Flexible Option Valuation Methods

Yun YIN

School of Economics

University of East Anglia

A thesis submitted for the degree of

Master of Philosophy

Abstract

This thesis is concerned with methods of option valuation that fall completely outside of the Black-Scholes-Merton (BSM) framework. Data on S&P500 Index options are used to demonstrate the proposed methods. Some of our favoured methods are based on semi-parametric regression; others on simulation. The thesis consists of a number of chapters. In Chapter 2, we outline existing option valuation methods, with particular attention paid to the binomial-tree model and the Black-Scholes formula. We demonstrate that under some circumstances these two methods are equivalent. We also demonstrate using the binomial tree method that a "TGARCH effect" (which plays an important role in later chapters) can explain the well-known "smirk" pattern that is often observed in market option price data.

In Chapter 3, we use regression analysis to investigate the ways in which features of an option actually determine the market price. We start with polynomial regressions, and progress to additive models, with components obtained using the B-spline technique. The focus in these regression models is the role of volatility. Historical measures of volatility are used as explanatory variables in the regression, with one objective being to discover how far back into the past option traders are going when computing volatility. It is proposed that this approach gives rise to an alternative measure of implied volatility that is completely free of the Black-Scholes framework. We use the Practitioner Black Scholes (PBS) model as a Benchmark for comparison. The best of our regression models is found to perform better than the Black-Scholes formula in out-of-sample prediction of market prices.

In Chapter 4, we focus on the underlying (S&P500) Index, and consider a number of varying volatility models (ARCH, GARCH and TGARCH) of daily returns. We found that the TGARCH model is the best model to represent the volatility process. Then we simulated data from the models considered using the coefficients from the estimated models. After that, we found that the simulated ARCH family volatility models worked correctly, since the "true" parameter values are included in the confidence intervals.

Chapter 5 continues with the simulations of daily return data, in building a Monte Carlo program for the valuation of European options that allows for varying volatility. Of particular interest is whether superior models of the underlying stock price (i.e. ARCH, GARCH and TGARCH) result in option valuations that are superior to the Black-Scholes valuation. Superiority in this context is defined primarily in terms of ability to predict market prices. We find that all models perform better than the benchmark (PBS) model. Which model performs best depends on the type of market and time to

expiry: ARCH is the best model for predicting the short and medium term put options for both bear and calm market and GARCH is the best one for predicting the long term put options in the bear market. In the crash market, TGARCH Monte Carlo simulation is the best model for predicting the long term European call and put options.

Acknowledgement

Firstly, I would like to express my sincere gratitude to my supervisor Professor Peter Moffatt for the continuous support of my PhD study and related research, for his patience, motivation, and immense knowledge. His guidance helped me throughout the time of the research and the writing of this thesis. I could not have imagined having a better supervisor and mentor for my PhD study.

Besides my supervisor, I would like to thank the rest of my thesis committee, Dr. James Watson and Dr. Bahar Ghezelayagh, for their insightful comments and encouragement, and also for the hard questions which prompted me to widen my research from various perspectives.

Special thanks also go to my family. Words cannot express how grateful I am to my mother and father for all of the sacrifices that you've made on my behalf. Your prayer for me was what sustained me thus far. I would also like to thank all of my friends who supported me in writing, and encouraged me to strive towards my goal.

I am also very grateful to the examiners, Dr Panayotis Andreou and Dr Marta Wisniewska, for supplying excellent feedback on earlier versions of the thesis.

Contents

List of Figures

List of Tables

Chapter 1 Introduction

Options are financial instruments that can provide investors with the flexibility needed in a wide variety of investment situations. Although the history of options extends several centuries, a formal market for options was not established until the Chicago Board Options Exchange (CBOE) opened its doors in 1973.In the same year, two economists, Fischer Black and Myron Scholes, published an article proposing a model for calculating the theoretical estimate of an options price over time. Also in the same year, their colleague Robert Merton published an additional study providing a mathematical extensions of the Black-Scholes model.

With an exchange created and a solid model for pricing, the market flourished. New options contracts were issued subject to standardized terms, such as uniform expiry dates and established strike prices. The Options Clearing Corporation (OCC) became the central clearing house, which guaranteed trades, and was responsible for regulatory oversight on par with U.S. stock markets.

In 1973, options trading at the CBOE was restricted to call options in only 16 stocks.Over time, the listed options market expanded to additional exchanges and products, including put optionsand index options.

The popularity of options has steadily increased. The US is the largest option trading market, with total trading volume on U.S. options exchanges being 4.14 billion contracts in 2015 (16.4 million contracts per day).[1] However, the popularity of options has spread over the world. In the UK, the London International Financial Futures Exchange (LIFFE) introduced equity options to its product range in 1993. China eventually launched stock options on 9thFebruary 2015, aiming to develop broader markets and give investors a tool to manage risk. The options were written on an exchange-traded fund, the China ETF, which tracks 50 of China's largest listed firms, including banking giant ICBC and carmaker SAIC Motor.

Options are useful in hedging, since when held in conjunction with other assets, loss from holding the portfolio can be avoided with certainty. They are also useful for speculators: it is possible to make large profits by searching for favourably priced options, and then either holding them to expiry or selling them at a profit at some time before expiry.

Whether the trader is a hedger or a speculator, they need reliable methods for valuing options.If a hedger is purchasing an option in order to create a hedge, they need to know that the offer price is close to the true value of the option, otherwise the portfolio might result in a net loss. If the trader

[1] Acworth (2015)

is a speculator, an obvious strategy is to buy options for which the true value is higher than the offer price, and sell options for which the true value is less than the bid price.

Option valuation methods are usually related in some way to the Black-Scholes-Merton (BSM) framework (Black and Scholes, 1973; Merton, 1973). This framework is built on the assumption that the underlying asset price follows a geometric Brownian motion process with constant return volatility (or a random walk). Another key assumption is risk-neutral arbitrage. The great advantage of the BSM framework is that the value of certain options (e.g. European options) can be expressed in closed form, and the formula can be easily applied. Because of this, the BSM approach has become the industry standard in recent decades.

In 1997, the importance of this contribution was recognized when Robert Merton and Myron Scholes were awarded the Nobel Prize for Economics. Robert Merton, among many others, extended the Black-Scholes model in several important ways.As these studies have shown, option pricing theory is relevant to almost every area of finance.

While fully acknowledging the hugely significant contribution of Black, Scholes and Merton, the overall message of this thesis is that there is perhaps too great a reliance on the BSM framework. The main problem is that it is based on the assumption of constant return volatility of the underlying asset.There is a vast literature in Financial Econometrics establishing that this assumption is usually false. In addition, it is often suggested that the well-known “smiles” and “smirks” seen in option price data are a consequence of the violation of this assumption. Some authors have attempted to extend the Black-Scholes framework to incorporate varying volatility.In this thesis, we instead focus on option valuation methods that are completely outside of the Black-Scholes framework. One method is based on regression; the other is based on simulation.

Another message of the thesis relates to the question: what are option valuations actually useful for? For example, the predicted valuations from a regression model may be compared to actual option prices and hence a measure of the predictive performance of the model can be obtained. Predictive performance is clearly important in, for example, setting prices for options in thinly traded markets. Another use of option valuations is in the development of profitable trading rules. The decision to purchase an option is made on the basis of a comparison between the valuation and the market price. Purchased options may then be combined with appropriate units of the underlying to form a hedged portfolio. The average performance of the hedged portfolios can then be used as a measure of the performance of the model.

1.1 A broad framework

Here we will introduce a broad framework to which most of the analysis of the thesis can be related. We will not provide complete definitions here - this will be done in Chapter 2.

An option is a security written on an underlying asset or index (often just called the "underlying"). Two basic types of option are European and American options. The fundamental difference is that European options can only be exercised at expiry, while American options may be exercised at any date on or before expiry (i.e. "early exercise" is permitted). In this thesis we will mainly be concerned with European options because they are much easier to analyse.

Consider a European call option which has price (or value) c.It is well known that c depends on a small list of variables, namely the current (time t) price of the underlying (S_t), the strike price (K), the time to expiry (τ), the risk-free rate (r_f) and the volatility of the underlying (σ).[2] So we can write:

$$c = f\left(S_t, K, \tau, r_f, \sigma\right) \tag{1.1}$$

The first four arguments of f in (1.1), S_t, K, τ andr_f, are all known at time t (the time when the option is traded.The fifth argument, σ, is however unknown. This means that in order to apply the formula (1.1) it is necessary to estimate the volatility of the underlying index and use this estimate ($\hat{\sigma}$ say) in place of σ in (1.1). How this estimate of volatility should be obtained is one of the key issues of this thesis.

Other key ideas of the thesis can be introduced within the framework of (1.1). First of all, the binomial-tree formula and the Black-Scholes formula, which will both be derived in detail in Chapter 2, are both special cases of (1.1). Let us label the Black-Scholes formula as:

$$c = f_{BS}\left(S, K, \tau, r_f, \sigma\right) \tag{1.2}$$

An important concept is implied volatility, and this can be defined using (1.2).If the market price of the option is $\tilde{c}$, then the implied volatility is defined implicitly as:

$$\tilde{c} = f_{BS}\left(S, K, \tau, r_f, \tilde{\sigma}\right) \tag{1.3}$$

[2] Option prices generally also depend on the dividend rate, but since we will mainly be concerned with options written on stock market indexes, there are no dividends to consider, and the dividend rate can be assumed to be zero.

That is, the implied volatility is the volatility that is required in the Black-Scholes formula that makes the value of the option exactly equal to the market price of the option. Or, it is the volatility of the underlying asset that is implied by the market price of the option. The implied volatility is not the same as the true volatility, for two reasons: the market price of the option might not correctly represent the true value of the option; orthe Black-Scholes formula may be an invalid procedure for finding the true value of the option. Before 1987, implied volatility was a U-shaped function (known as a “volatility smile”) of the strike price, with minimum around the current price. This implies that both in-the-money and out-of-the money options were over-priced relative to at-the-money options. Since 1987, implied volatility has more commonly been a monotonically decreasing function of strike price (hence “volatility smirk”). This implies (e.g.) that in-the-money Calls are over-priced, but out-of-the-money calls are under-priced.

Another important topic of the thesis is regression analysis of option prices, which will be covered in Chapter 3. Here, we are performing least squares regressions with market prices of options as the dependent variable, and the five arguments off in (1.1) (and non-linear functions of them) as explanatory variables. An estimate $\hat{\sigma}$ is used in place of σ. We can write the fitted regression equation as:

$$\hat{c} = f_{LS}\left(S, K, \tau, r_f, \hat{\sigma}\right) \tag{1.4}$$

Regression models of the form (1.4) are useful for a number of reasons. Most importantly, by experimenting with different estimators $\hat{\sigma}$ for the volatility, σ, it is possible to find the estimator that optimises the fit of the regression.This optimal estimator could then be interpreted as the estimator of volatility that option traders are using when setting prices. Hence this provides an alternative means to compute the implied volatility of the option. However, unlike $\tilde{\sigma}$ in (1.3), this estimate of implied volatility is not based on the assumption that the Black-Scholes model is true. In this sense the implied volatility obtained using (1.4) is “model-free”.

All of the models considered above assume that the volatility of the underlying, σ, is fixed over time. Ways of relaxing this assumption are an important part of this thesis. A very important model which relaxes this assumption is the Practitioner Black-Scholes (PBS) model:

$$c = f_{PBS}\ (S, K, \tau, r_f, \widehat{\sigma}(K, \tau))$$

Where $\widehat{\sigma}(K,\tau)$ is a (usually non-linear) function of the strike price (K) and the time to expiry (τ). The PBS model is normally estimated in two stages.

Firstly, a least squares regression is performed with implied volatility as the dependent variable in order to estimate the function $\hat{\sigma}(K,\tau)$. Secondly, the predicted values from this regression are plugged into the Black-Scholes formula in order to obtain predicted option prices. The PBS has become a very popular model and will be used as a benchmark model for comparison with models of interest in this thesis.

We are also interested in models in which volatility simply varies over time (such as ARCH and GARCH). Let $\sigma(t)$ represent a model that describes the process followed by volatility over time, and let $\hat{\sigma}(t)$ be the estimated volatility model (estimated using historical data on the underlying price). Then:

$$c = f_{vv}\left(S, K, \tau, r_f, \hat{\sigma}(t)\right) \tag{1.5}$$

is a model of the option price that allows for varying volatility (vv).

The concept of implied volatility can be extended to varying volatility models.Let the market price of the option be $\tilde{c}$. If:

$$\tilde{c} = f_{vv}\left(S, K, \tau, r_f, \tilde{\sigma}(t)\right) \tag{1.6}$$

then $\tilde{\sigma}(t)$ is the *implied volatility process* of the option.Methods for finding $\tilde{\sigma}(t)$ have been considered by Engle and Mustafa (1992).

1.2 Outline of thesis

Chapter 2 surveys existing option valuation methods. The most basic of these are the Binomial-tree method and the Black-Scholes method, and these two methods are derived in detail and demonstrated. It is demonstrated that under some circumstances these two methods are equivalent. It is also demonstrated that including a “TGARCH effect” in the binomial tree is a way of explaining the well-known “smirk” phenomenon seen in option price data.

Chapter 3 considers regression analysis of the market prices of options.The central question is: what features of an option actually determine the market price? Since the function that is being estimated, (1.1) above, is a highly non-linear function of certain arguments, flexible regressions are required. We will start with polynomial regressions, and progress to additive models, using B-splines for the components.Using these sorts of models, we will see that it is possible to obtain a very good fit of the data with relatively few parameters.

The regression approach is very useful for finding out which measure of volatility is being used by traders, that is, how far they appear to go back into the past when computing historical volatility. This can be answered by considering which volatility measure gives rise to the best fit in the regression.

Having estimated the regression models and identified which measure of volatility is most useful, we need to assess the performance of the model. For this purpose we will consider both in-sample and out-of-sample predictive performance. One particularly important question is whether any of the regression models can predict option prices better than the benchmark Practitioner Black-Scholes (PBS) model, used by many researchers including Christoffersen and Jacobs (2004) and Andreou (2014).

Chapter 4 progresses to varying volatility models. The models considered are in theAutoregressive Conditional Heteroscedasticity (ARCH) family. In addition to ARCH itself, we consider Generalised ARCH (GARCH) and Threshold GARCH (TGARCH). We will first estimate all of these models using daily data on Index returns. We will then simulate data from the estimated models, and use the simulation routines in a Monte Carlo study of the three estimators. One reason for doing this is to verify that the simulations are being performed correctly. Another reason is that it prepares the ground for the more extensive Monte Carlo analysis of Chapter 5.

In Chapter 5, the objective is to estimate the value of options in ways that take account of the presence of varying volatility. This is done using simulation. In this chapter, first of all, we will use Monte Carlo simulation methods to simulate theARCH, GARCH and TGARCH. Secondly, we will calculate value the European call and put options from three different models ARCH, GARCH and TGARCH, and we will try to find which model is the best model for predicting the option price. After we have identified the best model, we will compare its predictive performance with the benchmark PBS model, by finding which model has predictions closest to option market prices. In order to compare the best Monte Carlo model with the benchmark PBS model, we will use the Root Mean Square Error (RMSE) from the out-of-sample test.

Chapter 6 concludes, and suggests directions for further research.One particular direction for future research is the use of Monte Carlo with varying volatility to value American options. American options play a minor part in the thesis because they are more complicated to model. However, in Chapter 6 we outline a routine that could be used to extend the techniques developed in Chapter 5 to the case of American options.

1.3 Data Extraction and Processing

The options data used in the thesis is data on S&P500 Index options, obtained from Optionmetrics.[3] Our data set covers the period from January 2000 to Oct 2009. From the Optionmetrics database, we obtain for each option a settlement date, an expiry date, a current price of underlying, and a strike price. The daily risk-free interest rate is the US 3-month Treasury-bill rate which is obtained from Datastream.[4] For the option price, we used the midpoint of the closing option bid-ask spread, since the midpoint price reduces the noise in the cross-sectional estimation of implied parameters (Dumas et al. 1998; Andreou et al. 2014). Also the midpoint of the closing option bid-ask spreadcorresponds to the closing value of the S&P 500 (Andreou et al. 2014). In our analysis, time to expiry, τ, is computed assuming that there are 360 calendar days per year. We also use time series data on the underlying S&P500 index, also obtained from Datastream. This analysis is to find estimates of volatility, and to estimate varying volatility models. Daily data starting on 1 January 1999 is used for this purpose.

Several filtering rules (following Bakshi et al. 1997; Andreou et al. 2014) are applied to construct the option data set. First, we eliminate all observations which the trading volume is equal to zero, since these options are not traded at all and do not represent actual trades. Second, in order to avoid observations on illiquid options, we eliminate all observations with either less than 6 or more than 253 trading days to maturity, or a moneyness ratio that is less than 0.75 or higher than 1.25. Third, we exclude observations with the price quotes lower than 1.0 or with midpoint option price lower than the bid-ask spread difference in order to minimise the impact of price discreteness on option valuation.

The final data set has a total of 381,265 observations: 165,648 call options and 215,617 put options.

We will divide the option data into several categories according to the term to expiration. An option contact can be classified as (i) short-term (< 60 days); (ii) medium term (60-80 days); and (iii) long term (> 180 days). Sample characterises for the whole data set are reported in Table 1.1. Summary statistics are reported for the average of midpoint price of bid-ask, daily average values of Black-Scholes implied volatility, daily average option volume and the total number of observations.

We will also divide the option data into different markets according to the time path of the S&P 500 Index. Figure 1.1 shows the time path of the S&P

[3] http://www.optionmetrics.com/

[4] http://financial.thomsonreuters.com

500 Index (daily data), daily return on the Index, and the squared daily return (volatility), from Jan 2000 to Oct 2009.

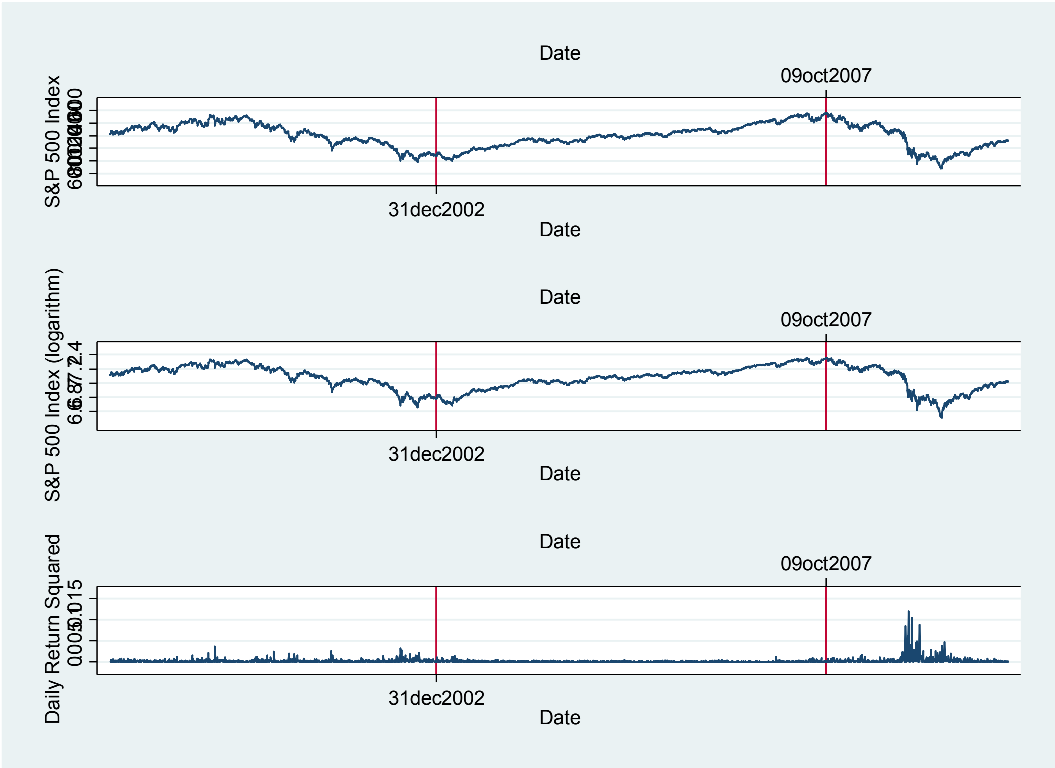


Figure 1.1 Time path of S&P 500 (daily data), daily return of S&P 500 and the squared daily return (volatility) of S&P 500; 1 Jan 2000 – 31 Oct 2009.

In Figure 1.1, we see that between Jan 2000 and 31 Dec 2002, the S&P 500 index has a tendency to move downwards. Between Jan 2003 and Oct 2007, it appears to rise steadily. Between Oct 2007 and Oct 2009 (the period including the financial crisis) the volatility appears to rise markedly. This is clearly verified in the third graph (squared daily return) which represents the volatility of the S&P 500 index. Based on these observations, the option data set will be divided into three time periods: (i) "Bear Market": 1 Jan 2000 to 31 Dec 2002; (ii) "Calm Market": 1 Jan 2003 to 09 Oct 2007; (iii) "Crash Market": 10 Oct 2007 to 31 Oct 2009. Sample characteristics for the data set are reported in Table 1.1.

	Bear Market	**Calm Market**	**Crash Market**	**ALL**
	Short term options: < 60 days			
Implied volatility	0.26	0.17	0.34	0.26
	(0.23)	(0.14)	(0.28)	(0.22)
Price	36.90	20.16	33.71	30.26
	(35.43)	(28.34)	(37.82)	(33.86)
Volume	873	2166	2336	1792
	(769)	(1610)	(4694)	(2358)
Observations	25,025	53,506	45,613	124,144
	(19,639)	(40,139)	(34,987)	(94,765)
	Medium term options: 60-180 days			
Implied volatility	0.26	0.19	0.31	0.25

	(0.21)	(0.14)	(0.25)	(0.20)
Price	51.54	29.03	53.23	44.60
	(41.37)	(36.73)	(42.21)	(40.10)
Volume	477	1190	1739	1135
	(427)	(924)	(1274)	(875)
Observations	17,042	33,222	23,359	73,623
	(13,545)	(23,332)	(20,227)	(57,104)
Long term options: > 180 days				
Implied Volatility	0.24	0.19	0.30	0.24
	(0.20)	(0.14)	(0.24)	(0.19)
Price	73.24	41.01	78.06	64.10
	(54.01)	(49.28)	(59.29)	(54.19)
Volume	234	767	1106	702
	(273)	(463)	(762)	(499)
Observations	4,254	9,227	4,369	17,850
	(3,326)	(6,626)	(3,827)	(13,779)
All	46,303	95,955	73,341	215,617
	(36,510)	(70,097)	(59,041)	(165,648)

Table 1.1 Sample descriptive statistics by time period and time to expiry
The table is divided into three blocks: short term (τ<60 days); medium term (60<τ<180 days); long term (τ>180 days).In each cell, the first number relates to put options, and the number below in parentheses relates to call options.The first row in each block contains average values of Black-Scholes implied volatility; the second row contains the average values of price (midpoint of bid-ask spread); the third row contains the average option volume; the last row contains the number of observations (i.e. number of options).

1.4 Literature Review

The purpose of this section is to provide a concise survey of the literature that is central to the themes of the thesis. The most important of these references will be cited again in later chapters.

<u>Texts and general references</u>

There is a huge number of textbooks on option theory, and they are often very similar to each other in content. One that is very popular, probably because it is comprehensive and well written, is Hull (2011). Another that we will refer to later is Adams et al. (2003). There is also a huge number of journal articles on option theory. Key references on the Black-Scholes framework for the valuation of options are Black and Scholes (1973) and Merton (1973). The key reference on the binomial-tree valuation method is Cox et al. (1979). One financial econometrics text that covers empirical analysis of option prices is Brooks (2014).

<u>Nonparametric Methods</u>

Hutchinson et al. (1994) used nonparametric methods to price European Options. The particular techniques used come under the heading of learning networks. The data is allowed to determine both the dynamics of the underlying and its relation to the option prices with minimal assumptions on either. The shortcomings of the approach is that it relies on large amount of historical prices, and therefore would not be useful for the analysis of thinly traded options.

Ait-Sahalia and Lo (1997) proposed a more advanced nonparametric kernel regression model to estimate the option price with no parametric restrictions on either the underlying asset's price dynamics or on the distribution assumed for the risk-neutral density. Ait-Sahalia and Duarte (2003) used a nonparametric option pricing model under shape restrictions to estimate the risk-neutral density. It was based on Mammen (1991)'s kernel regression model.

A more recent work related to the "model-free" regression is Pandher (2007). He presented a simple empirical approach to modelling and forecasting market option prices using localized option regression (LOR). He considered two classes of localized option regressions: structural and reduced form models. The model locally projects derivative prices on the state process of the underlying asset price, strike price, implied volatility and risk-free interest rate. The state space includes linear, quadratic and interaction terms arising among the variables. Both in-sample and out-of-sample test results showed that LOR is a relatively good model to price the options. To our best of our knowledge, it is the most closely related paper to the regression models we use in Chapter 3. Although the LOR is a good model for predicting option prices, it has limitations. In the LOR model, Pandher (2007) used implied volatility as an explanatory variable to explain option price. This seems to be a strange approach, because implied volatility is obtained directly from option price. Our regression models in Chapter 3 use historical volatility instead of implied volatility.

Andreou et al. (2008) consider non-parametric (Artificial Neural Network, ANN) and semi-parametric (CS, Corrado and Su, 1996) option pricing models, and compare these to the parametric model (BS). They use both historical and implied volatility as predictors; for historical volatility they use the past 60 days. They find that the Black-Scholes-based hybrid ANN models outperform (in predictive performance) the standard neural networks and the parametric ones.

Additive Models

Additive models (which will be used extensively in Chapter 3) were introduced by Stone (1985) and are explained in more simple terms by

Hastie and Tibshirani (1990). The components of the additive model will be obtained using B-splines, which are covered by de Boor (1987). A scatterplot smoothing technique which will be used repeatedly is the locally weighted scatterplot smoother (Lowess) due to Cleveland et al. (1979).

Implied Volatility functions and Practitioner's Black-Scholes (PBS)

Dumas at al. (DFW, 1998) considered the deterministic volatility function (DVF) option pricing model, which assumes that volatility is a deterministic function of asset price and time. They find that this approach is no better than a simple procedure that smooths Black-Scholes implied volatilities across strike price and time to expiry. They refer to this simpler technique as the Ad-hoc model. The Ad-hoc model has since become very popular and is commonly known as the Practitioner Black Scholes (PBS) model. The PBS has become a benchmark for model comparisons in the literature.

Aït-Sahalia and Lo (1998) apply non-parametric regression methods (in the form of Kernel regression) to the problem of estimating the implied volatility surface. Most other authors have used linear regression for this purpose.

Christoffersen and Jacobs (2004) consider the important question of which loss function should be used when estimating and evaluating option valuation models. They emphasise the importance of using the same estimation loss function for all models, and of using the same loss function for evaluation as for estimation. They also suggest that the choice of loss function should be guided by the objective of the exercise: that is, whether the objective is e.g. speculating, hedging or market-making. They illustrate the importance of the loss function in an application of the Practitioner Black-Scholes (PBS) model. Typically, the PBS is implemented using an implied volatility loss function, but evaluated using a pricing loss function. They demonstrate that the PBS performs much better when it is implemented using the same loss function as used for evaluation.

Christoffersen and Jacobs (2004, p.298), make a very important point when outlining the PBS procedure: "simply plugging [fitted implied volatility] into the Black-Scholes formula will yield a biased estimate of the observed call price". However they go ahead and do exactly this. Others who use PBS appear to do the same. This motivates us to apply the smearing technique (Duan, 1983) to correct this bias. We will do this in Chapter 3.

Berkowitz (2010) takes a close look at the PBS, which he refers to as the "ad-hoc Black-Scholes method". He shows that the PBS procedure can be used to provide an arbitrarily accurate approximation to the true option pricing formula, at a given point in time, given a polynomial of sufficient order, and a sufficiently large sample. The approximation cannot be expected to hold over time, and because of this he recommends frequent re-calibration of the volatility surface, which produces continually new approximations. He uses simulations to examine the importance of the sample size, the order of the polynomial, and the recalibration frequency. He finds that: (1) The best performing models require only linear and quadratic terms (although it seems strange that he includes an interaction variable K*T in the quadratic model but not in the cubic model); (2) A sample size of 64 options at each time-point in time is sufficient to generate reasonable pricing accuracy; (3) Frequent recalibration is more important than the correctness of the model.

Hull and Suo (2003) find that the PBS approach based on European options does not necessarily price other types of options accurately (e.g. American Options). They define “model risk” as the risk arising from the use of an inadequate model.

Andreou et al. (2014) consider a number of regression-based implied volatility models. They consider symmetric and asymmetric models. Asymmetric means assuming the function is different for ITM and OTM options. Asymmetric models estimated separately for calls and puts are found to provide the best in-sample performance. Symmetric models using the log of the strike price are found to provide the best out-of-sample performance.

Stochastic volatility models, jump-diffusion, and the GARCH family

Bates (1991) investigates whether option prices over the period 1985-7 contained evidence of expectations of the October 1987 crash. He finds evidence of this, both in over-pricing of OTM put options, and in the jump-diffusion parameters, which indicated that implicit distributions were negatively skewed over a period which included the crash. Bates (1996) develops an efficient method for pricing American options on a stochastic volatility/jump-diffusion process. He finds that the volatility smile is explained by “jump fears”.

Bakshi et al. (1997) develop an option pricing model, named SVSI-J, that allows volatility, interest rates, and jumps all to be stochastic. They claim that the pricing formula is closed form, although one of the equations (equation (9) in the article) contains an integral so it does not seem to be “closed-form”. The model contains standard models as special cases, including Black-Scholes (BS), stochastic interest rate (SI), and stochastic volatility (SV), and stochastic volatility random-jump models (SV-J). The SVSI-J model is evaluated in comparison to the other models on three different criteria: consistency of implied volatility with relevant time series data; out-of-sample pricing accuracy (based on predicted price of each option using the previous day's implied parameters and implied volatility); hedging performance (where optimal hedges are obtained using the current implied parameters, and then the hedges are liquidated after one day or five days). They consider pricing accuracy to reflect a models *static* performance, while hedging accuracy reflects *dynamic* performance. The overall finding is that the modelling of stochastic volatility is much more important than that of interest rates and jumps.

Heston and Nandi (2000) claim to develop a closed-form option valuation formula for the situation in which the underlying follows a GARCH process. However, this claim is confusing because the formula is presented in Equation (11) of the article and this formula involves an integral which can only be evaluated using a numerical procedure. Hence the formula is not "closed-form".

