Charles University in Prague

Faculty of Social Sciences

Diploma Thesis

2010

Viktor Chrobok

Charles University in Prague

Faculty of Social Sciences

Institute of Economic Studies

DIPLOMA THESIS

Option Pricing Methods

Author: Viktor Chrobok
Consultant: PhDr. Petr Gapko

Academic Year: 2009/2010

Declaration of Authorship

The author hereby declares that he compiled this thesis independently, using only the listed resources and literature.

The author grants to the Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, May 21, 2010 Viktor Chrobok

Acknowledgements:

I would like to thank PhDr. Petr Gapko for his supervision and the valuable comments about the topic. I also want to express my thanks to Doc. Ing. Oldřich Dědek, CSc. for introducing me to the theory of financial derivatives and recommending a suitable consultant for this thesis.

Abstrakt:

Diplomová práce je zaměřena na metody oceňování opcí. Jsou zde popsány základní charakteristiky a druhy opčních kontraktů. Poté je uveden popis 6 oceňovacích modelů - Black-Scholes model, French Black-Scholes model, binomický model, model kvadratické aproximace, Bjerksund-Stersland model a Jump-Diffusion model. Empirická část obsahuje zhodnocení jednotlivých modelů při aplikaci na reálná data. Bylo ukázáno, že všechny modely kromě Jump-Diffusion modelu poskytují velmi přesné výsledky, avšak bylo nemožné rozhodnout, který je nejlepší. Evidence ukazuje, že je vhodnější použít implikovanou volatilitu namísto historické. Navíc bylo prokázáno, že modely pro oceňování evropských opcí jsou vhodné i pro ocenění amerických opcí.

Abstract:

The diploma thesis is focused on the option pricing methods. There are described basic features of the option contracts and the types of them. Then a description of 6 pricing methods is given – the Black-Scholes model, the French Black-Scholes model, the Binomial Model, the Quadratic approximation model, the Bjerksund-Stersland model and the Jump-Diffusion model. The empirical part contains an analysis of the performance of all models on the real market data. It was shown that all models except for the Jump-Diffusion one fit the data very well, yet it was impossible to determine the best one. The evidence suggests that it is better to plug a few-days-delayed implied volatility than the historical one into all of the models. It was observed that the models for pricing European options are suitable even for the American ones.

Klíčová slova:

Opce, Black-Scholes model, French Black-Scholes model, binomický model, model kvadratické aproximace, Bjerksund-Stersland model, Jump-Diffusion model

Keywords:

Option, Black-Scholes model, French Black-Scholes model, Binomial Model, Quadratic approximation model, Bjerksund-Stersland model, Jump-Diffusion model

Content

Teze diplomové práce

Diplomant: Ing. Viktor Chrobok

Vedoucí: PhDr. Petr Gapko

Předpokládaný název práce: Metody oceňování opcí

Stručná charakteristika tématu

Práce se bude zabývat oceňováním opčních kontraktů a srovnáním jednotlivých způsobů ocenění. Základním modelem oceňování je Black-Scholes model, který předpokládá konstantní volatilitu podkladového aktiva. Sofistikovanějším modelem, uvažujícím volatilitu popsanou stochastickým Cox-Ingersoll- Ross procesem, je Hestonův model. Pro ocenění nestandardních typů opcí (asijské, exotické, knock-out) se využívají simulační techniky. Další možností pro zpřesnění metod ocenění je použití jiného než normálního rozdělení v oceňovacích modelech (např. hyperbolické, stabilní rozdělení).

Hypotézy

Cílem práce by mělo být porovnání přesnosti různých metod za pomoci ekonometrických simulačních technik ex post. Práce by měla poskytnout odpověď na otázku, kdy je ještě možné použít relativně výpočetně nenáročný Black-Scholes model a kdy je výhodnější přistoupit k modelům se stochastickou volatilitou, které používají metod diferenciálního počtu. Atraktivním předmětem zkoumání může být také vliv aktuálně probíhající finanční krize na způsob oceňování a hledání důvodu neschopnosti exaktních modelů předpovědět globální recesi.

Metody práce

Pro popis oceňovacích modelů je vhodné použít metody matematické analýzy a statistiky, případně simulačních technik. Srovnání jednotlivých metod lze provést pomocí ekonometrie, stejně tak, jako určení vlivu finanční krize na metody oceňování.

Předpokládaná osnova práce

1. Úvod

2. Oceňování pomocí Black-Scholes modelu

3. Modely se stochastickou volatilitou

4. Použití simulačních technik

5. Empirické srovnání jednotlivých metod

6. Vliv finanční krize na oceňování opcí

7. Závěr

Základní literatura

• Scholes, Myron; Black, Fischer: *The Pricing of Options and Corporate Liabilities*, Journal of Political Economy, Vol. 81, No. 3, 1973

• Merton, Robert C.: *Theory of Rational Option Pricing*, The Bell Journal of Economics and Management Science, Vol. 4, No. 1, 1973

• Heston, Steven, L.: A *Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options*, The Reveiw of Financial Studies, Vol. 6, No. 2

• Gibson, Rajna: *Option valuation: analyzing and pricing standardized option contracts*, New York, McGraw-Hill, 1991 Blake David: *Financial Market Analysis*, New York, McGraw-Hill, 1990

• Fouque, Jean-Pierre: *Derivatives in financial markets with stochastic volatility*, Cambridge, Cambridge University Press, 2000

• Zaffaroni, Paolo: *Nonlinear time series with long memory: a model for stochastic volatility*, London, London School of Economics and Political Science, 1997

• Gujarati, Damodar N.: *Basic econometrics*, Boston, McGraw-Hill, 2003

• Dlouhý, Martin a kol.: *Simulace podnikových procesů*, Brno: Computer Press, 2007

• Turnbull, Stuart M.: *Option valuation*, Holt, Rinehart and Winston of Canada, 1987

Diplomant | Vedoucí

1. Introduction

An option is a security giving its owner the right but not the obligation to sell or buy an underlying asset at a fixed price at or before specific date. The possibility to decide whether buy/sell the underlying asset is definitely an advantage for an owner and a disadvantage for the seller that is why options are sold for a so-called option premium. Determining this premium is crucial since the price of the option is the most important factor for an investor deciding whether buy the option or not. Essentially the investors seek options with high profit potential and issuers try to sell low profitable options for high prices.

The most important factor for determining the option premium is expected development of the underlying asset. Since the future development of the underlying asset's price is uncertain, it is needed to build some prediction model to determine it. There are various types of options providing different features enhancing or restricting pay-off possibilities. These modifications are naturally affecting the option prices as well.

There could be used several methods for determining the option prices. We shall try to describe some of them and evaluate their ability to determine the price precisely in this thesis. The second chapter will provide basic characteristics of options and shortly discuss the basic types of those derivatives.

In the third chapter six option pricing models are described. Those models are the Binomial model, the Black-Scholes model, the French Black-Scholes model the Quadratic Approximation model (BAW), the Bjerksund-Stersland (BJST) model and the Jump Diffusion model. The first two models are used to price European options; the binomial model is suitable for European and also American options. The BAW and BJST models were developed especially to price American options. The Jump Diffusion model in its basic form is suitable only for European options.

All six models are tested on the real dataset from Chicago Board of Options Exchange. This is the content of the fourth chapter. Firstly the whole dataset is analyzed to obtain basic summary statistics of the data. The only model which is applied to the whole dataset is the Black-Scholes one due to its computational simplicity and the huge amount of the available data. The other models are tested just on the several subsets containing 1000 observations.

The core of the fourth chapter is to find notifiable facts in various subsets. These subsets are chosen according to the traded volume of the options, time to expiration and the relationship between the underlying spot and the strike price. The second objective is to evaluate the pricing methods and their performance on the real data.

The available dataset contains just American options, but there are implemented methods pricing those options as the European ones. So the next question is whether it is possible to use those methods in this situation.

It is widely known that there are two types of prices at each market: bid price and ask price. We shall determine which one is more suitable when dealing with the option pricing models. The next possible question is whether it is better to use historical or implied volatility in the option pricing models to obtain usable results.

1.1 Main Hypotheses, Objectives and Methods

This thesis determines pros and cons of the described option pricing techniques. Not only does it mean deciding which one works more often, but also finding conditions under which each method work. The quality of the pricing model depends not only on accuracy, but also on the volatility of the differences between the real and estimated price and computational complexity of the pricing algorithm.

The pricing methods are described using mathematical techniques. Accuracy and volatility can be calculated and presented by the methods of descriptive and summary statistics. Those methods demand intermediate knowledge of differential equations calculus, probability theory, statistics and econometric analysis.

Currently most discussed research areas in economics are connected to the financial crisis which started in 2007. Many investors in call options suffered significant losses due to the current crisis and on the other hand investors in put options earned huge profits. We shall find out if any of the described option pricing techniques was affected by the crisis so badly that it is impossible to use it in its traditional state anymore.

2. Basic Characteristics of Options

An option is a security that belongs to a group of financial instruments called derivatives. This means that its price is derived from the price of something else (e.g. equities, bonds, commodities etc.) and it is a timed contract. An option gives its holder the right but not the obligation to purchase or sell specified amount of an asset at a certain price, which is usually higher than the asset's price at the time of issue.

Main characteristics of any option are:

- premium = option's price,
- amount of asset bought/sold in the future,
- price of asset bought/sold in the future,
- strike price,
- expiration date.

There are three types of options according to possibilities of exercising them:

European option – this option can be exercised only at the expiration date.

American option – this option can be exercised anytime before the expiration date.

Bermudan option – this option can be exercised only on specified dates on or before the expiration date.

Price of option is determined by two factors:

- **Intrinsic value –** this is the difference between strike price and underlying stock price. Compounding of intrinsic value is different for put and call options and will be explained shortly. Intrinsic value obviously cannot be negative. If this value is positive, we say that option is in the money. If the intrinsic value is zero, we say that option is out of the money[1].
- **Time value –** reflects possible fluctuations till the expiration date. It declines as the expiry of the option gets closer. Time value is zero at the expiration date.

2.1 Long Position in the Option Contract

The position when one is buying an option is called a long position. The following provides a closer look at the position of an investor at the expiration day. Cost of option

[1] In the described situation an option could also be at the money – to be explained shortly.

is C, strike price is K and spot price of the underlying is S. Situation is different for call and put options, both of them are shown in the figures 1 and 2.

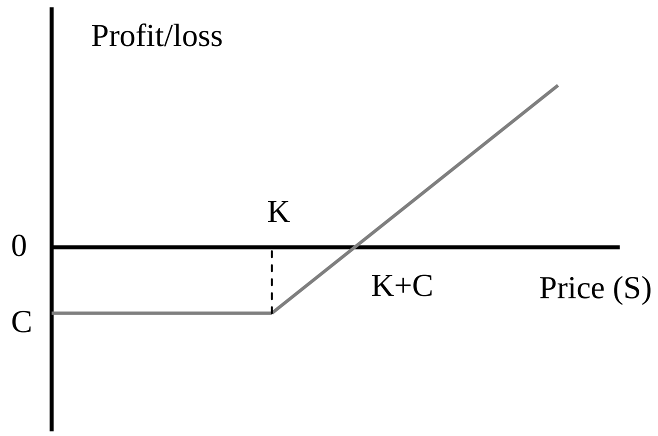

Fig. 1: *Position of an investor buying a call option at the expiration date (long call)*

We can see that if the price of bought asset is lower than the strike price S an investor will not exercise the option and bear a loss C. If the price S is higher than K but lower than K+C, investor will exercise the option, but still bear a loss S-K-C. Finally, if the price S is higher than K+C, the investor will exercise the option and gain a profit S-K-C.

Point K+C is called a break-even. When the price S is lower than strike price K, an option is out of the money. When the underlying price S equals the strike price K, the option is at the money. When the strike price K is higher than price S, the option is in the money.

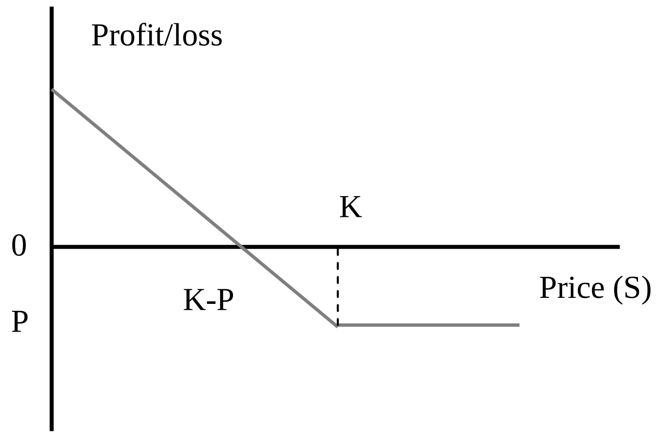

Fig. 2: *Position of an investor buying a put option at the expiration date (long put)*

We can see that if the price of an underlying asset S is lower than K-P an investor will exercise the option and gain a profit S-K-P. If the price S is higher than K-P but lower than K, the investor will exercise the option, but bear a loss S-K-P. Finally, if the price S is higher than K, the investor will not exercise the option and bear a loss P.

Point K is called the break-even. When the price S is lower than the strike price K, the option is in the money. When the underlying price S is the same as the strike price K, the option is at the money. When the strike price K is higher than the price S, the option is out of the money.

2.2 Short Position in the Option Contract

The seller's position in the option contract will be described now. This position is called a short position. We shall also differ between a put and a call option. We shall firstly plot the call option as in the previous text.

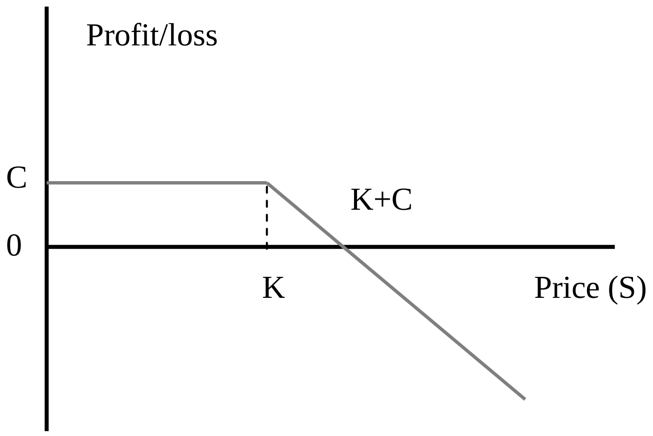

Fig. 3: *Position of an investor selling a call option at the expiration date (short call)*

One can see that if the price of sold asset is lower than the strike price S the option will not be exercised and a seller will obtain a profit C. If the price S is higher than K but lower than K+C, the option will be exercised, but a seller will still have a profit C-S+K. Finally, if the price S is higher than K+C, the option will be exercised and the seller will bear a loss C-S+K. Break-even point is the same as for the long position.

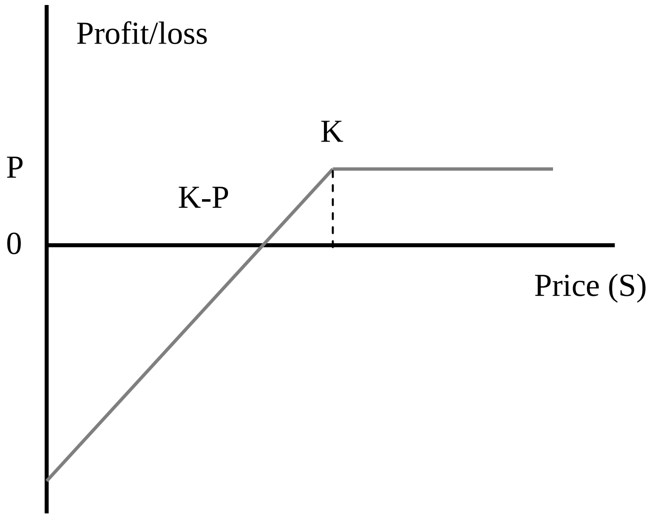

Fig. 4: *Position of an investor selling a put option at the expiration date (short put)*

If the price of an underlying asset S is lower than K-P an option will be exercised and the seller suffers a loss P-S+K. If the price S is higher than K-P but lower than K, the option will be exercised and the seller will gain a profit P-S+K. And at last, if the price S is higher than K, the seller will gain a profit P.

2.3 Put-Call Parity

Put-call parity is defined as a relationship between put option and call option with same expiration date and strike price. Derivation of this relationship assumes that options are European type. The relationship can be quantified[2] as:

$$C(t) + K * B(t, T) = P(t) + S(t),$$

Where C(t) is the value of a call option at time t, K is the strike price, B(t, T) is the value of bond that matures at time T, P(t) is the value of put option at time t and S(t) is the price of underlying at time t.

This identity implies that put and call options with combination of buying/selling underlying can be used as perfect substitutes in any portfolio neutral to the price of an underlying asset[3]. It means that buying a call and selling a certain amount[4] of the underlying asset has exactly the same effect as buying a put and buying a certain amount[5] of the underlying asset.

Contravention of put-call parity leads to opportunities of arbitrage. When assuming effective markets with rational players, the arbitrage will move prices of call and put options to the equilibrium given by the put-call parity equation.

2.4 Greeks

There are several option characteristics determining the sensitivity of options to the change of underlying parameters. These characteristics make the comparison of different options easier and some of them even figure in the option pricing formulas. Since various Greek letters are used for those characteristics, they are called "Greeks"

---

[2] More in Hull 2009

[3] This kind of portfolio is called delta neutral. Option's delta is a sensitivity of option's price for movements of underlying asset; it is obviously positive for call options and negative for put options. Delta neutral portfolio could be portfolio consisting of same numbers of call and put options with the same strike price and expiration date.

[4] If one call is bought then this amount equals call's delta d.

[5] This amount is in our case 1-d.

The most common sensitivity indicators are[6]:

i. **Option delta (Δ)** – sensitivity of the option premium to the price change of the underlying asset[7]. It is calculated as $\Delta = \frac{\partial C}{\partial S}$, where C is option premium and S is the price of the underlying asset. The value of option delta is always between 0 and 1. This is especially useful in hedging, when a so-called delta neutral portfolio is build. This portfolio contains a combination of options and underlying assets so that its delta is zero. It means that the value of the portfolio does not change with small price changes of the underlying asset.

ii. **Option gamma (Γ)** – sensitivity of the option delta to the change of underlying asset[8]. The calculation formula is $\Gamma = \frac{\partial^2 C}{\partial S^2}$. A gamma neutral portfolio is such portfolio, which has overall gamma equal to zero. This portfolio does not change its value even with large price changes.

iii. **Option speed** – sensitivity of the option to the price change of underlying asset. The calculation formula is $speed = \frac{\partial^3 C}{\partial S^3}$. Knowledge of the option speed is useful when determining whether to use gamma or delta hedging.

iv. **Option theta (Θ)** – sensitivity of the option premium to the change in the time remaining to maturity of the option. The calculation formula is $\Theta = \frac{\partial C}{\partial T}$, where T is time remaining to maturity.

v. **Option vega (ν)** – sensitivity of the option premium to the change in volatility of the underlying asset. It is calculated as $\nu = \frac{\partial C}{\partial \sigma^2}$, where σ^2 is a volatility of the underlying asset. A vega-hedged portfolio is a portfolio containing such assets so that their vega is zero. It means that the value of the portfolio does not change with the change in volatility.

vi. **Option rho (ρ)** – sensitivity of the option premium to the change in market interest rate. It is calculated as $\rho = \frac{\partial C}{\partial r}$, where r is a market interest rate.

vii. **Option vanna** – is the sensitivity of the option delta with respect to the option volatility. It can be calculated as $vanna = \frac{\partial \Delta}{\partial \sigma^2} = \frac{\partial \nu}{\partial S} = \frac{\partial^3 C}{\partial \sigma^2 \partial S}$. This Greek can be

[6] based on Dědek - lecturer's notes for the course "Financial Market Instruments" at IES FSV UK
[7] It is the first order approximation of change in the option price when applying the Taylor's theorem.
[8] It is the second order approximation of change in the option price when applying the Taylor's theorem.

useful when determining a reaction of delta-hedge to volatility changes or a vega-hedge to underlying price changes.

viii. **Option vomma** – sensitivity of the option vega to volatility changes. It is sometimes referred as vega convexity. It is in fact a second order sensitivity to volatility. It is calculated as $vomma = \frac{\partial v}{\partial \sigma^2} = \frac{\partial^2 C}{\partial \sigma^2 \partial \sigma^2}$.

ix. **Option zomma** – sensitivity of the option gamma with respect to its volatility. The calculation formula is $zomma = \frac{\partial \Gamma}{\partial \sigma^2} = \frac{\partial^3 C}{\partial S^2 \partial \sigma^2}$.

x. **Option charm** – sensitivity of the option delta to time. The calculation formula is $charm = \frac{\partial \Delta}{\partial T} = \frac{\partial \Theta}{\partial S} = \frac{\partial^2 C}{\partial T \partial S}$.

xi. **Option color** – sensitivity of the option gamma with respect to time. It can be calculated as $color = \frac{\partial \Gamma}{\partial T} = \frac{\partial^2 \Theta}{\partial S^2} = \frac{\partial^3 C}{\partial T \partial S^2}$.

The following table summarizes all Greeks and their relationship to four basic characteristics of options used in Black-Scholes model[9], which are spot price of the underlying security, time to expiration, volatility and risk-free investment.

	Spot price	**Volatility**	**Time to expiration**	**Risk free interest rate**
Option premium	Delta	Vega	Theta	Rho
Delta	Gamma	Vanna	Charm	
Gamma	Speed	Zomma	Color	

Tab. 1: *Greeks*

2.5 Exotic Options

All types of options mentioned above are called vanilla options. There is also another category of options called exotic options.

Exotic options could be divided into three groups[10]:

- **Path-dependent options** – these options' payout reflects a particular path taken by the option. This group could be divided to the following subgroups:
  - **Asian options –** their strike price or underlying price are replaced by arithmetic or geometrical average.

[9] To be defined shortly.
[10] based on Dědek - lecturer's notes for the course "Financial Market Instruments" at IES FSV UK

- **Barrier options –** these options are valid only if some lower or upper barrier is reached or not reached.
- **Look-back options** – the payoff of these ones depends on maximum/minimum asset price reached during the life of the option.

- **Correlation options –** the value of those options depends on the price of two or more assets. There are also several types of these options:
  - **Basket option –** the payout depends on the cumulative performance of a portfolio of securities.
  - **Spread option** – exercise price of this option is calculated as a difference between two securities.
  - **Dual-strike option –** this option has two underlying assets and pays the higher intrinsic value.
- **Hybrid options –** these options are simply any other options with non-conventional features. Some of these options are:
  - **Compound option –** this option is simply an option on option.
  - **Package option –** this is a combination of an option and a forward contract.
  - **Swaption –** this is an option on a swap.

3. Pricing Models

This chapter provides a theoretical descriptions of all pricing models used in the empirical part.

3.1 Binomial model

The binomial model was firstly proposed by Cox, Ross and Rubinstein in 1979[11]. This model is based on the assumption that the price of the underlying asset evolves discretely in time. It assumes that the option is European[12] and that the underlying asset does not pay any dividends and its price follows random walk. The discreetness of the time is the most significant shortcoming of the model, which makes the binomial based price determining usable only theoretically. However this model is important for deriving more sophisticated methods.

3.1.1 Derivation of the Model

The price of share has given probabilities and magnitudes for up and down jumps. We will demonstrate the method by the following figure:

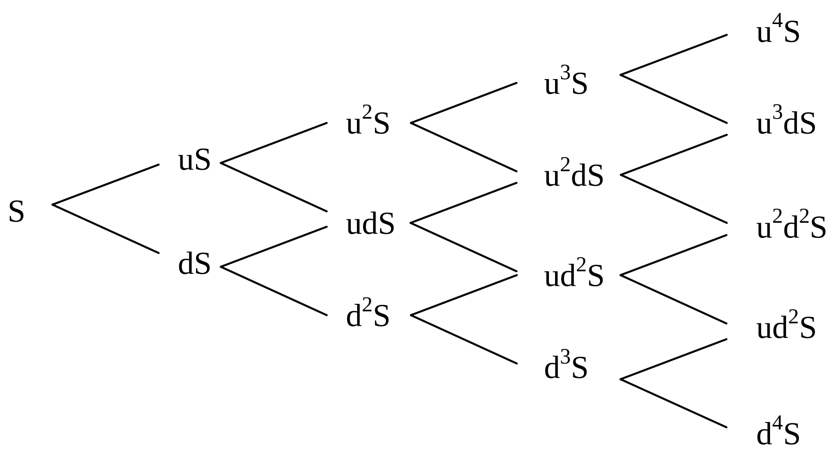

Fig.5: *Binomial scheme – underlying price*

[11] More in Cox, Ross and Rubinstein, 1979.

