Option Pricing and Bayesian Learning

Jönsson, Ola

2007

Citation for published version (APA):
Jönsson, O. (2007). *Option Pricing and Bayesian Learning*. [Doctoral Thesis (monograph), Department of Economics]. Department of Economics, Lund University.

Total number of authors:
1

Option Pricing and Bayesian Learning

Ola Jönsson

Acknowledgements

First of all, I would like to thank my advisor Anders Borglin, whose commitment and dedication, along with his impressive knowledge, has been truly inspiring. His clarity of thought helped me overcome obstacles that at times caused me to despair; problems that without the his help very well could have proven to be insurmountable. Never dodging difficulties, but rather confronting them head on he displays a trait common with my favorite boxers: grace under pressure. For his friendliness and invaluable help I owe him greatly. I also owe him uncountable cups of coffe.

I would also like to extend my thanks to Allan Timmermann whose work inspired my research and was helpful in arranging my stay at UCSD during the fall of 2004 and Rolf Poulsen who gave helpful comments at my final seminar. Erik Lindström also generously commented on a draft.

While Economics has been described as "the dismal science", my fellow colleagues at the Department of Economics possess decidedly good-natured dispositions. The friendly atmosphere provided by seniors such as Björn Hansson, Hossein Ashgarian, Bo Larsson and David Edgerton has been greatly appreciated. Others who have contributed in making the Department an enjoyable workplace are Birger Nilsson, Viktor Tanaka, Per Hjertstrand, Joakim Westerlund, Thomas Elger, Kristian Jönsson, and Tommy Andersson. During my early time at the Department of Economics I was fortunate to share office with Sonnie Karlsson, Christoffer Bengtsson, Sirly Palenmark, Ulf Erlandsson and Fredrik Hansen. Sharing office space with Wolfgang Hess as I have done during the last, stressful months of trying to complete my thesis, has been great fun. There is nothing like a good laugh to get your spirits up. Apparently his cheerful nature comes from his method of relieving his anxiety with micro-wave

cooking (is it a Bavarian thing?). During my stay at UCSD it was a pleasure to share office with Jennifer Poole.

Lunches at Gladimat and barbeques in my garden with my friends Peter Gustafsson, Hans Byström and Dag Rydorff has been a recurring joy. We always had interesting discussions and they have all generously helped out whenever I was puzzled by something. Peter always had an answer to all things related to Matlab and statistics and I owe him greatly. Lunches at Finn In with Kent Johansson and his gang of economic historians has also been great fun.

Life is not only about Economics - it's about boxing to. At the local boxing gym I also met people who became good friends: Fredrik Hermann, the twins Martin and Mattias Ekstrand, Johan Navarro and Martin Christensson. They have all contributed to making the spare time enjoyable, as have the parties at Casa de Antonio Maranon.

Lund, December 2006
Ola Jönsson

Contents

Introduction and Summary

In 1973 Fischer Black and Myron Scholes published their seminal paper on option valuation. Although more than three decades has passed the Black-Scholes option pricing formula remains a cornerstone in modern finance. With its aide it became possible to price financial assets giving the holder the right but not the obligation to buy (call option) or sell (put option) an underlying asset-say, a stock - in the future, at a prespecified price. Their work gave rise to what is sometimes called financial engineering. While traditional economics attempts to price assets by modelling individuals preferences with utility functions and is concerned with equilibrium outcomes, financial economists working in the vein of Black and Scholes rely on the concept of arbitrage-free pricing, relating the price of an asset to the prices of other assets so that no risk free profits can be made. Thus, assuming that markets do not allow arbitrage opportunities, the price of an asset such as a call option can be found by forming a replicating portfolio giving the same payoff in all states of the world. A key feature of this approach is that one is able to price derivatives such as call options on stocks without specification of individuals risk aversion, implying that any two individuals will agree on the price regardless of their attitude towards risk. Although this may seem remarkable, the answer lies in the fact that the risk premium is reflected in the price of the stock. Thus, whereas in traditional economics the price of a stock is derived in an equilibrium setting, the approach by Black and Scholes was to take the equilibrium price of the stock as given.

One assumption underlying the Black and Scholes option pricing formula is that the price of the underlying asset is lognormally distributed, so that the returns are normally distributed. Furthermore, they assumed

that the volatility of the returns is constant over time. These assumptions have been tested empirically and found wanting in realism. For example, stock returns have been shown to display fatter tails than implied by the normal distribution, meaning that very low and very high returns are more likely than the assumption of normally distributed returns allows. A related observation is that the volatility implied by the observed option price differs among options with different exercise price, so that often a smile shaped pattern occurs when the implied volatilities are plotted against the exercise prices. Implied volatilities also differ for options having different time to expiration. Consequently, much research has been devoted to amend these shortcomings, the majority of which has operated within the financial engineering framework, such as stochastic volatility models and jump-diffusion models. Such models can be regarded as piecemeal improvements of the original Black and Scholes theory as they operate within the same basic framework. While Black and Scholes formulated their theories in a continuous-time setting using relatively advanced mathematics, in 1979 Cox, Ross and Rubinstein published a paper showing that the Black and Scholes option pricing formula could be derived within a discrete-time setting. This made the intuition behind the theory more accessible as it only required mathematical tools that are standard in economic theory.

The aim of this thesis is to relate the observed pattern of implied volatilities to the behavior of individuals using an equilibrium, discrete-time setting.

Summary of Chapter 1

Chapter 1 analyses daily data on put and call options covering the period 1993 to 2000, on the Swedish OMX index, an index comprising the 30 stocks which have the largest volume of trading on the Stockholm Stock Exchange. The options are European options, meaning that they can only be exercised at the expiration date. For both calls and puts the implied volatilities are extracted using the aforementioned discrete time option pricing formula due to Cox, Ross and Rubinstein. These implied volatilities were then related to time to expiration and moneyness. The

implied volatilities varied substantially over the eight year period. When the implied volatilities were related to time to expiration the results were as follows. Throughout the sample period it was found that both calls and puts exhibit the largest implied volatility for short maturities and that the implied volatility tended to be decreasing in time to expiration. In general the volatilities implied by puts exceeded those for calls, for all maturities. In order to measure how much the implied volatilities varies for a given time to expiration the notion of strength of smile is introduced. For both calls and puts the strength of smile is decreasing in time to expiration, implying that options with short maturity are the ones most mispriced by the (discrete time) Black and Scholes formula. The strength of smile for puts was in general much higher than for calls. When the same analysis was performed with regard to moneyness, it was seen that the implied volatilities were the highest for calls with extreme values of moneyness. This pattern was duplicated for puts, although the implied volatilities for puts in general were higher. For both calls and puts, the strength of smile for a given interval of moneyness was also the highest for extreme values of moneyness. On average, the implied volatilities for puts clearly exceeded the volatilities implied by calls. Chapter One also contains a test of the no-arbitrage condition known as Put-Call Parity, showing that it holds well. For those options eligible to use in the test, the difference in implied volatilities was significantly smaller, indicating that the arbitrage powers underlying Put-Call Parity indeed is in effect.

Summary of Chapter 2

Chapter 2 presents the theory of how Bayesian learning affects asset pricing. The model is due to Timmermann and Guidolin [2003] and applies Lucas [1978] representative agent, equilibrium model. Thus, the point of departure is an economy with a bond in zero net supply and a stock paying an infinite stream of dividends evolving in a binomial tree. The solution of the consumers problem provides us with state prices with which any payoff can be priced. After having equipped the representative agent with power utility, specification of how the probabilities of up and down movements in the tree evolve over time distinguishes the state prices of

4

the Full Information model from those of the model with Bayesian learning. As a pedagogical tool the urn model due to Polya is introduced. The Polya distribution can be used to model Bayesian learning and also encompasses as a special case the binomial distribution which applies to the Full Information model. It is shown that Bayesian learning implies fatter tails of the probability distribution as compared to the binomial distribution. In the beta-binomial representation of the Polya distribution, the parameters reflect the Bayesian learners beliefs. Thus equipped with state prices the bond, stock and call prices are derived for the Full Information model and the model with Bayesian learning, reproducing the results of Timmermann and Guidolin. The bond prices with Bayesian learning implies that the present value of future income becomes path dependent, making computation of the risk neutral probabilities more complicated. A key feature of the stock price with learning is that the stock-dividend ratio is no longer constant, as it is for the Full Information case. Compared to the Full Information case, learning affects the call price both by widen the support and by increasing the state prices with which the payoffs at exercise date are valued. The contribution of this chapter is a derivation of closed form approximations for the stock and call price in the case of Bayesian learning, improving analytical and computational tractability of the model. Furthermore, by normalizing the model, the strength of learning is captured in one single parameter. Also, a precise description of the relationship between the call price of the Full Information model and the call price of the (discrete time) Black and Scholes model is given, presenting an equilibrium rationalization of the Black and Scholes model.

Summary of Chapter 3

In Chapter 3 the closed form approximation of the Bayesian learning call price formula is applied to the OMX data analyzed in Chapter 1 We also apply the Black and Scholes model, the Full Information model and a Spline model, where the volatility input to the Black and Scholes formula is modeled by a polynomial in the exercise price and time to expiration. Except for the Spline model, all models are also estimated allowing the

parameters to depend on moneyness and time to maturity, respectively. Accordingly, we have three classes of models: unstratified models and models stratified according to moneyness and time to expiration, respectively. The estimation of the Full Information model is performed using its relationship with the Black and Scholes model described in Chapter Two. In-sample results show that the model with Bayesian learning is superior to all its unstratified competitors in its ability to match the data, displaying an average in-sample error only 2/3 that of the Black and Scholes model, and is also the best model in 2/3 of the dates. It is followed in both regards by the Spline model. Out-of-sample predictions were performed using the parameter estimates obtained from the in-sample exercise. Both one-day ahead and one-week ahead predictions were performed. In both cases, the model with Bayesian learning stratified according to time to expiration scored the highest when the percentage of dates it outperformed its competitors was considered. It also displayed the lowest prediction error for one-day ahead predictions, closely followed by the unstratified model with Bayesian learning. For one-week ahead predictions, the Full Information model stratified according to moneyness displayed the lowest prediction error, closely followed by the unstratified Full Information model. For both prediction lengths the Spline model performed the worst in terms of prediction error, raising the issue of parameter stability. Consistent with the findings by Timmermamm and Guidolin, it was found that the model with Bayesian learning was the best of the unstratified models when out-of-the-money contracts were considered.

6

Chapter 1

The OMX Index and Derivatives

Introduction

The Black-Scholes option pricing model assumes that the underlying asset follows a geometric brownian motion with constant volatility, implying that all options on the same asset should provide the same implied volatility. However, it has been well documented that the implied volatilities tend to differ both across remaining time to expiration and exercise prices, resulting in "smiles" or "smirks" when plotted against these variables. For ease of reference we will refer to any such distribution of implied volatilities over the exercise price or remaining time to expiration as a smile. By the strength of the smile we denote the volatility of the implied volatilities. This notion is meant to measure to which degree the assumption of constant volatility is violated.

Lando and Poulsen [2006] provide an economic explanation of the existence of these volatility smiles in terms of portfolio insurance and crash fears: managers of portfolios mimicking the market stock index wanting to insure against a possible crash will want to purchase out-of-the-money put options, causing their prices to rise, and hence their implied volatilities.

The purpose of this chapter is to investigate whether the pricing of options on the Swedish OMX index causes the implied volatility to be a

7

non-constant function of the exercise price and the remaining time to expiration. Furthermore, we will determine how the strength of the smiles vary over time and how it depends on the exercise price and remaining time to expiration. This is done for calls and puts separately. In extracting the implied volatility the discrete time Black and Scholes model due to Cox, Ross and Rubinstein [1979] is used. Much research has been dedicated to analyze different option markets with regard to the patterns of implied volatilities; often cited papers include Edgerington and Guan [2002] and Das and Sundaram[1999].

When pricing an option one has to take into consideration whether the underlying asset pays dividends during the life of the option. For an option on an index such as the OMX index which is not adjusted for dividends one thus has to estimate the dividend yield during the life of the option. However, when the dividend yield was computed using Spot-Futures Parity the result was ambiguous. Although for some years the dividend yield clearly peaked during the dividend period, this pattern was not consistent throughout the sample. Furthermore, the dividend yield often assumed negative values. For these reasons the options affected by dividends were discarded.

The arbitrage relation known as Put-Call Parity was tested by computing the relative difference between the call and a portfolio constructed by borrowing the present value of the exercise price, buying a put and the underlying asset.

The outline of the paper is as follows. In section 1.1 we introduce the OMX index and describe the derivatives on the OMX and the data set used. Section 1.2 describes the numerical procedure used in solving the discrete time Black and Scholes model for the implied volatility. The implied volatility as a function of time to expiration is studied in Section 1.3 while section 1.4 investigates the implied volatility as a function of moneyness. Finally, a summary is given in Section 1.5.

Section 1.1 concludes with the results from the investigation of Put-Call Parity are presented and the volatility smiles are examined.

The convergence of the discrete Black and Scholes model to its continuous counterpart is described in Appendix A. In Appendix B the relationship between the spot price of an asset and the future price is

derived. The Put-Call Parity is described in Appendix C.

1.1 About the Data

The Swedish OMX Index

Introduced in 1986 by the Swedish exchange for options and other derivative securities (OM), the Swedish OMX index comprises the 30 stocks which have the largest volume of trading. It was constructed with the objective of creating an index based on a limited number of stocks, which develops in close correlation with the stocks listed on the Stockholm Stock Exchange. The index is reviewed semi-annually. The OMX is a price index, meaning that no cash dividend is reinvested in the index. Hence, the OMX only yields the performance of stock price movements.

Derivatives on the OMX Index

Three types of derivative assets on the OMX index are traded at the Stockholm Stock Exchange. These are futures, call options and put options. Each fourth Friday is an *expiration (delivery) date.* Trade is most frequent in derivatives expiring within the next three months although there is some trade in derivatives with expiration dates up to a year or more. Prices for the derivatives are quoted per unit of the index but trade is in contracts for 100 such units. For the derivatives on the OMX index the deals are settled in cash rather than by delivery of the portfolio underlying the index.

Purchasing a *futures* contract on the OMX index at the current date, t, with delivery date T and price $F_{t,T}$, at date t, one thus commits to pay $100 \times F_{t,T}$ SEK at the delivery date and is, on the other hand entitled to receive $100 \times S_T$ SEK, where S_T is the value of the OMX index at date T. There are no margin accounts but a deposit of 10 % of the futures price, the security amount, is required. Generally, the day following a delivery date a futures contract with delivery date in three months is created, so that each day except for delivery dates there are three futures contracts differing in delivery dates.

The OMX call and put options are European style options, so that exercise prior to the expiration date is not possible.

A *call option* bought at the current date, t, with expiration date T is paid at date t. It gives the owner the right, but not the obligation, to buy the underlying asset at strike price K at the expiration date. Since the underlying asset is here the OMX index there is a cash settlement and the owner receives $\max(S_T - K, 0)$ at the expiration date.

Similarly, a *put option* bought at date t is paid at the same date. It entitles the owner to sell the underlying asset at strike price K at the given expiration date and thus there is a cash settlement and the owner receives $\max(K - S_T, 0)$ at the expiration date.

In order to make the OMX options more accessible, the 27:th of April 1998 there was a split in the OMX index 4:1. This implies that beginning with this date, the number of option- and futures contract are multiplied by 4, while the exercise price and the future price is divided by 4. Thus a holder of one call option with exercise price $K = 600$ prior to the split, when the OMX was reported at, say, 640 would make a profit of

$$100(640 - 600) \quad = \quad 4000$$

since an option comprises 100 contracts. Since after the split the OMX will record at 160, in order to make the post-split position equivalent, the exercise price is divided by 4 and the number of contracts is multiplied by 4, so that the profit can now be written as

$$400(160 - 150) \quad = \quad 4000$$

Dividends and the Spot-Futures Parity

Cash dividends from companies traded on the Stockholm Stock Exchange are concentrated to the months March, April, May and June., with roughly 90% of the dividends occurring in April and May. Hence, the index must be adjusted for dividends only when the March-, April-, May- and June-option prices are analyzed. The standard procedure for doing this is using spot-futures parity in calculating the expected dividend yield during the life of the option. Spot-futures parity is described

in Appendix B

Using all futures with remaining days until delivery between 7 and 90, the average dividend yield was computed for each day in the entire period, including the non-dividend periods for which one would expect the computed dividend yield to be zero. However, the dividend yield was found to take both positive and negative values regardless of the period considered. Furthermore, even though for some years the dividend yield peaks during the dividend period, for some years no such pattern could be found. This is illustrated in Figure 1.1.A, where the average dividend yield is plotted for 1996 and 2000, respectively.

As a consequence of the inconsistent pattern displayed by the dividends, only dividend free July to February option contracts were analyzed.
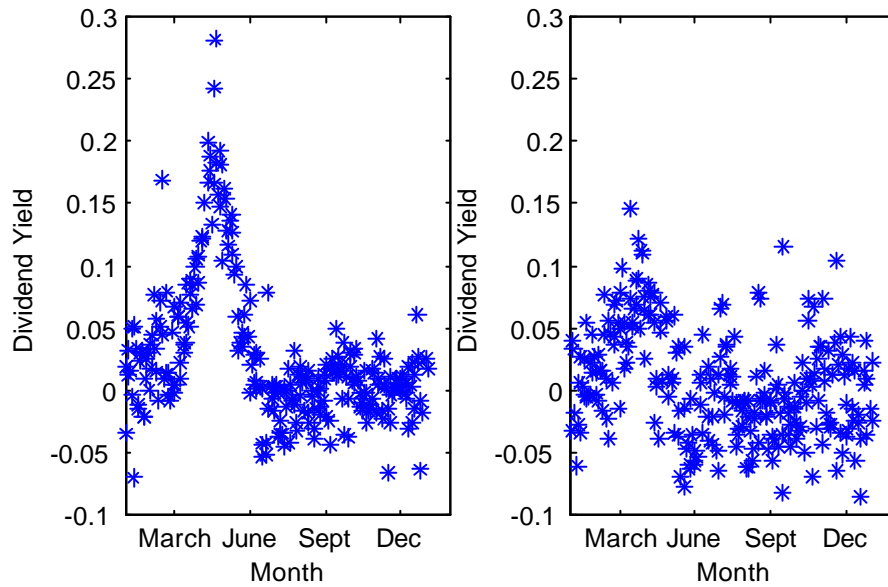
Figure 1.1.A: *Average dividend yields for the years 1996 (left) and 2000 (right)*

The Data Set for OMX and Derivatives

The data set used in this study consists of daily observations of the OMX index and the corresponding call and put option prices covering the period 1993 to 2000. In order to avoid unreliable prices, options with volume less than 20 contracts, strike prices deviating from the current

OMX index by a factor of more than 1.2 or less than 0.8 and options with remaining days until expiration less than 8 and greater than 90 were excluded, resulting in a data set of 31762 calls and 34416 puts. Discarding those options that were affected by dividends further reduced the original dataset to 20215 call options and 21783 put options.

For the sample period 1993 to 2000, Figure 1.1.B shows the distribution of contracts over time to expiration for calls and puts, respectively.
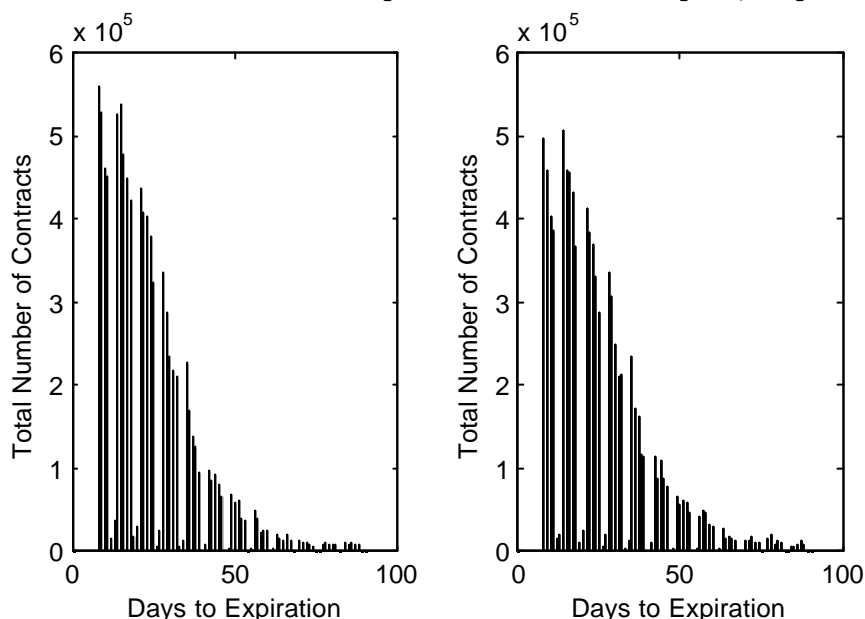
Figure 1.1.B: *The left figure shows that the number of traded contracts for calls is decreasing in time to expiration. The right figure shows the same pattern for puts.*

The distribution of contracts in the sample over moneyness, defined as OMX/K for calls and K/OMX for puts, is shown in Figure 1.1.C, revealing that little trade occurs for options with extreme values of moneyness. The relevant interest rates were attained by linear interpolation of the 1-month and 3-month, Swedish treasury bills. Thus, for an option with less than 31 days to expiration the interest rate on the 1-month Treasury bill was used. For options with $T > 30$ days to expiration the relevant interest rate was computed as

$$r = \frac{r_3 - r_1}{60} (T - 30) + r_1$$

where r_1 and r_3 denotes the 1-month and 3−month interest rates, respectively. Figure 1.1.D shows how the interest and OMX index have evolved during the sample period.
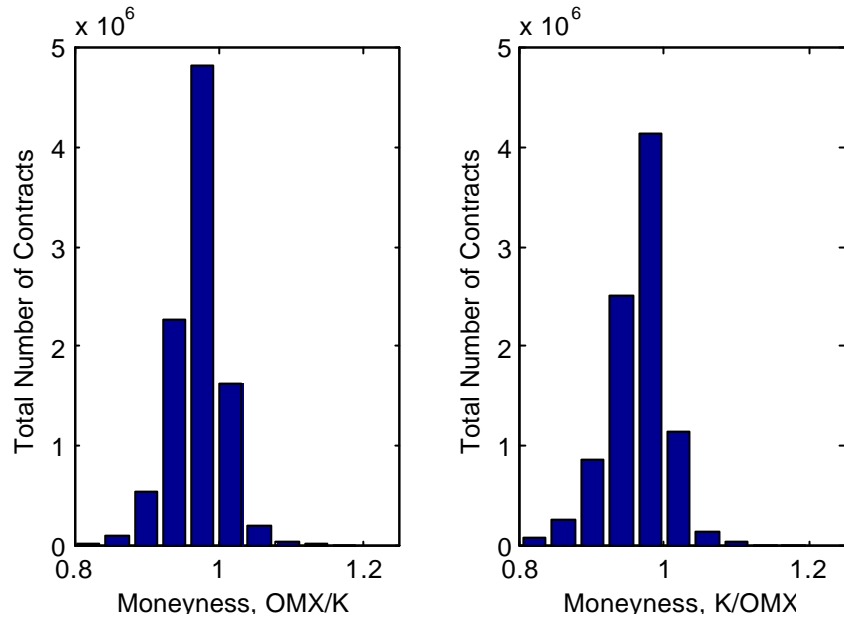
Figure 1.1.C: *The distribution of contracts over moneyness for calls (left) and puts (right) in the sample,*
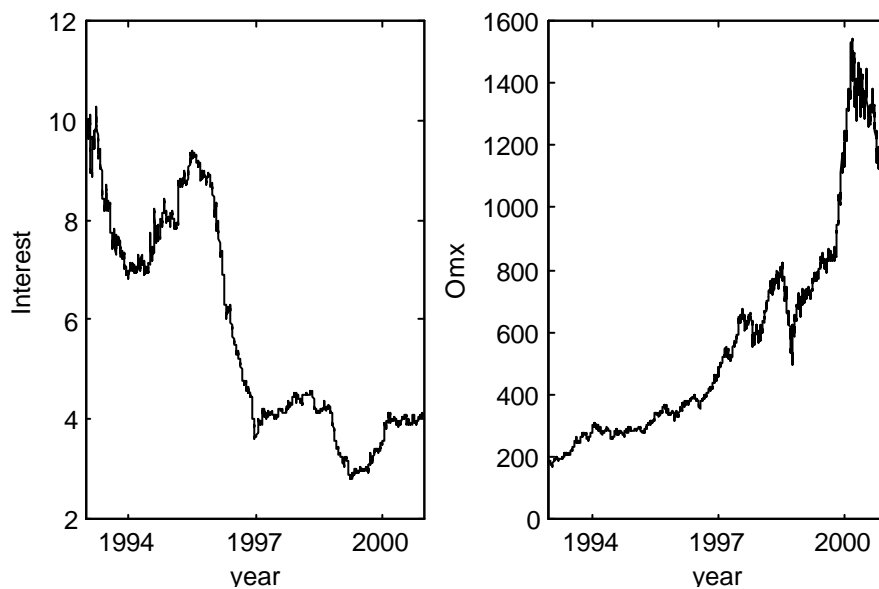
Figure 1.1.D: *Interest on the Swedish 3-month Treasury bills (left) and the OMX index (right)*
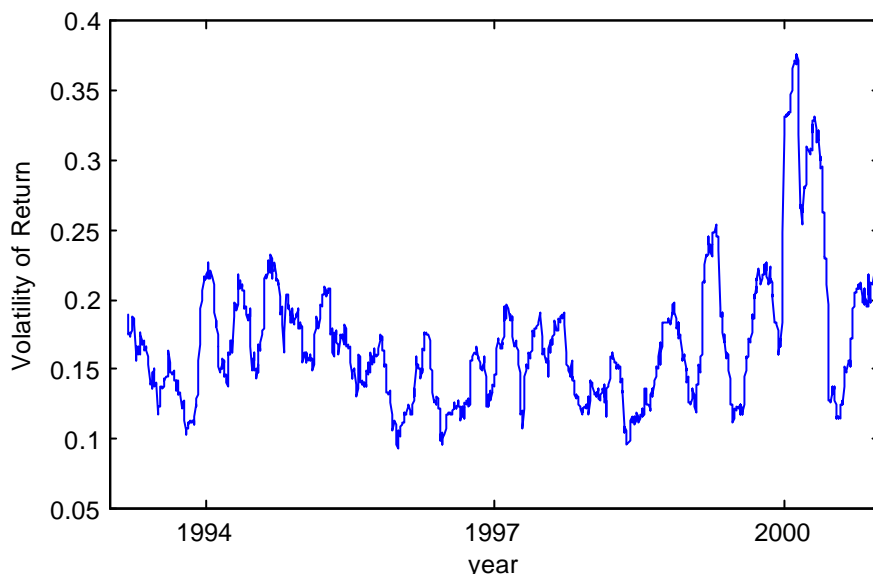
Figure 1.1.E: *Standard deviation of annual return on the OMX index*

In the Black and Scholes formula, all parameters are directly observable except for the standard deviation of the annual return. In Figure 1.1.E the standard deviation of the annual return on the OMX index is shown for each day in the sample period. Following common practice, the standard deviation for each day was calculated using the daily returns of the previous 30 days.

Put-Call Parity

In order to measure how well Put-Call Parity holds, we study the absolute difference between the observed call price, C_{obs}, and the price of the syntethic call option, C_{synt}, which is constructed by borrowing the present value of the exercise price, buying a put with the same exercise price and time to expiration as C_{obs} and buying the underlying asset. In the data set there are 10469 pairs of puts and calls traded at the same date with identical strike price and time to expiration. Although the underlying asset in this case is an index it is possible to buy the stocks comprising the index. Also, since there exists stock funds, OMXACT and Erik Penser Aktieindex Sverige OMX, mimicking the OMX index, another way of purchasing the equivalent of the underlying asset would be to

invest in such a fund. Dividing this quantity by the exercise price K we obtain the *relative difference*

$$\frac{|C_{obs} - C_{synt}|}{C_{obs}}$$

The distribution of the relative difference is shown in Figure 1.1.F, indicating that roughly 90% of the options have a difference less than 0.03%. Thus, disregarding other factors affecting arbitrage opportunities, transaction costs in excess of this would prevent profits to be made from this slight deviation from Put-Call Parity. It should be noted however that there exists an asymmetry in transaction costs due to the fact that the syntethic call option is created by buying the stock (index), for which the price, and hence the transaction cost, is higher.
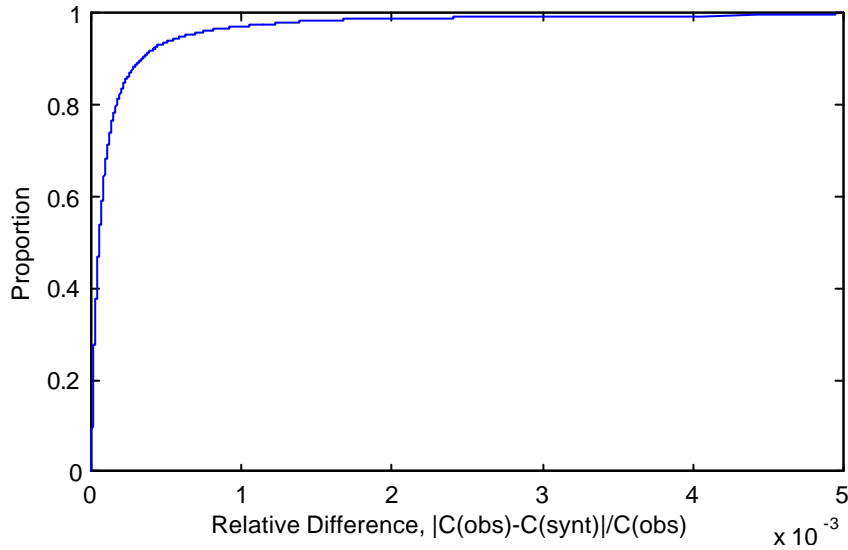
Figure 1.1.F: *Cumulative distribution of the the relative difference over the* 10469 *observations*

1.2 Numerical Procedure

For the purpose of obtaining the implied *annual* volatility we apply the discrete time BS model to the daily OMX data. Thus, viewing the one period model (Cf. Appendix A) as pertaining to one year, we now let the discrete time Black and Scholes model correspond to an M period

model, with one period equal to one day. Hence, in analyzing a call
option C_t at date t with T days to expiration, R, U and D refer to the
daily returns. The implied *daily* volatility associated with a call option
C_t with T periods to expiration, strike price K and current index value S_t
is extracted by numerically solving for the upfactor U in the non-linear
equation given by putting the observed option price C_{obs} equal to the
value given by the right hand member of the discrete time BS formula.
Cf. Appendix A

$$C_t \quad = \quad \frac{1}{R^T} \sum_{k=0}^{T} \binom{T}{k} p^k (1-p)^{T-k} \max(S_t U^k D^{T-k} - K, 0)$$

From Appendix A we have, setting q equal to $1/2$

$$U = e^{\mu_q/M + \sigma_q/\sqrt{M}} \text{ and } D = e^{\mu_q/M - \sigma_q/\sqrt{M}}$$

where M now denotes the number of business days in a year. Further-
more, by assuming the drift μ to be zero we have $U = \exp(\sigma_q/\sqrt{M})$ and
$D = 1/U$. Using that $p = (R-D)/(U-D)$ we can now write

$$C_t = \frac{1}{R^T} \sum_{k=0}^{T} \binom{T}{k} \left(\frac{RU-1}{U^2-1} \right)^k \left(\frac{U^2 - RU}{U^2 - 1} \right)^{T-k} \left(U^{2k-T} S_t - K \right)^+$$

where $\left(U^{2k-T} S_t - K \right)^+ = \max(U^{2k-T} S_t - K, 0)$. Hence, given the known
parameter values T, R, S_t and K, the price of a call option is a function
solely depending on U. Getting a solution consistent with absence of
arbitrage requires that the search is constrained to the interval where
$U > R$. The assumption of zero drift is a reasonable approximation since
we are using daily data; over the sample period the mean daily μ is
$8.72 * 10^{-4}$. However, in applying the numerical procedure we can not
be sure to find a U which satisfies the equation $C(U) - C_{obs} = 0$. For
given values of T, R,and S_t the function $h(U) = C(U) - C_{obs}$ is depicted
in Figure 1.2.A for two different values of the strike price K. The figure
shows that for options with high moneyness the minimum of $C(U)$ might
well exceed the observed option price, in which case $h(U) = 0$ does not
have a solution with $U > 1$.
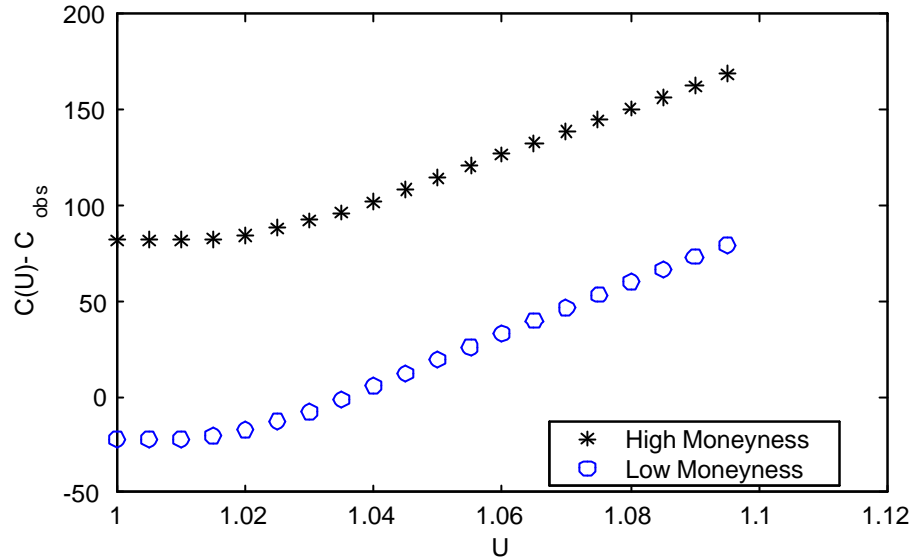
Figure 1.2.A: *The function $h(U) = C(U) - C_{obs}$ is depicted for two call options with strike prices 900 and 1100, respectively. They both have $C_{obs} = 22$, $S_t = 1002.5$, $R = 1.0001$ and $T = 15$.*

For approximately 1.5% of the calls and 0.2% of the puts in the dataset considered it was not possible to solve for U. These options differ from options for which U could be extracted with respect to volume, time to expiration and moneyness. Specifically, they were considerably less traded, had much higher moneyness and on average their time to expiration was half that of options with an implied U. Having solved for the daily upreturn U, the estimate of the daily standard deviation becomes $\ln U$. Cf. Appendix A. Assuming 252 business days, the annual volatility σ is then given by

$$\sigma = \sqrt{252}\ln U$$

Figure 1.2.B illustrates that letting one day constitute one step provides a good approximation of the implied volatility from the continuous BS formula. For the call options in the sample, the average difference between the volatilities extracted from the continuous BS model and the volatilities extracted from the discrete BS model with a period taken to be one day was 0.0019 When each day was divided into two subperiods

the average difference was 0.0011.Figure 1.2.B illustrates that letting one
day constitute one step provides a good approximation of the implied
volatility from the continuous BS formula. It is also seen that letting
each day comprise two steps in the discrete BS model only marginally
improves the approximation of the continuous implied volatility.
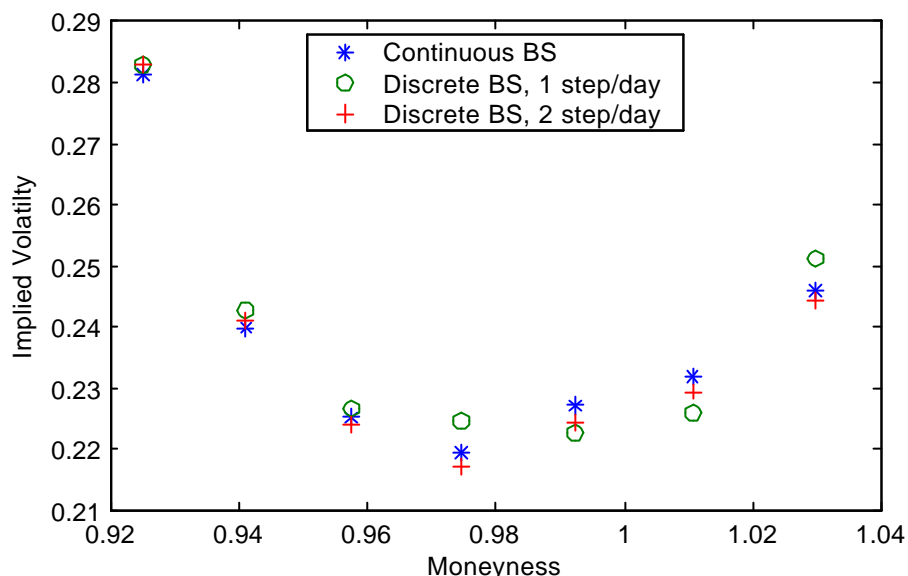
Figure 1.2.B: *Volatilities implied for a call option with 8 days*
to expiration by the continuous BS, the discrete
BS with one step/day and the discrete BS with 2
steps/day

Extracting the implied volatility for put options was done analogously.
Table 1.2.A shows the number of options for which it was not possible
to solve for U and the number of options which violated the arbitrage
condition.

