1493
		1.08	0.5307	4.7083	2.1299	4.7574
		1.11	2.7652	5.3974	3.7281	5.4220
	Deep-In-the-money	1.12	1.0646	3.1720	1.8532	3.1884
		1.17	2.4545	3.4474	2.8787	3.4521
		1.29	3.7298	3.8515	3.7985	3.8518
T = 59/252	Deep-out-of-the-money	0.86	22.6917	119.5118	148.5000	125.0559
		0.93	22.9524	53.6051	50.7790	55.1348
		0.93	23.9124	53.7180	50.5609	55.1857
	At-the-money	1.00	5.3325	16.2658	12.1033	16.6808
	In-the-money	1.03	3.1441	9.9069	6.3625	10.1655

3.5
Conclusions

In this work, we propose a method to obtain European option prices under a GARCH framework with non-Gaussian innovations. We used a new class of models, the dynamic conditional score, proposed by Harvey (2013), for modeling the volatility (and heavy tails) of observed underlying asset prices. These models replace the observations, or their squares, by the score of the conditional distribution. They are more robust in extreme events, allowing the modeling of leverage effect, adding components of short and long-term volatility.

To avoid the formulation of a restrict model, the risk-neutralization is applied to the empirical distribution of the sample paths generated from the assumed model, as Liu et al (2015) and Duan (2002) have done. To identify a risk-neutral measure, we use the empirical Esscher transform. The sample paths are reweighted, giving rise to a risk-neutralized sample from which option prices can be obtained by a weighted sum of the options' pay-offs in each path.

We empirically compare our approach to competing benchmarks: Black-Scholes (1973) and Heston and Nandi (2000). In general, the proposed method, with an assumed model to describe the empirical distribution, presented the lowest MAPE. When we increase the size of the empirical distribution, only in lower maturities the MAPE was reduced.

Future research would benefit from conducting an extensive empirical study on the performance of our proposed pricing method, considering asset returns of different frequencies, multiple cross-sections of market option prices and long-dated options.

4
Breeden and Litzenberger Method to Uneven Spaced States

Chapter Abstract:

This work proposes a new approach for indirect estimation of the risk-neutral distribution. We develop a discrete version of the Breeden and Litzenberger (1978) theorem based on the use of a combination of options to synthetize Arrow-Debreu securities and obtain their prices from the prices of said options. These prices are called the 'risk-neutral probabilities mass function', RNPMF. Then we generalize this derivation for the case where states are not equally spaced. Finally, we consider that the risk-neutral distribution is obtained by an Empirical Esscher Transform, with 'flexible' Esscher parameter, in the same spirit of Shimko (1993).

Keywords: Risk-neutral probability, empirical Esscher transform, indirect estimation, state-price.

4.1
Introduction

The importance of option-implied information was seen in Bates (1991). He studied the behavior of S&P 500 future options' prices prior to the crash of October 1987, and found unusually negative skewness in the option-implied distribution leading to the conclusion that the crash was expected by the market. This fact led to the development of methods that seek more information to explain the behavior of assets, markets and investors. The assumption is based on financial assets which are continually updated and thus incorporate more recent information than other economic indicators. Several studies have emerged with varying objectives:

- Pricing of illiquid derivatives and exotic options – Pérignon and Villa (2002);
- Extract indicators of the level of uncertainty and future trends of the economy – Almeida, Ardison, Garcia, Vicente (2016), Kitsul and Wright (2013), Hui, Lo and Lau (2013), Birru and Figlewski (2012), Ornelas and Takami (2011), Bakshi, Panayotov, and Skoulakis (2011);
- Estimation of parameters of the stochastic process (the assumption is based on cross-section option prices containing forward-looking information beyond historical returns) –Christoffersen, Jacobs, and

Ornthanalai (2012), Santa-Clara and Yan (2010), Bams, Lehnert, and Wolff (2009), Figlewski (2010);

- Recover the asset price’s stochastic process (building an implied binomial tree) – Rubinstein (1994), Derman and Kani (1994), Jackwerth (1999);
- Estimate implied risk aversion (describes risk preferences of a representative agent in an economy) – Figlewski and Malik (2014), Bollerslev and Todorov (2011), Duan and Zang (2013), Bondarenko (2003, 2009), Jackwerth (2000), Aït-Sahalia and Lo (2000), Rosenberg and Engle (2002), Bliss and Panigirtzoglou (2002).
- Estimate empirical risk-neutral density – Almeida and Azevedo (2014), Monnier (2013), Grith, Härdle and Schienle (2012), Markose and Alentorn (2011), Cheng (2010), Figlewski (2010);
- Risk management (portfolio management and selection, implied betas of the capital asset pricing model) – Chang, Christoffersen and Jacobs (2013), De Miguel, Plyakha, Uppal and Vilkov (2013), Giamouridis and Skiadopoulos (2012), Kostakis, Panigirtzoglou and Skiadopoulos (2011), Chang, Christoffersen, Jacobs and Vainberg (2011), Buss and Vilkov (2012);
- Implied information indices – They are published by the Chicago Board of Options Exchange (CBOE) for S&P 500 firms: the implied volatility index (market indicator), the implied correlation index (risk management) and the implied skew index (interpreted as indicator of a possible market crash).

We can observe that most studies are concerned about the implied moments of risk-neutral distributions. According to Jackwerth (2004), these methods can be segregated into two general groups: parametric and nonparametric.[23]

Parametric methods can be sub-divided into three groups: expansion, generalized distribution and mixture methods. The risk-neutral distribution is obtained by criterion optimization which minimizes the sum of squared errors given by the difference between occurred and predicted values.

[23] The work of Christoffersen, Jacobs and Chang (2011) presents several techniques.

Nonparametric methods allow greater flexibility when fitting the risk-neutral distribution. Rather than requiring a parametric form of distribution, they allow general functions. These methods can be divided in three groups: kernel, maximum entropy and curve-fitting methods. Curve-fitting methods can be divided into other two sub-classes: fitting a function between implied volatilities and strike prices or a risk-neutral distribution being approximated by some general function.

The advantages of parametric methods, according to Bondarenko (2003), are analytical expressions and the possibility of extracting parameters to perform hedge. However, effectiveness depends on the correct modeling of the data generating process. The main advantage of nonparametric methods is that they do not require a specific format for the probability distribution, they are more flexible and adaptable to any function class. Although these techniques reduce the misspecification risk, they require larger sample sizes and are affected by irregularities such as data sparsity and problems to complete tails.

Jackwerth (2004) argues that both methods suffer with negative probabilities and integrability to one. Another important constrain is the number of options that are traded in the market. Moreover, these methods are specific for a given maturity, because statistical properties of observed prices cannot be used for other maturities (Duan, 2002).

The robustness of the option-implied risk-neutral distribution can be compared between several methods by perturbing actual option prices. Bliss et al (2001) derived risk-neutral distributions based on many different perturbed sets of prices. The result is an indicative of how much risk-neutral distributions can differ from each other and that the confidence intervals around the moments can be large. Santos and Guerra (2015), Jackwerth (2004) and Bliss et al (2001) cite that the easiest and most stable methods tend to be in the group of curve-fitting methods. However, Jackwerth (2004) says that a largely unresolved area of study is the development of statistical tests. Much of the current work lacks statistical rigor and is merely descriptive.

This work introduces a new approach to indirect estimation of implicit risk-neutral probability. We generalize the discrete version of the Breeden and Litzenberger (1978) method for the case where states are not equally spaced. We suppose that the risk-neutral probability mass function is given by Empirical

Esscher transform and it depends on the strike price. We use the historical distribution of the underlying asset price and the observed option prices to estimate the implicit Esscher parameter. Then, we fit a polynomial between the implied Esscher parameter and the strike price in the same spirit of Shimko (1993).

Indirect estimation combines the historical information of the underlying asset with the option-implied information. According to Christoffersen et al (2011), option prices contain useful information that are not easily extracted using econometric models, and combining historical information may be effective and provide new insights about how information and risk preferences are incorporated into prices on financial markets.

The works of Aït-Sahalia et al (2000), Jackwerth (2000) and Rosenberg et al (2000) explore the idea of indirect estimation to obtain empirical pricing kernel, the relationship between probabilities (risk-neutral and physical) and the implied risk aversion. Grith et al (2012) uses the indirect estimation to obtain empirical pricing kernel and the risk-neutral probabilities. In our case, we suppose that the empirical pricing kernel is known and given by an empirical version of the Esscher transform (1932). This assumption is reasonable, because it is well known in the information theory that a problem of maximum entropy has a solution in the form of the Esscher transform (Buchen and Kelly, Stutzer, 1996, Duan, 2002).

We ran simulation experiments under different situations which seek to highlight the differences and similarities between the methods. We compare our method to two approach alternatives: Double Lognormal proposed by Baha (1997) and the Shimko (1993) method.

The remainder of this work is organized as follows. In Section 4.2 we describe the theoretical framework of Breeden et al (1978). In section 4.3 we introduce the new methodology. The validity of this methodology is tested with numerical experiments in section 4.4. The conclusions and future researches are presented in the section 4.5.

4.2 Breeden and Litzenberger method

The time-state preference approach to general equilibrium in an economy, as developed by Arrow (1964) and Debreu (1959), is one of the most general frameworks available for the theory of finance under uncertainty. An Arrow-Debreu security is a security associated with a particular state of the economy which pays \$1 if that state occurs, and nothing otherwise. The price of an Arrow-Debreu security is referred to as state-price. According to Cox, Ross and Rubinstein (1979), the state-price associated with a particular state is simply the risk-neutral probability of that state discounted at the risk-free rate.

Breeden and Litzenberger (1978) implements the time-state preference model in a multiperiod economy, deriving the prices of Arrow-Debreu securities from prices of call options on aggregate consumption. Given the prices of Arrow-Debreu securities, the value of any cash flow is calculated, that is, these prices permit an equilibrium evaluation of assets with uncertain payoffs at many future dates.

Suppose that the value of the underlying asset in T periods has a discrete probability distribution with possible values of: $S_T = \$1.00, \$2.00, \dots, \$N$. Let $C(K,T) = max(S_T - K, 0)$, the vector of payoffs of a European call option at time T and strike price of K. For calls with strike prices of \$0.00, \$1.00, and \$2.00, its payoffs are as shown in table 4.1.

Table 4.1: Payoffs on Call Options with Equally Spaced States.

Underlying Asset	$C(0,T)$	$C(1,T)$	$C(2,T)$
$S_T = 1$	1	0	0
$S_T = 2$	2	1	0
$S_T = 3$	3	2	1
$(\dots)$	$\vdots$	$\vdots$	$\vdots$
$S_T = N$	N	$N-1$	$N-2$

Note that, as the strike price of a call option is increased from K to $K+1$, two changes occur in the payoff vector: (1) the payoff in the set of states with $S_T = K+1$ becomes zero, and (2) the payoffs in all states with $S_T \geq K+2$ are reduced by the change in the strike price. The difference between $C(K,T) -$

$C(K+1,T)$ gives a payoff of \$1.00 in every state with $S_T \geq K+1$, and $C(K+1,T) - C(K+2,T)$ gives a payoff of \$1.00 in every state for which $S_T \geq K+2$:

$$\begin{pmatrix}1\\2\\3\\\vdots\\N\end{pmatrix} - \begin{pmatrix}0\\1\\2\\\vdots\\N-1\end{pmatrix} = \begin{pmatrix}1\\1\\1\\\vdots\\1\end{pmatrix} \ and \ \begin{pmatrix}0\\1\\2\\\vdots\\N-1\end{pmatrix} - \begin{pmatrix}0\\0\\1\\\vdots\\N-2\end{pmatrix} = \begin{pmatrix}0\\1\\1\\\vdots\\1\end{pmatrix}.$$

A payoff of \$1.00 for $S_T = 1$ may be constructed as $[C(0,T) - C(1,T)] - [C(1,T) - C(2,T)]$, since this combination of calls would have a payoff vector of:

$$\begin{pmatrix}1\\1\\1\\\vdots\\1\end{pmatrix} - \begin{pmatrix}0\\1\\1\\\vdots\\1\end{pmatrix} = \begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix}.$$

Given the call prices, $C(K,T)$, prices of elementary claims, $\pi(S_T,T)$, must be computed from the replicating portfolio of calls that consists of one long call with $K = S_T - 1$, one long with $K = S_T + 1$ and two short calls with $K = S_T$.

In general, if the step size between potential values is ΔS_T, then $[C(K,T) - C(K + \Delta S_T, T)]$ has a payoff vector with zeros for $S_T \leq K$, and with ΔS_T for all levels greater than or equal to $K + \Delta S_T$. Therefore, the portfolio of call options that produces a payment of \$1.00 if the market is S_T, and zero otherwise, is:

$$\pi(S_T,T) = \frac{[C(\,S_T - \Delta S_T, T) - C(S_T,T)] - [C(S_T,T) - C(\,S_T + \Delta S_T, T)]}{\Delta S_T}. \qquad (4.1)$$

The payoff of portfolio $\pi(K,T)$ is the same as one Arrow-Debreu security that payoffs one unit of cash if, and only if, S_T is equal to K. Thus, it turns out that a complete set of options at all strike prices is equivalent to a complete set of Arrow-Debreu securities.

Consider $1/\Delta S_T$ shares of this portfolio:

$$\frac{1}{\Delta S_T}\pi(S_T,T) = \frac{[C(S_T-\Delta S_T,T)-C(S_T,T)]-[C(S_T,T)-C(S_T+\Delta S_T,T)]}{(\Delta S_T)^2}. \tag{4.2}$$

Assuming that the asset S_T has a continuous payoff and taking the limit of expression (4.2) as ΔS_T goes to zero, we have:

$$\lim_{\Delta S_T\to 0}\frac{\pi(S_T,T;\Delta S_T)}{\Delta S_T} = \frac{\partial^2 C(K,T)}{\partial K^2}\Big|_{K=S}. \tag{4.3}$$

Thus (4.2) gives the pricing function for an elementary claim on S_T maturing in T periods in the discrete case, and (4.3) gives the pricing function for the continuous case.

The result (4.3) can also be obtained from the relationship proposed by Cox and Ross (1976).[24] Under the risk-neutral distribution $q(S_T)$, the payoff is discounted at the deterministic risk-free rate r:

$$C(K,T) = e^{-rT}\int_{-\infty}^{\infty}(S_T-K)^+ q(S_T)d\,S_T \tag{4.4}$$

or,

$$C(K,T) = e^{-rT}\int_{-\infty}^{\infty}(S_T-K)^+ m(S_T)f(S_T)d\,S_T \tag{4.5}$$

where $f(S_T)$ is the physical distribution and $m(S_T)=q(S_T)/f(S_T)$ is the pricing kernel, characterizing the change of measure $f(S_T)$ to $q(S_T)$. Take the partial derivative of C with respect to K to get:

[24] See appendix 6.5.

$$\frac{\partial C(K,T)}{\partial K} = -e^{-rT}[1 - Q(K)] \tag{4.6}$$

which yields the cumulative distribution function denoted by $Q(.)$,

$$Q(K) = 1 + e^{rT}\frac{\partial C(T,K)}{\partial K} \tag{4.7}$$

or

$$Q(S_T) = 1 + e^{rT}\frac{\partial C(T,K)}{\partial K}|_{K=S_T}. \tag{4.8}$$

The probability distribution function $q(.)$ can be obtained by taking the derivative of (4.7) or (4.8) with respect to K:

$$q(K) = e^{rT}\frac{\partial^2 C(T,K)}{\partial K^2} \tag{4.9}$$

or

$$q(S_T) = e^{rT}\frac{\partial^2 C(T,K)}{\partial K^2}|_{K=S_T}. \tag{4.10}$$

According to Christoffersen et al (2011), an approximation to $Q(.)$ and $q(.)$ can be made using finite differences. Following the put-call parity, we can replace call prices by put prices (P) in the formulas above.