[12] An extension of the American option is also possible. The idea is that the value at the each node is calculated as the maximum of the exercise value and the binomial value. For more information see for example Gibson 1991.

The price of an underlining asset can rise (i.e. be multiplied by “u”) with probability q or fall (i.e. be multiplied by “d”) with probability 1-q. One can build a similar scheme for corresponding option price:

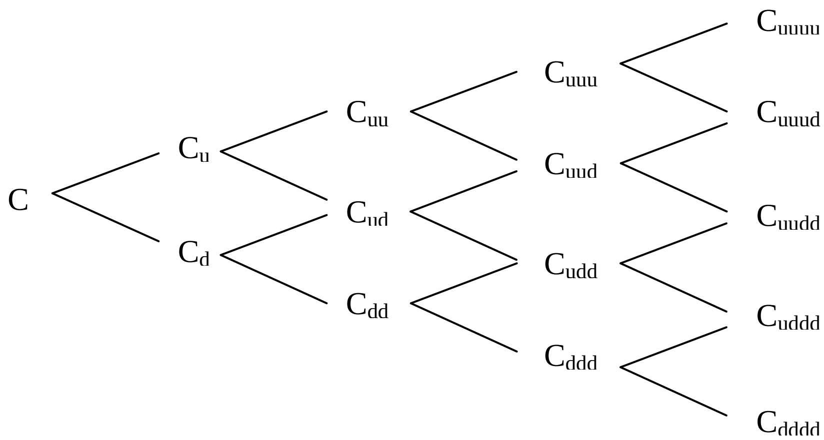

Fig. 6: *Binomial scheme – option price*

The C_u in the Fig. 6 is calculated as [0, uS - X], C_{uu}= [0, u^2S - X] etc. The binomial pricing method is based on the fact that the value of the option is known at the time of expiration. If one knows this, it is possible to go backwards and determine the option value one period sooner.

We shall use a delta neutral portfolio consisting of a long position in one underlying asset and a short one in 1/Δ call options on the very same asset to determine the option value in the sooner time period. The value of this portfolio does not change with the change of underlying price since it is compensated by the change of the option value.

The value of such portfolio if the price of the underlying price rises is

$$uS - \frac{C_u}{\Delta}. \qquad \text{(EQ 3.1.1)}$$

The portfolio value if the price of the underlying asset falls is

$$dS - \frac{C_d}{\Delta}. \qquad \text{(EQ 3.1.2)}$$

Since the portfolio is delta neutral one which does not reflect price changes EQ 3.1.1 must equal EQ 3.2.1.2. Furthermore take into account risk free interest rate (r) to give

$$S - hC = \frac{uS - \frac{C_u}{\Delta}}{1+r} = \frac{dS - \frac{C_d}{\Delta}}{1+r}, \qquad \text{(EQ 3.1.3)}$$

where C is a price of the option in the previous time period. One can get the price of the call option from EQ 3.2.1.3, which is

$$C = \frac{\frac{(1+r)-d}{u-d} C_u + \left[1 - \frac{(1+r)-d}{u-d}\right] C_d}{1+r}. \qquad \text{(EQ 3.1.4)}$$

Let us denote

$$\frac{(1+r)-d}{u-d} = q,$$

To give

$$C = \frac{qC_u + (1-p)C_d}{1+r}.$$

One can continue going backwards to get the general formula given by

$$C = \frac{\sum_{j=0}^{n} \frac{n!}{(n-j)! * j!} p^j (1-q)^{n-j} * max\{0, u^j d^{n-j} S - X\}}{(1-r)^n}.$$

3.1.2 Usage of the Model

As was stated above the binomial model is based on some unreal assumptions. The discreteness of time is undoubtedly one of them, but on the other hand the binomial model with discrete time could be a powerful tool if dealing with some exotic options, where other continuous models would be extremely complicated. The other unrealistic assumption is that all agents are perfectly informed and that they cannot influence the underlying price.

Pricing European options without any dividends of an underlying asset with binomial model is currently not used since there are more sophisticated techniques. However this model is the key one for simulation based option pricing techniques. Because of its relative simplicity the model could be easily implemented to various software even a spreadsheet editor such as MS Excel.

The binomial model is usually used to determine prices of American type options, options on assets paying dividends, complicated options such as Asian options or so called real options. The binomial model is a basic tool for Monte Carlo simulations in finance[13]. When the time steps are very small the binomial model converges to the Black-Scholes one[14].

3.2 The Black-Scholes Model

This is the well-known and widely used model for determining option prices. It was firstly introduced by Myron Scholes and Fischer Black in 1973[15]. The derivation of the model is based on partial differential equations, but the formula is quite easy to use and it became the basic tool at option markets all over the world.

3.2.1 Derivation of the Model

The model is based on the following assumptions:

- the underlying price follows a geometric Brownian motion with constant volatility and drift;
- the risk-free interest rate is known and constant in time;
- all securities are perfectly divisible;
- there are no transaction costs, neither arbitrage opportunities;
- options and underlying assets are traded on the ideal markets.

A geometric Brownian motion followed by the underlying price could be described by the Wiener process given by the following formula:

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \qquad \text{(EQ 3.2.1)}$$

where W is Brownian[16], σ is the volatility of the underlying asset and μ is a drift rate of S. The price of a call option at maturity is known. To determine the earlier price Ito's lemma[17] is used to get

$$dC = \left(\mu S \frac{\partial C}{\partial S} + \frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}\right) dt + \sigma S \frac{\partial C}{\partial S} dW. \qquad \text{(EQ 3.2.2)}$$

[13] See Dlouhý 2007.
[14] Find more in Nekula 2004.
[15] More in the original paper Scholes and Black 1973.
[16] i.e. it is a stochastic process with zero expected value and independent increments.
[17] More in Ito 1951

Now again construct a delta-neutral portfolio given by

$$\Pi = C - S\frac{\partial C}{\partial S}, \qquad \text{(EQ 3.2.3)}$$

where Π is the profit from such portfolio. Now introduce R as the accumulated profit or loss from a portfolio given by EQ 3.2.3. So the profit obtained during a time period [t, t + dt] is

$$dR = dC - dS\frac{\partial C}{\partial S}. \qquad \text{(EQ 3.2.4)}$$

Now substitute dC from EQ 3.2.2 to EQ 3.2.4 to get

$$dR = \left(\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}\right) dt. \qquad \text{(EQ 3.2.5)}$$

Since this does not contain any stochastic variable it must equal a risk free return, which means that following identity holds

$$r\Pi \mathrm{dt} = dR = \left(\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2}\right) dt. \qquad \text{(EQ 3.2.6)}$$

Now substitute from EQ 3.2.3 and divide by dt to obtain the Black-Scholes partial differential equation (PDE):

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} - rC = 0. \qquad \text{(EQ 3.2.7)}$$

Finally solution of the Black-Scholes PDE gives the Black-Scholes pricing formula:

$$C = S * N(d_1) - Xe^{-rT}N(d_2), \qquad \text{(EQ 3.2.8)}$$

where

$$d_1 = \frac{\ln\frac{S}{X} + rT}{\sigma\sqrt{T}} + \frac{1}{2}\sigma\sqrt{T} \qquad \text{(EQ 3.2.9)}$$

and

$$d_2 = d_1 - \sigma\sqrt{T}. \qquad \text{(EQ 3.2.10)}$$

The N stands for the probability distribution function of the normal distribution.

After plugging into the put-call parity one gets the following formula to obtain the price of a put option:

$$P = Xe^{-rT}N(-d_2) - S * N(-d_1). \qquad \text{(EQ 3.2.11)}$$

3.2.2 The Greeks under the Black-Scholes Model

The Greeks could be derived by using formulas provided in 2.4. When the exact prescription is known it is possible to get their values straightforward. We shall provide only values of the most important Greeks since the others could be easily derived from them.

- **Delta**

  The value of the option delta is

$$\Delta = N(d_1),$$

  for call options and

$$\Delta = -N(-d_1),$$

  for puts.

- **Gamma**

  The option gamma is same for calls and puts under Black-Scholes and is given by

$$\Gamma = \frac{N(d_1)}{S\sigma\sqrt{T-t}}.$$

- **Vega**

  The sensitivity to volatility changes for calls and puts under Black-Scholes could be calculated as

$$\nu = SN(d_1)\sqrt{T-t}.$$

- **Theta**

  The value of the option theta is

$$\Theta = -\frac{SN(d_1)}{2\sqrt{T-t}} - rXe^{-r(T-t)}N(d_2),$$

  for calls and

$$\Theta = -\frac{SN(d_1)}{2\sqrt{T-t}} + rXe^{-r(T-t)}N(-d_2),$$

for puts.

- **Rho**

The value of the call option rho is

$$\rho = (T-t)Xe^{-r(T-t)}N(d_2).$$

Rho of a put option is

$$\rho = -(T-t)Xe^{-r(T-t)}N(-d_2).$$

The other Greeks could by easily derived from the above and the formulae provided in 2.4.

3.2.3 Volatility under the Black-Scholes

The only unknown parameter in the Black-Scholes is volatility. There are two ways of estimating volatility from the data. The first one is a historical approach. It suggests that the volatility could be calculated as a standard deviation of the underlying asset's historical value. The problem is that Black-Scholes formula suggests that the volatility is constant over the time, which is generally not true. It means that the evaluation must be recalculated always when the volatility changes.

The second approach is calculation of implied volatility. The implied volatility can be calculated from the Black-Scholes formula if all other parameters including the option premium are known. If the implied volatility significantly differs from the historical volatility it means that the market expects some major change or that it is an imperfect market, the other possible explanation is that the implied volatility is different from the historical one because of the volatility smile[18].

3.2.4 Shortcomings of Black-Scholes

In the real world security prices do not follow stationary log-normal process[19], neither the risk-free interest rate is known and as was mentioned above the volatility of an underlying asset is not constant. All of these facts are assumed by Black-Scholes. The

[18] More avout volatility smile could be found in 4.4.3.

[19] See for example Eberlein 1995 for empirical evidence.

model also does not give good results for options far out of the money or far in the money. The log-normal distribution implies that extreme events are very rare, yet they occur more often in the real world. It means that there should be used some heavy-tailed distribution instead of the log-normal one.

The assumption of zero transaction costs is also unrealistic and it yields a significant liquidity risk, which is difficult to hedge. Other risks connected to unrealistic assumptions such as volatility risk[20], tail risk[21] or gap risk[22] could be hedged by using volatility hedging, out of the money options or gamma hedging. But this hedging also brings other costs which must be taken into account.

Nevertheless the Black-Scholes is the most popular tool for option pricing in spite of its shortcomings. The proper knowledge of Greeks under Black-Scholes leads to a significant lowering of the connected risks. If one considers parameters such as volatility or risk free rate as variables, not constants, the more realistic model is obtained. Then the volatility could be described for example by GARCH process and the model instantly becomes more realistic and gives better results.

3.2.5 French Black-Scholes Model

The French Black-Scholes model is a modification which is used when calendar days to expiration and trading days to expiration differs. This model was firstly proposed by D. French in 1984[23]. The original Black-Scholes uses just calendar days to evaluate an option. Under the real circumstances volatility is related to the number of trading days and the interest is compounded according to calendar days.

This difference is significant especially in the pricing of the short life options. The price of an option calculated by French Black-Scholes model is calculated using the following formulae.

$$C = S * N(d_1) - Xe^{-rT_c}N(d_2), \qquad \text{(EQ 3.2.12)}$$

where T_c is a number of calendar days until the expiration divided by a number of days in a year,

---

[20] The risk that the volatility would sinifficantly change.
[21] The risk of an occurence of an abnormal change of the underlying's price.
[22] The risk that the price will change form one level to another without any trading in between.
[23] See French 1984.

$$d_1 = \frac{\ln\frac{S}{X} + rT_t}{\sigma\sqrt{T}} + \frac{1}{2}\sigma\sqrt{T_t} \qquad \text{(EQ 3.2.13)}$$

and

$$d_2 = d_1 - \sigma\sqrt{T_t}, \qquad \text{(EQ 3.2.14)}$$

where T_t is a number of trading days divided by a number of trading days in a year. The N stands for the probability distribution function of the normal distribution again.

After plugging into the put-call parity one gets the following formula to obtain the price of a put option:

$$P = Xe^{-rT_c}N(-d_2) - S * N(-d_1). \qquad \text{(EQ 3.2.15)}$$

3.3 Quadratic Approximation Model

A quadratic approximation model was derived by Barone, Adesi and Whaley in 1987[24], it is referred as BAW or Whaley Model. It is a method to evaluate American type options. Since this method was introduced in the times when computers were much slower than nowadays; the main focus was on the computational complexity. This model is computationally much easier than the binomial model, but still precise.

3.3.1 Derivation of the Pricing Formula

The idea behind the BAW model is that the option price follows the stochastic differential equation given by:

$$1/_2\,\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + bS\frac{\partial V}{\partial S} - rV + \frac{\partial V}{\partial t} = 0, \qquad \text{(EQ 3.3.1)}$$

where V is a price of the option and b is a continuous rate of payout of the underlying asset[25]. Then the authors suggest that the price of an American option can be described by:

[24] See Barone, Adesi and Whaley 1987

[25] It is calculated as an interest rate minus dividend payment. The Black-Scholes model derived above assumed that this payout is same as the interest rate since it is possible to sell the asset immediately after exercising the option. The other possibility is that a holder of the underlying keeps the underlying and receives a dividend payout, which can be also zero. In this case there is a need to change the Black-Scholes pricing formula.
The price of a European call under the Black-Scholes is then

$$C(S,T) = c(S,T) + \varepsilon_c(S,T), \qquad \text{(EQ 3.3.2)}$$

where $C(S,T)$ is a price of an American call option for an underlying with a spot price S and time to maturity T, $c(S,T)$ is a price of European call option for the same underlying and $\varepsilon_c(S,T)$ is the early exercise premium. The rationale behind the model is that the early exercise premium must satisfy EQ 3.3.1, which means

$$1/2\,\sigma^2 S^2 \frac{\partial^2 \varepsilon}{\partial S^2} - r\varepsilon + bS\frac{\partial \varepsilon}{\partial S} + \frac{\partial \varepsilon}{\partial t} = 0. \qquad \text{(EQ 3.3.3)}$$

Now let us denote $M = 2r/\sigma^2$, $N = 2b/\sigma^2$ and $T = t^* - t$ where t^* is the time to expiration of the option and multiply the whole equation by $2/\sigma^2$ to get

$$S^2 \frac{\partial^2 \varepsilon}{\partial S^2} - M\varepsilon + NS\frac{\partial \varepsilon}{\partial S} - \frac{M}{r}\frac{\partial \varepsilon}{\partial T} = 0. \qquad \text{(EQ 3.3.4)}$$

Now the early exercise premium could be described by $\varepsilon_c(S,K) = K(T)f(S,K)$, so the derivatives are

$$\frac{\partial^2 \varepsilon}{\partial S^2} = K\frac{\partial^2 f}{\partial S^2} \text{ and } \frac{\partial}{\partial T} = \frac{\partial K}{\partial T}f + K\frac{\partial K}{\partial T}\frac{\partial f}{\partial T}$$

Now choose $K(T) = 1 - e^{-rT}$and plug into EQ 3.3.4 to get

$$S^2 \frac{\partial^2 f}{\partial S^2} + NS\frac{\partial f}{\partial S} - Mf/K - (1-K)\frac{\partial f}{\partial K} = 0. \qquad \text{(EQ 3.3.5)}$$

Now the approximation comes: the last term of LHS is assumed to be zero since the limit of this term goes to zero for T going to zero or infinity[26]. So we get

$$S^2 \frac{\partial^2 f}{\partial S^2} + NS\frac{\partial f}{\partial S} - Mf/K = 0, \qquad \text{(EQ 3.3.6)}$$

Which is a second order ODE. There are two linear independent solutions, which could be found by substituting $f = aS^q$:

$$aS^q[q^2 + (N-1)q - M/K] = 0. \quad \text{(EQ 3.3.7)}$$

$$c(S,T) = Se^{b-r} * N(d_1) - Xe^{-rT}N(d_2),$$

where

$$d_1 = \frac{\ln\frac{S}{X} + (b + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}$$

and

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

[26] It means that the option is very close or very far to its expiration.

Solve for q to get

$$q_{1,2} = (1-N) \pm \frac{\sqrt{(N-1)^2+4M/K}}{2}. \qquad \text{(EQ 3.3.8)}$$

Now EQ 3.3.8 determines the general solution:

$$f(S) = a_1 S^{(1-N)-\frac{\sqrt{(N-1)^2+4M/K}}{2}} + a_2 S^{(1-N)+\frac{\sqrt{(N-1)^2+4M/K}}{2}}. \qquad \text{(EQ 3.3.9)}$$

Since $M/K > 0$ implies $\lim_{S\to O} f(S) = 0$ a_1 must be zero, which means that the approximate value of the American call option is

$$C(S,T) = c(S,T) + K a_2 S^{(1-N)+\frac{\sqrt{(N-1)^2+4M/K}}{2}}. \qquad \text{(EQ 3.3.10)}$$

The last task is to determine a_2. If $S = 0$ then all terms of EQ 3.3.10 are equal to zero. However if S rises then the value of the European call rises and also the early exercise premium rises[27]. This means that $a_2 > 0$. Since the value of the American option is always at least as high as the price of the European option the curve given by RHS of EQ 3.3.10 should be touching, but not intersecting the line given by $S - X$[28]. So the constant a_2 could be obtained by solving

$$S - X = c(S,T) + K a_2 S^{(1-N)+\frac{\sqrt{(N-1)^2+4M/K}}{2}}. \qquad \text{(EQ 3.3.11)}$$

Since the LHS of the EQ 3.3.11 is linear in S and RHS is convex in S there exist just one unique S^* solving the equation

$$S^* - X = c(S^*, T) + K a_2 S^{*\,(1-N)+\frac{\sqrt{(N-1)^2+4M/K}}{2}}. \qquad \text{(EQ 3.3.12)}$$

Take into account that the tangents must be the same in the intersection.

$$1 = e^{(b-r)T} N[d_1(S^*)] + K\left[(1-N) + \frac{\sqrt{(N-1)^2+4M/K}}{2}\right] a_2 S^{*\,\frac{\sqrt{(N-1)^2+4M/K}}{2}-N}.$$
(EQ 3.3.13)

There are two equations and two unknowns. Rewrite EQ 3.3.13.

---

[27] More precisely - it does not fall.

[28] This also means that a function given by RHS of EQ 3.3.10 is convex.

$$a_2 = \frac{1-e^{(b-r)T}N[d_1(S^*)]}{K\left[(1-N)+\frac{\sqrt{(N-1)^2+4M/K}}{2}\right]S^{*\frac{\sqrt{(N-1)^2+4M/K}}{2}-N}} \qquad \text{(EQ 3.3.14)}$$

and plug into EQ 3.3.13 to finally get

$$S^* - X = c(S^*, T) + \frac{\{1-e^{(b-r)T}N[d_1(S^*)]\}S^*}{(1-N)+\frac{\sqrt{(N-1)^2+4M/K}}{2}}. \qquad \text{(EQ 3.3.15)}$$

In spite of the fact that S^* is the only unknown variable in EQ 3.3.15 it cannot be analytically determined so the iterative solution for S^* will be provided below.

The pricing formula for an American call option calculated from EQ 3.3.15 is given by the following formula:

$$C(S,T) = c(S,T) + S^* A_2 * \left(\frac{S}{S^*}\right)^{(1-N)+\frac{\sqrt{(N-1)^2+4M/K}}{2}}, \qquad \text{(EQ 3.3.16)}$$

where

$$A_2 = \frac{S^*}{(1-N)+\frac{\sqrt{(N-1)^2+4M/K}}{2}}\{1-e^{(b-r)T}N[d_1(S^*)]\}^{29}, \qquad \text{(EQ 3.3.17)}$$

for $S < S^*$ and

$$C(S,T) = S - X, \qquad \text{(EQ 3.3.18)}$$

for $S \geq S^*$.

One can analogically get pricing equations for put options. Now the S^{**} is determined by

$$S^{**} - X = p(S^*, T) + K a_1 S^{**(1-N)-\frac{\sqrt{(N-1)^2+4M/K}}{2}} \qquad \text{(EQ 3.3.19)}$$

and the pricing formulas are

$$P(S,T) = p(S,T) + S^{**} A_1 * \left(\frac{S}{S^*}\right)^{(1-N)-\frac{\sqrt{(N-1)^2+4M/K}}{2}}, \qquad \text{(EQ 3.3.20)}$$

[29] Note that $A_2 > 0$ since the costs of carry rate is lower or equal to the interest rate and the quintile of normal distribution is always lower or equal to 1.

where

$$A_1 = \frac{-S^{**}}{(1-N)-\frac{\sqrt{(N-1)^2+4M/K}}{2}}\left\{1-e^{(b-r)T}N[-d_1(S^{**})]\right\}^{30}, \qquad \text{(EQ 3.3.21)}$$

for $S > S^{**}$ and

$$P(S,T) = X - S, \qquad \text{(EQ 3.3.22)}$$

for $S \leq S^{**}$.

3.3.2 Newton's Algorithm

The only task left is determining S^* (resp. S^{**}) now. In case of pricing a call option one needs to find a root of equation EQ 3.3.15, which means solving:

$$g(S^*) = S^* - X - c(S^*, T) - \frac{\{1-e^{(b-r)T}N[d_1(S^*)]\}S^*}{(1-N)+\frac{\sqrt{(N-1)^2+4M/K}}{2}} = 0. \qquad \text{(EQ 3.3.23)}$$

The EQ 3.3.23 cannot be analytically solved so it is needed to implement an approximation algorithm. A Newton Algorithm seems as a rational choice here[31]. It is an iterative method following the formula below.

$$S_{i+1} = S_i - \frac{g(S_i)}{g'(S_i)}. \qquad \text{(EQ 3.3.24)}$$

The process described by EQ 3.3.24 is repeated until a value of S_i is sufficiently close to the root of EQ 3.3.23.

Lastly determine the derivative g':

$$g'(S) = \left(1 - {}^{1}/_{q_2}\right)\left\{1 - e^{(b-r)T}N[d_1(S^*)]\right\} + \frac{1}{q_2}e^{(b-r)T}N[d_1(S^*)]\frac{1}{\sigma\sqrt{T}}. \qquad \text{(EQ 3.3.25)}$$

3.3.3 Special Cases

For the practical part will be considered just two special cases, which simplify the general valuation formula described above. The first one is that costs of carry are zero it

---

[30] Note that $A_1 > 0$ since the costs of carry rate is lower or equal to the interest rate and the quintile of normal distribution is always lower or equal to 1.

[31] More about this method could be found for example in Pánková 2009.

means that the payout from holding the underlying is the same as an interest rate. This is true if a buyer immediately sells the underlying and gets cash for it. The pricing formulas are:

$$C(S,T) = c(S,T) + \frac{S^*\{1-e^{-r}N[d_1(S^*)]\}}{-1+\frac{\sqrt{1+4M/K}}{2}} * \left(\frac{S}{S^*}\right)^{-1+\frac{\sqrt{1+4M/K}}{2}}, \qquad \text{(EQ 3.3.26)}$$

for $S < S^*$ and

$$C(S,T) = S - X, \qquad \text{(EQ 3.3.27)}$$

for $S \geq S^*$.

The function $g(S)$ is now given by

$$g(S^*) = S^* - X - c(S^*,T) - \frac{\{1-e^{-r}N[d_1(S^*)]\}S^*}{-1+\frac{\sqrt{1+4M/K}}{2}}. \qquad \text{(EQ 3.3.28)}$$

And the derivative is

$$g'(S) = \left(1 - \frac{1}{-1+\frac{\sqrt{1+4M/K}}{2}}\right)\{1 - e^{-rT}N[d_1(S^*)]\} + \frac{1}{-1+\frac{\sqrt{1+4M/K}}{2}} e^{-rT}N[d_1(S^*)]\,\frac{1}{\sigma\sqrt{T}}.$$

(EQ 3.3.29)

One could analogically obtain the pricing formulas for American put options.