Table 1.2.A: *Results from Numerical Calculations*

	Calls	Puts
# obs	20215	21783
No solution in U	319	63
Violating no-arbitrage	3	1
# volatilities	19893	21719

1.3 Volatility Smiles and Time to Expiration

According to the BS model, the volatility should be constant across both moneyness and time to expiration. Geometrically, the volatility surface at a given date then corresponds to a plane parallel to the moneyness - time-to-expiration plane, as seen in Figure 1.3.A. In case the parameters of the BS model were also constant over calender time the same volatility plane would apply to each date.
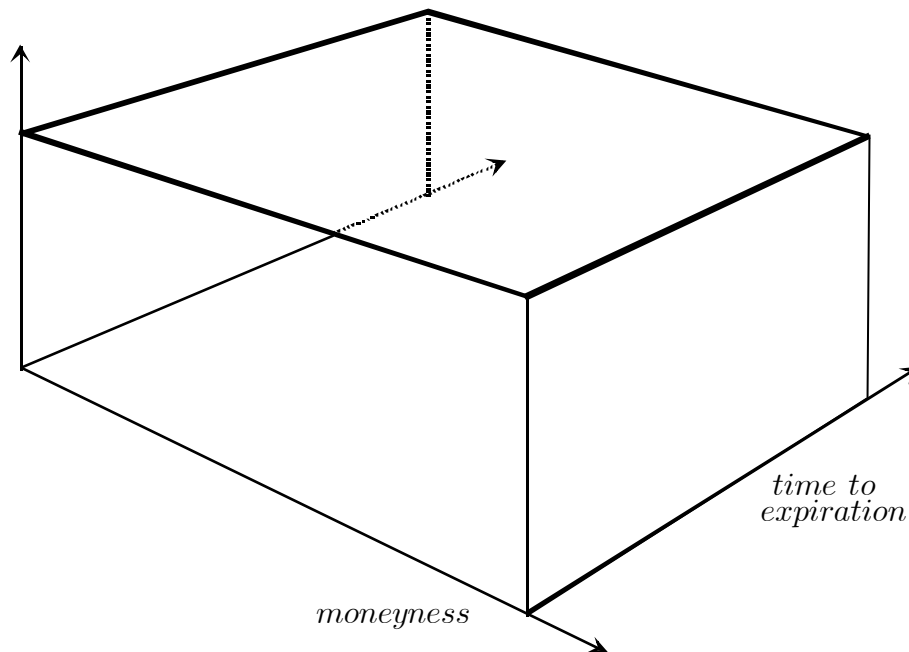
Figure 1.3.A: *At date t the implied volatility, according to the BS model is constant as a function of exercise price and time to expiration*

However, it is well documented that observed option prices usually fail to conform to the structure predicted by the BS model, so that data plots often exhibit "smiles" or "smirks". We will refer to such deviations simply as "smiles". A first step towards the construction of models better suited to explain option prices is to study closer the deviations from the BS model. Thus for the purpose of characterizing the nature and extent of the violations of the predictions by BS we introduce the notions of

level of smile and *strength of smile*.

Depending on the nature of the deviations from the BS assumptions, the volatility surface assumes different shapes. The volatility surface will be analyzed with respect to moneyness and time to expiration separately.

Let t denote the calender date, T the time to expiration and M moneyness (defined as the ratio of the OMX index and the exercise price for calls and the inverse ratio for puts). Consider a given calender date, $\bar{t}$. For the moment assume that time to expiration and moneyness are continuous variables. If the implied volatility for varying moneyness but fixed time to expiration, $\bar{T}$, is constant this will show up as a line in the volatility surface, at date $\bar{t}$. Thus varying $\bar{T}$ the volatility surface will be traced out by a family of lines as shown in Figure 1.3.B. Let $\sigma_{t,T,M}$ denote the implied volatility for an option at date t with T days to expiration and moneyness M.
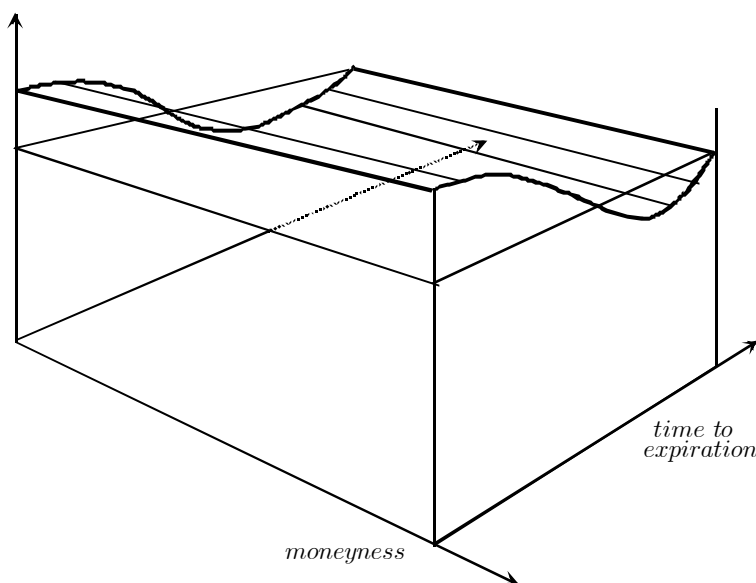
Figure 1.3.B: *Level of smile, at date $\bar{t}$, varies with time to expiration. Strength of each smile, given $\bar{T}$, is 0*

The date $\bar{t}$ *level of smile*, given time to expiration $\bar{T}$, is

$$\frac{1}{\left|\left(\bar{t},\bar{T},M\right)\right|} \sum_{M} \sigma_{\bar{t},\bar{T},M} = \sigma_{\bar{t},\bar{T}}$$

where $\left|\left(\bar{t},\bar{T},M\right)\right|$ denotes the number of observations at date $\bar{t}$ with $\bar{T}$ days to expiration as M varies. Cf. Figure 1.3.B where the volatility is

constant for each time to expiration so that there exists a valid BS model for each time to expiration but where, in general, the valid model will vary with time to expiration. The implied volatilities at date $\bar{t}$ for options with $\bar{T}$ days to expiration will typically also vary with moneyness so that these options give rise to a "smile" around the "level of the smile". Cf. Figure1.3.C.

The date $\bar{t}$ *strength of smile*, given time to expiration $\bar{T}$, is

$$\left(\frac{1}{\left| \left(\bar{t}, \bar{T}, M \right) \right|} \sum_M \left(\sigma_{\bar{t},\bar{T},M} - \sigma_{\bar{t},\bar{T}} \right)^2 \right)^{1/2} = z_{\bar{t},\bar{T}}$$

where $\left| \left(\bar{t}, \bar{T}, M \right) \right|$ denotes the number of observations at date $\bar{t}$ with $\bar{T}$ days to expiration as M varies.
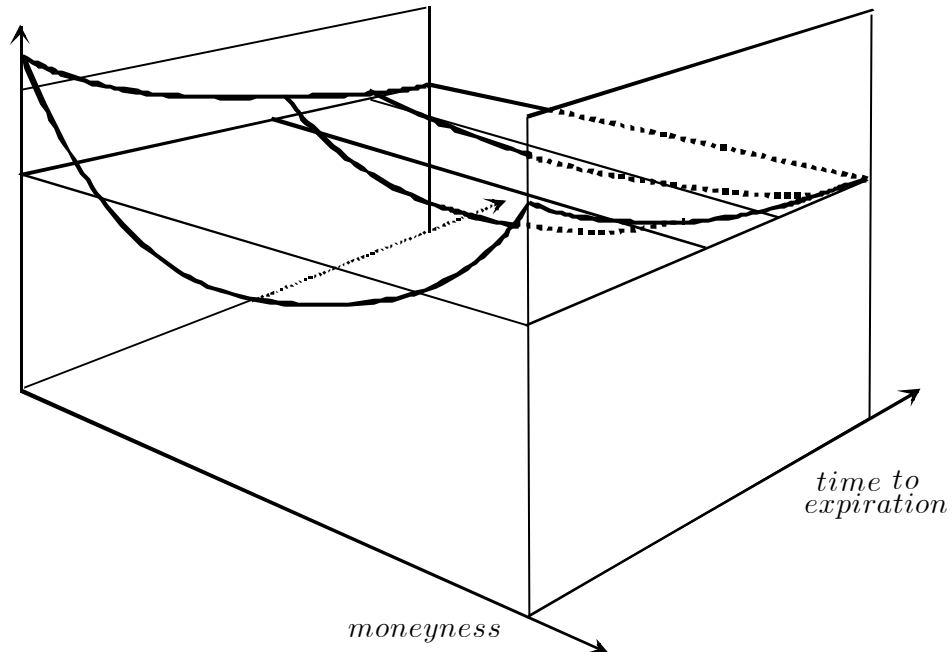
Figure 1.3.C: *Level of smile constant as time to expiration varies. Strength of smile decreasing with time to expiration*

Figure 1.3.C illustrates a case where the strength of the smiles decreases with time to expiration, in such a way as to keep the level of the smile constant.

Table 1.3.A shows the number of observations used in calculating the levels, for calls and puts respectively. While the level was calculated if

there were at least one observation, the strength was only calculated if
there were at least 2 observations. It follows that the number of observa-
tions used in calculating the strength is one less that of the number used
in computing the level.

Table 1.3.A: *number of observations used*

	Number of levels													
	1	2	3	4	5	6	7	8	9	10	11	12	$\geq$13	Σ
# obs/level Calls	6	10	12	17	14	11	22	26	21	15	16	4	19	1468
# obs/level, Puts	14	10	7	7	16	14	18	22	25	20	15	8	21	1550

Calls and Smiles given Time to Expiration

In this section we study the level and strength of smiles generated by
call options with times to expiration $T = 10, 30, 50$ and 70. This gives a
sample size of 1522 implied volatilities.

Recall that the *level of smile* at date $\bar{t}$ given time to expiration $\bar{T}$
was defined as

$$\frac{1}{\left|(\bar{t}, \bar{T}, M)\right|} \sum_M \sigma_{\bar{t}, \bar{T}, M} = \sigma_{\bar{t}, \bar{T}}$$

Figures 1.3.D and 1.3.E show how the level of smile for call options
has evolved over calendar time for times to expiration $T = 10, 30, 50$
and 70. The total number of levels were 193. The median number of
observations used in calculating the level was 8. As a consequence of the
relatively sparse trade in options with long maturity, the median value of
$\left|(\bar{t}, \bar{T}, M)\right|$ was only 3 for calls with 70 days to expiration. For each time
to expiration the level of the smiles have varied considerably over time
with the annual levels ranging between 0.15 and 0.30. Another feature
is the marked covariation between the level of smiles for different times
to expiration.

Figure 1.3.D: *Level of smile for calls with* 10 *days to expiration (left) and for calls with* 30 *days to expiration (right)*

Figure 1.3.E: *Level of smile with for calls with* 50 *days to expiration (left) and for calls with* 70 *days to expiration (right)*

To summarize the information in Figures 1.3.D and 1.3.E we define the *average level of smile,* given time to expiration $\bar{T}$, as

$$\frac{1}{|t,\bar{T}|} \sum_t \sigma_{t,\bar{T}} = L_{\bar{T}}$$

where $|t,\bar{T}|$ is the number of $\sigma_{t,\bar{T}}$ as t varies. Figure 1.3.F shows how the average level of smile is decreasing with time to expiration.

The *standard deviation of level of smile,* given time to expiration $\bar{T}$, is

$$\left[\frac{1}{|t,\bar{T}|} \sum_t \left(\sigma_{t,\bar{T}} - L_{\bar{T}}\right)^2\right]^{1/2} = SL_{\bar{T}}$$

The standard deviation of level of smile, as T varies, is given in Figure 1.3.F, which confirms the impression from Figures 1.3.D and 1.3.E that the standard deviation is relatively constant and the average level is decreasing about one half of the standard deviation as time to expiration increases.
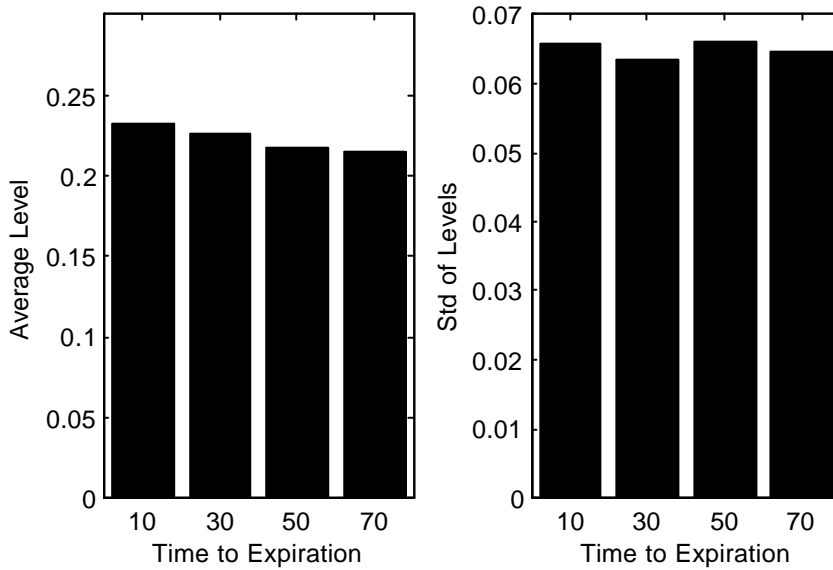
Figure 1.3.F: *Average level of smile for calls (left) and its standard deviation (right)*

The date $\bar{t}$ *strength of smile*, given time to expiration $\bar{T}$, was defined above as

$$\left(\frac{1}{|(\bar{t},\bar{T},M)|} \sum_M \left(\sigma_{\bar{t},\bar{T},M} - \sigma_{\bar{t},\bar{T}}\right)^2\right)^{1/2} = z_{\bar{t},\bar{T}}$$

Figure 1.3.G: *Strength of smile for calls with 10 days to expiration (left) and for calls with 30 days to expiration (right)*
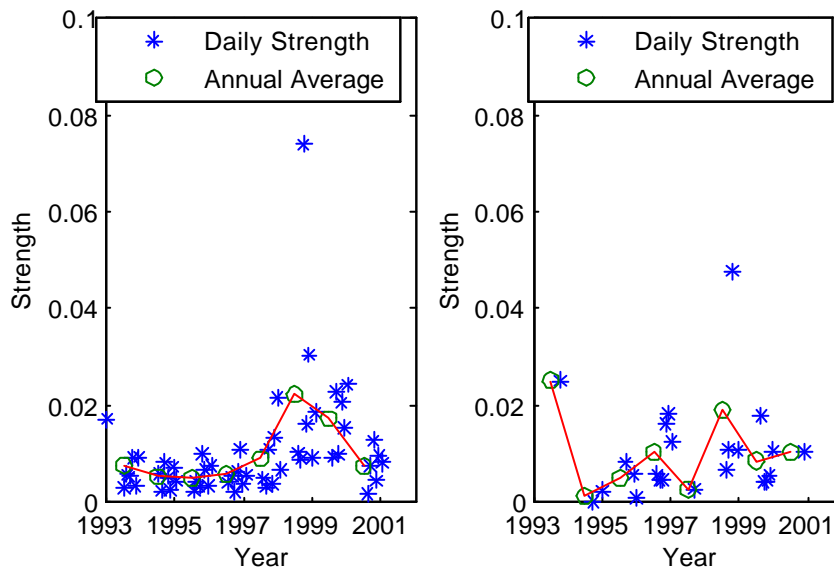
Figure 1.3.H: *Strength of smile for calls with 50 days to expiration (left) and for calls with 70 days to expiration (right)*

Computing the strength of smile as time to expiration varies yields

189 observations. Figures 1.3.G and 1.3.H show how the strength of smile has evolved over calender time for times to expiration $T = 10, 30, 50$ and 70. The figures indicate that the strength of smile decreases as time to expiration increases. They also suggest that the variations in strength of smiles over calendar time is smaller than for the levels of smiles. This difference is more pronounced for longer maturities. It should be noted that the number of observations for which the strength was calculated varied between 2 and 24.

To summarize the information in Figures 1.3.G and 1.3.H we define the *average strength of smile,* given time to expiration $\bar{T}$, as

$$\frac{1}{\left|\left(t, \bar{T}\right)\right|} \sum_{t} z_{t,\bar{T}} = Z_{\bar{T}}$$

where $\left|t, \bar{T}\right|$ is the number of $z_{t,\bar{T}}$ as t varies. Figure 1.3.I shows that the average strength of smile decreases strongly with time to expiration.
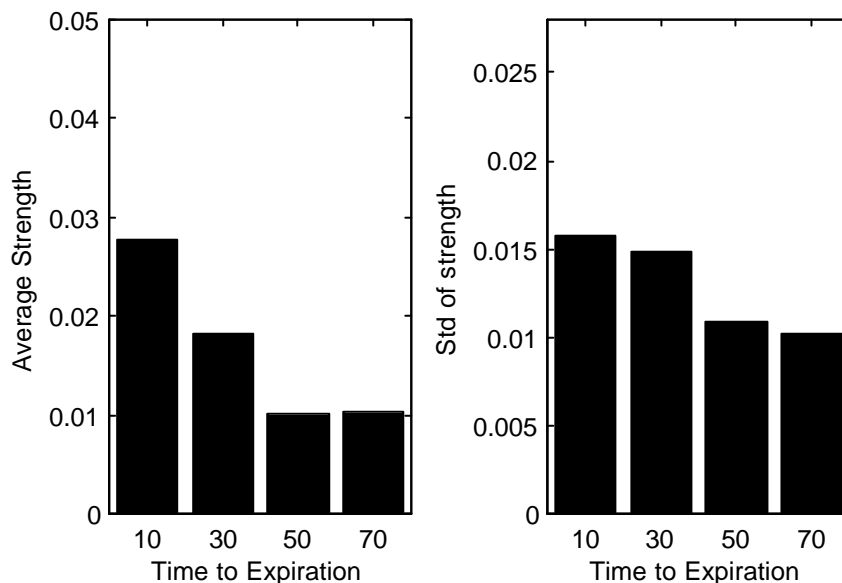
Figure 1.3.I: *Average strength of smile for calls (left) and standard deviation of strength (right)*

The *standard deviation of strength of smiles,* for time to expiration $\bar{T}$ is

$$\left[\frac{1}{\left|\left(t, \bar{T}\right)\right|} \sum_{t} \left(z_{\bar{t},\bar{T}} - Z_{\bar{T}}\right)^{2}\right]^{1/2} = SZ_{\bar{T}}$$

Figure 1.3.I shows the standard deviation of strength of smiles, as T varies.

Unlike the case with levels of smiles, the standard deviation of strength of smiles is decreasing with time to expiration which indicates that the deviation from the constant volatility assumption is less severe for calls with long maturity.

Puts and Smiles given Time to Expiration

In this section the previous analysis for calls is repeated for puts. For puts with time to expiration $T = 10, 30, 50$ and 70 we get 1622 observations, for which we can compute 197 levels. As for calls, the median number of observations used in calculating the level was 8. From Figures 1.3.J and 1.3.K, depicting the levels over time, it is seen that options with longer maturity exhibit lower level of smile. This is the same pattern exhibited by calls, although the level of smile for puts are in general higher for all times to expiration.
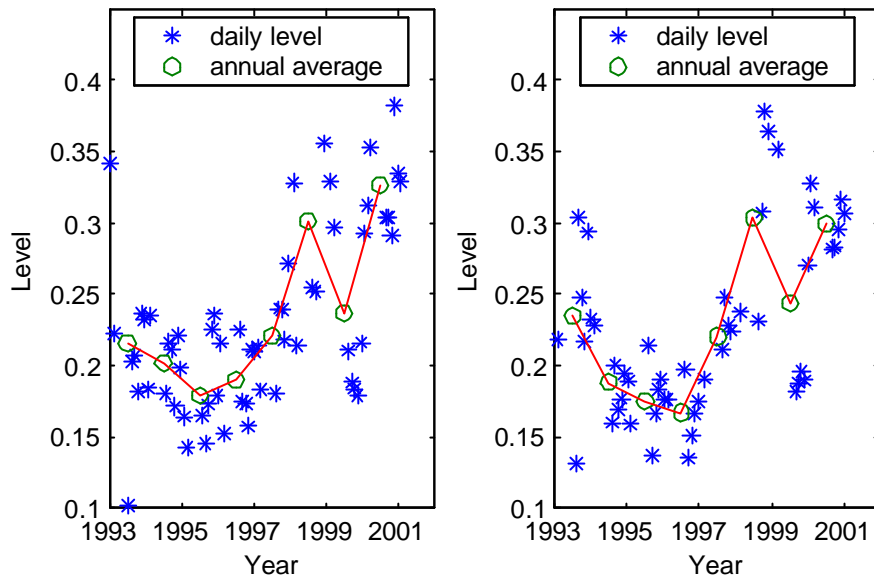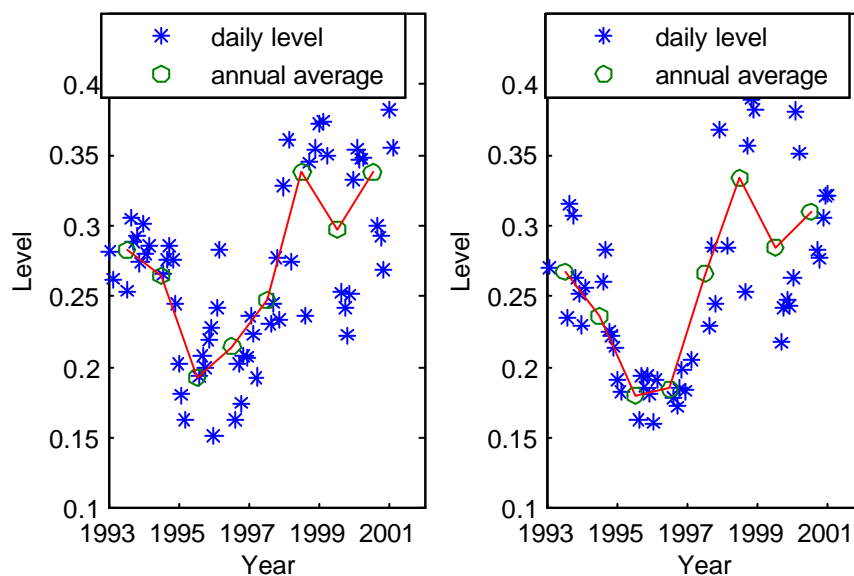

Figure 1.3.J: *Level of smile for puts with* 10 *days to expiration (left) and* 30 *days to expiration (right)*
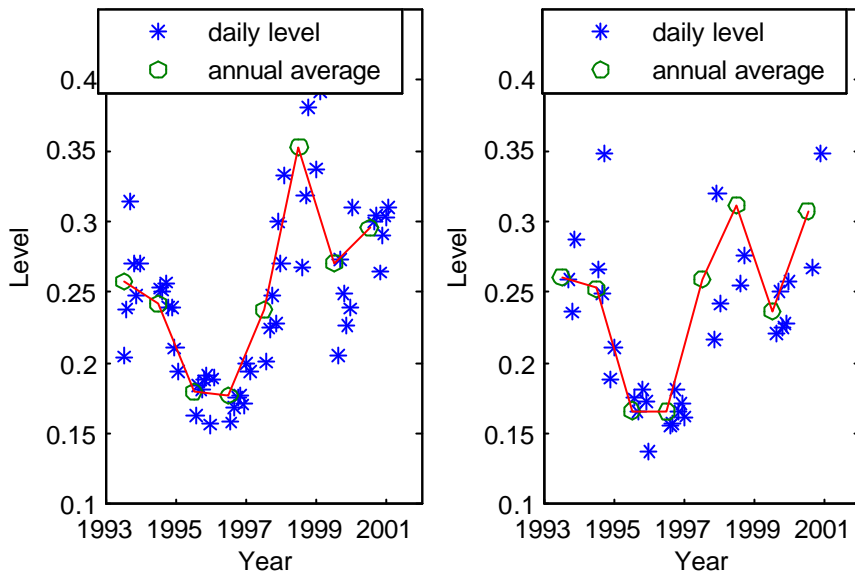
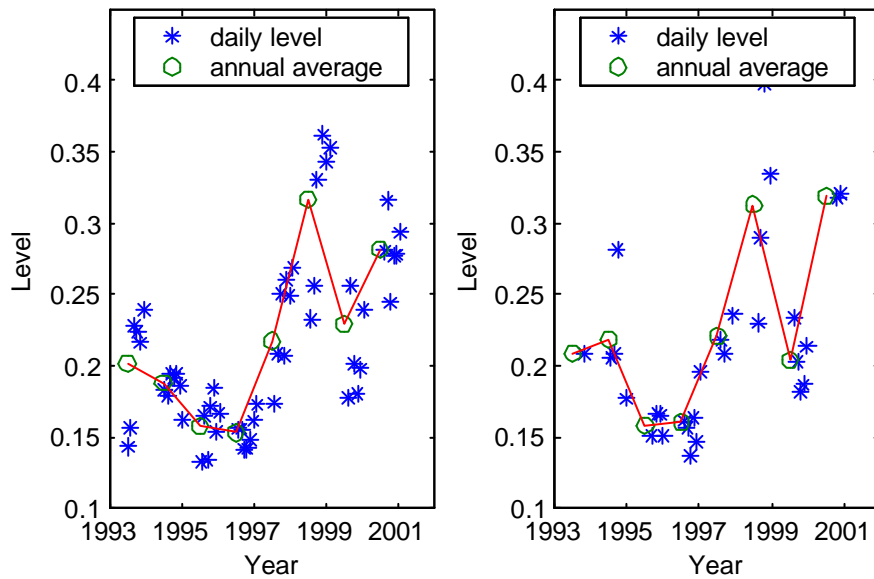**Figure 1.3.K:** *Level of smile for puts with* 50 *days to expiration*
(left) and 70 *days to expiration (right)*
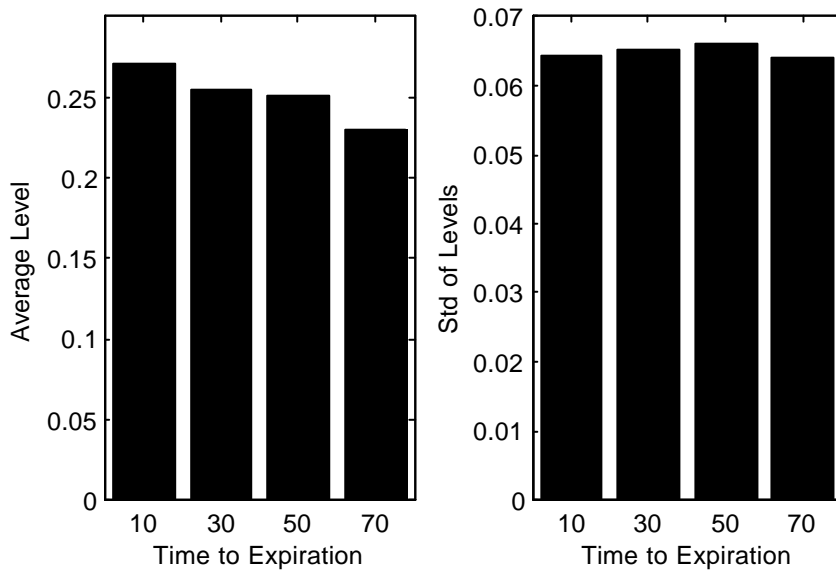
Figure 1.3.L: *Average level of smile for puts (left) and standard*
deviation of level for puts (right)

Computing the average level of smile and its standard deviation,
shown in Figure 1.3.L , reaffirms that the level is decreasing in time

to expiration and that the level is higher for puts than for calls, for all times to expiration. The standard deviation of level of smile is roughly constant over time to expiration and similar to those for calls.

In Figures 1.3.M and 1.3.N the strength of smile is displayed. Similar to the case for level of smile, the strength of smile is clearly decreasing in time to expiration. Comparing the strength of smile for puts with that of calls (Cf Figures 1.3.G and 1.3.H) it is seen that the strength for puts is markedly higher, especially for short maturities and for the early years in the sample. For example, in 1993 the annual average of strength of smile for puts with 10 remaining days to expiration is 0.06, which is about three times larger than for the corresponding calls.

Figure 1.3.O summarizes the findings in Figures 1.3.M and 1.3.N and shows how strength of smile and the standard deviation of strength of smile varies with time to expiration. It is seen that the pattern for calls, both with regard to average strength and its standard deviation is duplicated for puts.

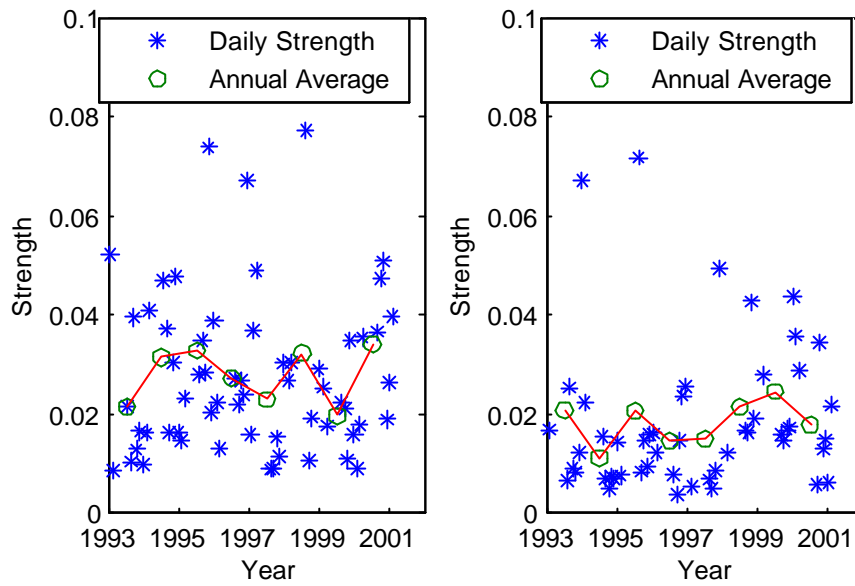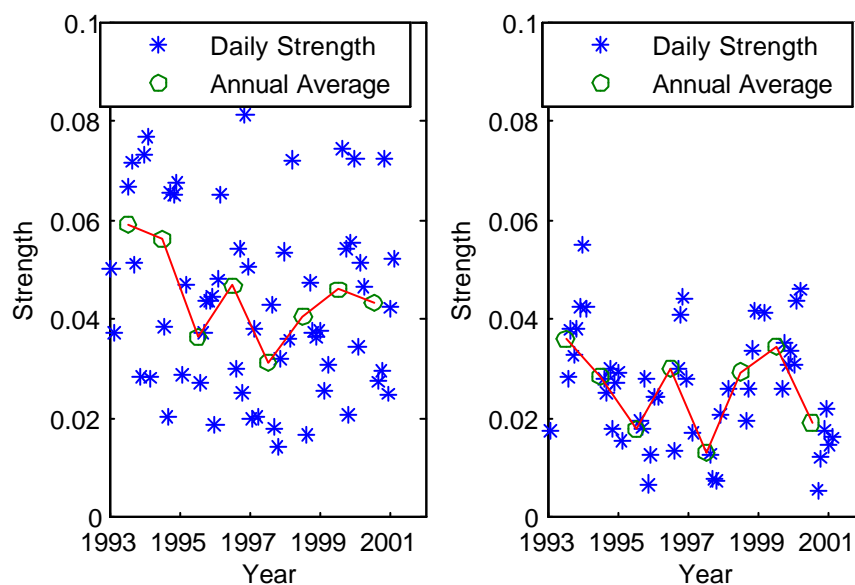**Figure 1.3.M:** *Strength of smile for puts with 10 days to expiration (left) and 30 days to expiration (right)*
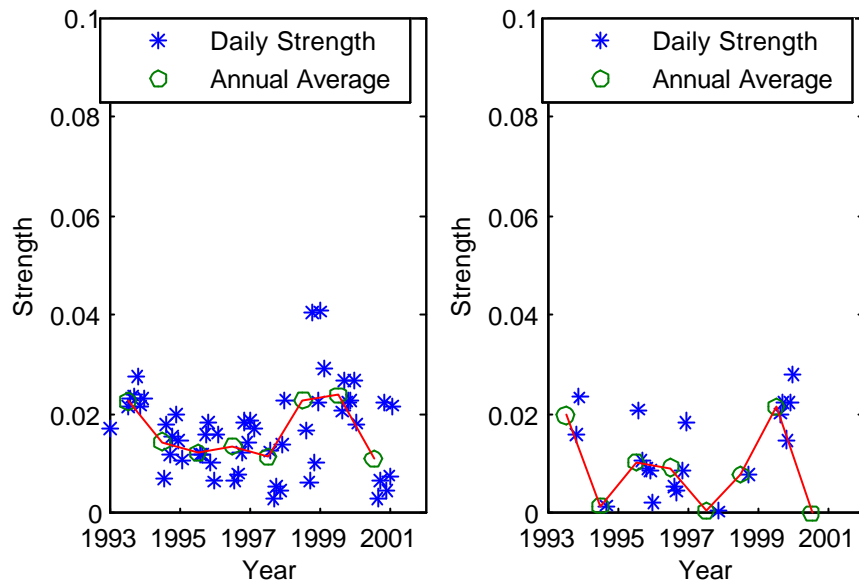
Figure 1.3.N: *Strength of smile for puts with* 50 *days to expira-*
tion (left) and 70 *days to expiration (right)*
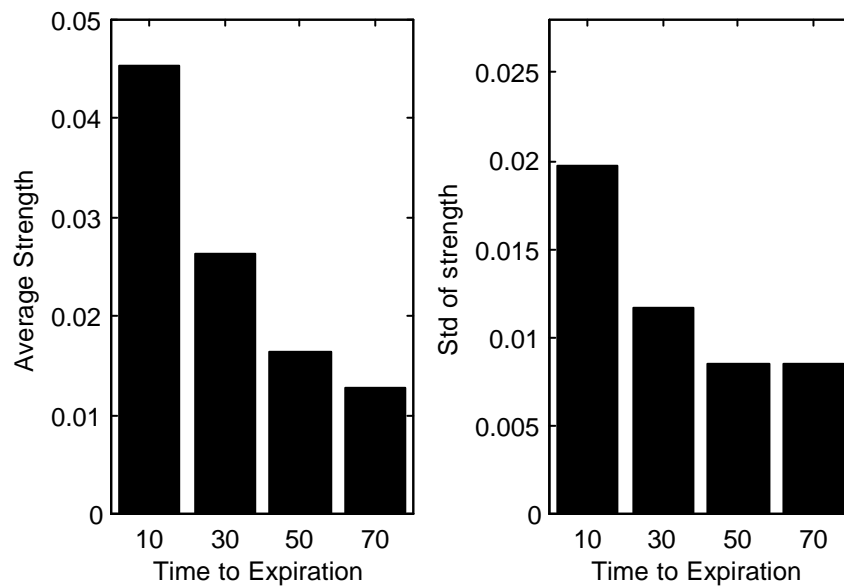
Figure 1.3.O: *Average strength of smile and standard deviation*
of strength of smile for puts

1.4   Volatility Smiles and Moneyness

When we applied the notions of strength and level of smile for a given time to expiration we considered the intersection of the volatility surface with a plane parallel with the moneyness-volatility axes.