According to Aparicio and Hodges (1998), one important characteristic of the Breeden and Litzenberger (1978) approach is that no assumptions are made about the underlying asset price dynamics and also, market participants preferences are not restricted as they are reflected in the call option prices. Moreover, it is assumed that there are no restrictions on short sales, that there are no transaction costs or taxes, and that investors may borrow at the riskless rates of interest. Their work is the starting point of a line of research addressed to the recovery of relevant aspects of the underlying asset distribution from option market data.

4.2.1
No-arbitrage constraints

According to Carr (2001), there are three essential properties to obtain a well-defined risk-neutral distribution (RND). The RND is nonnegative:

$$q_{S_T}(S_T) \geq 0, \tag{4.11}$$

it integrates to one:

$$\int_0^\infty q_{S_T}(S_T) dS_T = 1, \tag{4.12}$$

and the RND reprices all calls (or martingale property):

$$\int_0^\infty max[S_T - K; 0] q_{S_T}(S_T) dS_T = e^{r(T-t)} C(K,T), \qquad K \geq 0. \tag{4.13}$$

The nonnegativity (4.11) and integrability (4.12) properties ensure that the risk-neutral distribution is a probability distribution. The martingale property (4.13) ensures that the means of distribution is the forward price, which includes the special case $K = 0$. That is, this option is guaranteed to be in-the-money and at maturity we exercise the option and buy the underlying asset.

According to Brunner and Hafner (2003), these three properties, (4.11), (4.12) and (4.13), can be formulated in terms of a call option (or implied volatilities). A set of equivalent conditions is, at first,[25]

$$S_0 \geq C(K,T) \geq max\left[S_0 - e^{-r(T-t)} K; 0\right] \tag{4.14}$$

that the value of a call option never be greater than the asset price and never less than its intrinsic value. Second,

[25] See appendix 6.6.

$$-e^{-rT} \leq \frac{\partial C(K,T)}{\partial K} \leq 0 \tag{4.15}$$

that the call price function monotonically decrease. Finally,

$$\frac{\partial^2 C(K,T)}{\partial K^2} \geq 0, \qquad K \geq 0. \tag{4.16}$$

that the call option price function be convex.

Äit-Sahalia and Duarte (2003) proposed a method of option-implied density estimation based on locally polynomial regressions that incorporate shape restriction. The theory-imposed restrictions are that the price of a call option must be a decreasing (4.15) and convex function of the option's strike price (4.16).

4.3 Proposed method

Consider the notation: N is the number of possible states for the underlying at maturity $t = T$; S_i are the possible prices at time $t = T$; p_i are the physical probabilities of price i; q_i are the risk-neutral probabilities of state i; O_i is the fair value for derivative i; P_0 is the price of underlying at $t = 0$ (present day), where $i = 1, \dots, N$, and K is the strike price.

4.3.1 Breeden and Litzenberger Method with Uneven Spaced States

The price of a call option under a discrete risk-neutral distribution can be expressed as:

$$C(P_0, K) = e^{-rT} \sum_{i, S_i > K}^{N} (S_i - K) q_i \tag{4.17}$$

with $K \in \{S_i\}$. Now, consider that options O_i, for every possible strike prices S_i, are available. Possible payoffs for option O_i are one of the $N + 1$ values $\max[0, S_i - K]$, $i = 0, \dots, N$, according to the state at $t = T$.

Table 4.2: Payoffs on Call Options with Uneven Spaced States.

	0	1	2	…	N-1	N
0_0	0	$S_1 - S_0$	$S_2 - S_0$		$S_{N-1} - S_0$	$S_N - S_0$
0_1	0	0	$S_2 - S_1$		$S_{N-1} - S_1$	$S_N - S_1$
0_2	0	0	0		$S_{N-1} - S_2$	$S_N - S_2$
0_3	0	0	0		$S_{N-1} - S_3$	$S_N - S_3$
…	…	…	…		…	…
0_{N-1}	0	0	0		0	$S_N - S_{N-1}$
0_N	0	0	0		0	0

So, one can assemble a portfolio formed by one long position on option O_i and a short position on O_{i+1}. The possible payoffs are then given by a sequence of $i+1$ zeros followed by $N-1$ of difference between the payoffs, given by S_i:

Table 4.3: Portfolios of Call Options with Uneven Spaced States.

	0	1	2	…	N-1	N
$0_0 - 0_1$	0	$S_1 - S_0$	$S_1 - S_0$		$S_1 - S_0$	$S_1 - S_0$
$0_1 - 0_2$	0	0	$S_2 - S_1$		$S_2 - S_1$	$S_2 - S_1$
$0_2 - 0_3$	0	0	0		$S_3 - S_2$	$S_3 - S_2$
$0_3 - 0_4$	0	0	0		$S_4 - S_3$	$S_4 - S_3$
……	…	…	…		…	…
$0_{N-2} - 0_{N-1}$	0	0	0		$S_{N-1} - S_{N-2}$	$S_{N-1} - S_{N-2}$
$0_{N-1} - 0_N$	0	0	0		0	$S_N - S_{N-1}$

Consider now that we take $\frac{1}{(S_i - S_{i-1})}, i = 1, \dots, N$, shares of these portfolios:

Table 4.4: Arrow-Debreu Securities.

	0	1	2	…	N-1	N
$(0_0 - 0_1)/(S_1 - S_0)$	0	1	1		1	1
$(0_1 - 0_2)/(S_2 - S_1)$	0	0	1		1	1
……	…	…	…		1	1
$(0_{N-2} - 0_{N-1})/(S_{N-1} - S_{N-2})$	0	0	0		1	1
$(0_{N-1} - 0_N)/(S_N - S_{N-1})$	0	0	0		0	1

Now, it is easy to see that the portfolio is:

$$\frac{1}{(S_1 - S_0)}[O_0 - O_1] - \frac{1}{(S_2 - S_1)}[O_1 - O_2] \qquad (4.18)$$

and pays 1 if S_1 occurs. Then, the portfolio:

$$\frac{1}{(S_N - S_{N-1})}\left[O_{N-1} - \underbrace{O_N}_{=0}\right] \tag{4.19}$$

and pays 1 if S_N occurs. In general:

$$\frac{1}{(S_i - S_{i-1})}[O_{i-1} - O_i] - \frac{1}{(S_{i+1} - S_i)}[O_i - O_{i+1}], i = 1, \dots, N \tag{4.20}$$

is an Arrow-Debreu security giving a payoff of 1 when, taking $S_{N+1} = \infty$, final state is i. Note that i represents the strike price around which is calculated the second difference above. Also, remember that the price of an Arrow-Debreu security is the 'risk-neutral probability' multiplied by e^{-rT} (see equation 4.10). The price of this portfolio is known, since options O_i are available in the market.

Using the definition (4.17) in (4.20):

$$\begin{aligned} &\frac{1}{(S_i - S_{i-1})}[C(P_0, S_{i-1}) - C(P_0, S_i)] \\ &\qquad - \frac{1}{(S_{i+1} - S_i)}[C(P_0, S_i) - C(P_0, S_{i+1})], \end{aligned} \tag{4.21}$$

and we have:

$$\begin{aligned} &C(P_0, S_{i-1}) - C(P_0, S_i) \\ &\qquad = e^{-rT}\sum_{j=i}^{N}(S_j - S_{i-1})q_j - e^{-rT}\sum_{j=i+1}^{N}(S_j - S_i)q_j. \end{aligned} \tag{4.22}$$

The solution of (4.21) is:

$$C(P_0, S_{i-1}) - C(P_0, S_i) = e^{-rT}(S_i - S_{i-1})\sum_{j=i}^{N} q_j \tag{4.23}$$

and we also have:

$$C(P_0, S_i) - C(P_0, S_{i+1}) = e^{-rT}(S_{i+1} - S_i) \sum_{j=i+1}^{N} q_j. \quad (4.24)$$

Replacing (4.23) and (4.24) in (4.21):

$$e^{-rT} \sum_{j,S_j > S_{i-1}}^{N} q_j - e^{-rT} \sum_{j,S_j > S_i}^{N} q_j = e^{-rT} q_i. \quad (4.25)$$

We have proven that the second 'modified' difference gives the risk-neutral probability mass function.

4.3.2
Breeden and Litzenberger Method with uneven spaced states and q_i depending on K

Consider now that the probability mass function q depends on K. The price of a call option would then be written as:

$$C(P_0, K) = e^{-rT} \sum_{i,S_i > K}^{N} (S_i - K) G(S_i, K),\ K \in \{S_i,\ i = 1, N\} \quad (4.26)$$

where:

$$q = G(S_i, K) = \frac{e^{h(K)\ln\frac{S_j}{P_0}}}{\sum_{k=1}^{N} e^{h(K)\ln\frac{S_k}{P_0}} p_m} p_i$$

is the empirical Esscher transform. In our case, $p = p_m = 1/N$ and can be cancelled:

$$G(S_j, K) = \frac{e^{h(K)\ln\frac{S_j}{P_0}}}{\sum_{k=1}^{N} e^{h(K)\ln\frac{S_k}{P_0}}}. \quad (4.27)$$

As in (4.21):

$$C(P_0, S_{i-1}) - C(P_0, S_i) = e^{-rT} \sum_{j=i}^{N} (S_j - S_{i-1}) G(S_j, S_{i-1}) - e^{-rT} \sum_{j=i+1}^{N} (S_j - S_i) G(S_j, S_i) \tag{4.28}$$

and according to the precedent section in equation (4.22):

$$q_i = e^{rT} \left\{ \frac{1}{(S_i - S_{i-1})} [C(P_0, S_{i-1}) - C(P_0, S_i)] - \frac{1}{(S_{i+1} - S_i)} [C(P_0, S_i) - C(P_0, S_{i+1})] \right\} \tag{4.29}$$

$$q_i = \left\{ \frac{1}{(S_i - S_{i-1})} \left[\sum_{j=i}^{N} (S_j - S_{i-1}) G(S_j, S_{i-1}) - \sum_{j=i+1}^{N} (S_j - S_i) G(S_j, S_i) \right] - \frac{1}{(S_{i+1} - S_i)} \left[\sum_{j=i+1}^{N} (S_j - S_i) G(S_j, S_i) - \sum_{j=i+2}^{N} (S_j - S_{i+1}) G(S_j, S_{i+1}) \right] \right\} \tag{4.30}$$

4.3.3 No-arbitrage constraints for the proposed method

Nonnegativity property

This property depends on two factors in the formula (4.30): the difference between prices, $(S_i - S_{i-1})$ and $(S_{i+1} - S_i)$, and the impact that $h(K)$ causes in the terms $G(.)$ in the second bracket. Assuming that the difference in the second ratio, $1/(S_{i+1} - S_i)$, occurs in the tenth decimal, i.e, it is much smaller than the difference in the first ratio, $1/(S_i - S_{i-1})$. Suppose that the amount generated in the first bracket is approximately equal to the value generated in the second bracket. Thus, the nonnegativity property is violated due to the quantity generated by the second ratio. But if we fit, in $h(K)$, a function that reduces the importance of the second ratio, that is, the difference in the second bracket becomes the smallest, then it does not violate the property of nonnegativity.

Integrability property

Suppose that the scenarios S_i are equally spaced, $i = 1, \ldots, N$, we have:

$$\sum_{i=1}^{N} q_i = \sum_{i=1}^{N} \left[\sum_{j=i}^{N} (j-i+1) G(j, i-1) - 2 \sum_{j=i+1}^{N} (j-i) G(j,i) + \sum_{j=i+2}^{N} (j-i-1) G(j, i+1) \right]$$

then,

$$\sum_{i=1}^{N} q_i = \sum_{i=1}^{N} \sum_{j=i}^{N} (j-i+1) G(j, i-1) - 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (j-i) G(j,i) + \sum_{i=1}^{N-2} \sum_{j=i+2}^{N} (j-i-1) G(j, i+1)$$

$$\sum_{i=1}^{N} q_i = \sum_{i=0}^{N-1} \sum_{j=i+1}^{N} (j-i) G(j,i) - 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (j-i) G(j,i) + \sum_{i=2}^{N-1} \sum_{j=i+1}^{N} (j-i) G(j,i)$$

$$\sum_{i=1}^{N} q_i = \overbrace{\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (j-i) G(j,i) - 2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (j-i) G(j,i) + \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (j-i) G(j,i)}^{=0} + \sum_{j=1}^{N} j G(j,0) - \sum_{j=2}^{N} j G(j,1) + \sum_{j=2}^{N} G(j,1).$$

It doesn’t sum to one. But if $G(j,0) = G(j,1)$, it sums to one. Note that the formula is imperfect because $h(K)$ is not perfect.

According to Casella and Berger (2010), from a purely mathematical viewpoint, any nonnegative function with a finite positive integral (or sum) can be turned into a probability distribution function (or probability mass function). For example, if $w(x)$ is any nonnegative function that is positive on a set A, 0 elsewhere, and:

$$\int_{\{x \in A\}}^{0} w(x)dx = W < \infty$$

for some constant $W > 0$, then the function $f_X(x) = w(x)/W$ is a probability distribution function of a random variable X taking values in A. Most methods use this standardization for risk-neutral distribution sum to one.

Martingale property

Note that call option prices vary greatly across strike prices. That is, deep-in-the-money (DITM) are valued as high as the underlying asset itself, whereas deep-out-of-the-money (DOTM) calls are valued close to zero. According to Jackwerth (2004), the particular exercise of fitting the risk-neutral distribution is rarely undertaken because some of the requirements lead to numerical difficulties. Thus, Jackwerth (2004) concludes that the choice of option prices, for each method, can be somewhat of an art.

In our case, we include the special case by calculating the h-implied for strike equal to zero. Moreover, as noted in the literature, we obtain the martingale condition depending on a set of options used for each maturity. For example, for short periods, we can use two types of combinations: DITM, in-the-money (ITM) and at-the-money (ATM) or ATM and out-of-the-money (OTM). For long periods, just ATM, OTM and DOTM. Obviously, these combinations can yield a range of risk-neutral distributions.[26]

4.4 Methodology

The algorithm for our method:

1. Simulate the physical distribution for S_T;
2. Compute the h-implied for each observed option (including $K = 0$);

$$h_j = \arg_{h \in R} \left\{ C(K_j) = e^{-rT} \sum_{i,\, S_i > K}^{N} (S_{T,i} - K_j) G(S_{T,i}, S, h) \right\},\ j = 1, \dots, n$$

3. Fit the function $h(K)$;
4. Compute $G(.)$ for $j = 1, .., N$ and $i = 1, \dots, N$ with the equation (4.26);
5. Compute q_j for $j = 1, \dots, N$ with the equation (4.30).

In step one, our analysis is performed based on an unknown data generator process for the underlying asset. First, we simulate 5,000 prices, for each maturity, using the data generator process of the Heston (1993) model. After, we performed 5,000 bootstraps with replacement on historical returns to simulate an unknown stochastic process for the underlying asset.[27] In step two, we consider the market price given by the Heston (1993) model. We substitute these prices in step two's equation to calculate the h-implied. In step three, we fit a polynomial of degree two in $h(K)$.[28] We calculate $G(.)$ in step four, and in step five, q_j.

The parameter values of the Heston (1993) model follow Almeida and Azevedo (2014). They use the objective and risk-neutral parameters estimated in Garcia, Lewis, Pastorello and Renault (2011) adopting S&P 500 daily data from January 1996 to December 2005.

---

[26] Some researchers prefer robustness to precision. That is, they provide a range of risk-neutral distribution and impose restrictions on characteristics of admissible risk-neutral distributions to reduce no-arbitrage bounds. For example, Jackwerth (2004) suggests that the call price functions themselves need to be fairly smooth.

[27] It is not possible to write options that distinguish between two states if the underlying assets pay identical returns in those states. Hence, after we had performed bootstraps with replacement on historical returns, we used only 4,500 scenarios.

[28] Any other function could be used. However, we use a polynomial of degree 2 to compare with Shimko's method. He fits a polynomial with the same order between the strike price and the implied volatility.