The other restriction, which will occur in the practical part, is that the costs of carry are equal to the interest rate. It means that a buyer holds an underlying till the expiration of the original option and they do not receive any dividend payout. The pricing formulas are now

$$C(S,T) = c(S,T) + S^*A_2 * \left(\frac{S}{S^*}\right)^{(1-N)+\frac{\sqrt{(N-1)^2+4M/K}}{2}}, \qquad \text{(EQ 3.3.30)}$$

where

$$A_2 = \frac{S^*}{(1-N)+\frac{\sqrt{(N-1)^2+4M/K}}{2}}\{1 - N[d_1(S^*)]\}^{32}, \qquad \text{(EQ 3.3.31)}$$

for $S < S^*$ and

$$C(S,T) = S - X, \qquad \text{(EQ 3.3.32)}$$

for $S \geq S^*$.

The function g

$$g(S^*) = S^* - X - c(S^*,T) - \frac{\{1-N[d_1(S^*)]\}S^*}{(1-N)+\frac{\sqrt{(N-1)^2+4M/K}}{2}}, \qquad \text{(EQ 3.3.33)}$$

and its derivative

$$g'(S) = \left(1 - {}^{1}\!/_{q_2}\right)\{1 - N[d_1(S^*)]\} + \frac{1}{q_2}N[d_1(S^*)]\frac{1}{\sigma\sqrt{T}}, \qquad \text{(EQ 3.3.34)}$$

3.4 Bjerksund-Stersland Model

This model is another approximation for pricing American options. It was firstly described by Bjerksund and Stersland in 1993[33]. So it is younger method than BAW model. It also deals with the analytical impossibility of determining the price form of the American options.

The authors of the model assume that any American option has the same value as an European up-and-out call with a knock-out barrier I and a rebate calculated as $X - I$. The exercise of the option then depends whether the underlying price hits a flat boundary given by the trigger price I. According to the model the price of an American call option is

$$C = \alpha S - \alpha\phi(S,T,\beta,I,I) + \phi(S,T,1,I,I) - \phi(S,T,1,X,I) - X\phi(S,T,0,I,I) + X\phi(S,T,0,X,I), \qquad \text{(EQ 3.4.1a)}$$

with

$$\alpha = (I - X)I^{\beta}, \qquad \text{(EQ 3.4.1b)}$$

[32] Note that $A_2 > 0$ since the costs of carry rate is lower or equal to the interest rate and the quintile of normal distribution is always lower or equal to 1.
[33] See Bjerksund and Stersland 1993.

$$\beta = \left(\frac{1}{2} - \frac{b}{\sigma^2}\right) + \sqrt{\left(\frac{b}{\sigma^2} - \frac{1}{2}\right)^2 + 2\frac{r}{\sigma^2}}, \qquad \text{(EQ 3.4.1c)}$$

The function ϕ is given by

$$\phi(S,T,\gamma,H,I) = e^{\lambda} S^{\gamma} \left[N(d) - \left(\frac{I}{S}\right)^{\kappa} N\left(d - \frac{2\ln\left(\frac{I}{S}\right)}{\sigma\sqrt{T}}\right)\right], \qquad \text{(EQ 3.4.2a)}$$

where

$$\lambda = [\gamma b - r + 0.5\gamma(\gamma - 1)\sigma^2]T^{34}, \qquad \text{(EQ 3.4.2b)}$$

$$d = \frac{\ln\left(\frac{S}{H}\right) + [b + (\gamma - 0.5)\sigma^2]T}{\sigma\sqrt{T}}, \qquad \text{(EQ 3.4.2c)}$$

$$\kappa = \frac{2b}{\sigma^2} + (\gamma - 1). \qquad \text{(EQ 3.4.2d)}$$

The trigger price is calculated as

$$I = B_0 + (B_\infty - B_0)(1 - e^f), \qquad \text{(EQ 3.4.3a)}$$

$$f = -(Tb + 2\sigma\sqrt{T})\frac{B_0}{B_\infty - B_0}, \qquad \text{(EQ 3.4.3b)}$$

$$B_\infty = \frac{\beta}{\beta - 1}X, \qquad \text{(EQ 3.4.3c)}$$

$$B_0 = X * \max\left[1, \frac{r}{r-b}\right]. \qquad \text{(EQ 3.4.3d)}$$

To price an American put one can use a put-call parity given by

$$C(S,X,T,r-b,-b,\sigma) = P(S,X,T,r,b,\sigma). \qquad \text{(EQ 3.4.4)}$$

3.5 Jump Diffusion Model

This model was firstly introduced in 1976 by Merton[35], who was awarded the Nobel Prize in Economics for this work. This model addresses the “heavy” tails of the observed distributions of the underlying prices. The assumption of this model is that an underlying price can jump to a different level at any time. Merton proposed a model which assumes that the extra randomness due to jumps can be diversified away.

---

[34] Note that γ stands for β, 1 or 0 depending on the function ϕ, which is currently calculated.
[35] See Merton 1976.

The point of this model is that the stock's return can be described by the sum of Wiener process and Poisson process[36] given by

$$\frac{dS}{S} = (\alpha - \lambda k)dt + \sigma dz + dq, \qquad \text{(EQ 3.5.1)}$$

Where α is the instantaneous expected return on the stock, q is the independent Poisson process with parameter λ, which stand for the mean number of arrivals per unit of time and $k = E(Y - 1)$, where Y is the random variable percentage change in the stock price if the Poisson event occur. So the dz describes just "normal" price fluctuations and the dq describes unexpected events.

Merton then arrived to the following equation for the option pricing:

$$F(S,T) = \sum_{n=0}^{\infty} \frac{e^{-\lambda T}(\lambda T)^n}{n!} E_n\left[BS(SQ_n e^{-\lambda k T}, Y, \sigma^2, r)\right], \qquad \text{(EQ 3.5.2)}$$

where $E_n[\ldots]$ stands for the expected value at the n-th step, Q_n is the random variable identically distributed as the product of n random variables Y defined above, there is a restriction that $Y_0 \equiv 1$, and $BS(\ldots)$ is the price calculated by the Black-Scholes formula.

There are two new parameters entering the pricing formula: λ, which is determining the number of Poisson events per year and k which reflects how much is the underlying price affected by each Poisson event.

[36] The Poisson process is a discrete time process, where occur so called Poisson events with a probability $P\left[\left(N(t+\tau) - N(t)\right) = k\right] = \frac{e^{-\lambda\tau}(\lambda\tau)^k}{k!}$

4. Empirical Part

In this part the functionality and efficiency of methods described in the theoretical chapter will be tested. The aim of this is to compare the performance of the pricing formulas on the real data. We will evaluate options from the available dataset to decide which methods are efficient and suitable to use.

4.1 Description of the Dataset

The dataset contains the data for all options listed at CBOE[37] from 31st and 28th July 2009. The dataset contains a huge amount of data – there were 316 245 different options listed at CBOE on 31st July 2009.

The pricing will be done for the 31st July the dataset from 28th July was used just to assign values of the historical implied volatility. The next dataset which was used contains historical volatilities of all underlying assets on which were issued options at CBOE. This was used to calculate the number of jumps in the underlying price during a year and to assign a historical volatility to each underlying asset.

Each option has the following specifications:

- **Underlying Symbol –** the shortcut of the underlying stock (usually 1-4 letters),
- **Underlying Price** – the spot price of the underlying asset,
- **Option Ext** – two letters code implying the date of expiration and the strike price[38],
- **Type –** call or put,
- **Expiration –** the expiration date,
- **Datadate –** the date when the data were acquired,
- **Strike** – strike price,
- **Last** – last ask price,
- **Bid –** market bid price,
- **Ask** – market ask price,
- **Volume** – number of contracts traded during a day,

[37] Chicago Board of Options Exchange – for more details see www.cboe.com
[38] Each month has assigned a different letter and there is also a different letter indicating last two digits of the option price.

- **Open Interest –** number of open contracts on the specified option
- **IV –** implied volatility calculated from the Black-Scholes formula
- **Delta –** option delta
- **Gamma –** option gamma
- **Theta –** option theta
- **Vega –** option vega

All options in the dataset are American type according to the available description of the dataset. All Greeks were calculated by the basic Black-Scholes algorithm[39] using the Fed funds rate, which was 0.18 % on 31st July 2009[40].

Firstly a general description of the whole dataset will be provided. We will also estimate option prices by Blask-Scholes model for all available data. Other pricing methods are analytically more demanding so they will be tested on the subsequent datasets. Since the dataset contains a large amount of rows there will be provided characteristics of pricing methods on various samples of the original datasets.

All calculations connected to option pricing, summary statistics, descriptive statistics and plotting figures were made in MS Excel. It was necessary to use Add-in feature and macros implemented by VBA code to obtain all results. Econometric analysis was performed by Gretl software.

4.2 General Description of the Dataset

It was needed to omit some options from the dataset because of incomplete information or not available implied volatility from 28th July 2009. The rest of the dataset contains 273 152 options, which is still more than a sufficient size of the sample.

The analyzed dataset contains 136 544 call options and 136 608 put options. 225 842 options were not traded during the trading day and 187 383 had zero open interest. The

[39] The whole dataset and the information about the variables and their description was obtained at http://www.historicaloptiondata.com. One should also take into account the future dividend payout when pricing an option. This information was unfortunately not provided in the dataset and it is extremely hard to obtain such information for more than 22 000 underlying assets of options traded at CBOE. Further more this information is also hard to gather for a prospective investor, so we shall evaluate all options as if they have no future dividend income.

[40] We shall use this as an interest rate during the whole pricing process in this chapter. The yield rate was considered to be zero for all models. It was tried to set the yield rate equal to the interest rate, but the change in the results was nearly zero.

average bid ask spread was approximately 0.49 USD. The average time to expiration was 158 days the median value was 140 days.

There was calculated the price of all options by the Black-Scholes model[41] and the ask price was subtracted from this value. This difference posses the characteristics provided below. Two calculations were made: the first one computes with the historical volatility and the second one was obtained using the implied volatility from three days before.

	HV	IV
mean	16.29441	3.878228
median	0.949896	-0.02
max	1602.76	1393.439
min	-3585.2	-3586.2
Q1	-1.50069	-2.5
Q3	22	7.8259
stddev	89.63391	58.39637
skew	7.996411	4.687983
kurt	147.7505	295.1921
underpriced	6767	7000
overpriced	22598	8025

Tab. 2: *Differences between estimated and the market price (whole dataset, Black-Scholes model)*

HV stands for the historical volatility approach, IV stands for the implied volatility. The mean is calculated as a simple arithmetic average, Q1 and Q3 stands for quartiles, “stddev” is a standard deviation, “skew” stands for skewness, “kurt” for kurtosis, “underpriced” stands for the number of options whose value was estimated more than 50 USD[42] lower than the market price and “overpriced” indicates the number of options whose price was estimated more than 50 USD higher than the market price.

It is clear that the implied volatility approach brought better results than the historical one. The mean and median are closer to zero, the quartile spread and the standard deviation is lower, the kurtosis is higher. The historical volatility approach mispriced more than 10 % of all options, the implied volatility one mispriced just 5 % of the options from the whole dataset.

[41] Note that the Black-Scholes formula is suitable just for European options, but the dataset contains the American ones. Let’s consider the difference between the American and European option small enough missing index to be omitted. The results of the methods pricing American options will be provided below, but it cannot be done on the whole dataset due to the computational complexity.

[42] 50 USD can look as an enormous value, but even those values can occur in such a large dataset. A relative-difference approach would probably provide different results. Find out more about it in 4.4.3.

There are provided log-histograms below to find out more about the distribution of differences[43].

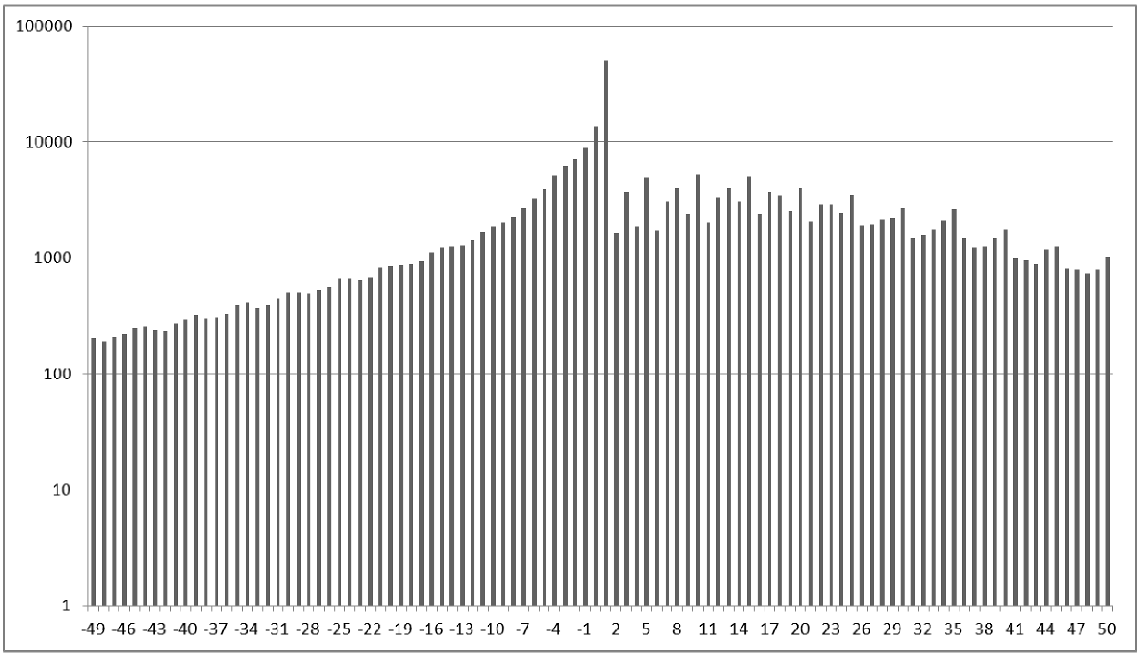

Fig. 7: *Log-histogram of deviances of the Black-Scholes Model from the market data (the whole dataset, historical volatility)*

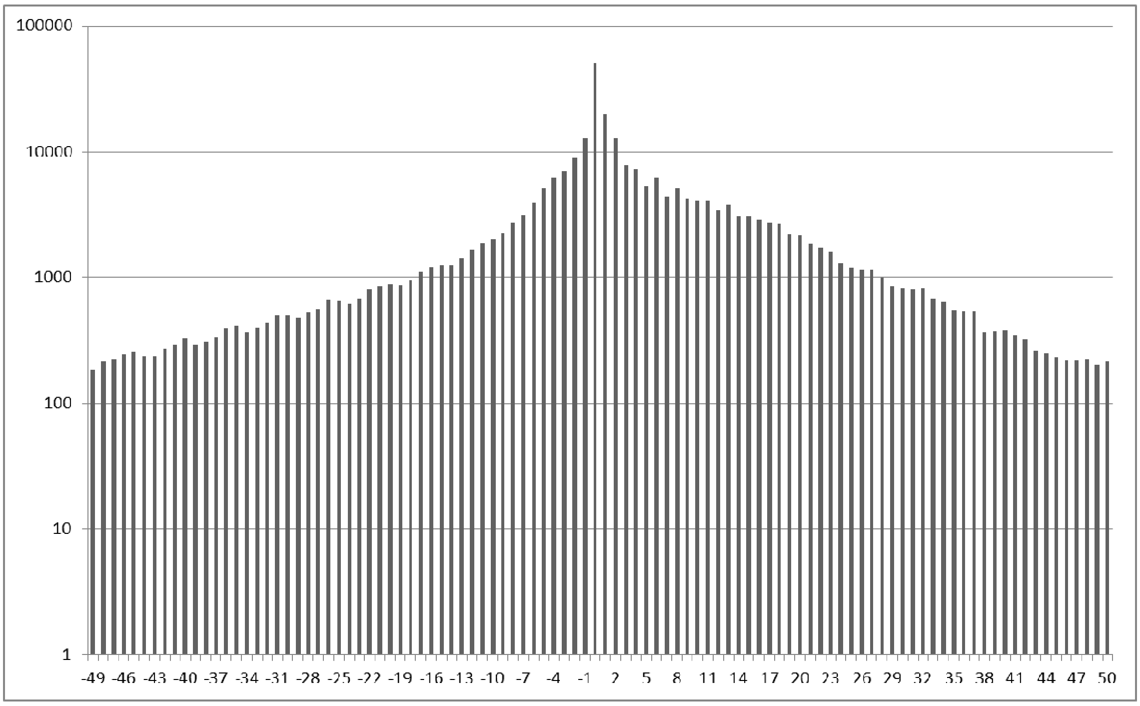

Fig. 8: *Log-histogram of deviances of the Black-Scholes Model from the market data (whole dataset, implied volatility)*

[43] The mispriced observations are omited.

Note the strong asymmetry for the estimates obtained by the model using historical volatility. The modal interval for the historical volatility model is between 0 USD and 1 USD and between -1 and 0 for the implied volatility model. Both modal intervals consist of more than 50 000 observations.

This section showed that the Black-Scholes formula for the European option works quite well even for the American options.

4.3 Comparison of the Pricing Methods on the Subsamples of the Dataset

Several subsets of the data were chosen for this part. The criteria for choosing the data was based on the RAND() function in MS Excel and specific requirements on the datasets. All subsets in this chapter contain 1000 rows.

We shall provide a general description and commentary of the results in the first subchapter and then just show the results and point out the most important facts of the results obtained in the other datasets. There will be used 6 different pricing methods:

- **Black-Scholes Model (BS) –** we provide the results of pricing by this model even though all options in the dataset are American type and this model evaluates just European options. The European price is the lower bound of the American price, which means that this model should systematically underestimate the option prices.
- **French Black-Scholes Model (FBS) –** this model is also for European options, so there should be the same problem with pricing as above. The number of trading days was calculated as the number of days to expiration *5/7 rounded down.
- **Binomial Model (bin) –** this is the first model which takes into account the fact that the options are American type. The number of iterations was calculated as the number of trading days divided by 7, rounded down, plus 5.
- **BAW Model (BAW) –** the quadratic approximation model evaluates American type options. The parameters are the same as for the Black-Scholes model.
- **Bjerssund-Stersland Model (BJST) –** This is another approximation model for American type options. The parameters are the same as for the Black-Scholes model.

- **Jump Diffusion Model (JD) –** This model is used to price European type options, but we shall use it experimentally for evaluating American options to determine whether the results are applicable even to the American options or not. There is a need to estimate 2 more parameters to price an option as was stated in 3.5. The proper estimation of those parameters would require a detailed analysis of the underlying price time series. Providing this for the whole dataset would demand much more data that are currently not available. So these parameters were experimentally "guessed" from the available data. The number of jumps (λ) was estimated from the historical volatility by the following algorithm: monthly data from previous two years were available, standard deviation was calculated from these data, the number of changes in the monthly volatility data higher than one standard deviation was counted and finally this number was divided by two and rounded as an estimate of number of jumps a year. It was assumed that the percentage of the total volatility explained by jumps (k) is 50 % for all options. These estimates probably do not correspond with the parameters acquired by the proper analysis of the underlying time series, but we shall try those values and observe the implications of this.

4.4 Random Sample

The first random sample contains options selected just by randomness there was no other criteria for the selection. The set contains 156 options which were traded during the data date and 686 options with nonzero open interest. There are 521 call options and 479 put options. The average bid/ask spread was 0.6 USD. 433 options were in the money[44], 482 options were out of the money[45] and the remaining options are considered to be at the money.

Before estimating the price by various models there is an essential need to define volatility which will be used for an estimate. There are 2 possibilities:

- **Historical volatility –** this characteristic is provided every month by CBOE for all underlying assets. Since this is in USD it must be divided by stock price to get the percentage amount. This volatility is same for all options issued on the same underlying.

[44] Strike/Underlying < 0.95 for calls and Underlying/ Strike < 0.95 for puts
[45] Strike/Underlying > 1.05 for calls and Underlying/ Strike > 1.05 for puts

- **Implied volatility –** this characteristic differs for various options. The implied volatility in the analyzed dataset was calculated as the volatility which after plugging in the Black-Scholes model gives the market price. This volatility is known everyday at the same time as the market price. Since the aim of usage of option pricing models is to predict the option prices it is impossible to use this volatility in the models. Fortunately there are available the same data from 28th July 2009. The value of implied volatility from three days before for each option in the dataset from 31st July was looked up in the older dataset and loaded to the younger one. Now it is rational to suppose that one knows three-days-old implied volatility for each option.

We shall show the results for both volatilities and compare with the market prices. Since the prices of options differ we shall use the difference between the market price and the estimated price to compare the methods. There will be recognized three different market prices in the following text:

- **Bid price –** this is the lowest price that a seller is willing to sell.
- **Ask price –** this is the highest price that a buyer is willing to buy.
- **Mean price –** the arithmetic average of bid and ask.

4.4.1 Historical Volatility Approach

The first evaluation of the options is calculated using the mean price to compute the difference and historical volatility. The results are in the following table.

	BS	FBS	Bin	BAW	BJST	JD
Mean	3.372011	3.45300318	3.1746936	3.3740771	3.201811	5.647096
median	2.906937	2.99219056	2.7451826	2.9108578	2.789229	2.771154
Max	261.6902	261.515	261.69017	261.69017	261.6902	463.4188
Min	-207.3	-207.3	-207.3	-207.3	-207.3	-207.3
Q1	0.730264	0.78003664	0.4791547	0.7304803	0.653325	0.756851
Q3	5.774186	5.95788727	5.6321502	5.7745587	5.588952	5.80694
Stddev	13.68631	13.6999055	13.69138	13.686355	13.59274	27.18006
Skew	2.232468	2.20150579	2.265608	2.2324621	2.247393	9.670222
Kurt	197.467	196.293128	197.34844	197.46396	202.6752	148.8149

Tab. 3: *Differences between estimated and the mean market price (random data, historical volatility)*

On the first line there is an average difference between the mean market price and estimated price by each model. On the second line is the median difference. There is a maximum and minimum on the third and fourth line. Next two lines contain first and third quartile. On the last three lines are provided moment characteristics of the data: standard deviation, skewness and kurtosis.

It seems that all models are overestimating the market price because the mean and the median are positive for all models. Also the first quartile is positive for all models. The possible explanation is that the models are build for investors willing to buy options, so here are calculated all outputs again with ask price as a market price:

	BS	FBS	bin	BAW	BJST	JD
mean	3.072201	3.15319318	2.874884	3.0742671	2.902001	5.347286
median	2.696053	2.79298568	2.486749	2.696053	2.610428	2.585035
max	261.6552	261.48	261.6552	261.65517	261.6552	462.8188
min	-219.2	-219.2	-219.2	-219.2	-219.2	-219.2
Q1	0.430087	0.46662451	0.153727	0.4301078	0.337644	0.477441
Q3	5.617426	5.76029006	5.474221	5.617932	5.460521	5.669728
mtddev	13.92373	13.9374976	13.92692	13.923744	13.83352	27.29623
mskew	1.460188	1.43206845	1.494368	1.4601912	1.461516	9.451771
kurt	197.4106	196.279704	197.3575	197.40991	202.3281	147.0438

Tab. 4: *Differences between estimated and the ask price (random data, historical volatility)*

The systematic overestimation of the price is still present so the problem could be the wrong model selection or that the historical volatility is not suitable for estimating option prices very well. But before proceeding to the estimations with implied volatility the density of the differences will be analyzed by histograms.

It seems that the binomial model fitted the data the best - the mean is the closest to the zero and there is the lowest standard deviation from all the models. On the other hand the Jump-Diffusion model is the worst fit, this could be caused by the fact that the additional parameters were chosen laxly without proper consideration.

The following figures show the log-density of all models. The modal intervals have length 1 USD and are from -20 USD to +20 USD. For each interval a number of occurred differences between its bounds were counted and then a common logarithm was calculated from this number. Plotting these data provides a histogram. We shall firstly provide a histogram for methods evaluating options as a European type[46].