"Implied" GARCH parameters have been estimated using option prices by Engle and Mustafa (1992).

Duan et al. (2006) use the Edgeworth expansion to derive analytical approximations for the pricing of European options in the GARCH framework, including GJR-GARCH (TGARCH) and EGARCH. The approximation is essentially the Black-Scholes formula with two additional terms, one for the skewness and one for the kurtosis of the cumulative return on the underlying. They assess the performance of the approximation on a "test pool" of randomly generated options. They compare the approximation to the Monte Carlo value using RMSE. They find evidence that their approximation is adequate for shorter-maturity options.

The three varying volatility models we will use in Chapters 4 and 5 are ARCH (Engle, 1982), GARCH (Bollerslev et al., 1986) and TARCH (Zakoian, 1993; Glosten et al., 1993).

Monte Carlo Methods

The Monte Carlo method applied to the problem of option valuation has been discussed by Glasserman (2003) and Fink & Fink (2006). However, the method has mainly been used for options for which closed-form valuations are not available, for example American options. The studies usually rely on the standard assumption of a geometric Brownian motion process with constant return volatility.

Some studies go further and apply the Monte Carlo method to situations of varying volatility, for example Duan et al. (2006). However, that study applies the method to randomly generated options. To our knowledge, the Monte Carlo method has not been used to value real market options under assumptions of varying volatility. This task is undertaken in Chapter 5 of this thesis.

Chapter 2 Overview of Existing Option Valuation Methods

There are a number of established methods for option valuation.The main purpose of this chapter is to survey these methods, and to focus on the two most popular, namely the binomial-tree method, and the Black-Scholes method. The relationship between these two methods is discussed, and it is demonstrated using a specific example that they can give the same valuation under some circumstances.

It is particularly important to describe the Black-Scholes method in some detail because it is used as a benchmark for comparison when we evaluate the methods of option valuation developed in later chapters. Our discussion of the Black-Scholes method includes a heuristic derivation of the Black-Scholes formula for a European call option.

We start by introducing definitions and notation.This material is thoroughly covered in Financial Mathematics textbooks such as Hull (2011). However the notation chosen here sometimes differs from Hull (2011).

2.1 Definitions and notation

A *(European) Call Option* is a security that gives its owner the right, but not the obligation, to *purchase* a specified asset for a specified price, known as the *strike price* or *exercise price*, at some date in the future (the *expiry date*).

A *(European) Put Option* is a security that gives its owner the right, but not the obligation, to *sell* a specified asset for a specified price, known as the *strike price* or *exercise price*, at some date in the future (*expiry*).

We will refer to the asset on which the option is written as the *underlying*.

The owner, or *holder*, of an option – who is said to adopt a *long* position – acquires the option by paying the *option price* to the *writer* – who is said to adopt a *short* position. If the *holder* of a *call option* chooses to *exercise* the *option*, he pays the *strike price* to the *writer* in exchange for the asset, at expiry. If the *holder* of a *put option* chooses to *exercise* the *option*, he delivers the asset to the *writer* at expiry, who simultaneously pays the *strike price* to the *holder*.

European Options can only be exercised on the *expiry date*.*American options*, in contrast, can be exercised at any time up to the *expiry date*. In this thesis, we are mainly interested in *European Options* because they are more straightforward to analyse.

Options that expire unexercised are said to *die*, and are worthless.

The following notation will be used:

t is the current date

T is the expiry date

$\tau = T - t$ is the time to expiry.

S_t is the current (underlying) stock price.

S_T is the stock price at expiry.

K is the strike price.

r is the risk-free rate of interest.

c_t is the current price (or the current value) of a Call Option

p_t is the current price (or the current value) of a Put Option

If you are the holder of a call option, you want the stock price at expiry to exceed the strike price.Then, you exercise the option to buy at the strike price, and immediately sell at a profit S_T - K.If the stock price at expiry is less than the strike price, you let the option die, and your payoff is zero.

The payoff from holding a call option is therefore:

$$\max\left(0, S_T - K\right)$$

Payoff diagrams are graphs showing the payoff from holding an option against the stock price at expiry, S_T. Figure 2.1 shows the payoff diagram for a call option with strike K.

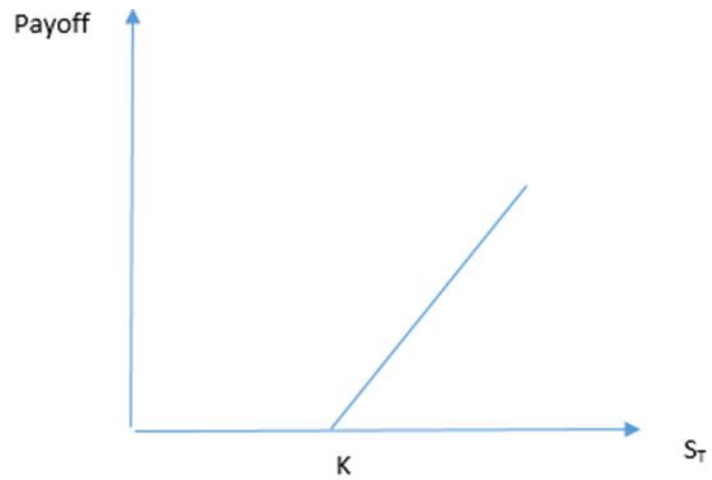


Figure 2.1 Payoff diagram for a Call Option with strike K

A call option for which the current price S_t is *above* the strike price *K* is said to be *in the money (ITM)*. A call option for which the current price S_t is *below* the strike price *K* is said to be *out ofthe money (OTM)*. A call option for which the current price S_t equals the strike price *K* is said to be *at the money (ATM)*.

Moneyness (for a call option) is defined as the ratio of the current price to the strike price: $m = S_t / K$.

The payoff from holding a Put Option is : $\max\left(0, K - S_T\right)$

Figure2.2 shows the payoff diagram for a Put Option with strike *K*.

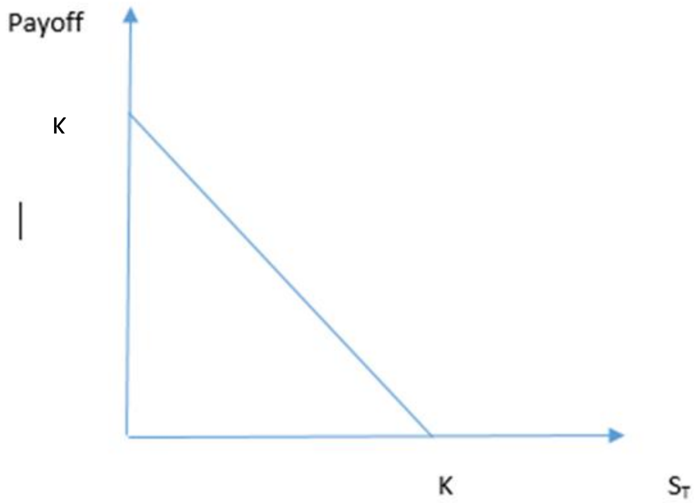


Figure 2.2 Payoff diagram for a put option with strike *K*

One difference of put options from call options is that the payoff from holding a put option cannot be above *K*, while the payoff from holding a call is unlimited.

Moneyness for a put option is $m = K / S_t$.

The sign of moneyness (for both types of option) tells us whether the option is ITM (*m>1*), ATM (*m=1*), or OTM (*m<1*).

Finally, the *current value* of a call option is the expected value of the payoff at expiry, at the risk-free interest rate:

$$c_t = \exp(-r\tau) E\left[\max(0, S_T - K)\right]$$

And the *current value* of a put option is the expected value of the payoff at expiry, discounted at the risk-free interest rate:

$$p_t = \exp(-r\tau) E\left[\max(0, K - S_T)\right]$$

In order to compute the expectations in these two formulae, we need to make an assumption about how the price of the underlying S_t evolves over time. One such assumption gives rise to the binomial formula; another assumption gives the Black-Scholes formula. These formulae are derived in the following sections.

2.2 Binomial Tree Model

Probably the simplest technique for valuing an option is based on the construction of a binomial tree. A binomial tree is a simple method for modelling the behaviour of the underlying stock price or stock index. It simply assumes that in any discrete time period, the stock price has a fixed probability of going up by a fixed proportion, and a fixed probability of going down by a fixed proportion. We will refer to this sort of process as a "simple random walk". The simple random walk is a discrete time version of geometric Brownian motion.

The number of steps in the binomial tree is set, and then the probability distribution of the price at expiry can be computed, and hence the value of an option can be computed. The Binomial tree method can be used to value both European and American options. In this section, we will take a close look at the method, and explain how it can be used to value European options.

We will again denote current stock price as S_t. First of all, we will assume only two steps in the binomial tree; then we will consider trees with more than two steps; finally, we will demonstrate the very useful "online binomial tree calculator" in the valuation of European options.

2.2.1 Two-step binomial tree

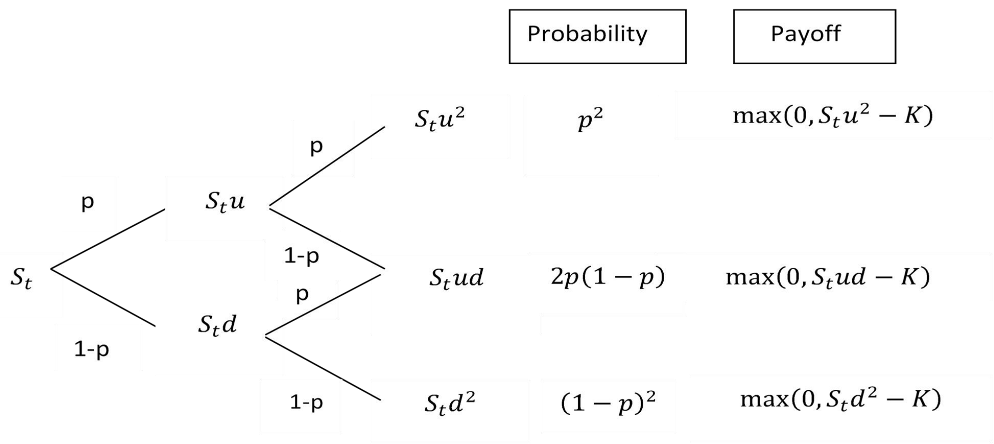


Figure 2.3 A two-step Binomial Tree

Figure 2.3 shows a two-step binomial tree that could be used to value a European call option with strike *K*. The current price of the underlying is S_t. The time to expiry is divided into two periods of length $\tau/2$. In each period, the stock price is assumed either to rise by a multiple *u* *(u>1)* with probability *p*, or decrease by a multiple *d (d<1)* with probability *(1-p)*. The possible pay-offs at expiry are shown in the final column of Figure 2.3, and the corresponding probabilities are shown in the penultimate column.

2.2.2 Calculation of the value of a European option using the binomial tree

The value of an option is the expected payoff at expiry, discounted using the risk-free rate. For the call option used as an example in Section 2.2.1, the value is:

$$c = \Big[\left(\max\left(0, S_t u^2 - K\right)p^2\right) + \left(\max\left(0, S_t ud - K\right)\left(2p\left(1-p\right)\right)\right) \\ + \max\left(0, S_t d^2 - K\right)\left(1-p\right)^2\Big]\exp\left(-r\tau\right)$$

If the option was instead a put option with strike K, the value of the option would be:

$$p = \Big[\left(\max\left(0, K - S_t d^2\right)p^2\right) + \left(\max\left(0, K - S_t ud\right)\left(2p\left(1-p\right)\right)\right) \\ + \max\left(0, K - S_t d^2\right)\left(1-p\right)^2\Big]\exp\left(-r\tau\right)$$

If the time to expiry is divided into more than two steps, the method for calculating the option value would be the same as above but the formulae would contain more terms.

The above option values are functions of *u*, *d*, and *p*. However, it is possible to write the formulae without *p*. This is because the value of *p* can be deduced from *u* and *d*. This is demonstrated in the next sub-section.

2.2.3 Calculation of the probability and volatility

2.2.3.1 Calculation of probability of "up"

Consider the one-step binomial tree shown in Figure2.4.

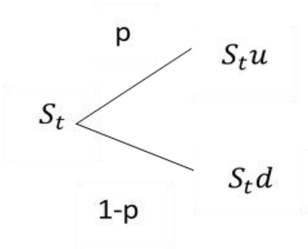


Figure 2.4 A one-step Binomial Tree

If we assume risk-neutrality, and no-arbitrage, it must be the case that the expected price of the underlying at expiry must be equal to the value of the investment resulting from investing S_t at the risk-free rate between *t* and expiry. In other words:

$$\left[S_t u \times p\right] + \left[S_t d \times (1-p)\right] = S_t \times \exp(r\tau)$$

It follows that:

$$p = \frac{\exp(r\tau - d)}{u - d} \tag{2.1}$$

(2.1) shows how the "up" probability *p* can be deduced from knowledge of *u* and *d*.

In practice, the values of *u* and *d* are not known. They need to be deduced from the annual volatility σ which is usually known. This issue is considered in the next sub-section.

2.2.3.2 Relationship between binomial-tree volatility parameters and annual volatility parameter

Consider the stock return measured as the proportion by which the stock price changes between t and $t+\tau$. This is $S_{t+\tau}/S_t$. The variance of this return is given by:

$$\mathrm{V}\left(\frac{S_{t+\tau}}{S_t}\right)=\sigma^2\tau$$

In the two-step binomial tree model used in the last sub-section, this variance is:

$$\mathrm{V}\left(\frac{S_{t+\tau}}{S_t}\right)=E\left(\frac{S_{t+\tau}}{S_t}\right)^2-\left[E\left(\frac{S_{t+\tau}}{S_t}\right)\right]^2=pu^2+(1-p)d^2-\left[pu+(1-p)d\right]^2$$

The tree's parameters should match the volatility of stock price, hence we have:

$$pu^2+(1-p)d^2-\left[pu+(1-p)d\right]^2=\sigma^2\tau \tag{2.2}$$

Substituting equation (2.1) into (2.2), we have:

$$\left[\frac{\exp(r\tau)-d}{u-d}*u^2\right]+\left[\frac{u-\exp(r\tau)}{u-d}*d^2\right]-\left[\frac{\exp(r\tau)-d}{u-d}*u+\frac{u-\exp(r\tau)}{u-d}*d\right]^2=\sigma^2\tau$$

Then we have:

$$\begin{aligned}&\frac{1}{u-d}*\left[\exp(r\tau)*u^2-du^2+ud^2-\exp(r\tau)*d^2\right]\\&-\frac{1}{u-d}*\left[\exp(r\tau)*u-du+ud-\exp(r\tau)*d\right]^2\\&=\sigma^2\tau\end{aligned}$$

Then we get:

$$\frac{1}{u-d}*(u-d)*\left[\exp(r\tau)*(u+d)-ud\right]-\frac{1}{u-d}*(u-d)*\exp(r\tau)^2=\sigma^2\tau$$

Then we have:

$$\left[\exp(r\tau)*(u+d)-ud\right]-\left[\exp(r\tau)\right]^2=\sigma^2\tau$$

Finally we have:

$$e^{r\tau}(u+d)-ud-e^{2r\tau}=\sigma^2\tau$$

Using Taylor's series expansion, we have:

$$u=e^{\sigma\sqrt{\tau}};d=e^{-\sigma\sqrt{\tau}} \tag{2.3}$$

(2.3) shows how the tree's volatility parameters (*u* and *d*) can be deduced from knowledge of the annual volatility σ and the time to expiry τ.

2.2.3.3 General binomial-tree option price formula

Consider a tree with n steps.

Over the first time interval:

$$S_{t+\frac{\tau}{n}}=S_t u \quad with \quad probability \quad p$$

$$S_{t+\frac{\tau}{n}}=S_t d \quad with \quad probability \quad (1-p)$$

Over the first two intervals:

$$S_{t+\frac{2\tau}{n}}=S_t u^2 \quad with \quad probability \quad p^2$$

$$S_{t+\frac{2\tau}{n}}=S_t d^2 \ with \ probability (1-p)^2$$

$$S_{t+\frac{2\tau}{n}}=S_t u^k d^{n-k} \ with \ probability \ 2p(1-p)$$

After n intervals, expiry is reached, since $S_T=S_{t+\tau}$.

The distribution of the stock price at expiry is given by:

$$S_T=S_t u^k d^{n-k} \quad with \quad probability \quad {}_nC_k p^k(1-p)^{n-k}, k=0,\ldots,n$$

The values of a call option and a put option are given by:

$$c_t=\left\{\sum_{k=0}^{n}\max\left(0,S_t u^k d^{n-k}-K\right){}_nC_k p^k(1-p)^{n-k}\right\}\exp(-r\tau)$$

$$p_t=\left\{\sum_{k=0}^{n}\max\left(0,K-S_t u^k d^{n-k}\right){}_nC_k p^k(1-p)^{n-k}\right\}\exp(-r\tau)$$

2.2.4 Online binomial tree calculators

If there are only a few steps in the tree, it is easy to apply the above method in order to value an option. If there are many steps, say 100, it becomes hard to apply the above method manually. However we can use an on-line "binomial tree calculator". The original Cox, Ross, & Rubinstein (1979) tree is a popular example. This calculator returns the value of a European or American option for given parameter values, and also provides a graphical display of the tree structure used in the calculations. A screenshot of the input window is shown in Figure 2.5.

The user inputs numbers into the boxes (strike price, volatility, underlying asset price, interest rate, days to expiry, dividend, number of tree steps), selects the option type and exercise type, then clicks on "calculate". One limitation of this on-line calculator is that a single user is not permitted to use it (for free) more than six times in a single day.

Strike price:		**Volatility:**		%
Underlying asset price:		**Interest rate:**		%
Days to expiration:				
Dividend: Enter an amount ($.cc) for discrete dividend, or an annual yield (eg 3.5 = 3.5% pa)		**Days to ex-dividend:** Enter days for discrete dividend; leave blank or zero for yield		
Option type:	Call Put	**Exercise style:**	American European	
No. Tree Steps (1- 150):		Calculate		

Source: http://www.hoadley.net/options/binomialtree.aspx?tree=B

Figure 2.5 Binomial Tree on-line Calculator

Let us take as a real example a European call option traded on 14th Jan 2002 written on the S&P 500 Index, with strike 1075 and 44 days to expiry. On 14th Jan 2002, the current underlying Index is 1138, the risk-free interest rate is 0.0213, and the volatility (based on the 60-day standard deviation of returns) is 0.1605. Since we are dealing with an Index and not a single stock, the dividend is 0.

When inputting the above numbers into the online calculator, the resulting numbers are too large to appear on the screen. For this reason, we divide the price variables by 10 before inputting them. Hence, we are considering a current price of S_t=113.8 and a strike of 107.5.

We start with 2 steps, then we extend to 10 steps, 50 steps, 100 steps and 150 steps. Figure 2.6 shows us the calculation result with 2 steps. Figure 2.7 shows us the calculation result with 10 steps. If we use the on-line calculator, the maximum number of steps that can be shown on the screen is 10.

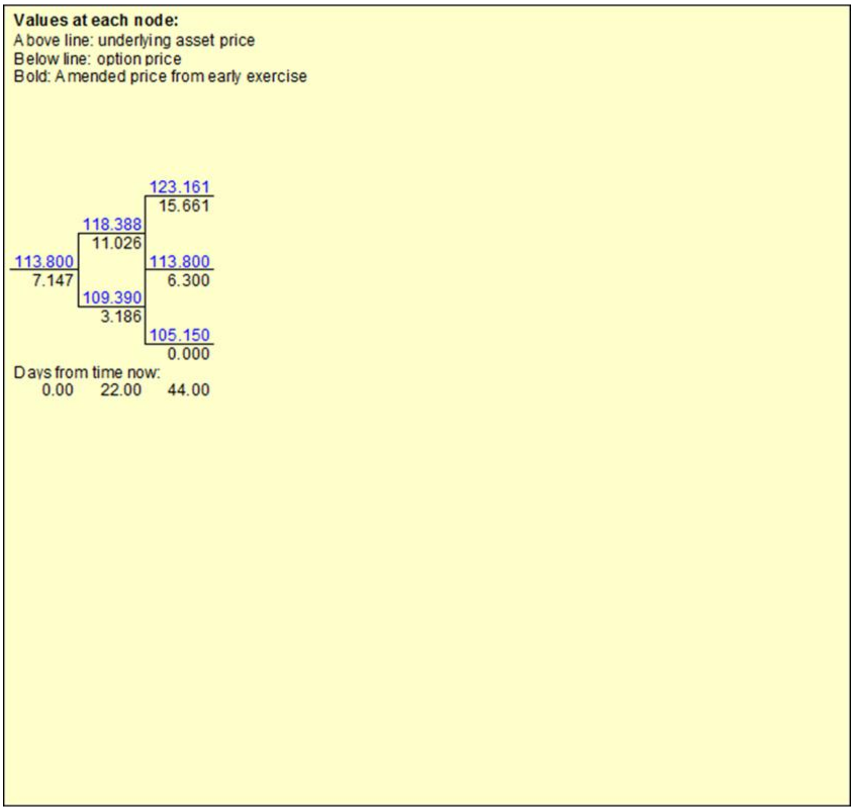


Figure 2.6 Binomial Tree on-line Calculator Result with 2 steps

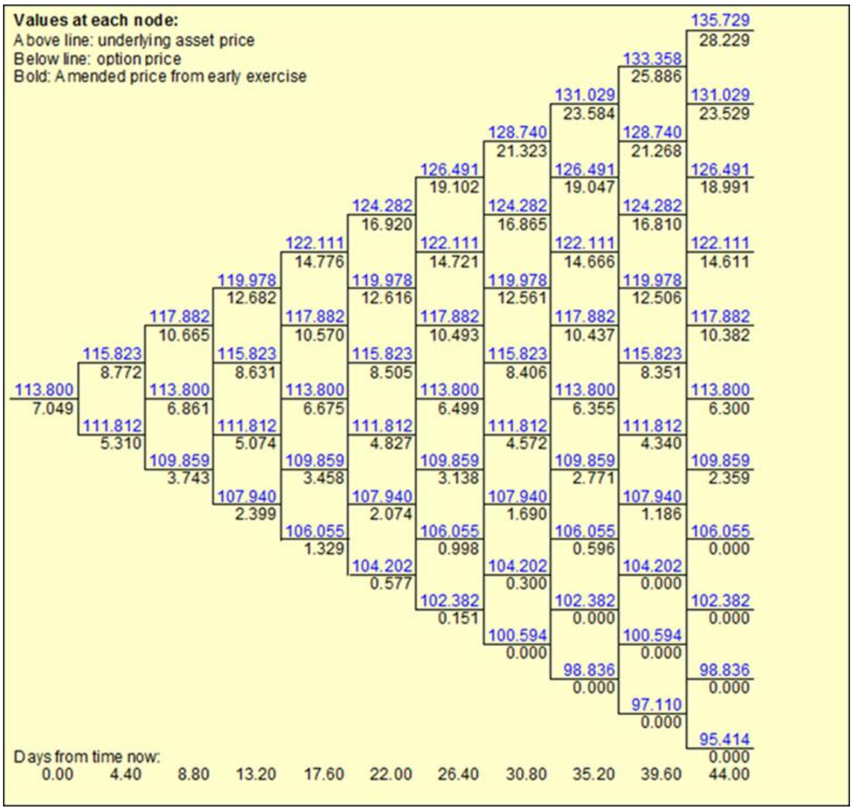


Figure 2.7 Binomial Tree on-line Calculator Result with 10 steps

From Figure 2.6, we see that the value of this European call option with 2 steps is 7.15; from Figure 2.7, we see that the result with 10 steps is slightly lower at 7.05. Remember that we divided the underlying stock price and the strike price by 10, so if we convert to the original units, the option value with 2 steps is 71.5 and with 10 steps 70.5. Does the binomial tree online calculator give the same result as the formula derived in section 2.2.3? Let us check this. The information we have about the option is shown in Table 2.1.

S_t	K	r_f	τ	σ
1138	1075	2.13%	44 days	16.05%

Table 2.1 Information for the Option Example

Following the stages of the binomial tree calculationas discussed in section 2.2.3, we first need to calculate the values of *u* and *d*. According to equation 2.3, we have:

$$u = e^{\sigma\sqrt{\tau}} = e^{0.1605\sqrt{44/360}} = 1.04$$

$$\mathrm{d} = e^{-\sigma\sqrt{\tau}} = e^{-0.1605\sqrt{44/360}} = 0.96$$

Secondly, we need to distribute a two-step binomial tree.

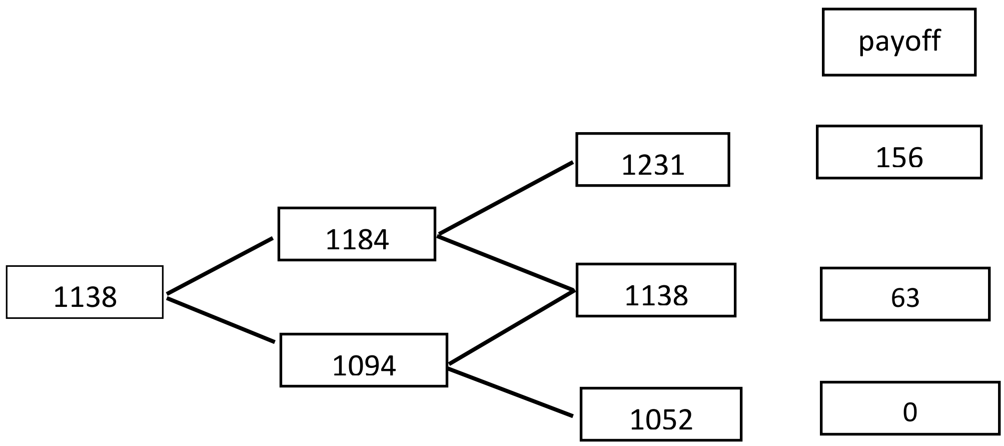


Figure 2.8 A two-step Binomial Tree example

Finally, we calculate the value of call option:

$$\left[156*0.25+63*0.5+0*0.25\right]*\exp\left(-0.0213*\frac{44}{360}\right)=70.5$$

When we use the formula to calculate the same European call option, we find the value is \$70.5 which is close to the result from online calculator.

We will use the same calculator to calculate the value of the European call option with 10 steps, 50 steps, and 100 steps. The results are shown in Table 2.2. We see that the call value falls when the number of steps rises, but it seems to be converging to a value around 70.2.

steps	2	10	50	100
Call Value	71.5	70.5	70.3	70.2

Table 2.2Binomial Tree on-line Calculator Result Results with different numbers of tree-steps

One major advantage of the binomial model is that it can be used to value American Options. This is because with the binomial model, it is possible to check at every node of the tree for the profitability of early exercise. When a node is found to have profitable early exercise, the part of the tree coming off that node is removed and replaced with the intrinsic value at that point. A limitation is that the speed of calculation is relatively slow. It is not a practical approach for the calculation of thousands of prices in a few seconds.

2.2.5 Binomial Tree with TGARCH Effect

In Sections 2.2.1-2.2.4, we discussed the binomial tree on the basis of the assumption of a simple random walk. However, later in the thesis, we use the TGARCH model to predict the option price. The TGARCH model (Threshold ARCH, or "GJR-GARCH", Glosten et al., 1993) is a model that allows the effects of good and bad news to have different effects on volatility. A feature of stock prices is that "bad" news tends to have a larger effect on volatility than "good" news. The tendency for volatility to fall when price rises and to rise when price falls is known as the "leverage effect" (Enders, 2004). In this section, we consider whether the assumption of a TGARCH volatility process instead of a simple random walk, in the context of a binomial tree, changes the valuation of some options. We will use a 2-step binomial tree model.

We assume that time to expiry is one year, and the year is divided into two 6-month steps. We first assume a simple random walk: the current stock price is 100, and at each step the stock price either rises by amount 10 with probability 50%, or falls by amount 10 probability 50%. The risk-free interest rate is 5%. The following diagram illustrates the tree.

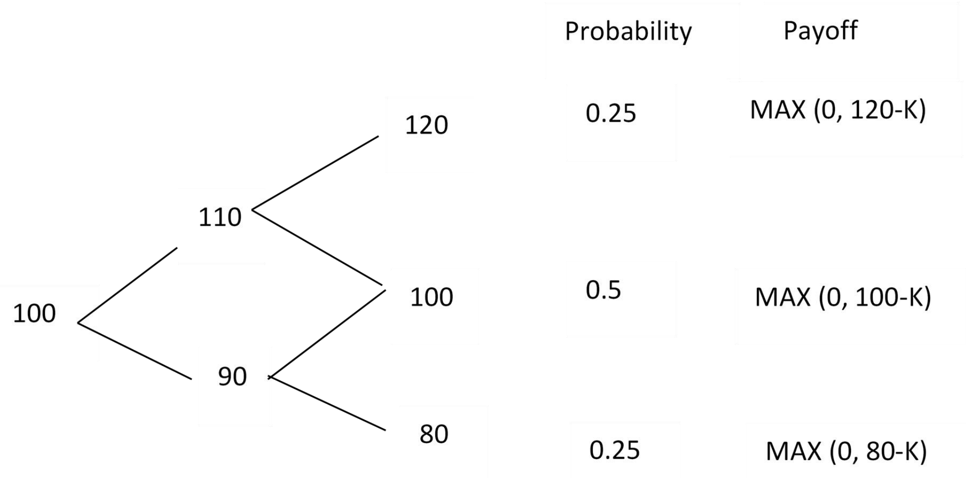


Figure 2.9 (a) A two-step binomial tree with simple random walk assumption

Consider two call options: an OTM call with strike K=110, and an ITM call with strike K=90. The values of these two options are computed to be:

$$OTM call(K=110):\quad c=\left[10*0.25\right]*\exp\left(-0.05*1\right)=2.38$$

$$ITM call(K=90):\left[30*0.25+10*0.5\right]*\exp\left(-0.05*1\right)=11.89$$

Next we assume that the stock price process follows a TGARCH process. We assume that if the price rises in the first step, the second step is the same as for the simple random walk. However, if the price falls in the first step, volatility in the second step doubles, and the price changes by 20 instead of 10. This gives the following tree.