Garcia et al (2011) does not report the estimated drift and the risk-free rate. Almeida et al (2014) imputes arbitrarily the drift to 10% and the risk-free rate to the year 2000 average of the 1-year T-Bill, $r = 5.93\%$. The same value for drifts has also been used in Gray and Newman (2005) and Harley and Walker (2010). According to latter authors, these values represent typical estimates from market data.

We consider four scenarios for pricing in table 4.4. We introduce three scenarios (scenario 1, 2 and 3) beyond the estimated in Garcia (2011) et al (scenario 4). That is, we change ρ to a positive value (scenarios 1 and 2) and we reduce the value of σ in 50% (scenario 1 and 3). The initial value of the volatility was equal to its long-term average.

Table 4.5: Parameter Values.

Scenary	κ	θ	σ	ρ
1	6.48	0.0203	0.2232	0.215
2	6.48	0.0203	0.4464	0.215
3	6.48	0.0203	0.2232	-0.215
4	6.48	0.0203	0.4464	-0.215

Finally, we compare our method to two approach alternatives: mixture method (following Ait-Sahalia and Duarte, 2003, Baha, 1997, Ornelas et al, 2011, Santos et al, 2015) and the Shimko (1993) method. The prices of the mixture method, or mixture of double lognormal (DLN), and the prices of the Shimko (1993) method (SHM) are calculated using expression (4.4). The strike prices were 70, 80, 90, 92.5, 95, 97.5, 100, 102.5, 105, 107.5, 110 and 120. The spot price was 100 and we use three maturities, 30, 90 and 180.

Lognormal Mixture

The mixture of lognormals was proposed by Bahra (1997) and Melick and Thomas (1997). Instead of specifying a dynamic for the underlying asset price, it is possible to make assumptions about the functional form of the risk-neutral distribution and then obtain the parameters of the distribution by minimizing the distance between observed option prices and theoretical prices. Consider the weighted sum of lognormal distribution functions:

$$q(S_T) = \sum_{i=1}^{k} [\theta_i L(\mu_i, \sigma_i; S_T)] \tag{4.31}$$

$$L(\mu_i, \sigma_i; S_T) = \frac{1}{S_T \sigma \sqrt{2\pi}} e^{\left[-\frac{(lnS_T - \mu)^2}{2\sigma^2}\right]} \tag{4.32}$$

$$\mu_i = ln(S_t) + \left(\mu_i - \frac{1}{2}\sigma_i^2\right)(T - t) \tag{4.33}$$

$$\sigma_i = \sigma_i \sqrt{(T - t)} \tag{4.34}$$

where $L(\mu_i, \sigma_i; S_T)$ is the *ith* lognormal distribution with parameters μ_i and σ_i. The term k defines the number of mixtures describing the risk-neutral distribution. In order to guarantee that $q(S_T)$ is a probability distribution, $\theta_i \geq 0$ for $i = 1, \dots, k$., and $\sum_i^k \theta_i = 1$. Replacing (4.48) in (4.4), we have the theoretical prices of the European call option:

$$\hat{c}(K,T) = e^{-rT} \sum_{i=1}^{k} \theta_i \int_X^{\infty} (S_T - K) L(\mu_i, \sigma_i, S_T) dS_t. \tag{4.35}$$

Applying the mixture of two lognormals used by Bahra (1997), named double lognormal (DLN), we get the following formula for theoretical prices of European call options:

$$\begin{aligned} \hat{c}(K,T) = e^{-rT} \Big\{ \theta \left[e^{\alpha_1 + \frac{1}{2}\beta_1^2} N(d_1) - XN(d_2) \right] \\ + (1 - \theta) \left[e^{\alpha_2 + \frac{1}{2}\beta_2^2} N(d_3) - XN(d_4) \right] \Big\} \end{aligned} \tag{4.36}$$

where $N(.)$ accumulated normal and:

$$d_1 = \frac{-ln(K) + \mu_1 + \sigma_1^2}{\sigma_1} \tag{4.37}$$

$$d_2 = d_1 - \sigma_1 \tag{4.38}$$

$$d_3 = \frac{-ln(K) + \mu_2 + \sigma_2^2}{\sigma_2} \tag{4.39}$$

$$d_4 = d_3 - \sigma_2. \tag{4.40}$$

For theoretical prices of European put options:

$$\hat{p}(X_i, T) = e^{-rT}\left\{\theta\left[-e^{\alpha_1+\frac{1}{2}\beta_1^2}N(-d_1) - KN(-d_2)\right] + (1-\theta)\left[-e^{\alpha_2+\frac{1}{2}\beta_2^2}N(-d_3) - KN(-d_4)\right]\right\}. \tag{4.41}$$

In order to find the parameters of the implied risk-neutral distribution we have to solve the minimization problem:

$$\min_{\mu_i\sigma_{i,},\theta} \sum_{i=1}^{n}[c-\hat{c}]^2 + \sum_{i=1}^{n}[p-\hat{p}]^2 + \left[\theta e^{\alpha_1+\frac{1}{2}\beta_1^2} + (1-\theta)e^{\alpha_2+\frac{1}{2}\beta_2^2} - e^{rT}S_0\right] \tag{4.42}$$

where the first and second terms refer to the sum of the squared deviation between theoretical prices and the observed market prices (c and p). The third term of the equation states that the expected value of the risk-neutral distribution must be equal to the underlying asset's forward price ($F = S_0e^{rT}$).

Shimko Method

The Shimko method (1993) assumes that the volatility smile is a function of strike price:

$$\sigma(K) = a_0 + a_1K + a_2K^2. \tag{4.43}$$

The annual implied volatility function $v(K)$ is given by:

$$v(K) = \sigma(K)\sqrt{\tau}. \tag{4.44}$$

Shimko (1993) applies Breeden et al (1978) in the Black-Scholes (1973) model. Taking the first and second derivatives with respect to K, we obtain:

$$v' = (a_1 + 2a_2K)\sqrt{\tau} \tag{4.45}$$

$$v'' = 2a_2\sqrt{\tau}. \tag{4.46}$$

The probability distribution function is calculated by:

$$q(K_t) = Fn(d_1)[v'' - v'd_1d_{1K}] - n(d_2)d_{2K} \tag{4.47}$$

where F is the forward price and $n(d_i)$ is the Gaussian normal distribution function:

$$d_1 = \frac{ln(F/K) + \frac{1}{2}\sigma^2(K)\tau}{\sigma(K)\sqrt{\tau}} \tag{4.48}$$

$$d_2 = d_1 - \sigma^2(K)\sqrt{\tau} \tag{4.49}$$

$$d_{1K} = \left(1 - \frac{d_1}{v}\right)v' - \frac{1}{Kv} \tag{4.50}$$

$$d_{2K} = -\left(\frac{1}{K} + d_1v'\right)\frac{1}{v} \tag{4.51}$$

$$n(d_i) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}d_i^2}, \qquad i = 1,2. \tag{4.52}$$

In the original formulation of Shimko (1993), the volatilities are interpolated by the range containing the observed prices. Probabilities in the sections located above and below the range of interpolated values are estimated by lognormal distributions. In this work, we keep a constant volatility for these points, as Campa, Chang and Reider (1998) and Malz (1997).

4.5
Results

The following Figures illustrate the h-implied behavior for the different proposed scenarios. In Figures (4.1) to (4.6), we obtain the h-implied based on the data generator process of Heston (1993) to analyze several effects about the empirical parameter. In tables 4.6 to 4.9, we use bootstrap with replacement under historical prices of the data generator process by Heston (1993).

Effect of the sign

Figure 4.1 presents the effects of changing the h-implied using the same distribution. We can observe that an increase in the h-implied changes the risk-neutral distribution to the right. That is, the growth of the h-implied heightens the probabilities of the right side of the distribution and reduce the probabilities of the tail side. When we reduce the h-implied, the opposite happens.

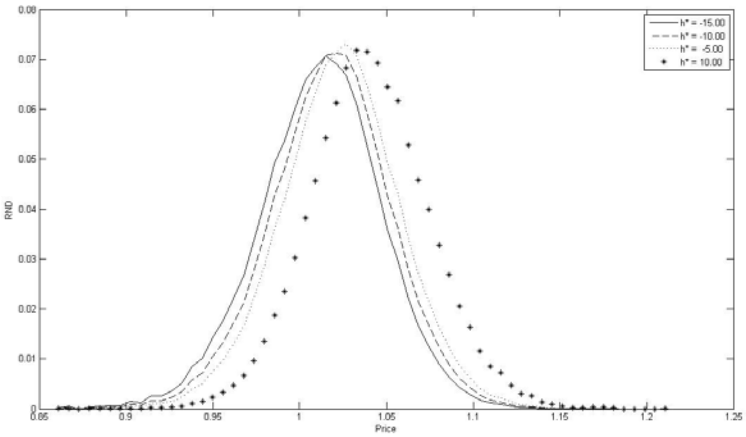


Figure 4.1: Effect of the sign

Effect of the asymmetry and kurtosis

In Figure 4.2, we analyze the impact of positive skewness on the h-implied. The correlation parameter ρ controls the skewness return's distribution. When $\rho > 0$, the probability distributions will be positively skewed. This has the effect of fattening the right tail of distribution, and thinning the left tail. In our case, to reproduce the market prices with greater strikes, the value of the h-implied grows and in consequence increases the probabilities of the underlying asset's price in the right tail.

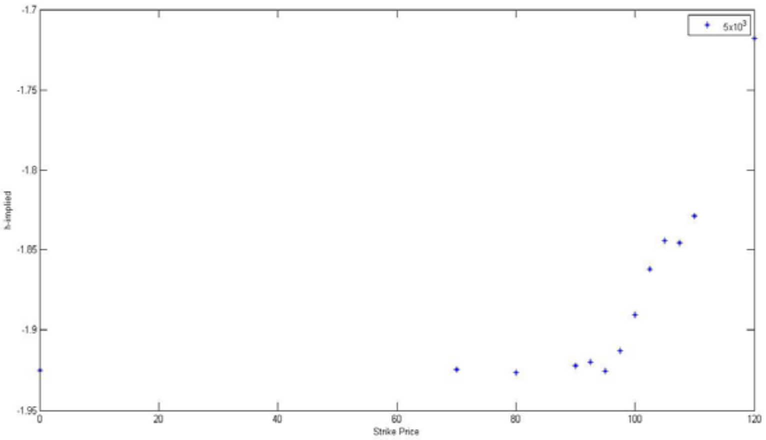


Figure 4.2: Effect of the asymmetry – $T = 90, \rho = 0.215, \sigma = 0.2232$

In Figure 4.3, we analyze the impact of negative skewness on the h-implied. When the correlation is negative, the distribution is negatively skewed and more weight goes to the left tail, and less to the right tail. In our case, to reproduce the market prices with bigger strikes, the value of the h-implied reduces and in doing so, it decreases the probabilities of the underlying asset's price in the right tail.

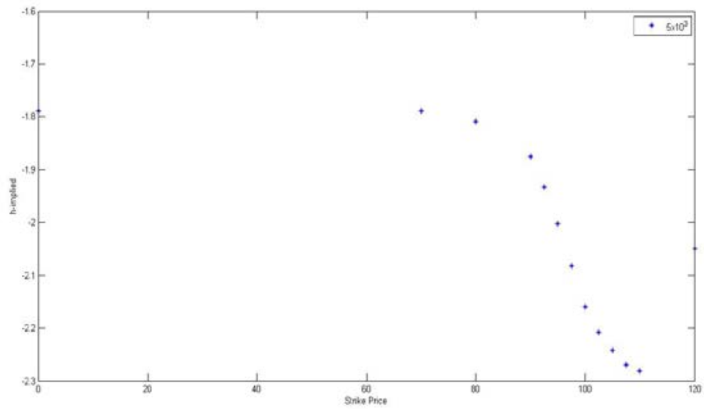


Figure 4.3: Effect of the asymmetry – $T = 90, \rho = -0.215, \sigma = 0.2232$.

In Figures 4.4 and 4.5, we consider the impact of increasing the volatility. The volatility of the variance parameter controls the kurtosis. When σ is high, the variance process is highly dispersed and the distribution of returns has higher kurtosis and fatter tails than when σ is small. Thus, a high variance volatility will increase the range of terminal stock price values. At maturities of 90 and 180 days, the increase of sigma changes the h-implied along the strike price according to the sign of ρ. At a maturity of 30 days, the increase of sigma only makes the value of the h-implied grow.

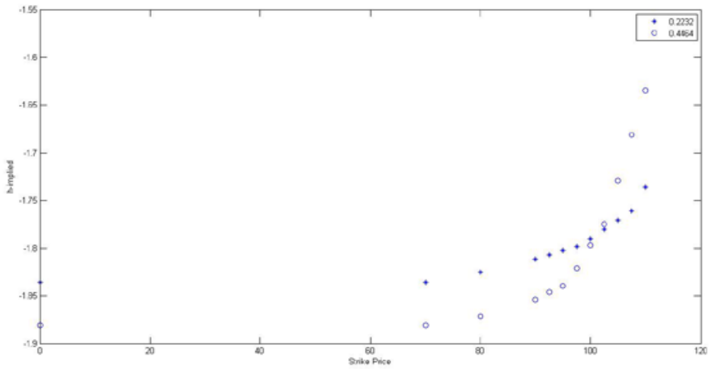


Figure 4.4: Effect of the kurtosis – $T = 180, \rho = 0.215, \sigma_1 = 0.2232, \sigma_2 = 0.4464$.

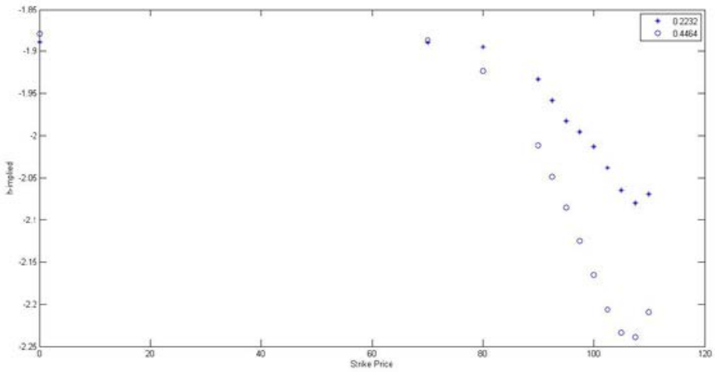


Figure 4.5: Effect of the kurtosis – $T = 180, \rho = -0.215, \sigma_1 = 0.2232, \sigma_2 = 0.4464$.

We can conclude that the behavior of the h-implied depends on the deformation required in the distribution to reproduce the market price.

Effect of the sample size

The sample size affects only the intensity of the h-implied value in the tails' distribution. That is, if we have more mass in the tails, the h-implied value is small. In the table 4.5, we calculate the probability of exercising the option, for a given strike price, and we analyze the impact that the sample size (5,000 scenarios) has on the value of h-implied. In the maturity $T = 30$ and $K = 120$, the probability of exercising the option is very small and the value of h-implied increases (see Figure 4.6). Therefore, to minimize the effect of the sample size on the h-implied, we can discard option (D)OTM in the short term.

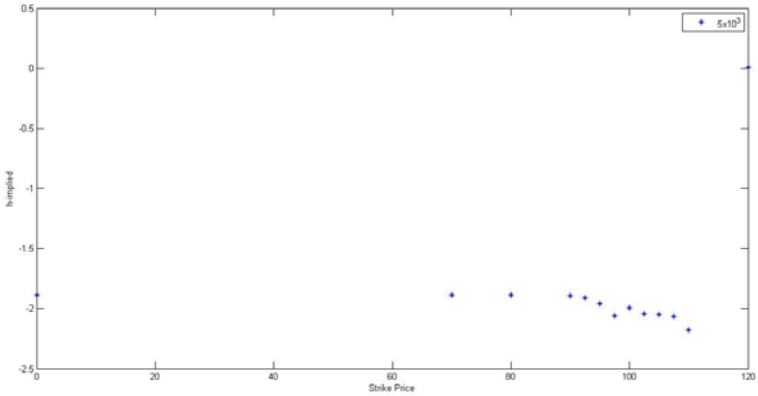


Figure 4.6: Effect of the sample size – $T = 30, \rho = -0.215, \sigma = 0.2232$.