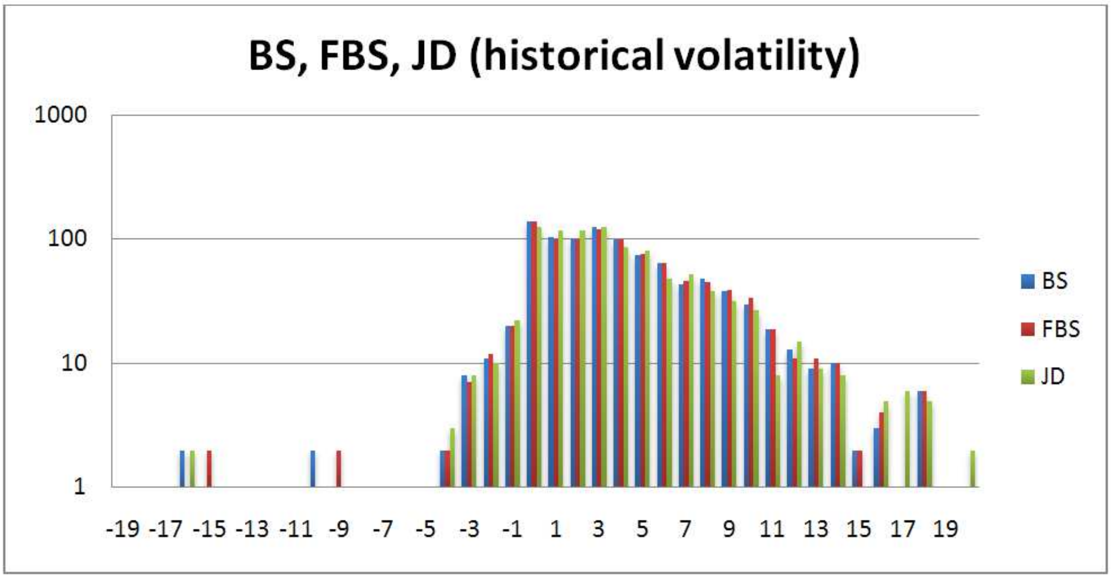

Fig. 9: *Log-histogram of deviances of the Black-Scholes, French Black-Scholes and Jump-Difussion Model from the market data (random data, historical volatility)*

The positive skew of all estimates is evident from the figure above. Note that the most estimates[47] were in the interval (0;1).

The following plot shows the log density of the remaining three models:

[46] There are ommited outliers for all density graphs from now on – the number of ommited observations varies for each histogram, but those are the same points as the ones described as mispriced in the following text.
[47] Precisely 138 for BS and FBS and 125 for JD Model.

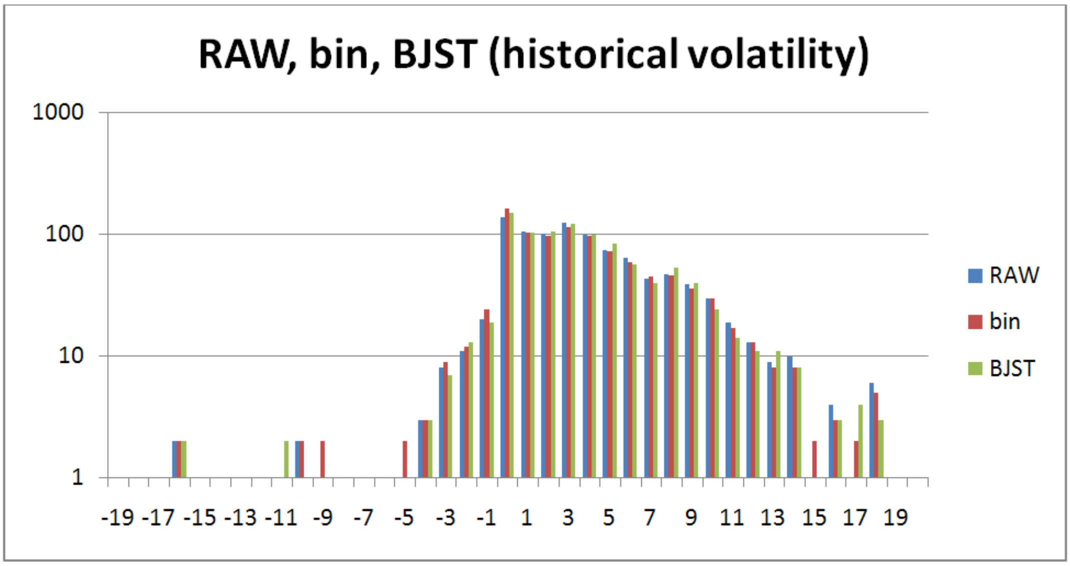

Fig. 10: *Log-histogram of deviances of the BAW, Binomial and BJST Model from the market data (random data, historical volatility)*

The modal interval is also (0;1) for these models. More than 950 options were priced with a difference between -5 USD and +15 USD.

All models except for Jump-Diffusion Model provide very similar results. To decide which one is the best could be used the R-squared statistics[48]. It is calculated as

$$R^2 = \frac{RSS}{TSS} = 1 - \frac{ESS}{TSS}. \qquad \text{(EQ 4.4.1)}$$

where RSS is the regression sum of squares, TSS is the total sum of squares and ESS is the error sum of squares, which could be calculated as

$$RSS = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2, \qquad \text{(EQ 4.4.2a)}$$

$$TSS = \sum_{i=1}^{n} (y_i - \bar{y})^2, \qquad \text{(EQ 4.4.2b)}$$

$$ESS = \sum_{i=1}^{n} (\hat{y}_i - y_i)^2, \qquad \text{(EQ 4.4.2c)}$$

The values of R-squared for all models are in the following table:

MODEL	BS	FBS	bin	BAW	BJST	JD
R-squared	0.816472	0.815671	0.817451	0.81646	0.81965	0.301597

Tab. 5: *R-squared of the used models (random data, historical volatility)*

[48] The R-squared is just one of the possible approaches to measure the quality of the model and it could be biased here because of the differences in the prices and because of the outliers in the dataset.

The R-squared of the used models except for the Jump-Diffusion one is very similar, but it suggests that the BJST model fits the best. From now on the R-squared will be automatically the part of the models outputs table.

4.4.2 Implied Volatility Approach

Now we shall provide the same calculations, but with a usage of implied volatility observed three days before the date of acquiring the dataset. As a market price was considered an ask price as in the previous section.

The table with summary statistics follows:

	BS	FBS	bin	BAW	BJST	JD
mean	0.026391	0.07398192	0.027869	0.027923	-0.01709	1.541238
median	-0.013889	0.00964	-0.01596	-0.01367	-0.02335	0.008826
max	261.6552	261.48	261.6552	261.6552	261.6552	540.9037
min	-219.2	-219.2	-219.2	-219.2	-219.2	-219.2
Q1	-0.105162	-0.0697959	-0.10982	-0.10446	-0.13936	-0.09234
Q3	0.058469	0.10790915	0.060754	0.059931	0.040929	0.105051
stddev	12.85127	12.8479727	12.85183	12.85131	12.85134	22.94104
skew	2.607804	2.58177416	2.607122	2.607424	2.617801	13.9947
kurt	274.5797	274.363612	274.5299	274.575	274.6089	334.3385
R-squared	0.850913	0.850985	0.8509	0.850912	0.850912	0.522769

Tab. 6: *Differences between estimated and the ask price (random data, implied volatility)*

The results look much better now. The mean moves around zero. The BJST has the lowest mean difference in the absolute value, but on the other hand the lowest median in the absolute value occurred for the FBS model and the BS model posses the lowest quartile spread. The FBS has the highest R-squared.

The histograms of the log-density are provided below.

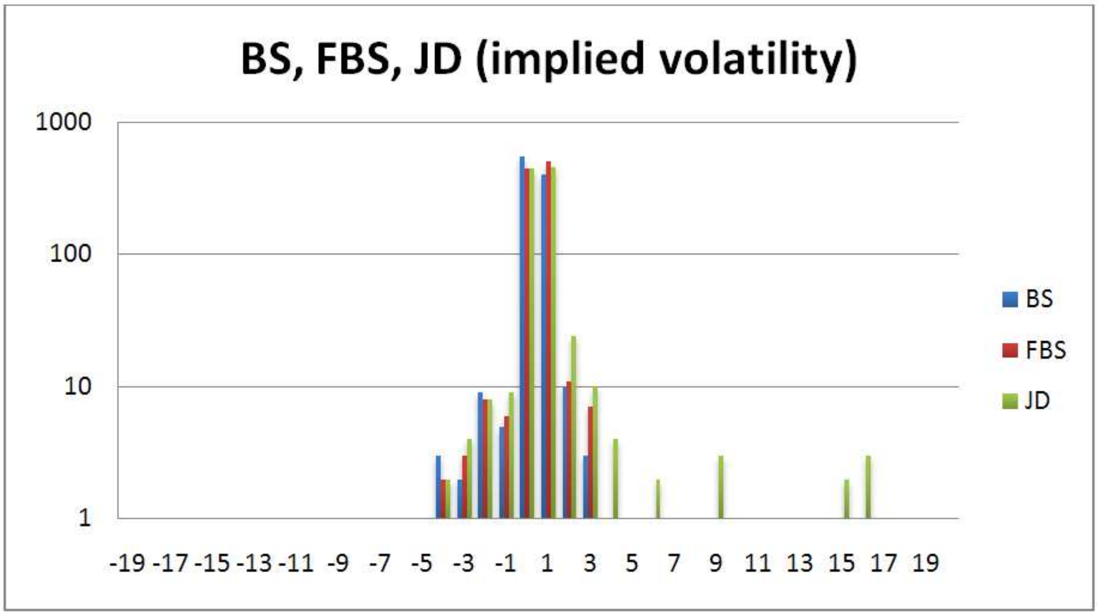

Fig. 11: *Log-histogram of deviances of the Black-Scholes, French Black-Scholes and Jump-Difussion Model from the market data (random data, implied volatility)*

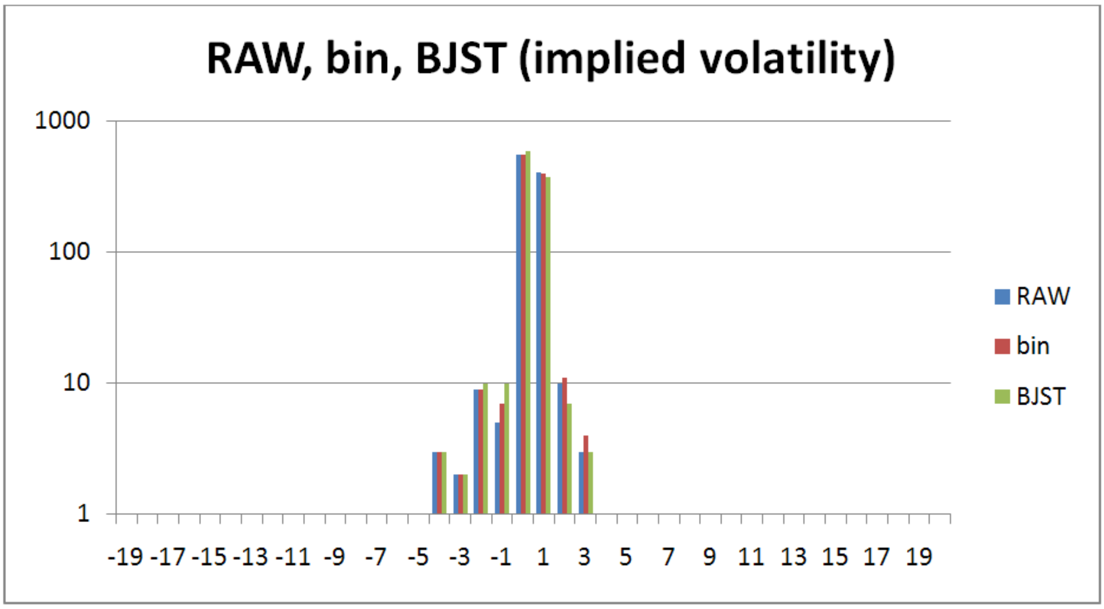

Fig. 12: *Log-histogram of deviances of the BAW, Binomial and BJST Model from the market data (random data, implied volatility)*

Note that the majority of models now fits very well so it would be reasonable to divide the observations into the smaller intervals. The following histogram plots the log-density between -1.02 USD and 1.02 USD with 0.04 USD long intervals.

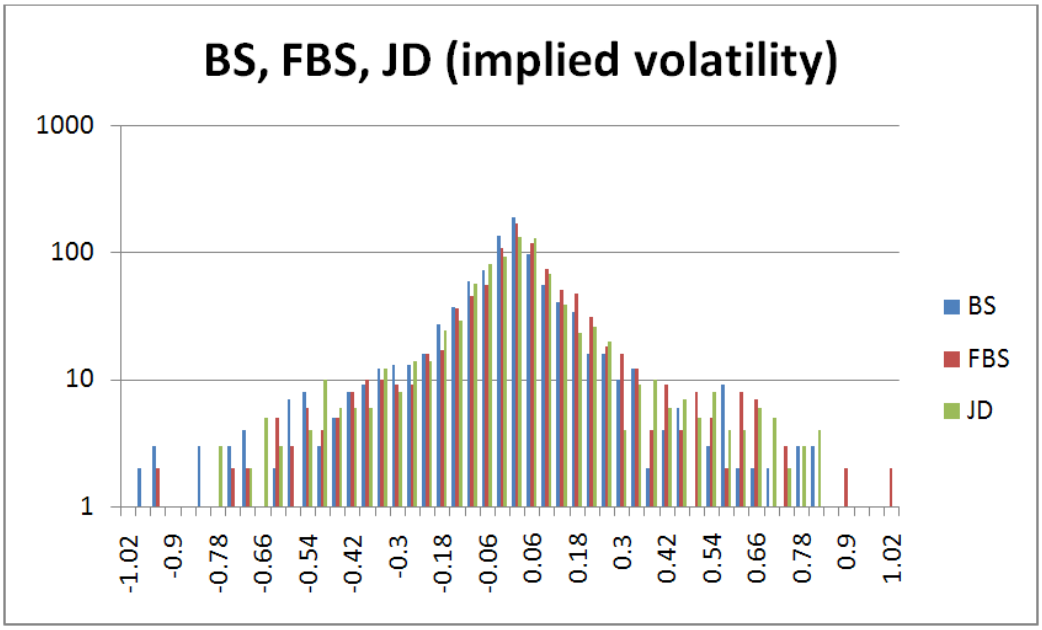

Fig. 13: *Log-histogram of deviances of the Black-Scholes, French Black-Scholes and Jump-Difussion Model from the market data (random data, implied volatility)*

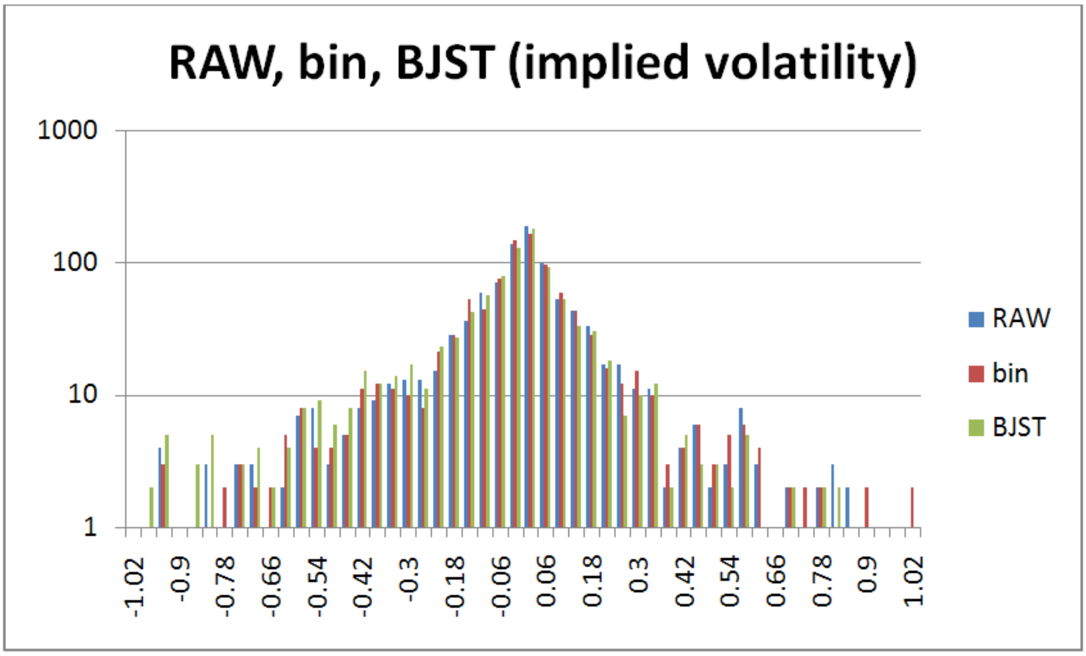

Fig. 14: *Log-histogram of deviances of the BAW, Binomial and BJST Model from the market data (random data, implied volatility)*

More than 130 options from the dataset were priced according to the ask price ±0.02 USD by all models. The most precise was the Black-Scholes Model which priced accurately 190 options. The skew is close to zero, which means that the histograms are nearly symmetric.

At end of the pricing spectrum there were just 48 options mispriced by all methods by more than 1 USD. Just 2 of them were traded on 31st July 2009[49], only 12 of those 48 had non-zero open interest and only 4 of them were out of the money!

The number of options priced further than 1 USD from the ask price is provided in the following table. It will become the essential output of all pricing comparisons from now on.

	BS	FBS	bin	BAW	JD	BJST
mispriced	48	53	52	48	101	51

Tab. 7: *The number of mispriced options (random data, implied volatility)*

There are some characteristics of mispriced options differing from the others in the dataset. The most significant one is that the most of mispriced options[50] was not traded and more than a half of them have zero open interest. The average ask price is about 60 USD which is for more than 50 USD higher than for the original dataset. The average bid-ask spread is about 3 USD. The put and call options in the mispriced sub-set are nearly uniformly distributed. Since there are just 1 000 options from 270 000 large dataset there were not found any two options issued on the same underlying.

4.4.3 Relative Differences

It seems that relative differences could be also useful for the analysis of the performance of the pricing methods. The relative difference is calculated according to the following formula.

$$relative\ difference = \frac{estimated\ price - ask\ price}{ask\ price}. \qquad \text{(EQ 4.4.3)}$$

This reflects the different prices of all options, but on the other hand it is more sensitive for the options with low ask price[51]. The results are in the following table.

[49] One was overpriced and one was underpriced.
[50] 90-100% depending on the pricing method.
[51] There are 346 options with a price under 1 USD in the dataset.

	BS	FBS	bin	RAW	BJST	JD
mean	0.451584	0.48519312	0.434103	0.451683	0.449607	0.840205
median	-0.004154	0.00358979	-0.00459	-0.00403	-0.00824	0.002013
max	261.6552	261.48	261.6552	261.6552	261.6552	261.6552
min	-1	-1	-1	-1	-1	-1
Q1	-0.048715	-0.0256148	-0.05708	-0.04871	-0.05074	-0.03348
Q3	0.025328	0.05216461	0.025507	0.025328	0.023382	0.089756
stddev	8.987129	8.98201816	8.987793	8.987126	8.987224	9.375395
skew	25.99481	25.9791099	25.99479	25.9948	25.99464	23.20535
kurt	728.6188	727.970292	728.6057	728.6187	728.611	613.6516
R-squared	0.926906	0.92696003	0.926909	0.926906	0.926906	0.920016
mispriced	67	68	77	67	67	142
underpriced	27	37	25	27	27	124
overpriced	40	31	52	40	40	18

Tab. 8: *Relative differences (random data, implied volatility)*

Options with estimated price differing by more than 50 % of ask price are considered as mispriced ones. The histograms are below.

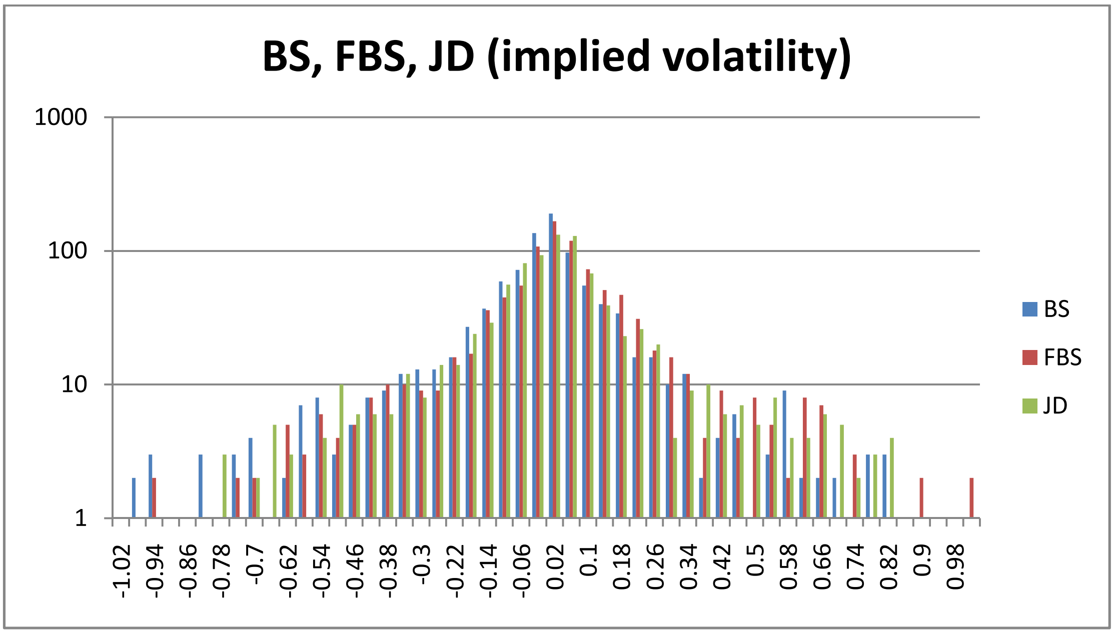

Fig. 15: *Log-histogram of relative deviances of the Black-Scholes, French Black-Scholes and Jump-Difussion Model from the market data (random data, implied volatility)*

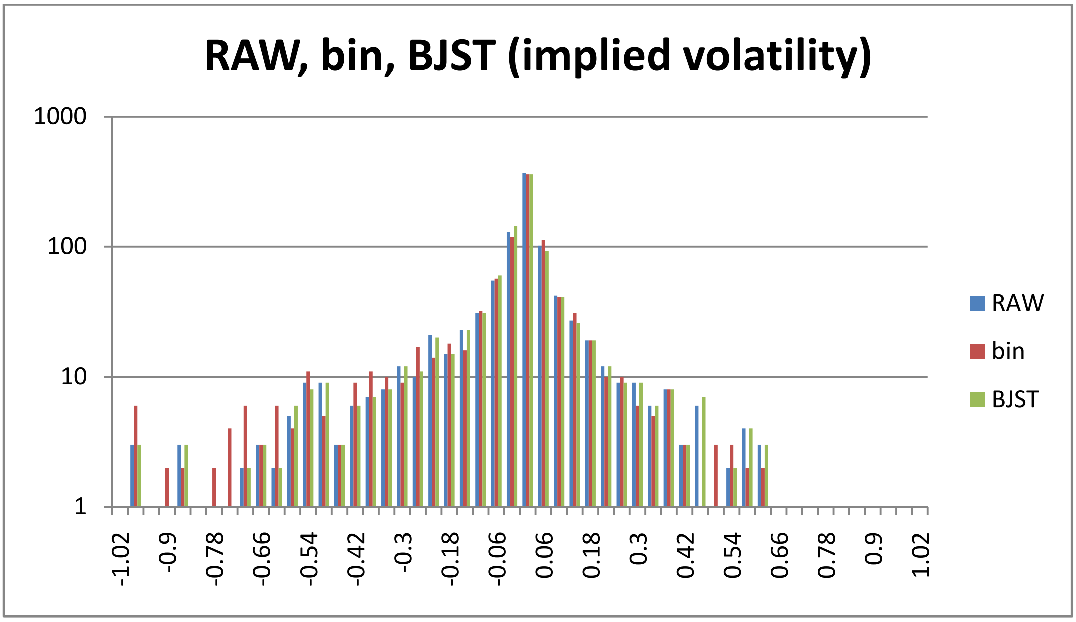

Fig. 16: *Log-histogram of relative deviances of the BAW, Binomial and BJST Model from the market data (random data, implied volatility)*

Remember that the scale cannot be interpreted in USD, but in the relative differences from the ask price. The histograms are quite similar to the previous ones, but the summary statistics slightly differ. Relative differences provide better R- squared values, but there are more mispriced options. Most of the options[52] which were mispriced absolutely by more than 1 USD were also mispriced relatively. Furthermore some options with low ask price were relatively mispriced too.

Since more than a third of the dataset are options cheaper than 1 USD and the relative pricing approach is penalizing them more than necessary we shall use absolute differences. This naturally brings a risk of considering expensive options as mispriced even if the estimated price is relatively close to the market one. This should not bring any serious bias since the results for absolute and relative differences are nearly the same.

4.4.4 Volatility Smile

There is a widely known observation that the implied volatility of the at the money options tends to be lower than the volatility of the options in the money or out of the money. This trend displays different characteristics for different markets because of the different probabilities of extreme moves.