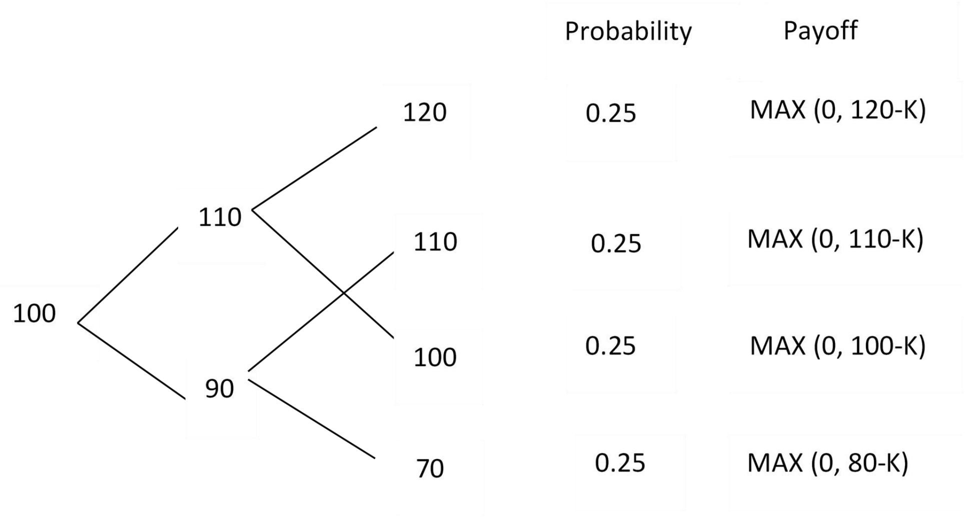


Figure 2.10 (b) A two-step binomial tree with TGARCH assumption

Under the TGARCH assumption, the values of the two call options are:

$$OTM call(K=110): \quad c=[10*0.25]*\exp(-0.05*1)=2.38$$

$$ITM call(K=90): [30*0.25+20*0.25+10*0.25]*\exp(-0.05*1)=14.27$$

We see that the value of the OTM call is unaffected by the TGARCH assumption. However, the value of the ITM call is around 20% higher under the TGARCH assumption.

Next, we consider put options. If the stock price follows a simple random walk, let us consider two put options: an OTM put with strike K=110, and an ITM put with strike K=90. The values of these two options are computed to be:

$$OTM\ put(K=90): p=[10*0.25]\exp(-0.05*1)=2.38$$

$$ITM\ put\ (K=110): p=[10*0.5+30*0.25]\exp(-0.05*1)=11.89$$

Under the TGARCH assumption, the values of the two put options are:

$$OTM\ put\ (K=90): p=[20*0.25]\exp(-0.05*1)=4.75$$

$$ITM\ put\ (K=110): p=[10*0.25+40*0.25]\exp(-0.05*1)=11.89$$

We see that the value of the ITM put is unaffected by the TGARCH assumption. However, the value of the OTM put is around 50% higher under the TGARCH assumption.

In Chapter 4, we find strong evidence of a TGARCH process in daily stock index data. This means that, by the above analysis, ITM call options and OTM put options are likely to appear over-priced if the random walk (e.g. Black-Scholes) is assumed. This is the well-known "volatility smirk" that is indeed seen in option price data: ITM calls tend to have higher implied volatilities than OTM calls, and OTM puts tend to be higher implied volatilities than ITM puts, which is equivalent to saying that they are over-priced under Black-Scholes assumptions. The simple analysis in this section is proposing a simple explanation for the volatility smirk and table 2.3 gives us a summary of the results.

	ITM Call	OTM Call	ITM Put	OTM Put
Simple Random Walk	11.89	2.38	11.89	2.38
TGARCH	14.27	2.38	11.89	4.75

Table 2.3 comparison results of simple random walk with TGARCH

2.3 The Black-Scholes Method

2.3.1 Heuristic derivation of Black-Scholes formula

The Black-Scholes formula is considered as the most prominent achievement in option pricing theory. The formula is used to calculate the theoretical European Option price (ignoring dividends paid during the life of the option) using the five key determinants: price of underlying, strike price, volatility, time to expiry and risk-free interest rate.

Most textbooks use differential calculus and Ito's lemma to derive the Black-Scholes valuation formula. In this section, we will derive the Black-Scholes valuation formula, using the "heuristic derivation", presented by Adams et al. (2003). This method is more straightforward since all that is required is a formula for the truncated mean of the lognormal distribution.

We will start the derivation the Black-Scholes formula by listing its assumptions. The Black-Scholes formula is based on 7 important assumptions:

(1) Stock Returns are normally distributed with known mean and variance;

(2) The annual volatility is constant;

(3) No arbitrage argument;

(4) The risk-free interest rate on the money market is known and constant;

(5) No trading cost;

(6) No taxes;

(7) No dividend

A consequence of assumption 2 is that the value of the stock price changes between t and $t+\Delta t$ according to the following equation:

$$S_{t+\Delta t} = S_t \exp\left(\mu\Delta t + \sigma\sqrt{\Delta t}Z\right)$$

Where $Z \sim N(0,1)$

The price at expiry is:

$$S_T = S_{t+\tau} = S_t \exp\left(\mu\tau + \sigma\sqrt{\tau}Z\right)$$

It follows that:

$$\frac{S_T}{S_t} = \exp\left(\mu\tau + \sigma\sqrt{\tau}Z\right)$$

So:

$$\frac{S_T}{S_t} \sim \log normal\left(\mu\tau, \sigma^2\tau\right)$$

From the properties of the lognormal distribution, we know that:

$$E\left(\frac{S_T}{S_t}\right) = \exp\left(\mu\tau + \frac{\sigma^2\tau}{2}\right) = \exp\left[\left(\mu + \frac{\sigma^2}{2}\right)\tau\right]$$

This is a formula for the expected proportional increase in the stock price from the present to expiry.

Another expression for the same expectation is:

$$E\left(\frac{S_T}{S_t}\right) = \exp\left(r\tau\right)$$

This is because, in a risk-neutral world, the average return of all stocks equals the risk-free return.

For both of these formulae to be correct, it must be the case that:

$$\mu + \frac{\sigma^2}{2} = r$$

Let us first consider a European call option. The value of a European call option with strike K and time to expiry τ is:

$$\exp\left(-r\tau\right)E\left[\max\left(0, S_T - K\right)\right]$$

The expectation in this formula can be re-written:

$$\begin{aligned}
E\left[\max\left(0,S_T-K\right)\right] &= E\left(0\middle|S_T<K\right)\times P\left(S_T<K\right)+E\left(S_T-K\middle|S_T>K\right)\times P\left(S_T>K\right)\\
&= E\left(S_T-K\middle|S_T>K\right)\times P\left(S_T>K\right)\\
&= \left[E\left(S_T\middle|S_T>K\right)\times P\left(S_T>K\right)\right]-\left[K\times P\left(S_T>K\right)\right]\\
&= \left[S_t\times E\left(\frac{S_T}{S_t}\middle|\frac{S_T}{S_t}>\frac{K}{S_t}\right)\times P\left(\frac{S_T}{S_t}<\frac{K}{S_t}\right)\right]-\left[K\times P\left(\frac{S_T}{S_t}>\frac{K}{S_t}\right)\right]
\end{aligned}$$

Since we know that $\frac{S_T}{S_t}\sim \log normal\left(\mu\tau,\sigma^2\tau\right)$, at this point, we require a formula for the truncated mean of the lognormal distribution. This formula is known to be:

If $Y\sim Lognormal\left(\mu,\sigma^2\right)$, then:

$$E\left(Y\,|\,Y>c\right)=\frac{\Phi\left(\frac{\mu-\ln c}{\sigma}+\sigma\right)}{\Phi\left(\frac{\mu-\ln c}{\sigma}\right)}\exp\left(\mu+\frac{\sigma^2}{2}\right)$$

Applying this formula to the conditional expectation, we obtain:

$$E\left(\frac{S_T}{S_t}\middle|\frac{S_T}{S_t}>\frac{K}{S_t}\right)=\frac{\Phi\left(\frac{\mu\tau-\ln\left(\frac{K}{S_t}\right)}{\sigma\sqrt{\tau}}\right)+\sigma\sqrt{\tau}}{\Phi\left(\frac{\mu\tau-\ln\left(\frac{K}{S_t}\right)}{\sigma\sqrt{\tau}}\right)}\exp\left(\mu\tau+\frac{\sigma^2\tau}{2}\right) \tag{2.4}$$

We also require:

$$P\left(\frac{S_T}{S_t}>\frac{K}{S_t}\right)=P\left(Z>\frac{\ln\left(\frac{K}{S_t}\right)-\mu\tau}{\sigma\sqrt{\tau}}\right)=\Phi\left(\frac{\mu\tau-\ln\left(\frac{K}{S_t}\right)}{\sigma\sqrt{\tau}}\right) \tag{2.5}$$

Then we obtain:

$$E\left[\max\left(0,K-S_T\right)\right]=K\Phi\left(\frac{\ln\left(\frac{K}{S_t}\right)-\mu\tau}{\sigma\sqrt{\tau}}\right)-S_t\Phi\left(\frac{\ln\left(\frac{K}{S_t}\right)-\mu\tau}{\sigma\sqrt{\tau}}-\sigma\sqrt{\tau}\right)\exp\left(\mu\tau+\frac{\sigma^2\tau}{2}\right)$$

At the beginning of this section, we assume that there is no arbitrage being used, so we have the following equation:

$$\mu+\frac{\sigma^2}{2}=r, so \mu=r-\frac{\sigma^2}{2}, \exp\left(\mu\tau+\frac{\sigma^2\tau}{2}\right)=\exp\left(r\tau\right)$$

Also, since it is desirable to write the formula without the parameter μ, we substitute $\mu=r-\frac{\sigma^2}{2}$. We obtain:

$$E\left[\max\left(0,S_T-K\right)\right]=S_t\Phi\left(\frac{\left(r-\frac{\sigma^2}{2}\right)\tau-\ln\left(\frac{K}{S_t}\right)}{\sigma\sqrt{\tau}}+\sigma\sqrt{\tau}\right)\exp\left(r\tau\right)-K\Phi\left(\frac{\left(r-\frac{\sigma^2}{2}\right)\tau-\ln\left(\frac{K}{S_t}\right)}{\sigma\sqrt{\tau}}\right)$$

All that is now needed is the discount factorexp(-rτ). This completes the formula for the present value of a call option:

$$
\begin{aligned}
c_t &= \exp(-r\tau)E[\max(0, S_T - K)] \\
&= \exp(-r\tau)S_t\Phi\left(\frac{\mu\tau - \ln\left(\frac{K}{S_t}\right)}{\sigma\sqrt{\tau}} + \sigma\sqrt{\tau}\right)\exp(r\tau) - \exp(-r\tau)K\Phi\left(\frac{\mu\tau - \ln\left(\frac{K}{S_t}\right)}{\sigma\sqrt{\tau}}\right) \\
&= S_t\Phi\left(\frac{\mu\tau - \ln\left(\frac{K}{S_t}\right)}{\sigma\sqrt{\tau}} + \sigma\sqrt{\tau}\right) - \exp(-r\tau)K\Phi\left(\frac{\mu\tau - \ln\left(\frac{K}{S_t}\right)}{\sigma\sqrt{\tau}}\right) \\
&= S_t\Phi\left(\frac{\left(r - \frac{\sigma^2}{2}\right)\tau - \ln\left(\frac{K}{S_t}\right)}{\sigma\sqrt{\tau}} + \sigma\sqrt{\tau}\right) - \exp(-r\tau)K\Phi\left(\frac{\left(r - \frac{\sigma^2}{2}\right)\tau - \ln\left(\frac{K}{S_t}\right)}{\sigma\sqrt{\tau}}\right) \\
&= S_t\Phi\left(\frac{\left(r + \frac{\sigma^2}{2}\right)\tau - \ln\left(\frac{K}{S_t}\right)}{\sigma\sqrt{\tau}}\right) - \exp(-r\tau)K\Phi\left(\frac{\left(r - \frac{\sigma^2}{2}\right)\tau - \ln\left(\frac{K}{S_t}\right)}{\sigma\sqrt{\tau}}\right) \\
&= S_t\Phi\left(\frac{\left(r + \frac{\sigma^2}{2}\right)\tau + \ln\left(\frac{S_t}{K}\right)}{\sigma\sqrt{\tau}}\right) - \exp(-r\tau)K\Phi\left(\frac{\left(r - \frac{\sigma^2}{2}\right)\tau + \ln\left(\frac{S_t}{K}\right)}{\sigma\sqrt{\tau}}\right)
\end{aligned}
$$

The Black-Scholes formula of a European call option is therefore:

$$
\begin{aligned}
c_t &= S_t\Phi(d_1) - \exp(-r\tau)K\Phi(d_2) \\
d_1 &= \frac{\ln\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}} \\
d_2 &= \frac{\ln\left(\frac{S_t}{K}\right) + \left(r - \frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}} \\
&= d_1 - \sigma\sqrt{\tau}
\end{aligned}
\tag{2.6}
$$

The process of deriving the value of European put option is similar to the process just used for the call option.

One of the most important by-products of the BS formula is the implied volatility. In financial markets, the market price of an option is known, and all of the parameters except σ are known. Hence we can find the value of σ that gives rise to an option value equal to the market price according to the BS formula. This value of σ is known as the implied volatility. However, in the financial market, skilled option traders do not rely solely on the implied volatility but will look behind the estimates to see whether they are higher or lower than the historical volatility and current volatility. Basically, implied volatility gives you an indication of the price of the option; while historical volatility gives you an indication of its value. We will discuss measures of historical volatility in detail in chapter 3.

2.3.2 The sensitivities (the "GREEKS") of the Black-Scholes Formula

The second important by-product of the BS formula are what are known as the "sensitivities". A sensitivity is the change in the option value resulting from a *ceteris paribus* change in one of the model parameters. The sensitivities are also known as the "Greeks", and are named: delta, gamma, theta, vega and rho. The Delta of an option is the rate of change of its price with respect to the price of the underlying asset. Gamma is the second derivative of the Option value with respect to the underlying price. Theta represents the rate of change of the value of the option with respect to the time to expiry. Vega represents the rate of change the price of option with respect to the volatility. Rho is the rate of change of option value with respect to the risk-free interest rate.

The following table shows us the Greeks of Black-Scholes Call Formula.

Name (symbol)	Formula	Sign
Delta (Δ)	$\frac{\partial C_t}{\partial S_t}\Phi(d_1)$	+
Gamma (Γ)	$\frac{\partial C_t^2}{\partial S_t^2}=\frac{\Phi(d_1)}{\sigma S_t\sqrt{\tau}}$	+
Theta (Θ)	$\frac{\partial C_t}{\partial \tau}=-\frac{S_t\sigma\phi(d_1)}{2\sqrt{\tau}}-\exp(-rt)Kr\Phi(d_2)$	-
Vega $(vega)$	$\frac{\partial C_t}{\partial \sigma}=S_t\sqrt{\tau\phi}(d_1)$	+
Rho (ρ)	$\frac{\partial C_t}{\partial r}=\exp(-rt)\tau K\Phi(d_2)$	+

Table 2.4 Greeks of Black-Scholes Call Formula

The main advantage of the Black-Scholes formula is that you can calculate a very large number of option prices in a very short time compared with the calculation of binomial model. The major limitation of Black-Scholes formula is that it cannot calculate the American style options. However, researchers have made adjustments to the Black-Scholes formula to enable it to approximate American option prices (Fischer Black Pseudo-American method) but these only work well with certain limits.

2.4 Comparison of Black-Scholes Formula and Binomial Tree Model

In this section, we are going to compare the two models. First of all, we will look at the link between the two models; secondly, we will discuss differences of the two models according to the assumptions of models; finally, we will compare the valuations of options computed using the two models.

2.4.1 Link between Black-Scholes Formula and Binomial Tree Model

The link between Black-Scholes Formula and Binomial Tree Model is in the volatility. In the BS formula, volatility is defined as σ, and it is constant annual volatility. For the binomial tree model, volatility is viewed as the movements of option prices (the option price goes up by a certain proportion (u) with a certain probability (p) or goes down by a certain proportion (d) with a certain probability (1-p)(see section 2.3).

As we explain in section 2.4, for the BS model, we assume that

$$\frac{S_T}{S_t} \sim \log normal\left(r\tau, \sigma^2\tau\right)$$

It indicates that standard deviation of return on stock price in a short time period of length of $\tau$is $\sqrt{\tau}$. So the variance of return is $\sigma^2\tau$.

On the tree figure2.3.1, the variance of stock return is:

$$pu^2 + (1-p)d^2 - \left[pu + (1-p)d\right]^2$$

The tree's parameters should match the volatility of stock price, hence we have:

$$pu^2 + (1-p)d^2 - \left[pu + (1-p)d\right]^2 = \sigma^2\tau \qquad (2.7)$$

Substituting equation (2.1) into (2.7), we have:

$$e^{r\tau}\left(u+d\right) - ud - e^{2r\tau} = \sigma^2\tau$$
$$e^{r\tau}\left(u+d\right) - ud - e^{2r\tau} = \sigma^2\tau$$

Using Taylor's series expansion, we have:

$$u = e^{\sigma\sqrt{\tau}}; d = e^{-\sigma\sqrt{\tau}} \qquad (2.8)$$

Equation (2.8) shows us the relationship between up and down parameters with the volatility, and it is also an important link between the returns of binomial tree model and the returns of Black-Scholes formula.

2.4.2 Differences between Black-Scholes Formula and Binomial Tree Model

Analytically, the Black-Scholes model is a continuous-time model, and the binomial tree model can be viewed as its discrete-time version. This analytical difference, however, does not affect the crucial idea that underlies the derivation of each model. The actual difference of the two models comes from the assumption of the process of stock return.

In Section 2.3.1, we listed the 7 assumptions underlying the Black-Scholes model. The assumptions underlying the binomial tree model are the same, except for Assumption 1 (stock returns are normally distributed with known mean and variance). For the binomial tree model, this assumption is replaced with: stock price follows a simple random walk (meaning that return is either u-1 with probability p, or d-1 with probability 1-p).

A (conventional) random walk is a sequence for which the change from one period to the next period is a white noise process:

$$X_t = X_{t-1} + u_t$$
$$u_t \sim N\left(0, \sigma^2\right)$$

Often, random walk includes a drift term ϒ, leading to the “random walk with drift”:

$$x_t = \gamma + x_{t-1} + u_t$$

It is conventional to assume that the natural logarithm of price follows a random walk (or similar process):

$$\ln X_t = \ln X_{t-1} + u_t$$
$$or \Delta \ln X_t = \gamma + u_t$$

ΔlnX_t represents (approximately) the proportional change in X between t-1 and t. Hence the term $\Upsilon + u_t$ represents the proportional change between t-1 and t. The parameter Υ represents the expected proportional change. The standard deviation of u_t is the standard deviation of daily proportional change, and this is the standard measure of the daily volatility of a price variable.

Consider daily data on the S&P 500 index, from 26 May 1998 to 23 May 2013. The time path of the S&P 500 Index is presented in Figure 2.11:

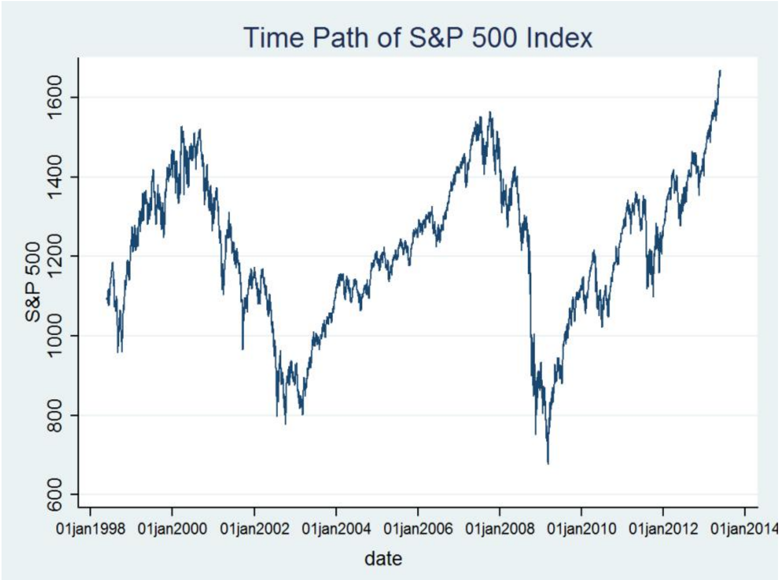
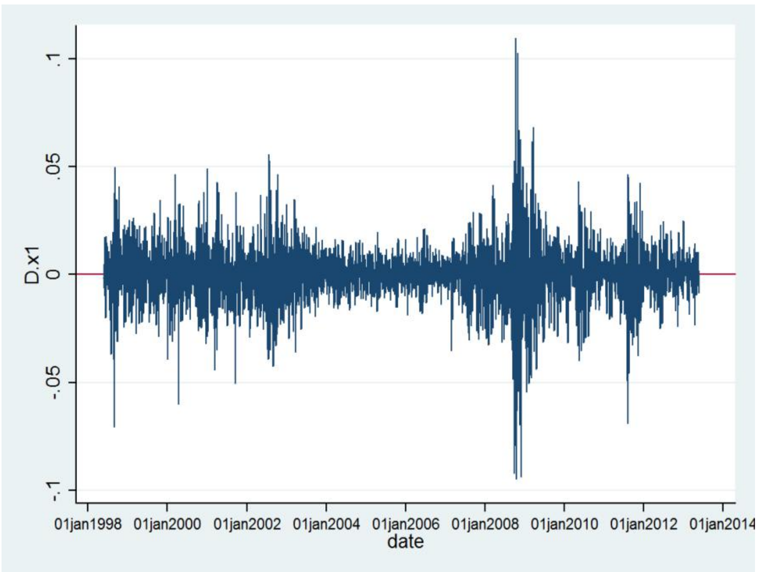


Figure 2.11 Time Path of S&P 500 Index

To find the daily return, we require the following sequence of STATA commands:

```
gen x1=ln( sp500)
tsset t
line d.x1 date, yline(0)
summ d.x1
```

The result is:



Variable	Obs	Mean	Std. Dev.	Min	Max
x1					
D1.	3912	.0001059	.0130693	-.0946975	.1095791

Figure 2.12 Time-Series of daily volatility

We note that the daily volatility of S&P 500 is 0.01307.

The plot of the daily returns seen in Figure 2.12 strongly suggest that the volatility is non-constant, and it is clearly inappropriate to measure volatility over 15 year period by a single number. This leads us to consider varying volatility models (ARCH, GARCH, TGARCH) which are discussed in chapters 4 and 5.

However, sometimes a single volatility measure is required. If this is the case, it is conventional to report the annualised volatility. This is obtained by multiplying the daily volatility by $\sqrt{252} = 15.87$ (252 being the number of trading days in the year).

When we apply this rule, we find that the annualised volatility of S&P 500 is 0.2074.

Let us attempt to simulate a series that has similar properties to the price index examined above. The properties we established were (1) that the mean of *d.logX* was around 0.0001-this is required drift term ϒ; (2) that the (daily) volatility was around 0.013. This is the required standard deviation of the white noise process in the random walk.

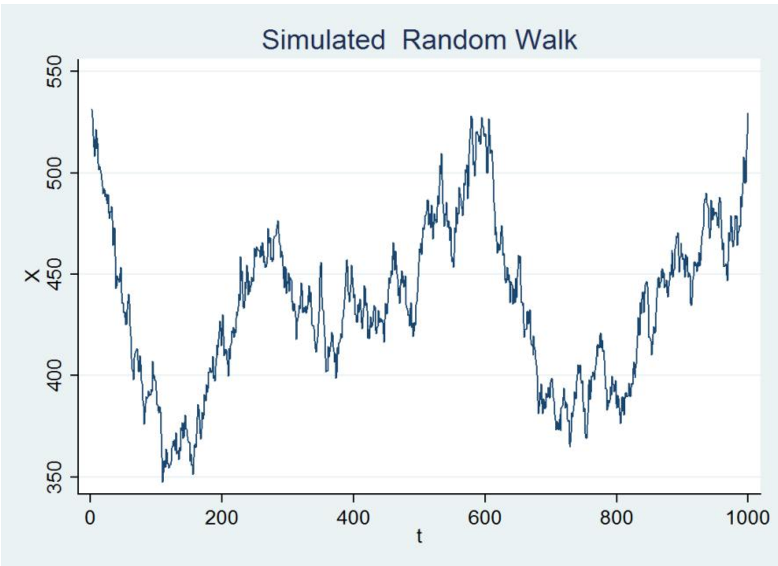


Figure 2.13 Simulated Random Walk

Recall from section 2.4, the stock price follows a lognormal distribution,

$$\ln S_T \sim N\left(\ln S_t + \left(\mu - \frac{\sigma^2}{2}\right)\tau, \sigma^2\tau\right)$$

We will attempt to simulate a “Black-Scholes” series that has similar properties to random walk examined above, the mean of *lnS*$_T$is $\ln S_t + \left(\mu - \frac{\sigma^2}{2}\right)\tau$ and the standard deviation is $\sqrt{\tau}$.

We do the simulation in STATA, and the result is shown in Figure 2.14.

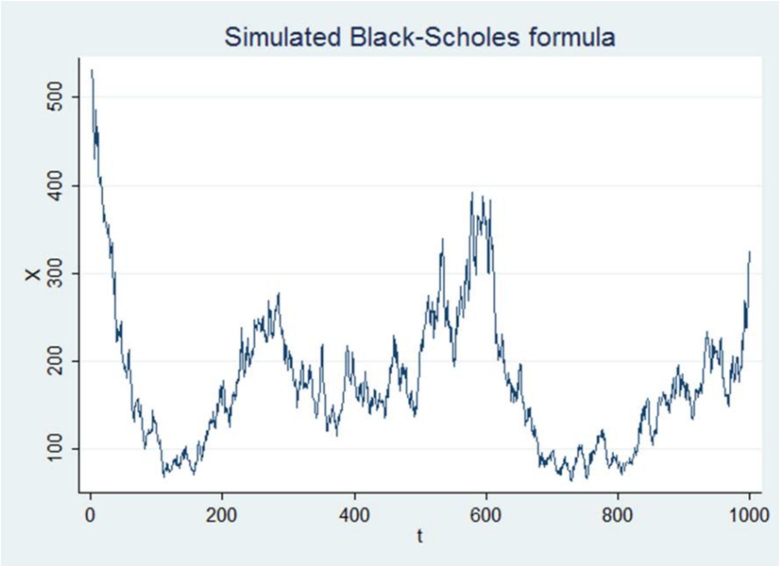


Figure 2.14 Simulated “Black-Scholes” series

We will also simulate a Binomial Tree series (i.e. simple random walk) based on the same parameters as used above. The result is shown in Figure 2.15.

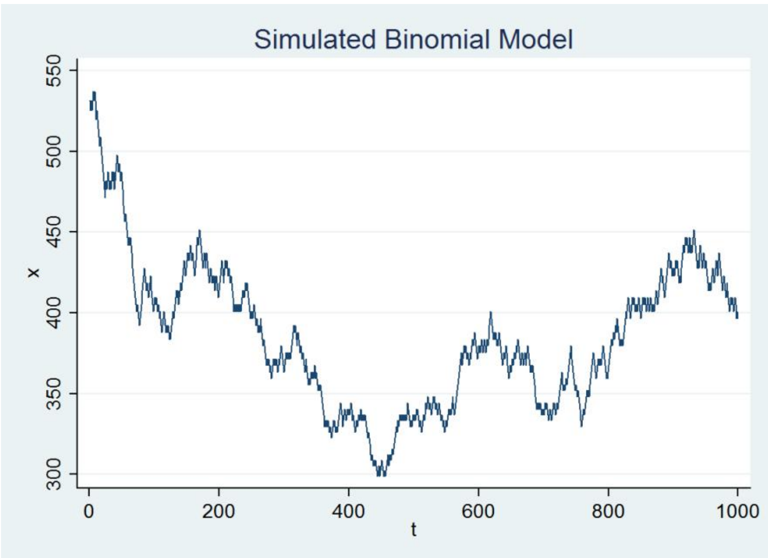


Figure 2.15 Simulated Binomial Model

From figures 2.14 and 2.15, we can see that the simulated stock return process of each model are different. However, according to the central limit theorem, when the number of Bernoulli trials increases, the binomial distribution approaches the normal distribution. As a result, for European options, the binomial model converges on the Black-Scholes formula as the number of steps of the tree increases.

In fact, the Black-Scholes formula for European options is really a special case of the binomial model where the number of binomial steps is infinite. In other words, the binomial model provides a discrete approximation to the continuous process underlying the Black-Scholes formula.

2.5 Comparison of calculated valuations from Black-Scholesand Binomial Tree Models

In this section, we will use a specific example to compare the value of an option obtained using the two models. We use the same example as we used in section 2.3.

S_t	K	r_f	τ	σ
1138.41	1075	2.13%	44 days	16.05%

The value of European Options of Black-Scholes formula and Binomial Model is given by the following table:

	Call Option	**Put Option**
Binomial Model with 2 steps	\$71.05	\$5.6
Binomial Model with 10 steps	\$70.5	\$4.8
Binomial Model with 50 steps	\$70.3	\$4.6
BinomialModel with100 step	\$70.2	\$4.5
Black-Scholes Formula	\$70.2	\$4.5

Table 2.5 Value of European Options of Black-Scholes formula and Binomial Model

From table 2.5, we see that Black-Scholes formula (bottom row) gives a slightly different value from the binomial model. However, as expected, when we add more steps to the binomial model, the valuation converges to the Black-Scholes valuation.

2.6 Other Methods for Valuing Options

We have introduced two very popular models for option valuation. However, there are many other option valuation models in the literature. Methods for valuing options can be divided into two groups: analytical and numerical.

The difficulty in pricing American options using analytical methods (such as the Black-Scholes equation) is the free boundary value problem. Han and Wu (2003) found the accurate boundary conditions for American options. A different analytical method was introduced by MacMillan (1986) and Barone-Adesi and Whaley (1987). They use quadratic approximation techniques combined with the finite difference method. The advantage of the method is fast speed of calculation; however, the results are of dubious accuracy.

A third analytical method is the interpolation method introduced by Johnson (1983) and Broadie and Detemple (1996). This method is faster than MacMillan and Barone-Adesi and Whaley's method. The disadvantage is that this method requires regression, which is time consuming. A fourth analytical method is the partial differential equation (PDE) method. Here, more terms are added in approximation compared with the above methods, increasing the accuracy of the method. Geske and Johnson (1984) use infinite series of multivariate cumulative normal functions to approximate American options (Ju, 1998). The computational demands of such methods are high, requiring the computation of multidimensional integrals. Bunch and Johnson (1992) introduced a modified two-point Geske-Johnson approach to avoid the multidimensional integrals. Thanks to their efforts, Carr (1997), Myneni (1992), Jacka (1991), and Kim (1990) finally find the formulas contain early exercise boundary. But the method is still time consuming, since it is necessary to calculate many early exercise points. In order to avoid this problem, Huang, Subramanyam, and Yu (1996) find a four-point extrapolation scheme (Ju, 1998). Only six early exercise points are required. Ju (1998) proposed another approximation based on the integral representation method in order to improve accuracy. Ju (1998) approximated the early exercise boundary as a mutipiece exponential function; the resulting integral can be evaluated in closed form.