Table 4.6: Probability of Exercising the Option.

$\boldsymbol{\rho = -0.215 - \sigma = 0.2232}$

I	0.00	0.00	0.00	2.50	5.00	7.50	00.00	02.50	05.00	07.50	10.00	20.00
$\Delta_{T=30}$	.0000	.0000	.9859	.9560	.8865	.7587	.5760	.3744	.2040	.0926	.0353	.0002
$\Delta_{T=90}$	.9999	.9963	.9345	.8875	.8202	.7318	.6259	.5102	.3947	.2891	.2005	.0282
$\Delta_{T=180}$	.9990	.9850	.9014	.8586	.8051	.7417	.6697	.5919	.5114	.4315	.3556	.1299

$\boldsymbol{\rho = -0.215 - \sigma = 0.4464}$

$\Delta_{T=30}$	.9999	.9994	.9822	.9537	.8914	.7715	.5885	.3726	.1933	.0852	.0323	.0008
$\Delta_{T=90}$	.9996	.9937	.9355	.8938	.8324	.7478	.6407	.5182	.3930	.2795	.1876	.0266
$\Delta_{T=180}$	.9977	.9821	.9056	.8664	.8162	.7547	.6829	.6029	.5182	.4332	.3521	.1201

$\boldsymbol{\rho = 0.215 - \sigma = 0.2232}$

$\Delta_{T=30}$	.0000	.0000	.9907	.9629	.8896	.7506	.5585	.3601	.2014	.0989	.0434	.0007
$\Delta_{T=90}$	.0000	.9985	.9413	.8904	.8159	.7191	.6069	.4895	.3776	.2792	.1986	.0377
$\Delta_{T=180}$	.9997	.9899	.9045	.8576	.7990	.7303	.6540	.5736	.4927	.4148	.3426	.1351

$\boldsymbol{\rho = 0.215 - \sigma = 0.4464}$

$\Delta_{T=30}$	.9997	.0000	.9900	.9660	.8977	.7587	.5528	.3453	.1912	.0964	.0470	.0021
$\Delta_{T=90}$	.0000	.9977	.9469	.9001	.8268	.7258	.6048	.4783	.3612	.2630	.1865	.0418
$\Delta_{T=180}$	.9994	.9899	.9117	.8658	.8064	.7346	.6537	.5680	.4826	.4017	.3286	.1304

Comparison with other methods

In Tables 4.7 and 4.8, we present the biases between the prices of the methods and the market price for different scenarios. In the tables 4.9 and 4.10, we present the Mean Absolute Percentage Error (MAPE) of the methods. Finally, in the Figures 4.8 – 4.13, we present the volatility smiles.

It is shown in tables 4.7 and 4.8, under assumptions of the scenarios 1 and 2 of the table 4.4, in general, that the biases of EET are small, except for $K = 120$, at time $T = 30$. Biases tend to increase when volatility increases, independent of the method. There is a change signal for all methods, indicating that the recovered risk-neutral density provides prices above or below the market.

From the ratio between the underlying asset price and the strike price, we can classify the options' prices by moneyness. For example, we separate strike prices in ITM ($K \leq 92.50$), ATM ($95.00 \leq K \leq 105.00$) and OTM ($K \geq 107.50$) and we calculate the MAPE by Moneyness (MM). The general MAPE (GM) is calculated for all strike prices.

In the table 4.8, under assumptions of the scenarios 1 and 2 of the table 4.4, we get the smaller MM in the scenario 1. For the GM, the smaller errors were obtained by the Shimko method. Our method outperforms the DLN method in maturities T=90 and T=180. When we increase the volatility, the lowest GM are among the proposed method and Shimko's method.

Table 4.9, under assumptions of the scenarios 3 and 4 of the table 4.4, shows that the Shimko method outperforms the other methods with the lowest MM and GM. In scenario 3, the proposed method outperforms the DLN method. When we increase the volatility, the lowest MM and GM are among the Shimko method and the DLN method.

Although percentage bias allows detecting differences in call prices, it gives no indication of the relative difference between them. Implied volatilities are important because they are embedded in option prices, and the mentioned prices reflect future expectations of market participants. This means that implied volatilities constitute a forward-looking estimate of the underlying asset's volatility.

Market returns and prices are more skewed, and show greater kurtosis, than the normal distributions allow. Both of these distributional features are thought to explain smiles and smirks. If returns were normal, then implied volatility would be constant across moneyness and maturity according to the Black-Scholes (1973) model. Smiles can occur, however, because returns show greater kurtosis than stipulated under normality, so extreme returns are more likely. This implies that DITM and DOTM options are more expensive related to the Black-Scholes' price. Smirks can occur because returns often show asymmetry, which again the normal distribution does not allow. For example, large negative returns are more likely, leading to implied volatilities for ITM calls that are higher than implied volatilities for OTM calls.

In general, smiles and smirks are more pronounced for short-term options and less pronounced for long-term options. This is synonymous with long-term returns being closer to being normally distributed than short-terms returns. Moneyness and maturity are the two factors thought to influence the most the shape of smiles and smirks.

When $\rho > 0$, the volatility smile is positively sloped and when $\rho < 0$, the volatility smile is negatively sloped. The increase of sigma also causes an increase of the volatility smile's curvature. In this work, we can see that the proposed method follows market expectations and reproduces the volatility smile. In some situations, the smile of the proposed method alternates above and below the market volatility smile.

Table 4.7: Option Pricing under scenarios 1 and 2.

Percentage bias for different strike prices (*K*) and maturities (*T*) from EET (empirical Esscher transform), DLN (double lognormal) and SHM (Shimko method), under assumptions of the scenarios 1 and 2 in the table 4.4. Underlying asset current price = $ 100.00; risk-free rate = 5.93%.

		Scenario 1: $\rho = 0.215, \sigma = 0.2232$			Scenario 2: $\rho = 0.215, \sigma = 0.4464$		
T	**K**	**Bias%-EET**	**Bias%-DLN**	**Bias%-SHM**	**Bias%-EET**	**Bias%-DLN**	**Bias%-SHM**
30	70.0000	0.0031	-0.0313	-0.0031	0.0064	-0.0618	0.0423
	80.0000	0.0045	-0.0464	-0.0040	0.0088	-0.0924	0.0639
	90.0000	-0.0307	-0.0711	0.0874	-0.0640	-0.2064	0.2440
	92.5000	0.0010	0.0099	0.1738	-0.1896	-0.1632	0.4381
	95.0000	0.2306	0.1336	0.1064	-0.5674	0.0673	0.6876
	97.5000	0.4190	0.5435	-0.6543	-1.3970	0.9778	0.6572
	100.0000	0.3955	0.1074	-2.9374	-3.1398	1.0437	-0.8449
	102.5000	-0.2602	-0.7569	-7.2345	-5.7790	-1.1828	-5.6653
	105.0000	-0.3810	-5.8111	-12.8871	-8.7376	-10.9678	-13.3601
	107.5000	-0.1618	-12.3768	-18.0461	-13.7118	-25.5579	-21.2308
	110.0000	2.2337	-25.8865	-20.2418	-17.9238	-46.9184	-25.5478
	120.0000	180.7643	-81.5024	64.8157	85.4722	-95.1619	78.2250
90	70.0000	0.0000	-0.0637	-0.0057	0.0010	-0.1104	0.8126
	80.0000	0.0214	-0.0912	-0.0078	0.0136	-0.1787	1.1744
	90.0000	-0.0073	0.0166	0.0178	0.3650	-0.1974	2.1113
	92.5000	-0.0152	0.1732	0.0338	0.5975	0.0463	2.6489
	95.0000	-0.0730	0.2850	0.0378	1.0227	0.3934	3.4400
	97.5000	-0.3903	0.4573	0.0045	1.4916	0.9174	4.5328
	100.0000	-0.6187	0.2467	-0.0972	1.6133	0.9636	5.9255
	102.5000	-0.7268	-0.1285	-0.2906	1.5146	0.3771	7.6149
	105.0000	-0.7883	-1.5395	-0.5736	0.8694	-2.1879	9.7302
	107.5000	-0.5319	-3.4365	-0.9021	0.1625	-6.4734	12.6780
	110.0000	-0.7140	-7.0825	-1.1795	-0.1085	-13.7626	17.2144
	120.0000	-1.9399	-32.8712	1.8915	-0.5792	-56.3341	76.8027
180	70.0000	0.0054	-0.6449	-0.0804	-0.0035	-0.1265	0.1883
	80.0000	0.0006	-0.9026	-0.1099	0.1297	-0.1958	0.2639
	90.0000	0.1223	-1.3952	-0.1580	0.5332	0.0005	0.4677
	92.5000	0.3715	-1.5890	-0.1861	0.7496	0.2009	0.5556
	95.0000	0.6914	-1.9125	-0.2314	1.1004	0.3851	0.6482
	97.5000	0.9901	-2.2943	-0.3015	1.2923	0.5969	0.7271
	100.0000	1.2045	-2.9375	-0.4028	1.1952	0.5709	0.7692
	102.5000	1.2877	-3.7003	-0.5393	0.8785	0.3529	0.7557
	105.0000	1.4649	-4.8984	-0.7102	0.1913	-0.4941	0.6836
	107.5000	1.4914	-6.2778	-0.9089	-0.7723	-1.8360	0.5773
	110.0000	1.2494	-8.2693	-1.1221	-1.7651	-4.2131	0.4938
	120.0000	0.4846	-20.4278	-1.6609	-5.7578	-22.6493	2.4148

Table 4.8: Option Pricing under scenarios 3 and 4.

Percentage bias for different strike prices (*K*) and maturities (*T*) from EET (empirical Esscher transform), DLN (double lognormal) and SHM (Shimko method), under assumptions of the scenarios 3 and 4 in the table 4.4. Underlying asset current price = $ 100.00; risk-free rate = 5.93%.

		Scenario 3: $\rho = -0.215, \sigma = 0.2232$			Scenario 4: $\rho = -0.215, \sigma = 0.4464$		
T	**K**	**Bias%-EET**	**Bias%-DLN**	**Bias%-SHM**	**Bias%-EET**	**Bias%-DLN**	**Bias%-SHM**
30	70.00	0.0118	0.0886	0.0001	0.0052	0.2000	0.0104
	80.00	0.0173	0.1329	0.0000	0.0047	0.2967	0.0154
	90.00	-0.0197	0.1403	-0.0373	-0.0153	0.2759	0.0241
	92.50	-0.0285	0.0128	-0.0844	0.1044	-0.0393	-0.0376
	95.00	0.0183	-0.3116	-0.1193	0.4483	-0.6123	-0.1040
	97.50	0.2776	-0.3958	-0.0287	1.0428	-0.8211	-0.1070
	100.00	0.0334	-0.5719	0.3852	1.2365	-0.4930	0.2803
	102.50	0.3775	0.9976	1.2949	1.1667	2.2241	1.0711
	105.00	-0.3749	1.8924	2.6369	0.7526	2.4515	0.8754
	107.50	-2.0098	6.0797	4.0586	1.8946	1.6984	-0.6581
	110.00	-8.2485	6.3500	4.8899	8.1050	-10.1213	-3.3494
	120.00	30.8831	36.7322	-18.9177	13.4011	-69.8823	-26.6836
90	70.00	0.0020	0.2531	0.0000	-0.0020	0.5059	0.0245
	80.00	-0.0339	0.3300	-0.0015	-0.0597	0.6127	0.0150
	90.00	-0.2779	0.0504	-0.0150	0.0156	-0.0170	-0.1009
	92.50	-0.1363	-0.0911	-0.0141	0.2549	-0.3172	-0.1040
	95.00	-0.0288	-0.2800	-0.0028	0.6814	-0.5993	-0.0179
	97.50	0.0445	-0.2558	0.0244	1.3207	-0.5000	0.2328
	100.00	0.0896	-0.2173	0.0688	2.2239	-0.1314	0.7072
	102.50	-0.4343	0.3555	0.1226	3.2997	1.0319	1.3799
	105.00	-0.9195	0.8923	0.1657	4.3332	2.1643	2.0700
	107.50	-0.6579	2.2416	0.1666	5.4273	3.6877	2.4343
	110.00	-0.5760	3.2109	0.0896	7.2129	3.6065	2.0838
	120.00	-0.9089	7.0713	-1.0688	12.8573	-17.7861	-5.7523
180	70.00	-0.0104	0.0760	-0.0093	0.0127	-0.0178	-0.0187
	80.00	-0.0719	-0.0137	-0.0124	0.1632	0.0183	-0.0421
	90.00	0.2093	-0.3354	-0.0160	0.9195	-0.0102	-0.0969
	92.50	0.4211	-0.3740	-0.0184	1.2317	-0.0324	-0.0985
	95.00	0.6919	-0.4029	-0.0229	1.4857	-0.0895	-0.0836
	97.50	0.8193	-0.2807	-0.0311	1.7666	-0.0611	-0.0475
	100.00	0.8661	-0.1203	-0.0447	1.9949	-0.0707	0.0088
	102.50	0.9890	0.3041	-0.0653	2.1031	0.0774	0.0737
	105.00	0.8553	0.7731	-0.0939	1.8854	0.1005	0.1217
	107.50	0.5733	1.6156	-0.1297	1.3091	0.2657	0.1122
	110.00	0.1607	2.4654	-0.1695	1.0791	0.0493	-0.0061
	120.00	-0.5427	8.2436	-0.1546	0.0286	-5.0617	-2.3583

Table 4.9: MAPE under scenarios 1 and 2.

MAPE over different strike prices (*K*) and maturities (*T*) of EET (empirical Esscher transform), DLN (double lognormal) and SHM (Shimko method), under assumptions of the scenarios 1 and 2 in the table 4.4.

		Scenario 1: $\rho = 0.215, \sigma = 0.2232$			Scenario 2: $\rho = -0.215, \sigma = 0.4464$		
T	**Moneyness**	**EET**	**DLN**	**SHM**	**EET**	**DLN**	**SHM**
30	ITM	0.0098	0.0397	0.0671	0.0672	0.1309	0.1971
	ATM	0.3373	1.4705	4.7639	3.9242	2.8479	4.2430
	OTM	61.0533	39.9219	34.3679	39.0359	55.8794	41.6679
	General	15.4071	10.6064	10.5997	11.4165	15.2001	12.2694
90	ITM	0.0110	0.0862	0.0163	0.2443	0.1332	1.6868
	ATM	0.5194	0.5314	0.2007	1.3023	0.9679	1.6868
	OTM	1.0619	14.4634	1.3244	0.2834	25.5234	35.5650
	General	0.4856	3.8660	0.4204	0.6949	6.8285	12.7183
180	ITM	0.1249	1.1329	0.1336	0.3540	0.1309	0.3689
	ATM	1.1277	3.1486	0.4371	0.9315	0.4800	0.7168
	OTM	1.0751	11.6583	1.2307	2.7651	9.5662	1.1620
	General	0.7803	4.6041	0.5674	1.1974	2.6352	1.0564

Table 4.10: MAPE under scenarios 3 and 4.

MAPE over different strike prices (*K*) and maturities (*T*) of EET (empirical Esscher transform), DLN (double lognormal) and SHM (Shimko method), under assumptions of the scenarios 3 and 4 in the table 4.4.