In this section we want to study how level of smile and strength of smile varies, for given moneyness. Figure 1.4.A exhibits a case where, for each moneyness, the implied volatility is constant across time to expiration, but is decreasing in moneyness.

time to expiration
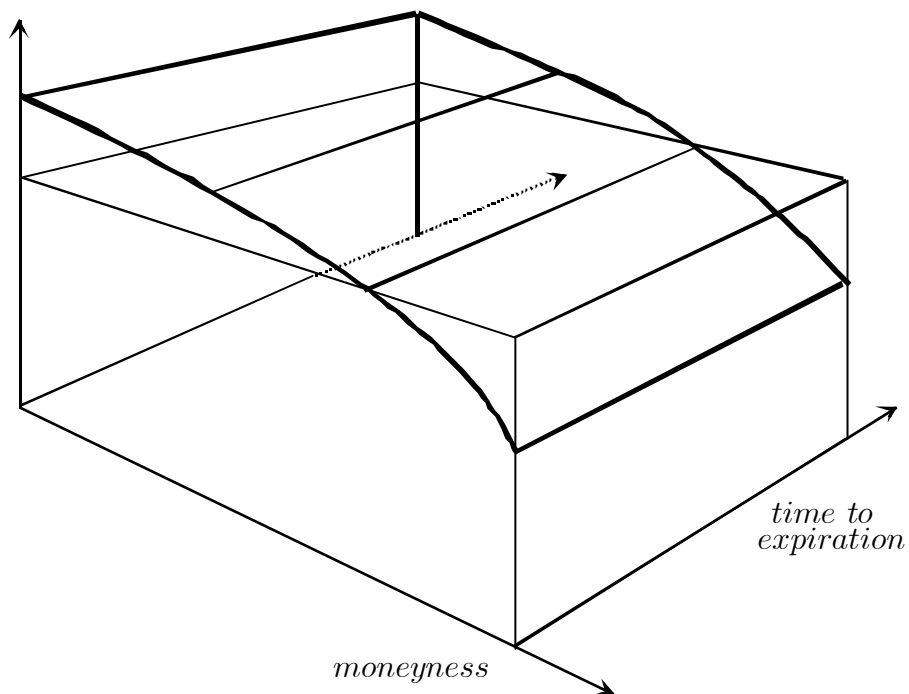
moneyness

Figure 1.4.A: *Strength of smile with respect to time to expiration is zero, but level of smile varies with moneyness*

Figure 1.4.B illustrates a case where the level of smile is decreasing with moneyness and where the strength of smile varies from high to low and back to high. Here the strength of smile attains a minimum for options at the money, indicated by the straight line. It follows that for options at the money a BS model would be a good approximation.

Table 1.4.A: *Observations used in computations od levels*

	Number of levels													
	1	2	3	4	5	6	7	8	9	10	11	12	$\geq$13	Σ
# obs/level, Calls	383	320	348	328	259	253	197	207	154	151	126	104	317	19566
# obs/level, Puts	527	459	407	340	306	248	256	188	126	145	115	110	349	20926
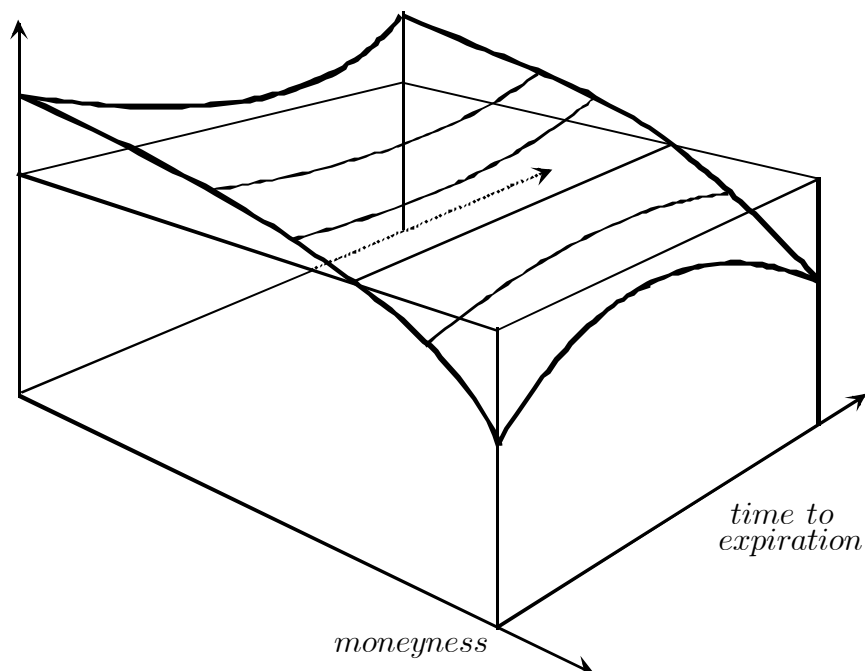
Figure 1.4.B: *Level of smile is decreasing with moneyness. Strength of smile varies from high to low and back to high*

Calls and Smiles given Moneyness

Since moneyness is a continuous variable we can not quite duplicate the computations done when analyzing the volatility surface given time to expiration.

In order to resolve this difficulty we let M denote moneyness belonging to an interval. Thus $\sigma_{\bar{t},\bar{M}}$ denotes the average of implied volatilities,

at calendar date $\bar{t}$, with moneyness in some interval $\bar{M}$. In this section we study the level and strength of smile generated by call options with moneyness in the intervals $(0.8, 0.9)$, $(0.9, 1.0)$, $(1.0, 1.1)$ and $(1.1, 1.2)$.

For a given calendar date $\bar{t}$ and moneyness $\bar{M}$ we thus study the implied volatility for call options with varying time to expiration.

The date $\bar{t}$ *level of smile*, given moneyness $\bar{M}$, is

$$\frac{1}{\left|\left(\bar{t}, T, \bar{M}\right)\right|} \sum_T \sigma_{\bar{t}, T, \bar{M}} = \sigma_{\bar{t}, \bar{M}}$$

where $\left|\left(\bar{t}, T, \bar{M}\right)\right|$ denotes the number of observations at date $\bar{t}$ with moneyness $\bar{M}$. The total number of levels were 3147. The median number of observations used to compute the level was 5.

Since moneyness is a continuous variable there is a high correlation between the number of options belonging to a certain moneyness interval at date t and the number of options belonging to the same interval at date $t + 1$. Plotting the daily level over time results in a highly cluttered picture. Hence, Figures 1.4.C and 1.4.D depict not the daily level, but rather its annual average, for different intervals of moneyness. The figures show that the level of smile given moneyness display the same pattern over calender time as the level of smile given time to expiration. For calls deeply out-of-the-money and deeply in-the-money the levels are located higher than for calls with moneyness close to unity.

The *average level of smile*, given moneyness $\bar{M}$, is

$$\frac{1}{\left|\left(t, \bar{M}\right)\right|} \sum_t \sigma_{t, \bar{M}} = L_{\bar{M}}$$

where $\left|\left(t, \bar{M}\right)\right|$ is the number $\sigma_{t, \bar{M}}$ as t varies.

The *standard deviation of level of smile*, given moneyness $\bar{M}$, is

$$\left(\frac{1}{\left|\left(t, \bar{M}\right)\right|} \sum_t \left(\sigma_{t, \bar{M}} - L_{\bar{M}}\right)^2\right)^{1/2} = SL_{\bar{M}}$$

where $\left|\left(t, \bar{M}\right)\right|$ denotes the number of $\sigma_{t, \bar{M}}$.as t varies.
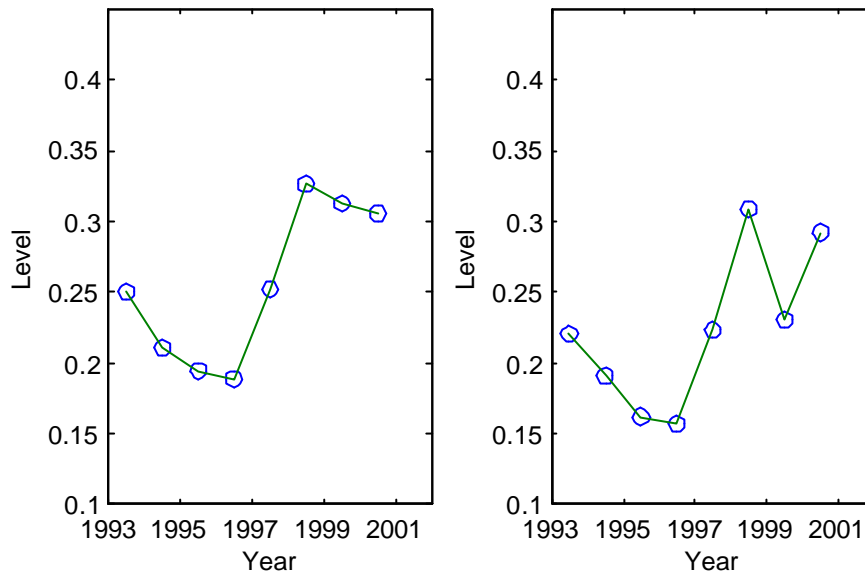
Figure 1.4.C: *Annual average of daily level for calls with mon-
eyness between 0.8 and 0.9 (left) and between 0.9
and 1 (right)*
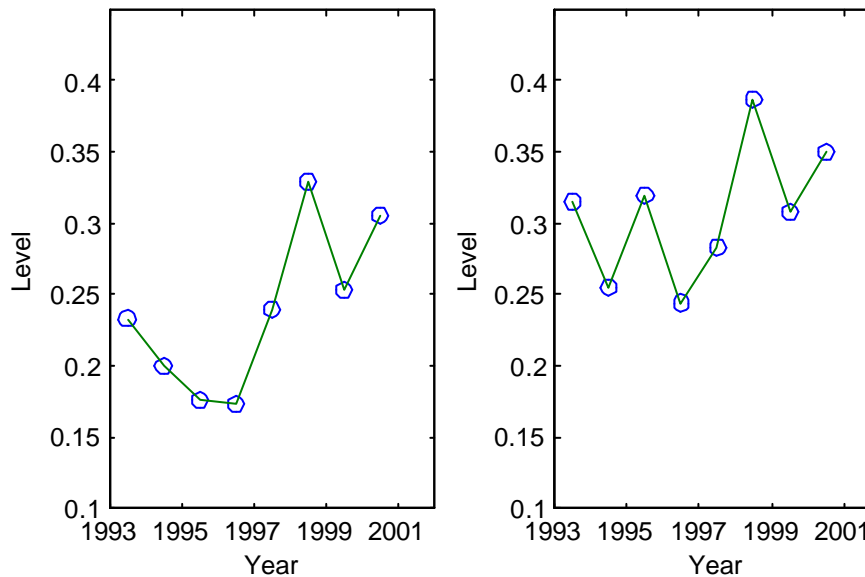
Figure 1.4.D: *Annual average of daily level for calls with mon-
eyness between 1 and 1.1 (left) and between 1.1
and 1.2 (right)*

Figure 1.4.E confirms the findings of Figures 1.4.C and 1.4.D, with the highest average level for extreme values of moneyness. By contrast, the standard deviation of level of smile is increasing in moneyness.
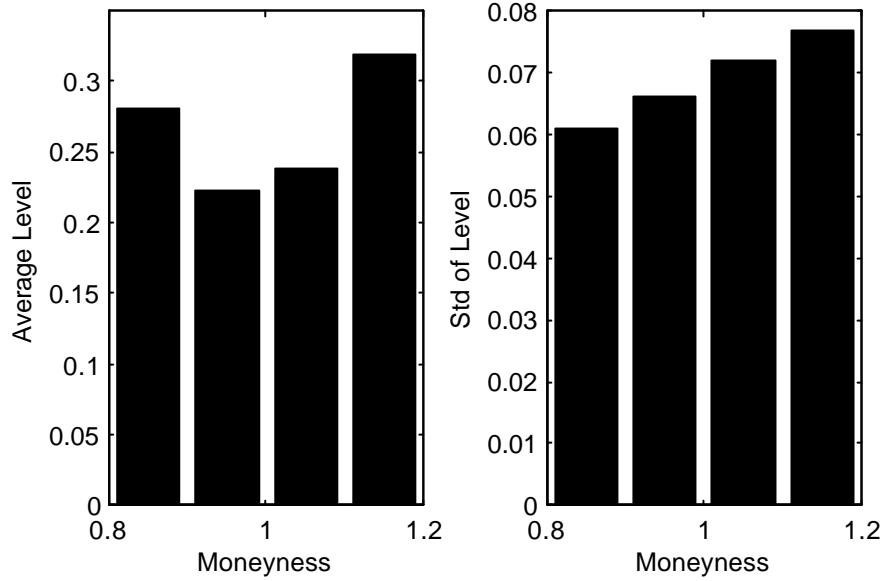
Figure 1.4.E: *Average level of smile for calls (left) and standard deviation of level of smile for calls (right)*

Recall that the date $\bar{t}$ *strength of smile*, given moneyness $\bar{M}$, is

$$\left(\frac{1}{\left|\left(\bar{t}, T, \bar{M}\right)\right|}\sum_{T}\left(\sigma_{\bar{t},T,\bar{M}} - \sigma_{\bar{t},\bar{M}}\right)^2\right)^{1/2} = z_{\bar{t},\bar{M}}$$

where $\left|\left(\bar{t}, T, \bar{M}\right)\right|$ denotes the number of observations, at date $\bar{t}$, with moneyness $\bar{M}$ as T varies. The total number of observations were 2764. Figures 1.4.F and 1.4.G show how the strength of smile has evolved over time. For extreme values of moneyness the strength and its variation over time is considerably higher than for calls with moneyness closer to unity.
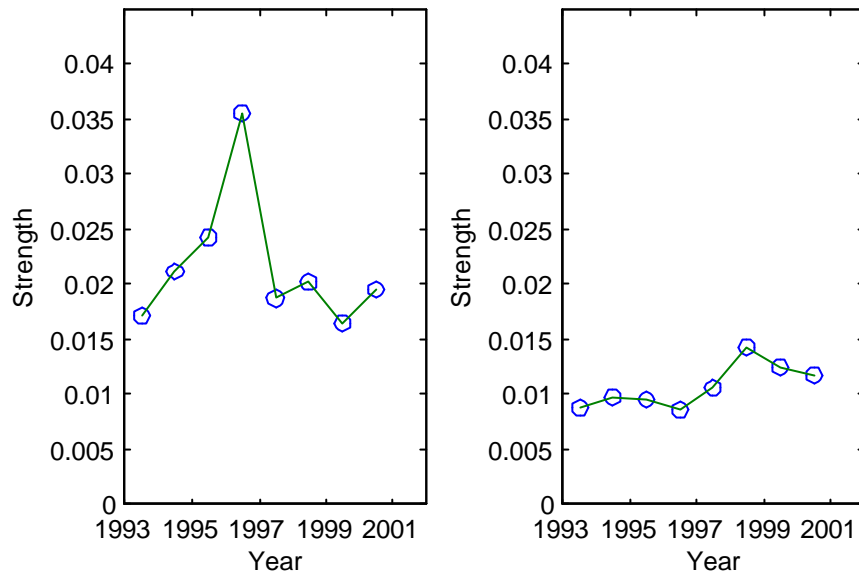
Figure 1.4.F: *Strength of smile for calls with moneyness between* 0.8 *and* 0.9 *(left) and between* 0.9 *and* 1 *(right)*
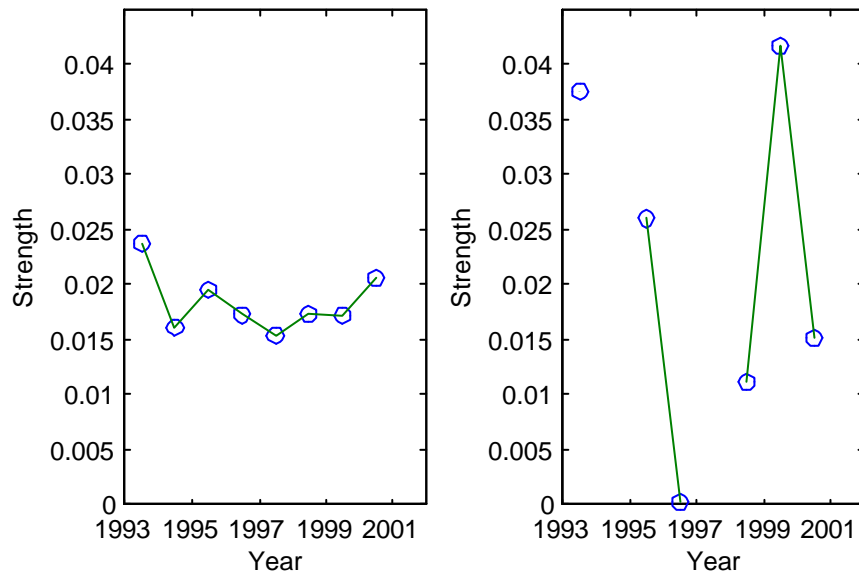
Figure 1.4.G: *Strength of smile for calls with moneyness between* 1 *and* 1.1 *(left) and between* 1.1 *and* 1.2 *(right)*

The *average of strength of smile*, given moneyness $\bar{M}$, is

$$\frac{1}{\left|(t, \bar{M})\right|} \sum_t z_{t,\bar{M}} = Z_{\bar{M}}$$

where $\left|(t, \bar{M})\right|$ denotes the number of $z_{t,\bar{M}}$ as t varies.

The *standard deviation of strength of smile*, given moneyness $\bar{M}$, is

$$\left(\frac{1}{\left|(t, \bar{M})\right|} \sum_t \left(z_{t,\bar{M}} - Z_{\bar{M}}\right)^2\right)^{1/2} = SZ_{\bar{M}}$$

where $\left|(t, \bar{M})\right|$ denotes the number of $z_{t,\bar{M}}$ as t varies. Figure 1.4.H shows that a large average strength is accompanied by a large standard deviation. Also, it is clear that the largest strength is found for calls deep in-the-money, implying that the Black and Scholes model would be particularly ill suited to price these option.
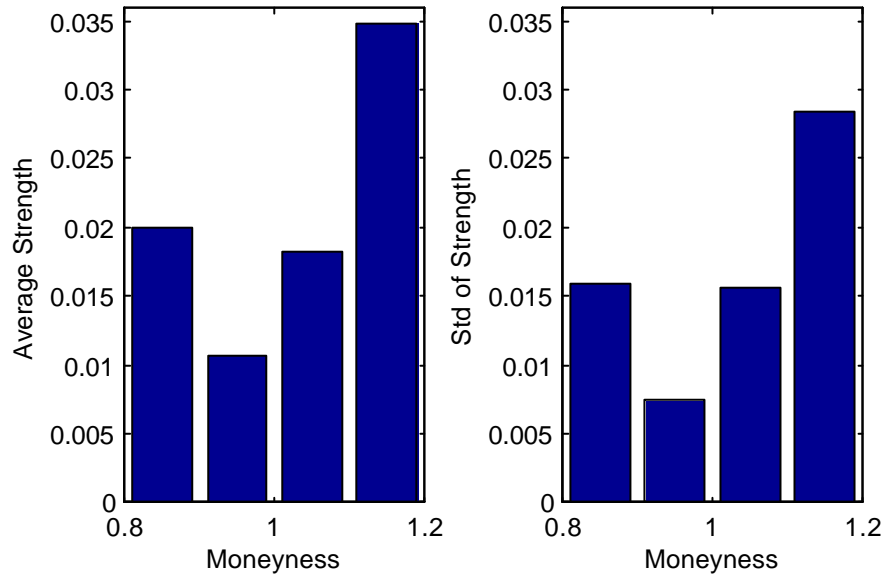
Figure 1.4.H: *Average strength of smile for calls (left) and standard deviation of strength of smile for calls (right)*

Puts and Smiles given Moneyness

Again, the analysis made for calls is repeated for puts. Computing the level of smile for the puts in the sample yields 3576 observations, where

the median number of observations used to compute each level was 5. The annual averages of these daily levels are depicted in Figures 1.4.I and 1.4.J. The preceding sections showed that the levels of smile, given time to expiration, displayed the same pattern over time for calls and puts, with the levels for puts significantly higher. Comparing the levels of smile given moneyness for calls and puts, we see that for all years and all intervals of moneyness, the levels for puts are considerably higher for puts than for calls.

In Figure 1.4.K it is seen that the highest average level is found for extreme values of moneyness, while the standard deviation is increasing in moneyness. Again , this mimics the pattern displayed by puts. By contrast, while the average level is higher for puts, its standard deviation seems slightly lower than the corresponding for puts.

Computing the strength of smile yields 3049 (daily) observations, with only 41 of these observations having moneyness in excess of 1.1. The annual averages of these are shown in Figures 1.4.L and 1.4.M, indicating that the strength of smile varies considerably less over time than the levels.
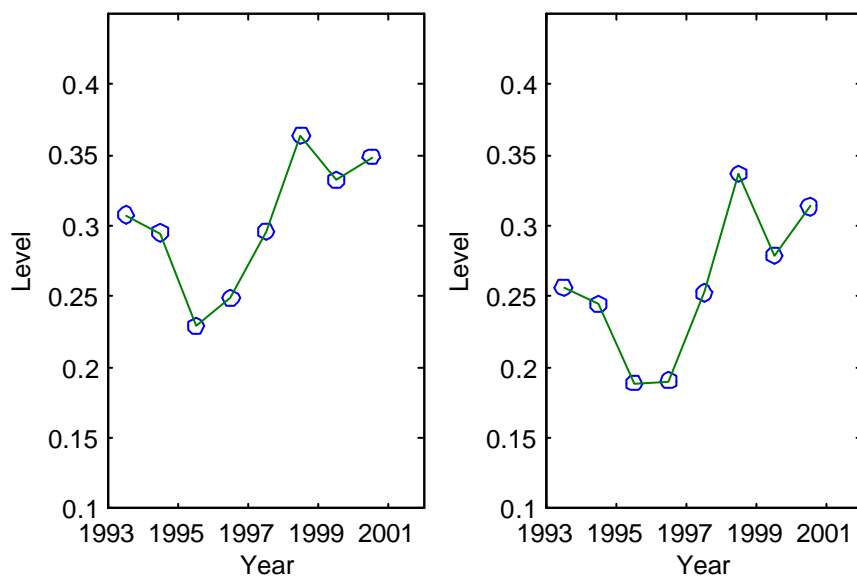
Figure 1.4.I: *Level of smile for puts with moneyness between* 0.8 *and* 0.9 *(left) and between* 0.9 *and* 1 *(right)*
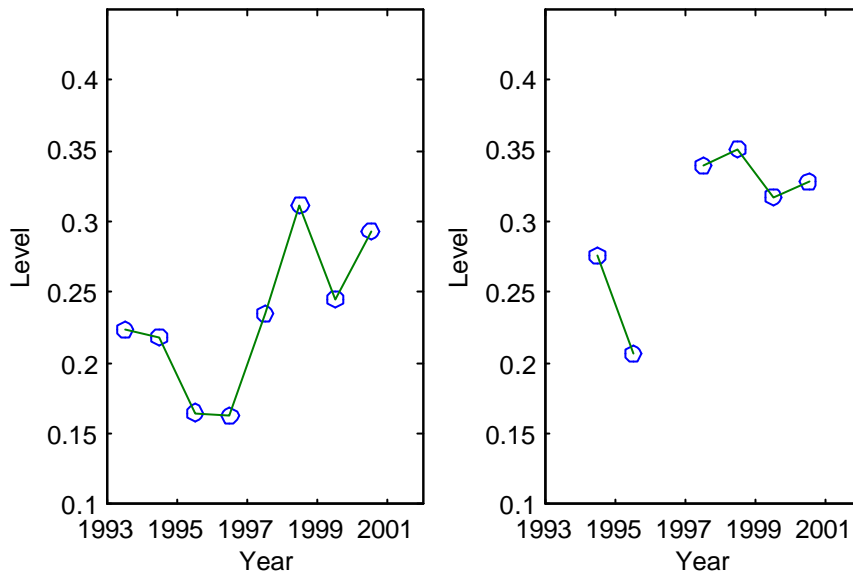
Figure 1.4.J: *Level of smile for puts with moneyness between* 1 *and* 1.1 *(left) and between* 1.1 *and* 1.2 *(right)*
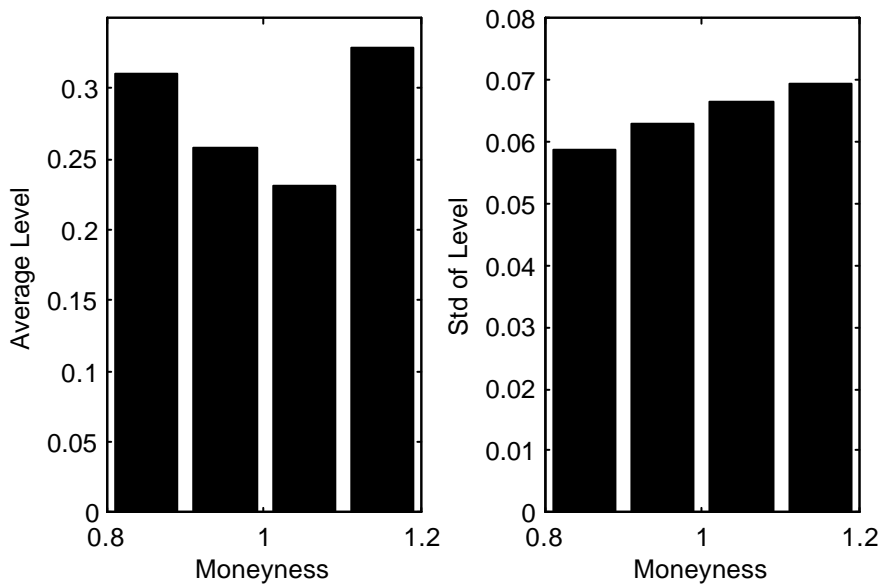
Figure 1.4.K: *Average level for puts (left) and standard devia- tion of levels of smile for puts (right)*
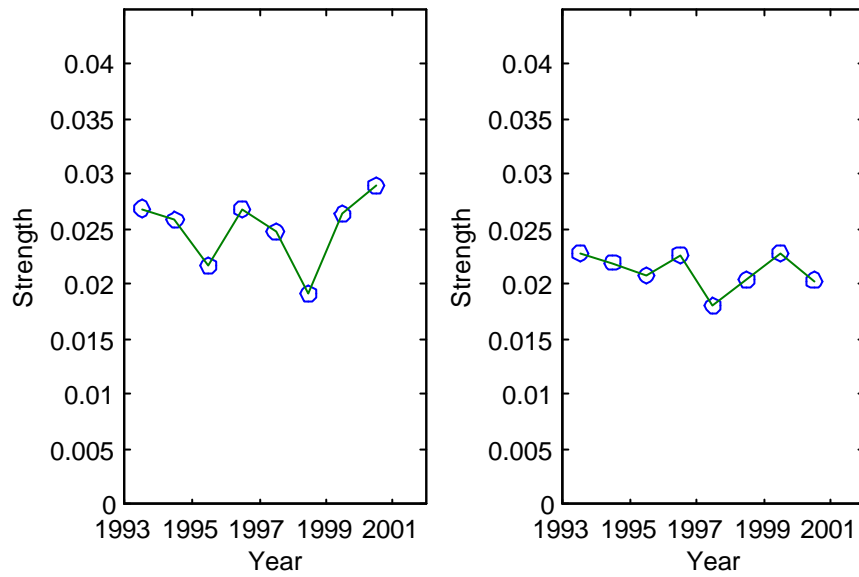
Figure 1.4.L: *Strength of smile for puts with moneyness between* 0.8- 0.9 *(left) and* $0.9 - 1$ *(right)*
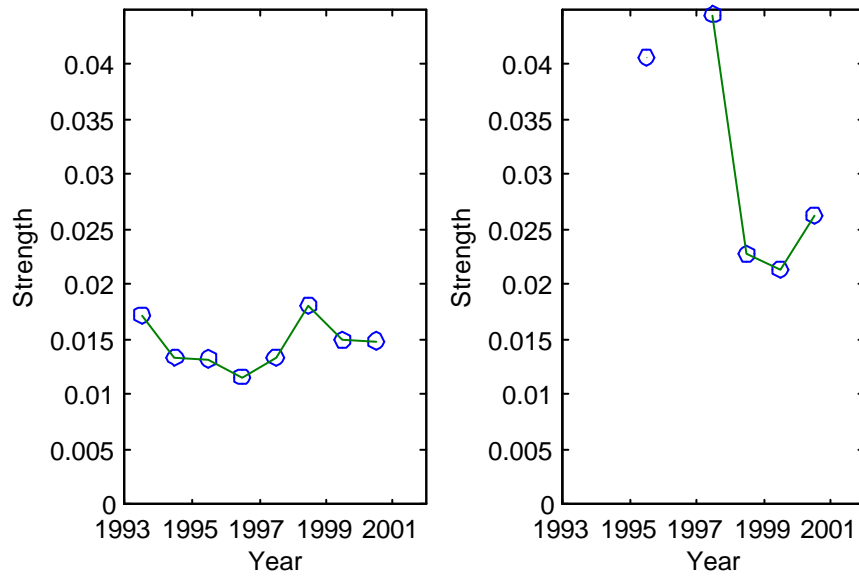
Figure 1.4.M: *Strength of smile for puts with moneyness between* $1- 1.1$ *(left) and* 1.1-1.2 *(right)*

The average strength of smile and their standard deviation is shown in Figure 1.4.N. Comparing with the average strength for calls, it is seen

that the average strength is higher for puts for out-of-the-money options, while the opposite is true for in-the-money options. As was the case for calls, the largest strength is obtained for extreme values of moneyness. The standard deviation of strength for out-of-the-money puts are higher than for the corresponding calls, while the standard deviation of strength for in-the-money puts is lower than for the corresponding calls.
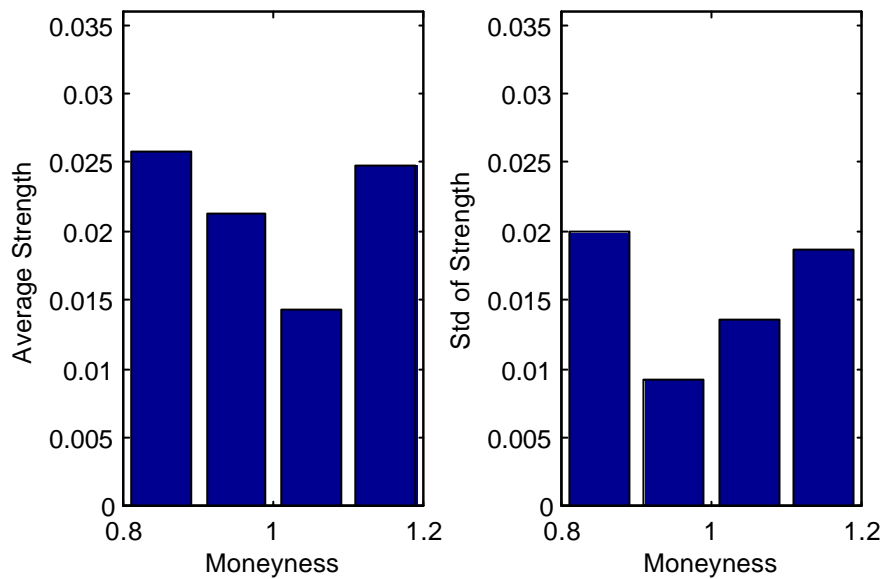
Figure 1.4.N: *Average strength of smile for puts and standard deviation of strength for puts*

1.5 Comparison of Volatility Surfaces for Calls and Puts

As a final step we estimate the volatility surfaces for calls and puts using kernel regression. This exercise provides us with three-dimensional graphs linking moneyness and time to expiration to implied volatility, thus summarizing the findings in the preceding sections.

The volatility surface for calls is shown in Figure 1.5.A and the volatility surface for puts is shown in Figure 1.5.B. An inspection of the surfaces

 reveals the marked strength of smile with respect to moneyness for puts, which is notably higher than the strength of smile with respect to moneyness for calls. While the strength of smile with respect to moneyness decreases with time to expiration for both puts and calls, the strength of smile for puts with long maturity is still significant. By contrast, the strength of smile for calls with long maturity appears marginal.

For both calls and puts the strength of smile with respect to time to expiration is the highest for extreme values of moneyness. Finally, the volatility surface for puts seems to be located higher than the volatility surface for calls, reflecting the fact that the average volatility implied by puts is 0.2659 while the average volatility implied by calls is 0.2368. For the 10469 pair of options for which Put-Call-Parity was applied, the difference in implied volatility was less significant; the average implied volatility was 0.2389 for calls and 0.2474 for puts. This is taken as evidence that the arbitrage forces underlying Put-Call-Parity indeed is in effect.
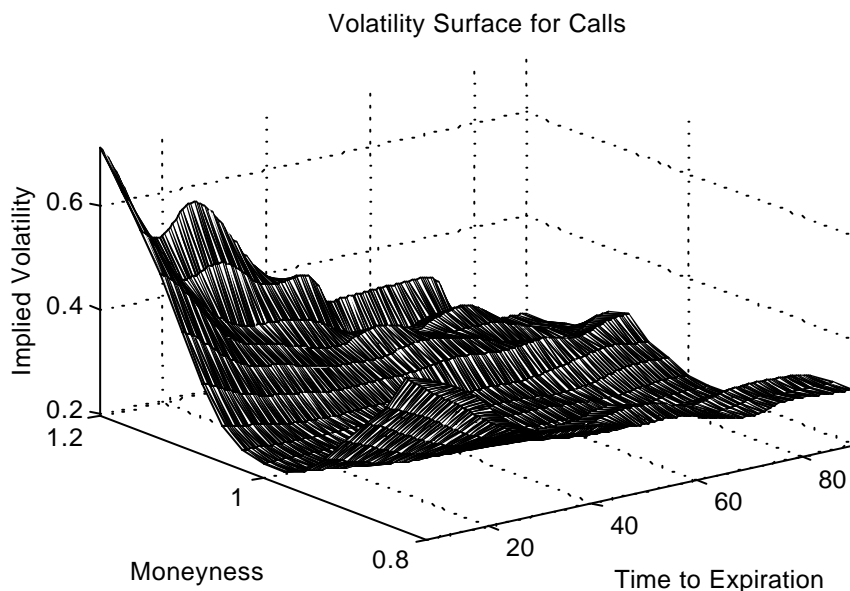
Figure 1.5.A: *Implied volatility surface for calls*

Volatility Surface for Puts
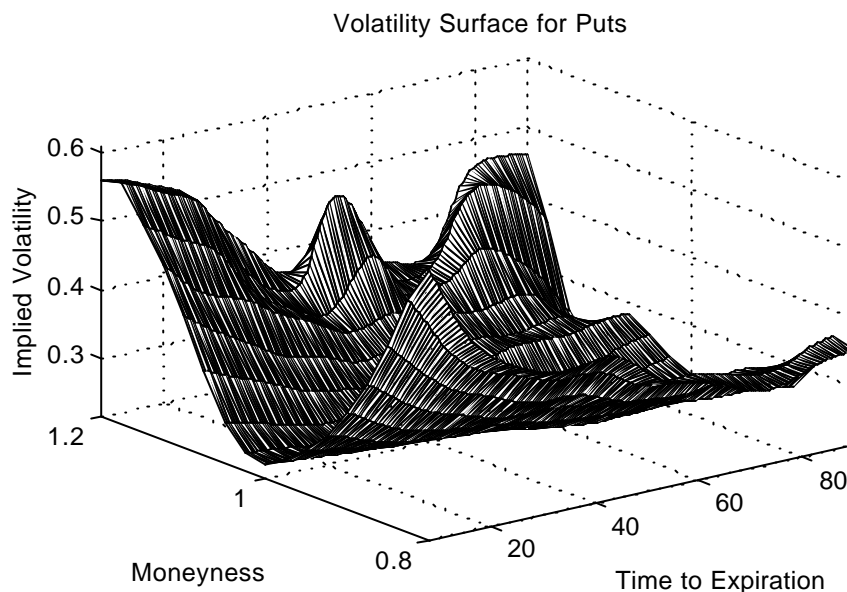
Figure 1.5.B: *Volatility surface for puts*

Summary

When pricing an option one has to take into consideration whether the underlying asset pays dividends during the life of the option. For an option on an index such as the OMX index which is not adjusted for dividends one thus has to estimate the dividend yield during the life of the option.