[52] 50-80 % depending on the model.

Let's find out whether our available data poses the same characteristics. We have chosen all options issued on the P&G stock[53] available for trading at CBOE during the 31st July 2009. The ratio between the strike price and the spot price of the underlying was calculated for all options.

The plot of the data is below. There is implied volatility in dollars on the vertical axis and strike/underlying ratio on the horizontal one.

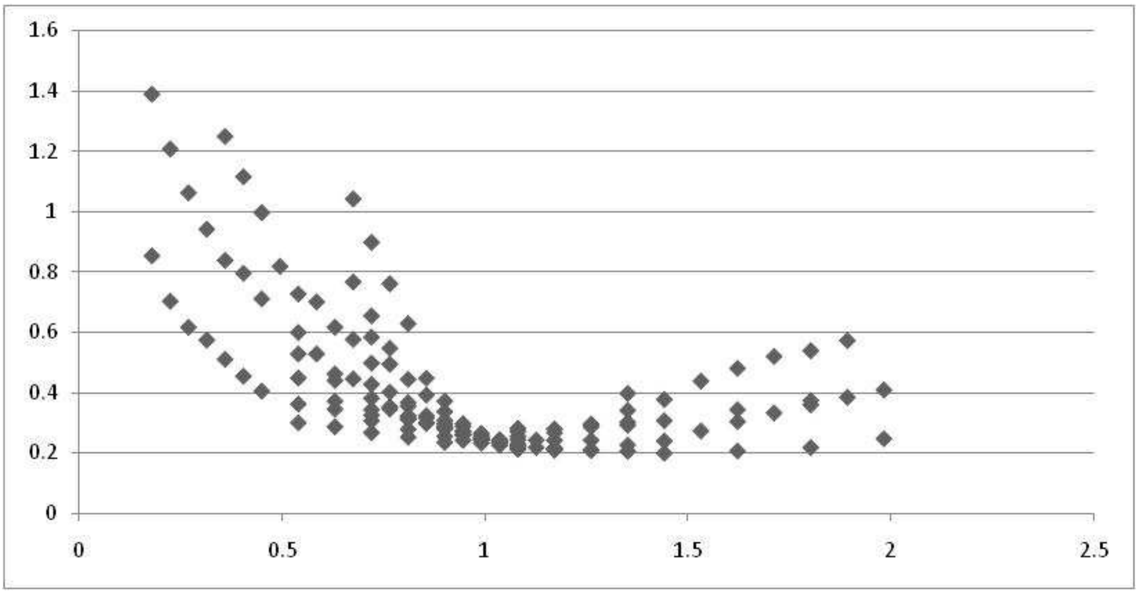

Fig. 17: *Volatility smile data*

We shall estimate "the smile" by polynomial regression assuming that there is a quadratic relationship, which means estimating the following equation by the OLS method.

$$IV = b_0 + b_1 SU + b_2 SU^2 + \varepsilon \,, \qquad \text{(EQ 4.4.4)}$$

where IV stands for the implied volatility and SU stands for the strike/underlying ratio.

The regression output is in the table below.

[53] This stock was also chosen randomly.

Model 1: OLS estimates using the 150 observations 1-150
Dependent variable: IV

	coefficient	std. error	t-ratio	p-value
const	1.31097	0.0598669	21.90	3.96e-048 ***
SU	-1.66115	0.119828	-13.86	1.78e-028 ***
SU2	0.638405	0.0553823	11.53	2.64e-022 ***

Mean dependent var	0.410205	S.D. dependent var	0.234439
Sum squared resid	3.014494	S.E. of regression	0.143202
R-squared	0.631899	Adjusted R-squared	0.626891
F(2, 147)	126.1733	P-value(F)	1.25e-32
Log-likelihood	80.19947	Akaike criterion	-154.3989
Schwarz criterion	-145.3670	Hannan-Quinn	-150.7296

White's test for heteroskedasticity -
Null hypothesis: heteroskedasticity not present
Test statistic: LM = 35.5581
with p-value = P(Chi-Square(4) > 35.5581) = 3.56717e-007

Test for normality of residual -
Null hypothesis: error is normally distributed
Test statistic: Chi-square(2) = 19.1092
with p-value = 7.08749e-005

Tab. 9: *Polynomial regression for the estimate of the volatility smile*

The estimates of the parameters are significant, the R-squared is quite high and the p-value for the F-test is very low. The autocorrelation test is not provided since the regression is run on the cross-sectional data. The multicollinearity is not tested since there are just two exogenous variables and the relationship between them is quadratic, not linear. The imperfections are that there is a serious heteroskedasticity and the residuals are not normally distributed.

Fortunately it is possible to use GLS to obtain an estimate which is not contaminated by heteroskedasticity. One of the easiest ways is to use “Heteroskedasticity corrected”[54] method in the Gretl software. The outputs are below.

[54] This method is just a WLS with the inverse values of residuals as weights.

Model 2: WLS estimates using the 150 observations 1-150
Dependent variable: IV

	coefficient	std. error	t-ratio	p-value	
const	1.32486	0.0900153	14.72	1.04e-030	***
SU	-1.76445	0.156343	-11.29	1.15e-021	***
SU2	0.719226	0.0662657	10.85	1.60e-020	***

Statistics based on the weighted data:

Sum squared resid	519.2548	S.E. of regression	1.879454
R-squared	0.470578	Adjusted R-squared	0.463375
F(2, 147)	65.33055	P-value(F)	5.01e-21
Log-likelihood	-305.9727	Akaike criterion	617.9455
Schwarz criterion	626.9774	Hannan-Quinn	621.6148

Statistics based on the original data:

Mean dependent var	0.410205	S.D. dependent var	0.234439
Sum squared resid	3.161715	S.E. of regression	0.146657

Test for normality of residual -
Null hypothesis: error is normally distributed
Test statistic: Chi-square(2) = 20.8323
with p-value = 2.99445e-005

Tab. 10: *Heteroskedasticity corrected method outputs*

The estimated equation is

$$\widehat{IV} = 1.32 - 1.76 * SU + 0.72 * SU^2. \qquad \text{(EQ 4.4.5)}$$

All parameters are statistically significant, the model explains nearly 50% of the original variance and it the p-value of the F-test is still very low.

The estimated smile is plotted in the figure below. The volatility smile is sometimes referred as volatility skew due to its characteristic shape. The problem with the non-normality of the residuals persists, so it would be better to use the MLE method, but the task here is just to show the relationship between the implied volatility and the strike/underlying ratio, which could be done despite of loosing asymptotic characteristics of the estimate.

The volatility smile implies that the prices of the options far in the money should be higher than the prices of the options at the money. This is not just because of a higher difference between the strike and the underlying price, but also because in the money options are expected to have higher volatility. Similarly the far out of the money options are not absolutely useless since there is a possibility of significant changes in the future, which could lead to the higher price of the option.

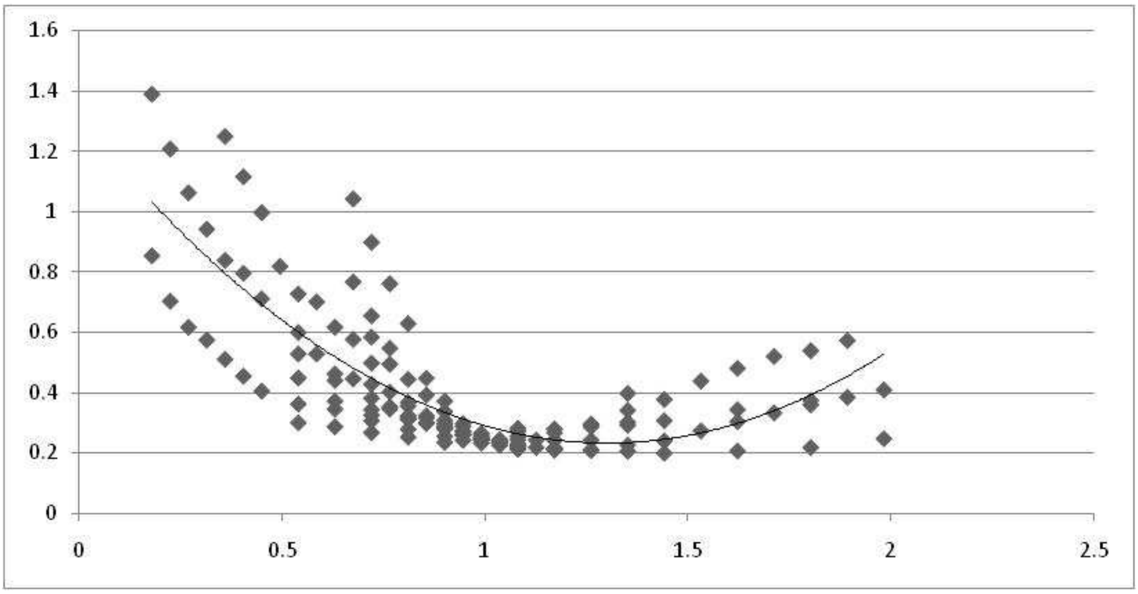

Fig. 18: *Volatility smile (skew) plot*

4.4.5 Examining Calls and Puts Separately

The midpoint of this part is determining whether the results are the same for call and put options or if they significantly differ.

Firstly there are given the results for call options. There are 479 of them, 420 were not traded during the day and 159 options have zero open interest. The bid ask spread is 0.52 USD, which is not statistically significant difference from the value of the full dataset since the standard deviation of the original bid ask spread is 0.75 USD. The time structure of the time to maturity did not change.

The following table contains the results for all calls in the dataset.

	BS	FBS	bin	BAW	BJST	JD
mean	0.00695	0.07468004	0.010943	0.010149	-0.08382	1.726445
median	0.002637	0.03986688	-0.00074	0.003041	-0.01467	0.027093
max	103.62	103.62	103.62	103.62	103.62	540.9037
min	-219.2	-219.2	-219.2	-219.2	-219.2	-219.2
Q1	-0.05111	-0.0226577	-0.05403	-0.04958	-0.11789	-0.0411
Q3	0.098297	0.16590129	0.097146	0.100387	0.058072	0.159366
stddev	12.13999	12.1424432	12.1408	12.14009	12.13983	28.44996
skew	-10.91021	-10.920025	-10.909	-10.9107	-10.8884	13.686
kurt	236.2743	236.326828	236.2253	236.2784	235.9585	278.9046
R-squared	0.77597	0.77587065	0.77594	0.775966	0.775965	-0.23489
mispriced	28	31	31	28	31	54
underpriced	11	11	13	11	17	15
overpriced	17	20	18	17	14	39

Tab. 11: *Differences between estimated and the market price (call options, implied volatility)*

The means and medians are nearly the same as for the models with all options. There are slightly more overpriced options than the underpriced ones. The R-squared values are little lower than for the general models. There is a negative skew when the first five models are used. These facts indicate that there is a difference between pricing calls and puts by used models will be observable but not significant.

The histograms are provided below.

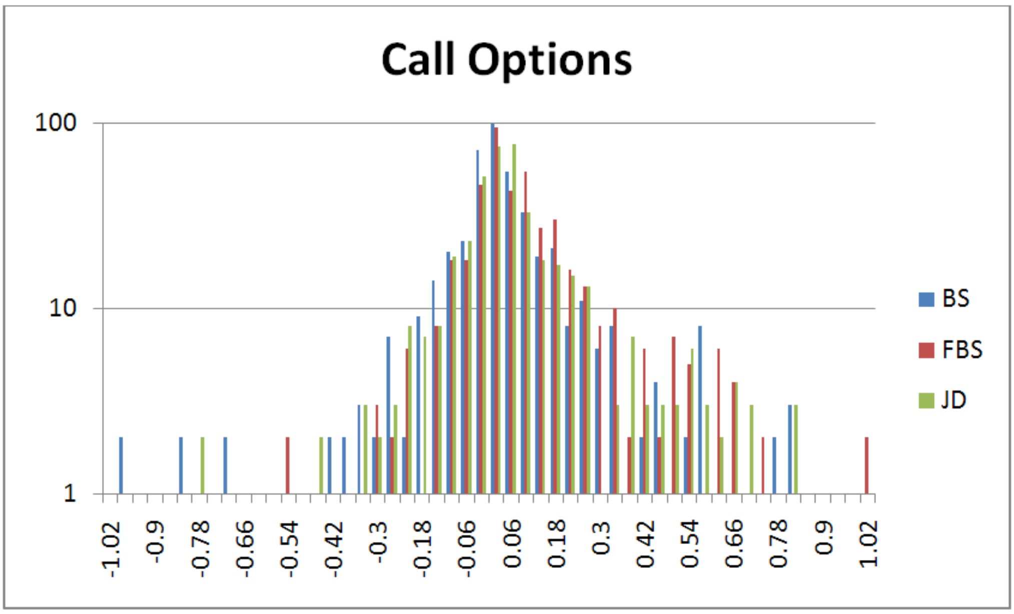

Fig. 19: *Log-histogram of deviances of the Black-Scholes, French Black-Scholes and Jump-Difussion Model from the market data (call options, implied volatility)*

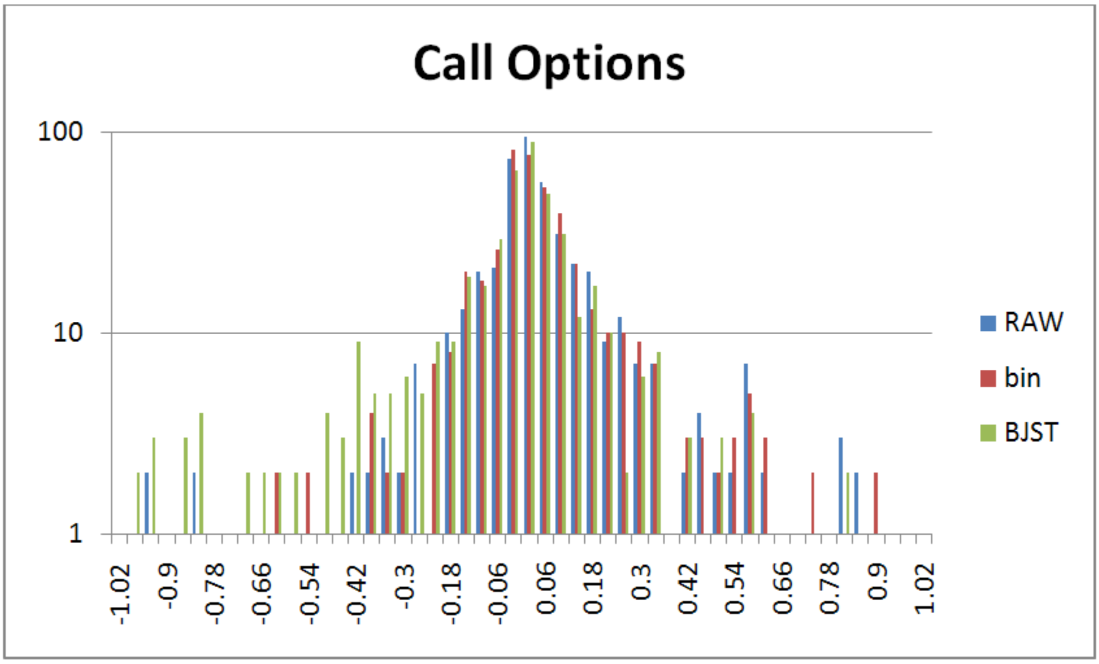

Fig. 20: *Log-histogram of deviances of the BAW, Binomial and BJST Model from the market data (call options, implied volatility)*

The modal interval for all models contains zero. The histograms are slightly different than for the general data, especially the BJST model brought different results than the rest of the models.

Now the same flow of facts will be provided for the put options. There are 521 options in the dataset, 441 were not traded during the day 155 of them had zero open interest. The average bid ask spread was 0.45 USD. The time structure of the time to maturity did not change.

The results are below.

	BS	FBS	bin	BAW	BJST	JD
mean	0.044265	0.07334008	0.043431	0.044265	0.044265	1.370961
median	-0.036441	-0.0171815	-0.03705	-0.03644	-0.03644	-0.01982
max	261.6552	261.48	261.6552	261.6552	261.6552	261.6552
min	-135	-135	-135	-135	-135	-135
Q1	-0.153096	-0.137426	-0.16324	-0.1531	-0.1531	-0.12805
Q3	0.025981	0.0506313	0.024397	0.025981	0.025981	0.078921
stddev	13.47208	13.4640458	13.47245	13.47208	13.47208	16.3078
skew	11.68797	11.6739443	11.68718	11.68797	11.68797	8.026463
kurt	295.1472	295.01884	295.1173	295.1472	295.1472	142.8852
R-squared	0.880618	0.88075857	0.880612	0.880618	0.880618	0.823838
mispriced	20	22	21	20	20	47
underpriced	12	12	12	12	12	13
overpriced	8	10	9	8	8	34

Tab. 12: *Differences between estimated and the market price (put options, implied volatility)*

This is an opposite story to the call options so we shall provide just brief comments. Means and medians are nearly zero. Quartile spreads contain zero, but they are not symmetric about it. The skewness is positive for the data. The values of R-squared are higher than for the general model. The skew is positive for all models

The histograms are below.

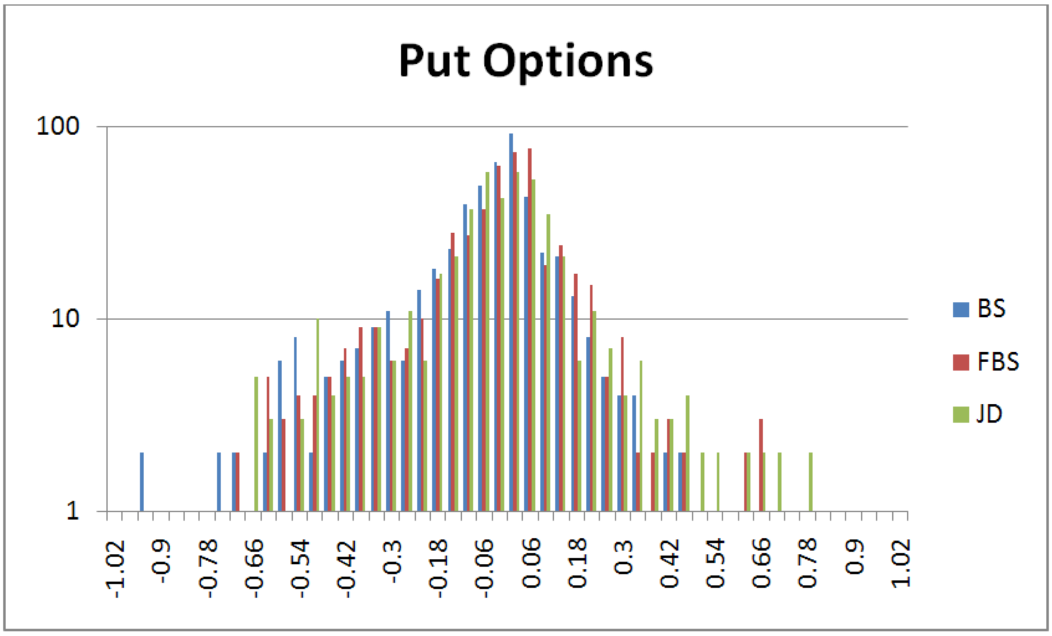

Fig. 21: *Log-histogram of deviances of the Black-Scholes, French Black-Scholes and Jump-Difussion Model from the market data (put options, implied volatility)*

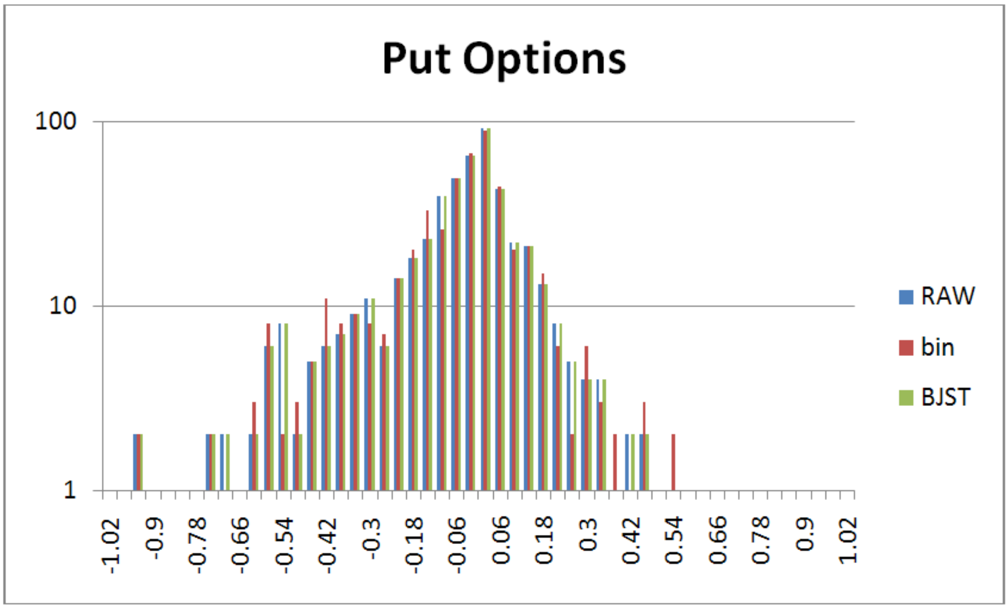

Fig. 22: *Log-histogram of deviances of the BAW, Binomial and BJST Model from the market data (put options, implied volatility)*

The modal intervals contain zero. The French Black-Scholes and the binomial model gave slightly different results. It was evident from the beginning that the histogram for all options will be the sum of the separate histograms for the put and call options, but an interesting fact is that there is a different skew for both datasets and that the put options are priced little more accurately.

4.5 Influence of the Traded Volume on the Pricing Methods

For this part were chosen two subsets – 1000 options with zero traded volume and 1000 options with the highest traded volume during the day. The aim of this analysis is to determine whether the traded volume positively or negatively influences the option pricing method.

4.5.1 Most Traded Options

The options in this dataset were mostly traded from the whole day[55]. There are 561 call options and 439 puts. 481 options were in the money, 110 out of the money and 409 could be considered as at the money. For 751 options an immediate issuance would

[55] It means that the traded volume was the highest for them.

bring a profit. The average expiration time was 66 days. The median expiration time[56] was 21 days. The average bid ask spread was 0.14.

The essential characteristics are provided in the following table.

	BS	FBS	Bin	BAW	BJST	JD
mean	-0.055301	-0.0001239	-0.04542	-0.05512	-0.05587	0.043416
median	-0.006151	0.01683436	-0.00123	-0.00615	-0.00652	-0.03854
max	3.518018	3.72237331	3.538522	3.518198	3.51579	168.0132
min	-67.13519	-67.193568	-66.3359	-67.1352	-67.1352	-67.1165
Q1	-0.067211	-0.0344343	-0.05977	-0.06721	-0.06782	-0.13637
Q3	0.069687	0.10398997	0.076445	0.070019	0.068812	0.032849
stddev	2.156182	2.15863925	2.132809	2.156144	2.156151	5.771398
skew	-30.14102	-30.185599	-30.0558	-30.1428	-30.1415	23.13035
murt	938.4983	940.586076	934.9727	938.5743	938.5206	736.5321
R-squared	0.906279	0.90612725	0.908319	0.906283	0.906281	0.328933
mispriced	25	17	23	25	25	50

Tab. 13: *Differences between estimated and the ask price (most traded options, implied volatility)*

According to the results the options were underestimated by all models except for the JD model. The reason, why the market price is a little higher than the estimated price, could be that the high demand for those options is forcing market prices up. The Binomial Model possess the most favorable R-squared and median value. The lowest mean in the absolute value and also the lowest number of mispriced options occurred for the French Black-Scholes Model. Another interesting fact is the distribution of mispriced options, it is provided in the table below.

	BS	FBS	Bin	BAW	BJST	JD
underpriced	13	3	11	13	13	32
overpriced	12	14	12	12	12	18

Tab. 14: *Distribution of the mispriced options (most traded options, implied volatility)*

The distribution of mispriced options is the same for BS, BAW and BJST models; the binomial model performed very similarly. On the other hand the ratio of underpriced and overpriced options changed from approximately 1:1 for the four previously mentioned models to more than 1:4 for the FBS model. The good performance of the FBS model could be explained by the fact that the differentiation between trading days and days to expiration is more significant for options closer to maturity.

[56] And also the minimum.

The histograms are below.