		Scenario 3: $\rho = -0.215, \sigma = 0.2232$			Scenario 4: $\rho = -0.215, \sigma = 0.4464$		
T	**Moneyness**	EET	DLN	SHM	EET	DLN	SHM
30	ITM	0.0193	0.0936	0.0305	0.0324	0.2030	0.0219
	ATM	0.2163	0.8339	0.8930	0.9294	1.3204	0.4876
	OTM	13.7138	16.3873	9.2887	7.8002	27.2340	10.2304
	General	3.5250	4.4755	2.7044	2.3481	7.4263	2.7678
90	ITM	0.1125	0.1811	0.0076	0.0831	0.3632	0.0611
	ATM	0.3033	0.4002	0.0768	2.3718	0.8854	0.8816
	OTM	0.7143	4.1746	0.4417	8.4992	8.3601	3.4234
	General	0.3425	1.2708	0.1450	3.1407	2.5800	1.2466
180	ITM	0.1782	0.1998	0.0140	0.5818	0.0197	0.0640
	ATM	0.8443	0.3762	0.0516	1.8471	0.0798	0.0671
	OTM	0.4256	4.1082	0.1512	0.8056	1.7922	0.8255
	General	0.5176	1.2504	0.0682	1.1649	0.4879	0.2562

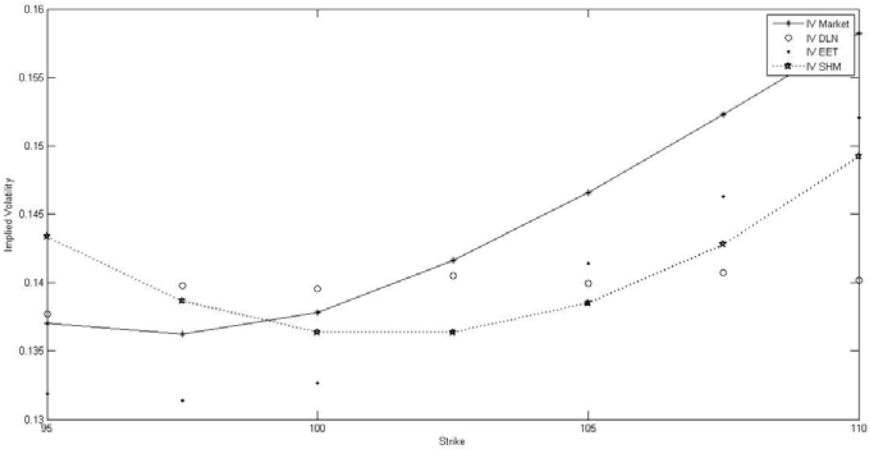


Figure 4.7: Volatility smile – $T = 30, \rho = 0.215, \sigma = 0.4464$.

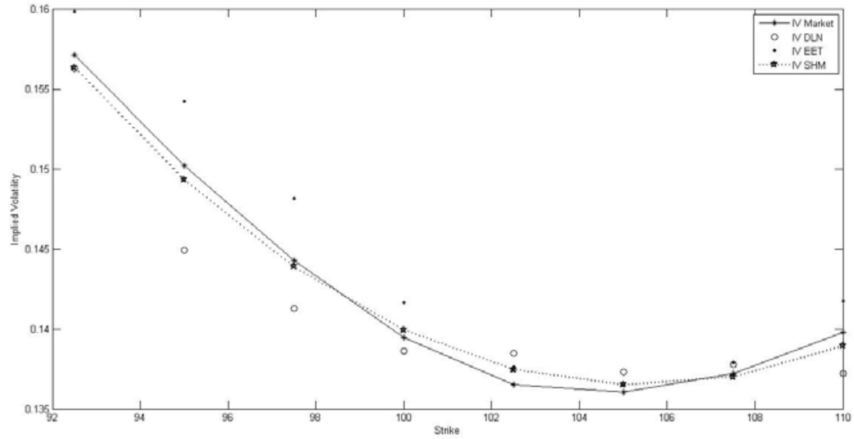


Figure 4.8: Volatility smile $\mathrm{T} = 30, \rho = -0.215, \sigma = 0.4464$.

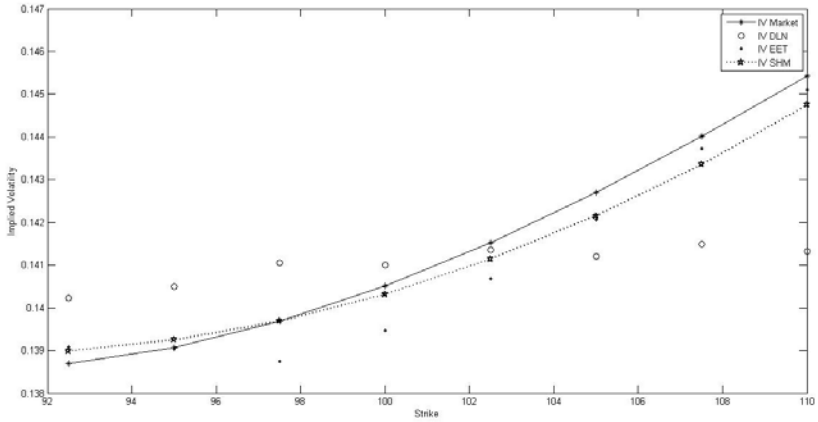


Figure 4.9: Volatility smile – $T = 90, \rho = 0.215, \sigma = 0.2232$.

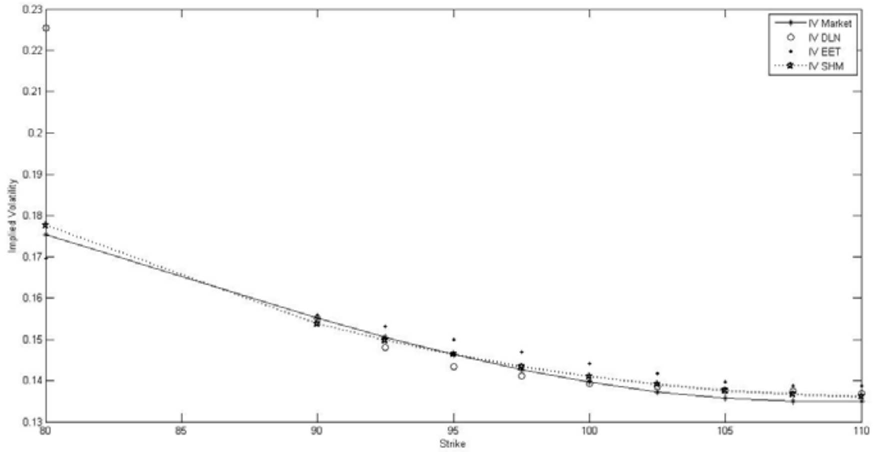


Figure 4.10: Volatility smile – $T = 90, \rho = -0.215, \sigma = 0.4464$.

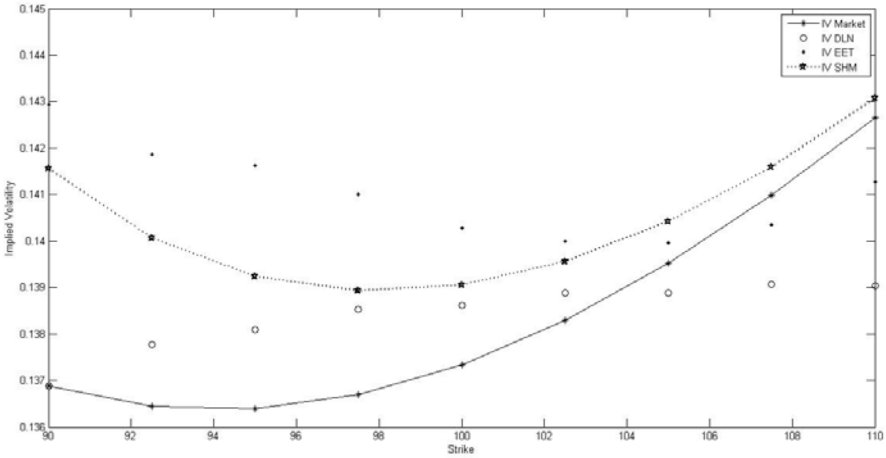


Figure 4.11: Volatility smile – $T = 180, \rho = 0.215, \sigma = 0.4464$.

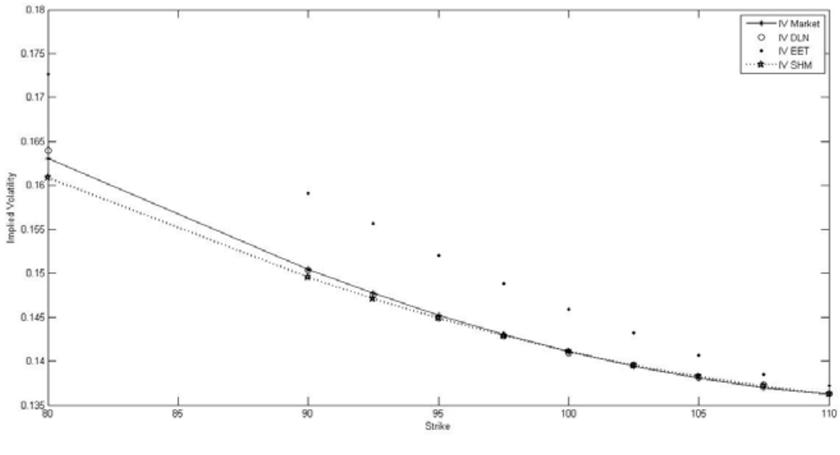


Figure 4.12: Volatility smile – $T = 180, \rho = -0.215, \sigma = 0.4464$.

4.6 Conclusions

This work proposes a new approach for the estimation of risk-neutral distribution. We develop a discrete version of Breeden and Litzenberger (1978) where states are not equally spaced. Our method can be classified as an indirect way of estimation because we estimate the physical distribution of the underlying asset's historical prices and the empirical Esscher parameter from option market prices. Following, we fit a polynomial of degree two, between the h-implied and the strike price, in order to estimate the state price.

The most straightforward application of the risk-neutral distribution is pricing any payoff with the same time until expiration (including illiquid and exotic options). We ran simulation experiments under different situations, which seek to highlight the differences and similarities between the methods. We compare our method to two approach alternatives: Double Lognormal proposed by Baha (1997) and the Shimko (1993) method. We calculate the European call option prices for each method according to various scenarios and maturities. Our method shows better results in various scenarios and, when we analyze the volatility smile (or smirk), our method reproduces market asymmetry.

Further research can be done comparing the proposed method to others methodologies, studying other option-implied information and its applications, verifying the results with parametric data generating processes, and using other functions instead of polynomials to help with pricing accuracy.

5
Conclusion

One of the central questions in quantitative finance is how to get a measure of the risk-neutral probability that provides theoretical prices closer to those observed in the market. The literature highlights two approaches to this problem: models based on the general equilibrium (Arrow, 1964, Debreu, 1959, Lucas, 1978, Rubinstein, 1976) and the models based on absence of arbitrage (Black-Scholes, 1973, Cox and Ross, 1976, Harrison and Kreps, 1979, Harrison and Pliska, 1981).

In both cases, these approaches require the formulation of an explicit risk-neutral model and are restricted to a few probability distributions for modeling the economy's uncertainty. However, empirical observations of asset returns showed several stylized facts, which highlight the parametric misspecification risk for the used stochastic process. Hence, due to the poor empirical performance of parametric methods, the nonparametric option pricing techniques have expanded rapidly in recent years, because they offer an alternative by avoiding possibly biased parametric restrictions (Haley and Walker, 2010).

The main objective of this thesis is to verify if simple assumptions on empirical pricing kernel are able to generate a measure Q that produces theoretical prices closer to those observed in the market. From our investigation, we are able to derive three articles.

The first article (Chapter 2) introduces the empirical Esscher transform and studies the nonparametric option pricing. In our proposal, we make only mild assumptions on the price kernel and there is no need for the formulation of a risk-neutral model. First, we simulate sample paths for the returns under the physical measure P. Then, based on the empirical Esscher transform, the sample is reweighted, giving rise to a risk-neutralized sample from which derivative prices can be obtained by a weighted sum of the options' pay-offs in each path.

We conduct artificial experiments in Black-Scholes and Heston worlds and real experiments to explore the potential usefulness of the proposed method. Artificial results show that the EET prices improve along with the sample size.

Real data results show that, when the stochastic process of underlying asset is unknown, the lowest pricing errors are between the nonparametric methods. For a maturity equal to 17, the nonparametric methods present similar results. For others maturities, the proposed method presents the lowest MAPE for moneyness equal to deep-out-of-the-money, out-of-the-money and at-the-money.

We also analyze the behavior of the empirical Esscher parameter. We can highlight that the standard deviation decreases along with the maturity and with the increase in the sample size, and the values of the descriptive statistics begin to converge to a constant value in larger samples. When we compare only between the empirical Esscher parameter obtained for synthetic and real data, the more important change is the signal. That is, the Esscher parameters obtained with synthetic data are simulated with a drift ($\mu = 10.00\%$) greater than the risk-free rate ($r = 5.00\%$). Thus, the negative parameter shifts the risk-neutral distribution to the left, which eliminates the risk premium and assures the average yield equal to the risk-free rate. With real data, the opposite happens. The positive parameter shifts the risk-neutral distribution to the right. This is contrary to financial theory. However, this does not constitute an arbitrage opportunity, because the daily risk-free rate is between the worst and the best daily return. When price time series are in falling, applications in risk-free interest rates are paying more than in these stocks.

In the second article (Chapter 3), we demonstrate that our proposal is flexible and performs very well in the presence of realistic financial time series. We use a recently proposed dynamic conditional score models, developed by Harvey (2013), which offers an alternative to model the volatility (and heavy tails) of observed underlying asset price using GARCH models and analyzes the nonparametric option pricing method.

We empirically compare our approach to competing benchmark approaches, like Black-Scholes (1973) and Heston and Nandi (2000), with real data. In general, the proposed method with an assumed model to describe the empirical distribution presents the lowest MAPE. When we increase the size of the empirical distribution, only for lower maturity the MAPE was reduced.

In our third contribution (Chapter 4), we introduce a new approach for indirect estimation of the implicit state-price in financial asset prices using the empirical Esscher transform. First, we generalize the discrete version of the

Breeden and Litzenberger (1978) method for the case where states are not equally spaced. Second, we use the empirical Esscher transform to include underlying assets' and derivatives' data. We use the historical distribution of the underlying asset prices and the observed option prices to estimate the implicit empirical Esscher parameter. Then, we fit polynomials between the implied Esscher parameter and the strike price, as in Shimko (1993), to obtain our measure Q.

We run simulation experiments under different situations, which seek to highlight the differences and similarities between the methods. We compare our method to two approach alternatives: Double Lognormal proposed by Baha (1997) and the Shimko (1993) method. We calculate the European call option prices for each method according to various scenarios and maturities. Our method shows better results in various scenarios and, when we analyze the volatility smile (or smirk), our method reproduces market asymmetry.

Bibliography

Aït-Sahalia, Y. and Duarte, J. (2003). Nonparametric Option Pricing Under Shape Restrictions. **Journal of Econometrics**, 116, 9-47.

Aït-Sahalia, Y. and A. Lo. (1998). Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices. **Journal of Finance**, vol. 53, no. 2 (April): 499–547.

———. (2000). Nonparametric Risk Management and Implied Risk Aversion. **Journal of Econometrics**, vol. 94, nos. 1-2 (January/February): 9-51.

Alcock, J. and Carmichael, T. (2008). Nonparametric American option pricing. **Journal of Futures Markets**, 28, 717–748.

Almeida, C.; Ardison, K.; Garcia, R. and Vicente, J. (2016). Nonparametric Tail Risk, Stock Returns and the Macroeconomy. Working Paper. FGV.

Almeida, C. and Azevedo R. (2014). Nonparametric Option Pricing with Generalized Entropic Estimators. Working Paper. FGV.

Aparicio, S.D. and Hodges, S. (1998). Implied Risk-Neutral Distribution: A Comparison of Estimation Methods. Working paper, University of Warwick.

Arrow, K.J. (1964). The Role of Securities in the Optimal Allocation of Risk-bearing. **Review of Economic Studies**. v.31, p.91-96.

Badescu A. and Kulperger R., (2007). GARCH option pricing: A semiparametric approach. **Insurance: Mathematics and Economics**, 43(1), 69-84.

Badescu, A.; Elliott, R.; Kulperger, R.; Miettinen, J. and Siu, T.K. A Comparison of Pricing Kernels for GARCH Option Pricing with Generalized Hyperbolic Distributions. International Journal of Theoretical and Applied Finance 14. (5) (2011): 669-708.