However, when the dividend yield was computed using Spot-Futures Parity the result was ambiguous. Although for some years the dividend yield clearly peaked during the dividend period, this pattern was not consistent throughout the sample. Furthermore, the dividend yield often assumed negative values. For these reasons the options affected by dividends were discarded.

The arbitrage relation known as Put-Call Parity was tested by computing the relative difference between the call and a portfolio constructed by borrowing the present value of the exercise price, buying a put and the underlying asset. It was shown that for about 90% of the observations the relative difference was less than 0.03%, suggesting that the possibility of exploiting the relation for arbitrage profits is remote.

The volatility surface was analyzed with respect to time to expiration and moneyness for calls and puts separately.

When the development of levels of the smiles over time for calls was studied, it was found that the highest levels over the sample period was found for calls with the shortest time to expiration.

Summarizing the information over time for calls, the *average level of smile given time to expiration* was seen to be monotonically decreasing in time to expiration, whereas the standard deviation of the average levels remained constant.

The strength of smile given time to expiration for calls were relatively stable over time, for all years and times to expiration.

The *average strength of smiles* for calls is decreasing in time to expiration, with the average strength of smile for calls with 10 days to expiration almost 3 times that of calls with 70 days to expiration. This would indicate that the assumption of constant volatility is relatively more appropriate for calls with long time to expiration, although this conclusion is made somewhat more uncertain by the *standard deviations of the strength of smiles* being higher for calls with short time to expiration.

The analysis of the *average levels* given moneyness for calls revealed that the average level was the highest for calls with extreme values of moneyness. The average strength of smiles given moneyness was the highest for extreme values of moneyness . This fact, in conjunction with the observation that the strength of smiles given time to expiration was the highest for calls with short time to expiration implies that we would expect the Black and Scholes model to be particularly flawed when pricing calls with high moneyness and short time to expiration.

Analyzing the volatility surface for puts we note both some similarities and dissimilarities as compared to the case for calls. The results for given time to expiration were as follows.

For puts, the highest levels throughout the sample period was found for puts with the shortest time to expiration. Again, the pattern for calls is duplicated in the distribution of the average level of the smiles given time to expiration for puts, with the average levels for puts being higher than for calls.

When we consider the strength of smiles given time to expiration for puts, the pattern is similar to that for calls, with relatively stable annual means for each maturity.

The average strength of smile given time to expiration for puts is decreasing in time to expiration and higher than for calls, for all maturities.

The analysis of the smile structure for puts given moneyness gave the following results. The average level of smiles for puts given moneyness were the highest for extreme values of moneyness, mimicking the pattern for calls, although the average levels were higher than for calls, for all intervals of moneyness.

The average strength of smile given moneyness was the highest for extreme values of moneyness. Interestingly, for out-of-the-money puts the average strength was higher than for the corresponding calls, whereas the opposite was true for in-the-money options.

Chapter 2

Bayesian Learning and Asset Pricing

Introduction

Drawing on the work by Timmermann and Guidolin [2003] this chapter investigates the consequences of introducing learning in asset pricing models. The starting point is a special case of the asset pricing model due to Lucas [1978], the Full Information (FI) model. It is shown that this model give identical option price predictions as the (discrete time) Black and Scholes (BS) option pricing model, given that certain conditions on the parameters are met. Whereas the BS model takes as given an equilibrium where the price of the stock and the bond are determined, the FI model is a representative agent, equilibrium model. Thus, the FI model can be seen as an equilibrium rationalization of the BS model.

The basic construction for (the discrete version of) the model is a binomial tree and the dividend pattern of the stock and the bond over the nodes of the tree. The given prices for the stock and bond induce state prices for contingent income and risk-neutral probabilities.

The specification of the dividend pattern of the stock and the bond implies that the risk-neutral probabilities are invariant over time. Other assets may then be priced by arbitrage. Since markets are complete one can, on the one hand, construct a portfolio which is varied over time so as to mimic the dividend pattern of, for example a European

call option. Since the valuation of such a portfolio is determined by the assumed prices of the stock and the bond an assumption of non-existence of arbitrage induces an implicit valuation of the option.

Another way to arrive at a price for the option is to make use of the fact that complete markets and the non-existence of arbitrage induces uniquely determined, positive state prices for contingent income at each of the nodes in the tree. The price of the option is simply the current value of the future uncertain income stream of the option.

Given the state prices it is possible to construct risk neutral probabilities. The current price of the option is then the expected value of the (time) discounted future dividends with respect to the probability measure given by the risk-neutral probabilities.

Observed deviations from predictions by the Black and Scholes model

It has been noted that the prices predicted by the BS pricing formula deviate from observed prices in a systematic way. The BS model implies a time invariant volatility. Thus one can use options with different exercise dates or exercise prices to infer the implied volatility. For a given exercise date it turns out that options with extreme exercise prices, "far out-of-the-money" or "deep in-the-money" implies volatility in excess of that implied by options "at-the-money". Drawing diagrams relating exercise price to implied volatility one gets "smiles" or "smirks".

A related observation is that the risk-neutral probability distribution predicted by the BS model appears to be less spread out than the risk-neutral probability distribution, which can be inferred from option prices.

Introducing learning in asset pricing models

Here we will study the consequences of introducing learning in the model. To do so it is useful to construct explicitly an equilibrium model, in the vein of Lucas [1978]. Such an equilibrium model then induces the stock and bond prices taken as given in the BS model. When learning is introduced the agents are equipped with some initial probability belief,

their "priors", and after observing the rise or fall in the stock price they update these probabilities. Agents are rational and are assumed to take into account that they will learn from future price movements.

We show that this type of learning actually can be described by an urn model due to Polya. Although the model is a modest modification of the BS model, Timmermann and Guidolin [2003] reports that it agrees much better with observed prices than the BS model. While we derive the same option pricing formulas implied by the FI model and the model with Bayesian learning (BL), although in a different manner, as do Timmermann and Guidolin, the contribution here is twofold. First, we derive a closed form approximation of the stock an call price in the case of Bayesian learning using properties of the Polya distribution. Second, a precise description of the relationship between the BS model and the FI model is presented, thus giving an equlibrium rationalization of the BS model.

The introduction of learning is interesting and the method is suitable also for studying the price formation in other financial markets. In particular, it is relevant for the reinsurance market where the trade occurs in contracts whose value crucially depend on the probabilities of some underlying events. These probabilities are not invariant over time and thus a model incorporating learning is useful.

The Polya distribution and its relation to Bayesian learning is introduced in Section 2.1. In Section 2.2 the dividend process is described and the consumer problem is stated and state prices are derived for both the learning and the Full Information case. These state prices are then used in Section 2.3 to derive the stock and bond prices for the learning and Full Information case. Section 2.3 also presents approximations of the stock and call prices in the case of learning as well as a description of the relationship between the BS model and the FI model. To illustrate the effects of learning an example is presented in Section 2.4.

2.1   Bayesian Learning and Probability

The concept of Bayesian learning and how it differs from the Full Information case can be illustrated with the aid of the Polya distribution. For references on the Polya distribution see Johnson and Kotz [1977], Barton and David [1962] and Fisz [1963].

Ultimately we wish to arrive at a description of how fundamentals evolve in a binomial tree where at each point in time either an upstep U or a downstep D can occur. In what follows we present the Polya urn model and the beta-binomial model, which can be seen as a special case of the Polya model. The framework provided by the beta-binomial model facilitates the description of Bayesian learning as it allows specification of strength of learning and the probability beliefs of individuals.

Consider an urn initially containing n white balls and s black balls, so that the total number of balls in the urn is $N = n + s$. One draws a ball from the urn, notices its color and puts it back together with c balls of the same color, where

$$c = \{-1, 0, 1, 2, ...\}$$

Thereafter the procedure is repeated. Put

$$X_i = \begin{cases} 1, \text{ if you get a white ball at the i:th drawing} \\ \\ 0, \text{ otherwise} \end{cases}$$

The number of white balls in τ drawings is then

$$Y_\tau = X_1 + X_2 + ...X_\tau$$

The probability of getting a white ball in the first drawing is

$$\Pr(X_1 = 1) = \frac{n}{N}$$

The probability of getting a white ball in the second drawing is

$$\begin{aligned} \Pr(X_2 = 1) = \quad &\Pr(X_1 = 0)\Pr(X_2 = 1 \mid X_1 = 0) + \\ &\Pr(X_1 = 1)\Pr(X_2 = 1 \mid X_1 = 1) = \\ &= \frac{s}{N}\frac{n}{(N+c)} + \frac{n}{N}\frac{n+c}{(N+c)} = \\ &= \frac{(N+c)n}{N(N+c)} = \frac{n}{N} \end{aligned}$$

indicating that the $X_i : s$ are identically distributed.

To show that the $X_i : s$ are dependent, we note that

$$\Pr(X_1 = 1, X_2 = 1) = \Pr(X_1 = 1)\Pr(X_2 = 1 \mid X_1 = 1) =$$

$$= \frac{n}{N}\frac{n+c}{N+c}$$

which differs from

$$\Pr(X_1 = 1)\,P(X_2 = 1) = \frac{n^2}{N^2}$$

The probability of getting white balls in the first k drawings, when the total number of drawings is τ is

$$\Pr(X_1 = 1 = X_2 = \ldots = X_k = 1, X_{k+1} = X_{k+2} = \ldots = X_\tau = 0) =$$

$$= \frac{n}{N}\frac{n+c}{N+c}\frac{n+2c}{N+2c}\cdots$$

$$\cdot\frac{n+(k-1)\,c}{N+(k-1)\,c}\frac{s}{N+kc}\cdots\frac{s+(\tau-k-1)\,c}{N+(\tau-1)\,c}$$

Using the factor $\binom{\tau}{k}$ to account for the number of sequences containing k white balls we get the probability function for the Polya distribution

$$\Pr(Y_\tau = k) = \binom{\tau}{k}\frac{n}{N}\frac{n+c}{N+c}\frac{n+2c}{N+2c}\cdots$$

$$\cdot\frac{n+(k-1)\,c}{N+(k-1)\,c}\frac{s}{N+kc}\cdots\frac{s+(\tau-k-1)\,c}{N+(\tau-1)\,c} =$$

$$= \binom{\tau}{k}\frac{\prod_{i=1}^{k} n+i-1\prod_{j=k}^{n-1} s+\tau-1-j}{\prod_{q=1}^{n} N+\tau-q}$$

2.1.A

where $N = n + s$. Notice that, for $c > 0$, as the total number of balls in the urn increases the effect of adding new balls diminishes; we say that the strength of learning decreases. Put differently, the strength of learning is affected both by the initial urn size N, and the number of balls we add each drawing, c. Thus, for given c, a large initial urn N implies weak learning effects. Y_τ is Polya distributed with parameters n, s, τ and c. The Polya distribution encompasses some well known distributions, for example

- $c = 0 : Y_\tau \in bin\left(\tau, \dfrac{n}{N}\right)$

- $c = -1 : Y_\tau \in hyp\left(N, \tau, \dfrac{n}{N}\right)$

The expectation and variance of the Polya distribution is given by

- $E\left(Y_\tau\right) = \tau \dfrac{n}{N}$

- $Var\left(Y_\tau\right) = \tau \dfrac{n}{N} \dfrac{s}{N} \dfrac{N + \tau c}{N + c}$

We now turn to the beta-binomial representation of the Polya distribution; it is defined for $c > 0$ and yields the same probabilities as Equation 2.1.A. In addition to providing a closed form expression for the probability distribution, it has the advantage of illuminating the Bayesian approach employed by the representative agent (to be introduced when we present the economic model in Section 2.2). To see the relation to the urn model, suppose that the agent does not know for sure the composition of the initial urn. His uncertainty is reflected by a probability distribution, a prior distribution $P(\pi)$, over probabilities π of attaining a white ball in the first drawing. After observing the outcome of the first drawing he then updates his probability beliefs by using Bayes rule and obtains his posterior distribution as

$$P\left(\pi \mid X_1\right) = \frac{P\left(X_1 \mid \pi\right) P\left(\pi\right)}{\int P\left(X_1 \mid \pi\right) P\left(\pi\right) d\pi}$$

Here we assume that the prior employd is the beta distribution with parameters $\alpha = n/c$ and $\beta = (N - n)/c$ reflecting the composition of the initial urn and the strength of learning. Since the drawings are Bernoulli trials, the beta distribution is a conjugate prior, which means that the posterior distribution is also a beta distribution . Specifically, with a prior given by $\text{Beta}(\alpha, \beta)$, after observing k white balls in τ drawings, the posterior distribution is $\text{Beta}(\alpha + k, \beta + \tau - k)$ (See DeGroot[1986]). At each point in time, the representative agent uses the mean of the current distribution to estimate the probability of getting a white ball in the next drawing. Thus, initially his probability estimate is given

by $E(\pi) = \alpha/(\alpha + \beta) = n/N$ and his uncertainty is reflected by the variance of the prior beta distribution

$$Var(\pi) = \frac{\alpha\beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}$$

which is decreasing in the urnsize N and increasing in the strength of learning c. Thus, given c, a large initial urn implies extensive prior information so that the probability estimate of the representative agent will be very precise. After observing k white balls in τ drawings his new estimate $p_{k,\tau}$ is given by the mean of a Beta$(\alpha + k, \beta + \tau - k)$ distribution

$$p_{k,\tau} = \frac{\alpha + k}{\alpha + \beta + \tau}$$

Casella and Berger [2002] points out that $p_{k,\tau}$ can be written as a linear combination of initial estimate $\alpha/(\alpha + \beta) = n/N$ and the sample mean, k/τ, that would serve as an estimate if one ignored the prior information. Thus,

$$p_{k,\tau} = \left(\frac{\tau}{\alpha + \beta + \tau}\right)\frac{k}{\tau} + \left(\frac{\alpha + \beta}{\alpha + \beta + \tau}\right)\frac{\alpha}{\alpha + \beta}$$

Recalling that $\alpha = n/c$ and $\beta = (N - n)/c$ and letting $\delta = c/N$ we can write this as

$$p_{k,\tau} = \left(\frac{\tau\delta}{1 + \tau\delta}\right)\frac{k}{\tau} + \left(\frac{1}{1 + \tau\delta}\right)\frac{n}{N}$$

Notice that δ now measures the strength of learning, and as $\delta \to 0$, $p_{k,\tau} \to n/N$, the estimate used if no learning effects were present.

To see how the Polya distribution can be expressed using the beta-binomial representation, we start by assuming that the $X_i : s$ are independent Bernoulli trials with probability of drawing a white ball equal to p, so that the probability of any ordered sequence is $p^k(1-p)^{n-k}$. Letting p be a realization of a stochastic variable $\pi \in [0, 1]$, we can write

$$\Pr(Y_\tau = k \mid \pi = p) = \binom{\tau}{k} p^k (1 - p)^{\tau - k}$$

Letting $f_\pi(p)$ denote the probability density function for π we get, using the law of total probability,

$$\Pr(Y_\tau = k) = \int_0^1 \binom{\tau}{k} p^k (1-p)^{\tau-k} f_\pi(p) dp \qquad \textbf{2.1.B}$$

Assuming that $f_\pi(p)$ is the beta distribution with parameters α and β ($\alpha > 0$ and $\beta > 0$), $\pi \in \text{Beta}(\alpha, \beta)$ that is

$$f_\pi(p \mid \alpha, \beta) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1} & \text{for } 0 < p < 1 \\ \\ 0 & \text{otherwise} \end{cases}$$

expression (2.1.B) becomes

$$\Pr(Y_\tau = k) = \int_0^1 \binom{\tau}{k} p^k (1-p)^{\tau-k} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1} dp =$$

$$= \binom{\tau}{k} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 p^{\alpha+k-1}(1-p)^{\beta+\tau-k-1} dp$$

$$\textbf{2.1.C}$$

Using that

$$\int_0^1 p^{\alpha+k-1}(1-p)^{\beta+\tau-k-1} dp = \frac{\Gamma(\alpha+k)\Gamma(\beta+\tau-k)}{\Gamma(\alpha+\beta+\tau)}$$

and

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

(2.1.C) simplifies to

$$\Pr(Y_\tau = k) = \binom{\tau}{k} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+k)\Gamma(\beta+\tau-k)}{\Gamma(\alpha+\beta+\tau)} =$$

$$= \binom{\tau}{k} \frac{B(\alpha+k, \beta+\tau-k)}{B(\alpha,\beta)}$$

$$\textbf{2.1.D}$$

where $\alpha = n/c$ and $\beta = s/c$. The ratio of beta functions in Equation 2.1.D reflects the assumption that the initial beliefs on the probability of getting a white ball in the first drawing is beta distributed with parameters α and β.

To see that expression 2.1.D is equivalent to expression 2.1.A, divide the numerator and denominator in 2.1.A by c to get

$$\Pr(Y_\tau = k) \;=\; \binom{\tau}{k} \frac{\left(\frac{n}{c}\right)}{\left(\frac{N}{c}\right)} \frac{\left(\frac{n}{c}+1\right)}{\left(\frac{N}{c}+1\right)} \cdot \ldots \cdot \frac{\left(\frac{n}{c}+(k-1)\right)}{\left(\frac{N}{c}+(k-1)\right)} \frac{\left(\frac{s}{c}\right)}{\left(\frac{N}{c}+k\right)} \cdot \ldots$$

$$\ldots \cdot \frac{\left(\frac{s}{c}+(\tau-k-1)\right)}{\left(\frac{N}{c}+(\tau-1)\right)}$$

Letting $\alpha = n/c$ and $\beta = s/c$ we get

$$\Pr(Y_\tau = k) = \binom{\tau}{k} \frac{\alpha}{(\alpha+\beta)} \frac{(\alpha+1)}{[(\alpha+\beta)+1]} \cdot \ldots \cdot \frac{[\alpha+(k-1)]}{[(\alpha+\beta)+(k-1)]} \cdot$$

$$\cdot \frac{\beta}{[(\alpha+\beta)+k]} \ldots \frac{[\beta+(\tau-k-1)]}{[(\alpha+\beta)+(\tau-1)]}$$

2.1.E

Using that

$$\frac{\Gamma(\alpha+k)}{\Gamma(\alpha)} = \alpha(\alpha+1)\ldots(\alpha+k-1)$$

(2.1.E) transforms into

$$\Pr(Y_\tau = k) = \binom{\tau}{k} \frac{\dfrac{\Gamma(\alpha+k)}{\Gamma(\alpha)} \dfrac{\Gamma(\beta+\tau-k)}{\Gamma(\beta)}}{\dfrac{\Gamma(\alpha+\beta+\tau)}{\Gamma(\alpha+\beta)}} =$$

$$= \binom{\tau}{k} \frac{\Gamma(\alpha+k)\,\Gamma(\beta+\tau-k)\,\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(\beta)\,\Gamma(\alpha+\beta+\tau)}$$

which equals expression (2.1.D), showing that the Polya distribution can be expressed with the beta distribution, where the parameters reflect the initial distribution.

The Polya distribution can be used to model how the probabilities evolve in a binomial tree by interpreting "the probability of getting k white balls in τ drawings" as "the probability of getting k upsteps in τ steps".

As an example, consider an individual whose beliefs on the up and down movements of, say, the price of a stock, are characterized by Bayesian

learning. Let his initial beliefs be given by n and N, so that initially he believes that the probability that the stock price goes up in the next period is $\pi = n/N$. After having experienced the outcome he will revise his probability beliefs in accordance with the parameter c which, given N, is taken to measure the strength of learning. Thus, if an upstep occurs, Bayesian updating of his prior implies that he now believes the probability of an upstep to be $\pi_U = (n+c)/(N+c)$. Similarly, if the first step is a downstep he will take the probability of an upstep in the second period to be $\pi_D = n/(N+c)$. Alternatively, these probabilities can be obtained using the beta-binomial representation by noting that π_U is the expectation of a $Beta\left(\frac{n}{c}+1, \frac{N-n}{c}\right)$ distributed random variable and π_D is the expectation of a $Beta\left(\frac{n}{c}, \frac{N-n}{c}+1\right)$ distributed random variable. Figure 2.1.A depicts the conditional probabilities at each node for a two-period tree when $n = 6$, $N = 10$ and $c = 1$. Note that the strength of learning in this setting also depends on the size of the urn initially, N, and that the effect of the updating will diminish as more balls are added.
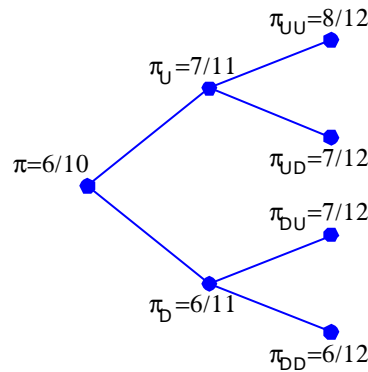
Figure 2.1.A: *Conditional probabilities of upstep when $N = 10$,*
$n = 6$ and $c = 1$

It is clear that when we have Bayesian learning, so that the true probability distribution has $c > 0$ and constant probabilities are assumed, as in the BS model, inaccurately setting $c = 0$ will underestimate the

variance, something that empirical studies of biases of the BS-formula have shown to be the case. In terms of the Polya urn, if we replace the ball drawn and add one ball of the same color, i.e. if we indeed have a Polya distribution with the parameter $c = 1$, and then calculate the probability of getting k white balls/up-steps in τ drawings by using the binomial distribution, we will underestimate the probabilities of getting extreme values; this is the consequence of ignoring the Bayesian learning effects.

Figure 2.1.B shows two (continuous representations of) Polya distributions with $c = 0$ and $c = 1$, respectively. The graphs show the probabilities of getting white balls in 100 drawings when the urn initially contains 30 white balls and 20 black balls. Although the expected number of white balls is 60 for both distributions, the learning distribution with $c = 1$ displays much fatter tails.
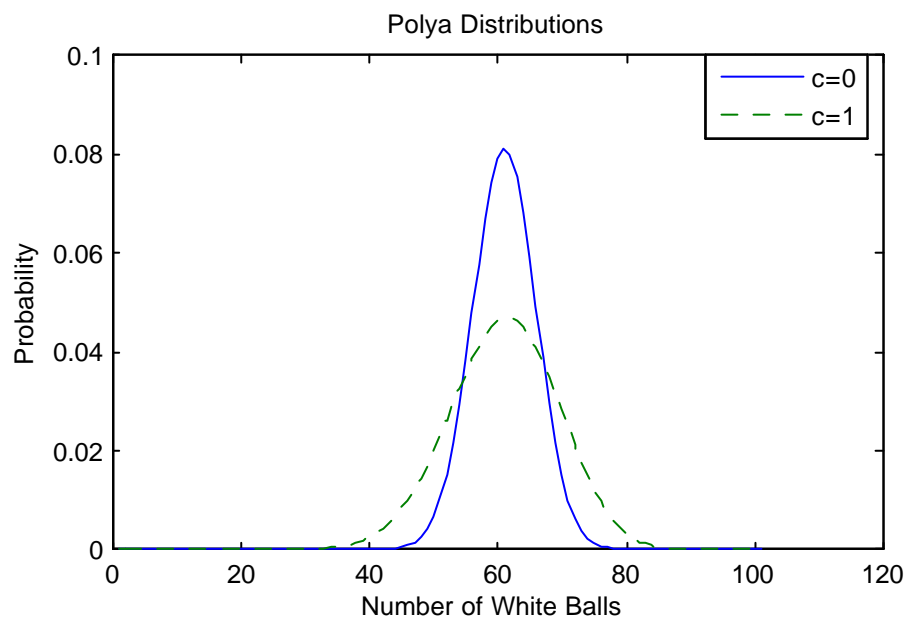
Figure 2.1.B: *Polya distributions with $\tau = 100, n = 30$, $s = 20$ and c is zero and one, respectively*

Bayesian Learning and Recombinant Trees

A recombinant tree is a process such that the probabilities are path independent, so that the probability of getting a white ball/upstep followed

by a blackball/downstep equals the probability of the reversed sequence.

To illustrate that Bayesian learning induces a recombinant tree, we note that for the first two drawings

$$\Pr(X_1 = 1,\, X_2 = 0) = \quad \Pr(X_1 = 1)\Pr(X_2 = 0 \mid X_1 = 1) =$$
$$= \frac{n}{N}\frac{s}{N+c}$$

and

$$\Pr(X_1 = 0,\, X_2 = 1) = \quad \Pr(X_1 = 0)\Pr(X_2 = 1 \mid X_1 = 0) =$$
$$= \frac{s}{N}\frac{n}{N+c}$$

so that

$$\Pr(X_1 = 1, X_2 = 0) = \Pr(X_1 = 0, X_2 = 1)$$

For a formal proof start by considering an urn which initially contains n white balls and s black balls. A sequence of $\tau - 2$ drawings in which a white ball has been drawn l times will thus contain $n + lc$ white balls and a total number of balls equal to $N + (\tau - 2)c$; denote the probability of this composition after $\tau - 2$ drawings by $\Pr(A)$. Now there are two ways for the urn to contain $n + lc + c$ white balls after τ drawings, given the composition after $\tau - 2$ drawings; getting a white ball in drawing $\tau - 1$ and a black ball in drawing τ, or the other way around. The first probability we can write as

$$\Pr(A)\Pr(X_{\tau-1} = 1, X_\tau = 0 \mid A) =$$
$$= \Pr(A)\, \frac{(n + lc)}{N + (\tau - 2)c}\frac{[N + (\tau - 1)c] - (n + lc + c)}{N + (\tau - 1)c} = \qquad \textbf{2.1.F}$$
$$= \Pr(A)\frac{(n + lc)[N + (\tau - 2)c] - (n + lc)^2}{[N + (\tau - 2)c][N + (\tau - 1)c]}$$

Similarly, the probability of getting a black ball in drawing $\tau - 1$ and white ball in drawing τ, is

$$\Pr(A)\Pr(X_{\tau-1} = 0, X_\tau = 1 \mid A) =$$
$$= \Pr(A)\frac{[N + (\tau - 2)c] - (n + lc)}{[N + (\tau - 2)c]}\frac{(n + lc)}{[N + (\tau - 1)c]} \qquad \textbf{2.1.G}$$

The proof is completed by noting that the rightmost members of (2.1.F) and (2.1.G) are equal.

2.2   The Lucas Model

The point of departure is the infinite horizon model studied by Lucas [1978]. Thus, we consider a representative agent economy where output each period is exogenous and perishable. In equilibrium consumption must then equal output and asset prices can be found by solving the consumers problem.

The Dividend Process

There are two assets in the economy, a single-period, risk-free discount bond in zero net supply trading at a price of B and earning interest of $r = (1/B) - 1$, making the interest rate $f = \ln(1/B)$, and a stock traded at ex-dividend price S_t in net supply of one. Following Lucas [1978] the stock is a claim at an infinite stream of perishable, real dividends. The first dividend, coming at date t, will be denoted by d_t, while a subsequent dividend at node σ in the binomial lattice will be denoted d_σ. To identify the nodes, we will use Σ to denote the set of nodes, while $\Sigma_{t+\tau}$ refers to the set of nodes at date $t + \tau$, so that

$$\Sigma = \Sigma_t \cup \Sigma_{t+1} \cup \ldots \cup \Sigma_{t+\tau} \cup \ldots$$

The set of nodes at date $t + \tau$ itself can be partitioned, so that

$$\Sigma_{t+\tau} = \Sigma_{t+\tau}^0 \cup \Sigma_{t+\tau}^1 \cup \ldots \cup \Sigma_{t+\tau}^k \cup \ldots \cup \Sigma_{t+\tau}^\tau$$

where $\Sigma_{t+\tau}^k$ is the set of nodes at date $t + \tau$ with k upsteps, $0 \leq k \leq \tau$.

At each point in time the growth rate will be either $U - 1$ if an upstep occurs or $D - 1$ if a downstep occurs, so that $d_\sigma = d_t U^k D^{\tau-k}$ for $\sigma \in \Sigma_{t+\tau}^k$.

When identifying nodes in time, we will use $\mathbb{T} = \{0, 1, 2, \ldots\}$ to denote the set of dates and $\mathbb{T}_1 = \{1, 2, \ldots\}$ to denote the set of future dates so that $\sigma \in \cup_{\tau \in \mathbb{T}_1} \Sigma_{t+\tau}$ will refer to some node other than the initial node.

The Consumer Problem

Given the dividend process, the asset prices are determined in equilibrium by the representative agent's first order conditions. The representative agent maximizes expected utility of an infinite stream of dividends. Because of the equilibrium setting, there exists a unique and positive vector of state prices, normalized so that $\beta_t = 1$, with which we can state the maximization problem

$$\text{Max}_c \quad \sum_{\tau=0}^{\infty} \sum_{\sigma \in \Sigma_{t+\tau}} \lambda^\tau u(c_\sigma) P_\sigma$$

$$\text{s.t} \quad \sum_{\tau=0}^{\infty} \sum_{\sigma \in \Sigma_{t+\tau}} \beta_\sigma c_\sigma \leq \sum_{\tau=0}^{\infty} \sum_{\sigma \in \Sigma_{t+\tau}} \beta_\sigma d_\sigma$$

where

c_σ is consumption at node σ
λ is the subjective discount factor
P_σ is the probability of reaching node σ
β_σ is the state price of node σ
d_σ is the dividend at node σ

A note on terminology is appropriate : some authors reserve the term state price for the one period model and use the term event prices for multi period models. Here the state price β_σ is used to denote the date t value of receiving 1 crown at any node $\sigma \in \Sigma$.

Assuming that the consumer problem has a solution, and using the market balance condition $c_\sigma = d_\sigma$, we get the first order conditions

$$\begin{cases} u'(d_t) - \gamma = 0 & \text{for } \sigma \in \Sigma_t \\ \lambda^\tau u'(d_\sigma) P_\sigma - \gamma \beta_\sigma = 0 & \text{for } \sigma \in \cup_{\tau \in \mathbb{T}_1} \Sigma_{t+\tau} \end{cases} \qquad \textbf{2.2.A}$$

State Prices

Using relation (2.2.A) we get $\gamma = u'(d_t)$ and can solve for the state prices

$$\beta_\sigma = \frac{u'(d_\sigma)}{u'(d_t)} \lambda^\tau P_\sigma \quad \text{for } \sigma \in \cup_{\tau \in \mathbb{T}_1} \Sigma_{t+\tau}$$

State prices; CRRA utility function

Assume that the representative agent's utility function exhibits constant relative risk aversion and has preferences given by

$$
u(x) = \begin{cases} \dfrac{x^{1-\alpha}}{1-\alpha} & \text{if } \alpha \geq 0, \alpha \neq 1 \\[3mm] \ln x & \text{if } \alpha = 1 \end{cases}
$$

Then

$$
Du(x) = x^{-\alpha}
$$

so that

$$
\beta_\sigma = \frac{d_\sigma^{-\alpha}}{d_t^{-\alpha}} \lambda^\tau P_\sigma = \left[U^k D^{\tau-k} \right]^{-\alpha} \lambda^\tau P_\sigma
$$

Probabilities in case of Full Information

In the Full Information case the probabilities are constant, and with π being the probability of an upstep and $(1-\pi)$ the probability of a downstep we get, for $\sigma \in \Sigma_{t+\tau}^k$

$$
P_\sigma = \pi^k (1-\pi)^{\tau-k}
$$

Summing over all nodes at date $t+\tau$ with k upsteps we get

$$
\sum_{\sigma \in \Sigma_{t+\tau}^k} P_\sigma = \binom{\tau}{k} \pi^k (1-\pi)^{\tau-k}
$$

which is the probability of getting k upsteps until date $t+\tau$, i.e.

$$
\Pr \left\{ d_\sigma = d_t U^k D^{\tau-k} \right\}
$$

Probabilities in case of learning

In the learning case, the probability is a function of the number of upsteps, n, and the total number of steps N up to date t, constituting the prior belief of the representative agent, as well as the strength of learning c

$$
P_\sigma = \frac{B(n/c + k,\ (N-n)/c + \tau - k)}{B(n/c,\ (N-n)/c)} \quad \text{for } \sigma \in \Sigma_{t+\tau}^k
$$

so that the probability of getting $d_\sigma = d_t U^k D^{\tau-k}$ is

$$\sum_{\sigma \in \Sigma^k_{t+\tau}} P_\sigma = \binom{\tau}{k} \frac{B(n/c + k, \ (N-n)/c + \tau - k)}{B(n/c, \ (N-n)/c)} \text{ for } \sigma \in \Sigma^k_{t+\tau}$$

State prices for a constant risk aversion utility function with Full Information

In this case

$$P_\sigma = \pi^k (1-\pi)^{\tau-k} \text{ for } \sigma \in \Sigma^k_{t+\tau}$$

It follows that the state price is

$$\beta_\sigma = \frac{d_\sigma^{-\alpha}}{d_t^{-\alpha}} \lambda^\tau \pi^k (1-\pi)^{\tau-k} \ \ \text{for } \sigma \in \Sigma^k_{t+\tau} \qquad \textbf{2.2.B}$$

State prices for a constant risk aversion utility function with learning

In this case

$$P_\sigma = \frac{B(n/c + k, \ (N-n)/c + \tau - k)}{B(n/c, \ (N-n)/c)} \ \ \text{for } \sigma \in \Sigma^k_{t+\tau}$$

It follows that

$$\beta_\sigma = \frac{d_\sigma^{-\alpha}}{d_t^{-\alpha}} \lambda^\tau \frac{B(n/c + k, \ (N-n)/c + \tau - k)}{B(n/c, \ (N-n)/c)} \ \ \text{for } \sigma \in \Sigma^k_{t+\tau} \qquad \textbf{2.2.C}$$

2.3   Stock, Bond and Call Price

Assuming that the price of the stock is the discounted value of the dividends, then the ex-dividend, date t price of the stock is

$$S_t = \sum_{\tau=1}^\infty \sum_{\sigma \in \Sigma_{t+\tau}} \beta_\sigma d_\sigma$$

Since the bond is single-period, the date t bond price is

$$B_t = \sum_{\sigma \in \Sigma_{t+1}} \beta_\sigma$$

The payoff to a European call option at the expiration date $t + T$ at some node $\sigma \in \Sigma^k_{t+T}$ is the difference between the stock price S_σ and the exercise price K,whenever this difference is positive

$$(S_\sigma - K)^+ = \max(S_\sigma - K, 0)$$

The current price of a European call option, C_t, is given by discounting this amount with the state prices β_σ and summing over all nodes σ such that $(S_\sigma - K) > 0$. Hence, the date t price of a call option with T periods to expiration and exercise price K is

$$C_t = \sum_{\sigma \in \Sigma_{t+T}} \beta_\sigma (S_\sigma - K)^+$$

The Full Information Case

Given that a transversality condition holds, so that the discounted value of the stock converges to zero, and that $\lambda < 1/\left[D^{1-\alpha} + \pi \left(U^{1-\alpha} - D^{1-\alpha}\right)\right]$, ensuring that the infinite sum of the discounted dividends converges and that the stock price is positive, a closed form expression for the stock price can be found.