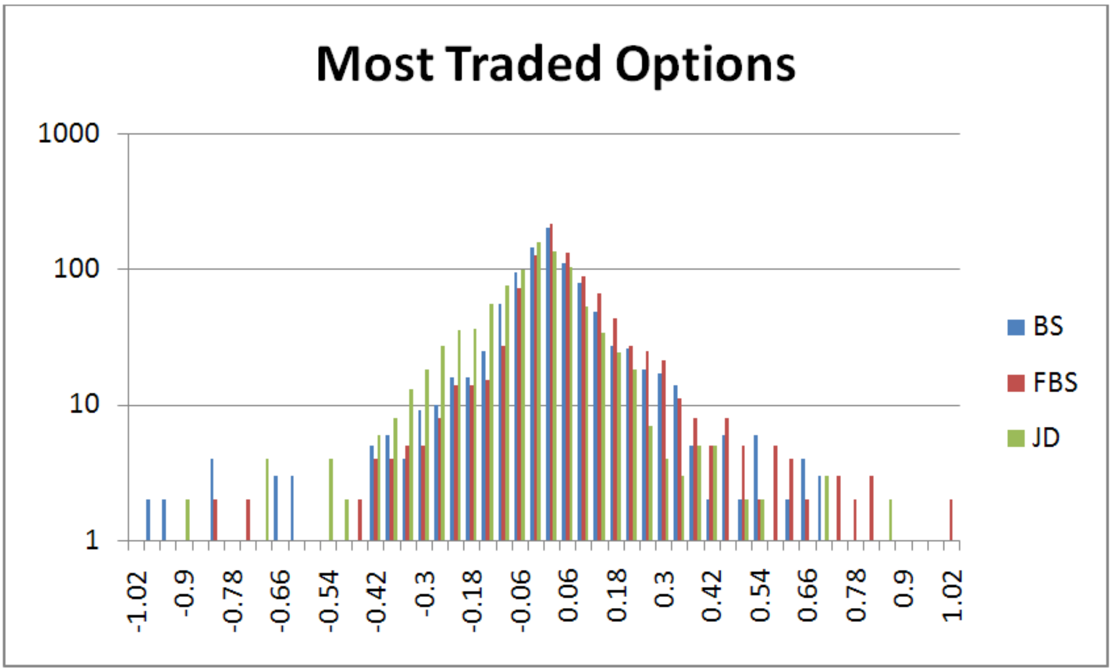

Fig. 23: *Log-histogram of deviances of the Black-Scholes, French Black-Scholes and Jump-Difussion Model from the market data (most traded options, implied volatility)*

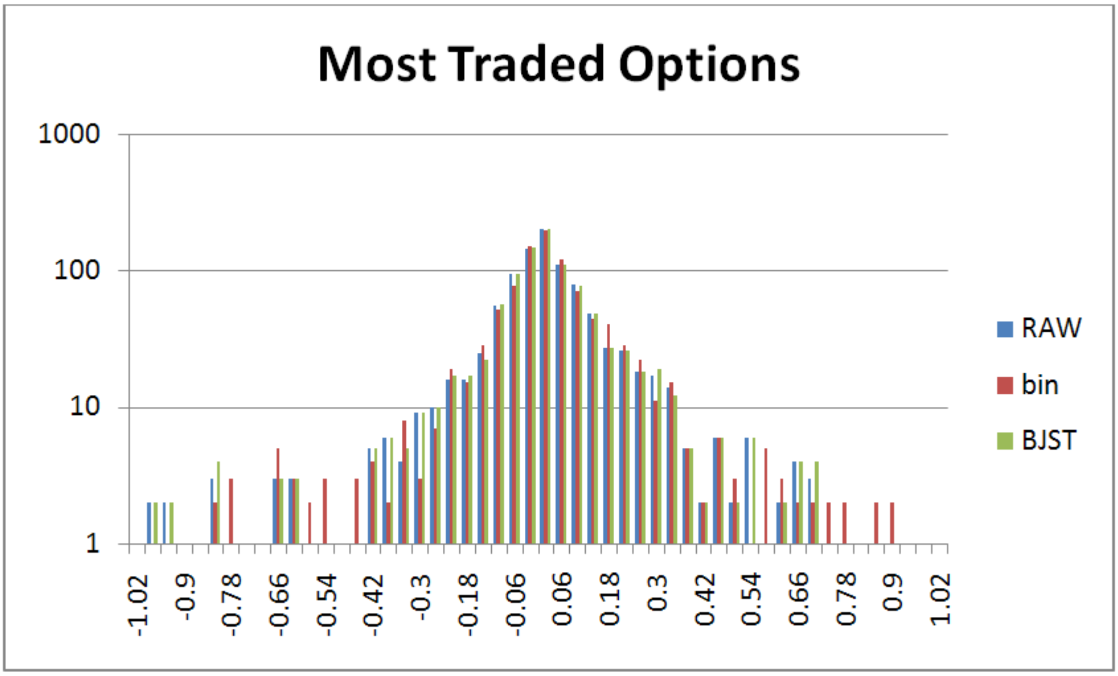

Fig. 24: *Log-histogram of deviances of the BAW, Binomial and BJST Model from the market data (most traded options, implied volatility)*

More than 200 options were priced with less than 0.02 USD discrepancy by 5 most successful models. Note the negative skew of the data and much higher kurtosis than for the general sample.

4.5.2 Non-Traded Options

Now the non-traded options will be analyzed. Since there are more than 1000 options which were not traded during 31st July 2009 we used a random generator to select 1000 of them as an illustrative sample.

There are 461 calls and 539 puts in the dataset. All of them were not traded, but 394 of them have positive open interest. The average time to expiration is 158 days. 454 options were out of the money, 465 were in the money and just 81 were at the money. The average bid ask spread was 0.62 USD.

The table showing the results of the evaluation follows.

	BS	FBS	Bin	BAW	BJST	JD
mean	0.12645	0.18185988	0.130665	0.129078	0.073862	4.219481
median	-0.010489	0.01080482	-0.01059	-0.01017	-0.0214	0.017831
max	344.8099	344.8099	344.8099	344.8099	344.8099	897.2187
min	-182.6	-182.6	-182.6	-182.6	-182.6	-182.6
Q1	-0.085468	-0.0601499	-0.08889	-0.0842	-0.14125	-0.07287
Q3	0.063809	0.11268238	0.070627	0.067636	0.0376	0.106928
stddev	13.81509	13.8162952	13.81525	13.81511	13.81603	42.56029
skew	11.2844	11.2696799	11.28314	11.28381	11.2928	13.27062
kurt	444.2823	443.946108	444.2488	444.2714	444.332	230.3686
R-squared	0.770638	0.77057767	0.770632	0.770637	0.77062	-1.19803
mispriced	52	56	50	51	59	106
underpriced	27	24	27	26	35	31
overpriced	25	32	23	25	24	75

Tab. 15: *Differences between estimated and the ask price (least traded options, implied volatility)*

The averages suggest that the models overestimate option prices, but one could assume that this is caused just by a few outliers since the medians are nearly zero. Note that according to the R-squared the JD model really failed[57], all other models gave similar

[57] Note that this is not a linear regression so the R-squared can be negative. This must be caused by those 75 overpriced options. It would still be better to use this model, because it fitted about 900 options quite well.

results according to the R-squared. The results suggest that the BJST model could be the best fit because of the lowest mean and the second most favorable quartile spread.

Below could be found histograms.

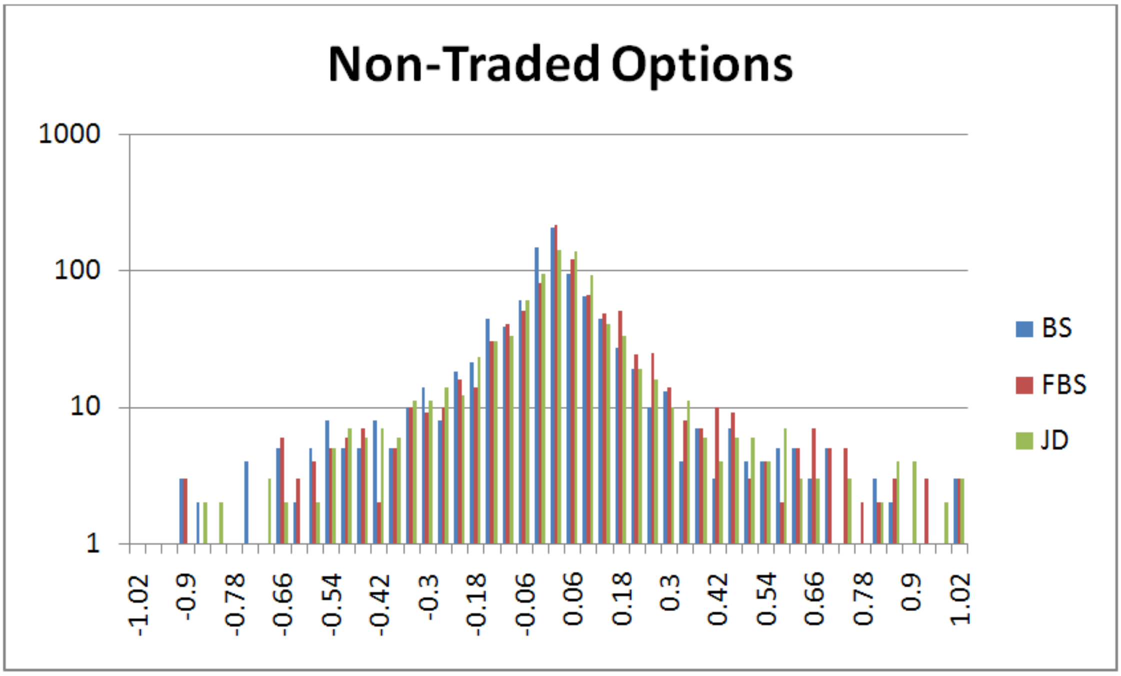

Fig. 25: *Log-histogram of deviances of the Black-Scholes, French Black-Scholes and Jump-Difussion Model from the market data (non-traded options, implied volatility)*

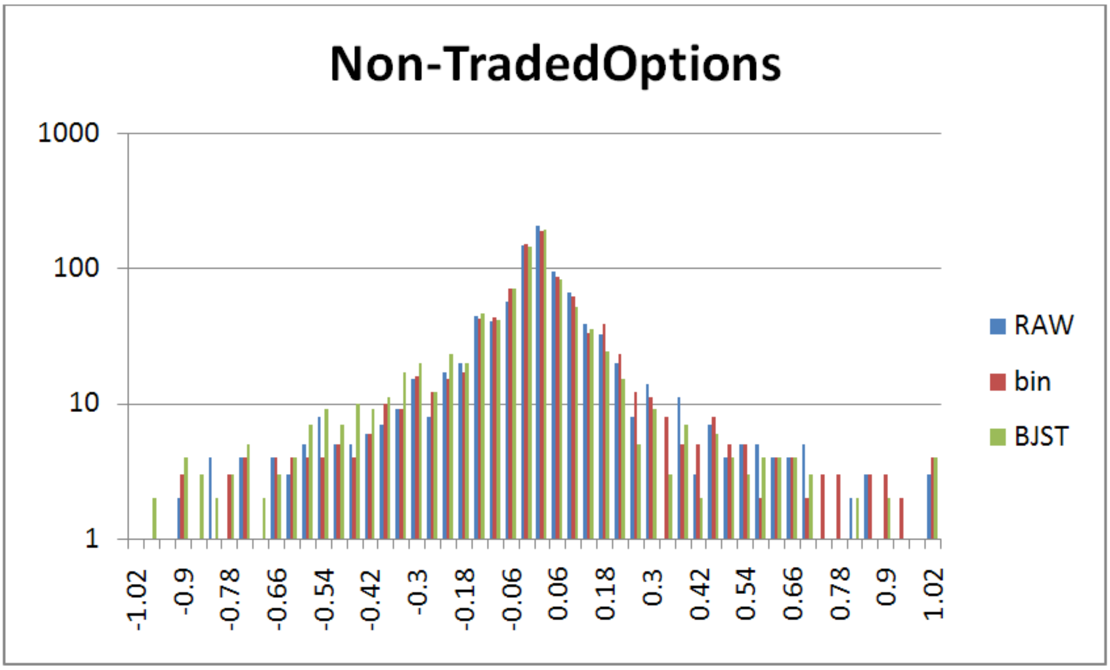

Fig. 26: *Log-histogram of deviances of the BAW, Binomial and BJST Model from the market data (non-traded options, implied volatility)*

The JD Model estimated 142 options with 0.02USD accuracy, other models varied from 187 (bin) to 213 (FBS). The skew is positive, but close to zero which means, that the histograms are nearly symmetrical, also the mispriced options are uniformly distributed.

4.5.3 Observed Differences

The option pricing models generally fit more traded options better than the non-traded ones. The most traded options are usually very close to its expiration and their bid ask spread is lower than for the non-traded ones. There are much more options at the money in the most traded options. The French Black-Scholes seems to be the best alternative to price the most traded options; on the other hand the BJST model provides the best results for the non traded options. No difference has been proven between call and put options estimated prices.

An interesting fact is that a model for pricing European options provides the best results for price American options. This could be caused by the previously mentioned observation that more than a half of the most traded options expired in three weeks – the real differences between European and American option are minimal in such a short time. On the other hand the BJST model was confirmed to be able to precisely price American options with long time to expiration.

4.6 Influence of the Time to Expiration on the Pricing Methods

For this part were chosen 1000 options with the shortest time to maturity and 1000 options with the longest time to maturity. The aim is to find the impact of this to the analyzed option pricing methods.

4.6.1 Options with Short Time to Expiration

All of these options have 21 days to expiration, so there were obviously more than 1000 such options in the original dataset and it was necessary to implement a randomly choosing algorithm. There were chosen 504 call options and 496 puts, 771 of them were not traded during the day and 346 had zero open interest. The average bid ask spread was approximately 0.5 USD. 438 options were out of the money, 440 in the money and the rest at the money.

The table with results is provided below.

	BS	FBS	bin	BAW	BJST	JD
mean	0.089653	0.10923952	0.085131	0.089738	0.076584	0.095379
median	-0.027658	-0.0208628	-0.03589	-0.0276	-0.0297	-0.00426
max	62.5	62.5	62.5	62.5	62.5	62.49275
min	-23.8	-23.8	-23.8	-23.8	-23.8	-23.8
Q1	-0.102888	-0.088573	-0.11058	-0.10282	-0.10885	-0.09565
Q3	0.024467	0.04095275	0.026229	0.024542	0.01766	0.059597
stddev	2.730656	2.73065144	2.734544	2.730654	2.729761	2.741958
skew	14.05429	14.0435537	14.00428	14.0543	14.06691	13.84211
kurt	302.8545	302.483585	301.2438	302.8542	303.48	297.6291
R-squared	0.996335	0.99633297	0.996325	0.996335	0.996338	0.996304
mispriced	63	64	65	63	63	66
underpriced	36	35	37	36	37	36
overpriced	27	29	28	27	26	30

Tab. 16: *Differences between estimated and the ask price (3 weeks to expiration, implied volatility)*

The lowest median, quartile spread and standard deviation were observed for the French Black-Scholes Model. The BJST model posses the best value of R-squared and the mean difference between the estimated and the market price.

The histograms follow.

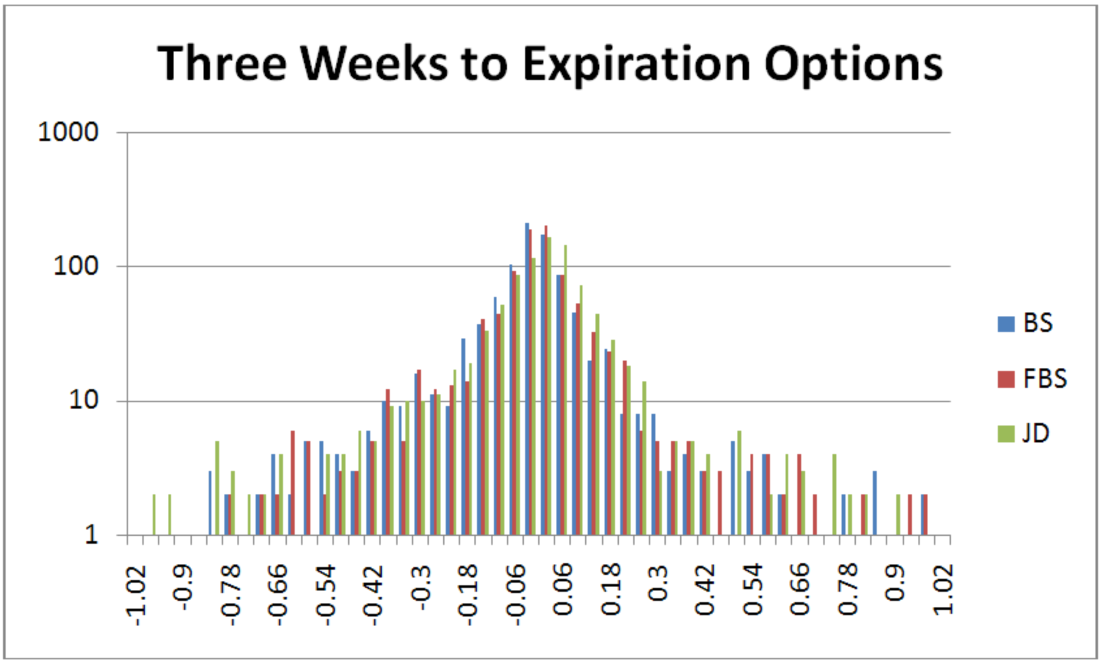

Fig. 27: *Log-histogram of deviances of the Black-Scholes, French Black-Scholes and Jump-Difussion Model from the market data (three weeks to expiration, implied volatility)*

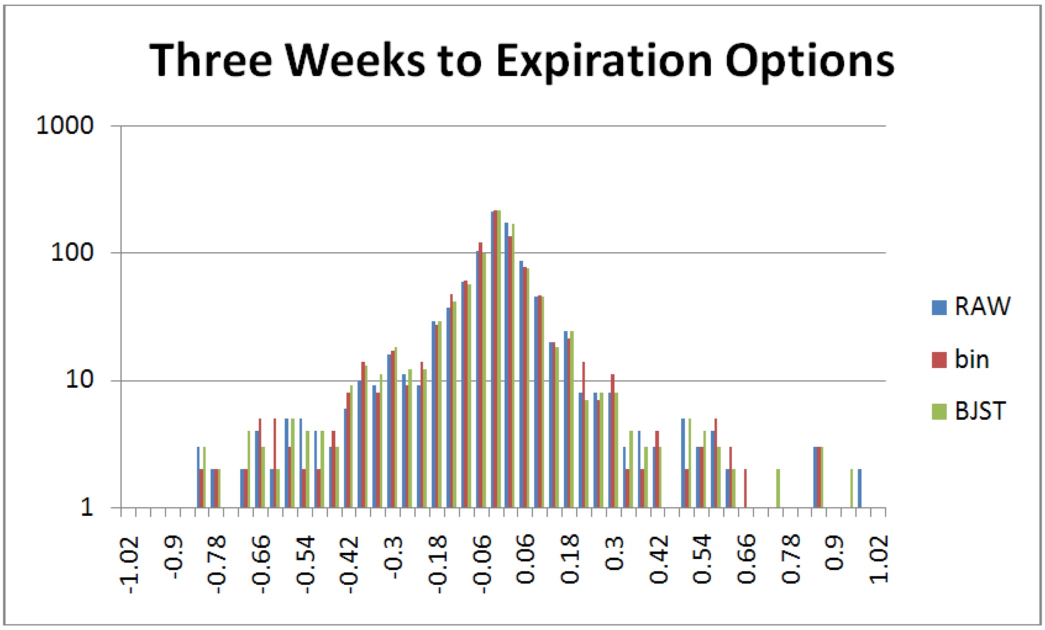

Fig. 28: *Log-histogram of deviances of the BAW, Binomial and BJST Model from the market data (3 weeks to expiration, implied volatility)*

Note that the interval with the highest frequency of observations does not contain zero for most of the models. The binomial model with 216 observations has the highest density in the modal interval (-0.06;-0.02). The JD Model is the only one for which the modal interval contains zero. The histograms are nearly symmetrical, but there is a slight positive skew.

4.6.2 Options with Long Time to Expiration

For this part was obtained the subset containing exactly 500 calls and 500 puts, 956 of them were not traded and 524 had zero open interest. Most of the options in this subset have 686 or 868 days to expiration. The average bid ask spread is approximately 1.5 USD. 454 options are in the money, 457 are out of the money and 89 are at the money.

The results of all pricing methods are below.

	BS	FBS	bin	BAW	BJST	JD
mean	-1.041866	-0.7186147	-0.95984	-1.00239	-1.13173	273.5027
median	-0.027812	0.16897666	-0.00342	-0.01549	-0.05514	85.90407
max	47.0192	46.96	47.0192	47.0192	47.0192	3107.826
min	-156.58	-154.77847	-154.292	-158.009	-156.445	-145.641
Q1	-0.172661	-0.1086713	-0.1665	-0.16365	-0.27787	19.60119
Q3	0.094956	0.66391546	0.235344	0.13812	0.057805	291.1828
stddev	15.7856	15.7806073	15.74759	15.80273	15.78386	453.1945
skew	-6.313089	-6.4267394	-6.29948	-6.32659	-6.29573	2.646384
kurt	50.03736	50.7125655	49.79084	50.22247	49.88389	7.723117
R-squared	0.98612	0.98616039	0.986196	0.986095	0.986112	-14.5392
Mispriced	131	224	146	132	154	985
underpriced	63	68	62	60	86	12
Overpriced	68	156	84	72	68	973

Tab. 17: *Differences between estimated and the ask price (long time to expiration, implied volatility)*

The most evident fact is that a Jump-Diffusion Model failed to price these options. The other models underpriced the options by an average about 1 USD. According to the median the best model was the binomial one. Quite good performance showed also BAW Model with the lowest quartile spread and the second lowest number of mispriced options.

It is reasonable to plot the log-histograms with -20 to 20 USD scale because of the high number of mispriced options.

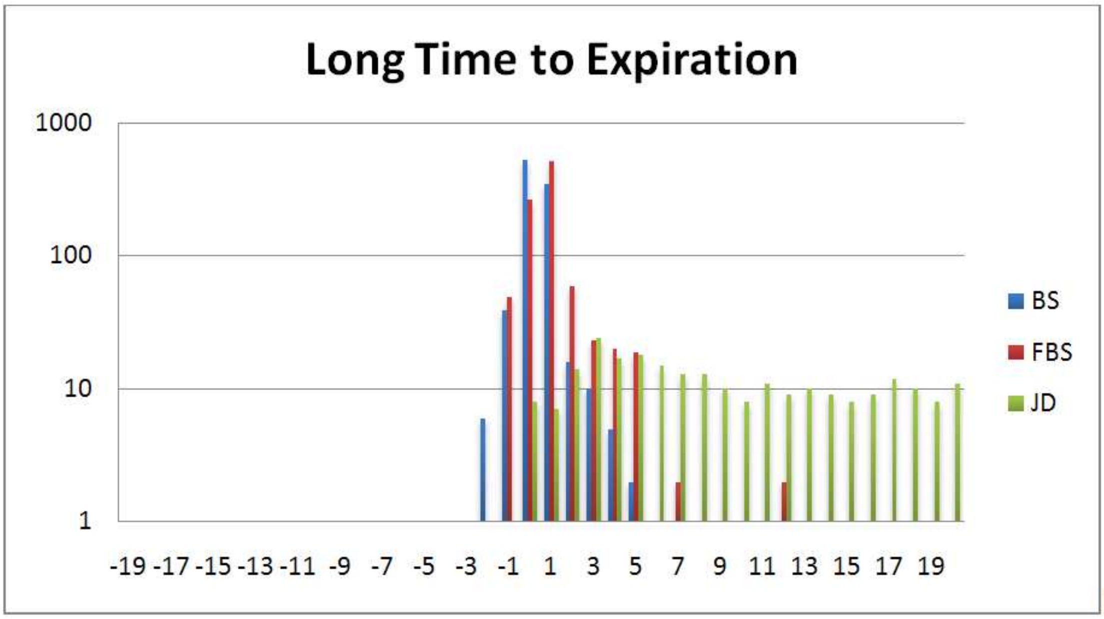

Fig. 29: *Log-histogram of deviances of the Black-Scholes, French Black-Scholes and Jump-Difussion Model from the market data (long time to expiration, implied volatility)*

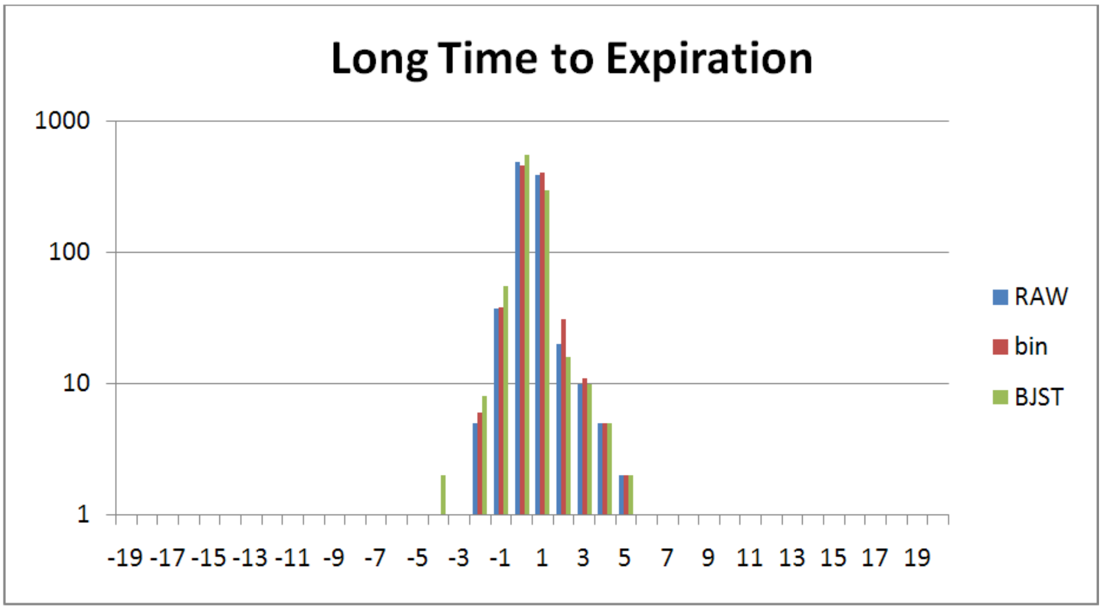

Fig. 30: *Log-histogram of deviances of the BAW, Binomial and BJST Model from the market data (3 weeks to expiration, implied volatility)*

The figures above clearly show the central tendency of all models except for the Jump-Diffusion one. Since the high kurtosis of the histogram is clear the more precise histograms even for this data are provided.