The stochastic volatility model was first introduced by Heston (1993). He derived a closed-form solution for the price of a European call option on an asset with stochastic volatility instead of the annual constant volatility of Black-Scholes formula.

Numerical methods are more important than analytical methods for the pricing of American options. There are three basic numerical methods: binomial trees, Monte Carlo Simulation, and Finite difference methods.

The original binomial tree model of Cox, Ross, and Rubinstein (1979), was improved by Breen (1991), who developed the Accelerated Binomial Tree. Subsequently, Chang, Chung and Stapleton (2001) extended Breen’s model. They use a repeated Richardson extrapolation. The result was that they had a higher degree of accuracy (Global Derivatives, 2010).

Finite difference methods were first applied to option pricing by Schwartz (1977). Generally, finite difference methods are used to price options by approximating the differential equation that describes how an option price evolves over time by a set of (discrete-time) difference equations.The finite difference approach is applied to the backward partial differential equation (this equation was based on the Black-Scholes model). The existence of a formal expression for the solution of this partial differential equation allows evaluating the value of different kinds of options. The basic idea of the model is: replace the continuous derivatives in the PDF (Partial Differential Equation) by dividend differences and solve the corresponding matrix system at each time level.

Monte Carlo simulation methods introduced by Glasserman (2003) can be viewed as the most realistic models for valuing options other than the most simple options. Traditional valuation techniques become intractable as soon as the option becomes complex. Rather than attempting to solve the differential equations that describe the option's value in relation to the underlying price, a Monte Carlo model uses simulation to generate random price paths of the underlying, each of which results in a payoff for the option. The average of these payoffs can be discounted to yield an expectation value for the option. The advantage of Monte Carlo method is that it can be used to value complex options including American Options. We will discuss the Monte Carlo method in detail in chapters 4 and 5.

2.7 Summary

In this Chapter, first of all, we reviewed two basic methods for valuing European Call options: Binomial Tree model and Black-Scholes-Merton formula. We firstly discussed the binomial tree on the basis of the assumption of a simple random walk. After that, we extended the binomial tree model using the TGARCH assumption. We found that: ITM calls tend to have higher implied volatilities than OTM calls, and OTM puts tend to be higher implied volatilities than ITM puts, which is equivalent to saying that they are over-priced under Black-Scholes assumptions. This is the well-known “volatility smirk”.

Then, we compared the two basic models. The difference between the models is seen in the assumptions underlying the models. The Black-Scholes formula assumes that stock returns are normally distributed with known mean and variance; while the Binomial Model assumes that the stock price follows a simple random walk. Despite this, the Black-Scholes and Binomial Tree modelsare closely connected. The Black-Scholes formula for European options is a special case of the binomial model where the number of binomial steps is infinite.

Based on the two models, we found that there are five factors affecting the value of European Call Options: the stock price, strike price, time to expiry, risk-free interest rate and the volatility. In the next chapter, we are going to investigate whether these five factors affect the value of the European Options in the ways predicted by the models. We will use semi-parametric regression on market option data to identify the true determinants of European option prices.

Chapter 3 Regression analysis of Option Prices

3.1 Introduction

The Black-Scholes (BS; 1973) model and its variants are considered to be some of the most prominent achievements in financial theory in recent decades. However, empirical research has shown that the BS formula has significant shortcomings, in the form of biases (Bates, 2003; Bakshi et al., 1997; Andreou et al., 2008). The bias of the BS model stems from the assumptions such as geometric Brownian motion of stock price, constant variance of the underlying returns, constant interest rates, etc. For example, Rubinstein (1985) found that the implied volatility exhibited as U shape ("volatility smile"), which meant that implied volatility was a function of moneyness (Underlying/Strike) and time to expiry. Corrado and Su (1996) and Bakshi et al. (1997) showed that is stock returns are skewed and display excess kurtosis relative to the lognormal distribution. Other types of bias have been reported by many other researchers (see Bates, 1991; Bakshi et al., 2000; Andersen et al., 2002).

In view of the limitations of the BS formula, scholars and practitioners have accounted for deviations from its assumptions by using a regression-based deterministic volatility function (DVF) method to generate volatility values that depend on the option moneyness level and time to expiry (see Christoffersen and Jacobs, 2004; Christoffersen et al., 2009; Andreou et al., 2014). Such approaches often include the estimation of an "implied volatility function". However, to obtain implied volatilities from option prices requires use of the Black-Scholes framework. Since the objective of this thesis is to use methods completely outside the Black-Scholes framework, we will focus on historical volatility measures instead of implied volatility. There are many different ways to measure historical volatility. One important objective of this chapter will be to determine the best measure of historical volatility.

We will apply straightforward regression analysis to option prices, using various measures of historical volatility, and also the other standard determinants of option prices. We will focus on the regression coefficients. We consider this to be important because the regression coefficients correspond to the "greeks" in the BS model. By using regression analysis, we are obtaining estimates of the "greeks" without making any of the assumptions underlying the BS model. One previous article that performs regression analysis of option prices and focuses on the regression coefficients is Pandher (2007).

In accordance with the Black-Scholes formula, there are five factors expected to affect European option prices: current underlying price; strike price; risk-free interest rate; time to expiry; and volatility. Following previous research, we will define "moneyness" (for call options) as current price divided by strike price, and for put options as strike price divided by current price. It is obvious that the effects of certain variables (particularly moneyness) are highly non-linear. For this reason, we will make use of polynomial regression models, including linear regression, quadratic regression and cubic regression. Quadratic and cubic models are sometimes useful in capturing non-linearities. However, a superior approach is additive models, introduced by Stone (1985) and made more accessible by Hastie and Tibshirani (1990). These models estimate an additive approximation of the multivariate regression function. The components in the additive model will be B-splines. This will be discussed in detail in section 3.3.2.

A key feature of the regression analysis in this chapter is that annual volatility of the underlying index, σ, is measured in a number of different ways, and the results appear to be highly sensitive to this choice. The principal approach is to measure volatility using the standard deviation of daily returns for a period previous to the date on which the option is traded. But the question is how many days to use in this calculation. We try many different measures and find which one has the strongest effect on option price. This one will be considered the "correct" measure of volatility. The coefficient on this volatility measure is an estimate of the "vega" sensitivity parameter.

Apart from moneyness and volatility, the regressions also include risk-free rate and time to expiry. The coefficients on these variables will be estimates of "rho" and "theta" respectively.

The models estimated by Pandher (2007) are similar to ours. One difference is that he uses the option's implied volatility as the measure of volatility for the regression. This seems strange because the implied volatility has been computed from the option price.

Having estimated the regression models and identified which measure of volatility is most useful, we need to assess the performance of the model. For this purpose we will consider both in-sample and out-of-sample predictive performance. One particularly important question is whether any of the regression models can predict option prices better than the benchmark model. The benchmark model used in this chapter is the Practitioner Black-Scholes (PBS) model, used by many researchers including Christoffersen and Jacobs(2004) and Andreou (2014). It is widely used as a benchmark because it provides an effective means of mitigating the BS volatility smile or smirk anomaly. In the section on prediction, we will use the Box-Cox transformation of the option price as the dependent variable in regressions in order to ensure that the predicted option price is always positive.

3.2 Data Description

The data in this chapter covers the period from January 2000 to Oct 2009. The data was described in detail with tables and graphs in Chapter 1 Section 1.3. We applied several filtering rules to construct the option data set; these filtering rules were also discussed in detail in Chapter 1 Section 1.3. After filtering, the data set consists of 165,648 call options and 215,617 put options.

3.3 Regression Models

The purpose of this section is to use regression analysis to investigate the true determinants of option prices. We aim to discover which regression model is best in terms of in-sample fit. In addition, we are particularly interested in finding which measure of volatility is best for explaining option prices.

The dependent variable in all regressions in this section is option price itself. Later in the chapter, we will consider predictive performance of models. There, we will use a transformation of the option price as the dependent variable so as to ensure that the predicted prices are always positive. In this section, we are not concerned with prediction. We are simply concerned with the determinants of option prices. One advantage from using option price itself as the dependent variable is that the coefficients are easy to interpret and moreover they provide (model-free) estimates of the well-known "Greeks" of Black-Scholes theory.

From the Black-Scholes formula and the Binomial Tree model (see Chapter 2), we already know that there are five factors affecting the option price: underlying stock price; strike price; risk-free interest rate; time to expiry; and volatility. From financial market data, we may extract the current underlying price, strike price, risk-free interest rate and time to expiry. The only factor we do not know is volatility. A key task is therefore to find the best way to measure volatility. We will focus on measures of historical volatility.

Historical (daily) volatility is usually defined as the standard deviation of historical stock price daily returns for a period up to the current date (Hull, 2011). Annual volatility is computed as daily volatility times the square root of the number of trading days in a year.

According to the definition of historical volatility, first of all, we need to find the stock daily return, which denoted byR_t

$$R_t = \ln\left(\frac{S_t}{S_{t-1}}\right)$$

Where S_t is the index on day t. Annual volatility is then given by:

$$\sigma = s.d.(R_t) * \sqrt{252}$$

Note that 252 is the number of trading days in a year.

The range of returns used (i.e. the number of days over which the standard deviation is computed) in the calculation clearly influences the result. We will treat this as an empirical question. We will try 12 different time ranges: 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110 and 120 days. For each model we will therefore estimate 12 different regressions, one with each of these measures of volatility. The best measure of volatility will be determined as the one that gives rise to the highest positive coefficient when used in the regression.

Apart from determining the best measure of volatility, the other important task is to find which regression model provides the best option pricing model. In this part of the chapter, we will simply be looking at in-sample goodness-of-fit (i.e. R^2 and adjusted R^2).

3.3.1 Polynomial Regression Models

The general regression model for call optioncan be written as:

$$c = f(m, \tau, r_f, \sigma) \tag{3.1}$$

The general regression model for put option can be written as:

$$p = f(m,\tau,r_{f,}\sigma) \tag{3.2}$$

In both (3.1) and (3.2), m is moneyness. Moneyness of the call option is defined as $m=S_t/K$ and moneyness of the put option is defined as $m= K/S_t$. A central issue is that c and p are both highly non-linear function of moneyness (m), and possibly of the other variables as well.

We will specify (3.1) and (3.2) using polynomial functions: linear, quadratic, and cubic. The most general of these is the cubic:

$$c = \beta_1 + \beta_2 * m\, \beta_3 * m^2 + \beta_4 * m^3 + \beta_5 * tau + \beta_6 * r_f + \beta_7 * \sigma + u \tag{3.3}$$

$$p = \beta_1 + \beta_2 * m + \beta_3 * m^2 + \beta_4 * m^3 + \beta_5 * tau + \beta_6 * r_f + \beta_7 * \sigma + u \tag{3.4}$$

Where:

c is the mid-point of ask and bid call option price.

p is the mid-point of ask and bid put option price.

In all regressions, the "robust" option is used in order to correct the standard errors for heteroscedasticity.

Model 1: Linear Regression Model

The linear model is obtained by imposing restrictions on equation (3.3) and equation (3.4). The linear model assumes that option price depends linearly on moneyness, which means that $\beta_3=\beta_4=0$ in equation (3.3) and (3.4).

Model 2: Quadratic Regression Model

The quadratic regression model is also obtained by imposing restrictions on equation (3.3) and (3.4). It assumes that option price depends quadratically on moneyness, but linearly on other determinants. The required restriction on (3.3) and (3.4) is $\beta_4=0$.

Model 3: Cubic Regression Model

The cubic regression model is given by (3.3) and (3.4) with no restrictions imposed.

Table 3.2 presents sets of regression results from all three models, applied to the complete set of put options (N=215,617). Firstly note that, as expected, the coefficients on m^2 and m^3 are both highly significant in the cubic model, indicating that the cubic model is the best of the three. This is confirmed by the adjusted R^2 which is highest under the cubic model, at 0.922. It is therefore appropriate to use the results of the cubic model for interpretation. The coefficients provide model-free estimates of the "Greeks" of Black-Scholes theory. We see that the model-free estimate of "theta" is 67.80, "rho" is 34.67, and "vega" is 55.27. The signs of these are all in agreements with theory (see section 2.4.3: the Greeks of Black-Scholes formula).

Note that the measure used for volatility is V60, which is 60-day volatility. This measure is chosen because it is found to give the highest coefficient of all volatility measures – see Table 3.3 in Section 3.3.2.3.

	Linear	Quadratic	Cubic
m	397.6***	-2622.9***	-2.653***
	[0.523]	[4.538]	[48.93]
m2		1563.1***	-1131.1***
		[2.343]	[50.15]
m3			915.1***
			[17.01]
Tau	87.83***	67.84***	67.80***
	[0.285]	[0.166]	[0.165]
rf	-23.38***	34.79***	34.67***
	[2.826]	[1.617]	[1.606]
V60	86.26***	55.33***	55.27***
	[0.417]	[0.243]	[0.241]
_cons	-381.0***	1075.0***	233.3***
	[0.527]	[2.203]	[15.80]
N	215617	215617	215617
R^2	0.758	0.921	0.922
adj. R^2	0.758	0.921	0.922

Standard errors in brackets

$^{*}p< 0.05$, $^{**}p< 0.01$, $^{***}p< 0.001$

Table 3.1 Polynomial Regression Results for complete set of put options. Dependent variable: put option price.

3.3.2 Additive model: B-spline regression model

As noted above, assuming a cubic effect of moneyness appears to give a reasonable fit to the data, with an adjusted R^2 of 0.922. However, it may be that more flexibility is required. One approach to providing such flexibility is the use of additive models. Stone (1985) proposed additive models, and Hastie and Tibshirani (1990) made them more accessible. These models estimate an additive approximation of the multivariate regression function.

For the case of a dependent variable y_i and a total of m available predictors, the model takes the following form:

$$\left(y_i \middle| x_{1i}, \cdots, x_{mi}\right) = \beta_0 + \sum_{j=1}^{m} f_j\left(x_{ji}\right) \qquad (3.5)$$

In general the functions f_j are piecewise polynomials (or "splines") although they do not have to be so. Some predictors are better modelled linearly (so $f_j\left(x_{ji}\right) = \beta_j x_{ji}$ in the context of (3.5)).

3.3.2.1 The B-spline

Spline regression is superior to polynomial regression because it fits a polynomial (usually cubic) regression to each segment of the data, and joins these curves together in a way that results in a single smooth curve. Hence it is more flexible. Splines form a useful compromise between the global fit of polynomial regression, and the local fit of kernel smoothers.The "pieces" of the piecewise polynomials are separated by a sequence of K "knots", $\xi_1, \ldots, \xi_K$, and are forced to join smoothly at these knots.

Cubic splines are usually chosen, and the smoothness requirement is that the piecewise cubic functions are continuous and have continuous first and second derivatives at the knots[5].

The more knots are used, the more flexible the smoother. However more knots also means more parameters to estimate, and therefore fewer degrees of freedom. Clearly the choice of the number of knots must depend on the sample size: the larger the sample, the more knots can be used. Another choice that needs to be made is the positioning of the knots. An obvious strategy is to spread the knots uniformly over the range of the predictor variable, giving rise to what is known as a "cardinal spline". A more adaptive approach is one that places knots at appropriate quantiles of the predictor variable; for example, three knots, one at each of the three quartiles of the predictor. A more adventurous scheme would be one that places a higher concentration of knots in parts of the range of the predictor in which the nonlinearities are seen to be most marked. Such a scheme must of course be conducted by trial and error, since a smoother must be seen in order for nonlinearities to be identified.

The necessity of using ad hoc procedures for choosing the number and position of the knots is often seen as one of the (few) shortcomings of the spline approach.

---

[5]According to Hastie and Tibshirani (1990, p.22) "our eyes are skilled at picking up 2^{nd} and lower order discontinuities, but not higher".

The most popular approach for obtaining a piecewise cubic smoother with the required properties is the B-spline approach (de Boor, 1978). This amounts to a linear regression of the dependent variable on a set of *basis functions*, with no intercept. If there are K knots, there are *K+4* basis functions in total, although for practical reasons, only *K+2* of them are used in the regression. The positions of *K* interior knots are determined as the *(K+1)*th quantiles of the variable being measured on the horizontal axis, *x*. Let these be $q_1, q_2, \ldots, q_k$. Let q_0 is the minimum value of *x*, and q_{k+1} is the maximum value of *x*. In order to construct the basis functions, additional knots are placed at q_0 and q_{k+1}.

The set of eight fourth-order basis functions for *K=4* with equally spaced x-values on [0, 1] is shown in figure 3.2. Basis functions can be computed directly using the "bspline" command in STATA. The command of generating basis functions in STATA is:

```
bspline, xvar(x) power(3) knots(.00 .20 .40 .60 .80 0.10) gen(bs)
```

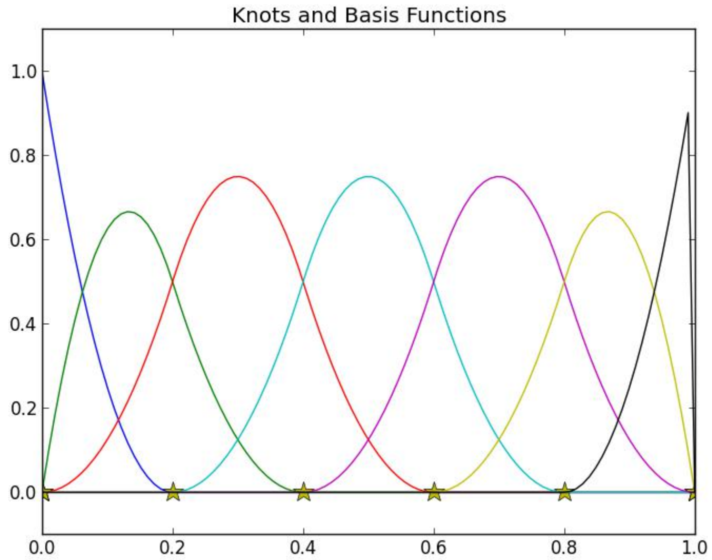


Figure 3.1 Basis Functions with four knots

The `gen(bs)` option in the command causes the basis functions to be stored as bs1-bs7.

If the basis functions to be used in the B-spline regression are $B_1(.), \ldots, B_{k+2}(.)$, then the piecewise cubic functions $f_j(.)$ appearing in (3.5) may be expressed as:

$$f_j\left(x_{ji}\right)=\sum_{k=1}^{K+2}\gamma_{jk}B_k\left(x_{ji}\right) j=1,\cdots,m \tag{3.6}$$

It is important that (3.6) does not contain an intercept. This is necessary for the model intercept (β_0 in (3.5)) to be identified.

Note also that (3.6) would lead to a fully general Additive Model in the sense that all *m* of the predictors are being assumed to have flexible effects. As noted in the discussion following (3.5), there are strong reasons for *not* modelling the effect of every predictor in accordance with (3.6).

3.3.2.2 B-spline of moneyness

For illustration, let us consider the basis functions applying to the variable moneyness.

First of all, we have to find the knots. We will use the quantiles of moneyness (from the put sample) to determine the knots. The quantiles are obtained as follows:

Smallest	*1%*	*5%*	*10%*	*25%*	*50%*	*75%*	*90%*	*99%*	*Largest*
0.75	0.76	0.81	0.84	0.91	0.97	1.01	1.06	1.19	1.25

We will use five knots. The first knot is the minimum value of moneyness, which is 0.75. The second knot is the first quartile of moneyness, which is 0.91. The third knot is the median of moneyness, which is 0.97. The fourth knot is third quartile of moneyness, which is 1.01. The fifth knot is maximum of moneyness, which is 1.25.

Figure 3.3 shows the basis functions of moneyness with these 5 knots.

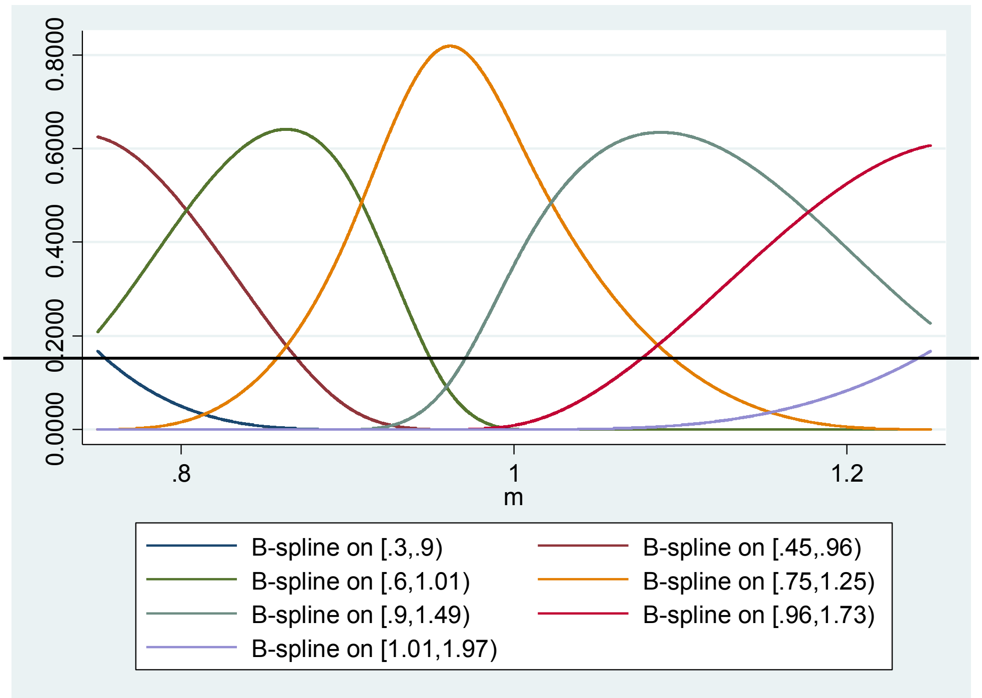


Figure 3.2 Basis functions for moneyness

These are the functions of moneyness used as explanatory variables in the B-Spline regression.

In a similar way, we also generate the basis functions of risk-free interest rate and time to expiry.

3.3.2.3 B-spline regression model with volatility

The B-spline model is estimated using a linear regression of the dependent variable on a set of basis functions of the independent variables. B-splines are used for three variables: moneyness; risk-free rate; and time to expiry. For volatility, we use a single variable. However, since the objective is to discover which volatility measure is best, we perform 12 different regressions with a different volatility measure in each.

Table 3.2 presents the results from five of these twelve regressions. Firstly notice that using the B-spline method causes the adjusted R^2 to increase from 0.922 (for the cubic model, Table 3.2), to 0.938. This indicates that the B-spline model provides a better in-sample fit even allowing for the greater number of parameters in the model.

Note also that the adjusted R^2 is not very useful in choosing between the models of Table 3.2, because it appears to be the same for all models. For this reason, and also because we are focusing on measures of volatility, the criterion we use here for choosing between models is the magnitude of the coefficient on the volatility variable. The model that has the highest positive coefficient of the volatility variable is considered to be the preferred model. In Table 3.3, we see that V60 gives the highest positive coefficient, of 57.14. Hence we conclude that V60 (60-day volatility) is the best (historical) volatility measure for explaining the complete set of put option prices.

We apply this procedure separately by sample range and by time to expiry. The results are presented in Table 3.3. The best volatility measure in each case is indicated by a rectangle placed around the coefficient. We see some interesting patterns. As expected, the number of days used in the volatility measure seems to increase with the time to expiry: for short term options, the number of days varies between 30 and 60, while for long term options it is usually 70. However, we also see a tendency for the number of days to vary according to the type of market: even for short-term options, the number of days used in the volatility measure is higher in the crash market than in the calm market or the bear market. This suggests that option traders appear to take account of a longer range of historical data when the market is more volatile.

	V20	V30	V40	V50	V60
bsm1	-34.02***	-32.29**	-33.79***	-35.73***	-38.26***
	[8.893]	[8.899]	[8.916]	[8.943]	[8.978]
bsm2	-60.99***	-59.96***	-61.44***	-64.24***	-67.06***
	[9.035]	[9.042]	[9.059]	[9.086]	[9.122]
bsm3	-38.48***	-37.35***	-38.85***	-41.50***	-44.27***
	[8.856]	[8.862]	[8.879]	[8.906]	[8.941]
bsm4	-30.71***	-29.58***	-31.06***	-33.72***	-36.48***
	[8.940]	[8.947]	[8.964]	[8.990]	[9.026]
bsm5	29.21***	30.40***	28.94***	26.27***	23.53***
	[8.746]	[8.753]	[8.770]	[8.795]	[8.830]
bsm6	330.4***	331.6***	330.2***	327.5***	324.6***
	[10.01]	[10.02]	[10.04]	[10.07]	[10.11]
bsrf2	-1.305***	-0.152***	1.068***	2.195***	3.068***
	[0.0915]	[0.0934]	[0.0958]	[0.0985]	[0.101]
bsrf3	0.0239***	1.175***	2.339***	3.432***	4.347***
	[0.104]	[0.107]	[0.109]	[0.112]	[0.116]
bsTau1	-81.51***	-81.52***	-81.34***	-82.76***	-81.04***
	[2.730]	[2.732]	[2.738]	[2.746]	[2.756]
bsTau2	-45.96***	-46.04***	-46.09***	-45.88***	-46.05***
	[0.346]	[0.347]	[0.347]	[0.348]	[0.350]
bsTau3	-36.54***	-36.52***	-36.39***	-36.47***	-36.32***
	[0.235]	[0.235]	[0.236]	[0.236]	[0.237]
bsTau4	-27.94***	-27.89***	-27.85***	-27.84***	-27.81***
	[0.164]	[0.164]	[0.165]	[0.165]	[0.166]
bsTau5	-13.31***	-13.30***	-13.22***	-13.30***	-13.17***
	[0.302]	[0.302]	[0.302]	[0.303]	[0.304]
V20	47.79***				
	[0.196]				
V30		50.71***			
		[0.209]			
V40			53.17***		
			[0.221]		
V50				55.35***	
				[0.233]	
V60					57.14***
					[0.245]
_cons	67.04***	64.30***	64.17***	65.44***	66.99***
	[8.906]	[8.912]	[8.929]	[8.955]	[8.991]
N	215617	215617	215617	215617	215617
R^2	0.938	0.938	0.938	0.938	0.938
adj. R^2	0.938	0.938	0.938	0.938	0.938

Standard errors in brackets

Table 3.2 B-spline Regression Results with different measures of volatility. Dependent variable: put option price. All put options included in sample.

	Bear Market		Calm Market		Crash Market	
Short term options: < 60 days						
	Put	Call	Put	Call	Put	Call
V10	6.933	13.24	48.77	50.09	27.23	33.45
V20	10.46	15.70	54.12	58.09	28.53	35.35
V30	$\boxed{11.28}$	15.93	$\boxed{54.52}$	62.24	29.23	36.19
V40	11.13	17.42	53.56	63.37	30.06	37.17
V50	8.585	$\boxed{18.39}$	51.43	$\boxed{63.63}$	30.96	38.45
V60	6.375	17.63	49.48	62.57	$\boxed{31.25}$	$\boxed{39.22}$
V70	4.037	15.32	46.32	58.76	30.53	39.19
V80	1.513	10.91	41.79	54.78	29.22	38.18
V90	-1.262	6.963	36.82	51.31	27.39	36.32
V100	-5.099	3.348	33.38	48.09	24.99	34.01
V110	-8.592	1.174	30.94	45.51	22.88	31.51
V120	-11.35	-0.318	29.00	41.93	21.72	29.13
Medium term options: 60~180 days						
V10	9.883	22.44	75.68	82.79	37.32	48.55
V20	16.30	30.22	85.92	98.47	41.59	53.48
V30	19.83	35.79	90.20	107.6	44.34	56.17
V40	$\boxed{21.31}$	40.51	$\boxed{91.21}$	113.3	46.17	58.13
V50	20.57	44.01	90.93	117.1	47.14	59.43
V60	20.27	$\boxed{44.52}$	88.82	$\boxed{117.3}$	$\boxed{47.41}$	60.29
V70	19.45	42.68	83.63	110.5	47.05	$\boxed{60.61}$
V80	17.64	38.15	78.74	105.2	46.58	60.58
V90	15.12	34.26	74.52	102.8	45.75	59.75
V100	12.62	31.48	71.63	101.0	44.43	59.02
V110	10.18	30.35	69.88	99.53	42.76	57.66
V120	10.53	31.26	66.40**	93.86	41.77	56.05
Long term options: >180 days						
V10	10.63	8.076	100.4	100.6	41.86	58.97
V20	17.28	12.06	120.3	119.5	46.70	64.79
V30	23.90	15.88	$\boxed{123.7}$	124.4	50.55	70.08
V40	30.98	21.95	118.7	123.6	54.70	74.96
V50	35.51	27.18*	115.3	125.4	58.67	79.65
V60	37.80	28.66	115.0	$\boxed{128.0}$	61.78	84.56
V70	$\boxed{38.26}$	$\boxed{29.02}$	112.4	126.0	$\boxed{61.80}$	$\boxed{87.50}$
V80	36.35	26.93	108.4	121.8	59.50	87.47
V90	35.10	26.64	102.6	118.2	56.88	85.67
V100	32.35	26.78	98.84	114.2	52.60	83.54
V110	31.02	26.41	95.00	110.4	48.25	81.20
V120	30.81	27.66	92.31	106.6	46.79	80.47

Table 3.3 Volatility coefficients from B-spline Regression models estimated separately by time period and time to expiry. Each coefficient in the table is from a different regression. Rectangles placed around coefficients indicate the best-performing volatility measure.

3.3.2.4 B-spline regression model with volatility and skewness

In this Section, we will add historical skewness to the B-spline regression model. Skewness is a measure of the asymmetry of the distribution of a variable about its mean. It can be positive or negative or zero. Skewness of return is calculated by the third standardized moment of daily return, denoted by:

$$\gamma = \frac{n}{(n-1)(n-2)} \sum_{i=1}^{n} \left(\frac{r - \overline{r}}{s} \right)^3$$

Where $\overline{r}$ is mean of daily return and s is sample standard deviation of daily return.

Recall the assumptions of the Black-Scholes formula. It assumes that stock returns are normally distributed with known mean and variance, implying that the skewness of the return is zero. There is much evidence that returns are skewed. In fact, the primary reason skew is important is that the analysis based on the normal distributions incorrectly estimates expected returns and risk. In the option market, skewness is likely to have a positive effect on call option prices and a negative effect on put option prices.