Bahra, B. (1997). Implied Risk-Neutral Probability Density Functions From Option Prices: Theory and Application. Working paper, Bank of England.

Bakshi, G.; Cao, C. and Chen, Z. (1997). Empirical Performance of Alternative Option Pricing Models. **Journal of Finance**, 52, 2003-2049.

Bakshi, G.; Panayotov, G. and Skoulakis, G. (2011). Improving the Predictability of Real Economic Activity and Asset Returns with Forward Variances Inferred from Option Portfolios. **Journal of Financial Economics**, 100, 475-495.

Bams, D.; Lehnert, T. and Wolff, C. (2009). Loss Functions in Option Valuation: A Framework for Model Selection. **Management Science**, 55: 853-862.

Barone-Adesi, G.; Engle, R.-F. and Mancini, L. (2008). A GARCH Option pricing model with filtered historical simulation. **Review of Financial Studies**, 21, 1223-1256.

Bates, D. (1991). The Crash of ‘87: Was It Expected? The Evidence from Options Markets. **Journal of Finance**, vol. 46, no. 3 (July): 1009-44.

———. (2000). Post-’87 Crash Fears in S&P 500 Futures Options. **Journal of Econometrics**, vol. 94, nos. 1-2 (January/February): 181-238.

Binghan, N. and Kiesel, R. (2004). Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives. Springer, 2004.

Birru, J. and Figlewski, S. (2012). Anatomy of a meltdown: The risk neutral density for the S&P 500 in the fall of 2008. **Journal of Financial Markets**, 15, 151-180.

Black, F. and Scholes, M. (1973). The Pricing of Options and Corporate Liabilities. **The Journal of Political Economy**, 81, 637-654.

Bliss, R. and N. Panigirtzoglou. (2001). Testing the Stability of Implied Probability Density Functions. **Journal of Banking and Finance**, vol. 26, no. 2-3 (March): 381-422.

———. (2002). Option-Implied Risk Aversion Estimates: Robustness and Patterns. Previously titled: Recovering Risk Aversion from Options. Working paper, Federal Reserve Bank of Chicago.

Bollerslev, T. (1987). A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. **Review of Economics and Statistics**, 69: 542-547.

Bollerslev, T. and Mikkelsen, H.O. (1996). Modeling and Pricing Long Memory in Stock Market Volatility. **Journal of Econometrics**, 73, 151-184.

Bollerslev, T. and Todorov, V. (2011). Tails, Fears, and Risk Premia. **The Journal of Finance**, 66, 2165-2211.

Bondarenko, O. (2003). Estimation of Risk-Neutral Densities Using Positive Convolution Approximation. **Journal of Econometrics**, 116, 85-112.

Bondarenko, O. (2000). Recovering Risk-Neutral Densities: A New Nonparametric Approach. Working paper, University of Illinois.

Breeden, D.T. and Litzenberger, R.H. (1978). Prices of State-Contingent Claims Implicit in Option Prices. **The Journal of Business**, 51, 621-651.

Breiman, L. (1968). Probability. Menlo Park, Calif.: Addsion-Wesley. Reprint. Philadelphia: Society of Industrial and Applied Mathematics, 1992.

Brennan, M.J. (1979). The Pricing of Contingent Claims in Discrete Time Models. **The Journal of Finance**, 34, 53-68.

Broadie, M.; Chernov, M. and M. Johannes. (2007). Model Specification and Risk Premia: Evidence from Futures Options. **Journal of Finance**, 62: 1453.1490.

Brunner, B. and Hafner, R. (2003). Arbitrage-Free Estimation of the Risk-Neutral Density From the Implied Volatility Smile. **Journal of Computational Finance**, 7(1): 75-106.

Buchen, P. and Kelly, M. (1996). The maximum entropy distribution of an asset inferred from option prices. **Journal of Financial and Quantitative Analysis**, 31, 143-159.

Buss, A. and Vilkov, G. (2011). Option-Implied Correlation and Factor Betas Revisited. Working paper, Goethe University.

Bühlmann, H.; Delbaen, F.; Embrechts, P. and Shiryaev, A.N. (1996). No-Arbitrage, Change of Measure and Conditional Esscher Transforms, CWI Quarterly 9(4): 291-317.

Campa, J.M.; Chang, P. and Reider, R.L. (1998). Implied Exchange Rate Distributions: Evidence from OTC Option Markets. **Journal of International Money and Finance**, 17, 117-160.

Casella, G.; Berger, R.L. (2001). Statistical Inference. Cengage Learning. 2nd edition.

Carr, P. (2001). Constraints on Implied Volatility. Working Paper. Bank of America Securities.

CBOE. (2009). The CBOE Volatility Index - VIX. Technical Report.

CBOE. (2011). The CBOE Skew Index - SKEW. Technical Report.

Chan, T. (1999). Pricing Contingent Claims on Stocks Driven by Levy Processes. **Annals of Applied Probability**, 9(2): 504-28.

Chang, B.Y.; Christoffersen, P. and Jacobs, K. (2013). Market Skewness Risk and the Cross-Section of Stock Returns. **Journal of Financial Economics**, 107, 46-68.

Chang, B.-Y.; Christoffersen, P.; Jacobs, K. and Vainberg, G. (2011). Option-Implied Measures of Equity Risk. **Review of Finance**, Forthcoming.

Cheng, K.C. (2010). A New Framework to Estimate the Risk-Neutral Probability Density Function Embedded in Options Prices. Working Paper. International Monetary Fund.

Chernov, M.; Gallant, A.R.; Ghysels, E. and Tauchen, G. (2003). Alternative Models for Stock Price Dynamics. **Journal of Econometrics**, 116, 225-257.

Chorro, C.; Guegan, D. and Ielpo, F. (2008), Documents de Travail du Centre d'Economie de la Sorbonne - Option Pricing under GARCH models with Generalized Hyperbolic innovations (II) : Data and results.

Christoffersen, P.; Elkamhi, R.; Feunou, B. and Jacobs, K. (2010). Option Valuation with Conditional Heteroskedasticity and Nonnormality. **Review of Financial Studies**, 23, 2139-2183.

Christoffersen, P.; Jacobs, K. and Ornthanalai, C., (2012). Dynamic jump intensities and risk premiums: Evidence from S&P500 returns and options. **Journal of Financial Economics**, 106: 447-472.

Christoffersen, P.; Jacobs, K. and Heston, S. (2011). A GARCH Option Model with Variance-Dependent Pricing Kernel. Working Paper. University of Toronto. Canada.

Christoffersen, P.; Jacobs, K. and Wang, Y. (2004). Option Valuation with Long-run and Short-run Volatility Components. Cirano Working Papers.

Christoffersen, P.; Jacobs, K. and Ornthanalai, C. (2013). GARCH Option Valuation: Theory and Evidence, Rotman, CBS and CREATES University of Houston and Tilburg Rotman School, University of Toronto June 13, 2013

Christoffersen, P.; Jacobs, K. and Chang (2011). Forecasting with option implied information. Business and Social Sciences, Departament of Economics and Business.

Christoffersen, P.; Jacobs, K. and Chang, K. (2006). Volatility components: affine restrictions and Nonnormal innovations. working paper, McGill University.

Christoffersen, P.; Heston, S. and Jacobs, K. (2006). Option Valuation with Conditional Skewness. **Journal of Econometrics**, 131: 253-84.

Cont, R. (1997). Beyond Implied Volatility. In Econophysics. Edited by J. Kertesz and I. Kondor. Dordrecht, Netherlands: Kluwer.

Cox, J.C. and Ross, S.A. (1976). The valuation of options for alternative stochastic processes. **Journal of Financial Economics**, v. 3, p. 145-146.

Cox, J.C.; Ross, S.A. and Rubinstein, M. (1979). M. Option pricing: A simplified approach. **Journal of Financial Economics**, 7: 229-63.

Creal, D.; Koopman, S.J. and Lucas, A. (2013). Generalized autoregressive score models with applications. **Journal of Applied Econometrics**, 28, 777-95.

Danthine, J.P. and Donaldson, J.B. (2015). Intermediate Financial Theory. Academic Press. 2º edition.

Debreu, G. (1959): Theory of Value, Cowles Foundation Monograph n. 17. EUA: John Wiley and Sons, p. 104.

De Miguel, V.; Plyakha, Y.; Uppal, R. and Vilkov, G. (2011). Improving Portfolio Selection Using Option-Implied Volatility and Skewness. Working paper, London Business School.

Derman, E. and Kani, I. (1994). Riding on a Smile. Risk, vol. 7, no. 2 (February): 32-39.

Duan, J. (1995). The GARCH Option Pricing Model. **Mathematical Finance**, vol. 5, no. 1 (January): 13-32.

Duan, J.-C. (1999). Conditionally Fat-Tailed Distributions and the Volatility Smile in Options. Working Paper, National University of Singapore.

Duan, J.-C. (2002). Nonparametric Option Pricing by Transformation., Working Paper, Rotman School of Management.

Duan, J.C. and Zhang, W. (2013). Forward-Looking Market Risk Premium. Working Paper. Risk Management Institute, NUS Business School, and Department of Economics, National University of Singapore, Singapore

Engle, R. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of UK Inflation. **Econometrica**, 50: 987.1008.

Engle, R. and Lee, G. (1999). A Long-Run and Short-Run Component Model of Stock Return Volatility in Cointegration, Causality and Forecasting. Edited by R. Engle and H. White, Oxford University Press.

Eraker, B.; Johannes, M. and Polson, N. (2003). The Impact of Jumps in Volatility and Returns. **The Journal of Finance**, 58, 1269-1300.

Esscher, F. (1932). On the probability function in the collective theory of risk. **Skandinavisk Aktuarietidskrift**, 15: 175-195.

Figlewski, S. (2010). Estimating the Implied Risk Neutral Density for the U.S. Market Portfolio. In Bollerslev, T., Russell, J. R., and Watson, M. (Eds.). Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle, Oxford University Press, 323-353.

Figlewski, S. and Malik, M.F. (2014). Options on Leveraged ETFs: A Window on Investor Heterogeneity. Working paper. New York University – Stern School of Business.

Garcia, R.; Lewis, M.; Pastorello, S. and Renault, P. (2011). Disasters implied by equity index option. **Journal of Econometric**, 160: 22-32.

Gerber, H. and Shiu, E. (1994). Option pricing by esscher transforms. **Transactions of Society of Actuaries**, 46: 99-191.

Gerber, H.U. and Shiu, E.S.W. (1996). Actuarial Bridges to Dynamic Hedging and Option Pricing. **Insurance: Mathematics and Economics**, 18: 183-218.

Giamouridis, D. and Skiadopoulos, G. (2012). The informational content of financial options for quantitative asset management. A Review. In B. Scherer and Winston (Eds.), The Oxford Handbook of Quantitative Asset Management (pp. 243-265). New York: Oxford University Press.

Gray, P. and Newman, S. (2005). Canonical valuation of options in the presence of stochastic volatility. **Journal of Futures Markets**, 25(1): 1-19.

Gray, P.; Edwards, S. and Kalotay, E. (2007). Canonical Pricing and Hedging of Index Options. **Journal of Futures Markets**, Forthcoming. 1, 16

Grith, M.; Härdle, W.K. and Schienle, M. (2012). Nonparametric Estimation of Risk-Neutral Densities. **Handbook of Computational Finance**.

Guegan, D., Ielpo, F. and Lalaharison, H. (2013). Option Pricing with Discrete Time Jump Processes. To appear in the **Journal of Economic Dynamics and Control**.

Haley, R. and Walker, T. (2010). Alternative tilts for nonparametric option pricing. **Journal of Future Markets**, 30(10): 983-1006.

Hansen, L.P and Jagannathan, R. (1991). Implications of Security Market Data for Models of Dynamic Economies. **Journal of Political Economy**, vol. 99, n. 2, 225-62.

Harrison, J.M.; Kreps, D.M. (1979). Martingales and arbitrage in multiperiod securities markets. **Journal of Economic Theory**, 20: 381-408.

Harrison, J.M. and Pliska, S.R. (1981). Martingales and stochastic integrals in the theory of continuous trading. **Stochastic Processes and Their Applications**, 11: 215-280.

Harrison, J.M. and Pliska, S.R. (1983). A stochastic calculus model of continuous trading: Complete markets. **Stochastic Processes and Their Applications**, 15: 313-316.

Harvey, A.C. (2013). Dynamic Models for Volatility and Heavy Tails. Econometric Society Monograph. Cambridge University Press.

Harvey, A.C. Chakravarty, T. (2009). Beta-t-EGARCH. Working paper. Earlier version appeared in 2008 as a Cambridge Working Paper in Economics, CWPE 0840.

He, C. and Teräsvirta, T. (1999). Properties of moments of a family of GARCH processes. **Journal of Econometrics,** 92, 173-92.

Heston, S. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. **Review of Financial Studies**, vol. 6, no. 2 (Summer): 327-343.

Heston, S.L. and Nandi, S. (2000). A Closed-Form GARCH Option Valuation Model. **Review of financial Studies**, 13 (3), 585-625.

Hui, C.H.; Lo, C.F. and Lau, C.S. (2013). Option-implied correlation between iTraxx Europe Financials and Non-Financials Indexes: A measure of spillover effect in European debt crisis. **Journal of Banking and Finance**, 37 (9), 3694-3703.

Hutchinson, J.A.; Lo and T. Poggio. (1994). A Nonparametric Approach to Pricing and Hedging Derivative Securities via Learning Networks. **Journal of Finance**, vol. 49, no. 3 (July): 851-889.

Jackwerth, J. (1997). Generalized Binomial Trees. **Journal of Derivatives**, vol. 5, no. 2 (Winter): 7-17.

———. (1999). Option Implied Risk-Neutral Distributions and Implied Binomial Trees: A Literature Review. **Journal of Derivatives**, vol. 7, no. 2 (Winter): 66-82.

———. (2000). Recovering Risk Aversion from Option Prices and Realized Returns. **Review of Financial Studies**, vol. 13, no. 2 (Summer): 433-451.

Jackwerth, J.C. (2004). Option-Implied Risk-Neutral Distribution and Risk Aversion – The Research Foundation of AIMR (Association for Investment Management and Research.

Jackwerth, J. and M. Rubinstein. (1996). Recovering Probability Distributions from Option Prices. **Journal of Finance**, vol. 51, no. 5: 1611–31.

Kitsul, Y. and Wright, J.H. (2013). The economics of options-implied inflation probability density functions. **Journal of Financial Economics**, 110, 696-711.

Kostaskis, A.; Panigirtzoglou, N. and Skiadopoulos, G. (2011). Market Timing with Option-Implied Distributions. A Forward-Looking Approach. **Management Science**, 57 (7), 1231-1249.

Li, S. and Badescu, A. (2012). University of Calgary – Risk neutral measures and GARCH model calibration – Department of Mathematics and Statistics.

Lin, Y.-N.; Strong, N. and Xu, X. (2001). Pricing FTSE 100 index options under stochastic volatility. **Journal of Futures Markets**, 21, 197-211.

Liu, Y.; Li , J.; S.H., Ng. A.C.Y. (2015). Option pricing under GARCH models with Hansen’s skewed-t distributed innovations. **North American Journal of Economics and Finance**, 31, 108-125.

Lucas, R.E. (1978). Asset Prices in an Exchange Economy. **Econometrica**, vol. 46, No. 6. Nov., 1978, pp. 1429-1445.

Malz, A.M. (1997). Option-Implied Probability Distribution and Currency Excess Returns. Staff Reports, Federal Reserve Bank of New York.

Markose, S, and Alentorn, A. (2011). The Generalized Extreme Value Distribution, Implied Tail Index, and Option Pricing. **Journal of Derivatives**, 18, 35-60.

Melick, W. and Thomas, C. (1997). Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil during the Gulf Crisis. **Journal of Financial and Quantitative Analysis**, vol. 32, no. 1 (March): 91-115.