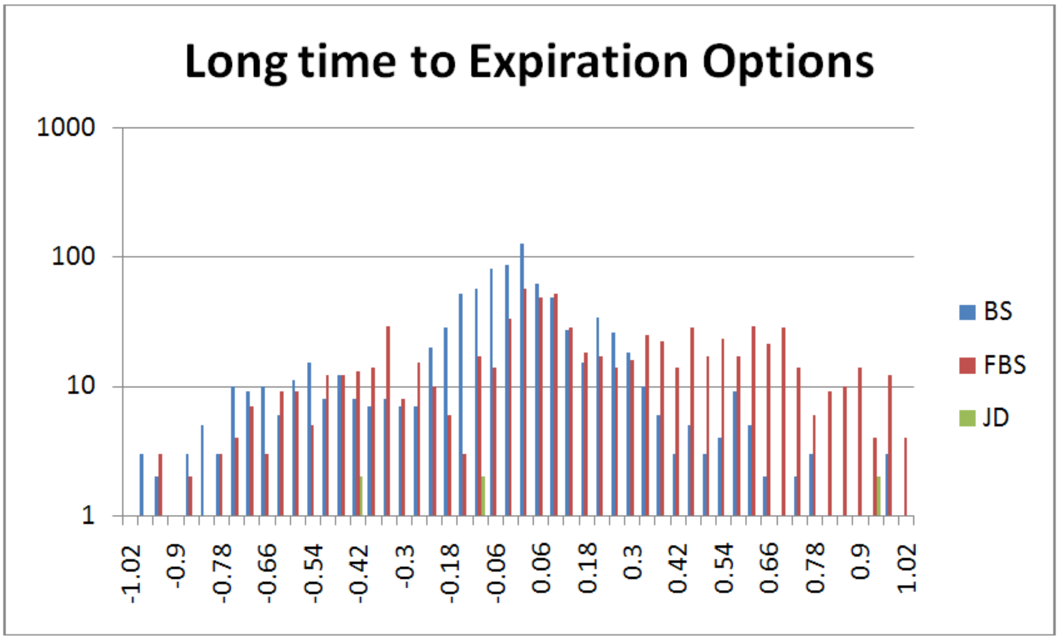

Fig. 31: *Log-histogram of deviances of the Black-Scholes, French Black-Scholes and Jump-Difussion Model from the market data (long time to expiration, implied volatility)*

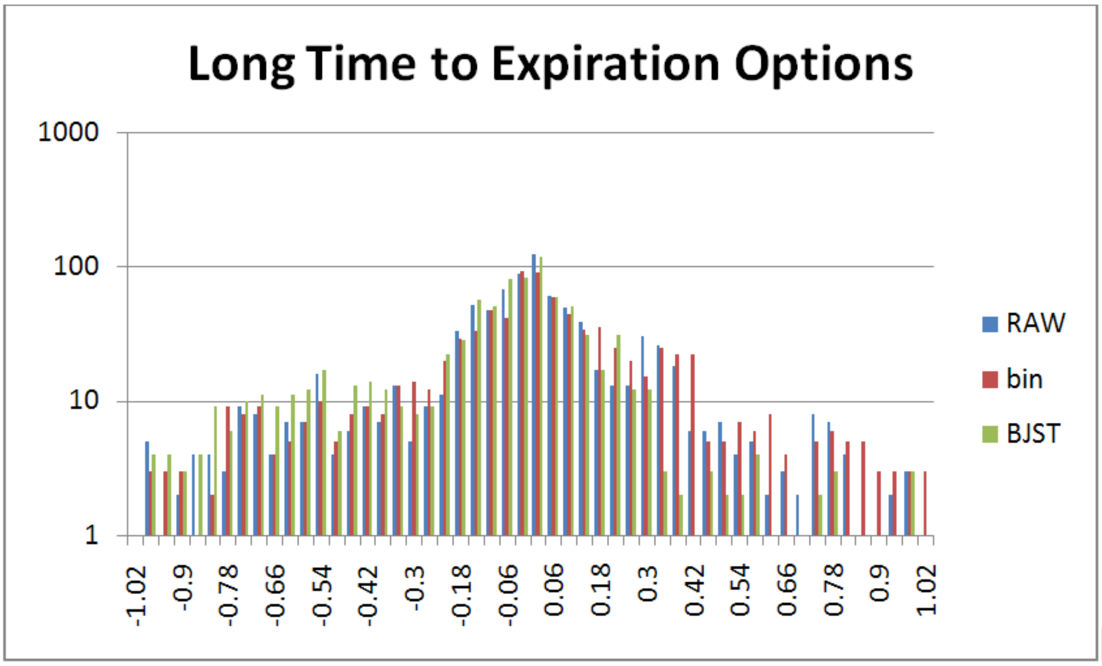

Fig. 32: *Log-histogram of deviances of the BAW, Binomial and BJST Model from the market data (long time to expiration, implied volatility)*

The Fig. 31 shows the largest discrepancy between the Black-Scholes and French Black-Scholes models observed here. It means that the French Black-Scholes formula is much more suitable for pricing options closer to expiration. The difference between pricing formulas for American and European options should be the most evident just now. Note that the kurtosis reaches its lowest values in comparison to the other models.

4.6.3 Observed Differences

The options with short time to expiration are obviously traded more than the options with long time to expiration. The average bid ask spread for options far from the expiration is approximately three times higher than for the options close to the expiration. The distribution of options at the money, out of the money and in the money is nearly the same for both subgroups.

The most significant differences were observed during an application of the pricing methods. The Jump-Diffusion model did very well for the first group but has absolutely failed to price options far from the expiration. Other methods performed not as bad as the JD and the results were generally better for the options close to the expiration. The kurtosis is much higher for the first group than for the second one. In the first group, there was also observed a positive skew. On the contrary the negative skew appeared in the second subset of the data.

An interesting fact is that the French Black-Scholes model did not perform significantly better for the options with shorter time to expiration. Analogically the BJST model did not out perform other models when dealing with options far from the expiration.

4.7 Influence of the Relationship between the Underlying and the Strike Price on the Pricing Methods

In this part will be analyzed the effect of an option being in the money, out of the money or just at the money to the option pricing. There will be provided three subsets for the proper determination of the impacts of the option states on the pricing methods.

4.7.1 In the Money Options

There are 497 calls and just 503 puts in this subset. The average time to maturity is 236 days, but the median is 168. The average bid ask spread is 1.95 USD. 977 options from this subset were not traded and 672 of them have zero open interest. Those options are not attractive to buy because the possible profit, which could be made from the immediate expiration, is overweighed by the high price.

The results of evaluating formulas are provided below.

	BS	FBS	bin	BAW	BJST	JD
mean	44.7904	44.7730499	44.78742	44.7931	44.77763	57.39062
median	-0.052533	-0.0711354	-0.05718	-0.05	-0.07742	0.10115
max	2785.93	2785.93	2785.93	2785.93	2785.93	2782.748
min	-24.075	-24.075	-24.075	-24.075	-24.075	-24.0854
Q1	-0.397876	-0.435	-0.40323	-0.39788	-0.40832	-0.22903
Q3	13.3941	13.3941	13.3941	13.3941	13.3941	26.74099
stddev	176.8741	176.878484	176.8748	176.8734	176.8773	191.2826
skew	8.752299	8.75192027	8.752234	8.752357	8.752021	7.228897
kurt	111.4247	111.416745	111.4233	111.4259	111.4189	80.29721
R-squared	0.098163	0.09816308	0.098163	0.098163	0.098163	-0.08042
mispriced	407	410	407	407	407	548
underpriced	118	121	118	118	118	107
overpriced	289	289	289	289	289	441

Tab. 18: *Differences between estimated and the ask price (in the money options, implied volatility)*

Just a short look at the table above should indicate that all pricing methods provide worse results than in all previous cases. On the other hand high mean differences and nearly zero medians indicate that the problem lies in extremely high maximum values. The amounts of mispriced options are also alarming. Note that there are more than two times more overpriced options than the underpriced ones.

The log-histograms follow.

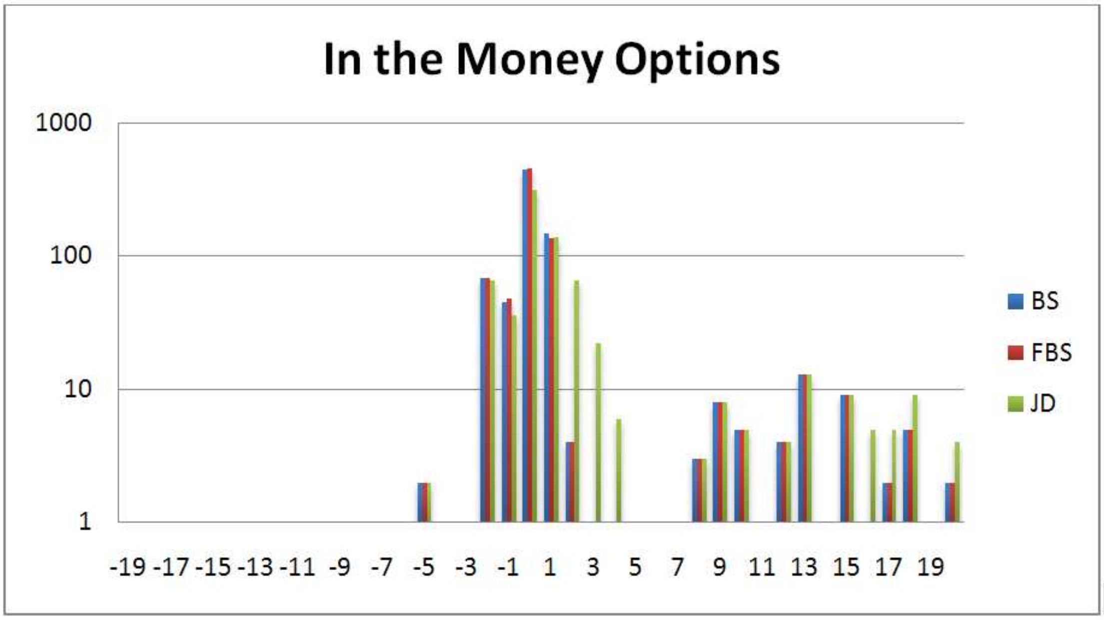

Fig. 33: *Log-histogram of deviances of the Black-Scholes, French Black-Scholes and Jump-Difussion Model from the market data (in the money options, implied volatility)*

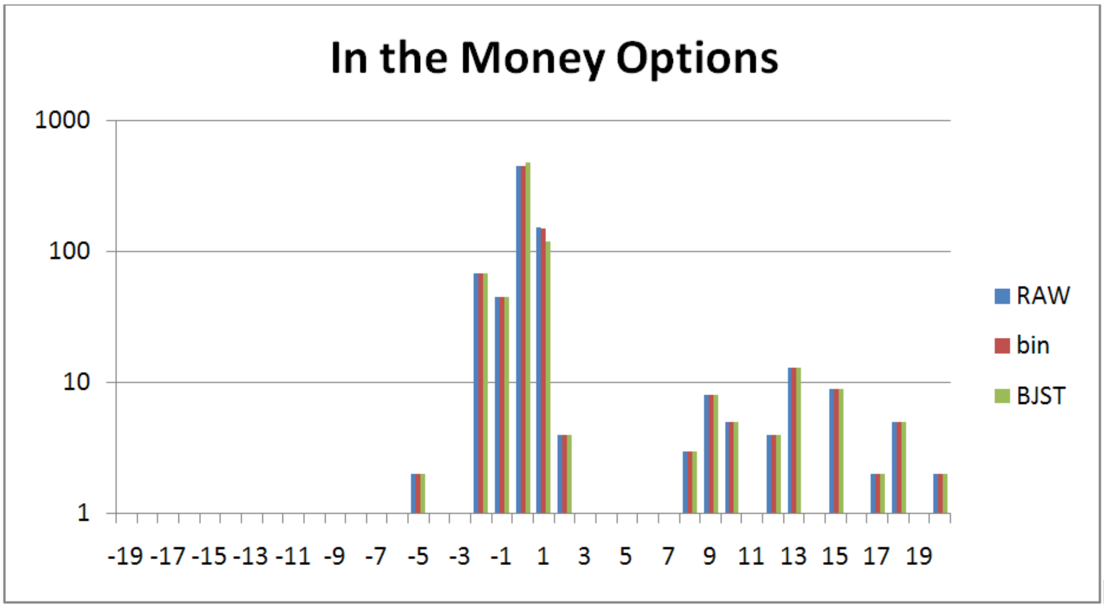

Fig. 34: *Log-histogram of deviances of the BAW, Binomial and BJST Model from the market data (in the money options, implied volatility)*

It looks like all log-histograms are bimodal, but There are still about 500 options priced so precisely that the difference from the market price is lower than 1 USD. So it seems that it is rational to look at the more detailed histograms, which are provided below.

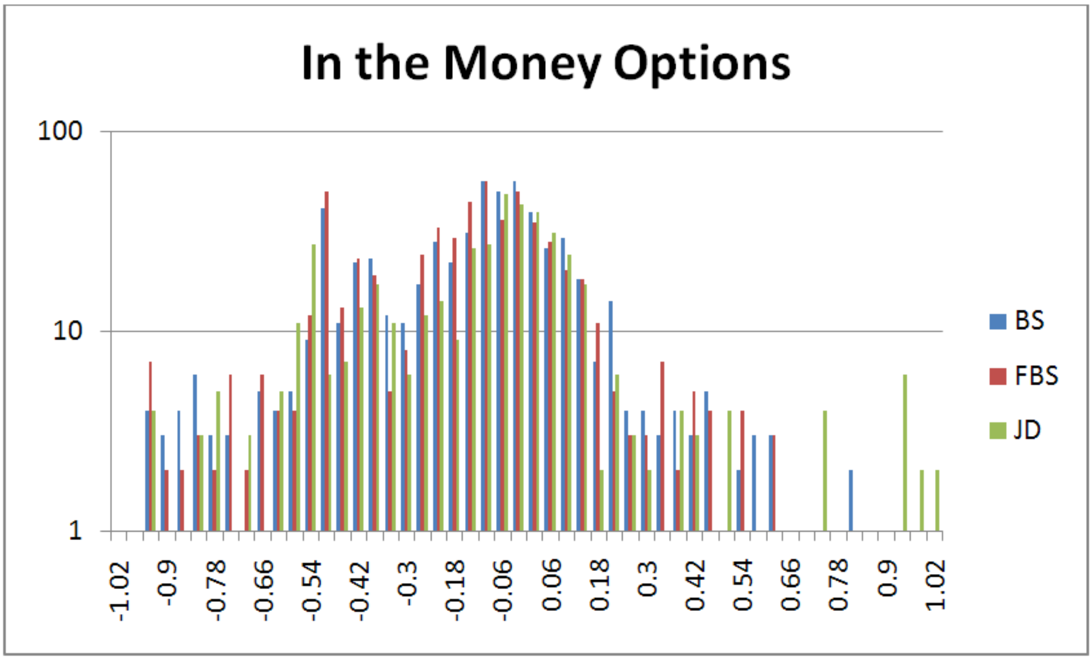

***Fig. 35:** Log-histogram of deviances of the Black-Scholes, French Black-Scholes and Jump-Difussion Model from the market data (in the money options, implied volatility)*

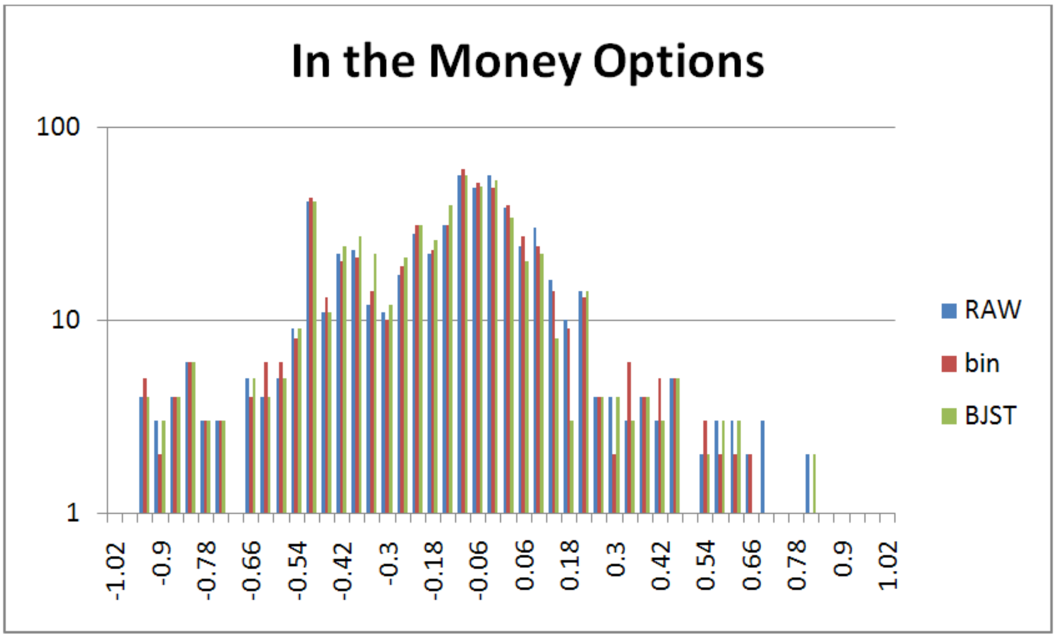

***Fig. 36:** Log-histogram of deviances of the BAW, Binomial and BJST Model from the market data (in the money options, implied volatility)*

One could conclude that the central tendency to price the options very close to the market price holds even for this subset, but the averages are contaminated by heavy tail on the positive side of the spectrum.

4.7.2 Out of the Money Options

For this part were chosen 1000 out of the money options exercising those options would bring immediate loss for its holder. There are 497 calls and 503 puts. 957 options from this subset were not traded and 344 of them had zero open interest. The average bid ask spread was 1.43 USD. The average time to expiration was 216 days, the median value 168 days.

The results are below.

	BS	FBS	bin	BAW	BJST	JD
mean	-12.06073	-12.053267	-12.0649	-12.0607	-12.0618	-11.9341
median	-0.020739	-0.0118357	-0.03175	-0.02053	-0.02096	0.002035
max	0.239465	0.26842208	0.239668	0.239768	0.23647	13.49368
min	-292.8	-292.8	-292.8	-292.8	-292.8	-292.8
Q1	-0.75	-0.75	-0.75	-0.75	-0.75	-0.75
Q3	-0.002552	0.00415642	-0.00676	-0.00254	-0.00259	0.039275
stddev	43.1569	43.1589774	43.15575	43.15691	43.15661	43.19791
skew	-4.351191	-4.3510423	-4.35127	-4.35119	-4.35121	-4.34665
kurt	19.19525	19.194016	19.19593	19.19524	19.19541	19.16258
R-squared	-0.065315	-0.065315	-0.06532	-0.06532	-0.06532	-0.06558
mispriced	227	227	227	227	227	255
underpriced	227	227	227	227	227	228
overpriced	0	0	0	0	0	27

Tab. 19: *Differences between estimated and the ask price (out of the money options, implied volatility)*

From the table above is clear that all models underprice the options. On the other hand the medians are also very close to zero. So one could suggest that it is going to be just an opposite situation than before, but note that the first 5 models did not overprice any option.

The following figures contain log-histograms.

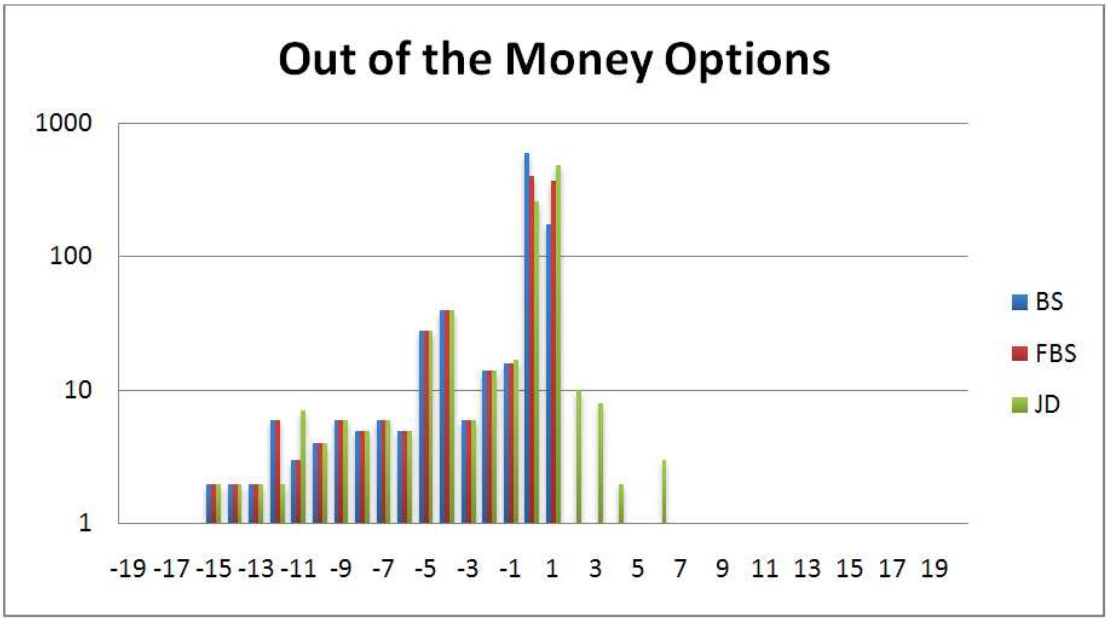

Fig. 37: *Log-histogram of deviances of the Black-Scholes, French Black-Scholes and Jump-Difussion Model from the market data (out of the money options, implied volatility)*

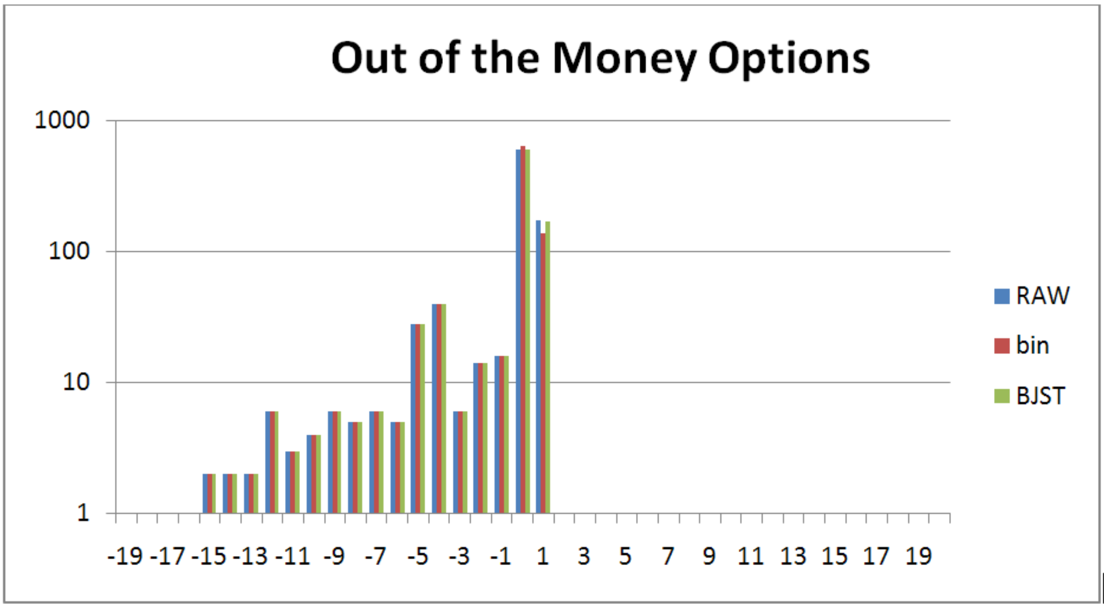

Fig. 38: *Log-histogram of deviances of the BAW, Binomial and BJST Model from the market data (out of the money options, implied volatility)*

There is a clear negative skew in the data. In spite of the biased averages more than 700 options were priced with just 1 USD deviation from the market price. So the more detailed log-histograms are plotted below.