We are going beyond the Black-Scholes assumptions by including skewness in the regression model, and testing whether skewness has a significant impact on the option price. If skewness has a significant effect on the option price, this provides direct evidence that one of the Black-Scholes assumptions is incorrect.

The results obtained from the B-spline regression model with (60-day) skewness included are shown in Table 3.4. Note that skewness has a strongly significant effect for both call and put options, and for put options, adjusted R^2 has increased from 0.938 to 0.963 as a result of including skewness. Note that in both cases skewness has a negative effect on option price. For put options, this is exactly what we expect. However, for call options we expected a positive effect.

The significant effect of historical skewness on option prices clearly suggests that more advanced models are required for modelling option prices. These methods will be discussed in chapters 4 and 5.

	Call(V60&SK60)	Put(V60&SK60)
bsm1	13.98***	23.39***
	[5.715]	[2.847]
bsm2	-28.70***	-0.576***
	[1.013]	[0.576]
bsm3	7.734***	20.74***
	[0.337]	[0.270]
bsm4	25.65***	29.28***
	[0.273]	[0.229]
bsm5	109.5***	87.85***
	[0.324]	[0.267]
bsm6	327.0**	398.2
	[0.682]	[0.501]
bsrf2	7.670	4.217
	[0.120]	[0.103]
bsrf3	20.68	4.598
	[0.136]	[0.115]
bsTau1	-95.49	-82.30
	[3.272]	[2.738]
bsTau2	-57.41	-46.03
	[0.422]	[0.347]
bsTau3	-45.51	-36.40
	[0.288]	[0.236]
bsTau4	-35.16	-27.86
	[0.201]	[0.165]
bsTau5	-16.73	-13.13
	[0.367]	[0.302]
V60	76.51	60.15
	[0.297]	[0.250]
Skew60	-3.554	-2.550
	[0.0593]	[0.0473]
N	165648	215617
R^2	0.955	0.963
adj. R^2	0.955	0.963

Table 3.4 B-spline Regression Results with 60-days range of volatility and 60-days range of skewness. Dependent variable: call option price and put option price. All call and put options included in samples.

3.4 Predictive Performance of Regression Models

In this section, we compare the predictive performance of various models, including the B-spline regression model developed earlier in the chapter. The prediction criterion we will use is the root mean square error (RMSE, see Christoffersen and Jacobs, 2004; Andreou et al., 2008&2014; Chang et al., 2012; Bates, 2000), and we will apply this to assessing predictive performance both in-sample and out-of-sample.

Importantly, we will compare the predictive performance of the B-spline model to that of the Practitioners Black Scholes (PBS) model, which will be treated as a benchmark model.

3.4.1 The Practitioner's Black-Scholes (PBS) model

One particularly important question addressed in this chapter is whether any of the regression models can predict option prices better than the benchmark model. In order to answer this question, we will compare the best of our regression models with the Practitioner Black-Scholes model (PBS). The PBS model is a popular benchmark option pricing model, which was introduced (as the "ad-hoc model") by Dumas et al. (1998) and considered in more detail Christoffersen and Jacobs (2004), Berkowitz (2010), and Andreou et al. (2014).

There are three steps in the implementation of the PBS model. First of all, implied volatility of each option in the sample is computed using (the inverse of) the Black-Scholes formula. Second, we will use simple ordinary least squares (OLS) to run a regression of implied volatilities on different polynomials in time-to-expiry and strike price. This regression model is sometimes referred to as the "DFW implied volatility (IV) equation", having been introduced by Dumans, Fleming, and Whaley (1998). The most general IV equation appearing in the literature is:

$$\sigma = \theta_0 + \theta_1 K + \theta_2 K^2 + \theta_3 T + \theta_4 T^2 + \theta_5 KT + \varepsilon_{IV} \quad (3.7)$$

where σ is implied volatility, K is the strike price, and T is time to expiry.

Finally, the fitted values for IV are obtained as:

$$\sigma(\widehat{\theta}) = \widehat{\theta}_0 + \widehat{\theta}_1 K + \widehat{\theta}_2 K^2 + \widehat{\theta}_3 T + \widehat{\theta}_4 T^2 + \widehat{\theta}_5 KT \qquad (3.8)$$

Where hats indicate OLS estimates. These fitted values are plugged back into the Black-Scholes formula to obtain the PBS prices.

3.4.2 The Smearing Method in Prediction

Consider the regression model:

$$Y = X^{'}\beta + u$$

Which is estimated by OLS, resulting in predictions $\widehat{Y}$ and residuals $\widehat{u}$. But assume that we are interested in predicting not *Y*, but some non-linear function $f(Y)$ of Y. We could just use $f(\widehat{Y})$ for this. However, this gives biased predictions.

Duan (1983) proposed the "smearing" formula which is as follows:

$$\widehat{f}_i = \frac{1}{n}\sum_{j=1}^{n} f\left(\widehat{Y}_i + \widehat{u}_j\right)$$

The important feature of the smearing formula is that each observation (*i*) is predicted using the mean of a quantity involving the OLS prediction of Y_i and all *n* of the residuals. It is classified as a non-parametric method, since it is valid whatever the distribution of *u* in the regression model.

We use the smearing method in two places. Firstly, when we use PBS, as described in the last sub-section, we obtain predicted values of IV using a linear regression. Then we plug these predictions into the Black-Scholes formula in order to obtain predictions of the option prices. However, smearing is required to make these predictions unbiased. Christofferson and Jacobs (2004, p.298) strongly suggested the use of this sort of procedure in PBS: "It is clear that simply plugging $\sigma(\theta_{IV})$ into the Black-scholes formula will yield a biased estimate of the observed call price. While OLS will ensure that E$[\varepsilon_{IV}] = 0$, the non-linearity of the dollar option price in volatility and thus in ε_{IV} implies that $E[C] \neq C^{BS}(\sigma(\theta_{IV}))$." However, to our knowledge, no correction has ever been made for this problem. We will apply the smearing method in order to correct this bias.

The STATA code required to apply smearing to PBS is as follows. The first three commands perform the IV regression and obtain the fitted values (ivhat) and the residuals (uhat). The main part of the code is a loop over the sample (using the forvalues command).

```
regress iv K K2 tau tau2 Ktau
predict ivhat, xb
predict uhat, resid

gen d1=.
gen d2=.
gen pbs_smearing=.

quietly{
local N=_N
forvalues j=1(1)`N' {
scalar ivhat_temp=ivhat in `j'
scalar S_temp=S in `j'
scalar K_temp=K in `j'
scalar r_temp=r in `j'
scalar tau_temp=tau in `j'
replace d1=(ln(S_temp/K_temp)+    \\\
       (r_temp+(ivhat_temp+uhat)^2/2)*tau_temp)/ \\\
       ((ivhat_temp+uhat)*sqrt(tau_temp))
replace d2=d1-(ivhat_temp+uhat)*sqrt(tau_temp)
egenw=mean(S_temp*normal(d1) \\\
       -exp(-r_temp*tau_temp)*K_temp*normal(d2))
quietly replace pbs_smearing=w if _n==`j'
drop w
}
}
```

The second application of smearing is when we use the Box-Cox transformation to transform the option price (see next sub-section). Estimation of the Box-Cox model gives predictions of the transformed option prices. To obtain predictions of the option prices themselves, we need to apply the smearing method when the transformation is reversed. Conveniently, when the “predict” option is used after the “boxcox” command in STATA, the smearing method is used as the default.

3.4.3 The Box-Cox Transformation

In earlier sections, we used the option price itself as the dependent variable in regressions. This was to enable easy interpretation of coefficients. In this section, because the emphasis has shifted to prediction, we need to find a way of ensuring that the predicted option prices are always positive. To do this, we need to use an appropriate transformation of option price as the dependent variable. We could use the logarithm of option price for this purpose. However, we have found that predictive performance is greatly improved by using instead the more general Box-Cox transformation.

The Box-Cox transformation is defined as:

$$Y^{(\lambda)} = \frac{Y^{\lambda} - 1}{\lambda}$$

Note that when λ=1, the variable is essentially not transformed. When λ=0, the Box-Cox transformation becomes the log-transformation. λ is a parameter to be estimated. It is usually estimated to be between 0 and 1. The Box-Cox transformation is therefore more flexible than the log-transform.

When the Box-Cox transformation is applied to the dependent variable in a regression, the resulting predictions of the dependent variable are always positive, as required.

Table 3.5 shows Box-Cox regression results for all put options and all call options. We see that the estimates of λ are 0.447 and 0.490, respectively. This indicates that the log-transform would be inappropriate because λ is significantly greater than zero. It also indicates that using an untransformed option price is also inappropriate because λ is significantly less than 1.

	Put	Call
Notrans		
bsm1	4.399	-13.00
bsm2	0.784	-0.905
bsm3	4.827	1.987
bsm4	6.623	7.968
bsm5	18.53	24.16
bsm6	30.11	29.34
bsrf2	0.546	1.116
bsrf3	0.163	2.611
bsTau1	-14.31	-17.34
bsTau2	-7.624	-9.890
bsTau3	-5.813	-8.103
bsTau4	-4.131	-6.052
bsTau5	-1.633	-2.510
V60	10.60	13.95
Skew60	-0.363	-0.504
λ	0.447***	0.490***
	(0.00122)	(0.00137)
σ	1.415	1.707
N	215617	165648

Standard errors in parentheses
$^{*}p< 0.05$, $^{**}p< 0.01$, $^{***}p< 0.001$

Table 3.5 Box-Cox regression results for all put options and all call options

3.4.4 Pricing Performance of the option models

In this section, we will summarize the pricing performances of the different (Box-Cox) regression models, in which their in-sample andout of sample, pricing accuracy are measured in terms of RMSE.

As Christofferson and Jacobs (2004) and Andreou et al. (2014) point out, the estimation and evaluation of a model should be based on the same error measure. They also suggest that, among different loss functions, Root Squared Error (RMSE) estimates may perform best (see Christoffersen and Jacobs, 2004;Andreou et al., 2008&2014 ; Chang et al., 2012; Bates, 2000).

The RMSE for a put option is given by:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\left(P_i - \widehat{P_i}\right)\right)^2}$$

Where P_i and $\widehat{P_i}$ are the data and model put option prices respectively, and n is the number of options used. We will do the same for the call options.

Table 3.6 reports the pricing performance of six different models: (Box-Cox) linear regression, (Box-Cox) quadratic regression, (Box-Cox) cubic regression, (Box-Cox) bspline regression, PBS, and PBS with smearing. The PBS models use an IV function of the form (3.7) in Section 3.4.1.

Both in-sample and out-of-sample predictions are considered. Panel A reports the in-sample results of all models. Panel B reports the out-of-sample results. For the in-sample predictions, we used the first 87.5% of the whole data set for both estimation and prediction. For the out-of-sample predictions, we used the first 87.5% for estimation, and then used the resulting estimates to predict the remaining 12.5% of the data. In the case of the bear market period, we therefore use the first 21 months of option prices to predict the last 3 months. In the calm period, we use the first 49 months of data to predict last 6 months. In the crash period, we again used the first 21 months to predict the last 3 months. Each panel reports the RMSE of each model.

	Bear Market		**Calm Market**		**Crash Market**	
Panel A: In-sample						
	Put	Call	Put	Call	Put	Call
Linear	1.0196	1.1076	1.1069	1.2083	1.1484	1.3953
Quadratic	0.9292	1.1059	1.0054	1.2450	1.0542	1.41088
Cubic	0.8438	1.0401	0.9605	1.1182	0.9321	1.2325
Bspline	0.5534	*0.6399	*0.5550	*0.7862	*0.6256	*0.9713
PBS	0.5876	0.9035	1.2870	0.8906	1.1102	1.4152
PBS_s	*0.5297	0.9328	1.1735	0.8621	1.0680	1.4824
Panel B: Out-of-sample						
Linear	0.8641	0.6788	0.8708	1.2748	1.2535	1.0416
Quadratic	0.6286	0.5983	0.7878	1.3164	1.0301	0.9982
Cubic	0.5993	0.5796	0.7447	1.0603	0.8963	0.8562
Bspline	*0.3085	*0.2983	*0.4617	*0.8358	*0.2583	*0.2427
PBS	1.0093	0.8129	2.0060	1.3513	1.6899	1.5061
PBS_S	1.0640	0.8209	1.9608	1.3137	1.7054	1.4613

Table 3.6 In-sample and out-of-sample pricing performance of (Box-Cox) Linear Regression, (Box-Cox) Quadratic Regression, (Box-Cox) Cubic Regression, (Box-Cox) B-spline Regression, PBS, and PBS with Smearing (PBS_S). Panel A reports all in-sample results and Panel B reports all out-of-sample results. The numbers in the table are the RMSE.

3.4.4.1 In-Sample Pricing Performance

In this section, we will compare the in-sample predictive performance of the various models considered above. We are referring to panel A of Table 3.6.

In all cases except one, the best in-sample fitting accuracy is obtained by the Box-Cox B-spline regression model. We can therefore conclude that, on an in-sample basis, the Box-Cox B-spline regression model is superior to all of the (Box-Cox) polynomial models, and also superior to the PBS model, with or without smearing.

In four out of six cases, PBS with smearing has a smaller RMSE than PBS. This suggests that the application of smearing improves in-sample predictive performance.

3.4.4.2 Out-of-sample pricing performance

In this section, we will compare the out-of-sample predictive performance of the various models considered above. We are referring to panel B of Table 3.6.

Clearly, the (Box-Cox) B-Spline model is destined to perform better than other (Box-Cox) regression models in term of in-sample prediction performance, as a consequence of the fact that the former contains additional parameters and is therefore more flexible. However, out-of-sample prediction only improves if the additional parameters represent a genuine improvement to the model; addition of unnecessary parameters would result in a worsening of out-of-sample forecasting performance. Such a problem is sometimes known as "over-fitting".

We see in panel B of Table 3.6 that, in every case without exception, the RMSE of the Box-Cox B-spline model is lower than that of all other models. Hence we conclude that the Box-Cox B-spline model is unambiguously superior to all other models. Moreover, there is no evidence of over-fitting.

In three out of six cases, PBS with smearing has a smaller RMSE than PBS. This suggests that the application of smearing does not have a major effect on out-of-sample predictive performance.

3.5 PBS with different Implied Volatility Functions

According to Christoffersen and Jacobs (2004), a sufficiently general DVF function is:

$$\sigma = \theta_0 + \theta_1 X + \theta_2 X^2 + \theta_3 T + \theta_4 T^2 + \theta_5 XT + \varepsilon_{IV}$$

Where X is the strike price, it is time to expiry.

Andreou at al. (2014) considered extensions to the model. They generated several alternative specifications of regression-based DVF models so as to determine which best characterizes the daily implied volatility functions of the S&P 500 index options. Basically, symmetric DVF specifications define the implied volatilities as quadratic polynomial functions of strike price and/or time to expiry and/or moneyness. Symmetric DVF models are widely used as benchmarks in the literature (for example, Andreou et al., 2010; Christoffersen et al., 2009; Linaras and Skiadopoulos, 2005). In order to give more flexibility to the DVF models, we will extend the benchmark IV functions by applying the B-spline method. We will use basis functions for strike price, time to expiry, and moneyness.

We will investigate whether the B-spline DVF model performs better than the benchmark DVF models.

Andreou et al. (2014) considered the following symmetric DVF models:

DVF1:

$$\sigma^S{}_X = \max(0.01, \alpha_0 + \alpha_1 X + \alpha_2 X^2 + \alpha_3 T + \alpha_4 XT + \alpha_5 T^2)$$

DVF2:

$$\sigma^S{}_{\ln X} = \max(0.01, \alpha_0 + \alpha_1 \, ln\, X + \alpha_2 (ln\, X)^2 + \alpha_3 T + \alpha_4 \, (ln\, X)\, T + \alpha_5 T^2)$$

DVF3:

$$\sigma^S{}_K = \max(0.01, \alpha_0 + \alpha_1 K + \alpha_2 K^2 + \alpha_3 T + \alpha_4 KT + \alpha_5 T^2)$$

Where X is the strike price and K is moneyness, which is (as usual) defined by the spot price S divided by the strike price X.

Basically, the symmetric DVF specifications define the implied volatilities as quadratic polynomial functions of different combinations of strike price, moneyness and time to expiry. Note that DVF2 uses the log of strike price instead of strike price. Andreou at al. (2014) find that DVF3 is the best model and DVF2 performs better than DVF1. DVF2 is inferior to DVF3 because the natural logarithm of moneyness doesn't offer a significant scaling benefit.

A novel feature of our analysis is that we apply the B-spline method to the DVF model. We generate the basis functions for strike price, time to expiry, and moneyness. The new symmetric DVF models will be:

DVF_bspline1:

$$\sigma^S{}_{bspline1} = \max(0.01, \alpha_0 + \alpha_1 f(X) + \alpha_3 f(T) + \alpha_4 XT)$$

DVF_bspline2:

$$\sigma^S{}_{bspline2} = \max(0.01, \alpha_0 + \alpha_1 f(K) + \alpha_3 f(T) + \alpha_4 KT)$$

Where: f(X) is the basis functions of strike price, f (K) is the basis functions of moneyness and f (T) is the basis functions of time to expiry. Note that the difference between the two B-Spline functions is that DVF_bspline1 uses basis functions for strike price and time to expiry, while DVF_bspline2 uses basis functions for moneyness and time to expiry.

We first consider the performance of the five DVF equations. How close are the predictions of IV to observed IV? Table 3.7 reports the results of the five different DVF models from in-sample and out-of-sample tests. Again, for the in-sample predictions, we used the first 87.5% of the whole data set for both estimation and prediction. For the out-of-sample predictions, we used the first 87.5% for estimation, and then used the resulting estimates to predict the remaining 12.5% of the data.

	Bear Market		**Calm Market**		**Crash Market**	
Panel A: In-sample						
	Put	Call	Put	Call	Put	Call
DVF1	0.2079	0.1878	*0.1774	0.1616	0.2049	0.2500
DVF2	0.2045	0.1830	0.1775	*0.1588	0.2037	0.2498
DVF3	0.2119	0.1972	0.2395	0.2924	0.33492	0.3906
DVFbspline1	0.2033	0.1930	0.2081	0.2556	0.2999	0.3322
DVFbspline2	*0.1995	*0.1739	0.1781	0.1597	*0.2068	*0.2506
Panel B: Out-of-sample						
DVF1	0.1844	*0.1559	0.3454	0.3567	0.3722	0.5142
DVF2	0.1871	0.1597	0.3493	0.3786	0.3671	0.5113
DVF3	0.1997	0.2185	*0.2370	*0.2654	*0.2622	*0.3923
DVFbspline1	*0.1650	0.1864	0.3417	0.4289	0.3154	0.4382
DVFbspline2	0.1945	0.1774	0.3405	0.4038	0.4014	0.5967

Table 3.7 In-sample and out-of-sample pricing performance of five IV equations for both call and put options. Panel A reports the in-sample results of five different DVF models. Panel B reports the out of sample results. The numbers in the table are the RMSE (comparing predicted IV to observed IV); the red rectangle shows the lowest RMSE which gives us the best pricing performance.

We begin by discussing the DVF models' in-sample pricing performance. From table 3.8 panel A, in the bear market, and the crash market, the B-spline DVF (with moneyness and time to expiry) gives us the lowest RMSE. In the calm market, it seems that the basic DVF's give the lowest RMSE.

Next we turn to out-of-sample predictions. Here we see a different pattern. Based on table 3.8 panel B, the B-spline only does best in the bear market for call options. In most other cases, DFV3 performs the best. This result is consistent with previous research (Andreou (2014); Christoffersen et al., 2009)).

The finding that the B-spline model fits performs best in-sample, but a simpler DVF equation performs best out-of-sample, suggests that we have a case of over-fitting.

However, we need to remember that the purpose of estimating the IV equations is to obtain predicted IV's to plug into the Black-Scholes formula in order to predict option prices. It remains to be seen how well the B-Spline DVF's predict option prices. We consider this next.

Table 3.8 shows us the in-sample and out-of-sample results of the five PBS models resulting from the five DVF equations. Here we are comparing predicted option prices to actual option prices using RMSE. Note that the first row of each panel in Table 3.8 is the same as the "PBS" row in Table 3.7.

	Bear Market		**Calm Market**		**Crash Market**	
Panel A: In-sample						
	Put	Call	Put	Call	Put	Call
PBS_DVF1	0.5876	0.9035	1.2870	0.8906	1.1102	1.4152
PBS_DVF2	0.5067	0.6671	1.3238	0.9204	1.0629	1.3379
PBS_DVF3	1.1121	1.6208	0.9082	1.1334	2.0101	2.5119
PBS_DVFB1	*0.3445	0.6035	0.5698	0.6723	*0.4604	1.8164
PBS_DVFB2	0.4490	*0.4473	*0.5039	*0.4916	0.9294	*0.7374
Panel B: Out-of-sample						
PBS_DVF1	1.0093	0.8129	2.0060	1.3513	1.6899	1.5061
PBS_DVF2	0.9679	0.7609	2.0109	1.2129	1.5275	1.4936
PBS_DVF3	0.7295	0.9116	1.0153	1.1624	2.5003	2.3869
PBS_DVFB1	*0.2364	*0.3023	0.8510	*0.5334	*0.7756	2.0625
PBS_DVFB2	0.2806	0.3598	*0.5197	0.5895	1.0854	*1.4101

Table 3.8 In-sample and out-of-sample pricing performance of PBS using different DVF models of both call and put options. PBS_DVFB1 is the model using DVFbspline1; PBS_DVFB2 is the model using DVFbspline2. Panel A reports the in-sample results of five different DVF models. Panel B reports the out of sample results. The numbers in the table are the RMSE; the red rectangle shows the lowest RMSE which gives us the best pricing performance.

From Table 3.8 Panel A, we see that, in-sample, the best performing model for prediction of option prices is always one of the two B-spline models, PBS_DVFB2and PBS_DVFB1.

From Panel B, we see that out-of-sample, the best fitting model is again always one of the two B-spline models, PBS_DVFB2 and PBS_DVFB1.

It is interesting that using the B-spline approach to estimate the IV equation seems to result in the problem of over-fitting, but when the predictions of IV from these over-fitted equations are used to predict option prices, the out-of-sample predictions are superior to those from simpler models. This perhaps indicates that we should not be concerned with the problem of overfitting when estimating IV equations.

3.6 Summary

The overall objective of this chapter is to use regression analysis to investigate which model is the best model for predicting the European option price. First of all, we generate four regression models, including linear regression, quadratic regression, cubic regression and B-spline regression models to find the true determinants of option prices. In addition, we were particularly interested in finding which measure of volatility is best for explaining option prices.

Having estimated the regression models and identified which measure of volatility is most useful, we set out to assess the predictive performance of the models. Since the emphasis shifted to prediction, we needed to use an appropriate transformation of option price as the dependent variable, to ensure that predictions of price are always positive. We found that predictive performance is greatly improved by using the Box-Cox transformation. Therefore, we used box-cox regressions instead of linear regressions to compare the predictive performance of the models. We assessed in-sample and out-of-sample performance using the RMSE to compare models. We found that the best in-sample fitting accuracy is obtained by the Box-Cox B-spline regression model. Regarding out-of-sample performance, we found that the Box-Cox B-spline model is unambiguously superior to all other models. Moreover, there is no evidence of over-fitting (with regard the prediction of prices).

We found that the PBS with smearing model has a smaller RMSE than PBS in the majority of cases. This suggests that the application of smearing does have an effect, if not a major effect, on out-of-sample predictive performance.

Chapter 4 Analysing Financial Volatility

4.1 Introduction

In the history of humankind financial markets are to be known as one of the most exclusive and complicated systems. Volatility of asset returns has been one of the primary concerns in the risk analysis research over the past decades.

The research on variation and co-variation of financial time series has attracted different groups of audience, ranging from hedge fund managers to macroeconomists. Over the past decades, appealing properties of financial asset returns have been uncovered and have assisted to build appropriate models. At the heart of investment finance, there is the concept of the Efficient Market Hypothesis (EMH). No other concept in finance has been discussed and scrutinized as much as EMH. The history of EMH spans more than a century[6] and the building blocks of the past 50 years of research depend on whether EMH is rejected or not.

The Efficient Market Hypothesis (EMH) proposes that future prices cannot be predicted based on past information and one cannot achieve excessive returns using past information. Even if EMH holds, the vast majority of financial economists and risk managers are not interested in predicting the future price but they are rather interested in forecasting the future volatility, and have used for this purpose a variety of parametric econometric models such as generalized autoregressive conditional heteroscedasticity (GARCH) or stochastic regime switching models. Volatility measurement is of particular value to option traders as they are essentially trading in volatility.

As mentioned in previous chapters, five factors determine the value of a European option: underlying stock price, strike price, time to expiry, risk-free interest rate and volatility. It is straightforward to find the value of underlying stock price, strike price, time to expiry and risk-free interest rate. The only factor that is not known is the volatility.

The need for discovering and better understanding the dynamics of financial markets reaches its peak every time the world faces a financial crisis. The recent financial crises over the past 30 years, 1987's Black

---

[6]George Gibson in 1889 in his book "The Stock Markets of London, Paris and New York" explains the idea behind efficient markets: "When shares become publicly known in an open market, the value which they acquire may be regarded as the judgment of the best intelligence concerning them".

Monday, 1997's Asian Financial Crisis, early 2000's recession, and the very recent 2008's Global Financial Crisis, have brought further attention to the characteristics of financial time series. These events which are often known as "fat tail" events when do happen have catastrophic consequences which often last a long time and that is why they are sometimes given the "Black Swans" title[7].

It is discussed heavily in the literature ((Mandelbrot, 1971), (Bollerslev and Wright, 2000), (Cont, 2001)), that linear filtering statistical tool, such as autocorrelation analysis, are not capable of distinguishing between asset returns and white noise. Roughly speaking, the absence of autocorrelation in returns only gives some empirical evidence for random walk models of prices, but it does not mean that they are also independent. If the returns were to be independent, then any nonlinear functions (e.g. absolute returns, squared returns) of return should also show no signs of autocorrelation.

As it is accepted in the literature and also demonstrated in this chapter that the autocorrelation of absolute/squared returns exhibit significant positive autocorrelations. Among others Mandelbrot, (1963), explain that "large price changes are not isolated between periods of slow change" but large changes tend to be followed by large changes of either positive or negative sign. Also "small changes tend to be followed by small changes". That is why the autocorrelations of absolute returns are significantly positive. This is also in line with the findings of (Bollerslev and Wright, 2000) and (Cont, 2001) and shows that returns are not independently distributed.

What is useful here is that, the squared/absolute of future returns can be predicted (to some extent) based on past returns, even though the sign of price changes (returns) cannot be forecasted. This result is well known as clustered volatility and has been verified in many financial markets [to name a few: Cont et al. (1997), Bollerslev and Wright (2000), Cont (2001)].

These nonlinear functions of returns, which form different volatility models, show signs of predictability in contrast with return that does not show any signs of predictability. That is, there is hope in predicting volatility rather than the returns themselves. To what extent or how far into the future these non-linear transformations of returns are predictable is of great importance to risk managers.

---

[7]The term "Black Swan" is used by Nassim Taleb in capitalized form to refer an event with three attributes: rarity, extreme impact, and retrospective predictability. Please see (Taleb, 2008) and (Taleb et al., 2009) for more details.

There are a large number of models that can accommodate varying volatility. Engle (2001) provides a survey, and Bollerslev (2009) provides a useful glossary of "ARCH-acronyms". In recent years, the most interesting approaches for modelling the volatility are the "asymmetric" or "leverage" volatility models, in which good news and bad news have different predictability for future volatility (see Duan et al. (2006); Engle and NG (1993); Engle (2011)). There is large number of models proposed in the financial econometrics literature to represent this phenomenon, for example, Engle and NG (1993) introduced Asymmetric GARCH (AGARCH); Glosten, Jaganathan, and Runkle (1993) proposed GJR-GARCH; Zakoian (1994) introduced Threshold GARCH (TGARCH) and Nelson (1991) proposed the exponential GARCH (EGRCH). Researchers have previously compared the asymmetric conditional heteroscedastic models by comparing their predictive performance. For example, Loudon et al. (2000) conclude that several alternative asymmetric GARCH models are similar when representing leverage effect; Balaban (2004) and Alberg et al. (2008) show that EGARCH performs better among other models.

The structure of this chapter is organised as follows. After basic definitions, the data used in this section is described. Next, stylised facts of return are uncovered, a series of volatility models (from the GARCH family) are listed, and applied to daily data on the S&P500 Index.

4.2 Asset Returns Definitions

Because returns have much better statistical properties than price levels and they are often assumed stationary, risk modelling focuses on describing the dynamics of returns rather than prices. This thesis uses daily returns on the S&P500 from 1 January 2000, through 31 October 2009, to illustrate the features of daily return and aims to estimate the best-suited volatility model. The data has been downloaded from DataStream.

There are two definitions of the asset return. These are, the simple rate of return and the continuously compounded or log-return. It can be easily proved that the two returns are fairly equal.

The Simple Rate of Return(r_t)**:**

The daily simple rate of return of an asset is defined as:

$$r_t = \frac{S_t - S_{t-1}}{S_{t-1}} \qquad (4.1)$$

Where S_t is the daily closing price at time t and S_{t-1} is the daily closing price at time t-1.

The Continuously Compounded Rate of Return (R_t):

The continuously compounded rate of return on an asset is defined as:

$$R_t = \ln(S_t) - \ln(S_{t-1}) \qquad (4.2)$$

Where $\ln(x)$ is the natural logarithm of x , S_t is the daily closing price at time t and S_{t-1} is the daily closing price at time t-1.

Equation (4.2) can be re-written as:

$$R_t = \ln(S_t) - \ln(S_{t-1}) = \ln\left(\frac{S_t}{S_{t-1}}\right)$$

And equation (4.1) as:

$$r_t = \frac{S_t - S_{t-1}}{S_{t-1}} = \frac{S_t}{S_{t-1}} - 1$$

Replacing $\frac{S_t}{S_{t-1}}$ in equation (4.2) above:

$$R_t = \ln(S_t) - \ln(S_{t-1}) = \ln\left(\frac{S_t}{S_{t-1}}\right) = \ln(r_t + 1) \qquad (4.3)$$

The Taylor series expansion of an infinitely differentiable f(x) is $\sum_{n=0}^{\infty} \frac{f^n(a)}{n!}(x-a)^n$. Using the Taylor series expansion of $\ln(r_t + 1)$ equation (4.3) transforms to:

$$R_t = \ln(r_t + 1) = r_t - \frac{r_t^2}{2} + \ldots$$

If $|r_t| < 1$, the equation above can be an approximation using only the first term. The assumption that $|r_t| < 1$, is a plausible one as returns on financial assets are small numbers (the average of r_t is a number often very close to zero). If r_t is close enough to zero then:

$R_t \approx r_t$ (4.4)

Equation (4.4) confirms that the simple rate of return r_t and the log return R_t are approximately equal rates of return.