Proposition 2.3.A *In the case of Full Information the stock price, bond price, interest rate and call price are given by*

$$S_t^{FI} = d_t \frac{\lambda \left[\pi U^{1-\alpha} + (1-\pi) D^{1-\alpha}\right]}{1 - \lambda \left[\pi U^{1-\alpha} + (1-\pi) D^{1-\alpha}\right]}$$

$$B_t^{FI} = \lambda \left[\pi U^{-\alpha} + (1-\pi) D^{-\alpha}\right]$$

$$f_t^{FI} = \ln \frac{1}{\lambda \left[\pi U^{-\alpha} + (1-\pi) D^{-\alpha}\right]}$$

$$C_t^{FI} = \lambda^T \sum_{k=0}^{T} \binom{T}{k} \left[U^k D^{T-k}\right]^{-\alpha} \pi^k (1-\pi)^{T-k} \left(U^k D^{T-k} S_t^{FI} - K\right)^+$$

 Proof. Stock Price. Assume that the discounted price of the stock converges to 0. Then the stock price at date t, S_t^{FI}, equals the value of the future discounted dividends,

$$S_t^{FI} = \sum_{\tau \in \mathbb{T}_1} \sum_{\sigma \in \Sigma_{t+\tau}} \beta_\sigma d_\sigma = \sum_{\tau \in \mathbb{T}_1} \sum_{\sigma \in \Sigma_{t+\tau}} \left[\frac{u'(d_\sigma)}{u'(d_t)} \lambda^\tau P_\sigma\right] d_\sigma =$$

$$= \frac{1}{d_t^{-\alpha}} \sum_{\tau \in \mathbb{T}_1} \sum_{\sigma \in \Sigma_{t+\tau}} d_\sigma^{-\alpha} \lambda^\tau P_\sigma d_\sigma$$

Since, for $\sigma \in \Sigma_{t+\tau}^k$

$$d_\sigma = d_t U^k D^{\tau-k}$$

and

$$P_\sigma = \pi^k (1 - \pi)^{\tau - k}$$

we get

$$S_t^{FI} = \frac{1}{d_t^{-\alpha}} \sum_{\tau=1}^\infty \lambda^\tau \left\{ \sum_{k=0}^\tau \binom{\tau}{k} \left[d_t U^k D^{\tau-k} \right]^{1-\alpha} \pi^k (1 - \pi)^{\tau - k} \right\} =$$

$$= \frac{1}{d_t^{-\alpha}} d_t^{1-\alpha} \sum_{\tau=1}^\infty \lambda^\tau \left\{ \sum_{k=0}^\tau \binom{\tau}{k} \left[U^{1-\alpha} \pi \right]^k \left[D^{1-\alpha} (1 - \pi) \right]^{\tau - k} \right\}$$

Put $a = U^{1-\alpha} \pi$ and $b = D^{1-\alpha} (1 - \pi)$. Then

$$S_t^{FI} = d_t \sum_{\tau=1}^\infty \lambda^\tau \left\{ \sum_{k=0}^\tau \binom{\tau}{k} a^k b^{\tau-k} \right\} =$$

$$= d_t \sum_{\tau=1}^\infty \lambda^\tau \left\{ \sum_{k=0}^\tau \binom{\tau}{k} \left[\frac{(a+b)\, a}{(a+b)} \right]^k \left[\frac{(a+b)\, b}{(a+b)} \right]^{\tau-k} \right\} =$$

$$= d_t \sum_{\tau=1}^\infty \lambda^\tau \left\{ (a+b)^\tau \sum_{k=0}^\tau \binom{\tau}{k} \left[\frac{a}{a+b} \right]^k \left[\frac{b}{a+b} \right]^{\tau-k} \right\} =$$

$$= d_t \sum_{\tau=1}^\infty \lambda^\tau (a+b)^\tau$$

since

$$\sum_{k=0}^\tau \binom{\tau}{k} \left[\frac{a}{a+b} \right]^k \left[\frac{b}{a+b} \right]^{\tau-k} = 1$$

for $\tau \in \mathbb{T}_1$.

Finally, insert the expressions for a and b to get

$$S_t^{FI} = d_t \sum_{\tau=1}^\infty \lambda^\tau (a+b)^\tau$$

$$= d_t \frac{\lambda (a+b)}{1 - \lambda (a+b)}$$

$$= d_t \frac{\lambda \left[\pi U^{1-\alpha} + (1 - \pi) D^{1-\alpha} \right]}{1 - \lambda \left[\pi U^{1-\alpha} + (1 - \pi) D^{1-\alpha} \right]}$$

where in the second step we have used the condition

$$\lambda \left[\pi U^{1-\alpha} + (1 - \pi) D^{1-\alpha} \right] < 1$$

, to ensure that the sum converges.

The Bond Price. Recall that the state prices at date $t+1$ are given by

$$\lambda \pi U^{-\alpha} \text{ for } \sigma \in \Sigma^1_{t+1}$$

and

$$\lambda(1-\pi)D^{-\alpha} \text{ for } \sigma \in \Sigma^0_{t+1}$$

Hence

$$
\begin{aligned}
B_t^{FI} = \quad & \lambda \pi U^{-\alpha} + \lambda(1-\pi)D^{-\alpha} = \\
& \lambda \left[\pi U^{-\alpha} + (1-\pi)D^{-\alpha} \right] = \\
& \lambda \left[D^{-\alpha} + \pi \left(U^{-\alpha} - D^{-\alpha} \right) \right]
\end{aligned}
$$

Interest rate. With interest defined as $r = (1/B_t^{FI}) - 1$, the interest rate f_t^{FI} becomes

$$
\begin{aligned}
f_t^{FI} = \quad & \ln(1+r) = \ln \frac{1}{B_t^{FI}} = \\
& = \ln \frac{1}{\lambda \left[D^{-\alpha} + \pi \left(U^{-\alpha} - D^{-\alpha} \right) \right]}
\end{aligned}
$$

The Call Price. Since the stock price is homogenous in dividends the stock price at a node $\sigma \in \Sigma^k_{t+T}$ is

$$S_\sigma^{FI} = S_t^{FI} U^k D^{T-k}$$

Discounting the positive payoffs $\left(S_\sigma^{FI} - K \right)^+$ at exercise date $t+T$ by the state prices yields the date t call price

$$
\begin{aligned}
C_t^{FI} \quad = \quad & \sum_{\sigma \in \Sigma_{t+T}} \beta_\sigma \left(S_\sigma^{FI} - K \right)^+ \\
= \quad & \lambda^T \sum_{k=0}^{T} \binom{T}{k} \left[U^k D^{T-k} \right]^{-\alpha} \pi^k (1-\pi)^{T-k} \left(U^k D^{T-k} S_t^{FI} - K \right)^+
\end{aligned}
$$

$$\square$$

Due to the Lucas setting and the constant relative risk aversion by the representative agent, the FI stock price at date t will be linearly homogenous in d_t.

Since the probability is constant over time, so is the FI bond price. This means that the present value of a crown received at a future date is independent of the sequence of high and low growth of the dividends.

The Learning Case

In the learning case, the probability of an upstep at date t is given by the prior $\pi_t = n/N$. If an upstep were to occur the next period, the representative agent with strength of learning c will revise his beliefs and estimate the probability to $\pi_{t+1} = (n + c) / (N + c)$. The distinguishing feature of the learning model is that the representative agent's perception of probability distributions over future dividends is based on Bayesian updating and the realizations at the current date t that this updating is done recursively over time. Hence, the agent not only revises his probability beliefs over time, but also accounts for these future revisions at date t. As shown in Section 2.1, when the representative agent views the probability as a beta distributed random variable with parameters n/c and $(N - n)/c$ then he beliefs that probability of getting k occurrences of high growth in τ periods is

$$P_\sigma = \binom{\tau}{k} \frac{B(n/c + k, \ (N - n)/c + \tau - k)}{B(n/c, \ (N - n)/c)} \ \text{ for } \sigma \in \Sigma_{t+\tau}^k$$

Proposition 2.3.B presents the implications for asset prices of taking learning into account. In Proposition 2.3.B, let $G_\tau = \left[U^k D^{\tau-k}\right]^{1-\alpha}$ It follows that the expectation of G_τ at date t is

$$E_t \left[G_\tau \middle| n, N\right] = \sum_{k=0}^{\tau} \left(U^{1-\alpha}\right)^k \left(D^{1-\alpha}\right)^{\tau-k} \binom{\tau}{k} \frac{B(n/c + k, \ (N - n)/c + \tau - k)}{B(n/c, \ (N - n)/c)}$$

It is convenient to define the infinite sum of discounted expectations of G_τ. Thus

$$\psi_t \left(n, N\right) = \sum_{\tau=1}^{\infty} \lambda^\tau E_t \left[G_\tau \middle| n, N\right]$$

Note that the sufficient condition used in Proposition 2.3.B ensure convergence of the stock price independent of the initial composition of the urn. Thus we can safely assume that the prior is given by n/N.

Proposition 2.3.B *Assume that* $\lambda \max \left(U^{1-\alpha}, D^{1-\alpha} \right) < 1$. *Then*

$$S_t^{BL} = d_t \psi_t (n, N)$$

$$B_t^{BL} = \lambda \left[D^{-\alpha} + \pi_t [U^{-\alpha} - D^{-\alpha}] \right]$$

$$f_t^{BL} = -\ln B_t^{BL} = \ln \frac{1}{\lambda \left[D^{-\alpha} + \pi_t [U^{-\alpha} - D^{-\alpha}] \right]}$$

$$C_t^{BL} = \lambda^T \sum_{i=0}^{T} \binom{T}{i} (U^{-\alpha})^i (D^{-\alpha})^{T-i} \frac{B(n/c + i, \ (N-n)/c + T - i)}{B(n/c, \ (N-n)/c)} \times$$

$$\left(U^i D^{T-i} S_t^{BL} \frac{\psi_{t+T} (n + ic, N + Tc)}{\psi_t (n, N)} - K \right)^+$$

Proof. Stock Price. Assume that the discounted price of the stock converges to 0. Then the stock price at date t, S_t^{BL}, equals the value of the future discounted dividends

$$S_t^{BL} = \sum_{\tau \in \mathbb{T}_1} \sum_{\sigma \in \Sigma_{t+\tau}} \beta_\sigma d_\sigma = \sum_{\tau \in \mathbb{T}_1} \sum_{\sigma \in \Sigma_{t+\tau}} \left[\frac{u'(d_\sigma)}{u'(d_t)} \lambda^\tau P_\sigma \right] d_\sigma =$$

$$= \frac{1}{d_t^{-\alpha}} \sum_{\tau \in \mathbb{T}_1} \sum_{\sigma \in \Sigma_{t+\tau}} d_\sigma^{-\alpha} \lambda^\tau P_\sigma d_\sigma$$

Since, for $\sigma \in \Sigma_{t+\tau}^k$, $\tau \in \mathbb{T}_1$ we have $d_\sigma = d_t U^k D^{\tau-k}$ and it follows that

$$S_t^{BL} = \frac{1}{d_t^{-\alpha}} \sum_{\tau=1}^{\infty} \lambda^\tau \left\{ \sum_{k=0}^{\tau} \left[d_t U^k D^{\tau-k} \right]^{1-\alpha} \binom{\tau}{k} P_\sigma \right\} =$$

$$= d_t \sum_{\tau=1}^{\infty} \lambda^\tau \left\{ \sum_{k=0}^{\tau} \left[U^{1-\alpha} \right]^k \left[D^{1-\alpha} \right]^{\tau-k} \binom{\tau}{k} P_\sigma \right\}$$

$$= d_t \sum_{\tau=1}^{\infty} \lambda^\tau E \left[G_\tau | n, N \right]$$

$$= d_t \psi_t (n, N)$$

It remains to show that the condition $\lambda \max \left(U^{1-\alpha}, D^{1-\alpha} \right) < 1$ guarantees that S_t^{BL} is finite. Hence

$$
\begin{aligned}
S_t^{BL} = \ & d_t \sum_{\tau=1}^{\infty} \lambda^\tau \left\{ \sum_{k=0}^{\tau} [U^{1-\alpha}]^k \left[D^{1-\alpha} \right]^{\tau-k} \binom{\tau}{k} P_\sigma \right\} \\
\leq \ & d_t \sum_{\tau=1}^{\infty} \lambda^\tau \left\{ \sum_{k=0}^{\tau} \left[\max \left(U^{1-\alpha}, D^{1-\alpha} \right) \right]^\tau \binom{\tau}{k} P_\sigma \right\} \\
\leq \ & d_t \sum_{\tau=1}^{\infty} \left[\lambda \max \left(U^{1-\alpha}, D^{1-\alpha} \right) \right]^\tau \left\{ \sum_{k=0}^{\tau} \binom{\tau}{k} P_\sigma \right\} \\
\leq \ & d_t \frac{\lambda \max \left(U^{1-\alpha}, D^{1-\alpha} \right)}{1 - \lambda \max \left(U^{1-\alpha}, D^{1-\alpha} \right)}
\end{aligned}
$$

Bond price. Recalling that for $\sigma \in \Sigma_{t+1}$, the state prices will be

$$
\beta_\sigma = \ \lambda \pi_t U^{-\alpha} \text{ for } \sigma \in \Sigma_{t+1}^1
$$

and

$$
\beta_\sigma = \ \lambda (1 - \pi_t) D^{-\alpha} \text{ for } \sigma \in \Sigma_{t+1}^0
$$

Hence

$$
\begin{aligned}
B_t^{BL} = \ & \lambda [U^{-\alpha} \pi_t + D^{-\alpha}(1 - \pi_t)] = \\
= \ & \lambda \left\{ D^{-\alpha} + \pi_t [U^{-\alpha} - D^{-\alpha}] \right\}
\end{aligned}
$$

Interest rate. With interest defined as $r = (1/B_t^{BL}) - 1$, the interest rate f_t^{BL} becomes

$$
f_t^{BL} = \ \ln(1 + r) = \ln \frac{1}{B_t^{BL}} = \ln \frac{1}{\lambda \left[D^{-\alpha} + \pi_t [U^{-\alpha} - D^{-\alpha}] \right]}
$$

Call price. The date t price of a call option with T periods to expiration and exercise price K is given by discounting the positive payoffs at date $t + T$ by the state prices and sum over all nodes

$$
C_t^{BL} = \sum_{\sigma \in \Sigma_{t+T}} \beta_\sigma \left(S_\sigma^{BL} - K \right)^+
$$

where the state prices β_σ used to compute the date t value of the payoffs T periods ahead assume the values

$$
\lambda^T \left(U^{-\alpha} \right)^i \left(D^{-\alpha} \right)^{T-i} \frac{B(n/c + i, \ (N - n)/c + T - i)}{B(n/c, \ (N - n)/c)} \quad , \ 0 \leq i \leq T
$$

At expiration date the composition of the urn at a node $\sigma \in \Sigma_{t+T}^i$ reached after i upsteps is $(n + ic, N + Tc)$. Hence the stock prices S_σ^{BL} distributed over the nodes at expiration date $t + T$ assume the values

$$d_t U^i D^{T-i} \sum_{\tau=1}^{\infty} \lambda^\tau E_{t+T} \left[G_\tau | \ n + ic, N + Tc \right]$$

$$= d_t U^i D^{T-i} \psi_{t+T} \left(n + ic, N + Tc \right), \ 0 \le i \le T$$

Using that $d_t = S_t^{BL}/\psi_t \left(n, N \right)$, the stock prices at expiration date can be rewritten as

$$U^i D^{T-i} S_t^{BL} \frac{\sum\limits_{\tau=1}^{\infty} \lambda^\tau E_{t+T} \left[G_\tau | \ n + ic, N + Tc \right]}{\sum\limits_{\tau=1}^{\infty} \lambda^\tau E_t \left[G_\tau | \ n, N \right]}$$

$$= U^i D^{T-i} S_t^{BL} \frac{\psi_{t+T} \left(n + ic, N + Tc \right)}{\psi_t \left(n, N \right)}, \ 0 \le i \le T$$

Discounting the payoffs at exercise date with the stateprices we thus get the date t call price C_t^{BL} as

$$\lambda^T \sum_{i=0}^{T} \binom{T}{i} \left(U^{-\alpha} \right)^i \left(D^{-\alpha} \right)^{T-i} \frac{B(n/c + i, \ (N - n)/c + T - i)}{B(n/c, \ (N - n)/c)}$$

$$\times \left(U^i D^{T-i} S_t^{BL} \frac{\psi_{t+T} \left(n + ic, N + Tc \right)}{\psi_t \left(n, N \right)} - K \right)^+$$

$\square$

If $\alpha > 1$ then the model implies adverse pricing (Cf Abel [1988]), that is the stock price increases when the probability of a downstep increases. To avoid this intuitively unappealing case we restrict ourselves to study the case when $\alpha < 1$, in which case $U^{1-\alpha} > D^{1-\alpha}$. Thus the convergence condition in Proposition 2.3.B becomes $\lambda U^{1-\alpha} < 1$.

A key feature of the stock price with learning is that the stock-dividend ratio is no longer constant as was the case with FI model. In the learning case, in addition to the direct effect of dividends, there is an indirect effect of dividends on the stock price since it changes the probability beliefs reflected in the ψ function.

The BL bond price varies and depends at each date on the probability beliefs of the representative agent; since these are decided by the number of times high and low growth has occurred this means that the bond price varies not only over time but also, for each date $t + \tau$, across the nodes $\sigma \in \Sigma_{t+\tau}$.

The pricing of bonds with learning affects the discounting of future sure income in the following way. The present value of a crown received at future date depends not only on the number of occurrences of high and low growth, but also on the sequence that these occur. This path dependence has consequences for the computation of the risk neutral probabilities, as will be seen in Section 2.4.

Compared to the Full Information case, learning affects the call price in two ways. First, learning causes the highest stock price at exercise day to exceed the highest stock price in the FI case, while the lowest stock price with learning is lower than the lowest stock price in the FI case, that is, learning causes the support to widen. Second, the fatter tails of the learning distribution carries over to the state prices used to discount the payoffs at exercise day, so that the BL state prices discounting the high payoffs will exceed the corresponding FI state prices.

Approximation of BL Stock Price

With a prior given by (n, N), computing the stock price with Bayesian learning involves computing the expected value

$$
E_t \left[G_\tau \middle| n, N \right] = \sum_{k=0}^{\tau} \left(U^{1-\alpha} \right)^k \left(D^{1-\alpha} \right)^{\tau-k} \binom{\tau}{k} \frac{B(n/c + k, \ (N-n)/c + \tau - k)}{B(n/c, \ (N-n)/c)}
$$

for all integers $\tau \in [1, I]$, where I is a number large enough to guarantee convergence of the stock price. With daily data this is a demanding task given the computer power currently available. To address this issue we thus seek an approximation of $E_t \left[G_\tau \middle| n, N \right]$. The approximation results in a closed form expression for the stock price with Bayesian learning which not only is conveniant from a computational standpoint, but also facilitates analytic tractability. We start by letting $D = 1/U$ and consider

a second degree Taylor expansion of

$$G_\tau^* (k) = \left[U^{1-\alpha} \right]^{2k-\tau}$$

at some point y. Treating the number of upsteps k as a continuous variable and letting $\gamma = 2 \log U^{1-\alpha}$, the first and second derivatives are

$$DG_\tau^* (k) = \gamma G_\tau^* (k)$$

$$D^2 G_\tau^* (k) = \gamma^2 G_\tau^* (k)$$

and we arrive at the Taylor expansion at y by performing the computation

$$
\begin{aligned}
G_\tau^* (k) &= G_\tau^* (y + (k - y)) \\
&\approx G_\tau^* (y) + DG_\tau^* (y) (k - y) + \frac{1}{2} D^2 G_\tau^* (y) (k - y)^2 \\
&= G_\tau^* (y) + \gamma G_\tau^* (y) (k - y) + \frac{1}{2} \gamma^2 G_\tau^* (y) (k - y)^2 \\
&= G_\tau^* (y) \left[1 + \gamma (k - y) + \frac{1}{2} \gamma^2 (k - y)^2 \right] \\
&= \tilde{G}_\tau^* (k)
\end{aligned}
$$

Taking the expectation of the Taylor approximation $\tilde{G}_\tau^* (k)$ with respect to some, so far unspecified, measure yields

$$
\begin{aligned}
E_t \left[\tilde{G}_\tau^* (k) \right] &= E_t \left[\tilde{G}_\tau^* (y) \left(1 + \gamma (k - y) + \frac{1}{2} \gamma^2 (k - y)^2 \right) \right] \\
&= \tilde{G}_\tau^* (y) \left(E_t [1] + \gamma E_t [k - y] + \frac{1}{2} \gamma^2 E_t \left[(k - y)^2 \right] \right) \\
&= \tilde{G}_\tau^* (y) \left(1 + \gamma E_t [k - y] + \frac{1}{2} \gamma^2 E_t \left[(k - y)^2 \right] \right)
\end{aligned}
$$

2.3.A

Since

$$
\begin{aligned}
E_t \left[(k - y)^2 \right] &= E_t [k^2 - 2ky + y^2] \\
&= E_t [k^2] - 2y E_t [k] + y^2 \\
&= V (k) + (E_t [k])^2 - 2y E_t [k] + y^2
\end{aligned}
$$

we get that $E_t\left[\tilde{G}_\tau^*(k)\right]$ equals

$$\tilde{G}_\tau^*(y)\left(1+\gamma E_t\left[k-y\right]+\frac{1}{2}\gamma^2\left[V(k)+(E_t[k])^2-2yE_{\hat{p}}[k]+y^2\right]\right)$$

Now, let probability measure be the measure with Bayesian learning and recall from Section 2.1 that the expectation and variance of a Polya distributed variable k with parameters n, N, τ and c are

$$E_t(k) \;=\; \tau\frac{n}{N}$$

$$V_t(k) \;=\; \tau\frac{n}{N}\left(1-\frac{n}{N}\right)\frac{N+\tau c}{N+c}$$

Then, if we let $y=E_t(k)=\tau\dfrac{n}{N}$, 2.3.A reduces to

$$E_t\left[\tilde{G}_\tau^*(k)\right] \;=\; \tilde{G}_\tau^*(y)\left(1+\frac{1}{2}\gamma^2V_{\hat{p}}(k)\right)$$

$$\;=\; \tilde{G}_\tau^*(y)\left(1+\frac{1}{2}\gamma^2\tau\frac{n}{N}\left(1-\frac{n}{N}\right)\frac{N+\tau c}{N+c}\right)$$

Finally, using that $\tilde{G}_\tau^*(y)=[U^{1-\alpha}]^{\tau\left(\frac{2n}{N}-1\right)}$ and $\gamma=2\log U^{1-\alpha}$, we get that $E_t\left[\tilde{G}_\tau^*(k)\right]$ equals

$$\tau\left[U^{1-\alpha}\right]^{\tau\left(\frac{2n}{N}-1\right)}\left[\frac{1}{\tau}+2\left[(1-\alpha)U\right]^2\frac{n}{N}\left(1-\frac{n}{N}\right)\frac{N+\tau c}{N+c}\right]$$

Thus, replacing $E_t\left[G_\tau\right]$ in Proposition 2.3.B with $E_t\left[\tilde{G}_\tau^*(k)\right]$ we can write the approximated stock price with Bayesian learning as

$$\tilde{S}_t^{BL}=d_t\left\{\sum_{\tau=1}^\infty\lambda^\tau E_t\left[\tilde{G}_\tau^*(k)\right]\right\}=d_t\left\{\sum_{\tau=1}^\infty\lambda^\tau\tau\left[U^{1-\alpha}\right]^{\tau\left(\frac{2n}{N}-1\right)}\right.$$

$$\left.\times\left[\frac{1}{\tau}+2\left[(1-\alpha)\log U\right]^2\frac{n}{N}\left(1-\frac{n}{N}\right)\frac{N+\tau c}{N+c}\right]\right\}\quad(2.3.B)$$

We now proceed to show that this can be rewritten as a closed form expression. Thus, with a view to eliminate the summation sign in relation 2.3.B we first note that

$$\sum_{t=1}^\infty tx^t \;=\; \frac{x}{(1-x)^2}$$

and

$$\sum_{t=1}^\infty t^2x^t \;=\; \frac{(1+x)x}{(1-x)^3}$$

Letting $\hat{U} = U^{(1-\alpha)\left(\frac{2n}{N}-1\right)}$, the approximated stock price with learning, $\tilde{S}_t^{BL}$, can be rewritten as

$$d_t \left\{ \sum_{\tau=1}^{\infty} \left(\lambda\hat{U}\right)^{\tau} \tau \left[\frac{1}{\tau} + 2\left[(1-\alpha)\log U\right]^2 \frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N+\tau c}{N+c}\right] \right\}$$

$$= d_t \left\{ \sum_{\tau=1}^{\infty} \left(\lambda\hat{U}\right)^{\tau} + \left(\lambda\hat{U}\right)^{\tau} \tau 2\left[(1-\alpha)\log U\right]^2 \frac{n}{N}\left(1 - \frac{n}{N}\right)\frac{N+\tau c}{N+c} \right\}$$

and further calculations give that $\tilde{S}_t^{BL}$ equals

$$d_t \left\{ \sum_{\tau=1}^{\infty} \left(\lambda\hat{U}\right)^{\tau} + \frac{2\left[(1-\alpha)\log U\right]^2 \frac{n}{N}\left(1 - \frac{n}{N}\right)}{N+c} \right.$$
$$\left. \times \sum_{\tau=1}^{\infty} \left[\left(\lambda\hat{U}\right)^{\tau} \tau (N+\tau c)\right] \right\}$$

$$= d_t \left\{ \sum_{\tau=1}^{\infty} \left(\lambda\hat{U}\right)^{\tau} + \frac{2\left[(1-\alpha)\log U\right]^2 \frac{n}{N}\left(1 - \frac{n}{N}\right)}{N+c} \right.$$
$$\left. \times \left[N \sum_{\tau=1}^{\infty} \left(\lambda\hat{U}\right)^{\tau} \tau + c \sum_{\tau=1}^{\infty} \left(\lambda\hat{U}\right)^{\tau} \tau^2\right] \right\}$$

Finally, making use of the formulas on series above,

$$\tilde{S}_t^{BL} = d_t \left\{ \frac{\lambda\hat{U}}{1-\lambda\hat{U}} + \frac{2\left[(1-\alpha)\log U\right]^2 \frac{n}{N}\left(1 - \frac{n}{N}\right)}{N+c} \right.$$
$$\left. \times \left[N\frac{\lambda\hat{U}}{\left(1-\lambda\hat{U}\right)^2} + c\frac{\left(1+\lambda\hat{U}\right)\lambda\hat{U}}{\left(1-\lambda\hat{U}\right)^3}\right] \right\}$$

Comparison: FI, BL and Approximation of BL

If $\alpha = 1$ then the stock price without learning equals the stock price with Bayesian learning as well as its approximation and the common price is

$$S_t^{FI} = S_t^{BL} = \tilde{S}_t^{BL} = d_t\frac{\lambda}{1-\lambda}$$
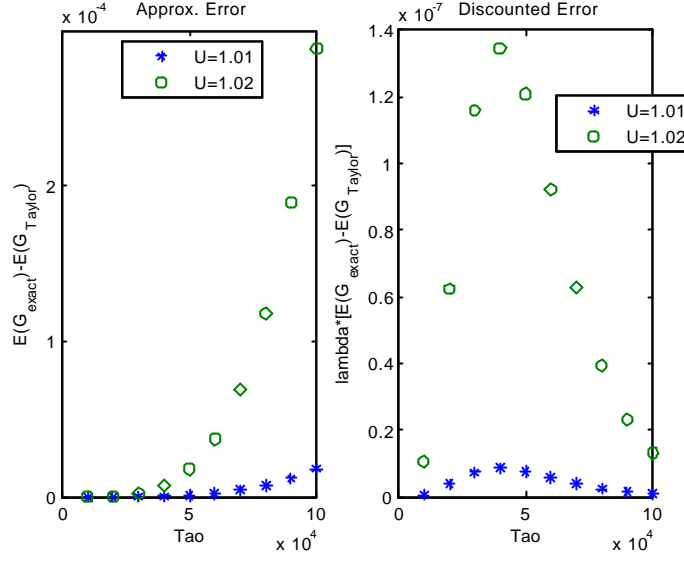
Figure 2.3.A: *Difference between exact and approximated value of E(G) (left). Discounted difference (right)*

When $\alpha \neq 1$ the stock price with learning will differ from its approximation. The error will depend on the values of the parameters α, λ, U, c, n and N as well as the number of terms I used to achieve convergence of the stock price. For a given value of I the exact stock price at date t is

$$S_t^{BL} = d_t \sum_{\tau=1}^{I} \lambda^\tau E\left[G_\tau \mid n, \ N\right]$$

An analytical expression of the difference between $E\left[G_\tau\right]$ and its approximation $E_t\left[\tilde{G}_\tau^*\right]$ is difficult to derive. Instead we compute the difference for values of τ ranging between 10^4 and 10^5 and set the parameters to

$$\alpha = 0.999 \quad \lambda = 0.9999 \quad c = 1 \quad n = N - n = 40$$

and let U assume the values 1.01 and 1.02. In Figure 2.3.A (left) the difference $E\left[G_t\right] - E_t\left[\tilde{G}_\tau^*(k)\right]$ is shown for different values of τ. With $d_t = 1$ the resulting difference in stock price is represented by the areas defined by the discounted differences shown in Figure 2.3.A (right).

It should be noted that setting $c = 0$ in the expression for the stock price with learning in 2.3.B, that is removing the possibility of learning effects, reproduces the expression for the stock price without learning in

2.3.A. Setting $c = 0$ in the approximation of the stock price with learning however does not result in the stock price without learning.

Generalizing Polya Urn Models

In Section 2.1 the concept of learning was illustrated by introducing Polya's urn model and in the subsequent derivation of the asset prices with Bayesian learning the parameters of the Polya distribution reemerged. In the expression for the stock price with Bayesian learning at date t the parameters of the Polya distribution, all integers, have the following interpretation:

n : number of upsteps prior to date t

N : total number of steps taken prior to date t

c : strength of learning

In this setting the strength of learning is affected both by N and c; an increase in c strengthens the learning whereas the opposite holds for an increase in N. Clearly, the precision of the estimated probability of an upstep is also effected by the size of N, since a large value of N implies extensive prior knowledge.

It would be preferable to have only one parameter steering the strength of learning, since this would facilitate estimation and benefit analytical tractability. We now proceed to show that this is in fact possible by introducing the parameter $\delta = c/N$. Hence we normalize the model by setting the total number of steps taken prior to date t, N, equal to unity.

Recall from Section 2.1 that the probability of getting k upsteps in τ steps is

$$
\begin{aligned}
P_{k,\tau} = \ & \binom{\tau}{k} \ \frac{n}{N} \ \frac{n+c}{N+c} \ \cdots \ \frac{n+(k-1)\,c}{N+(k-1)\,c} \\
& \times \ \frac{N-n}{N+kc} \ \frac{N-n+c}{N+(k+1)\,c} \ \cdots \ \frac{N-n+(\tau-k-1)\,c}{N+(\tau-1)\,c}
\end{aligned}
$$

Dividing all terms by N and setting $\delta = c/N$ and $\pi = n/N$ yields

$$P_{k,\tau} = \binom{\tau}{k} \frac{\pi}{1} \frac{\pi + \delta}{1 + \delta} \cdots \frac{\pi + (k-1)\delta}{1 + (k-1)\delta}$$
$$\times \frac{(1-\pi)}{1 + k\delta} \frac{(1-\pi) + \delta}{1 + (k+1)\delta} \cdots \frac{(1-\pi) + (\tau - k - 1)\delta}{1 + (\tau - 1)\delta}$$

It was shown in Section 2.1 that the urn model could be represented with the beta function if the probability of an upstep π was beta distributed with parameters $\alpha = n/c$ and $\beta = (N-n)/c$. For the normalized model we have

$$\alpha = \frac{n}{c} = \frac{\pi}{\delta} \quad \text{and} \quad \beta = \frac{N-n}{c} = \frac{1-\pi}{\delta}$$

Thus, if $\pi \in \text{Beta}\left(\frac{\pi}{\delta}, \frac{1-\pi}{\delta}\right)$ then the probability of getting k upsteps in τ steps is

$$P_{k,\tau} = \binom{\tau}{k} \frac{B\left(\frac{\pi}{\delta} + k, \frac{1-\pi}{\delta} + \tau - k\right)}{B\left(\frac{\pi}{\delta}, \frac{1-\pi}{\delta}\right)}$$

Whereas in the original urn model c was the number of balls to add if a white ball was drawn, we can interpret δ, being a rational number, to be the amount of sand to add to the normalized urn of size 1. Obviously π is the initial probability of getting a white ball/upstep.

Thus, with $\delta = c/N$ and $\pi = n/N$ the normalized approximation of $\psi_t(n, N) = \Sigma_{\tau=1}^{\infty} \lambda^\tau E_t[G_\tau | n, N]$ is

$$(\pi, \delta) = \left\{ \frac{\lambda \hat{U}}{1 - \lambda \hat{U}} + \frac{2[(1-\alpha)\log U]^2 \pi (1-\pi)}{1 + \delta} \right.$$
$$\left. \left[\frac{\lambda \hat{U}}{\left(1 - \lambda \hat{U}\right)^2} + \delta \frac{\left(1 + \lambda \hat{U}\right) \lambda \hat{U}}{\left(1 - \lambda \hat{U}\right)^3} \right] \right\}$$

$$\textbf{2.3.C}$$

where now $\hat{U} = U^{(1-\alpha)(2\pi - 1)}$. Hence the approximated stock price with learning is

$$\tilde{S}_t^{BL} = d_t \ (\pi, \delta) \qquad\qquad \textbf{2.3.D}$$

Similarly, the approximated date t price of a call with T periods to expiration and strike price K is

$$
\tilde{C}_t^{BL} = \lambda^T \sum_{k=0}^{T} \binom{T}{k} \left[U^{2k-T} \right]^{-\alpha} \frac{B\left(\dfrac{\pi}{\delta} + k, \dfrac{1-\pi}{\delta} + T - k \right)}{B\left(\dfrac{\pi}{\delta}, \dfrac{1-\pi}{\delta} \right)}
$$

$$
\times \left(U^{2k-T} \tilde{S}_t^{BL} \frac{\left(\dfrac{\pi + k\delta}{1 + T\delta}, \dfrac{\delta}{1 + T\delta} \right)}{(\pi, \delta)} - K \right)^+
$$

$$
\textbf{2.3.E}
$$

Rationalization of the BS Model

In this section we investigate the relationship between the Lucas model without learning (Full Information, FI) and the Black and Scholes model. We conclude with conditions on the parameters, which if met result in identical option price predictions of the two models. This can be seen as a rationalization of the Black and Scholes model, which takes the price of the underlying asset as given and uses arbitrage arguments to price the option. Alternatively, the Black and Scholes model can be viewed as a reduced form of the FI model.