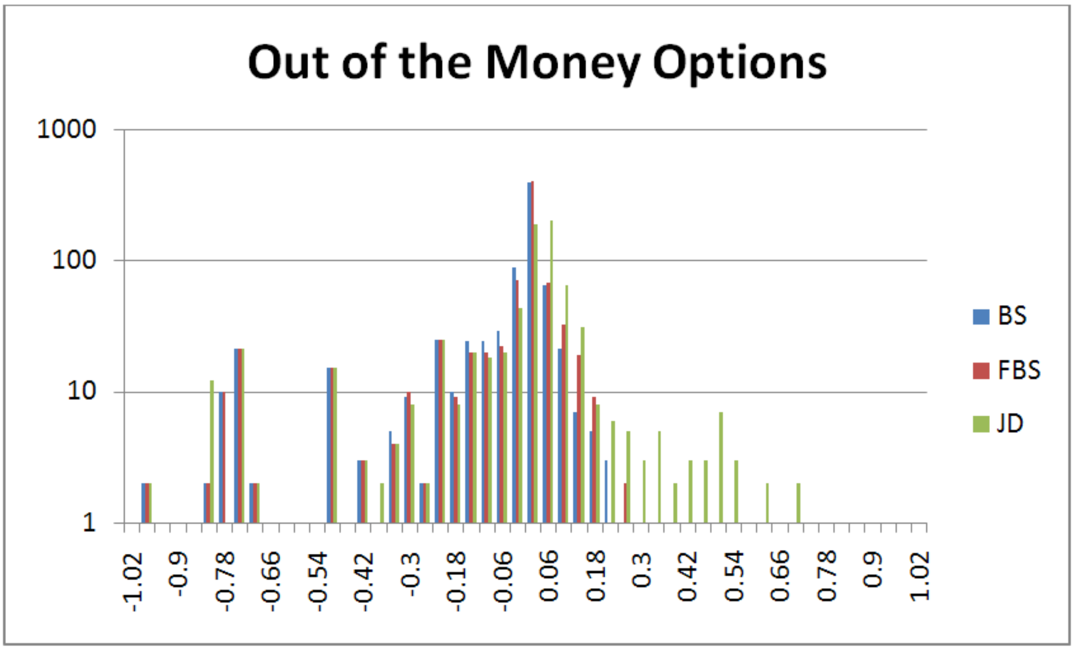

Fig. 39: *Log-histogram of deviances of the Black-Scholes, French Black-Scholes and Jump-Difussion Model from the market data (out of the money options, implied volatility)*

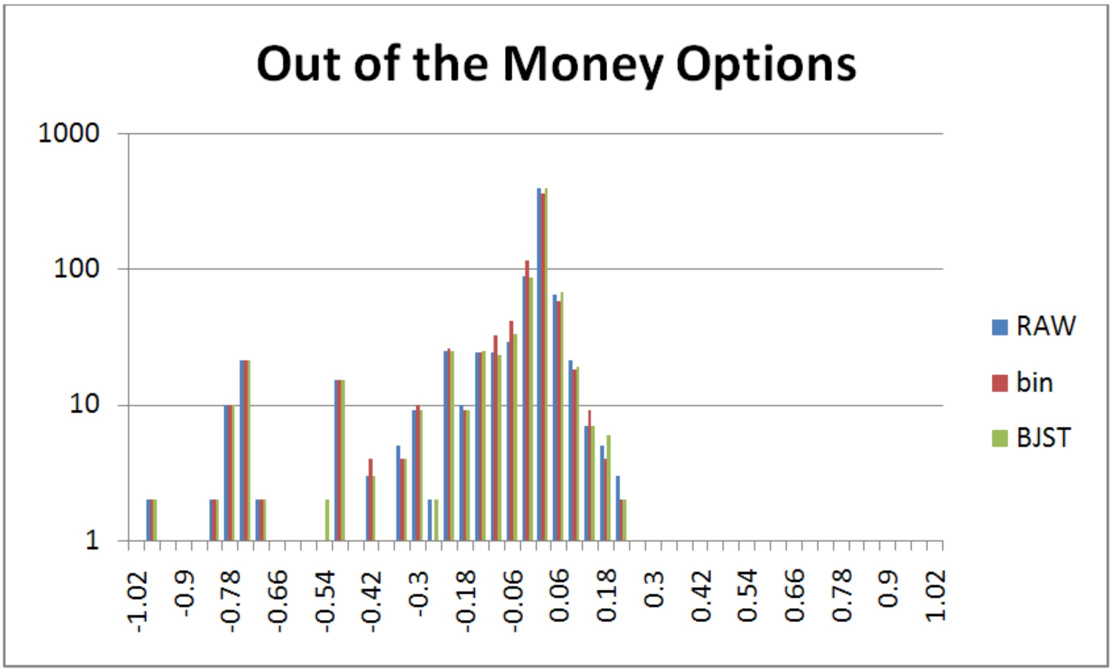

Fig. 40: *Log-histogram of deviances of the BAW, Binomial and BJST Model from the market data (out of the money options, implied volatility)*

The tendency of all models to price the options correctly is clear from the figures. Note that only JD Model overestimated the price by more than 0.25 USD.

4.7.3 At the Money Options

At the end the subset containing at the money options will be analyzed. There are exactly 500 calls and 500 puts. 626 options from this subset were not traded and 220 of them have zero open interest. The average time to expiration is 120 days and the median value is 77 days. The bid ask spread is approximately 1.36 USD.

The results of the estimations by all 6 models are below

	BS	FBS	bin	BAW	BJST	JD
mean	0.044204	0.12296939	0.022425	0.044941	0.04364	1.503987
median	0.009535	0.05211771	-0.00209	0.00973	0.009257	-0.06893
max	9.252433	9.49468058	9.192422	9.25523	9.248853	280.2144
min	-3.613434	-3.5923228	-3.63311	-3.61343	-3.61343	-3.83764
Q1	-0.091779	-0.0373271	-0.10718	-0.0916	-0.09373	-0.20953
Q3	0.115632	0.18678979	0.104315	0.115649	0.115043	0.030204
stddev	0.628884	0.63933559	0.628323	0.628555	0.628308	18.39032
skew	5.743993	5.81189062	5.606783	5.758223	5.751558	13.03512
kurt	84.11267	83.0188296	82.79137	84.31871	84.33439	175.8376
R-squared	0.99501	0.99467784	0.995037	0.995014	0.995019	-3.27491
mispriced	59	60	60	59	59	103
underpriced	22	17	26	22	22	58
overpriced	37	43	34	37	37	45

Tab. 20: *Differences between estimated and the ask price (at the money options, implied volatility)*

The results look very well; especially the binomial model provided outstanding outcome. This is the first time when one model is so close to the ideal one. The high numbers of R-squared, low standard deviations, nearly zero means and medians and rarely short quartile spreads suggest that the first five models fit the data very well. Note that the number of mispriced options is nearly the same for the first 5 models, but the distribution between the underpriced and overpriced options significantly differs.

Below are provided the histograms.

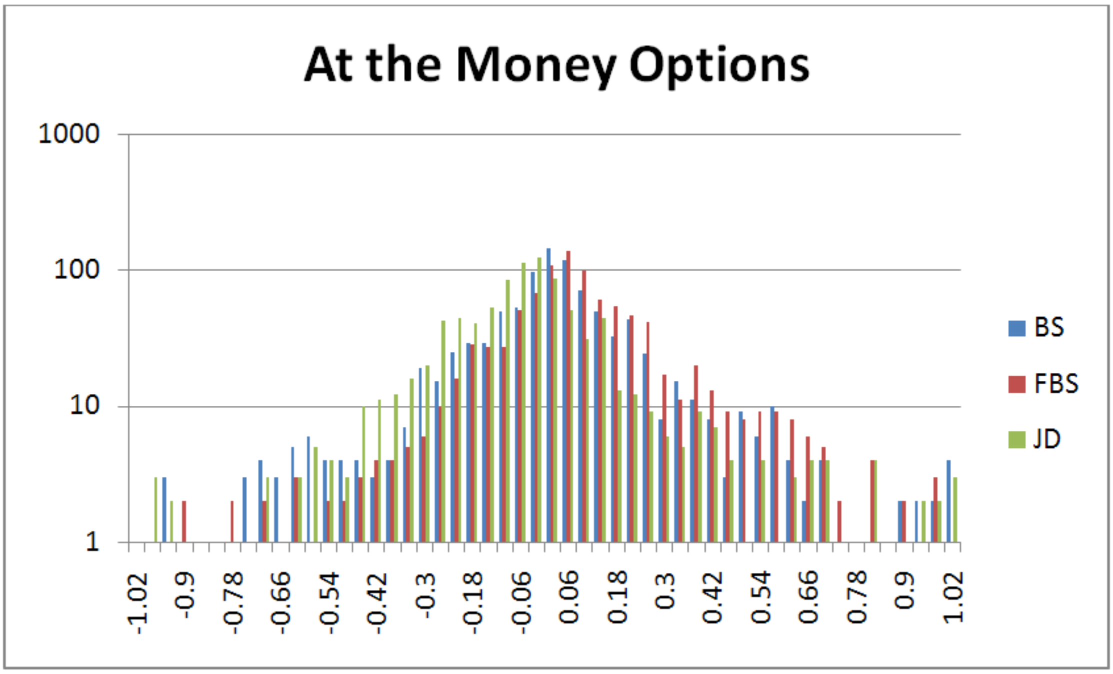

Fig. 41: *Log-histogram of deviances of the Black-Scholes, French Black-Scholes and Jump-Difussion Model from the market data (at the money options, implied volatility)*

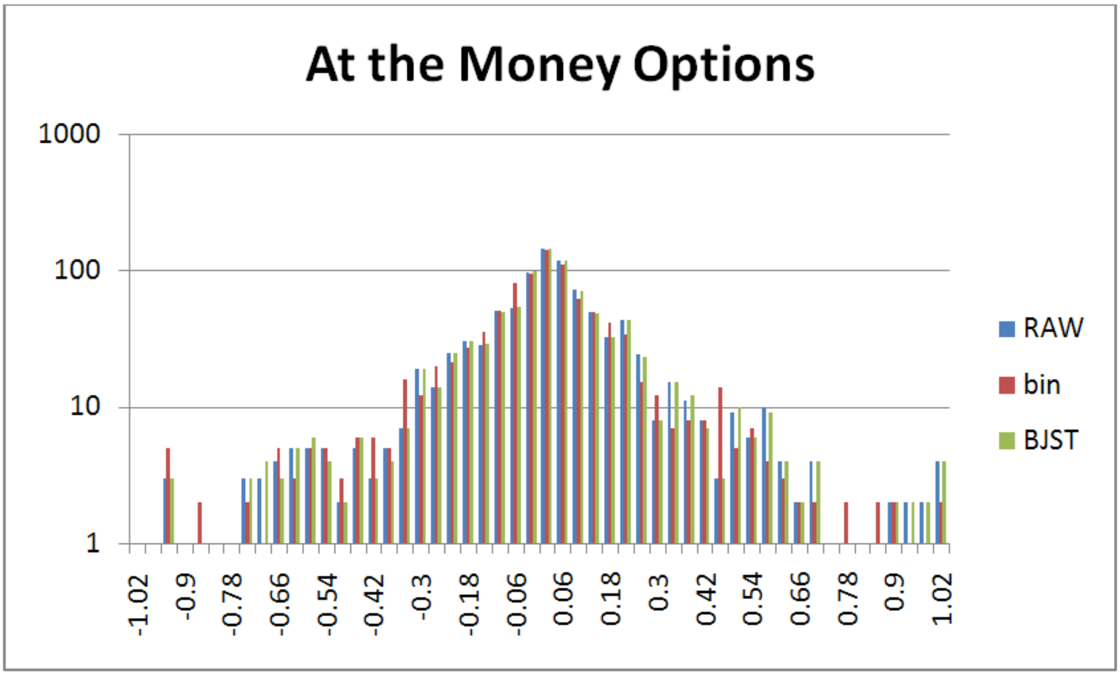

Fig. 42: *Log-histogram of deviances of the BAW, Binomial and BJST Model from the market data (at the money options, implied volatility)*

The modal interval for all models contains zero, but there are just about 140 observations in those intervals for the first 5 models. One could expect more according to the results provided above, but note that the kurtosis is not definitely as high as for the model run on the random data.

4.7.4 Observed Differences

All option pricing models were biased when dealing with options extremely out of the money or extremely in the money; this resulted in overpricing the options in the money and underpricing the options out of the money. Most of the options with extreme ratio of the spot and strike price were not traded. These observations correspond the facts stated when the implied volatility was discussed in 4.4.4.

All pricing methods worked very well on the subset containing at the money options. The only disadvantage of the estimation of at the money options was lower kurtosis of the difference between estimated and market price distribution than for the previous datasets.

5. Conclusion

In this thesis was provided a description of the financial derivatives called options. There were described the basic characteristics of the options and the methods of usage of those instruments in the second chapter. Since there are various types of options it was necessary to differ between these types to obtain the sufficient knowledge to understand the pricing models.

The third chapter provided a description of six pricing methods. The first one was the binomial model, which is the most intuitive way to determine an option price. This model could be used for pricing European and also American options. The next model was the famous Black-Scholes model for pricing European options. There is also a description of the submodel called the French Black-Scholes model, which should be more suitable for options close to the expiration since it takes into account the real number of trading days. The next was the Quadratic approximation model, which is built on the assumptions of the Black-Scholes and provides an extension for the pricing of American options. Then follows Bjerksund-Stersland model, which is another extension to price American options. The last model tested in this paper was the Jump Diffusion model, which was the only one of the used models that assumes the development of the underlying price is at least partially described by the Poisson process.

The fourth chapter provides the results of the empirical testing of all previously mentioned models. From the dataset were obtained the basic characteristics of the options. It was impossible to evaluate all of them since the dataset contains more than 270 000 items, so there were chosen several subsets to test all pricing models. The only model applied to the whole dataset was the Black-Scholes one and it gathered quite good results, but firstly it was needed to decide whether the historical or implied volatility is better to use in the pricing methods. The 3 days old implied volatility provided much better results not only for the Black-Scholes model, but later on even for the others. The Black-Scholes model evaluated the options quite well in spite of the fact that it is build for the European options. The central tendency to price the options correctly was very clear.

After the analysis of the whole dataset, all methods were implemented on the randomly chosen subset of 1000 options. The first important result of this analysis was that the options are priced to fit the ask price, not the bid one and not even anywhere in between. The results confirmed that the implied volatility is better for all pricing formulas, not just the Black-Scholes one. Most of the models showed a very good fit to the data so it was nearly impossible to determine which one is the best. On the other hand the Jump Diffusion model was undoubtedly the worst, which was caused by the need of naively guessing the parameters due to the computational complexity connected to estimating Poisson parameters for all underlying assets.

The first non-random selection of the subset was made on the basis of traded volume. It was observed that the pricing models generally fit traded options more than the non-traded ones. The French Black-Scholes model provided the best results for the most traded options since those options were very close to their expiration. On the contrary Bjerksund-Stersland provided the best results for the least traded options.

The next selection was based on the time to expiration. The bid ask spread for options far from the expiration was approximately three times higher than for the opposite extreme. The results showed that the Jump-Diffusion model as it was used in this thesis is not suitable for options far from the expiration. The results for options with shorter time to maturity were generally better. The options with a long time to maturity seem to be little underpriced due to the negative skew of the obtained distribution of the differences between the estimated and the ask price.

The last group of subsets was chosen according to the relationship between the underlying spot and strike price. All models provided biased results for options far in the money or out of the money, but one should mention that these options were traded very rarely. On the other hand all pricing methods worked quite well on the subset containing at the money options, but the distribution of differences possessed lower kurtosis than for the general subset.

Since all of the mentioned methods were used in the same way before the current financial crisis and they roughly worked now, it cannot be declared that these methods were affected by the contemporary financial situation. It was nearly impossible to determine the best method for most of the subsets since there are many criteria determining the quality of the pricing model and there simply was not any dominant alternative in this multi-criteria decision making process. However there are many significant results, which could help to decide about the option prices more effectively.

The first crucial finding of this thesis is the fact that it does not matter so much if an American option is priced as a European one. There was provided empirical evidence that the discrepancies between the different results of the used models are so tiny that the difference between European and American options could be almost neglected. The second important observation is that the option pricing methods usually calculate an ask price. The differences between the ask price and the results of the option pricing outputs were usually very close to zero. The last but not least revealed fact is that all option pricing methods work generally much better with implied volatility than the historical one. It is necessary to use a delayed implied volatility to estimate the option prices since it is impossible to know the actual implied volatility before the actual option price.

This paper does not outline any universal way how to make a profit at the option markets, but it could save a lot of work for prospective investors. The results show the clear picture of the robustness and the suitability of all analyzed methods in various situations.

References

Printed

[1] Barone-Adesi, G.; Whaley, R. E.: *Efficient Analytic Approximation of American Option Values*, The Journal of Finance, Vol. 42, No. 2, 1967

[2] Bjerksund, P.; Stensland, G.: *Closed Form Valuation of American Options*, The Scandinavian Journal of Economics, 2002

[3] Blaha, Z. S.: *Opce, swapy a futures - deriváty finančního trhu*,Praha Management Press, 1997

[4] Blake David: *Financial Market Analysis*, New York, McGraw-Hill, 1990

[5] Cont, R., Tankov, P.: *Non-Parametric Calibration of Jump–Diffusion Option Pricing Models*,Journal of Computational Finance, Vol. 7, No 3, 2004

[6] Cont, R.: *Empirical properties of asset returns: stylized facts and statistical issues*, Quantitative Finance, 1: 223-236, 2001

[7] Cox, J. C., Ross, S. A., Rubinstein, M.: *Option Pricing: A Simplified Approach*, Journal of Financial Economics 7: 229-263, 1979

[8] Dlouhý, M. a kol.: *Simulace podnikových procesů*, Brno: Computer Press, 2007

[9] Eberlein, E.; Keller, U.: *Hyperbolic Distributions in Finance*, Bernoulli 1 (3), 281-299, 1995

[10] Fabozzi, F. J.: *Fixed Income Analysis*, New Jersey,John Wiley & Sons, 2007

[11] Feng, L.; Linetsky, V.: *Pricing Options in Jump-Diffusion Models: An Extrapolation Approach*, Operations Research, Vol. 56, No. 2, 2008

[12] Fouque, J. P.: *Derivatives in financial markets with stochastic volatility*, Cambridge, Cambridge University Press, 2000

[13] French, D.: *The Weekend Effect on the Distribution of the Stock Prices: Implications for Options Pricing*, Journal of Financial Economics, Vol. 13, No. 4, 1984, 547 - 559

[14] Gapko, P.: *Warrant aneb „opce pro drobné investory*“,Rigorózní práce, Universita Karlova v Praze, 2006

[15] Gibson, R.: *Option valuation: analyzing and pricing standardized option contracts*, New York, McGraw-Hill, 1991

[16] Gujarati, D. N.: *Basic econometrics*, Boston, McGraw-Hill, 2003

[17] Hanson, F. B.; Westman, J. J.: *Jump-Diffusion Stock Return Models in Finance: Stochastic Process Density with Uniform-Jump Amplitude*, University of Illinois at Chicago; University of California, 2003

[18] Heston, S. L.: A *Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options*, The Reveiw of Financial Studies, Vol. 6, No. 2

[19] Hitchner, J. R.: *Financial Valuation Applications and Models*, New Jersey John Wiley & Sons, 2003

[20] Hoffmann, Christian: *Valuation of American Options,* Diploma thesis, University of Oxford, 2000

[21] Hull, J.C.: *Options Futures and Other Derivatives,* New Jersey,Prentice Hall, 2009

[22] Hušek, R.: *Ekonometrická analýza*,Praha, Oeconomica, 2007

[23] Ito, K.: On stochastic differential equations, Memoirs, American Mathematical Society 4, 1951

[24] Kou, S. G.: *A Jump-Diffusion Model for Option Pricing*, Management Science, Vol. 48, No. 8, 2002

[25] Merton, R. C.: *Option pricing when underlying stock returns are discontinuous,* Journal of Financial Economics, 125-144, 1976

[26] Merton, R. C.: *Theory of Rational Option Pricing*, The Bell Journal of Economics and Management Science, Vol. 4, No. 1, 1973

[27] Nekula, K.: *Finanční opce,* Diploma thesis, Universita Karlova v Praze, 2004

[28] Pánková, V.: *Nelineární optimalizace pro ekonomy*, Praha, Professional Publishing, 2009

[29] Scholes, M.; Black, F.: *The Pricing of Options and Corporate Liabilities*, Journal of Political Economy, Vol. 81, No. 3, 1973

[30] Tankov, P.; Voltchkova, E.: *Jump-diffusion models: a practitioner's guide*, Université Paris 7, Université Toulouse 1, 2008

[31] Toivanen, J.: *Numerical Valuation of European and American Options under Kou's Jump-Diffusion Model*, University of Jyvaskyla, 2007

[32] Turnbull, S. M.: *Option valuation*, Holt, Rinehart and Winston of Canada, 1987

[33] Zaffaroni, P.: *Nonlinear time series with long memory: a model for stochastic volatility*, London, London School of Economics and Political Science, 1997

[34] Zwillinger, D.: *Handbook of Differential Equations*, Boston, Boston Academic Press, 1997

Internet

[1] Baruník, J.: *Stock Price Evolution - Definitions, Markov Random Processes, Brownian Motion, (Lecturer's notes on Quantitative finance II)*, IES FSV UK
Available at:
http://staff.utia.cas.cz/barunik/files/QFII/QF_II_Lecture2.pdf

[2] Chance, D.M.: *Convergence of the Binomial to the Black-Scholes Model*, Louisiana State University
Available at:
http://www.bus.lsu.edu/academics/finance/faculty/dchance/Instructional/TN00-08.pdf

[3] Dědek, O.: *Option contracts (Lecturer's notes on Financial Markets Instruments)*, IES FSV UK,
Available at:
http://samba.fsv.cuni.cz/~dedek/Financial%20Market%20Instruments/he02%20Options.pdf

[4] Hogan, F.: *Applications of Stochastic Calculus to Option Pricing*, Columbia University, 2008
Available at:
http://www.columbia.edu/~mh2422/Applications.pdf

[5] Röman, J. R. M.: *American Options*, Mälardalen University
Available at:
http://janroman.dhis.org/doc/lnotes/AFI/American.pdf

[6] Wuertz, D.: *Basics of Option Valuation*
Available at:
http://cran.r-project.org/web/packages/fOptions/fOptions.pdf

Datasets sources

[7] http://www.historicaloptiondata.com

[8] http://www.cboe.com