Menn, C. and Rachev, S.T. (2005). Smoothly truncated stable distributions, GARCH-models, and option pricing. Working paper, Cornell University and University of Karlsruhe and UCSB.

Monfort, A. and Pegoraro, F. (2011). Asset pricing with second-order Esscher transform. **Journal of Banking and Finance**, 36, 1678-1687.

Monnier, J.B. (2013). Risk-Neutral Density Recovery via Spectral Analysis. **SIAM Journal on Financial Mathematics**, 4: 650-667.

Nelson, D. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. **Econometrica**, 59: 347.70.

Nelson, D.B. and C.Q. Cao (1992). Inequality constraints in the univariate GARCH model. **Journal of Business and Economic Statistics**, 10, 229-235.

Nelson, D. (1996). Asymptotically Optimal Smoothing with ARCH Models. **Econometrica**, 64, 561-573.

Ornelas, J.R H. and Takami, M.Y. (2009). **Brazilian Finance Review**, 1, 9-26.

Orthanalai, C. (2011). Levy Jump Risk: Evidence from options and returns. **Journal of Financial Economics**, 112 (1), 69-90.

Pérignon, C. and Villa, C. (2002). Extracting Information from Option Markets: Smiles, State-Price Densities and Risk Aversion. **European Financial Management**, vol. 8, 495-513.

Rosenberg, J. and R. Engle. (2002). Empirical Pricing Kernels. **Journal of Financial Economics**, vol. 64, no. 3 (June): 341-372.

Rubinstein, M. (1976). The Valuation of Uncertain Income Streams and the Pricing Options. **Bell Journal Econ. Management Sci**., 7, 407-425.

Rubinstein, M. (1985). Nonparametric Tests of Alternative Option Pricing Models Using All Reported Trades and Quotes on the 30 Most Active CBOE Option Classes from August 23, 1976 through August 31, 1978. **Journal of Finance**, vol. 40, no. 2 (June): 455-480.

———. (1994). Implied Binomial Trees. **Journal of Finance**, vol. 49, no. 3 (July): 771-818.

Santa-Clara, P. and S. Yan. (2010). Crashes, Volatility, and the Equity Premium: Lessons from S&P 500 Options. **Review of Economics and Statistics**, 92: 435.451.

Santos, A. and Guerra, J. (2015). Implied Risk Neutral Densities From Option Prices: Hypergeometric, Spline, Lognormal, and Edgeworth Functions. **Journal of Futures Markets**, 35: 655-678.

Satoyoshi, K. and Mitsui, H. (2011). Empirical study of Nikkei 225 options with the Markov switching GARCH model. **Asia-Pacific Financial Markets**, 18, 55-68.

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

Siu, T.K.; Tong, H. and Yang, H. (2004). On Pricing Derivatives Under GARCH Models: A Dynamic Gerber-Shiu Approach. **North American Actuarial Journal**, 8(3): 17-31.

Stutzer, M. (1996). A Simple Nonparametric Approach to Derivative Security Valuation. **Journal of Finance**, vol. 51, no. 5 (December): 1633-52.

Tsay, R.S. (2013). An introduction to Analysis of Financial Data with R. Wiley; 1a Edition.

Zhang, J.E. and Shu, J. (2003). Pricing S&P 500 index options with Heston’s model. Proceedings of IEEE 2003 International Conference on Computational Intelligence for Financial Engineering (CIFE, 2003), March 21-23, 2003, Hong Kong (pp. 85-92).

6
Appendix

6.1
Option Pricing by Gerber and Shiu

The price of a European call on a non-dividend-paying stock is obtained under the risk-neutral distribution $q(S_T)$ and the payoff is discounted at the deterministic risk-free rate r:

$$C = e^{-rT}E[max(S_T - k, 0)] = e^{-rT}\int_{x^*}^{+\infty} max(S_T - K, 0)f(x, T; h)dx \quad (6.1)$$

where T is the time to maturity, S_T is the underlying asset price, K is the strike price, $f(x, T; h)$ is the risk-neutral distribution of the asset price at the option's expiration. Consider:

$$S_T = S_0 e^{X_T} \quad (6.2)$$

$$f(x, T; h) = \frac{e^{hx}f(x, T)}{\int_{-\infty}^{+\infty} e^{hx}f(x, T)dx} = \frac{e^{hx}f(x, T)}{M_{X_T}(h, T)} \quad (6.3)$$

$$M_{X_T}(h,) = \int_{-\infty}^{+\infty} e^{hx}f(x, T)dx. \quad (6.4)$$

The lower bound is:

$$K = S_o e^x \rightarrow e^x = \frac{K}{S_o}$$

$$\ln(e^x) = \ln\left(\frac{K}{S_o}\right) \rightarrow \mathrm{x}^* = \ln(\mathrm{K}/\mathrm{S_o}). \quad (6.5)$$

Rewrite (6.1), using the equations (6.2)-(6.5):

$$C = e^{-rT}\left[S_o \int_{x^*}^{+\infty} e^{(1+h)x} \frac{1}{M_{X_T}(h,T)} f(x,T)dx - K \int_{x^*}^{+\infty} e^{hx} \frac{1}{M_{X_T}(h,T)} f(x,T)dx\right]. \tag{6.6}$$

Consider the probability distribution $X \sim N(\mu, \sigma^2)$:

$$f(x,T) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$

Then (6.4) will be:

$$M_{X_T}(h,T) = \int_{-\infty}^{+\infty} e^{hx} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx. \tag{6.7}$$

Let the changing of variable in (6.7):

$$y = hx \rightarrow x = y/h \tag{6.8}$$

$$\frac{dx}{dy} = \frac{1}{h} \rightarrow dx = \frac{1}{h} dy. \tag{6.9}$$

Replace (6.8) and (6.9) in (6.7):

$$M_{X_T}(h,T) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}h\sigma} e^{-\frac{1}{2}\left(\frac{y-h\mu}{h\sigma}\right)^2 + y} dy. \tag{6.10}$$

Solve only the exponential of (6.10):

$$= \frac{y^2 - 2yh(\mu + h\sigma^2) + h^2\mu^2}{2h^2\sigma^2}$$

add and subtract $h^2(\mu + h\sigma^2)^2$:

$$= -\frac{1}{2}\left[\frac{\left(y - h(\mu + h\sigma^2)\right)}{\sigma h}\right]^2 + \left[\mu h + \frac{1}{2}h^2\sigma^2\right]. \quad (6.11)$$

Rewritte (6.10) using (6.11):

$$M_{X_T}(h,T) = e^{\left[\mu h + \frac{1}{2}h^2\sigma^2\right]} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}h\sigma} e^{-\frac{1}{2}\left[\frac{\left(y-h(\mu+h\sigma^2)\right)}{\sigma h}\right]^2} dy. \quad (6.12)$$

We need to transform (6.12) in a standardized normal:

$$z = \frac{y - h(\mu + h\sigma^2)}{\sigma h}$$

$$\frac{dz}{dy} = \frac{1}{\sigma h} \therefore dy = \sigma h dz$$

then:

$$M_{X_T}(h,T) = e^{\left[\mu h + \frac{1}{2}h^2\sigma^2\right]} \underbrace{\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz}_{=1} = e^{\left[\mu h + \frac{1}{2}h^2\sigma^2\right]}. \quad (6.13)$$

Replacing (6.13) in (6.6), we have:

$$C = \frac{e^{-rT}}{e^{\left[\mu h + \frac{1}{2}h^2\sigma^2\right]}} \left[S_o \int_{x^*}^{+\infty} e^{(1+h)x} f(x,T)dx - K \int_{x^*}^{+\infty} e^{hx} f(x,T)dx\right]. \quad (6.14)$$

Parcels in (6.14) are expected value of the truncated normal:

$$E\left[e^{(1+h)x}\right] = \int_{x^*}^{+\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 + (1+h)x} dx$$

$$= e^{\left[\mu(1+h)+\frac{1}{2}(1+h)^2\sigma^2\right]} \int_{x^*}^{+\infty} \frac{1}{\sqrt{2\pi}(1+h)\sigma} e^{-\frac{1}{2}\left(\frac{y-(1+h)\left(\mu+(1+h)\sigma^2\right)}{\sigma(1+h)}\right)^2} dy. \qquad (6.15)$$

We need to transform (6.15) in standardized normal:

$$z = \frac{y-(1+h)(\mu+(1+h)\sigma^2)}{\sigma(1+h)}$$

$$\frac{dz}{dy} = \frac{1}{\sigma(1+h)} \therefore dy = \sigma(1+h)dz$$

and lower bound:

$$L_z = \frac{x^* - \mu^* T}{\sigma\sqrt{T}}, with\ \mu^* = \mu + (1+h)\sigma^2$$

then:

$$E\left[e^{(1+h)x}\right] = e^{\left[\mu(1+h)+\frac{1}{2}(1+h)^2\sigma^2\right]} N(L_z). \qquad (6.16)$$

By symmetry of the normal distribution, we have:

$$\int_{-\infty}^{-L_z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z)^2}\, dz = \int_{-L_z}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z)^2}\, dz.$$

Then (6.16) will be:

$$E\left[e^{(1+h)x}\right] = e^{\left[\mu(1+h)+\frac{1}{2}(1+h)^2\sigma^2\right]} N\left(-\frac{x^* - \mu^* T}{\sigma\sqrt{T}}\right)$$
$$= e^{\left[\mu(1+h)+\frac{1}{2}(1+h)^2\sigma^2\right]} N\left(\frac{-ln(K/S_o) + \mu T + (1+h)\sigma^2 T}{\sigma\sqrt{T}}\right). \qquad (6.17)$$

The expected value of the first parcel in (6.14) is:

$$\int_{x^*}^{+\infty} e^{(1+h)x} f(x,T)dx = e^{\left[\mu(1+h)+\frac{1}{2}(1+h)^2\sigma^2\right]} N(d_1) \tag{6.18}$$

$$d_1 = \frac{ln(S_o/K) + \mu\tau + (1+h)\sigma^2 T}{\sigma\sqrt{T}} \tag{6.19}$$

and of the second parcel in (6.14) is:

$$\int_{x^*}^{+\infty} e^{hx} f(x,T)dx = e^{\left[\mu h+\frac{1}{2}h^2\sigma^2\right]} N(d_2) \tag{6.20}$$

$$d_2 = \frac{ln(S_o/K) + \mu\tau + h\sigma^2 T}{\sigma\sqrt{T}}. \tag{6.21}$$

Replace the results of (6.18)-(6.21) in (6.14):

$$C = S_o e^{-rT+(\mu+h\sigma^2)+\frac{1}{2}\sigma^2} N(d_1) - Ke^{-rT} N(d_2). \tag{6.22}$$

The h^* depends on of the condition martingale by equation (2.8):

$$e^r = \left\{ \frac{M_{X_T}(1+h;T)}{M_{X_T}(h;T)} = e^{\mu+h\sigma^2+\frac{1}{2}\sigma^2} \right\}$$

$$ln(e^r) = ln\left(e^{\mu+h\sigma^2+\frac{1}{2}\sigma^2}\right)$$

$$h = \frac{r-\mu}{\sigma^2} - \frac{1}{2} \tag{6.23}$$

Replace (6.23) in (6.22):

$$C = S_o N(d_1) - Ke^{-rT} N(d_2)$$

$$d_1 = \frac{ln(S_o/K) + \left(r + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}$$

$$d_2 = \frac{ln(S_o/K) + \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}.$$

Finally, we have the Black and Scholes formula.

6.2
Jacobian Method

Let ε_t a continuous random variable with density $f(\varepsilon_t)$. We want to find the density of a new random variable $y_t = h(\varepsilon_t)$, where $h(.)$ is a function. The density of y_t, $f(y_t)$, can be found from following procedure:

$$f(y_t) = f(\varepsilon_t)\left|\frac{\partial \varepsilon_t}{\partial y_t}\right| \qquad (6.24)$$

where the absolute value $|\partial\varepsilon_t/\partial y_t|$, called Jacobian term, is to ensure that $f(y_t)$ is not negative.

Consider the equations:

$$y_t = \mu + \sqrt{h_t}z_t \qquad (6.25)$$

$$z_t = \left(\frac{\upsilon - 2}{\upsilon}\right)^{\frac{1}{2}}\varepsilon_t, \qquad \upsilon > 2 \qquad (6.26)$$

$$\varepsilon_t \sim t_\upsilon\left(0, \frac{\upsilon}{\upsilon - 2}\right) \qquad (6.27)$$

and the t-student distribution:

$$f(\varepsilon_t; \upsilon) = \frac{\Gamma((\upsilon+1)/2)}{\Gamma(\upsilon/2)\sqrt{\pi\upsilon}}\left(1+\frac{\varepsilon_t^2}{\upsilon}\right)^{-\frac{\upsilon+1}{2}}. \tag{6.28}$$

Rewrite (6.25), using (6.26), we have:

$$\varepsilon_t = \frac{y_t-\mu}{\left[\frac{h_{t-1}(\upsilon-2)}{\upsilon}\right]^{1/2}} \tag{6.29}$$

and

$$\left|\frac{\partial\varepsilon_t}{\partial y_t}\right| = \frac{1}{\left[\frac{h_{t-1}(\upsilon-2)}{\upsilon}\right]^{1/2}}. \tag{6.30}$$

Replace (6.28), (6.29) and (6.30) in (6.24):

$$f(y_t; F_{t-1}) = \frac{\Gamma((\upsilon+1)/2)}{\Gamma(\upsilon/2)\sqrt{\pi\upsilon}}\left(1+\frac{\left(\frac{y_t-\mu}{\left[\frac{h_{t-1}(\upsilon-2)}{\upsilon}\right]^{1/2}}\right)^2}{\upsilon}\right)^{-\frac{\upsilon+1}{2}}\left|\frac{1}{\left[\frac{h_{t-1}(\upsilon-2)}{\upsilon}\right]^{1/2}}\right| \tag{6.31}$$

$$f(y_t; F_{t-1}) = \frac{\Gamma((\upsilon+1)/2)}{\Gamma(\upsilon/2)\sqrt{\pi(\upsilon-2)}}\left(\frac{1}{h_t}\right)^{\frac{1}{2}}\left(1+\frac{(y_t-\mu)^2}{(\nu-2)h_t}\right)^{-\frac{\upsilon+1}{2}}. \tag{6.32}$$

The logarithm of (6.32):

$$lnf(y_t|Y_{t-1},\mu,h_{t-1},\upsilon) = ln\Gamma\left(\frac{\upsilon+1}{2}\right) - ln\Gamma\left(\frac{\upsilon}{2}\right) - \frac{1}{2}ln\pi - \frac{1}{2}ln(\upsilon-2) - \frac{1}{2}lnh_{t-1} - \frac{\upsilon+1}{2}ln\left(1+\frac{(y_{t-1}-\mu)^2}{(\upsilon-2)h_{t-1}}\right). \tag{6.33}$$

6.3
Changing the Measure

Under the risk-neutral probability, the discounted price process must be a martingale. The first condition that the risk-neutral probability structure must satisfy is that the discounted price process has zero drift under this structure and it must also be equivalent to the original structure (i.e., the same set of price paths must have positive probability under both structures). This section presents the key to achieving both results.

6.3.1
Girsanov Theorem

Since under continuous-time stochastic processes the events can occur over a continuous range of values, we define the physical probability space as $\{\Omega,\mathcal{F},\mathcal{P}\}$, where filtration $\mathcal{F}=\{\mathcal{F}_t, 0\leq t\leq T\}$ gives the information structure of events over a continuous interval $[0,T]$. The cumulative probability now defined as $\mathcal{P}(z_t)$, where z_t belongs to the N-dimensional continuous sample space of real number $\mathbb{R}^N=\{-\infty,\infty\}$. The differential of $\mathcal{P}(z_t)$ is given as $d\mathcal{P}(z_t)$, which can be used to obtain the time 0 probability density function of events at time t. Though $d\mathcal{P}(z_t)$ gives the time 0 probability density, we do not explicitly mention the subscript for time 0. Let a stochastic process $X(t)$ be adapted, which means that it is observable on $\mathcal{F}_t$. In other words, it is possible to deduce all possible values of $X(t)$ based on events in $\mathcal{F}_t$.