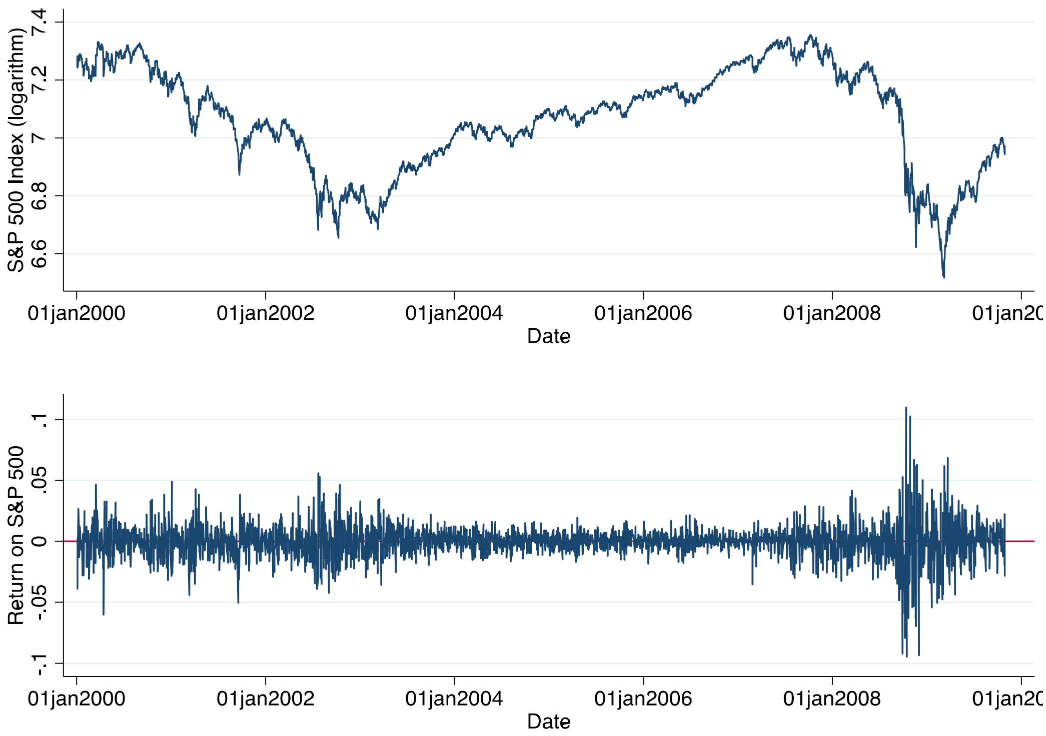


Figure 4.1 The return series and the closing prices against time from 1 January 2000, through 31 October 2009.

Figure (4.1) shows the return series and the closing prices against time from 1 January 2000, through 31 October 2009. We can see that the return series hovers around zero. The average of R_t is -0.00014, which is a small number, and so equation (4.4) is valid. This chapter will use the log return definition (R_t) unless otherwise stated.

And when using log returns (R_t) from equation (5.2), then tomorrow's closing price can be calculated from:

$$S_{t+1} = e^{R_{t+1}} S_t$$

4.3 Stylised Facts of Asset Returns

Covariance and Correlation

The sample covariance between two random variables can be estimated via:

$$cov(X,Y) = \frac{\sum_{t=1}^{T} (x_t - \overline{X})(\sum_{t=1}^{T} (y_t - Y)}{T}$$

The sample correlation between two random variables, x and y, is calculated as:

$$\rho_{x,y} = \frac{cov(X,Y)}{\sqrt{var(X)Var(Y)}}$$

Correlation measures the linear association between two different variables, while autocorrelation measures the linear association between the current value of a variable (time series) and the past value of the same variable. Autocorrelation is a crucial tool for detecting linear dynamics in time series analysis and it captures the linear relationship between today's value and the value τ days ago. Autocorrelation is widely used in economics and finance and it is a central tool for observing the dependence structure of the variable.

The autocorrelation for lag τ is defined as:

$$\rho_\tau = \frac{cov(R_t,R_{t-\tau})}{\sqrt{var(R_t)Var(R_{t-\tau})}} = \frac{cov(R_t,R_{t-\tau})}{var(R_t)} \qquad (4.5)$$

In the formula above it is assumed that $var(R_t) = Var(R_{t-\tau})$, which is satisfied as long as the variable is stationary. It is widely accepted that financial returns are stationary.

Assume we are using daily returns and we have observed today's return, then ρ_1 which is the autocorrelation between R_t and R_{t-1} is in fact calculating how closely today's return (R_t) and yesterday's return (R_{t-1}) are correlated. ρ_2 is the autocorrelation between today's return and the return the day before yesterday's return (return, two days ago). We can continue increasing the lags, for example ρ_{300} would be the autocorrelation between today's return and return 300 days ago. One can expect that as we increase the lags, the autocorrelations naturally decreases meaning that information (returns) far into the past are not very relevant to today's information (return).

Knowing what autocorrelations for several lags are can be very helpful for predicting what future autocorrelations can be.

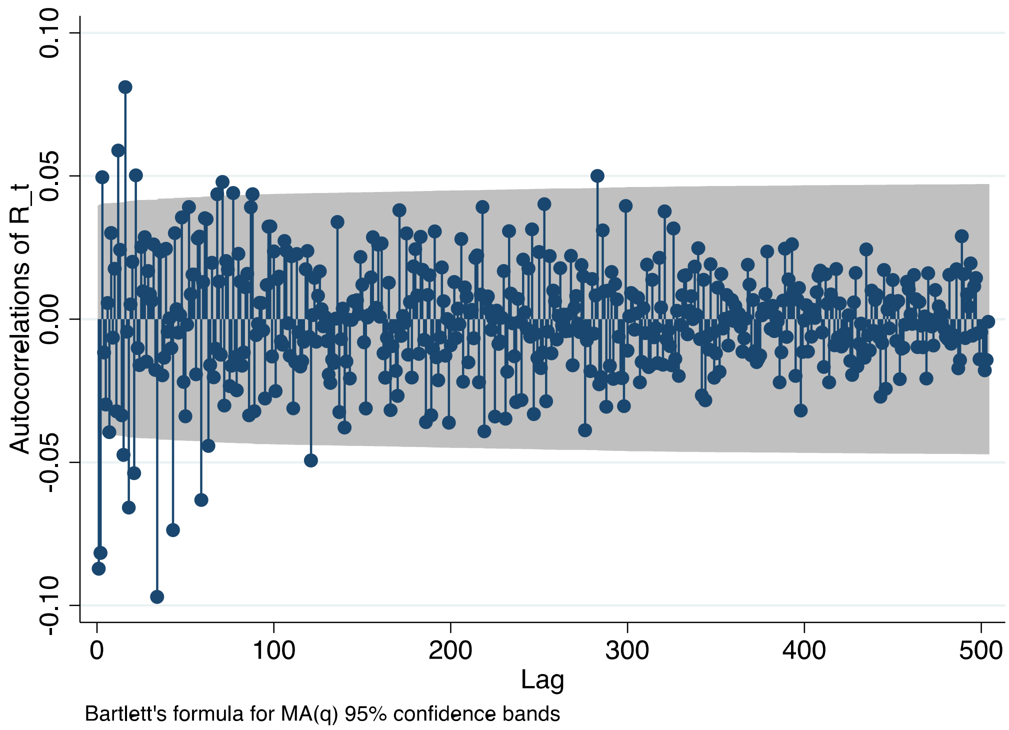


Figure 4.2 Autocorrelations of daily S&P500 returns for up to two years

In order to investigate the dependence structure in a variable, it is useful to see this dynamic through a graph, rather than by looking at autocorrelations, calculated from equation (4.5). What would be useful to look at is a plot of autocorrelations at different lags so that we can see what happens to autocorrelations as we increase the lag (travel into the past) and whether there are days that seem very relevant to today. The autocorrelation plot against τ, shows how important the value τ days ago is for today.

Among all the characteristics of daily returns, one of the most commonly known is that the daily returns have little autocorrelations. That is:

$$\rho_\tau = corr(R_t,R_{t-\tau}) \approx 0, \tau = 0,1,2,3,\ldots,n \qquad (4.6)$$

In other words, returns are almost impossible to predict from their own past. Figure (4.2) shows the correlation of daily S&P500 returns with returns lagged from 1 to 504 days (roughly two years). The grey area shows the insignificant area. The autocorrelation plot shows whether there is a significant dynamic linear relationship between today's return and past returns. If more autocorrelation coefficients are sticking out from the grey area, then it means that there are predictable patterns in the past values of the variable.

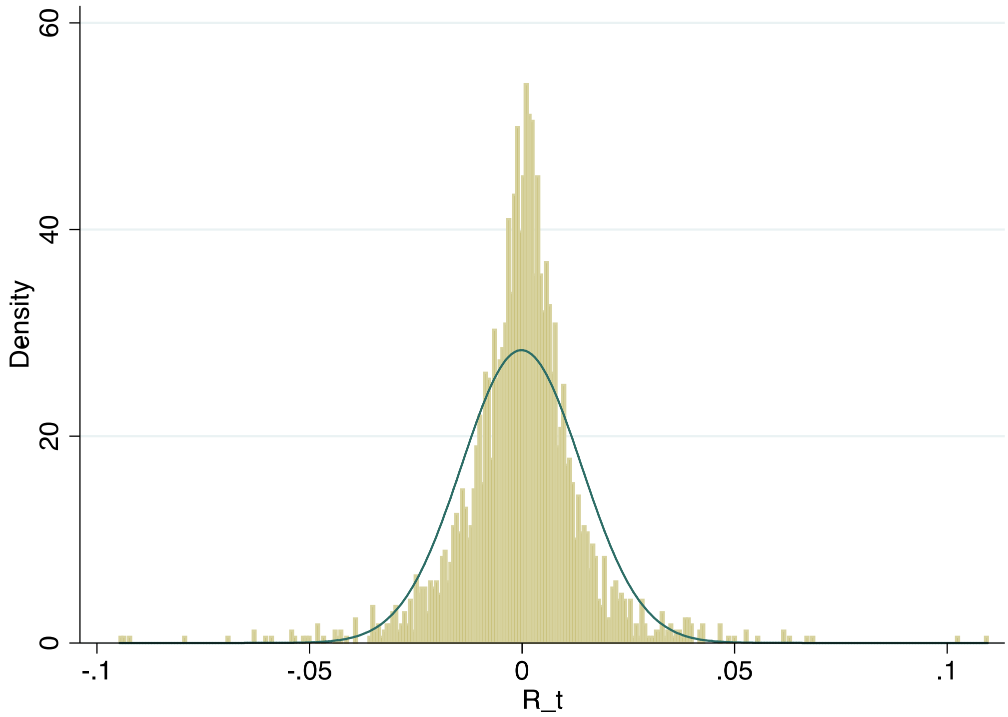


Figure 4.3 Histogram of daily S&P500 returns and the normal distribution

The unconditional distribution of daily returns does not follow the normal distribution. Figure (4.3) shows a histogram of the daily S&P500 return data with the normal distribution imposed. It is visible that the histogram does not match the histogram of the normal distribution.

The histogram of S&P500 is more peaked around zero, and there are some events in either sides (tails) of the histogram, which compared to the normal distribution shouldn't have existed. The horizontal axis starts from -0.1 and ends around 0.1, whereas, most of the events happen between -0.05 and 0.05. These events are called extreme events. While the kurtosis of the standard normal distribution is 3, we expect the kurtosis for the sample data to be greater than 3 due to all the extreme events apparent from the graph. This is called excess kurtosis.

Kurtosis measures the fatness of the tails of a distribution. Positive excess kurtosis means that distribution has fatter tails than a normal distribution. Fat tails means there is a higher than normal probability of big positive and negative returns. Positive excess kurtosis indicates a leptokurtic distribution. The higher the kurtosis coefficient is above 3, the more likely that future returns will be either extremely large or extremely small. In fact the kurtosis in this case is 10.58.

	N	mean	sd	skewness	kurtosis
R_t	2472	-0.000137	0.0141	-0.0993	10.58
Q (10)			50.9962		
Q 2(10)			2046.8341		
Breusch-Godfrey			93.314		
Correlation between $R_t$and R_t^2			-0.0384		

Table 4.1 Summary Statistics of daily return, S&P500

The skewness for our data is negative which resembles the fact that there are very large drops but not equally large increases (there are more negative returns than positive returns). This is also observed in figure (4.3). The return distribution is negatively skewed (skewed to the left) and it is asymmetric. The skewness is -0.01. Negatively skewed distributions have a long left tail, which for investors can mean a greater chance of extremely negative outcomes.

The standard deviation of returns completely dominates the mean of returns. The mean of the returns is much smaller than the standard deviation of returns. It is observable from figure (4.1) how the returns deviates from average mean.

Variance measured by squared returns, displays positive correlation with its own past. Figure (4.4) shows the autocorrelations of squared returns, that is:

$$corr\left(R_t^2, R_{t-\tau}^2\right) > 0, \qquad (4.7)$$

Financial indices display negative correlation between changes in variance and returns. As asset prices decline, companies become mechanically more leveraged.

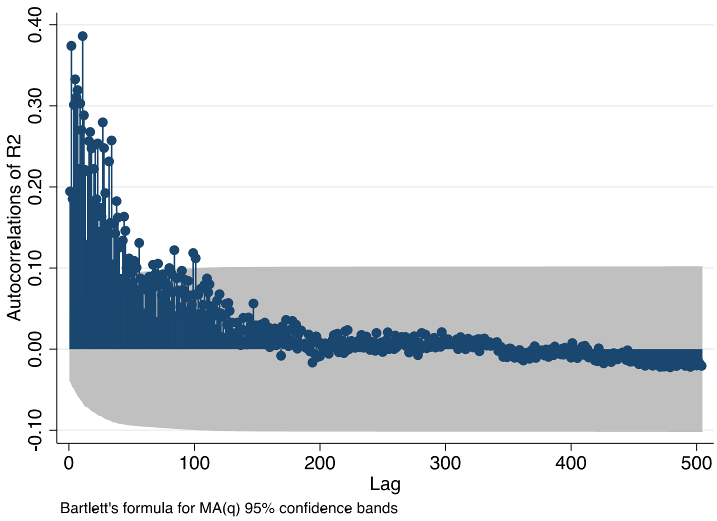


Figure 4.4 Autocorrelation of squared daily S&P500 returns for up to 504 days

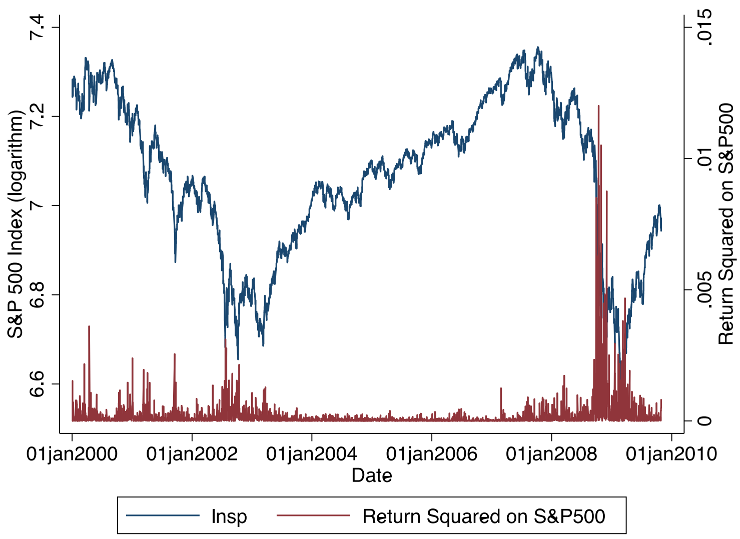


Figure 4.5 S&P500 Index (logarithm) and returns squared.

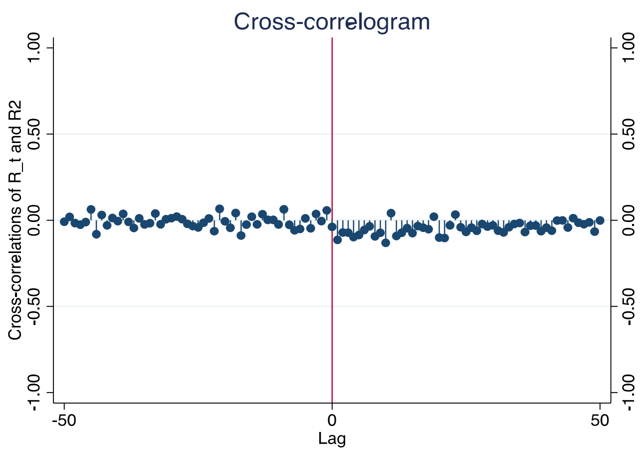


Figure 4.6 Cross-correlation between $R_t$and R_t^2

Since the relative value of companies' debt rises relative to that of their equity, it is natural to expect that their stock becomes riskier, hence more volatile. The effect is generally asymmetric: other things equal, declines in stock prices are accompanied by larger increases in volatility than the decline in volatility. This is often called the leverage effect, arising from the fact that a drop in a stock price will increase the leverage of the firm as long as debt stays constant. This increase in leverage might explain the increase in variance associated with the price drop. ``Bad'' news tends to have a larger effect on the variance than ``good'' news. The tendency for the variance to fall when returns rise and to rise when returns fall is known as the ``leverage effect''. See figure (4.5).

Consistent with stylized facts of asset returns, a generic form of individual asset return could be:

$$R_{t+1} = \mu_{t+1} + \delta_{t+1} z_{t+1},\ with\, z_{t+1} \sim N(0,1), \qquad (4.8)$$

The random variable z_{t+1} is an innovation term, which we assume is identically and independtly distributed (i.i.d) according to the distribution $N(0,1)$, which has a mean equal to zero and variance equal to one. The assumption that the innovation to asset return is normally distributed is not realistic but for the simplicity it is assumed here. The conditional mean of return, $E_t(R_{t+1})$ is μ_{t+1} and the conditional variance; $E_t(R_{t+1} - \mu_{t+1})^2$is σ^2_{t+1}. Most of the time we will assume that the conditional mean of the return, μ_{t+1} is simply zero since it is dominated by the standard deviation of the return. For daily data this is a reasonable assumption (see table (4.1)).

Given the assumptions made, we can write the daily return as:

$$R_{t+1} = \delta_{t+1} z_{t+1},\ with\, z_{t+1} \sim N(0,1),$$

These assumptions imply that once a model of the time-varying variance, σ^2_{t+1} is found, the entire distribution of the asset is known. Our focus will be to establish a model for forecasting variance.

4.4 Modelling Volatility

Definition of Volatility

A variable's volatility, δ, is defined as the standard deviation of the return per unit of time when the return is expressed using continuous compounding (R_t):

$$\widehat{VAR(R_t)} = \frac{1}{T}\sum_{t=1}^{T}(R_t - \overline{R})^2, \qquad (4.9)$$

Equation (4.9) shows that squared observations are the building blocks of the variance of the series, which suggests that the variation (volatility) of the series is time dependent.

Since variance, as measured by the squared returns, exhibits strong autocorrelation (figure (4.4)), if the recent period was one of high variance, then tomorrow is likely to be a high-variance day as well. The easiest way to capture this phenomenon is by letting tomorrow's variance be the simple average of the most recent m observations:

$$\sigma_{t+1}^2 = \frac{1}{m}[R_t^2 + R_{t-1}^2 + R_{t-2}^2 + \dots + R_{t+1-m}^2], \quad (4.10)$$

Equation (4.10) will be a simple forecast for σ_{t+1}^2 since the forecast for tomorrow's variance will be instantaneously available at the end of today when the return is realised. However, there is a big downside with this model, which is the fact that the model puts equal weights of $\frac{1}{m}$on the past m observations.

An extreme return (either positive or negative) which has happened today will affect (bump up) the variance by $\frac{1}{m}$ times the return squared for exactly m periods ahead. However this is not what is observed in figure (4.4), which is a gradual decline in the autocorrelation of the squared return series. A high m will lead to an excessively smooth σ_{t+1}^2, whereas, a low m will lead to a jagged pattern of σ_{t+1}^2 over time. There are two problems with using equation (4.10):

1. It is not clear how many lags, m, should be used.
2. Figure (4.4) shows a gradual decline, and $corr(R_t^2, R_{t-1}^2)$ is clearly more important than $corr(R_t^2, R_{t-504}^2)$, whereas equation (4.10) suggests otherwise.

The autocorrelation plot of the squared returns in figure (4.4) suggests that a more gradual decline is expected in the effect of past returns on today's variance.

An alternative approach could be the exponential smoother, which takes into account the fact that the weights on past squared returns decline exponentially as we move backward in time (what the value has been at time t-1 has far more effect to the value at time t+1 compared to the value at time t-504), giving more weights to the more recent data:

$$\sigma_{t+1}^2 = (1-\lambda)\sum_{\tau=1}^{\infty} \lambda^{(\tau-1)} R_{(t+1-\tau)}^2, for\ 0<\lambda<1, \quad (4.11)$$

It is easy to see that equation (4.11) can be written as:

$$\sigma_{t+1}^2 = \lambda\sigma_t^2 + (1-\lambda)R_t^2, \qquad (4.12)$$

The forecast for tomorrow's volatility can thus be seen as a weighted average of today's volatility and today's squared return. Equation (4.11) corresponds to weights that decrease exponentially. If we substitute for σ_t^2 to get:

$$\sigma_{t+1}^2=\lambda\,[\lambda\sigma_{t-1}^2+(1-\lambda)\ R_{t-1}^2]+(1-\lambda)\ R_t^2=(1-\lambda)\ (R_t^2+\lambda R_{t-1}^2)+\lambda^2\sigma_{t-1}^2$$

Substituting in a similar way for σ_{t-1}^2, we get:

$$\sigma_{t+1}^2=(1-\lambda)\ (R_t^2+\lambda R_{t-1}^2+\lambda^2 R_{t-2}^2)+\lambda^3\sigma_{t-2}^2$$

Continuing in this way, we see that:

$$\sigma_{t+1}^2 = (1-\lambda)R_t^2+\lambda R_{t-1}^2+\lambda^2 R_{t-2}^2+\lambda^3 R_{t-3}^2+\ldots+\lambda^m R_{t-m}^2)+\lambda^m\sigma_{t-m+1}^2$$

For a large m, the term $\lambda^m\sigma_{t-m+1}^2$ is sufficiently small and could be ignored. The weights for R_t^2 decline at the rate of λ.

This model has some advantages. First, it tracks variance changes in a way that is broadly consistent with observed returns. Recent returns matter more for tomorrow's variance than distant returns as λ is less than one and therefore the impact of the lagged squared return gets smaller when the lag τ gets bigger.

Second, the model only contains one unknown parameter, namely λ.

Third, this approach has the attractive feature that the data storage requirements are quite modest. At any given time, we need to remember only the current estimate of the variance rate and the most recent observation on the value of the market variable. When we get a new observation on the value of the market variable, we update our estimate of the variance rate. The old estimate of the variance rate and the old value of the market variable can then be discarded.

This approach is designed to track changes carefully in the volatility. Suppose there is a big move in the market variable on day t-1 so that R_t^2 is large. From equation (4.12) this causes our estimate of the current volatility to move upward. The value of λ decides how responsive the estimate of the daily volatility is to the most recent daily percentage change. A low value of λ leads to a great deal of weight being given to the R_t^2 when σ_{t+1}^2 is calculated. In this case, the estimates produced for the volatility on successive days are themselves highly volatile. A high value of λ (close to 1.0) produces estimates of the daily volatility that respond relatively slowly to new information provided by the daily change.

When estimating λ on a large number of assets, the estimates can be quite similar across assets. The research carried out by JPMorgan in 1989 which was made publicly available later simply set λ =0.94 for every asset for daily variance forecasting. In this case no estimation is required, which is a huge advantage in large portfolios. The weight on today's squared returns is 1- λ =0.06, and the weight is exponentially decaying to $(1-\lambda)\lambda^{99}$=0.000131 on the 100th lag of squared return. After including 100 lags of squared returns, the cumulated weight is 0.998, so that 99.8% of the weight has been included. Therefore we could assume that it is only necessary to store about 100 daily lags of returns in order to calculate tomorrow's variance, σ_{t+1}^2. This model however does have certain shortcomings.

4.5 The GARCH Variance

GARCH models are a set of models that can capture features of returns that are quite flexible. The simplest Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model of dynamic variance can be written as

$$\sigma_{t+1}^2 = \omega + \alpha R_t^2 + \beta \sigma_t^2, \ \alpha + \beta < 1, \quad (4.13)$$

Where $\omega = \sigma^2(1 - \alpha - \beta)$ and σ is the long-run average variance. Equation (4.12) can be viewed as a special case of the simple GARCH model if we force $\alpha = 1 - \lambda$ and $\beta = \lambda$, so that $\alpha + \beta = 1$ and further $\omega = 0$.

Tomorrow's variance is a weighted average of the long-run variance, today's squared return, and today's variance. Equation (4.12) ignores the long-run variance.

The term $\alpha + \beta$ is refered to as the persistence of the model. A high persistence ($\alpha + \beta$ close to 1) means that shocks that push variance away from its long-run average will persist for a long time, but eventually the long horizon forecast will be the long run average variance, σ^2.

When $\alpha + \beta = 1$ as it is in equation (4.12), persistence is 1, which implies that a shock to variance persists forever, meaning that an increase in variance will push up the variance forecast for all future forecast horizons.

Equation (4.12) ignores the long-run variance when forecasting which is one the shortcomings of this model.

If $\alpha + \beta$ is close to 1 as is typically the case, then the two models give similar predictions for short horizons, but their longer horizon forecasts are very different. If today is a high-variance, then equation (4.12) predicts that all future days will be high-variance. The GARCH model however more realistically assumes that eventually in the future variance will revert to the average value. For a stable GARCH process, we require $\alpha + \beta < 1$.

The model in equation (4.13) is GARCH(1,1) since there are only one lag of R_t^2 and σ_t^2 in the model.

The bottom part of figure (4.1) shows the daily returns and figure (4.5) shows the squared return series and the closing prices against time. We can see from these graphs that:

i) R_t^2 changes as t changes, i.e., varying volatility.

ii) R_t^2 changes in a rather unique pattern, i.e., volatility clustering (calm periods followed by volatile periods).

iii) as R_t^2 increases closing prices decreases, i.e., leverage effect.

Any model deployed should be able to at least accommodate these facts. This chapter introduces three volatility models from the GARCH family, namely ARCH, GARCH and TGARCH.

Consider equation (4.8) in which we will first assume the stock prices follow a random walk model:

$$R_t = \mu_t + u_t, \qquad (4.14)$$

$$VAR(u_t) = \sigma^2$$

u_t is a random term whose variance on day t is σ^2. We can now improve equation (4.14) to accommodate changes in the variance of u_t:

Let h_t be the variance of ut conditional on the value of u_{t-1}. That is:

$$h_t = V\left(u_t \middle| u_{t-1}\right)$$

The ARCH (1) model (Autoregressive Conditional Heteroscedasticity, Engle, 1982) is defined as:

$$h_t = \omega + \alpha u_{t-1}^2, \quad (4.15)$$

The ARCH (1) model (4.15) assumes that the variance in the current period depends on the value of the series in the previous period. The order of the

ARCH process is 1 because only 1 lag appears in (4.15). A positive value of the parameter αindicates an "ARCH effect", which essentially implies that the data contain volatility clusters.

The GARCH (1, 1) model (Generalised ARCH, Bollerslev, 1986) is defined as:

$$h_t=\omega + \alpha u_{t-1}^2+\beta h_{t-1}, \quad (4.16)$$

Here, it is assumed that the variance in the current period depends on the value of the series in the previous period and on the variance in the previous period. In (4.16), a positive value of the parameter α again indicates an "ARCH effect", while a positive value of the parameter β indicates a "GARCH effect". The latter can be interpreted as implying that the volatility clusters showing persistence.

In the basic GARCH model, since only squared residuals enter the conditional variance equation, the signs of the residuals have no effect on conditional volatility. However, a stylised fact of financial volatility is the leverage effect, which has already been discussed in section (4.2). The TGARCH model (Threshold ARCH, or "GJR-GARCH", Glosten et al., 1993), which is a further generalisation of ARCH and GARCH, allows the effects of good and bad news to have different effects on volatility.

The TGARCH (1, 1, 1) model is defined as:

$$h_t=\omega + \alpha u_{t-1}^2+\beta h_{t-1} + \gamma d_{t-1} u_{t-1}^2, \qquad (4.17)$$

In (4.17), dt-1 is a dummy variable that is equal to +1 if $u_{t-1} > 0$ ("good" news in previous period) and 0 if $u_{t-1} < 0$ ("bad" news in previous period). u_{t-1}=0 is a "threshold" in the sense that shocks greater than this threshold have a different effect on volatility from shocks below the threshold.

The intuition of (4.17) is as follows. If last period's news was "bad" ($u_{t-1} < 0$), its effect on this period's volatility is $+\gamma u_{t-1}^2$. If last period's news was "good" ($u_{t-1} > 0$), its effect on this period's volatility is lower at ($+\alpha u_{t-1}^2$ + βh_{t-1}). In accordance with this intuition, we expect the parameter $\gamma$to be negative. A significantly negative estimate of γ therefore provides evidence of the leverage effect.

As discussed in section (4.2) the return series have significant autocorrelations of squares and leverage effect. Therefore, we fit the ARCH, GARCH, and TGARCH models with generalized error distribution for the errors. The results are presented in table (4.2) and the fitted models are as follows:

ARCH (1):

$$h_t = 0.0001 + 0.3370u_{t-1}^2$$

GARCH(1,1):

$$h_t = 0.0000 + 0.0723u_{t-1}^2 + 0.9241h_{t-1}$$

TGARCH(1,1,1):

$$h_t = 0.0000 + 0.1129u_{t-1}^2 + 0.9464h_{t-1} - 0.1369d_{t-1}u_{t-1}^2$$

After estimating the conditional volatilities for each of the models, $\widehat{\sigma_t}$, the residuals are computed as $\widehat{\varepsilon_t} = \frac{u_t}{\widehat{\sigma_t}}$. Table (4.2) also reports several diagnostics based on these residuals. For the ARCH model the kurtosis hasn't reduced, but for the GARCH(1,1) and TGARCH models the kurtosis have been clearly reduced with respect to the sample kurtosis of S&P500 returns which is 10.58. Furthermore, the autocorrelations of squared residuals are significant for the ARCH model, and not significant for GARCH and TGARCH models. Therefore, it seems that the GARCH and TGARCH fitted models have been successful in representing the dynamic evolution of the squares and the kurtosis observed in the S&P500 returns. The ARCH coefficient is significant, so the magnitude of last period's error (u_t) has a positive effect on this period's volatility.