In the Lucas set-up, the stock pays a dividend. Thus, as a first step in relating the FI model to the BS model, in order to facilitate the comparison we give an account of the relationship between stocks with and without dividends.

A BS model with dividends is defined by

$\quad S \quad = \quad$ initial stock price

$\quad U \quad = \quad$ up factor

$\quad D \quad = \quad$ down factor

$\quad R \quad = \quad 1 + r =$ return on bond

$\quad d \quad = \quad$ initial dividend

Thus, the owner of a BS stock paying dividends d receives at date 1

$$
(S + d)\, U
$$

if an upstep occurs and

$$(S + d) D$$

if a downstep occurs.

It is possible to make a BS stock without dividends equivalent to a BS stock with dividends, in the sense that both stocks give the same return, by redefining the factors determining the return for the BS stock without dividends. To this end we define the upfactor for a BS stock without dividends as

$$\tilde{U} = \left(\frac{S + d}{S} \right) U$$

Similarly, the down factor becomes

$$\tilde{D} = \left(\frac{S + d}{S} \right) D$$

With the bond giving date 1 return R, it is easy to show that the state prices for both cases are identical. Hence, the state price in case of an upstep is

$$\beta_h = \frac{R - \left(\dfrac{S + d}{S} \right) D}{\left(\dfrac{S + d}{S} \right) (U - D) R} = \frac{R - \tilde{D}}{\left(\tilde{U} - \tilde{D} \right) R}$$

In case of a down step the state price is

$$\beta_l = \frac{\left(\dfrac{S + d}{S} \right) U - R}{\left(\dfrac{S + d}{S} \right) (U - D) R} = \frac{\tilde{U} - R}{\left(\tilde{U} - \tilde{D} \right) R}$$

Since the up and down factors are assumed constant over time the one period case is easily generalized to the T period case. A stock without dividends and upfactor $\tilde{U}$ and downfactor $\tilde{D}$ has price after k upsteps in T periods equal to

$$S \left[\frac{S + d}{S} U \right]^k \left[\frac{S + d}{S} D \right]^{T-k} = \left[\frac{S + d}{S} \right]^T S U^k D^{T-k}$$

This relation has the following interpretation. Buying 1 unit of a stock without dividends giving returns defined by $\tilde{U}$ and $\tilde{D}$ yields the same

payoff at a (k, T)-node as the strategy of buying $[(S + d) / S]^T$ amount of a stock paying dividends and giving one period returns U and D.

This reasoning carries over to the valuation of call options. Thus a call option on 1 unit of a stock without dividends and expiration date T and strikeprice K has date T value

$$\max\left(S\left[\frac{S + d}{S} U\right]^k \left[\frac{S + d}{S} D\right]^{T-k} - K, 0\right)$$

at a (k, T)-nod. On the other hand, consider a call option on $[(S + d) / S]^T$ units of a stock with dividends with expiration date T which requires the buyer to pay K at date T. It has the date T value

$$\max\left(\left[\frac{S + d}{S}\right]^T SU^k D^{T-k} - K, 0\right)$$

at a (k, T)-nod, which is the same as the value of a call option on a stock paying no dividends. Hence the two options have the same initial value. The FI stock price at a (k, T)-node is

$$SU^k D^{T-k}$$

and the corresponding state price is

$$\left(\lambda \pi U^{-\alpha}\right)^k \left(\lambda (1 - \pi) D^{-\alpha}\right)^{\tau - k} = \beta_h^k \beta_l^{\tau - k}$$

Letting $A = \pi U^{1-\alpha} + (1 - \pi) D^{1-\alpha}$ we know from Proposition 2.3.A that the initial dividend is

$$d = S\frac{1 - \lambda A}{\lambda A}$$

By Proposition 2.3.A the FI bond price is

$$\lambda \pi U^{-\alpha} + \lambda (1 - \pi) D^{-\alpha} = \frac{1}{R}$$

Setting $D = 1/U$ we thus obtain a non-linear equation system

$$\begin{cases} d &= S\dfrac{1 - \lambda A}{\lambda A} \\ \lambda \pi U^{-\alpha} + \lambda (1 - \pi) U^{\alpha} &= \dfrac{1}{R} \end{cases}$$

Given values of U, d and α solving this equation system thus provides values of λ and π which make the Full Information model and the Black and Scholes model equivalent in the following sense: the price of a call option on one unit of a BS stock with upfactor $(S+d)U/S$ is exactly that which an FI agent with subjective discount factor λ would be willing to pay for a call option on $[(S+d)/S]^T$ units of a stock paying initial dividend d and having upfactor U. This holds regardless of the agents probability beliefs and risk aversion, provided that these are consistent with convergence of the FI stock price.

2.4 Bayesian Learning vs. Full Information - an Example

The purpose of this section is to illustrate how the model with learning compares to the model without learning. Thus, with current stock price set to 30, we use the model with learning to generate the stock prices at date $t+5$. We then get the date $t+5$ value of 6 call options differing only in their exercise prices. The state prices from the model with learning is then used to calculate the date t price of the options. These prices are then compared to the prices predicted by the FI model.

Generating Stock Price from the BL model

We use the following parameters

$$U = 1.2, \qquad D = 0.9, \quad \alpha = 0.9 \ , \quad \lambda = 0.9615$$

By setting the parameters as above, convergence of the sum giving the stock price under Bayesian learning is guaranteed. With a prior given by $n = 6$ and $N = 10$, making the initial probability of an upstep equal to 0.6, the spot prices of the stock at date $t + 5$ were calculated by approximating the infinite sum from Proposition 2.3.B with the first 100 terms. This gives six distinct prices shared among the 32 terminal nodes representing the different states of the world. Table 2.4.A shows these stock prices, along with the numbers that were used to calculate the state

prices according to relation (2.2.C). It also reports the values of the ψ function for the different configurations of the urn at expiration date.

Table 2.4.A: *Data generated at date $t+5$ by the model with learning*

# of nodes	Σ	S_{t+5}^{BL}	P_σ	d_σ	ψ_σ	β	$\beta*$# of nodes
1	Σ_{t+5}^5	82.8205	0.1259	2.4883	32.6986	0.0455	0.0455
5	Σ_{t+5}^4	58.8249	0.0503	1.8662	30.9665	0.0236	0.1180
10	Σ_{t+5}^3	41.8465	0.0280	1.3997	29.3716	0.0170	0.1700
10	Σ_5^2	29.8138	0.0210	1.0498	27.9013	0.0165	0.1650
5	Σ_{t+5}^1	21.2726	0.0210	0.7873	26.5440	0.0214	0.1070
1	Σ_{t+5}^0	15.2003	0.0280	0.5905	25.2893	0.0369	0.0369

In Table 2.4.B the payoffs at exercise day to six call options expiring at date $t + 5$, with strike prices K ranging from 5 to 80 are shown. The date t prices, calculated using the state prices from Table 2.4.A, are shown in the last row.

Table 2.4.B: *Payoffs to call options for the BL model at expiration date. The strike prices range from 5 to 80. The last row gives the date t learning prices.*

Σ	C_{80}^{BL}	C_{65}^{BL}	C_{50}^{BL}	C_{35}^{BL}	C_{20}^{BL}	C_5^{BL}
Σ_{t+5}^5	2.8205	17.8205	32.8205	47.8205	62.8205	77.8205
Σ_{t+5}^4	0	0	8.8249	23.8249	38.8249	53.8249
Σ_{t+5}^3	0	0	0	6.8465	21.8465	36.8465
Σ_{t+5}^2	0	0	0	0	9.8138	24.8138
Σ_{t+5}^1	0	0	0	0	1.2726	16.2726
Σ_{t+5}^0	0	0	0	0	0	10.2003
Price date t	0.1284	0.8115	2.5358	6.1515	12.9076	22.3636

Generating Stock Price from the FI model

When computing the stock and call prices for the FI model we maintain the values of U, D, λ and α above. Table 2.4.C reports the stock prices

at exercise date as well as the probabilities and state prices for the final
nodes. It is seen that the FI stock prices at expiration date varies less
than the corresponding BL stock prices. Further, note that the proba-
bilities of extreme stock prices are lower in the FI case, a feature which
carries over to the state prices.

Table 2.4.C: *Data generated at date t+5 by the FI model*

# of nodes	Σ	S_σ^{FI}	P_σ	d_σ	β	$\beta*$# of nodes
1	Σ_{t+5}^5	74.6496	0.0778	2.4883	0.0281	0.0281
5	Σ_{t+5}^4	55.9872	0.0518	1.8662	0.0243	0.1215
10	Σ_{t+5}^3	41.9904	0.0346	1.3997	0.0210	0.0210
10	Σ_5^2	31.4928	0.0230	1.0498	0.0181	0.1810
5	Σ_{t+5}^1	23.6196	0.0154	0.7873	0.0157	0.0785
1	Σ_{t+5}^0	17.7147	0.0102	0.5905	0.0135	0.0135

Table 2.4.D shows the payoffs to the call options in the FI model, with
current call prices in the last row. Compared to the BL case, payoffs for
states with many upsteps are lower, as are all current call prices. Note
that the lower variance of the FI stock prices results in zero payoffs in all
states for the option with $K = 80$.

Table 2.4.D: *Payoffs to call options for the FI model at expira-
tion date. The strike prices range from 5 to 80.
The last row gives the date t FI prices*

Σ	C_{80}^{FI}	C_{65}^{FI}	C_{50}^{FI}	C_{35}^{FI}	C_{20}^{FI}	C_5^{FI}
Σ_{t+5}^5	0	9.6496	24.6496	39.6496	54.6496	69.6496
Σ_{t+5}^4	0	0	5.9872	20.9872	35.9872	50.9872
Σ_{t+5}^3	0	0	0	6.9904	21.9904	36.9904
Σ_{t+5}^2	0	0	0	0	11.4928	26.4928
Σ_{t+5}^1	0	0	0	0	3.6196	18.6196
Σ_{t+5}^0	0	0	0	0	0	12.7147
Price date t	0	0.2715	0.8389	1.7720	3.1381	4.9178

Risk neutral probabilities

We conclude the example by comparing the risk neutral probabilities of the terminal nodes of the BL model and the FI model.

The risk neutral probability (Cf Leroy and Werner [2001]) of a state $\sigma \in \Sigma_{t+\tau}^k$ is

$$\pi_\sigma = \frac{\beta_\sigma}{\rho_\sigma}$$

where ρ_σ is the discount factor given by the product of the bond prices at the preceding nodes on the path corresponding to the state $\sigma \in \Sigma_{t+\tau}^k$. The discount factor is the same for those final nodes which have the same predecessor. As an example, consider a two-period tree. Then the discount factor used for the node reached by two upsteps and the node reached by one upstep followed by a downstep is

$$\rho_\sigma = B_0 B_u$$

where B_0 is the bond price at the original node and B_u is the bond price at the node reached after one upstep.
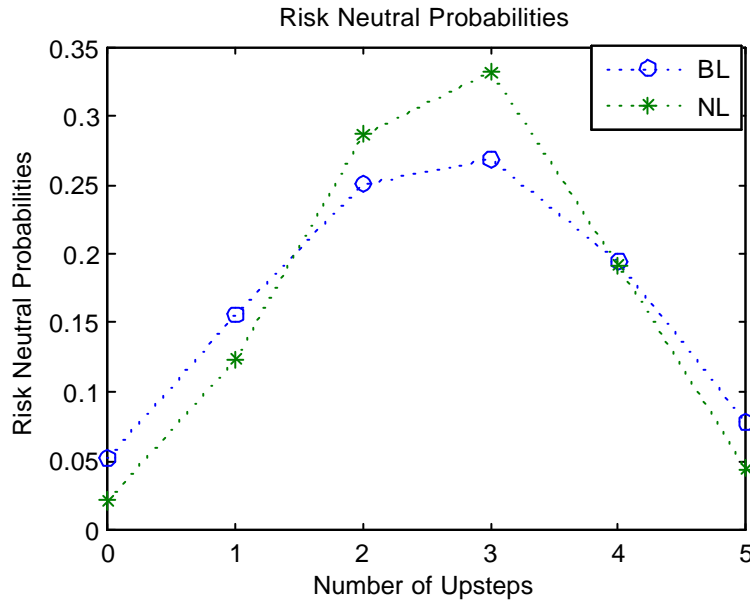
Figure 2.4.A: *Risk neutral probabilities for the BL and FI models for the 5 period example*

Computing the risk neutral probabilities is more complicated for the BL model than for the FI model, since for the BL model the present value of a crown received at a future node will depend not only on the number of high and low growth periods preceding that node, but the sequence in which they occur. Hence, with 32 final nodes, computing the risk neutral probabilities in the BL case requires the computation of 16 unique discount factors. Since the bond price in the FI model is constant, only one discount factor needs to be computed.

Figure 2.4.A shows the risk neutral distributions for the BL and FI models. The risk neutral probabilities for terminal nodes $\sigma \in \Sigma_{t+5}^k$ are summed. It is seen that the fatter tails of the (subjective) probability distribution for the learning case carries over to the risk neutral distribution.

Summary

This chapter introduced a special case of the representative agent model due to Lucas [1978], where dividends evolve on a binomial lattice and preferences are represented by power utility. Furthermore, having introduced the Polya distribution we could distinguish between the probability beliefs of the FI agent and the BL agent by letting the learning parameter c equal zero, in which case the binomial distribution is obtained, and belonging to $\mathcal{Z}_+$, respectively. Thus, having specified the dividend process, the utility function and the respective probability distributions we could give explicit expressions for the state prices for FI and BL cases, respectively. Then, with the state prices at hand it was straightforward to determine stock, bond and call prices with and without learning.

With learning, the implied probability distribution displays fatter tails as compared to the FI (binomial) case. This feature carries over to the state prices, implying that the BL agent values payoffs in extreme events higher than does the FI agent.

Introducing learning has several implications for the asset prices. The bond price is no longer constant but will depend on the sequence of occurrences of high and low growth. This in turn will cause the interest

rate process under learning to be path dependent. The higher volatility of the BL stock price obviously affects the valuation of call prices under learning, as does the valuation of the payoffs at exercise date with the BL call prices.

This chapter also investigated the relationship between the FI model and the BS model, describing the conditions under which the two models give identical call option price predictions. Thus, the FI model can be viewed as an economic rationalization of the BS model. Alternatively, the BS model can be viewed as a reduced form of the FI model.

In order to make the BL model more tractable, a closed form approximation of the BL stock price was derived, utilizing the properties of the Polya distribution. This approximation was further simplified by normalizing the urn so that the strength of learning was expressed through the single parameter δ.

A drawback of the learning model presented here is that in order to avoid learning to vanish as new balls are added, the initial urn needs to be repeatedly replaced, causing jumps in the strength of learning. Although sharp revisions in probability beliefs could conceivably be motivated by dramatic events, further research on how to keep the strength of learning from decreasing in a mechanical fashion is commendable.

Chapter 3

Application of BL to Option Pricing

Introduction

In Chapter 1 the shortcomings of the Black and Scholes (BS) option pricing model was explored and the resulting volatility smiles were investigated for data on calls and puts on the Swedish OMX index. A competing option pricing model utilizing the Lucas asset pricing model and Bayesian Learning was presented in Chapter 2.

Chapter 2 also provided an account of the relationship between the Black and Scholes model and the model combining the Lucas set-up and no learning, the Full Information model (FI). It was shown that by choosing the parameters appropriately, it was possible to obtain equality between the FI and BS call prices.

Ever since Black and Scholes presented their option pricing formula in 1973, a host of alternative option pricing models has been suggested. Bates [2003] gives an overview of the results

In this chapter, we turn to the empirical performance of the learning model and compare it with various versions of the BS/FI model. Thus we apply the various call pricing models to data on call options on the Swedish OMX index covering the period $1993 - 2000$ described in Chapter 1. Here we restrict our study to prices recorded on Wednesdays, yielding a sample of 4132 observed call prices distributed over 261 days.

The number of observed call prices each day range between 5 and 40 and the median number of observations is 15. The evaluation of the models is done by computing the in-sample pricing errors for each date t in the sample. In fitting the models to the data, two routes are taken. The first approach utilizes all observations on each date t to estimate the models. The second approach entails dividing the data into strata according to moneyness and time to expiration, and fitting a model for each strata. The stratas are defined as follows. Moneyness is divided into three subsets defined by the intervals $M_{OTM} = [0.8, 0.98]$, $M_{ATM} = (0.98, 1.02]$ $M_{ITM} = (1.02, 1.2]$. Similarly, options are classified according to time to expiration by the intervals $T_S = [8, 32]$, $T_M = [33, 55]$, $T_L = [56, 90]$.

Table 3.0.A shows the distribution of call options in the sample over these subsets.

Table 3.0.A: *Distribution of calls across moneyness and time to expiration*

	M_{OTM}	M_{ATM}	M_{ITM}	Total
T_S	1240	633	501	2374
T_M	720	347	187	1254
T_L	286	162	56	504
Total	2246	1142	744	4132

Due to the discrete time set-up, we are concerned with the (daily) upfactor U rather than the (annual) volatility σ, although at times it will be convenient to make the transformation $\sigma = U \ln \sqrt{252}$.

The estimated parameters resulting from the in-sample estimation is then used to assess the performance of the models out-of-sample. Section 3.1 describes the measures of fit applied and how the estimation is performed. Section 3.2 reports the in-sample results while Section 3.3 reports the out-of-sample results.

3.1   Estimation

The empirical performance of an option pricing model is assessed by choosing the parameters such that the deviation from observed prices,

C^{obs}, is minimized. Thus, for each date t in the sample, with n_t observations in the subset of interest, we compute the root mean squared valuation error

$$RMSVE_t = \left[\frac{1}{n_t}\sum_{i=1}^{n_t}\left(C_i^{obs}\left(v_t\right) - C_i\left(v_t, l\right)\right)^2\right]^{1/2}$$

where the current stock price S, exercise price K, interest rate r and time to expiration T at date t are collected in the vector $v_t = [S, K, r, T]$. $C_i\left(v_t, l\right)$ is the price prediction of the model we wish to evaluate having parameters l. The minimization of $RMSVE_t$ is performed with respect to the parameters in the vector l, although for the Full Information and the Bayesian Learning models some parameters are fixed.

We also compute the relative root mean squared valuation error, defined as

$$RRMSVE_t = \frac{RMSVE}{\left[\frac{1}{n_t}\sum_{i=1}^{n_t}\left(C_i^{obs}\left(v_t\right)\right)^2\right]^{1/2}} = \left[\frac{\sum_{i=1}^{n_t}\left(C_i^{obs}\left(v_t\right) - C_i\left(v_t, l\right)\right)^2}{\sum_{i=1}^{n_t}\left(C_i^{obs}\left(v_t\right)\right)^2}\right]^{1/2}$$

Summing over all dates t yields the measures

$$RMSVE = \sum_t RMSVE_t \quad \text{and}$$
$$RRMSVE = \sum_t RRMSVE_t$$

We now proceed to present the various call pricing models and give an account of the estimation procedures. Estimates of the relevant parameters were obtained by numerical minimization of the in-sample pricing errors for each date t in the sample. The presentation is focused on the unstratified models where n_t denotes all observations at date t used in estimating the parameters of the model. The estimation procedures are completely analogous for the stratified models where the BS, FI and BL models are estimated on subsets of the data described above. These models are referred to by adding arguments M and T, so that BS(M) refers to the the BS model where the estimation produces a set of estimated parameters for each interval of moneyness and BS(T) refers to the the BS model where the estimation produces a set of estimated parameters for each interval of time to expiration.

Estimation of the Black and Scholes model

We estimate the modified BS model presented in Chapter 2. It was shown that the call price prediction of this model equaled the call price a FI agent would be willing to pay for a call giving the right to $((S + d)/S)^T$ amount of stock in T days, given certain conditions on the parameters. One condition stipulated that with upfactor U in the FI model, the corresponding upfactor $\tilde{U}$ in the BS model should be set to

$$\tilde{U} = \frac{S + d}{S} U$$

where d is the daily dividend. Thus, we obtain estimates of U and d and the corresponding measure of fit of the BS model by numerically solving, for each date t in the sample,

$$\min_{U,d} \quad \left[\frac{1}{n_t}\sum_{i=1}^{n_t}\left(C_i^{obs}\left(v_t\right) - C_i^{BS}\left(v_t, \tilde{U}\right)\right)^2\right]^{1/2} \qquad \textbf{3.1.A}$$

where n_t is the number of observations at date t.

Being that the rationale for estimating this modified version of the BS model is to achieve equivalence with the equilibrium model without learning, the estimation procedure is performed with a view to the convergence condition of the FI model. Thus, while the dividend clearly must be nonnegative, we increase the lower bound and restrain the search to the interval $[0.05, 1.5]$.

Estimation of the Full Information model

In estimating the FI model for date t we fix the risk aversion, α, at 0.999 and retain the estimates of U and d received from fitting the BS model. Thus, equipped with these values of α, U and d and given the interest rate r we proceed to solve for each observation i at date t the equation system

$$\begin{cases} d &=& S\dfrac{1 - \lambda A}{\lambda A} \\[2ex] \lambda\pi U^{-\alpha} + \lambda\left(1 - \pi\right)U^{\alpha} &=& \dfrac{1}{R} \end{cases} \qquad \textbf{3.1.B}$$

where $R = 1 + r$, $A = \pi U^{1-\alpha} + (1 - \pi) U^{\alpha-1}$ and S is the current stock price. Recalling the results in Chapter 2, we know that the solution provided by probability π and subjective discount factor λ ensures the equality of the BS and the FI call prices in the sense that

$$C_i^{BS}(v_t, U, d) = C_i^{FI}(v_t, U, d, \alpha, \pi, \lambda)$$

where the FI call price $C_i^{FI}(v_t, U, d, \alpha, \pi, \lambda)$ is obtained by discounting with the FI state prices the payoff

$$\max\left([(S + d)/S]^T U^{2k-T} S - K, 0\right)$$

where $[(S + d)/S]^T$ is the amount of stock the FI agent has the right to purchase at price K, as owner of a call maturing in T days.

The resulting measure of fit is then

$$RMSVE_t = \left[\frac{1}{n_t}\sum_{i=1}^{n_t}\left(C_i^{obs}(v_t) - C_i^{FI}(v_t, U, d, \alpha, \pi, \lambda)\right)^2\right]^{1/2} \qquad \textbf{3.1.C}$$

and it should be clear from Chapter 2 that this yields the same result as 3.1.A.

Probability as function of dividend   Discount factor as function of dividend
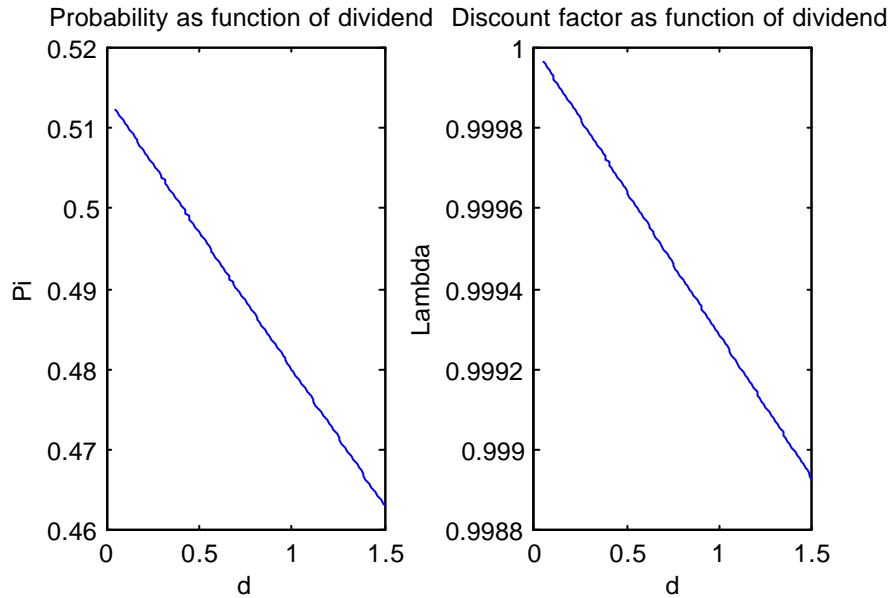
Figure 3.1.A: *Effect of initial dividend on the probability and and the subjective discount factor. Data are from October 11 1995 with S=1396, r=0.089 and U=1.01 implied by the observation*

It should be noted that for a given date t there will in general among the n_t observations be calls differing in remaining days to expiration T. It follows that the interest r used in 3.1.B will also differ for these calls, and hence the resulting values of π and λ. As a result we study the daily averages $\bar{\pi}$ and $\bar{\lambda}$.

In solving 3.1.B for fixed values of α, U, S and r, the resulting values of π and λ are decreasing in the dividend d. Figure 3.1.A shows the solutions π and λ to 3.1.B for a given day in the sample when U, S and r are given by the data and we let d vary from 0.05 to 1.5. This illustrates the need to restrict the values of d; high values of d could potentially result in unrealistically low probabilities, while to low values of d could cause the convergence criterion to be violated.

Estimation of the Bayesian Learning model

In estimating the learning model we employ the approximation presented in Chapter 2, i.e. Equation 2.3.E. In addition to facilitate the computations, this approach has the added advantage of supplying a single measure of the strength of learning, conveyed by the estimate of the parameter δ. Fixing the risk aversion α at 0.999 we thus face the task of finding values of π, U, d, λ and strength of learning δ so as to minimize $RMSVE_t$.

One possible approach would be to maintain the FI estimates of π, U, d, and λ and minimize $RMSVE_t$ with respect to $\delta \geq 0$. However, it turns out that proceeding in this fashion leaves little room for learning as the estimates of δ for most dates t then become zero, rendering the improvement over the BS/FI model insignificant. Instead, the approach followed here is to adopt the FI estimate of λ, while we set $d = 0.05$ to ensure convergence and minimize $RMSVE_t$ with respect to π, U and δ:

$$\min_{\pi, U, \delta} RMSVE_t = \left[\frac{1}{n_t} \sum_{i=1}^{n_t} \left(C_i^{obs}(v_t) - C_i^{BL}(v_t, U, d, \alpha, \pi, \lambda, \delta) \right)^2 \right]^{1/2}$$

3.1.D

In order to make the results comparable with the FI model, we postulate that the owner of a BL call maturing in T days has the right to purchase $[(S + d)/S]^T$ amount of stock at exercise date at price K. It

should be noted that this approach to estimating the parameters is inconsistent in that it does not take into account the restraints on how the updating occurs over time imposed by the Bayesian setting.

Estimation of Black and Scholes Spline model

Next we estimate the BS spline model. Here, for each date t we let the volatility input σ^S to the BS model depend on the exercise price K and time to maturity T in the following way:

$$\sigma_i^S \quad = \quad \max\left\{0.01, \gamma_1 + \gamma_2 K_i + \gamma_3 K_i^2 + (\gamma_4 + \gamma_5 K_i^2)T_i\right\}$$

We get estimates of the parameters $\gamma_t = (\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5)$ by solving

$$\min_{\gamma}\sum_{i=1}^{n_t}\left(\sigma_i^{imp}(v_t) - \sigma_i^S(K, T, \gamma)\right)^2$$

for each date t, where σ_i^{imp} is the volatility implied by call i at date t. Then, equipped with estimates of γ we can compute σ_i^S for each option and the corresponding upfactor $U_i = e^{\sigma_i^S/\sqrt{252}}$. The BS spline call prices $C_i^{BS}\left(v_t, U_i\left(\sigma_i^S\right)\right)$ result in the measure of fit

$$RMSVE_t = \left[\frac{1}{n_t}\sum_{i=1}^{n_t}\left(C_i^{obs}(v_t) - C_i^{BS}\left(v_t, U_i\left(\sigma_i^S\right)\right)\right)^2\right]^{1/2}$$

3.2 In-Sample Results

The preceding estimation procedures yield a set of estimated parameters and measures of fit for each of the models. In this section the results of the estimations are presented and the competing models are evaluated. We thus report for each model its ability to match the data as measured by $RMSVE_t$ and $RRMSVE_t$. While the fit of a model is of obvious importance, the estimates of the parameters are interesting in their own right. For each model we report statistics on its estimated parameters and how they compare between the models. The fact that the FI model can be viewed as an equilibrium representation of the BS model makes it interesting to compare the FI and BL estimates of the parameters. The one parameter only present in the BL model is δ, the strength of learning.

We first present the results for the unstratified models that were estimated using all data for each date and then proceed to the stratified models that were estimated when the data were divided in subsets according to moneyness and time to expiration.

Unstratified Models

Measure of fit

To compare the results across models, we start by examining the measures $RMSVE_t$ and $RRMSVE_t$ for the BS, FI, BL and BS spline models.

Figure 3.2.A shows how $RMSVE_t$ and $RRMSVE_t$ have evolved over time for the BS, BL and spline models. The pictures indicate that there is a clear tendency for the BL model to outperform the BS model and the spline model.
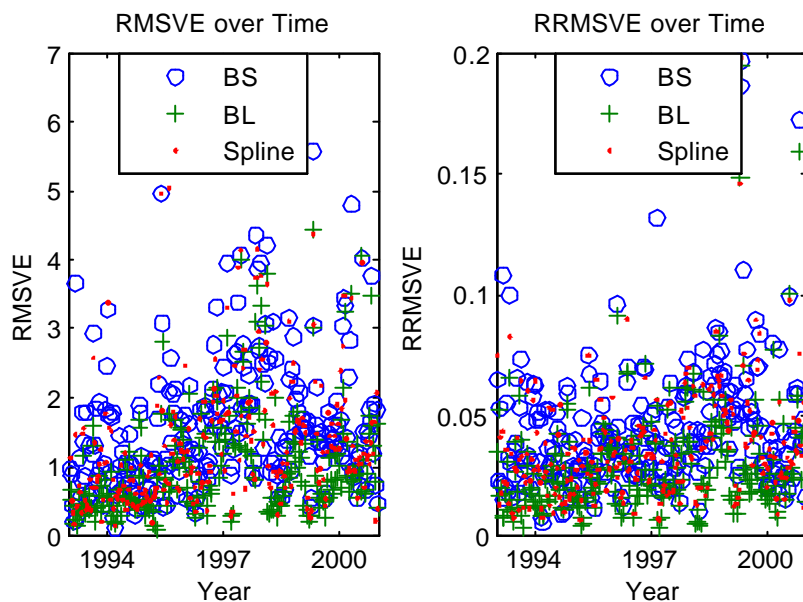
Figure 3.2.A: *RMSVE (left) and relative RMSVE (right) for the BS, BL and BS spline models over time*

Table 3.2.A presents a closer examination of the ability of the models to fit data. Along with statistics on $RMSVE_t$ and $RRMSVE_t$ it shows that the BL model gives the lowest overall error with a $RMSVE$ of 280.79 SEK, which is about 2/3 that of the $RMSVE$ of the BS model.

The second best model is the spline model, which has a $RMSVE$ of 344.99 SEK, which is about 4/5 that of the $RMSVE$ of the BS model.

Due to the set-up, in theory the BS model and the FI model should give identical measures of fit.

However, Table 3.2.A shows that there is indeed a slight discrepancy; we attribute this difference to the fact that for each date the BS call prices are computed using the same values of U and d, whereas the corresponding FI call prices are obtained by different values of π and λ, depending on the interest rate pertaining to the individual call. Also, it is possible that the fact that the parameter d was restricted to the interval $[0.05, 1.5]$ affects the results.

Table 3.2.A also presents the percentage of dates for which each model gives the best fit; it is seen that the BL model is able to outperform the other models about 67% of the dates in the sample. It should be pointed out that the fact that the BS/FI models are able to do better than the BL model in a few cases is in all likelihood a result of the route taken in estimating the BL model; had we retained the values of π, U and d from the FI model and minimized $RMSVE_t$ only with respect to $\delta \geq 0$, then the exact BL call price would equal the FI call price if $\delta = 0$.

Table 3.2.A: *Statistics on measure of fit for the BS, FI, Spline and BL models*

	BS	FI	Spline	BL
mean $RMSVE_t$	1.6049	1.6080	1.3414	1.0758
std $RMSVE_t$	1.2903	1.2900	0.9540	0.8666
min $RMSVE_t$	0.1100	0.1100	0.1531	0.0973
max $RMSVE_t$	15.81	15.811	6.7832	6.4831
mean $RRMSVE_t$	0.0437	0.04377	0.0366	0.0299
std $RRMSVE_t$	0.0254	0.0254	0.0218	0.0228
min $RRMSVE_t$	0.0055	0.0054	0.0066	0.0033
max $RRMSVE_t$	0.1967	0.1967	0.1977	0.1943
$RMSVE$	418.88	419.69	350.11	280.79
$RRMSVE$	11.40	11.42	9.5567	7.80
% best fit	3.83	3.45	24.52	66.67

In estimating the BS and FI models the parameter d was restricted to the interval $[0.05, 1.5]$. In order to evaluate the impact of d on the performance of the FI model, the estimation was also performed with d fixed at 0.05, resulting in a $RMSVE$ of 429.25. Since this is only moderately higher than the $RMSVE$ of 419.69 recorded when we allowed d to vary, it suggests that the treatment of d is of limited importance for the fit of the model.

Upfactor

While the daily upfactor in the FI model is U, the modified upfactor in the BS model was set to $\tilde{U} = U(S + d)/S$ in order to achieve equivalence between the two models. Obviously, $\tilde{U} > U$ and we now proceed to investigate how these upfactors compares with the upfactor in the BL model. Figure 3.2.B depicts the daily upfactors and the corresponding annual volatilities for each date in the sample. The figure shows that for most days in the sample the BL upfactor is smaller than the upfactors of the FI and BS models; in fact, for only about 4% of the dates does the BL upfactor exceed $\tilde{U}$ and for only about 13% of the dates does it exceed U.
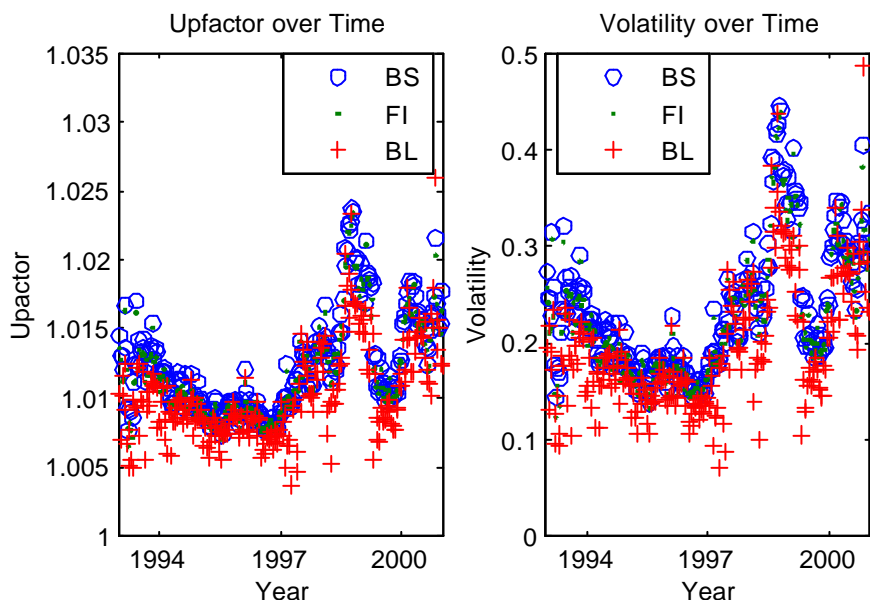
Figure 3.2.B: *Daily upfactor (left) and annual volatility (right) for the BS, FI and BL models*

Table 3.2.B presents statistics on the upfactor for all the models. While the average upfactor does not differ very much for the BS and FI models, corresponding to annual volatilities of 0.233 and 0.226 respectively, the average upfactor of the BL model is considerably lower, corresponding to an annual volatility of 0.194.