At least since the paper of Black and Scholes (1973), it has become commonplace in the study of derivative pricing to employ continuous-time models of asset price evolution. Most continuous-time models are based on Wiener process. Assume that the Ito process Y_t, with drift μ and variance σ^2:

$$dY_t = \mu dt + \sigma dW_t \tag{6.34}$$

where dW_t is Wiener process. Suppose we wish to change the drift of the (6.34) to another drift in an equivalent probability structure. We use the following procedures:

1. First, we define a new Wiener process, $\widehat{W}_t$, using the old Wiener process, W_t;
2. Next, we redefine a new process, $\widehat{Y}_t$, using the step 1.

Step 1: Define λ and $\widehat{W}_t$:

$$\mu - \eta = \sigma\lambda \tag{6.35}$$

Note that λ need not be a constant, since μ, σ and η are not required to be constants. Now define the process $\widehat{W}_t$ by:

$$d\widehat{W}_t = \lambda dt + dW_t. \tag{6.36}$$

Step 2: Now, let the process Ŷ be defined by:

$$d\widehat{Y}_t = \eta dt + \sigma d\widehat{W}_t. \tag{6.37}$$

By definition, the $\widehat{Y}_t$ process has a drift of η. A little bit of algebra shows that this process $\widehat{Y}_t$ is identical to the process Y_t defined in (6.34). Replace (6.36) in (6.37):

$$d\widehat{Y}_t = \eta dt + \sigma[\lambda dt + dW_t] = [\eta + \sigma\lambda]dt + \sigma dW_t \tag{6.38}$$

and use (6.35) in (6.38):

$$d\widehat{Y}_t = [\eta + (\mu - \eta)]dt + \sigma dW_t = dY_t. \tag{6.39}$$

Note that the variance σ^2 has not changed in going from the original process to the new process. Girsanov's Theorem ensures that the new probability structure is equivalent to the original structure only if following technical condition holds:

$$\xi(\lambda) = exp\left\{-\frac{1}{2}\int_0^t \lambda^2 ds + \int_0^t \lambda dW\right\}. \qquad (6.40)$$

and

$$E\left[exp\left\{-\frac{1}{2}\int_0^t \lambda^2 ds\right\}\right] < \infty \qquad (6.41)$$

This technical conditional ensures that the process $\xi(\lambda)$ must be a martingale under the original probability structure, and the process $\widehat{W}_t$ is a Wiener process under the new structure. Equation (6.41) is named Novikhov's Condition.

Consider the discounted price process obtained by discounting the growth in stock prices at the risk-free rate, i.e., by dividing the stock price, S_t, by the bond price, B_t:

$$Z_t = \frac{S_t}{B_t}. \qquad (6.42)$$

The bond price evolves from its initial value $B_0 = 1$ according to the ordinary differential equation:

$$dB_t = rB_t dt \qquad (6.43)$$

where r (expressed in continuously compounded terms) is constant. The stock price evolves from its initial value S_0 according to the stochastic differential equation:

$$dS_t = \mu dt + \sigma dW_t. \qquad (6.44)$$

Using Ito's lemma in conjunction with (6.43) and (6.44), we can derive the instantaneous drift and variance of the Z_t process. Let μ denote this drift and σ^2 the variance:

$$dZ_t = \mu dt + \sigma dW_t. \tag{6.45}$$

To find a risk-neutral probability, we have to find an equivalent probability structure in which Z_t has zero drift. That is, $\eta = 0$ in (6.35) and λ is given by:

$$\mu = \lambda\sigma. \tag{6.46}$$

Girsanov's Theorem ensures that the Z_t process may be rewritten using Ŵ as a process with zero drift:

$$dZ_t = \sigma d\widehat{W_t} \tag{6.47}$$

Since the new probability structure is equivalent to the original structure, and since Z_t is a martingale under the new structure.

6.4
Properties of First-Order Model

If $E\left[z_t^j\right] < \infty$, a necessary and sufficient condition for the existence of the $j - th$ moment of y_t is:

$$E[\phi + \alpha u_{t-1}]^{j/2} < 1, \qquad j = 2,4, \dots \tag{6.48}$$

Proof. The model is a member of the class of models defined by He and Terasvirta (1999) in which h_t, $h_t = \sigma_t^2$, is given by:

$$\tilde{\sigma}_t^d = a(z_{t-1}) + c(z_{t-1})\tilde{\sigma}_{t-1}^d. \tag{6.49}$$

He and Teräsvirta (1999) stated the condition for the existence of moments in terms of $c(z_t)$, that is, $E[c(z_t)]^{j/2} < 1, j = 2,4, \ldots$. In (3.30) $d = 2$, $a(z_t) = \delta$ and $c(z_t) = \phi + \alpha u_t$ ■

From He et al (1999), y_t strictly stationary and ergodic if $E[\phi + \alpha u_t] = \phi < 1$. Furthermore, y_t is second-order stationary if $\upsilon > 2$. For $j = 4$, the condition is:

$$\phi^2 + \alpha^2 E[u_t^2] = \phi^2 + \alpha^2 \frac{2\upsilon}{\upsilon + 3} < 1, \qquad \upsilon > 4 \tag{6.50}$$

or, if we write $c(z_t) = \beta + \alpha(\upsilon + 1)b_t$:

$$\beta^2 + 2\alpha\beta + \frac{3\alpha^2(\upsilon + 1)}{\upsilon + 3} < 1, \qquad \upsilon > 4. \tag{6.51}$$

In the limit as $\upsilon \to \infty$ the above expression tends to the standard GARCH(1,1) condition for the existence of fourth moments. As in the standard GARCH(1,1) model, the autocorrelation function of Beta-t-GARCH (1,1) is of the form $\rho(\tau; y_t^2) = \phi^{\tau-1}\rho(1; y_t^2)$ for $\tau \geq 1$, but $\rho(1; y_t^2)$ now depends on υ as well as ϕ and α, with $0 < \phi < 1$ is:

$$\rho(\tau; y_t^2) = \frac{\phi^{\tau-1}(\phi + 2\alpha)K_\upsilon^* - \phi^\tau}{\kappa_\upsilon - 1}, \qquad \tau = 1,2, \ldots \tag{6.52}$$

where,

$$K_\upsilon^* = \frac{1 - \phi^2}{1 - \phi^2 - \alpha^2\, 2\upsilon/(\upsilon + 3)} \tag{6.53}$$

$$\kappa_\upsilon = \frac{3(\upsilon - 2)}{(\upsilon - 4)}, \qquad \upsilon > 4. \tag{6.54}$$

The Beta-t-GARCH model may be extended to include leverage effects by adding the indicator variable $I(y_{t-1} < 0)(\upsilon + 1)b_{t-1}$. When (3.28), or (3.29), is modified in this way, the model is still a special case of (6.49) with:

$$c(z_t) = \beta + \alpha(\upsilon + 1)b_{t-1} + \alpha^* I(y_{t-1} < 0)(\upsilon + 1)b_{t-1}$$
$$= \beta + \{\alpha + \alpha^* I(y_{t-1} < 0)\}(\upsilon + 1)b_{t-1}. \quad (6.55)$$

Hence the condition for the existence of the second moment, assuming $\upsilon > 2$, is now $\alpha + \beta + \alpha^*/2 < 1$, whereas the fourth moment exists if:

$$\beta^2 + 2\alpha\beta + \frac{3(\alpha^2 + \alpha\alpha^* + \alpha^{*2}/2)(\upsilon + 1)}{\upsilon + 3} < 1, \qquad \upsilon > 4 \quad (6.56)$$

The properties of the Gamma-GED-GARCH model can be derived in a similar way.

6.5 Breeden and Litzenberger Method with Integral

According to Leibnitz rule, if we can derivative$f(x,\theta)$, $a(\theta)$ and $b(\theta)$ with relation at θ, we have:

$$\frac{d}{d\theta}\int_{a(\theta)}^{b(\theta)} f(x,\theta)dx = f(b(\theta),\theta)\frac{db(\theta)}{d\theta} - f(a(\theta),\theta)\frac{da(\theta)}{d\theta} + \int_{a(\theta)}^{b(\theta)} \frac{d}{d\theta} f(x,\theta)dx \quad (6.57)$$

where $-\infty < a(\theta)$ and $b(\theta) < \infty$.

From the relationship proposed by Cox and Ross (1976) to option pricing:

$$C(K,T) = e^{-rT}\int_{-\infty}^{\infty}(S_T - K)^+ q(S_T)dS_T. \quad (6.58)$$

If we rewrite (6.58) using (6.57), then:

$$\theta = K, a(K) = K, b(K) \to \infty \; and \; f(x,\theta) = (S_T - K)q(S_T). \tag{6.59}$$

Consider the first term in (6.57):

$$f(b(K),\theta)\frac{\partial b(\mathrm{K})}{\partial K} = (b(K) - K)f(K) = 0 \tag{6.60}$$

because $\partial b(\mathrm{K})/\partial K = 0$. Let the second term in (6.57):

$$-f(a(K),\theta)\frac{\partial a(K)}{\partial K} = -(a(K) - K)f(K) = 0 \tag{6.61}$$

because $a(K) = K$. Finally, the third term in (6.57):

$$\int_{a(K)}^{b(K)} \frac{\partial}{\partial K} f(x,K)dx = \int_{K}^{\infty} \frac{\partial}{\partial K}(S_T - K)q(S_T)\, dS_T = -\int_{K}^{\infty} q(S_T)\, dS_T. \tag{6.62}$$

Then, take the partial derivative of C with respect to K, expression (6.58), to get:

$$\frac{\partial C(K,T)}{\partial K} = -e^{-rT}\int_{K}^{\infty} q(S_T)\, dS_T. \tag{6.63}$$

But, we know that:

$$\int_{-\infty}^{\infty} q(S_T)dS_T = \int_{-\infty}^{K} q(S_T)dS_T + \int_{K}^{\infty} q(S_T)dS_T \tag{6.64}$$

then:

$$\int_{K}^{\infty} q(S_T)dS_T = 1 - \int_{-\infty}^{K} q(S_T)dS_T = 1 - Q(S_T). \quad (6.65)$$

Rewrite (6.65) in (6.63):

$$Q(S_T) = 1 + e^{rT}\frac{\partial C(T,K)}{\partial K}|_{K=S_T} \quad (6.66)$$

which yields the cumulative distribution function denoted by $Q(S_T)$. The probability distribution function $q(S_T)$ can be obtained by taking the derivative of (6.66):

$$q(S_T) = e^{rT}\frac{\partial^2 C(T,K)}{\partial K^2}|_{K=S_T}. \quad (6.67)$$

6.6
Boundary Conditions of Option Prices

In order to avoid arbitrage opportunities, there are several restrictions on the price of an option, which are discussed in the appendix. This section discusses boundary conditions of calls only.[29]

In the first restriction (4.14), the value of a call option is never greater than the asset price and never less than its intrinsic value:

$$S_0 \geq C(K,T) \geq max(S_0 - e^{-rT}K; 0). \quad (6.68)$$

Moreover, the value of a call option takes the value of the asset at a strike price of zero and converges to a value of zero for very large strike prices:

$$C(0,T) = S_T, \qquad \lim_{K\to\infty} C(K,T) = 0. \quad (6.69)$$

[29] Restrictions on puts can be derived in a similar way.

The price of a call cannot be negative $C(K,T) \geq 0$, as it needs not be exercised, because its intrinsic value is $max(S_T - e^{-rT}K; 0)$. If call has a negative price, $C(K,T) < 0$, then a riskless profit could be made by buying the call (receiving an instant positive profit equal to the value of the call) and holding it until expiration to make a nonnegative income equal to the value of the call at expiration. The price also cannot exceed the stock price $C(K,T) \leq S_0$, since ending up owning the stock is the best that can happen to the option holder. If $C(K,T) > S_0$, buying the stock and selling the call would create an instant profit of $C(K,T) - S_0$ and generate a nonnegative amount at expiration equal to $S_T - max(S_T - K; 0)$.

Now, consider the following portfolios: portfolio #1 consist of one call, $C(K,T)$, and Ke^{-rT} zero-bonds yielding the risk-free rate and portfolio #2 consist of one stock of value S_0. These portfolios are compared in table 6.1.

Table 6.1: Portfolio Comparison.

Portfolio	Value at $t = 0$	Value at expiration $t = T$	
		$S_T < K$	$S_T \geq K$
#1	$C(K,T) + Ke^{-rT}$	$0 + K$	$(S_T - K) + K$
#2	S_0	S_T	S_T
Portfolio comparison at expiration		#1 > #2	#1 = #2

At expiration, portfolio #1 has the same or a higher value than portfolio #2. Hence, an arbitrage-free call has to satisfy:

$$C(K,T) + Ke^{-rT} \geq S_0 \rightarrow C(K,T) \geq max(0, S_0 - Ke^{-rT}). \quad (6.70)$$

If $C(K,T) \leq S_0 - e^{-rT}K$, once again we can make an instant profit by buying the call and selling the portfolio $S_0 - e^{-rT}K$. Then at expiration date we receive a non-negative payoff equal to $max(0, K - S_T)$.

In the second restriction (4.15), the value of a vertical call spread is nonpositive or the call price function is monotonically decreasing:

$$-e^{-rT} \leq \frac{\partial C(K,T)}{\partial K} \leq 0. \quad (6.71)$$

The monotonicity of the call option prices establishes that the following inequality has to be satisfied:

$$(K_1 - K_2)e^{-rT} \geq C(K_2, T) - C(K_1, T), with\, K_1 > K_2. \qquad (6.72)$$

In case inequality above is not fulfilled, buy $C(K_1)$ short sell $C(K_2)$. This yields at least $(K_1 - K_2)e^{-rT}$, which is invested at the risk-free rate. This portfolio has a nonnegative value at expiration. Moreover, it requires no net investment, and in case of strict inequality even yields a positive amount of money at $t = 0$. A detailed analysis is shown in table 6.2.

Table 6.2: Arbitrage Portfolio.

Value at $t = 0$	Value at expiration $t = T$		
	$S_T < K_2 < K_1$	$K_2 \leq S_T < K_1$	$S_T \geq K_1 > K_2$
$-C(K_1, T)$	0	0	$S_T - K_1$
$+C(K_2, T)$	0	$-(S_T - K_2)$	$-(S_T - K_2)$
$-(K_1 - K_2)e^{-rT}$	$+(K_1 - K_2)$	$+(K_1 - K_2)$	$+(K_1 - K_2)'$
≥ 0	$+(K_1 - K_2) > 0$	$+(K_1 - S_T) > 0$	0

Finally, the third restriction (4.16), the value of a butterfly spread is nonnegative or the call option price function is convex. The risk-neutral distribution is given as a function of the first derivative of the call option pricing formula:

$$Q(S_T) = 1 + e^{rT} \frac{\partial C(T, K)}{\partial K} |_{K=S_T}. \qquad (6.73)$$

If C is twice differentiable, the risk-neutral density function is given by the following expression:

$$dQ(S_T) = e^{rT} \frac{\partial^2 C(T, K)}{\partial K^2} |_{K=S_T}. \qquad (6.74)$$

A direct implication of the convexity assumption is:

$$\int_0^{\infty} dQ(K = S_T)dK = \int_0^{\infty} e^{rT}\frac{\partial^2 C(T,K)}{\partial K^2} \\ = e^{rT}\left[\lim_{K\to\infty}\frac{\partial C(T,K)}{\partial K} - \lim_{K\to 0}\frac{\partial C(T,K)}{\partial K}\right] = 1. \tag{6.75}$$

Note that these assumptions are sufficient to ensure that the mean of the distribution is the forward price since:

$$E_Q = \int_0^{\infty} S_T dQ(S_T) = -\frac{1}{e^{-rT}}\int_0^{\infty}\frac{\partial C(T,K)}{\partial K} dS_T = \frac{1}{e^{-rT}}S, \tag{6.76}$$

i.e., the underlying forward price.