The predicted variances are plotted in figure (4.7) and we can see that the estimated ARCH variances are clearly different to the GARCH and TGARCH estimated variances.

ALL S&P 500			
VARIABLES	ARCH(1)	GARCH(1,1)	TGARCH(1,1,1)
L.ARCH	0.334**	0.0723***	0.113***
	(0.0511)	(0.0106)	(0.0133)
L.GARCH		0.924***	0.946***
		(0.0105)	(0.00827)
L.TGARCH			-0.137***
			(0.0149)
CONSTANT	0.0001***	8.08e-06**	8.84e-07
	(7.07e-06)	(42.69e-07)	(1.82e-07)
AIC	-14579.7	-15573.93	-15688.74
LOG-LIKELIHOOD	7292.849	7790.964	7849.368

Residuals			
Mean	-5.088	-0.0609	-0.0284
S.D.	79.9474	0.999	1.0014
Skewness	-0.722	-0.333	-0.3466
Kurtosis	10.2784	4.3084	4.09042
Q(10)	43.8912	9.9644	9.6889
Q2(10)	350.0440	16.5307	15.0006

Table 4.2 Estimated models for the daily returns of SP500 index under generalised errors.

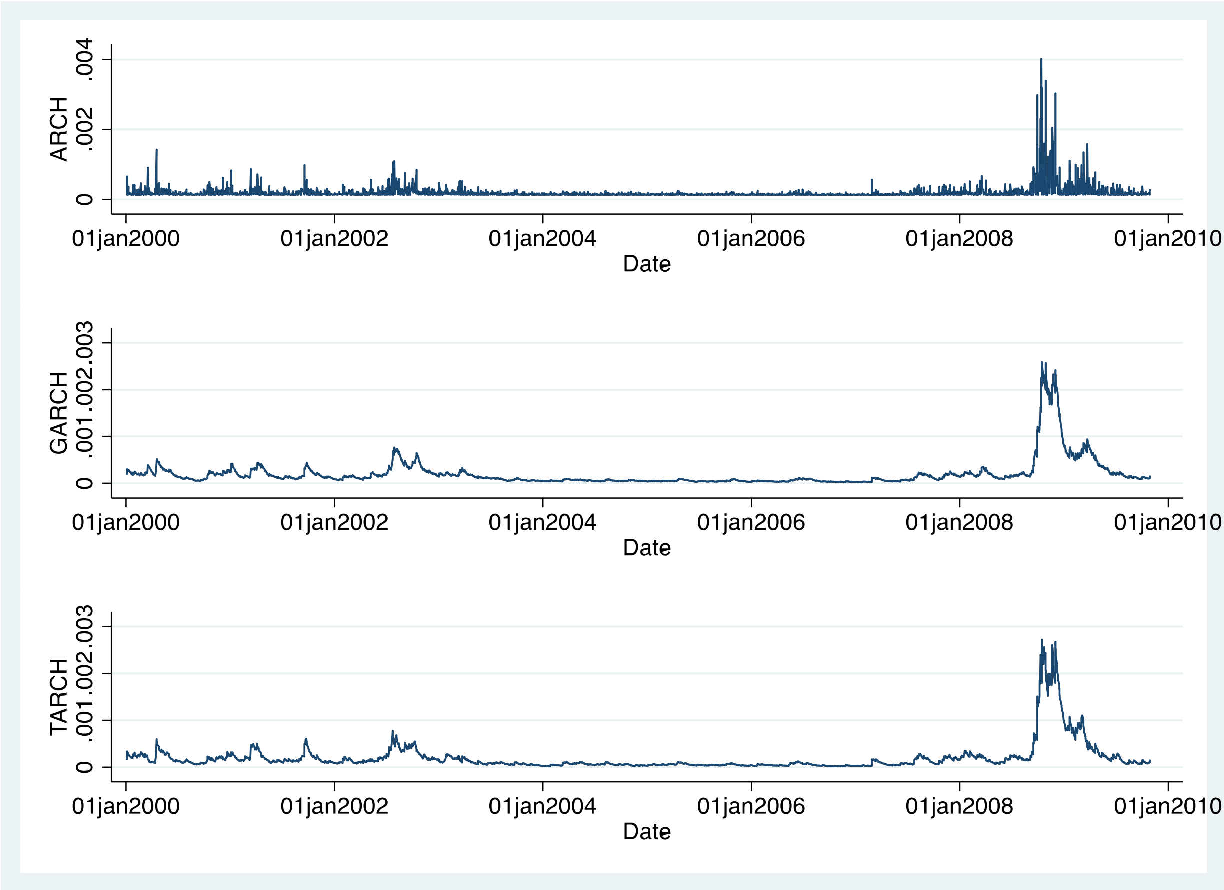
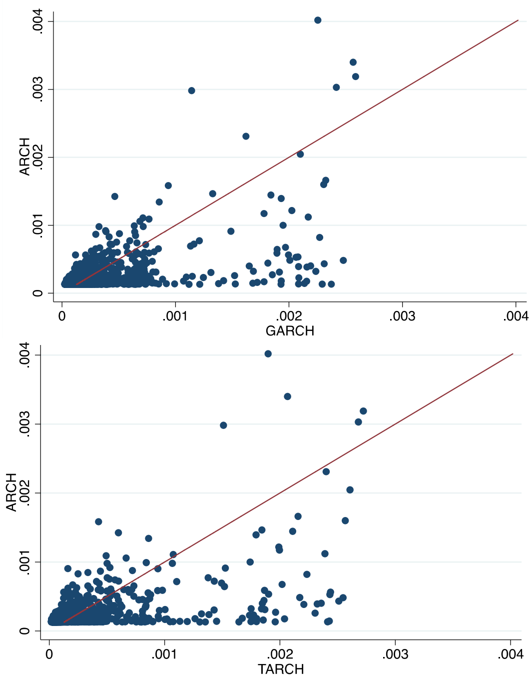


Figure 4.7 Conditional variances (from top to bottom: ARCH, GARCH, TGARCH), one-step ahead



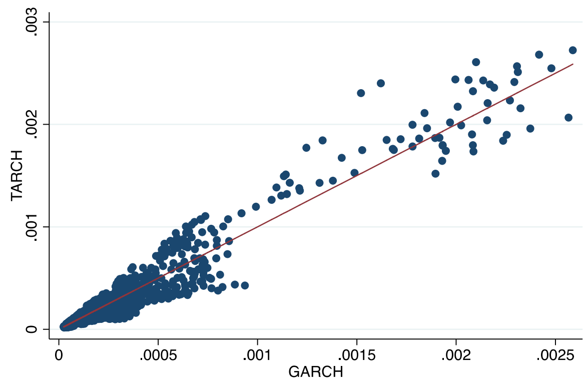


Figure 4.8 Scatter plots of conditional volatilities against each other, obtained after fitting each models for S&P500 returns

The final goal in this section when fitting conditionally heteroskedastic models is to obtain estimates of the underlying volatilities. The differences among the variances estimated by each models may be important when using in financial application. In fact it is important for our next chapter.

The main diagonal of figure (4.8) plots estimates of the volatility obtained after fitting each of the three models. The general shape of the estimated volatilities of GARCH and TGARCH is similar, while there is a clear difference between ARCH against GARCH and TGARCH. This is further in line with the outcome of table (4.2) that GARCH and TGARCH should be preferred to ARCH.

Model Comparisons Using Akaike Information Criterion (AIC)

One of the most popular methods to compare the three different models is to use Akaike information criterion (AIC). Given a collection of models for the data, AIC estimates the quality of each model, relative to each of the other models. If we assume that the ARCH model has K number of estimated parameters and the maximum value of the likelihood function is L, then the AIC value of the model is the following:

$$AIC = 2K - 2\ln L$$

Given a set of candidate models for the data, the preferred model is the one with the minimum AIC value.

According to table 4.2, the ARCH(1) model gives us the AIC value of -14579.7, the GARCH(1,1) model gives us the value of -15573.93 and the TARCH model gives us the value of -15688.74. Hence, the preferred model is TARCH model with the lowest AIC value.

4.6 Bear, Calm, and Crash Markets

We again divide the sample into the three "market periods", "bear market", "calm market" and "crash market", as explained in Chapter 1, Section 1.3. Figure 4.9 reminds us of the division.

Figure 4.9 S&P500 prices (logarithm)

	N	mean	sd	skewness	kurtosis	Q(10)	Q 2(10)
R_1	751	-0.001	0.015	0.191	4.134	5.959	130.4152
R_2	1199	0.0004	0.001	-0.136	4.696	28.981	280.095
R_3	520	-0.0008	0.022	-0.087	7.174	28.353	350.241

Table 4.3 Summary statistics of the data across the three ranges

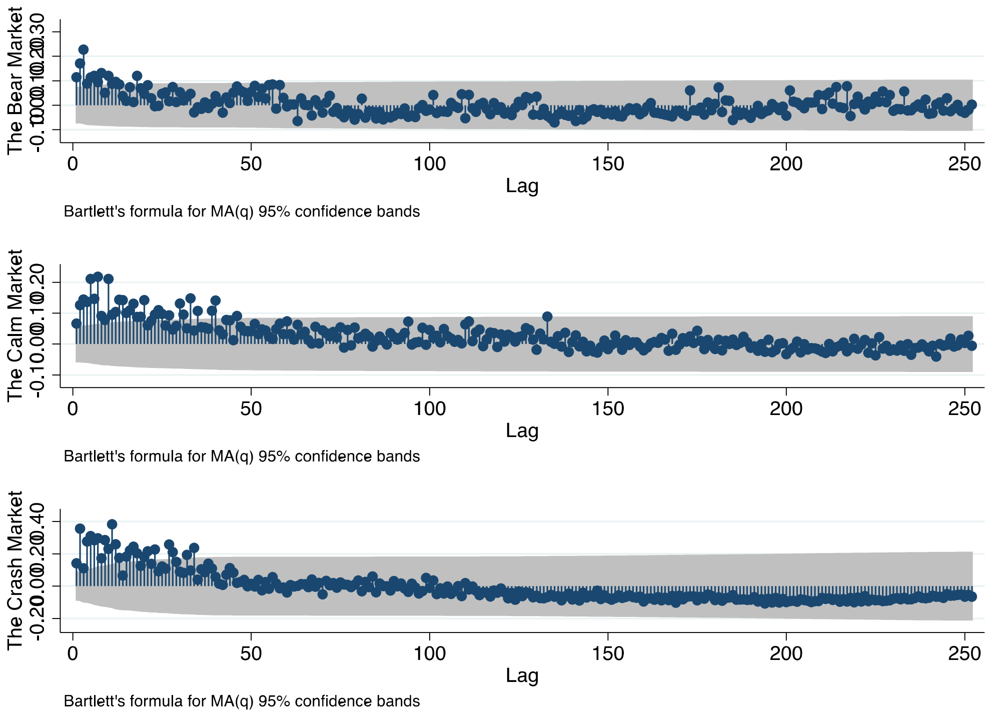


Figure 4.10 Autocorrelation of squared returns for each market separately

The summary statistics of the data in table (4.3) divided into three ranges show how different the return series are. While range 1 has the lowest kurtosis, range 3 has the highest kurtosis. Also figure (4.10) shows the autocorrelation of the squared return series for up to 100 lags. Notice how different the Bear market looks in comparison with the other two.

4.7 Estimation Results

We now estimate the three models of interest ARCH, GARCH, and TGARCH. The main purpose of the results will be in the construction of the data generating processes (DGPs) for the Monte Carlo analyses done in Chapter 5. For this reason, it is very important to use an estimation sample that is previous to the time period in which the options are traded. The options we consider in Chapter 5 will be from the years 2002, 2005 and 2008. Hence for estimation we will use daily data from the two years previous to each of these years. So, we will use: 2000-2001 ("pre-Bear market"); 2003-2004 ("pre-calm market"); 2006-2007 ("pre-crash market").

Range 1 (pre-Bear Market; 1 Jan 2000 – 31 December 2001):

TGARCH (1,1,1):

$$h_t = 0.0000 + 0.1508u_{t-1}^2 + 0.930h_{t-1} - 0.1897d_{t-1}u_{t-1}^2$$

GARCH(1,1):

$$h_t = 0.0000 + 0.105u_{t-1}^2 + 0.856h_{t-1}$$

ARCH(1):

$$h_t = 0.0002 + 0.148u_{t-1}^2$$

Range 2 (pre-Calm Market; 1 Jan 2003 – 31 December 2004):

TGARCH (1,1,1):

$$h_t = 0.0000 + 0.0915u_{t-1}^2 + 0.938h_{t-1} - 0.0989d_{t-1}u_{t-1}^2$$

GARCH(1,1):

$$h_t = 0.0000 + 0.0552u_{t-1}^2 + 0.925h_{t-1}$$

ARCH(1):

$$h_t = 0.0000 + 0.0573u_{t-1}^2$$

Range 3 (pre-Crash Market; 1 Jan 2006 – 31 December 2007):

TGARCH (1,1,1):

$$h_t = 0.0000 + 0.132u_{t-1}^2 + 0.931h_{t-1} - 0.158d_{t-1}u_{t-1}^2$$

GARCH(1,1)

$$h_t = 0.0000 + 0.0927u_{t-1}^2 + 0.895h_{t-1}$$

ARCH(1)

$$h_t = 0.0004 + 0.220u_{t-1}^2$$

Table 4.4 summarises the results of estimated ARCH, GARCH, and TGARCH for the bear, calm and crash markets respectively.

RANGE 1: PRE-BEAR MARKET (1 JAN 2000 – 31 DECEMBER 2001)			
VARIABLES	ARCH	GARCH	TGARCH
L.ARCH	0.148**	0.105***	0.151***
	(0.0577)	(0.0281)	(0.0242)
L.GARCH		0.856***	0.930***
		(0.0389)	(0.0183)
L.TGARCH			-0.190***
			(0.0264)
CONSTANT	0.000184***	8.64e-06**	2.30e-06
	(1.54e-05)	(4.30e-06)	(1.63e-06)
AIC	-2859.325	-2882.619	-2916.917
LOG-LIKELIHOOD	1432.662	1445.309	1463.459
RANGE 2: PRE-CALM MARKET (1 JAN 2003 – 31 DECEMBER 2004)			
VARIABLES	ARCH	GARCH	TGARCH
L.ARCH	0.0573	0.0552***	0.0915***
	(0.0367)	(0.0166)	(0.0222)
L.GARCH		0.925***	0.938***
		(0.0239)	(0.0193)
L.TGARCH			-0.0989***
			(0.0239)
CONSTANT	6.00e-05***	1.21e-06*	9.87e-07**
	(3.92e-06)	(6.21e-07)	(4.21e-07)
AIC	-3312.907	-3372.253	-3376.172
LOG-LIKELIHOOD	1659.453	1690.126	1693.086
RANGE 3: PRE-CRASH MARKET (1 JAN 2006 – 31 DECEMBER 2007)			
VARIABLES	ARCH	GARCH	TGARCH
L.ARCH	0.220**	0.0927***	0.132***
	(0.0974)	(0.0297)	(0.0345)
L.GARCH		0.895***	0.931***
		(0.0326)	(0.0232)
L.TGARCH			-0.158***
			(0.0395)
CONSTANT	0.000362***	5.21e-06	3.82e-06*
	(4.15e-05)	(3.81e-06)	(1.95e-06)
AIC	-3370.252	-3441.619	-3477.074
LOG-LIKELIHOOD	1688.126	1724.809	1743.537

Table 4.4 Parameter Estimates from ARCH, GARCH and TGARCH models.

4.8 Summary

The importance of volatility in option pricing is clear from previous chapters. In this chapter, we have used more advanced models to estimate volatility, in place of the constant volatility assumption (used in the Black-Scholes-Merton model). The models we have used are ARCH family models: ARCH (1); GARCH(1,1); and TGARCH (1, 1, 1) model.

We used daily data on the S&P 500 Index to estimate the Random Walk model, and the three ARCH family models. We found that the TGARCH model is the best model to represent the volatility process, since all of its coefficients are strongly significant, it has the lowest AIC of the four models, and it passes all the statistical diagnostic tests. Then we simulated data from the models considered using the coefficients from the estimated models. After that, we found that the simulated ARCH family volatility models worked correctly, since the "true" parameter values are included in the confidence intervals. Therefore, we are ready to use the estimated models as the basis for Monte Carlo simulations to estimate the values of European Options. This is the subject of chapter 5.

Chapter 5 Using the Monte Carlo method to value European Options

5.1 Introduction

In Chapter 4, we used the ARCH, GARCH and TGARCH models to analyse the volatility process of the S&P 500 Index. We found clear evidence that the volatility of the Index is time-varying and TGARCH is the best model to represent the volatility process. The objective of this chapter is to develop methods for valuing options that do not depend on the assumption of a random walk with constant return variance, but instead allow for the time-varying volatility patterns of the type encountered when analysing asset price data. Some theoretical work has been done to extend the BSM framework to allow for time varying volatility. For example, Duan (1995), Heston and Nandi (2000), and Christoffersen and Jacobs' (2017) all derive formulae for valuing options when a GARCH process is assumed for the underlying returns.

The Monte Carlo simulation approach has proved to be a valuable and flexible computational tool in security pricing, especially in valuing American style options. Monte Carlo simulation is a good method for valuing American options because it makes it straightforward to derive the expected return from exercising the American option early, as long as we assume that the distribution for the underlying assets' price is log-normal. Tilley (1993), Barraquand and Martineau (1995), and Broyle et al. (1997), all applied the Monte Carlo simulation methods to value the American options. However, little research has been done to use the Monte Carlo simulation to value the European options. Unpublished work by Alexandros Kyrtsos[8] used Monte Carlo simulation to price European options. For the simulation, he assumed a geometric Brownian motion process for underlying stock price. He then compared the value of European options from the Monte Carlo simulation with the value from Black-Scholes formula. He found that they are the same, as expected. He writes "The results from the BSM model and the simulation exhibit perfect agreement". Although the Monte Carlo method has been used extensively in option valuation (see Glasserman, 2003), to our knowledge, few papers have applied Monte Carlo models with time-varying volatility, in the context of European option valuation. Duan et al. (2006) provide analytical formulae for the EGARCH and GJR-GARCH (TGARCH) models and in their analysis, they benchmark the analytical implementation of the models against the Monte Carlo ones. However, their Monte Carlo simulation was applied to fictional European

[8] http://polymer.bu.edu/hes/py538kyrtsos.pdf

options. In this chapter, we will break new ground by applying the Monte Carlo simulation method to the problem of valuing real market options.

We pursue this objective using Monte Carlo simulation to estimate the values of S&P 500 options. First of all, we will use Monte Carlo simulation methods to simulate the ARCH, GARCH and TGARCH. All Monte Carlo simulations are run in STATA. Secondly, we will calculate values of European call and put options from three different models ARCH, GARCH and TGARCH, and we will try to find which model is the best model for predicting the option price. After we have identified the best model, we will compare its predictive performance with the benchmark PBS model, by finding which model has predictions closest to observed market option prices. In order to make this comparison, and also compare the best Monte Carlo model with the benchmark PBS model, we will use the Root mean square error (RMSE) criterion for the out-of-sample predictions.

Note that when the Monte Carlo method is applied to predict option prices, the data generating process (DGP) is obtained using an estimation sample that is previous to the dates on which the options are traded. For example, the prices of our 2002 options are obtained using a DGP consisting of estimates from a sample of daily data from 2000-2001. This means that predictions of prices resulting from the Monte Carlo method are always out-of-sample predictions.

5.2 Monte Carlo Simulations for valuing European options

The Monte Carlo procedure, applied to the problem of valuing a European option, consists of the following steps (Glasserman, 2003):

(1) Simulate a sample path of the underlying price over relevant time horizon. We will simulate these paths assuming either random walk, ARCH, GARCH, or TGARCH, and assuming the parameter estimates obtained in Chapter 4.
(2) For the simulated sample path, calculate the discounted payoff from holding the option to expiry. That is, calculate $\exp(-r^{\tau})\max[0,S_T-K]$ for the call options and calculate $\exp(-r^{\tau})\max[0, K-S_T]$ for the put options;
(3) Repeat step 1 and 2 to get many sample paths, and many discounted payoffs;
(4) Take mean of discounted payoffs over the replications.

In this section, we will carry out three different Monte Carlo simulations: one assuming ARCH; one assuming GARCH; and one assuming TGARCH. For the assumed parameter values for each model, we will use the estimates obtained in chapter 4, presented in Table 4.6. First of all, we will describe the option data to which we apply the Monte Carlo method. Secondly, we will use the Monte Carlo simulations of the ARCH, GARCH and TGARCH models to calculate the European option prices, and we will also use these three models to predict the market price of the European options.Thirdly, we compare the predictive performance of the ARCH, GARCH and TGARCH models. The prediction criterion we will use is the root mean square error (RMSE, see Christoffersen and Jacobs, 2004; Andreou et al., 2008&2014 ; Chang et al., 2012; Bates, 2000), and we will apply this to assessing predictive performance out-of-sample. Importantly, we will compare the predictive performance of these models to that of the Practitioners Black Scholes (PBS) model, which will again be treated as a benchmark model (see Chapter 3).

5.2.1 Option Data

Before we discuss the Monte Carlo simulations, we will describe the data used in this chapter.

In chapter 3, we mentioned that the option data set were divided into three time periods: (i) “Bear Market”: Jan 2000 to 31 Dec 2002; (ii) “Calm Market”: Jan 2003 to 09 Oct 2007; (iii) “Crash Market”: 10 Oct 2007 to Oct 2009. In order to keep things consistent, we will use these three time periods to carry out the Monte Carlo simulations. From the bear market, we will take options from 2002 only for the Monte Carlo simulation. From the calm market, we will take options from 2005 only. From the crash market we will take options from 2008 only. In 2002, there are 12,901 call options and 16,168 put options; in 2005, there are 13,468 call options and 18,185 put options; in 2008, there are 28,960 call options and 36,071 put options. Because of the limitations of the computation speed of Monte Carlo, we will randomly select 10% of these options to apply the Monte Carlo simulations. The eventual numbers of each type of options used for the Monte Carlo simulation are presented Table 5.1.

	Call Option	**Put Option**
2002	1,290	1,617
2005	1,349	1,819
2008	2,896	3,607
Total	5,535	7,403

Table 5.1 Numbers of options used in the Monte Carlo simulation

5.2.2 The Monte Carlo Simulations of ARCH, GARCH and TGARCH models

In this section, we will carry out three different Monte Carlo simulations: one assuming ARCH; one assuming GARCH; and one assuming TGARCH. For the assumed parameter values for each model, we will use the estimates obtained in chapter 4, presented in Table 4.4. Recall that the estimates in this table were obtained using daily data on the S&P500 index for three different markets: bear market (2002), calm market (2005) and crash market (2008).

We will run the Monte Carlo simulation of three different models for each market using the results from table 4.4.

ARCH model (simulation 1):

The data generating process (DGP) for the ARCH Monte Carlo simulation is given as follows:

$$\tilde{S}_{ARCH,T} = \exp\left(\ln\left(S_0\right) + \sum_{t=1}^{T} \ln\left(1 + \tilde{r}_{ARCH,t} + r_d\right) \right) \quad (5.1)$$

Where:

For the bear market (2002), $\tilde{r}_{ARCH,\,t}$ is given by:

$$\tilde{r}_{ARCH,\,t} = Z_t\sqrt{0.00184 + 0.148\tilde{r}_{ARCH,\,t-1}{}^2}$$

For the calm market (2005), $\tilde{r}_{ARCH,\,t}$ is given by:

$$\tilde{r}_{ARCH,\,t} = Z_t\sqrt{0.00006 + 0.0573\tilde{r}_{ARCH,\,t-1}{}^2}$$

For the crash market (2008), $\tilde{r}_{ARCH,\,t}$ is given by:

$$\tilde{r}_{ARCH,\,t} = Z_t\sqrt{0.0000362 + 0.220\tilde{r}_{ARCH,\,t-1}{}^2}$$

Where Z_t is a simulated standard normal.

And r_d(daily risk free rate) is given by:

$$r_d = \exp\left[\frac{\ln\left(1 + r_f\right)}{252}\right] - 1$$

Hence, the value of a European call option for the ARCH model is obtained using:

$$value_{call_{ARCH}} = E\left[\max\left(\tilde{S}_{ARCH,T} - K, 0\right)\right] * \exp\left(-r_f * \left(days / 252\right)\right)$$

The payoff of a European put option for the ARCH model is obtained using:

$$value_{put_{ARCH}} = E\left[\max\left(K - \tilde{S}_{ARCH,T} -, 0\right)\right] * \exp\left(-r_f * \left(days / 252\right)\right)$$

The DGPs for GARCH and TGARCH are clearly more complicated than the above DGP. However, below we present the STATA code for all three models.

For a put option (with S_t=1119.31; K=900; r_f = 0.0158):

```
clear
prog drop _all
program define TGARCH_sim, rclass
syntax [, days(integer 1 ) s0(real 1.0) k(real 1.0) a0(real 1.0) a1(real
1.0) a2(real 1.0) a3(real 1.0) rf(real 1.0) ]
drop _all
tempvar r S
set obs `days'
scalar rd=exp(ln(1+`rf')/260)-1
gen e=rnormal()
gen v=0 if _n==1
replace  v=`a0'+(`a1'+`a3'*(e[_n-1]>0))*(v[_n-1]*e[_n-1])+`a2'*(v[_n-1])
if v==.
gen r=sqrt(v)*e
gen S=exp(ln(`s0')+sum(ln(1+r+rd)))
return scalar payoff=max(`k'-S[_N],0)*exp(-`rf'*(`days'/260))
              end

**Monte Carlo Simulation of ARCH**
simulate payoff0=r(payoff) , nodots  reps(10000): TGARCH_sim, days(53)
s0(1119.31) k(900) a0(0.0002) a1(0.148) a2(0) a3(0) rf(0.0158)
qui sum payoff0
di "Mean " `format' r(mean)

**Monte Carlo Simulation of GARCH**
simulate payoff0=r(payoff) , nodots  reps(10000): TGARCH_sim, days(53)
s0(1119.31) k(900) a0(0.0000) a1(0.105) a2(0.856) a3(0) rf(0.0158)
qui sum payoff0
di "Mean " `format' r(mean)

**Monte Carlo Simulation of TGARCH**
simulate payoff0=r(payoff) , nodots  reps(10000): TGARCH_sim, days(53)
s0(1119.31) k(900) a0(0.0000) a1(0.1508) a2(0.930) a3(-0.190) rf(0.0158)
qui sum payoff0
di "Mean " `format' r(mean)
```

For a call option(with S_t=1100.64; K=1200; r_f = 0.025):

```
clear
prog drop _all
program define TGARCH_sim, rclass
syntax [, days(integer 1 ) s0(real 1.0) k(real 1.0) a0(real 1.0) a1(real
1.0) a2(real 1.0) a3(real 1.0) rf(real 1.0) ]
drop _all
tempvar r S
set obs `days'
scalar rd=exp(ln(1+`rf')/260)-1
gen e=rnormal()
gen v=0 if _n==1
replace  v=`a0'+(`a1'+`a3'*(e[_n-1]>0))*(v[_n-1]*e[_n-1])+`a2'*(v[_n-1])
if v==.
gen r=sqrt(v)*e
gen S=exp(ln(`s0')+sum(ln(1+r+rd)))
```

```
return scalar payoff=max(S[_N]-`k',0)*exp(-`rf'*(`days'/260))
              end

**Monte Carlo Simulation of ARCH**
simulate payoff0=r(payoff) , nodots  reps(10000): TGARCH_sim, days(120)
s0(1089.84) k(1025) a0(0.0002) a1(0.148) a2(0) a3(0) rf(0.0173)
qui sum payoff0
di "Mean " `format' r(mean)

**Monte Carlo Simulation of GARCH**
simulate payoff=r(payoff) , nodots  reps(10000): TGARCH_sim, days(58)
///
s0(1100.64) k(1200) a0(0.0000) a1(0.105) a2(0.856) a3(0) rf(0.025)
qui sum payoff
di "Mean " `format' r(mean)

**Monte Carlo Simulation of TGARCH**
simulate payoff=r(payoff) , nodots  reps(10000): TGARCH_sim, days(58)
///
s0(1100.64) k(1200) a0(0.0000) a1(0.1508) a2(0.930) a3(-0.190) rf(0.025)
qui sum payoff
di "Mean " `format' r(mean)
```

5.3 Predictive Performance of Monte Carlo Simulations of ARCH, GARCH and TGARCH

In this section, we compare the predictive performance of various Monte Carlo simulation models, including ARCH, GARCH and TGARCH models. The prediction criterion we will use is the root mean square error (RMSE, see Christoffersen and Jacobs, 2004; Andreou et al., 2008&2014 ; Chang et al., 2012; Bates, 2000), and we will apply this to assessing predictive performance only for out-of-sample. Importantly, we will compare the predictive performance of Monte Carlo simulation models to that of the Practitioners Black Scholes (PBS) model, which will be treated as a benchmark model.

5.3.1 Out-of-sample pricing performance of the Monte Carlo Simulations of ARCH, GARCH ,TGARCH, PBS and PBS with smearing

In this section, we will summarizes the pricing performances of the different Monte Carlo simulation models and the benchmark PBS and PBS smearing. We only apply the out-of-sample predictive performance to compare the different models. To determine the pricing accuracy of each model’s estimates, we examine the Root Mean Square Error (RMSE).

First of all, we will use all options to see which model performs best. Secondly, we will divided into three different markets (bear, calm and crash) to investigate which model predict the option price best. Last but not least, we will divide the options in different markets into different time-to-expiry. As mentioned in Section 1.3 of Chapter 1, we will divide the option data into several categories according to the term to expiration. An option contact can be classified as (i) short-term (< 60 days); (ii) medium term (60-80 days); and (iii) long term (> 180 days). We will see which model does best with different maturity dates in different markets.

Table 5.2, table 5.3 and table 5.4 reports the pricing performance (RMSEs) of out-of-sample predictions of all models: ARCH, GARCH, TGARCH, PBS and PBS with smearing. In each table reports the out of sample results. Table 5.2 summarizes the pricing performance of all models with all options. Table 5.3 reports the pricing performance of all models with different markets: Bear Market (2002), Calm Market (2005) and Crash Market (2008). Table 5.4 shows the pricing performance of all models of three markets with three maturity range: short term options, medium term options and long term options.