Table 3.2.B: *Statistics on the daily upfactor for the BS, FI and BL models*

	Upfactor		
	BS	FI	BL
min	1.0073	1.0065	1.0037
max	1.0237	1.0234	1.0259
mean	1.0124	1.012	1.0103
std	0.003612	0.003557	0.003489

Probability of an upstep

Estimation of the FI and the BL model yields competing estimates of the probability π. Although potentially interesting, a comparison is hampered by the sensitivity of FI estimates to the treatment of the parameter d. Our approach entailed allowing d to vary in the interval $[0.05, 1.5]$ in the FI model, while for the BL model we set $d = 0.05$. The resulting probabilities are shown in the left graph of Figure 3.2.C, while the right picture shows the probabilities obtained when d was set to 0.05 in the FI model as well.

Table 3.2.C: *Statistics on the FI and BL probabilities.*

	Probability		
	FI	BL, $d = 0.05$	FI, $d = 0.05$
min	0.4101	0.4833	0.5035
max	0.5151	0.5625	0.5163
mean	0.4921	0.5163	0.5083
std	0.0190	0.0083	0.0030

The figures show that the effect of allowing d to vary in the FI model

is to strongly increase the volatility of the probabilities, in that low prob-
abilities are then attainable. Conversely, when d is fixed, the FI proba-
bilities are much less volatile than the BL probabilities. In both cases,
the FI probabilities are in general lower than the probabilities generated
by the BL model. Table 3.2.C summarizes the findings.
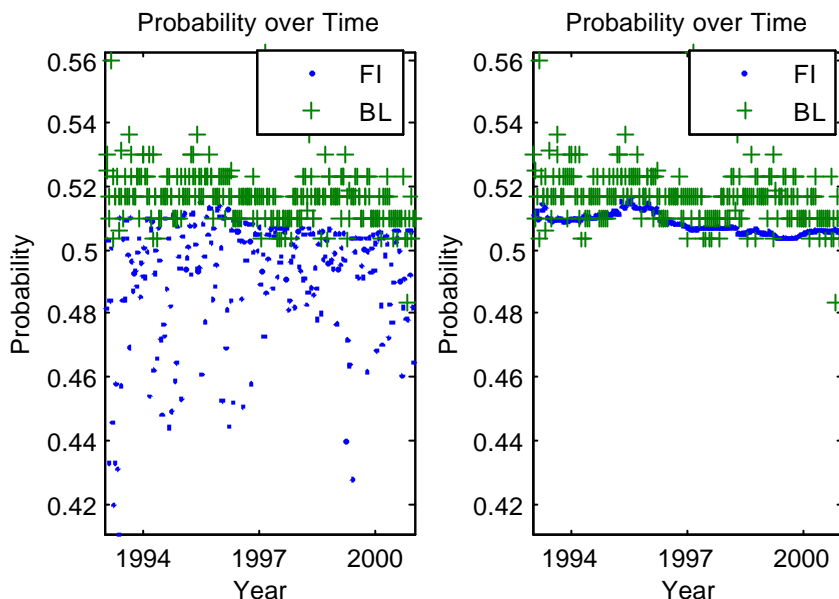
Figure 3.2.C: *Estimated probabilities from the FI model when
d varies and from the BL models when d is fixed
(left). Probabilities for the FI and BL models
when d is fixed for both models (right)*

Subjective discount factor

Fitting of the FI model yielded estimates of the subjective discount factor
λ, which were then used when we adapted the BL model to data. These
daily estimates were contained in the interval $[0.99789, 0.99998]$ with an
average λ of 0.9996. Figure 3.2.D shows that the variation of the daily
discount factor has remained relatively constant over time.

Daily dividend

In estimating the dividend d in the FI and BS models we restricted it to
the interval $[0.05, 1.5]$. As pointed out in the description of the estimation
procedure for the FI model, the lower bound is motivated by a concern

to ensure convergence of the FI model, since to low values of d will cause the discount factor λ to exceed 1. The lower bound was binding in 59 cases, while the upper bound was reached in 14 cases. The distribution of the dividend over time is shown in Figure 3.2.E. The average dividend was 0.478.
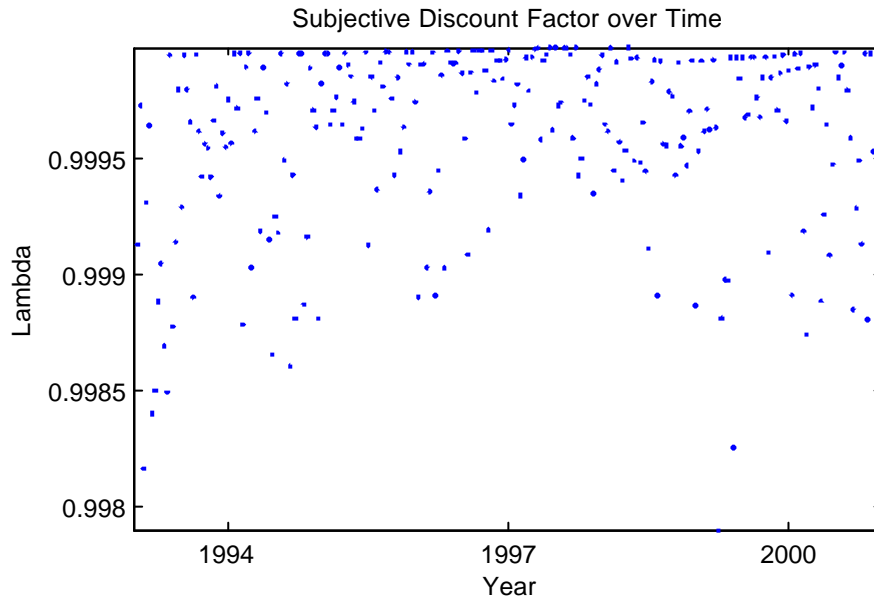
Figure 3.2.D: *The daily subjective discount factor over time*
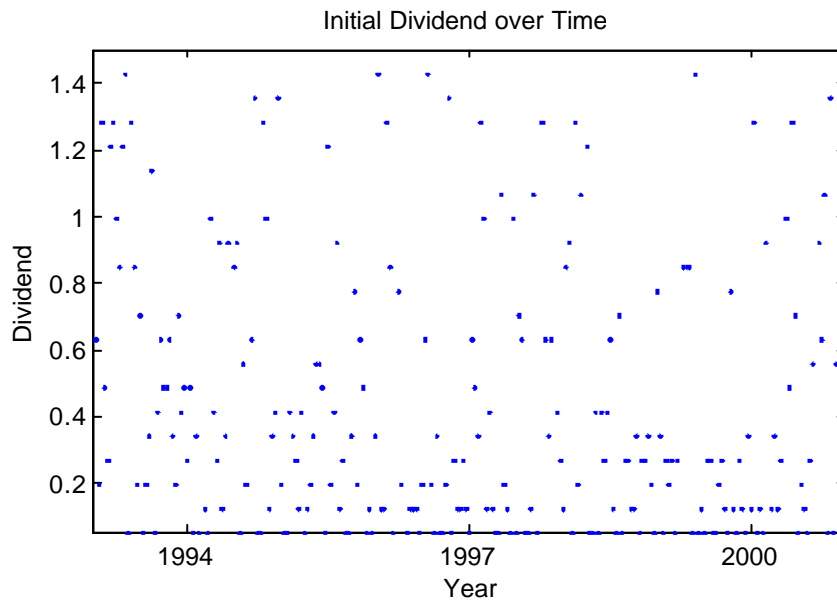
Figure 3.2.E:

Strength of learning

The strength of learning in the BL model is measured by the parameter δ. For 135 of the 261 days in the sample δ is zero. These dates, for which there were no learning effects, are evenly distributed over time. Inspection of Figure 3.2.F reveals that the variation of strength of learning is roughly constant over time. Furthermore, no unambiguous relationship between the strength of learning and the other parameters of the BL model, or even the fit of the model, could be established.
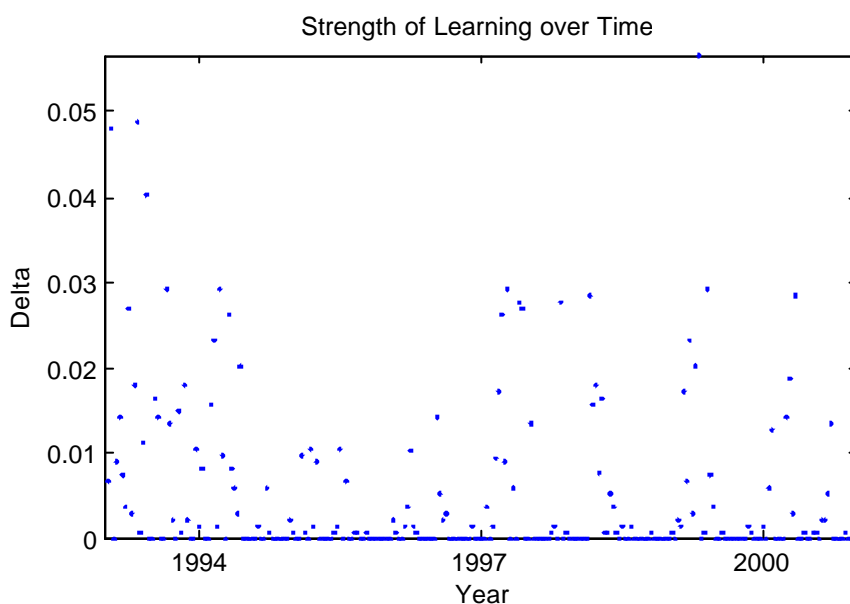
Figure 3.2.F: *Strength of learning over time*

The average δ was 0.0048 and the standard deviation was 0.0093 over the entire sample. Disregarding those dates for which $\delta = 0$ yields an average of almost 0.01 and a standard deviation of 0.011.

Parameters of the BS spline model

Fitting the BS spline model to data yields for each date a set of estimates of $\gamma = [\gamma_1 \ \gamma_2 \ \gamma_4 \ \gamma_5 \ \gamma_5]$ which were used in modelling the volatility as a function of moneyness and time to expiration. Figure 3.2.G shows the distribution over time of these estimates and Table 3.2.D reports their means and standard deviations.

Table 3.2.D: *Mean and standard deviation of parameters in the BS spline model*

	γ_1	γ_2	γ_3	γ_4	γ_5
mean	0.2343	$-8.888 * 10^{-5}$	$-1.504 * 10^{-8}$	0.0043	$-8.858 * 10^{-9}$
std	0.0729	0.0009	$8.8524 * 10^{-8}$	0.03175	$4.9703 * 10^{-8}$
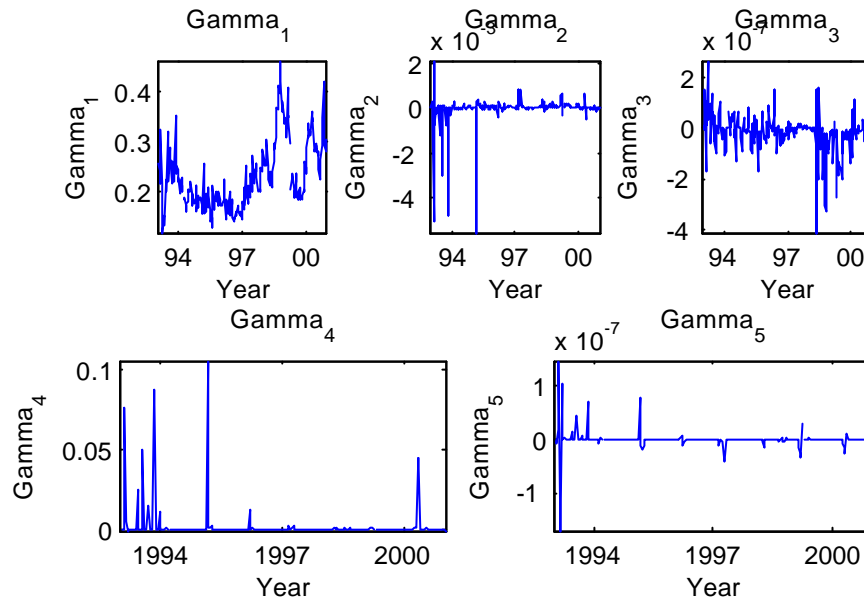
Figure 3.2.G: *The distribution of the parameters in the spline model over time (two outliers were removed)*

Stratified Models

We now proceed to present the results from fitting the BS, FI and BL models when the data was divided in subsets according to moneyness and time to expiration. This classification of models thus yields two set of models; we denote by $BS(M)$, $FI(M)$ and $BL(M)$ the models that were estimated when the data was divided in subsets of moneyness and by $BS(T)$, $FI(T)$ and $BL(T)$ the models that were estimated when the data was divided in subsets of time to expiration. In the following we will present both the aggregate results for each model and the results by intervals. This approach allows us to compare the estimated parameters with the results in Chapter 1, where the strength of smile was measured for intervals of moneyness and time to expiration. Of special interest

here is how the strength of learning δ varies with time to expiration and moneyness and how this variation compares with the strength of smile with respect to time to expiration and moneyness as studied in Chapter 1..

Stratification of the models implies that it is not always possible to obtain estimates of the parameters for each day and each strata, due to lack of data. Table 3.2.E shows for each strata the number of days it was possible to estimate the parameters of the stratified BS and BL model.

Table 3.2.E: *Number of days with sufficient number of observations, for each strata, for the BS and BL models, respectively.*

	S	M	L	OTM	ATM	ITM
BS	236	166	106	258	244	170
BL	234	162	85	250	204	118

Measure of fit

We start by comparing the models by intervals of moneyness and time to expiration. Since the results for the BS and FI models are practically identical, only the results for the FI model are displayed. Thus, comparing the average $RMSVE_t$ for the FI(T) and BL(T) models for each interval of time to expiration it is seen from Table 3.2.F that all models yield the best fit for options with long time to expiration. The improvement of the BL(T) model over the FI(T) model is roughly the same for all maturities; the mean $RMSVE_t$ of the BL(T) model is about half the mean $RMSVE_t$ of the FI(T) model. Both models perform the worst for options with medium time to expiration. However, the benefit of adding learning is the largest for these options with the average $RMSVE_t$ of the BL model is only 44.7% that of the FI model for options with medium time to expiration.

Turning to the relative measure of fit $RRMSVE_t$ the picture changes slightly. For both models RRMSVE$_t$ is decreasing in time to maturity. The benefit of allowing learning is still the greatest for options with medium time to expiration, with an average $RRMSVE_t$ only 43% that of

Table 3.2.F: *Statistics on RMSVE for the FI(T) and BL(T) models. The last column contains the ratio of the mean of RMSVE of BL(T) over the mean of RMSVE of FI(T)*

	$RMSVE_t$				
	FI(T)		BL(T)		
	mean	std	mean	std	Ratio
S	1.2032	0.9457	0.5463	0.5428	0.454 04
M	1.2450	0.8979	0.5570	0.7354	0.447 39
L	1.1288	0.9192	0.5371	0.7934	0.475 82

the FI model. However, both models now do worst for options with short time to expiration. The improvement of the BL model is still the least for options with long time to expiration, with an average $RRMSVE_t$ 52% that of the FI model. This result makes sense in the light of the findings in Chapter 1, where it was established that the strength of smile was decreasing in time to expiration.

Table 3.2.G: *Statistics on RRMSVE for the FI(T) and BL(T) models. The last column contains the ratio of the mean of RRMSVE of BL(T) over the mean of RRMSVE of FI(T)*

	$RRMSVE_t$				
	FI(T)		BL(T)		
	mean	std	mean	std	Ratio
S	0.0356	0.0223	0.0174	0.0150	0.488 76
M	0.0343	0.0249	0.0147	0.0182	0.428 57
L	0.0271	0.0195	0.0141	0.0207	0.520 30

The BL(M) yields a lower average $RMSVE_t$ for all intervals of moneyness than does the FI(M) model. While the performance of both models in terms of mean $RMSVE_t$ is decreasing in moneyness, the improvement of the BL(M) model over the FI(M) model is increasing in moneyness, with the greatest improvement for ITM options, which has an average $RMSVE_t$ only 65% that of the FI(M) model. For OTM options, adding learning provides the least improvement, with an aver-

age $RMSVE_t$ 77% that of the FI(M) model. Table 3.2.H presents the statistics on $RMSVE_t$ for the FI(M) and the BL(M) model.

Table 3.2.H: *Statistics on RMSVE for the FI(M) and BL(M) models. The last column contains the ratio of the mean of RMSVE of BL(M) over the mean of RMSVE of FI(M)*

	$RMSVE_t$				
	FI(M)		BL(M)		
	mean	std	mean	std	Ratio
OTM	0.7764	0.6525	0.6013	0.5785	0.774 47
ATM	0.9822	0.8999	0.7332	0.6665	0.746 49
ITM	1.1857	1.1709	0.7661	0.8098	0.646 12

The relative measure $RRMSVE_t$ reveals that the performance of both option pricing models is increasing in moneyness, yielding the lowest averages of $RRMSVE_t$ for ITM options. The improvement of the BL(M) model over the FI(M) model also increases with moneyness; for OTM options the average $RRMSVE_t$ of the BL(M) model is 73% that of the FI(M) model, while for ITM options the average $RRMSVE_t$ of the BL(M) model is only 64% that of the FI(M) model. The statistics on $RRMSVE_t$ are presented in Table 3.2.I.

The findings in Chapter 1 revealed that the average strength of smile was the highest for extreme values of moneyness, which would indicate that possibility of improvement would be the highest for these options. Disregarding the OTM options, the results here are compatible with that pattern, although differing stratification prevents a straightforward comparison.

Upfactor

The upfactor of the BS(T) model by design exceeds the upfactor of the FI(T) model. For the unstratified models we saw that the upfactors of the BS and FI models in general exceeded the upfactor of the BL model. This pattern carries over to the case when the estimation yields separate estimates of the upfactor for each interval of time to expiration. Table

Table 3.2.I: *Statistics on RRMSVE for the FI(M) and BL(M) models. The last column contains the ratio of the mean of RRMSVE of BL(M) over the mean of RRMSVE of FI(M)*

	$RRMSVE_t$				
	FI(M)		BL(M)		
	mean	std	mean	std	Ratio
OTM	0.0574	0.0389	0.0421	0.0302	0.73345
ATM	0.0270	0.0237	0.0193	0.0174	0.71481
ITM	0.0151	0.0124	0.0097	0.0100	0.64238

3.2.J shows that the upfactor of the FI(T) model clearly exceeds the upfactor of the BL(T) model for all times to expiration. The average annual volatility of the BL(T) model is at most about 70% that of the FI(T) model. This relative difference is most pronounced for options with medium time to expirations.

Table 3.2.J: *Statistics on the upfactor for the FI(T) and BL(T) models. The last column contains the ratio of the average annual volatilty of BL(T) over average annual volatilty o FI(T)*

	Upfactor						
	BS(T)		FI(T)		BL(T)		
	mean	std	mean	std	mean	std	Ratio
S	1.0124	0.0037	1.0120	0.0037	1.0086	0.0033	0.7179
M	1.0122	0.0037	1.0118	0.0035	1.0073	0.0028	0.6200
L	1.0109	0.0030	1.0105	0.0029	1.0068	0.0029	0.6488

In Table 3.2.K statistics on the upfactor of the BS(M), FI(M) and BL(M) models are reported. The average upfactor of the BS(M) and FI(M) models exceeds the average upfactor of the BL(M) model for all intervals of moneyness. Furthermore, while the average upfactor of the BS(M) and FI(M) models increases in moneyness, the average upfactor of the BL(M) model is decreasing in moneyness. Thus, while the average annual volatility of the FI(M) model increases from 0.22 for OTM options to 0.24 for ITM options, the corresponding annual volatility of the BL(M)

model decrease from 0.202 for OTM options to 0.17 for ITM options.

Table 3.2.K: *Statistics on the upfactor for the BS(M), FI(M) and BL(M) models. The last column contains the ratio of the average annual volatilty of BL(M) over the average annual volatilty of FI(M)*

	Upfactor						
	BS(M)		FI(M)		BL(M)		
	mean	std	mean	std	mean	std	Ratio
OTM	1.0123	0.0034	1.0116	0.0034	1.0107	0.0037	0.9228
ATM	1.0125	0.0039	1.0119	0.0036	1.0107	0.0042	0.7742
ITM	1.0133	0.0046	1.0128	0.0045	1.0092	0.0041	0.7200

Probability of an upstep

The probabilities generated from the FI(T) and BL(T) models mimic the pattern found for the unstratified models in that on average the probabilities from the BL(T) model are considerably higher for all times to expiration. While the probabilities in general are higher for the BL(T) model for all times to maturity, they vary less than the probabilities from the FI(T) model. Statistics on the probabilities of the FI(T) model and the BL(T) model are given in Table 3.2.L.

Table 3.2.L: *Statistics on the probabilities of the FI(T) and BL(T) models*

	Probability					
	S		M		L	
	mean	std	mean	std	mean	std
FI(T)	0.4914	0.0217	0.4952	0.0161	0.4920	0.0170
BL(T)	0.5211	0.0141	0.5242	0.0117	0.5252	0.0122

Table 3.2.M displays statistics on the probabilities generated by the FI(M) and BL(M) models. The average probability of the BL(M) model is increasing in moneyness and exceeds the average probability of the FI(M) model for all intervals of moneyness. For all levels of moneyness,

the volatility of the BL(M) probabilities is lower than the volatility of the FI(M) probabilities.

Table 3.2.M: *Statistics on the probabilities of the FI(M) and BL(M) models*

	Probability					
	OTM		ATM		ITM	
	mean π	std π	mean π	std π	mean π	std π
FI(M)	0.4813	0.0220	0.4859	0.0191	0.4895	0.0194
BL(M)	0.5098	0.0149	0.5134	0.0147	0.5212	0.0163

Subjective discount factor

The estimates of subjective discount factors generated by the FI(T) model were used when the BL(T) model was adapted to data. Similarly, the estimates of subjective discount factors generated by the FI(M) model were used when the BL(M) model was estimated. Table 3.2.N reports statistics on the discount factor for the FI(T) model.

Table 3.2.N: *Statistics on the subjective discount factor generated by the FI(T) model*

	Discount factor		
	S	M	L
mean	0.9996	0.9997	0.9996
std	0.0004338	0.0003464	0.0003587

Table 3.2.O reports statistics on the discount factor for the FI(M) model. Lacking any theoretical predictions for the behavior of the discount factor, we simply note that the average discount factor of the FI(T) model exceeds the average discount factor of the FI(M) model.

Daily dividend

The estimation of the dividend d is performed completely analogous to the case of the unrestricted models. Thus, d is restricted to the interval $[0.05, 1.5]$ for the BS(M), BS(T),FI(M) and FI(T) models, while we set

Table 3.2.O: *Statistics on the subjective discount factor gener-
ated by the FI(M) model*

	Discount factor		
	OTM	ATM	ITM
mean	0.9994	0.9994	0.9995
std	0.0004535	0.0004882	0.0004789

$d = 0.05$ for the BL(T) and BL(M) models. Table 3.2.P shows statistics
on the estimation of d, along with information on the number of times
the restriction were binding.

Table 3.2.P: *Statistics on dividend for the BS/FI(T) and
BS/FI(M) models. lb is the number of times the
estimate was restricted by the lower bound 0.05
and ub is the number of times the estimate was
restricted by the upper bound 1.5*

	Dividend					
	S	M	L	OTM	ATM	ITM
mean	0.4834	0.4331	0.5022	0.7396	0.6675	0.5667
std	0.4915	0.4156	0.4613	0.5067	0.4889	0.4730
lb	70	31	8	29	27	32
ub	20	5	4	25	18	9

Strength of learning

When the unstratified learning model was fitted to data it was found that
learning effects were absent in more than half of the 261 dates, yielding an
average δ less than 0.005. However, when the learning model is estimated
on intervals of time to expiration and intervals of moneyness, learning
effects increase both in magnitude and presence. Table 3.2.Q presents
statistics on the strength of learning δ for both the BL(T) and the BL(M)
model. It also shows the number of dates for which δ could be estimated
for the various categories , and the number of times learning effects were
present for each category.

For the BL(T) model the average δ is increasing in time to expiration.

This observation is in line with the fact that the average $RRMSVE_t$ is decreasing in time to expiration. Learning effects were present in the vast majority of the dates for all times to expiration; the least percentage of times learning effects were present occurred for options with short time to expiration, where δ was positive more than 88 % of the dates.

For the BL(M) model the average δ is increasing in moneyness. Again, this is compatible with the results on $RRMSVE_t$, which were seen to be decreasing in moneyness. It also agrees with the fact that the improvement over the FI(M) model, both in terms of $RMSVE_t$ and $RRMSVE_t$, is increasing in moneyness.

Compared to the BL(T) model, learning is weaker and less frequent for the BL(M) model.

Table 3.2.Q: *Statistics on the strength of learnng for the BL(T) and BL(M) models.*

	Strength of learning					
	S	M	L	OTM	ATM	ITM
# dates	234	162	85	250	204	118
# $\delta > 0$	207	157	76	115	104	78
% $\delta > 0$	88.46	96.91	89.41	46	50.98	0.6610
max δ	0.0885	0.1086	0.0900	0.0900	0.0980	0.0900
std δ	0.0213	0.0216	0.0222	0.0159	0.0177	0.0220
mean δ	0.0237	0.0283	0.0246	0.0082	0.0088	0.0164

We conclude this section by comparing the measure of fit for the unstratified models and the stratified models. Obviously, the stratified models will always outperform their unstratified counterparts, at the cost of more parameters.

We get $RMSVE_t$ for the models stratified according to time to expiration by weighing the $RMSVE_t$ for each interval by the fraction of observations in that interval. Hence

$$RMSVE_t = \left[\sum_{k=S,M,L} \frac{\sum_{j=1}^{n_t} I_k(j)}{n_t} \left(RMSVE_t(k)\right)^2 \right]^{1/2}$$

where I_k , $k = S, M, L$ is an indicator function taking value 1 if observation j is an option with time to expiration belonging to interval k and 0 otherwise. $RMSVE_t$ for models stratified according to moneyness is defined similarly.

The relative measure of fit $RRMSVE_t$ for the models stratified according to time to expiration is computed by weighing the $RRMSVE_t$ for each interval by the sum of squared observed call prices belonging to the interval divided by the sum of all squared call prices for date t. Hence

$$RRMSVE_t = \left[\sum_{k=S,M,L} \frac{\sum C_{j,k}^2}{\sum_{j=1}^{n_t} C_j^2} \left(RRMSVE_t \left(k \right) \right)^2 \right]^{1/2}$$

where the summation $\sum C_{j,k}^2$ is done over all options belonging to interval k, $k = S, M, L$. The relative $RMSVE_t$ for the models stratified according to moneyness is computed similarly.

Table 3.2.R shows the measures of fit for all models. Since the BS models give virtually identical results as the corresponding FI models, only the results of the FI model are presented. The table also reports the percentage of weeks that each model was able to outperform its competitors. The results show that the model with Bayesian Learning stratified according to time to expiration clearly dominates the alternatives; the BL(T) model has the best overall fit with an $RRMSVE$ of 4.30, while it is also the best model in about 57% of the 261 weeks in the sample. However, it should be noted that comparing the fit of the unstratified models to the fit of the stratified models is an exercise seriously flawed by the lack of data. Also, the discrepancy in sample points between the stratified models conveyed by Table 3.2.E makes the use of RMSVE and RRMSVE questionable as tools for ranking the stratified models.

Timmermann and Guidolin [2003] applies the (exact) BL option pricing model, as well as several alternative models, to data on S&P 500 index option prices over the period June 1988-December 1993. The approach they adopt in estimating the BL model differs from the one adopted here. Thus, they fix U corresponding to an annual volatility of 5% and set the

Table 3.2.R: *RMSVE and RRMSVE for the each model*

	FI	Spline	BL	FI(T)	BL(T)	FI(M)	BL(M)
RMSVE	419.69	344.99	280.79	357.57	157.83	275.19	180.51
RRMSVE	11.42	9.43	7.80	9.49	4.30	7.68	5.06
% best	0	1.15	0	1.92	57.47	3.07	36.4

risk aversion and annual subjective discount factor equal to 0.9 and 0.98, respectively. Also, whereas the procedure adopted here yields, for each week, estimates of π, δ and U, Timmermann and Guidolin produces weekly estimates of π, N and the frequency with which new information arrives, m. As opposed to the findings here, showing the BL model to outperform the alternative models, they report that the BS spline model is the best fitting model in-sample, followed by the model due to Heston and Nandi [2000] and the BL model. The mean probability for the unstratified BL model in this study was found to be 0.5163, which is considerably lower than the mean probability of 0.5589 reported by Timmermann and Guidolin.

3.3 Out-of-Sample Results

While the previous exercise demonstrates the ability of the option pricing models to fit the data in-sample, it also provides estimates of the parameters for each Wednesday in the sample. We now proceed to test the predictive capabilities of the models out-of-sample and use estimates from each Wednesday t to compute the call prices at dates $t + 1$ and $t + 7$. Furthermore, we assume that the index value and relevant interest rate at date $t + s$ is known at time t when the prediction is made so that the predicted call price is $C_{i,t+s}^{Pred} = C(v_{t+s}, l_t)$, where $v_{t+s} = [S_{t+s}, K_{t+s}, r_{t+s}, T_{t+s}]$ and l_t are the parameters estimated at date t. When applying the stratified models to predict the call price at date $t + s$ it will sometimes occur that no corresponding set of estimates from the stratified model exists at date t; when this is the case the estimates from the unstratified models are used. To gauge the out-of-sample fit we compute the Relative Root Mean Squared Prediction Error for each predicted date, RRMSPE$_t$, defined analogously to RRMSVE$_t$.

Table 3.3.A reports the mean $RRMSPE_t$ for all models and prediction lengths. It also shows for each model the percentage of dates it was able to outperform its competitors, in terms of $RRMSVE_t$, among all models.

The figures in parenthesis is the percentage of times the model outperforms its competitors in the same class of models, the classes being the unstratified models, the models stratified according to time to expiration and the models stratified according to moneyness.

For the one-day-ahead prediction the $BL(T)$ model emerges as the dominant model, as it displays both the lowest mean $RRMSPE_t$ of all models, 0.0645, as well as being the best model in 21.79 % of the predicted dates, which is the highest percentage of all models. The lowest mean $RRMSPE_t$ among the unstratified models is obtained by the BL model which is also the best of the unstratified models in 45.53 % of the predicted dates. In general, the stratified models display a lower mean $RRMSPE_t$ than their unstratified counterparts, although this does not hold for the $BL(M)$ model. The poor performance of the BS-spline model indicates that parameter stability is clearly an issue for this model even for one-day-ahead predictions.

Note that for one-day-ahead predictions, the models with the lowest mean $RRMSPE_t$ preserve their ranking when the percentage of dates they outperform the competing models is considered.

Turning to one-week-ahead predictions, for each class of models, the model with Bayesian learning is superior to the corresponding alternatives when it comes to the percentage of dates the models are able to outperform its competitors.

However, for each class of models, the model with Bayesian learning display higher mean $RRMSPE_t$ than the corresponding BS model. The lowest mean $RRMSPE_t$ for each class of models is obtained by the FI models. Thus, the lowest mean $RRMSPE_t$ is 0.0759 and is obtained by the $FI(M)$ model. The $FI(M)$ model is able to outperform its competitors 14.11 % of the dates and is only surpassed in this regard by the $BL(T)$ model, which outperforms its competitors 14.52 % of the dates. Extending the prediction length has drastic consequences for the performance of the BS-spline model. While for the other models the mean $RRMSPE_t$ is marginally higher for the one-week-ahead predictions than

for the one-day-ahead predictions for all models, for the BS-spline model
the mean RRMSPE_t increases by a factor of 2.8. With five parameters
to estimate, the poor performance of the BS-spline model is attributed
to overfitting

Table 3.3.A: *Mean Realtive Root Mean Squared Prediction Er-*
ror (RRMSPE) for all models. "% best" is the
percentage of dates the model has lower RRMSPE
than any other model. The value in parenthesis is
the percentage of dates the model has lower RRM-
SPE than any other model in the same class.

	Mean RRMSPE_t,Aggregated			
	t+1		t+7	
	mean RRMSPE_t	% best	mean RRMSPE_t	% best
BS	0.0694	3.11 (10.12)	0.0780	7.66 (14.52)
FI	0.0691	5.84 (20.23)	0.0768	6.45 (20.16)
BL	0.0660	16.34 (45.53)	0.0787	13.71 (43.14)
BS-spline	0.0981	10.90 (24.12)	0.2725	9.27 (22.18)
BS(T)	0.0691	3.50 (15.95)	0.0810	3.23 (21.37)
FI(T)	0.0688	5.84 (27.24)	0.0801	6.85 (35.89)
BL(T)	0.0645	21.79 (56.81)	0.0897	14.52 (42.74)
BS(M)	0.0674	6.23 (19.46)	0.0776	10.89 (27.42)
FI(M)	0.0666	11.28 (35.02)	0.0759	14.11 (35.08)
BL(M)	0.0699	10.12 (45.52)	0.0853	9.27 (37.50)
# dates	257		248	

In order to assess the performance of the various option pricing models
for different contracts, Tables 3.3.B and 3.3.C report the mean RRMSPE_t
for each model across moneyness and time to expiration, for one-day-
ahead and one-week-ahead predictions, respectively.

For one-day-ahead predictions, the BL model is the best unstratified
model for all strata except for ITM calls. Its superiority is marked for
OTM contracts. The BL(T) model is the best model for ATM, S and M
calls, while the FI(M) model is the best model for OTM and L calls.

Table 3.3.B: *Average Relative Root Mean Prediction Error by moneyness and time to expiration for prediction one day ahead. " dates" is the number of dates predicted and "median calls" is the median number of calls at each predicted date.*

	Average RRMSPE by strata, $t+1$					
	OTM	ATM	ITM	S	M	L
BS	0.1624	0.0712	0.0413	0.0682	0.0681	0.0854
FI	0.1618	0.0707	0.0413	0.0679	0.0679	0.0846
BL	0.1316	0.0681	0.0421	0.0658	0.0619	0.0800
Spline	0.2163	0.1012	0.0640	0.0876	0.0666	0.1228
BS(T)	0.1621	0.0710	0.0406	0.0707	0.0648	0.0794
FI(T)	0.1613	0.0705	0.0407	0.0704	0.0643	0.0787
BL(T)	0.1234	0.0657	0.0432	0.0661	0.0582	0.0762
BS(M)	0.1218	0.0724	0.0459	0.0679	0.0650	0.0768
FI(M)	0.1205	0.0713	0.0458	0.0673	0.0642	0.0759
BL(M)	0.1270	0.0723	0.0448	0.0665	0.0699	0.0841
# dates	257	257	229	230	164	154
median calls	8	4	2	9	5	1

For one-week-ahead predictions, the BL model is the best unstratified model except for ATM and L calls, for which the FI model provides the best fit. The best model overall is the FI(M) model, except for ITM contracts, for which the BL model provides the best fit, and ATM contracts, for which the FI model provides the best fit.