From Table 5.2, we see that when all options are used, the best out-of-sample fitting accuracy for the call option is the ARCH model, and the best out-of-sample fitting accuracy for the put option is the GARCH model.

From Table 5.3, we see that when we consider the three different markets separately, the results are different. In the bear market, the ARCH Monte Carlo simulation model predict both the European call option and put option prices more precisely than other models. When it comes to the calm years, the best fitting accuracy for European call is the ARCH Monte Carlo simulation model; while the best model for predicting the European put options is the PBS with the smearing. In the crash year, we find that the benchmark PBS and PBS with smearing model predict both the European call option and put option prices more precisely than other models.

In Table 5.4, we look at the pricing performance for the three markets separately by maturity ranges: short term options, medium term options and long term options. Again we reach different conclusions. In the bear market, different maturity ranges give us the different results. For the put options, the best out-of-sample prediction for both short term and medium term options is the ARCH model. But the GARCH model does best for the long term options. For the call options, the GARCH performs best for both short term and long term options.In the calm market, the ARCH Monte Carlo simulation model gives the lowest RMSE in the out-of-sample test for all maturity ranges for put options, while for call options, the best prediction is from the benchmark PBS model. When it comes to the crash years, the results differ from other markets. For put options, the best model to predict short term option price is the PBS smearing model, the bestmodel to predict medium term option price is the GARCH model, andthe bestmodel to predict long term option pricesis the TGARCH model. For the call options, the best model to predict short term option price is the PBS of smearing, the bestmodel to predict medium term option price is the GARCH model, andthe bestmodel to predict long term option pricesis the TGARCH Monte Carlo simulation model.

	Put	Call
ARCH	1.5568	*0.7083
GARCH	*1.1578	1.0831
TGARCH	1.8462	1.3274
PBS	1.7980	1.3977
PBS_Smearing	1.8232	1.3843

Table 5.2 Pricing Performance (RMSE) for the whole data set

	Bear Market		**Calm Market**		**Crash Market**	
	Put	Call	Put	Call	Put	Call
ARCH	*0.3822	*0.5904	*0.3782	0.6150	1.3196	3.0901
GARCH	2.1375	1.0819	2.2344	1.4948	1.8464	2.0098
TGARCH	3.2259	2.2759	2.5777	1.6055	2.1263	2.6795
PBS	1.2459	1.2183	1.4837	0.4980	*0.5441	1.7355
PBS_Smearing	1.2160	1.1758	1.3238	*0.4477	0.6042	*1.6359

Table 5.3 Pricing Performance (RMSE) of all models with different markets: Bear Market (2002), Calm Market (2005) and Crash Market (2008)

	Bear Market		**Calm Market**		**Crash Market**	
Short term options: < 60 days						
	Put	Call	Put	Call	Put	Call
ARCH	*0.5087	1.1497	*0.5828	0.5718	1.3301	2.5886
GARCH	1.7924	*1.1496	1.3720	0.7482	1.4768	2.0733
TGARCH	2.5332	1.9545	1.5652	0.9001	1.6294	2.5151
PBS	0.9186	1.2070	1.5183	0.5685	*0.8020	1.6347
PBS_Smearing	0.9250	1.2244	1.4241	*0.4661	0.8640	*1.5274
Medium term options: 60~180 days						
	Put	Call	Put	Call	Put	Call
ARCH	*0.2841	*0.5188	*0.2379	0.4627	0.6937	2.4210
GARCH	1.1056	1.0121	2.7927	1.2358	*0.4436	*1.0311
TGARCH	1.7569	2.1447	3.2354	1.6603	0.4487	1.4285
PBS	1.1279	1.2282	1.5046	0.5131	0.4732	1.8349
PBS_Smearing	1.0928	1.0924	1.1044	*0.4005	0.4824	1.6512
Long term options: > 180 days						
	Put	Call	Put	Call	Put	Call
ARCH	0.2632	1.0397	*0.2782	0.5196	0.6570	1.2977
GARCH	*0.2110	*0.4051	1.9446	1.3472	0.3749	0.5556
TGARCH	0.4906	1.7697	2.2327	1.8645	*0.3287	*0.4691
PBS	0.6889	0.8209	0.5826	*0.4007	0.4165	1.6665
PBS_Smearing	0.5729	0.8292	0.4249	0.4443	0.4011	1.3877

Table 5.4 Pricing Performance (RMSE) of all models of three markets with three maturity range: short term options, medium term options and long term options

5.4 Summary

In this chapter, we have attempted to develop option valuation methods that fall completely outside of the Black-Scholes-Merton (BSM) framework. The methods are based on the popular Monte Carlo technique, although a novel feature is that we have incorporated time-varying volatility of the underlying price into the simulations. We see this as essential because it is an undisputed fact that time-varying volatility is a perennial feature of asset price data, and this fact essentially invalidates the assumptions underlying the BSM framework.

The options of interest are 5,535 call options and 7,403 put options, spread over three different years: 2002 (bear market), 2005 (calm market) and 2008 (crash market).

The Monte Carlo valuation methods that we have developed are not difficult to apply. They require a few lines of STATA code. However, computer time is an issue. We could not value all of the options in the data because the required computer time would have been too high. Our sample size compares favourably with Duan et al. (2006) who performed a Monte Carlo study on 1000 (fictional) options.

We applied three different models for the Monte Carlo simulation to predict the option prices: ARCH, GARCH and TGARCH. The prediction criterion we used is the root mean square error (RMSE), and we applied this to assessing predictive performance only out-of-sample. What's more, we compare the predictive performance of Monte Carlo simulation models to the benchmark Practitioners Black Scholes (PBS) model. Since the option data is divided into the three different markets: bear market, calm market, and crash market, different market periods give us the different results. Basically, during the bear market period, ARCH Monte Carlo simulation is the best model for predicting the short and medium term put options and GARCH Monte Carlo simulation is the best one for predicting the long term put options. During the calm period, ARCH Monte Carlo simulation is always the best model for predicting the put options and benchmark PBS model is the best model for pricing the call options. In the crash market, TGARCH Monte Carlo simulation is the best model for predicting the long term European call and put options.

Chapter 6 Conclusion

The purpose of this chapter is to summarise the contents of the thesis, and to provide suggestions for future research.

The thesis is focused on ways of valuing options that completely avoid the well-known and popular Black-Scholes-Merton (BSM) framework. The principal motivation for this departure from convention is that many years of financial econometrics results have established that one of the key assumptions underlying the BSM framework, namely the assumption of constant volatility of the return of the underlying asset, is a false assumption. While some authors have attempted to extend the BSM framework to allow for violations of this assumption, our approaches are intended to be more direct.

In Chapter 2, we surveyed existing option valuation methods, including the BSM framework. The material covered in this chapter is important for a number of reasons. For example, it was demonstrated that the binomial tree method and the Black-Scholes method are equivalent provided the number of steps in the binomial tree is sufficiently large. It was also demonstrated that the binomial tree framework can be used to show that the introduction of a TGARCH effect gives rise to the well-known “smirk” pattern seen frequently in market option price data. This is important because the TGARCH effect plays an important role in later chapters. It also suggests a direction for further research. It would be possible to introduce a TGARCH parameter to the online binomial tree calculator, and this would be useful both for researchers and for option traders.

In Chapter 3, we applied regression models to market option price data. We generate four regression models, including linear regression, quadratic regression, cubic regression and B-spline regression models to find the true determinants of option prices. In addition, we are particularly interested in finding which measure of volatility is best for explaining option prices.As expected, the number of days used in the volatility measure seems to increase with the time to expiry: for short term options, the number of days varies between 30 and 60, while for long term options it is usually 70. However, we also see a tendency for the number of days to vary according to the type of market: even for short-term options, the number of days used in the volatility measure is higher in the crash market than in the calm market or the bear market. This suggests that option traders appear to take account of a longer range of historical data when the market is more volatile.

Having estimated the regression models and identified which measure of volatility is most useful, we needed to assess the performance of the models. We found that predictive performance was greatly improved by using the Box-Cox transformation. The Box-Cox B-spline regression model was found to be the best model.

As a benchmark model, we have extensively used the Practitioner Black-Scholes model (PBS). We extended the standard PBS model in two ways. Firstly, we made the estimation of the implied volatility equation more flexible by using the B-spline. Secondly we corrected the bias of standard PBS using a smearing estimator.

We found that, in practice, the application of smearing does not have a major effect on out-of-sample predictive performance.

It is interesting that using the B-spline approach to estimate the IV equation seems to result in the problem of over-fitting, but when the predictions of IV from these over-fitted equations are used to predict option prices (i.e. using the PBS model), the out-of-sample predictions are superior to those from simpler models. This perhaps indicates that we should not be concerned with the problem of overfitting when estimating IV equations.

In Chapter 4, we applied daily data on the S&P 500 Index to estimate the Random Walk model, and the three ARCH family models. We found that the TGARCH model is the best model to represent the volatility process, since all of its coefficients are strongly significant, it has the lowest AIC of the four models, and it passes all the statistical diagnostic tests. Then we simulated data from the models considered using the coefficients from the estimated models. After that, we found that the simulated ARCH family volatility models worked correctly, since the “true” parameter values were included in the confidence intervals. Therefore, we were able to use the estimated models as the basis for Monte Carlo simulations to estimate the values of European Options.

In Chapter 5, we applied the Monte Carlo method to the problem of valuing options, using the models estimated in Chapter 4. This was done using the simulate command in STATA. We applied three different models for the Monte Carlo simulation to predict the option prices: ARCH, GARCH and TGARCH. The prediction criterion we used is the root mean square error (RMSE). Once again, we compared the predictive performance of Monte Carlo simulation models to the benchmark Practitioners Black Scholes (PBS) model. As usual, the best-performing model varied according to market period and time to expiry of options.

<u>A specific suggestion for further research</u>

In this thesis, we have only considered European options. How might our methods be taken further? For example, could the methods be applied to the valuation of American options?

The purchaser of an American call (or put) option has the right but not the obligation to buy (or sell) the underlying asset *at any time between purchase of the option and its expiry*. The key difference from European options is that American options can be exercised at any time, while European options can only be exercised at expiry. This difference makes American options more difficult to analyse.

For standard American call options without dividends, there are reasons why the option should never be exercised before the expiry date (see Hull, 2011). For this reason, we will restrict attention to American put options.

The important decision for the holder of an American put option is: when is it rational to exercise the option. If the underlying price falls far below the strike price, it might be rational to exercise the option now and realise a sure profit, instead of holding on to the option and risking an increase in the price of the underlying. This reasoning gives rise to the concept of an "early exercise boundary". Whenever the price of the underlying falls below the "early exercise boundary", it is rational to exercise the American option. One significant problem is how to find where the early exercise boundary is.

Our suggestion for future research is to apply the Monte Carlo method to the problem of valuing American options. This way, the assumptions about varying volatility (e.g. ARCH, GARCH, and TGARCH) considered in chapters 4 and 5 could be incorporated.

There are four stages to the process.

Stage 1: Choose a model from Random Walk, ARCH, GARCH, and TGARCH. Simulate some time-paths of the underlying, from settlement to expiry, using the chosen model. For each time path, locate the position of the minimum point. Treat this point as the best early exercise point for the time path. Figure 6.1 shows 4 simulations from a TGARCH model, assuming 100 days to expiry. For each simulated path, a large red dot is used to indicate the lowest point.

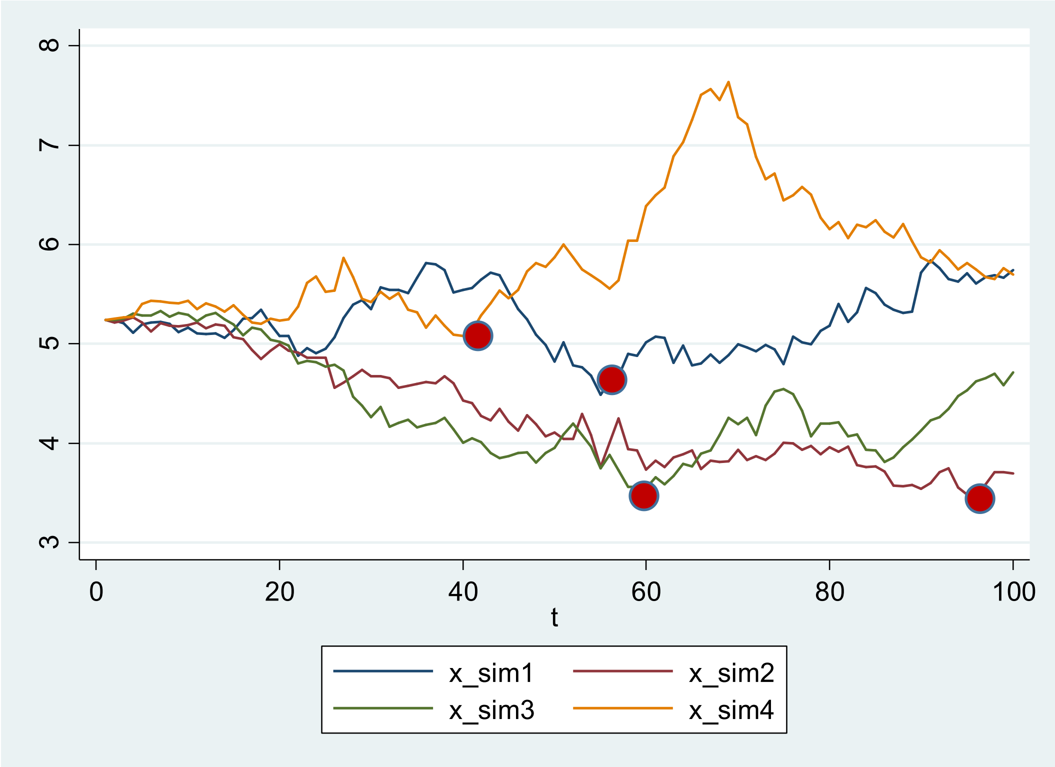


Figure 6.1 Four simulations of TGARCH models

Stage 2: Carry out a large number of simulations, and collect all of the early exercise points together on the same graph. Figure 6.2 shows such a graph obtained from 100 simulated paths. Use a non-parametric regression to fit a curve through these points. This curve will be the early exercise boundary. A Lowess smoother has been used in Figure 6.2 to obtain the early exercise boundary.

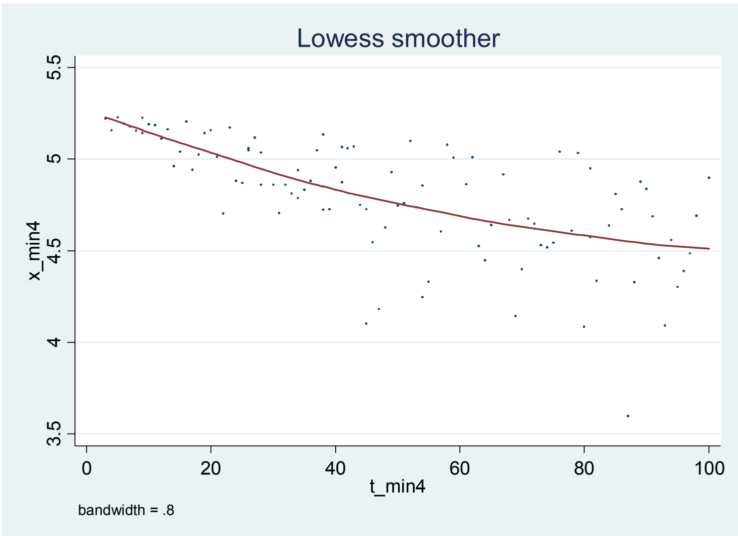


Figure 6.2 Early exercise boundary obtained from 100 simulated paths of the TGARCH model

Stage 3: Carry out a second set of simulated paths using the same model. Whenever the simulated path crosses the early exercise boundary, exercise the option early.

Stage 4: for each simulated path in stage 3, discount the actual payoffs by the time between settlement and exercise. Find the average discounted payoff. This is the estimated value of the option.

It is interesting to compare the early exercise boundaries obtained in Stage 2 from the four different models. Figure 6.3 shows this comparison. The Random Walk and ARCH seem to lead to the lowest early exercise boundary. GARCH gives the highest early exercise boundary. TGARCH gives the steepest early exercise boundary.

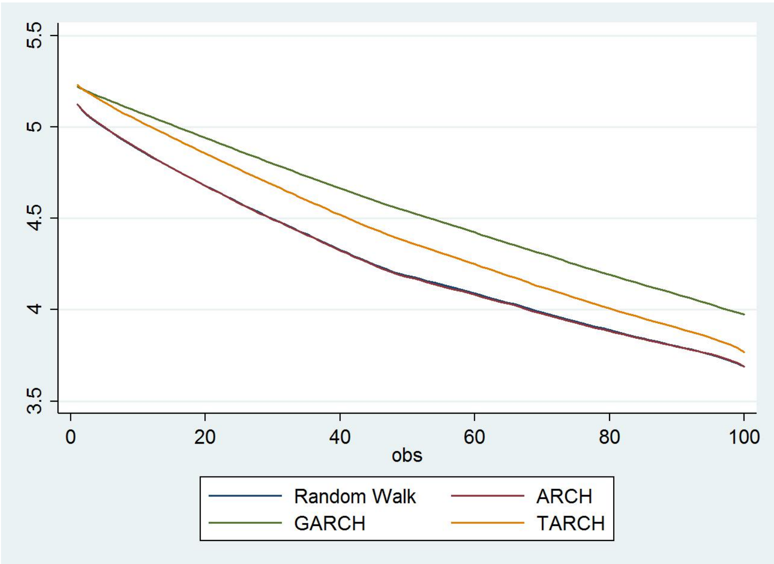


Figure 6.3 Early exercise boundaries from the four different models.

More research is needed to confirm that this pattern is a general result. This approach may be very useful in the valuation of American options.

References

Acworth Will, 2015, Annual Survey: Global Derivatives Volume. Magazine of the Global Futures, *Options and Cleared Swaps Markets*.

Adams A.T., Booth P.M., Bowie D.C., Freeth D.S., 2003, Investment Mathematics, *Wiley*.

Aït-Sahalia, Y., & Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *The Journal of Finance*, *53*(2), 499-547.

Ait-Sahalia, Y., Duarte, J., 2003, Nonparametric option pricing under shape restrictions. *Journal of Econometrics*, 116, 9-47.

Alberg, Dima, Haim Shalit, and Rami Yosef. "Estimating stock market volatility using asymmetric GARCH models." *Applied Financial Economics* 18.15 (2008): 1201-1208.

Andreou, P. C., Charalambous, C., & Martzoukos, S. H. (2008). Pricing and trading European options by combining artificial neural networks and parametric models with implied parameters. *European Journal of Operational Research*, *185*(3), 1415-1433.

Andreou, P. C., Charalambous, C., & Martzoukos, S. H., 2014, Assessing the performance of symmetric and asymmetric implied volatility functions. *Review of Quantitative Finance and Accounting*, *42*(3), 373-397.

Balaban, Ercan. "Comparative forecasting performance of symmetric and asymmetric conditional volatility models of an exchange rate." *Economics Letters* 83.1 (2004): 99-105.

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *The Journal of Finance*, *52*(5), 2003-2049.

Bates, D. S. (1991). The Crash of ’87: Was It Expected? The Evidence from Options Markets. *The journal of finance*, *46*(3), 1009-1044.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, *9*(1), 69-107.

Barraquand, J. and Martineau, D., 1995. Numerical valuation of high dimensional multivariate American securities. *Journal of financial and quantitative analysis*, *30*(03), pp.383-405.

Barone-Adesi., and R.Whaley, 1987, Efficient Analytical Approximation of American option Values.*Journal of Finance*, 42, 301-320.

Berkowitz, J. (2010). On justifications for the ad hoc Black-Scholes method of option pricing. *Studies in Nonlinear Dynamics and Econometrics*, *14*(1), 1-27.

Black, F., Scholes, M., 1973, The pricing of options and corporate liabilities. *Journal of Political Economy*,81,637-654.

Bollerslev, T., 1986, Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 31, 307-327.

Bollerslev T., 2009,*Glossary to ARCH (GARCH).In Volatility and Time Series Econometrics: Essays in Honour of Robert F. Engle, Bollerslev T, Russell JR, Watson M (eds)*. Oxford University Press: Oxford, UK; 137–164.

Bollerslev, Tim, and Jonathan H. Wright. "Semiparametric estimation of long-memory volatility dependencies: The role of high-frequency data." *Journal of econometrics* 98.1 (2000): 81-106.

Breen, R., 1991, The Accelerated Binomial Option Pricing Model. *The Journal of Financial and Quantitative Analysis*, 26(2),153-164.

Brennan M., Schwartz E., 1982, An Equilibrium model of Bond Pricing and a Test of Market Efficiency. *Journal of Financial and Quantitative Analysis*,17, 301-29.

Broadie, M., & Detemple, J., 1996, American option valuation: new bounds, approximations, and a comparison of existing methods. *Review of Financial Studies*, 9(4), 1211-1250.

Brooks, C., 2014. *Introductory econometrics for finance*. Cambridge university press.

Broyle, P., Broadie, M., & Glasserman, P., 1997, Monte Carlo methods for security pricing. *Journal of economic dynamics and control*, *21*(8), 1267-1321.

Bunch D., Johnson H., 1992, A Simple and Numerically Efficient Valuation Method for American Puts Using a Modified Geske-Johnson Approach.*Journal of Finance*, 47, 809–816.

Carr, Peter., 1998,Randomization and the American put.*Review of Financial Studies*, 11(3), 597-626.

Carr, Peter, Robert Jarrow, and Ravi Myneni., 1992,Alternative characterizations of American put options.*Mathematical Finance*, 2(2), 87-106.

Chang, C., Chuang, San-Lin., and Stapleton Richard C., 2001, Richardson extrapolation techniques for pricing American-style options. *Working paper*.

Chang, C.C., Chung, S.L, & Stapleton, R., 2001,Richardson Extrapolation Techniques for Pricing American-style Options.*Working Paper*.

Charnes, John M., 2000, Using Simulation for option pricing. *Working paper*.

Chen, N. F., & Johnson, H. (1985). Hedging options. *Journal of Financial Economics*, *14*(2), 317-321.

Christoffersen, P., & Jacobs, K. (2004). The importance of the loss function in option valuation. *Journal of Financial Economics*, *72*(2), 291-318.

Christoffersen, Peter, Mathieu Fournier, and Kris Jacobs. "The factor structure in equity options." *The Review of Financial Studies* (2017): hhx089.

Christoffersen, P., Heston, S., &Jacobs, K., 2009, The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well. *Management Science*, *55*, 1914-1932.

Cleveland, William S., 1979, Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association*, 74, 829-836.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223-236.

Cont, Rama, Marc Potters, and Jean-Philippe Bouchaud. "Scaling in stock market data: stable laws and beyond." *Scale invariance and beyond* 7 (1997): 75-85.

Corrado, C. J., & Su, T. (1996). Skewness and kurtosis in S&P 500 index returns implied by option prices. *Journal of Financial research*, *19*(2), 175-192.

Cox, J. C., Ross, S. A., & Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of financial Economics*, *7*(3), 229-263.

De Boor C., 2001, *Practical Guide to Splines*. Berlin: Springer Verlag.

Duan, J., 1995, The GARCH option pricing model. *Mathematical Finance*, 5, 13-32.

Duan, N. (1983). Smearing estimate: a nonparametric retransformation method. *Journal of the American Statistical Association*, *78*(383), 605-610.

Duan, J., Gauthier, G., Simonato, J., & Sasseville, C. (2006). Approximating the GJR-GARCH and EGARCH option pricing models analytically. *Journal of Computational Finance*, *9*(3), 41.

Dumas, B., Fleming, J., & Whaley, R. E. (1998). Implied volatility functions: Empirical tests. *The Journal of Finance*, *53*(6), 2059-2106.

Enders, W. (2004). *Applied Econometric Time Series*, 2nd Edition. Wiley Series in Probability and Statistics.

Engle, R., 1982, Autoregressive conditional Heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 50, 987-1008.

Engle, R., 2001. GARCH 101: The use of ARCH/GARCH models in applied econometrics. *The Journal of Economic Perspectives*, *15*(4), pp.157-168.

Engle, R. F., & Mustafa, C., 1992, Implied ARCH models from options prices. *Journal of Econometrics*, *52*(1), 289-311.

Engle, Robert F., Barone-Adesi, G., and Mancini L., 2008, A GARCH option pricing model with filtered Historical Simulation. *The Review of Financial Studies,* 21.

Fink, J., & Fink, K. (2005). Monte Carlo Simulation for Advanced Option Pricing: A Simplifying Tool.

Frey R., 1997, Derivative Asset Analysis in Models with Level-Dependent and Stochastic Volatility. *Discussing Paper*.

Geske, Robert, and Herb Johnson, 1994, The American put option valued analytically. *Journal of financial engineering,*53.

Glasserman, P., 2003, Monte Carlo methods in financial engineering. *Springer Science & Business Media*.

Glosten, L.R., Jagannathan, R., and Runkle, D.E., 1993, On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48, 1779-1801.

Global Derivatives, 2010, http://www.global-derivatives.com/index.php/pricing-models-othermenu-41/14?task=view

Gujarati, D.N. and Porter, D., 2009. *Basic Econometrics* Mc Graw-Hill International Edition.

Han H. and X. Wu, 2003, A fast numerical method for the Black-Scholes equation of American options.*SIAM J. Numer. Anal*., 41, 2081-2095.

Handscomb, D.C., 1964, A Monte Carlo method applied to the Heisenberg ferromagnet. *Mathematical Proceedings of Cambridge Philosophical Society*, 60, 115-122.

Hastie T.J. and Tibshirani R.J., 1990, Generalised Additive Models, Chapman and Hall.

Heston, S., 1993, A closed-form solution for options with stochastic volatility. *Review of Financial Studies*, 6, 327-344.

Heston, S., Nandi, S., 1997, A closed form GARCH option pricing model. *The Review of Financial Studies,* 13, 585-625.

Heston, S. L., & Nandi, S. (2000). A closed-form GARCH option valuation model. *The Review of Financial Studies*, *13*(3), 585-625.

Huang J., Subrahmanyam M., Yu G., 1996, Pricing and Hedging American Options: A Recursive Integration Method.*Review of Financial Studies*, 9, 277–300

Hull, J.C, and A. White, 1987, The Pricing of Options on Assets with Stochastic Volatilities.*Journal of Finance*, 42, 281-300.

Hull, J.C, 2011, *Fundamentals of Futures and Options Markets*. Seventh edition. Pearson.

Hull, J., & Suo, W. (2002). A methodology for assessing model risk and its application to the implied volatility function model. *Journal of Financial and Quantitative Analysis*, *37*(2), 297-318.

Hutchinson, J.M., Lo, A.W., and Poggio T., 1994, A nonparametric approach to pricing and Hedging derivative securities via learning networks. *Journal of Finance*, 49, 851-889.

Jacka, 1991, SD 1. Optimal stopping and the American put.*Mathematical Finance*, 1(2),1-14.

Ju, N. (1998). Pricing by American option by approximating its early exercise boundary as a multipiece exponential function. *The Review of Financial Studies*, *11*(3), 627-646.

Kim I. J., 1990, The Analytical Valuation of American Options.*Review of Financial Studies*, 3, 547–572.

Linaras, Charilaos E., and George Skiadopoulos. "Implied volatility trees and pricing performance: evidence from the S&P 100 options." *International Journal of Theoretical and Applied Finance* 8.08 (2005): 1085-1106.

Longstaff, F., Schwartz, E.S., 2001, Valuing American options by simulations: A simple least-squares approach. *Review of Financial Studies*, 14, 113-147.

Loudon, Geoffrey F., Wing H. Watt, and Pradeep K. Yadav. "An empirical analysis of alternative parametric ARCH models." *Journal of Applied Econometrics* (2000): 117-136.

Mandelbrot, Benoit B. "When can price be arbitraged efficiently? A limit to the validity of the random walk and martingale models." *The Review of Economics and Statistics* (1971): 225-236.

Mandelbrot, Benoit. "The variation of certain speculative prices." *The journal of business* 36.4 (1963): 394-419.

MacMillan, L.W., 1986, An Analytical Approximation for American Put Prices. *Advances in Futures and Options Research*, 1, 119-139.

Mammen, B.E., 1991, Estimating a smooth monotone regression function. *The Annals of Statistics*, 19, 724-740.

Nandi, S. (1996). *Pricing and hedging index options under stochastic volatility: an empirical examination* (No. 96-9). Working Paper, Federal Reserve Bank of Atlanta.

Nelson,D.B., 1991, Conditional heteroscedasticity in asset return: A new approach. *The Review of Financial Studies*, 3, 627-646, 347-370.

Pandher, G. S., 2007. Regression-based modeling of market option prices: with application to S&P500 options. *Journal of Forecasting*, *26*(7), 475-496.

Rubinstein, M., 1994. Implied binomial trees. *The Journal of Finance*, 49(3), 771-818.

Sanderson, Michael J.,2002, Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach.*Molecular biology and evolution*, 19(1),101-109.

Schwartz, E.S., 1977, The valuation of Warrants: Implementing a New Approach. *Journal of Financial Economics*, 4, 79-93.

Scott, L.O., 1987, Option Pricing When the Variance Changes Randomly: Theory Estimation and an Application. *Journal of Financial and Quantitative Analysis*, 22, 419-438.

Stein, E.M., Stein, J. C., 1991, Stock Price Distributions with Stochastic Volatility: An Analytic Approach. *Review of Financial Studies*, 4, 727-752;

Stone, C.J., 1985, Additive regression and other nonparametric models. *The Annals of Statistics*, 13(2), 689–705.

Taleb, N. N., 2008, The black swan: The impact of the highly improbable. Penguin, London, England.

Taleb, N. N., Goldstein, D. G., and Spitznagel, M. W., 2009, The six mistakes executives make in risk management. *Harvard Business Review*, 87(10):78-81.

Tilley, JA, 1993, Valuing American options in a path simulation model. *Transactions of the Society of Actuaries*, 45(83),104.

Whaley, R. E. (1982). Valuation of American call options on dividend-paying stocks: Empirical tests. *Journal of Financial Economics*, *10*(1), 29-58.

Wiggins, J.B., 1987, Option Values under Stochastic Volatilities.*Journal of Financial Economics*, 19, 351-372.

Wu L.X. and Kwok, Y.K., 1997, A front-fixing finite difference method for the valuation of American options. *Journal of Financial Engineering*, 6,83-97.

Zakoian, J.M., 1993, Threshold heteroskedastic Models. *Journal of Economic Dynamics and Control*, 65, 1-29.