In conclusion, for one-day-ahead predictions the BL(T) model provides the best mean fit, closely followed by the BL model. These models also give the best fit the highest percentage of the predicted dates. When the prediction length is extended to a week, the best model in terms of mean RRMSPE$_t$ is the FI(M) model, closely followed by the FI model. However, for each class of models, the model with Bayesian learning still provide the best fit the highest percentage of dates.

Furthermore, even for one-week-ahead predictions, the BL model is the best unstratified model for several types of contracts, especially for

OTM contracts. This is in agreement with the finding by Timmermann and Guidolin in their application of the BL model to data on the S&P500, showing that the BL model outperforms the competing models for OTM contracts.

An interesting observation is that for all prediction lengths and all classes of models, the FI model does slightly better out-of-sample, in terms of mean RRMSPE$_t$, than its BS counterpart. The advantage of the FI model is even more apparent when one considers the percentage of dates the models outperform the competitors; the FI model consistently scores higher in this regard. These results are not likely to be random: when predictions were made for 37 different prediction lengths, only once did the BS model obtain a lower mean RRMSPE$_t$ than the FI model.

Table 3.3.C: *Average Relative Root Mean Prediction Error by moneyness and time to expiration for prediction one week ahead. " dates" is the number of dates predicted and "median calls" is the median number of calls at each predicted date.*

	Average RRMSPE by strata, $t + 7$					
	OTM	ATM	ITM	S	M	L
BS	0.1738	0.0792	0.0445	0.0745	0.0830	0.0965
FI	0.1704	0.0779	0.0442	0.0739	0.0814	0.0941
BL	0.1596	0.0828	0.0440	0.0739	0.0804	0.1058
Spline	0.6776	0.2302	0.0637	0.1241	0.1404	0.4400
BS(T)	0.1800	0.0837	0.0445	0.0806	0.0822	0.0965
FI(T)	0.1785	0.0825	0.0444	0.0803	0.0813	0.0949
BL(T)	0.1834	0.0915	0.0471	0.0883	0.0839	0.1050
BS(M)	0.1462	0.0818	0.0495	0.0737	0.0812	0.0927
FI(M)	0.1426	0.0803	0.0488	0.0725	0.0786	0.0917
BL(M)	0.1644	0.0873	0.0492	0.0767	0.0893	0.1112
# dates	247	247	217	223	158	138
median calls	8	4	2	9	5	1

Summary

In this chapter we have performed in-sample tests of the BS, FI, BL and BS spline option pricing models. Among the unstratified models, it was seen that the best overall fit, as measured by RMSVE, was obtained by the BL model, which was only 2/3 that of the RMSVE of the BS model. Furthermore, the BL model displayed the best fit in 65% of the 261 dates, followed by the BS spline model which was the best model in 26% of the dates. The BL estimates of the upfactor U was in general considerably lower than the corresponding estimates for the BS and FI models. Although potentially interesting, a comparison of the BL and FI estimates of the probability π is hampered by the sensitivity of the probability to the dividend d.

The strength of learning δ in the unstratified model varied considerably over time, with only 48% of the estimates different from zero.

The BS, FI and BL models were also estimated on subsets according to time to expiration and moneyness. The averages of RMSVE_t and RRMSVE_t of the BL(T) model was about half that of the averages of RMSVE_t and RRMSVE_t of the BS(T)/FI(T) model, for all times to expiration. The strength of learning was considerably higher for all times to expiration than for the unstratified BL model, with an average δ in general 5 times that of the δ of the unstratified model. Furthermore, for the BL(T) model, δ was different from zero in 91% of the observations. These findings were related to the pattern of level and strength of smile described in Chapter 1.

Learning made less of a contribution in the models stratified according to moneyness, although it still provided a marked improvement. The improvement in terms of RMSVE_t and RRMSVE_t of the BL(M) model over the BS(M)/FI(M) model was increasing in moneyness, with an average RRMSVE_t 64% that of the average RRMSVE_t of the BS/FI model. The strength of learning for the BL(M) model was also increasing in moneyness and weaker than for the BL(T) model. Lack of data makes it difficult to use the overall measure of fit $RMSVE$ to compare the stratified models with the unstratified models, since the $RMSVE$ of the stratified models are biased downwards.

The out-of-sample tests of the option pricing models show that Bayesian learning is useful for one-day-ahead predictions, as $BL(T)$ is the best model, closely followed by the BL model. Bayesian learning is less successful for one-week-ahead predictions; here the $FI(M)$ model gives the lowest mean $RRMSPE_t$ and the FI model is a close second. For one-week-ahead predictions the stratified BL models gives the poorest fit, along with the BS-spline model, suggesting that parameter stability may be a concern.

APPENDIX

Appendix A The Discrete Black-Scholes Model

In order for the results from this study to be comparable with the implied volatilities from the discrete time option model with Bayesian learning, the discrete time BS (DBS) model is used to solve for the implied volatilities. We therefore want to make sure that the discrete BS provides a good approximation to its continuous counterpart. Specifically, we want to link the volatility implied by discrete BS to the annual standard deviation used in the continuous BS formula.

The model

The discrete BS takes as given an equilibrium resulting in a current date stock price S and a one period bond giving riskless return R. Denoting by U the up return and by D the down return, the stock price in equilibrium is assumed to either advance to SU or decline to SD with exogenous probabilities q and $1 - q$, respectively. The condition $U > R > D$ rules out the existence of arbitrage opportunities and provides the risk neutral probability of an upstep

$$p_1 = \frac{R - D}{U - D}$$

The up rate of return u is defined to be $\ln U$ and the down rate of return is defined to be $\ln D$.

One period expected rate of return and its variance

The expected return of the stock with respect to an exogenous probability q in the one period model is

$$q \ln U + (1 - q) \ln D = \mu_q$$

and the variance of the return is

$$q(1 - q)(\ln U - \ln D)^2 = \sigma_q^2$$

Solving for U and D we get

$$U = e^{\mu_q + \sigma_q \sqrt{\frac{(1-q)}{q}}}$$

$$D = e^{\mu_q - \sigma_q \sqrt{\frac{q}{(1-q)}}}$$

Dividing the given period into M subperiods

For the purpose of investigating the relationship between the parameters of the DBS and the moments of the return in the continuous BS model we now divide the one period model into M subperiods and denote by $\mathbb{M} = \{1, 2, \ldots, m, \ldots, M\}$ the set of subperiods. For the given q we want to choose subperiod returns U_M, D_M and R_M such that the expected return and variance of return of the one period model are maintained. An obvious choice is $R_M = e^{\ln R / M}$ and to get U_M and D_M we solve the equation system

$$\begin{cases} M \left[q \ln U_M + (1 - q) \ln D_M \right] & = & \mu_q \\ \\ M q (1 - q)(\ln U_M - \ln D_M)^2 & = & \sigma_q^2 \end{cases}$$

so that the moments from the one period model are intact. Doing this we obtain

$$U_M = e^{\frac{\mu_q}{M} + \frac{\sigma_q}{\sqrt{M}} \sqrt{\frac{(1-q)}{q}}}$$

$$D_M = e^{\frac{\mu_q}{M} - \frac{\sigma_q}{\sqrt{M}} \sqrt{\frac{q}{(1-q)}}} \qquad (*)$$

However, ensuring that the M-period model provides the same mean and variance as does the one-period model is not enough to show that the discrete BS model converges to the continuous BS model. For that purpose we define the risk neutral probability of an upstep to be

$$p_M = \frac{R_M - D_M}{U_M - D_M}$$

As a first step we show that the riskneutral probability in the M-period model converges to the exogenous probability q.

Convergence of riskneutral probabilities to exogenous

Let $a = \sigma_q^2 \left(1 - q\right)/q$ and $b = \sigma_q^2 q/\left(1 - q\right)$. Then

$$p_M = \frac{R_M - D_M}{U_M - D_M} = \frac{e^{r/M} - e^{\mu_q/M - \sqrt{b/M}}}{e^{\mu_q/M + \sqrt{a/M}} - e^{\mu_q/M - \sqrt{b/M}}}$$

where $r = \ln R_M$. Letting $x = 1/M$, $f\left(x\right) = e^{rx} - e^{\mu_q x - \sqrt{bx}}$, $g\left(x\right) = e^{\mu_q x + \sqrt{ax}} - e^{\mu_q x - \sqrt{bx}}$ and multiplying with $e^{-x\mu}$ yields

$$p_M = \frac{e^{x(r-\mu)} - e^{-\sqrt{bx}}}{e^{\sqrt{ax}} - e^{-\sqrt{bx}}} = \frac{f(x)}{g(x)}$$

In order to apply l'Hopitals rule to determine the limit of p_M as M tends to infinity we calculate the derivatives

$$f'(x) = (r - \mu)e^{x(r-\mu)} + \frac{\sqrt{b}}{2\sqrt{x}}e^{-\sqrt{bx}}$$

$$g'(x) = \frac{\sqrt{a}}{2\sqrt{x}}e^{\sqrt{ax}} + \frac{\sqrt{b}}{2\sqrt{x}}e^{-\sqrt{bx}}$$

so that

$$\frac{f'(x)}{g'(x)} = \frac{(r - \mu)e^{x(r-\mu)} + \frac{\sqrt{b}}{2\sqrt{x}}e^{-\sqrt{bx}}}{\frac{\sqrt{a}}{2\sqrt{x}}e^{\sqrt{ax}} + \frac{\sqrt{b}}{2\sqrt{x}}e^{-\sqrt{bx}}} = \frac{2\sqrt{x}(r - \mu)e^{x(r-\mu)} + \sqrt{b}e^{-\sqrt{bx}}}{\sqrt{a}e^{\sqrt{ax}} + \sqrt{b}e^{-\sqrt{bx}}}$$

Hence

$$\lim_{M \to \infty} p_M = \frac{f'(0)}{g'(0)} = \frac{0 + \sqrt{b}}{\sqrt{a} + \sqrt{b}}$$

$$= \frac{\sigma \dfrac{\sqrt{q}}{\sqrt{1-q}}}{\sigma \left[\dfrac{\sqrt{1-q}}{\sqrt{q}} + \dfrac{\sqrt{q}}{\sqrt{1-q}} \right]}$$

$$= \frac{\sqrt{q}\sqrt{q}\sqrt{1-q}}{\sqrt{1-q}} = q$$

Consequently, as $M \to \infty$ the riskneutral distribution approaches the fixed value of q. Figure A illustrates the convergence to $q = 1/2$ when $\mu = 0$, $r = 0.04$ and $\sigma = 0.3$.
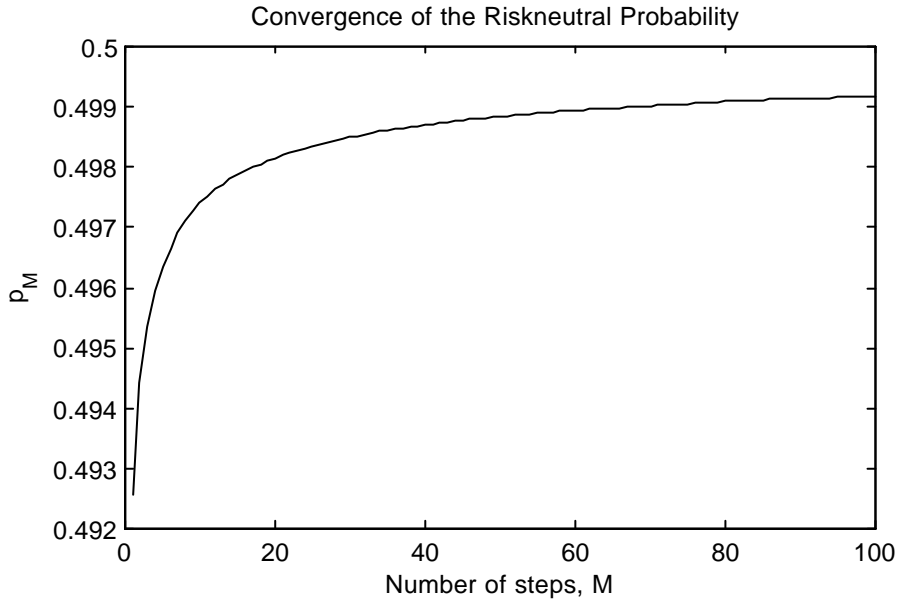

Figure A: *Convergence of the riskneutral probability as number of steps increases, when $\mu = 0$, $r = 0.04$ and $\sigma = 0.3$*

Moments of the limiting distribution

Equipped with the riskneutral probabilities p_M we can now show that the limiting value of the expected return and the variance of return with respect to the riskneutral probability measure p. Thus we will see that μ_p tends to $r - (1/2)\sigma_q^2$ and σ_p^2 tends to the variance with respect to

the subjective probability measure q. For this purpose we define the random variable $Y_M = z u_M + (1 - z) d_M$ where $z \in \text{Bin}(p_M, 1)$, enabling us to write the subperiod expected rate of return with respect to the probability p_M as

$$
\begin{aligned}
E[Y_M] &= E[z u_M + (1 - z) d_M] \\
&= E[z u_M] + E[(1 - z) d_M] = u_M p_M + d_M(1 - p_M)
\end{aligned}
$$

where $\ln U_M = u_M$ and $\ln D_M = d_M$. Since $V[z] = p_M(1 - p_M)$ and the z's are iid, the riskneutral variance of the subperiod rate of return is

$$
\begin{aligned}
V[Y_M] &= V[z u_M + (1 - z) d_M] \\
&= V[z u_M] + V[(1 - z) d_M] + 2\text{cov}[z u_M, (1 - z) d_M] \\
&= u_M^2 V[z] + d_M^2 V[(1 - z)] + 2 u_M d_M \text{cov}[z, (1 - z)] \\
&= (u_M^2 + d_M^2) V[z] + 2 u_M d_M (\text{cov}[z, 1] - \text{cov}[z, z]) \\
&= (u_M^2 + d_M^2) V[z] - 2 u_M d_M V[z] \\
&= (u_M - d_M)^2 V[z] \\
&= (u_M - d_M)^2 p_M (1 - p_M)
\end{aligned}
$$

Turning now to the riskneutral expected return and variance over all M periods we first note that since the return over all M periods is $Z_M = \sum_{i \in \mathbb{M}} Y_{M,i}$ where each $Y_{M,i}, i \in \mathbb{M}$, is distributed as Y_M we have

$$
E[Z_M] = E\left[\sum_{i \in \mathbb{M}} Y_{M,i}\right] = M E[Y_M] = M(p_M u_M + (1 - p_M) d_M)
$$

and

$$
V[Z_M] = V\left[\sum_{i \in \mathbb{M}} Y_{M,i}\right] = M V[Y_M] = M(u_M - d_M)^2 p_M (1 - p_M)
$$

It now remains to evaluate these expressions as $M \to \infty$.

Convergence of variance of the return

Recalling relation $(*)$ defining U_M and D_M we have

$$
(\ln U_M - \ln D_M)^2 = \frac{\sigma_q^2}{q(1 - q)}
$$

Hence

$$u_M - d_M = \frac{\sigma_q}{\sqrt{M}} \left(\sqrt{\frac{(1-q)}{q}} + \sqrt{\frac{q}{1-q}} \right) = \frac{\sigma_q}{\sqrt{M}} \frac{1}{\sqrt{q(1-q)}}$$

and since $p_M \to q$ as M tends to ∞ we get

$$\lim_{M \to \infty} V[Z_M] = \lim_{M \to \infty} M(u_M - d_M)^2 p_M(1 - p_M)$$
$$= \lim_{M \to \infty} \sigma_q^2 \frac{1}{q(1-q)} p_M(1 - p_M) = \sigma_q^2$$

Convergence of expectation of the return

Again letting $a = \sigma_q^2(1-q)/q$ and $b = \sigma_q^2 q/(1-q)$ the subperiod rates of return can be written

$$u = \frac{\mu_q}{M} + \sqrt{a/M}$$
$$d = \frac{\mu_q}{M} - \sqrt{b/M}$$

so that the riskneutral probability of an upstep an downstep becomes

$$p_M = \frac{R_M - D_M}{U_M - D_M} = \frac{e^{r/M} - e^{\mu_q/M - \sqrt{b/M}}}{e^{\mu_q/M + \sqrt{a/M}} - e^{\mu_q/M - \sqrt{b/M}}}$$
$$1 - p_M = \frac{U_M - R_M}{U_M - D_M} = \frac{e^{\mu_q/M + \sqrt{a/M}} - e^{r/M}}{e^{\mu_q/M + \sqrt{a/M}} - e^{\mu_q/M - \sqrt{b/M}}}$$

so that

$$E[Z_M] = M(p_M u_M + (1 - p_M)d_M)$$
$$= M\left[p_M\left(\mu_q/M + \sqrt{a}/\sqrt{M}\right) + (1 - p_M)\left(\mu_q/M - \sqrt{b}/\sqrt{M}\right) \right]$$
$$= \mu_q + \sqrt{M}\left(p_M\sqrt{a} - (1 - p_M)\sqrt{b} \right)$$

Putting $x = 1/\sqrt{M}$ we get

$$\sqrt{M}\left(p_M\sqrt{a}-(1-p_M)\sqrt{b}\right)$$

$$=\sqrt{M}\left(\frac{\sqrt{a}\left(e^{r/M}-e^{\mu_q/M-\sqrt{b/M}}\right)-\sqrt{b}\left(e^{\mu_q/M+\sqrt{a/M}}-e^{r/M}\right)}{e^{\mu_q/M+\sqrt{a/M}}-e^{\mu_q/M-\sqrt{b/M}}}\right)$$

$$=\sqrt{M}\left(\frac{\sqrt{a}\left(e^{r/M}-e^{\mu_q/M-\sqrt{b/M}}\right)-\sqrt{b}\left(e^{\mu_q/M+\sqrt{a/M}}-e^{r/M}\right)}{e^{\mu_q/M+\sqrt{a/M}}-e^{\mu_q/M-\sqrt{b/M}}}\right)$$

$$=\sqrt{M}\left(\frac{\sqrt{a}\left(e^{(r-\mu_q)/M}-e^{-\sqrt{b/M}}\right)-\sqrt{b}\left(e^{\sqrt{a/M}}-e^{(r-\mu_q)/M}\right)}{e^{\sqrt{a/M}}-e^{-\sqrt{b/M}}}\right)$$

$$=\frac{\left(\sqrt{a}+\sqrt{b}\right)e^{x^2(r-\mu_q)}-\left(\sqrt{a}e^{-x\sqrt{b}}+\sqrt{b}e^{x\sqrt{a}}\right)}{x\left(e^{x\sqrt{a}}-e^{-x\sqrt{b}}\right)}$$

Let f and g denote the functions with values given by the numerator and the denominator, respectively. Then

$$f'(x) = \left(\sqrt{a}+\sqrt{b}\right)2x\left(r-\mu_q\right)e^{x^2(r-\mu_q)}+\left(\sqrt{ab}e^{-x\sqrt{b}}-\sqrt{ab}e^{x\sqrt{a}}\right)$$

$$g'(x) = \left(e^{x\sqrt{a}}-e^{-x\sqrt{b}}\right)+x\left(\sqrt{a}e^{x\sqrt{a}}+\sqrt{b}e^{-x\sqrt{b}}\right)$$

Using l'Hopital's rule repeatedly we get

$$\lim_{x\to0}\frac{f'(x)}{g'(x)}$$

$$=\lim_{x\to0}\frac{\left(\sqrt{a}+\sqrt{b}\right)2\left(r-\mu_q\right)e^{x^2(r-\mu_q)}+\dfrac{\left(\sqrt{ab}e^{-x\sqrt{b}}-\sqrt{ab}e^{x\sqrt{a}}\right)}{x}}{\dfrac{\left(e^{x\sqrt{a}}-e^{-x\sqrt{b}}\right)}{x}+\left(\sqrt{a}e^{x\sqrt{a}}+\sqrt{b}e^{-x\sqrt{b}}\right)}$$

$$=\frac{\left(\sqrt{a}+\sqrt{b}\right)2\left(r-\mu_q\right)-\left(a\sqrt{b}+b\sqrt{a}\right)}{\left(\sqrt{a}+\sqrt{b}\right)+\left(\sqrt{a}+\sqrt{b}\right)}$$

$$=\left(r-\mu_q\right)-\frac{1}{2}\frac{\left(a\sqrt{b}+b\sqrt{a}\right)}{\left(\sqrt{a}+\sqrt{b}\right)}=\left(r-\mu_q\right)-\frac{1}{2}\sigma_q^2$$

so that $E\left[Z_M\right]$ tends to $r-(1/2)\sigma_q^2$ as M tends to infinity.

Although this expression does not explicitly contain the expected return,μ_q, calculated with respect to the probability q, it does depend on the expected return as perceived by a riskneutral agent. This is so because

$$pU + (1-p)D \quad = \quad \frac{R-D}{U-D}U + \frac{U-R}{U-D}D \quad = \quad R$$

and $\ln R = r$.

Convergence in distribution

The stochastic stock price at date 1 after dividing into M subperiods is

$$SU_M^k D_M^{M-k} = Se^{ku_M + (M-k)d_M} = Se^{Z_M}$$

where the number of upsteps $k \in \text{Bin}(p_M, M)$.

Let $x > 0$. We have

$$Pr\left(S_1 \le x\right) = \ Pr\left(Se^{Z_M} \le x\right) = \ Pr\left(Z_M \le \ln x - \ln S\right)$$

so that

$$Pr\left(\frac{Z_M - E\left(Z_M\right)}{\sigma_q} \le \frac{(\ln x - \ln S) - \left(r - \frac{1}{2}\sigma_q^2\right)}{\sigma_q}\right)$$

Using a generalization of the Central Limit Theorem, due to Lindeberg, it can be shown this probability tends to

$$\Phi\left(\frac{(\ln x - \ln S) - \left(r - \frac{1}{2}\sigma_q^2\right)}{\sigma_q}\right)$$

as M tends to infinity, where Φ is the Standard Normal Distribution.

Analogously the stock price at time T is distributed as the limit of Se^{TZ_M} as M tends to infinity. The limit of TZ_M has expectation $\left(r - (1/2)\sigma_q^2\right)T$ and standard deviation $\sigma_q T$.

Appendix B   Spot-Futures Parity

Future on stock with cash dividend; two dates

As a preliminary we first derive the price quoted at date t of a future with delivery date T, at the agreed price $F_{t,T}$, when the dividends on the stock are known and non-stochastic with date T- value D_T. We thus consider an economy with two dates and three assets: a futures contract stipulating the owner to pay the amount $F_{t,T}$ at date T, a bond with date-t price 1 each period giving return $(1+r)$ and a stock with postdividend prices S_t and S_T, respectively. The extended payoff matrix W (Cf Borglin [2004]) can then be written as

$$W = \begin{bmatrix} -S_t & -1 & 0 \\ S_T + D_T & (1+r)^{(T-t)} & S_T - F_{t,T} \end{bmatrix}$$

where the first row contains the amounts to be delivered at date t and the second row the payoffs at date T.

Now consider the portfolio θ

$$\theta = \begin{bmatrix} -1 \\ (1+r)^{-(T-t)}(F_{t,T} + D_T) \\ 1 \end{bmatrix}$$

resulting in the net income vector $W\theta$ equal to

$$\begin{bmatrix} -S_t & -1 & 0 \\ S_T + D_T & (1+r)^{(T-t)} & S_T - F_{t,T} \end{bmatrix} \begin{bmatrix} -1 \\ (1+r)^{-(T-t)}(F_{t,T} + D_T) \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} S_t - (1+r)^{-(T-t)}(F_{t,T} + D_T) \\ 0 \end{bmatrix}$$

Thus, in the absence of arbitrage we have

$$S_t - (1+r)^{(T-t)}(F_{t,T} + D_T) \quad = \quad 0$$

giving the price of the future

$$F_{t,T} \quad = \quad S_t(1+r)^{-(T-t)} - D_T$$

Letting $\ln(1 + r) = \rho$ we can write the future price as

$$F_{t,T} \quad = \quad S_t e^{(T-t)\rho} - D_T$$

Future on stock with cash dividend; several dates

When calculating the futures price on a stock index it is customary to view the index as a security that provides a known dividend yield, meaning that one assumes, in the discrete case, that the cash dividends are a constant proportion d of the predividend stock price. The continuous dividend yield is then defined by $\delta = \ln(1 - d)$.

Although similar to the case where the amount of cash dividend to be paid during the timespan is known, the arbitrage/replicating strategy here differs somewhat in that one does not buy one unit of the stock at time t, but rather a fraction $(1 - d)^{T-t}$. By repeatedly reinvesting the cash dividends in the stock, a holding of one unit of the stock at date T is accomplished.

We let an example with three dates illustrate the reasoning. Firstly, assume that at date t we buy one unit of the stock at the price S_t. Then the wealth in stock entering the date $t + 1$ node equals the predividend price, which is either $S_t U$ or $S_t D$. Furthermore, since the cash dividend is either $S_t U d$ or $S_t D d$ wealth is by dividends split up at $t + 1$

$$\begin{cases} S_t U & = & \underbrace{S_t U(1 - d)}_{post-dividend\ price} & + & \underbrace{S_t U d}_{cash\ dividend} \\[2em] S_t D & = & \underbrace{S_t D(1 - d)}_{post-dividend\ price} & + & \underbrace{S_t D d}_{cash\ dividend} \end{cases}$$

Accordingly, the postdividend price at $t+1$ is either $S_t U(1-d)$ or $S_t D(1-d)$ so that the amount of stock leaving the date $t + 1$ node is either

$$\frac{S_t U}{S_t U(1 - d)} = \frac{1}{(1 - d)} \quad \text{or} \quad \frac{S_t D}{S_t D(1 - d)} = \frac{1}{(1 - d)}$$

Entering the $t + 2$ node, the pre-dividend prices will be $S_t U^2(1 - d)$, $S_t U D(1 - d)$ or $S_t D^2(1 - d)$. Since we arrive at the date $t + 2$ node holding $1/(1 - d)$ of the stock, the wealth at date $t + 2$ is

$$
\begin{cases}
S_t U^2 & = & S_t U^2 (1-d) & + & S_t U^2 d \\
S_t U D & = & S_t U D (1-d) & + & S_t U D d \\
S_t D^2 & = & S_t D^2 (1-d) & + & S_t D^2 d
\end{cases}
$$

corresponding to $t+2$ postdividend prices $S_t U^2 (1-d)^2$, $S_t U D (1-d)^2$ and $S_t D^2 (1-d)^2$ implying that the amounts of stock at postdividend prices are

$$
\begin{aligned}
\frac{S_t U^2}{S_t U^2 (1-d)^2} &= \frac{1}{(1-d)^2} \\
\frac{S_t U D}{S_t U D (1-d)^2} &= \frac{1}{(1-d)^2} \\
\frac{S_t D^2}{S_t D^2 (1-d)^2} &= \frac{1}{(1-d)^2}
\end{aligned}
$$

We thus conclude that buying the amount $(1-d)^2$ of the stock at date t, and repeatedly reinvesting the cash dividends in the stock gives precisely 1 unit of the stock after dividends at date $t+2$. Generally, buying the amount $(1-d)^{T-t}$ of the stock at date t will result in a holding of one unit of the stock at date T.

To derive the arbitrage-free price of the future we consider the extended payoff matrix for three dates

$$
W =
\begin{bmatrix}
-S_t & -1 & 0 \\
\vdots & \vdots & \vdots \\
S_t U^2 & (1+r)^2 & -F_{t,T} + S_t U^2 (1-d)^2 \\
S_t U D & (1+r)^2 & -F_{t,T} + S_t U D (1-d)^2 \\
S_t D U & (1+r)^2 & -F_{t,T} + S_t D U (1-d)^2 \\
S_t D^2 & (1+r)^2 & -F_{t,T} + S_t D^2 (1-d)^2
\end{bmatrix}
$$

where only the payoffs for the initial date and last date have been written out. With several periods the portfolio has to be adjusted after each dividend date. Here we take the view of an investor who at the initial

date acquired the portfolio θ

$$\theta = \begin{bmatrix} -(1-d)^2 \\ \dfrac{1}{(1+r)^2} F_{t,T} \\ 1 \end{bmatrix}$$

Reinvesting the dividends in the stock will thus after two periods guarantee a holding of one unit of the stock. In the case of two upsteps, the final value of the portfolio θ is 0 since

$$-S_t U^2 (1-d)^2 + F_{t,T} - F_{t,T} + S_t U^2 (1-d)^2 \quad = \quad 0$$

and it is easy to check that this will also hold for the other three cases. Thus, assuming no arbitrage, the initial cost of this portfolio must be 0 so that

$$S_t (1-d)^2 \quad = \quad \frac{1}{(1+r)^2} F_{t,T}$$

In the general case, we have for an initial date t and final date T

$$S_t (1-d)^{(T-t)} \quad = \quad \frac{1}{(1+r)^{(T-t)}} F_{t,T}$$

implying that

$$F_{t,T} \quad = \quad S_t [(1-d)(1+r)]^{(T-t)}$$

which is equivalent to

$$F_{t,T} \quad = \quad S_t e^{(T-t)[\ln(1-d)+\ln(1+r)]}$$

Spot-future parity

Letting $\ln(1-d) = \delta$ and $\ln(1+r) = \rho$ we can thus state *spot-future parity* as

$$F_{t,T} = S_t e^{(T-t)(\rho+\delta)}$$

From this we can solve for the dividend intensity δ :

$$\delta = \frac{1}{(T-t)} \left[\ln F_{t,T} - \ln S_t\right] - \rho$$

and thus the spot-future parity can be used to infer the dividend intensity.

Appendix C   Put-Call Parity

Stock with known cash dividend; several dates

Analogously to the case of spot-futures parity, we first derive the arbitrage relation known as *put-call parity* when the dividends on the stock are known with date T- value D_T. Denoting the current call and put prices by C_t and P_t, the extended payoff matrix W is given by

Date	Stock	Bk acc	Call opt	Put opt
0	$-S_t$	-1	$-C_t$	$-P_t$
T	$S_T + D_T$	$(1+r)^{(T-t)}$	$(S_T - K)^+$	$(K - S_T)^+$

Thus

$$W = \begin{bmatrix} -S_t & -1 & -C_t & -P_t \\ S_T + D_T & (1+r)^{(T-t)} & (S_T - K)^+ & (K - S_T)^+ \end{bmatrix}$$

where $(S_T - K)^+ = \max(0, S_T - K)$. By acquiring the portfolio

$$\theta = \begin{bmatrix} -1 \\ \dfrac{K + D_T}{(1+r)^{(T-t)}} \\ 1 \\ -1 \end{bmatrix}$$

one gets the net income vector $W\theta$ given by

$$\begin{bmatrix} -S_t & -1 & -C_t & -P_t \\ S_T + D_T & (1+r)^{(T-t)} & (S_T - K)^+ & (K - S_T)^+ \end{bmatrix} \begin{bmatrix} -1 \\ \dfrac{K + D_T}{(1+r)^{(T-t)}} \\ 1 \\ -1 \end{bmatrix}$$

$$= \begin{bmatrix} S_t - \dfrac{1}{(1+r)^{(T-t)}} (K + D_T) - C + P \\ -S_T + K + (S_T - K)^+ - K - (K - S_T)^+ \end{bmatrix}$$

It is easily verified that regardless of whether $S_T > K$ or $S_T \leq K$, the last row, corresponding to the final dividend at date T, will equal 0. Hence,

in the absence of arbitrage, we have

$$S_t - \frac{K + D_T}{(1+r)^{(T-t)}} - C + P \;=\; 0$$

so that

$$C = S_t + P - \frac{K + D_T}{(1+r)^{(T-t)}}$$

Letting $\ln(1+r) = \rho$, we can state put-call parity as

$$C = S_t + P - e^{-(T-t)\rho}\left(K + D_T\right)$$

Stock with known cash dividend yield; several dates

Again letting d denote the constant proportion of the predividend stock price and $\hat{d} = 1 - d$, the extended payoff matrix, W, with three dates is given by

Date/State	Stock	Bk acc	Call opt	Put opt
0	$-S_t$	-1	$-C_t$	$-P_t$
1	0	0	0	0
2 — 1	$S_t U^2$	$(1+r)^2$	$(S_t U^2 \hat{d}^2 - K)^+$	$(K - S_t U^2 \hat{d}^2)^+$
2 — 2	$S_t UD$	$(1+r)^2$	$(S_t UD \hat{d}^2 - K)^+$	$(K - S_t UD \hat{d}^2)^+$
2 — 3	$S_t DU$	$(1+r)^2$	$(S_t DU \hat{d}^2 - K)^+$	$(K - S_t DU \hat{d}^2)^+$
2 — 4	$S_t D^2$	$(1+r)^2$	$(S_t D^2 \hat{d}^2 - K)^+$	$(K - S_t D^2 \hat{d}^2)^+$

The holding of one unit of the stock at the final date is achieved by repeatedly reinvesting the dividends in the stock. Acquiring at the initial date the portfolio

$$\theta = \begin{bmatrix} -(1-d)^2 \\ \dfrac{K}{(1+r)^2} \\ 1 \\ -1 \end{bmatrix}$$

gives a date 0 cost of

$$S_t(1-d)^2 - \frac{K}{(1+r)^2} - C - P$$

The portfolio θ gives after two upsteps the net income

$$-S_t U^2 (1-d) + K + (S_t U^2 (1-d)^2 - K)^+ - (K - S_t U^2 (1-d)^2)^+ = 0$$

It is easily verified that at the final date, also for the states $2, 3$ and 4, the net income received at date 2 is zero. Hence, in order to avoid arbitrage, the date 0 cost of the portfolio θ is zero.

Generally, for an initial date t and final date T we thus have

$$S_t (1-d)^{(T-t)} - \frac{K}{(1+r)^{(T-t)}} - C - P = 0$$

Again letting $\ln(1-d) = \delta$ and $\ln(1+r) = \rho$ we can write put-call parity for the case when dividends are a known proportion of the stock price as

$$S_t e^{(T-t)\delta} - K e^{-(T-t)\rho} - C - P = 0$$

Solving for δ, we get the dividend yield inferred by put-call parity:

$$\delta = \frac{1}{(T-t)} \ln \frac{e^{-\rho(T-t)} K + C - P}{S_t}$$

Bibliography

ABEL, A. (1988): "Stock Prices Under Time-Varying Dividend Risk. An Exact Solution in an Infinite-Horizon General Equilibrium Model," *Journal of Monetary Economics*, 22, 375–393.

BARTON, D., AND F. DAVID (1962): *Combinatorial Chance.* Charles Griffin Co., London.

BATES, D. (2003): "Empirical Option Pricing: A Retrospection," *Journal of Econometrics*, 116.

BLACK, F., AND M. SCHOLES (1973): "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637–654.

BORGLIN, A. (2004): *Economic Dynamics and General Equilibrium.* Springer Verlag.

CASELLA, G., AND R. BERGER (2002): *Statistical Inference.* Duxbury.

COX, J., S. ROSS, AND M. RUBINSTEIN (1979): "Option Pricing: A Simplified Approach," *Journal of Financial Economics*, 7, 229–263.

DAS, S., AND R. SUNDARAM (1999): "Of Smiles and Smirks: a Term Structure Perspective," *Journal of Financial and Quantitative Analysis*, 34, 211–239.

DEGROOT, M. H. (1986): *Probability and Statistics.* Addison Wesley Publishing Company.

D.LANDO, AND R. POULSEN (2006): "Topics in Finance," *Lecture notes, Copenhagen.*

EDGERINGTON, L., AND W. GUAN (2002): "Why Are Those Options Smiling?," *Journal of Derivatives*, 10, 9–34.

FISZ, M. (1963): *Probability Theory and Mathematical Statistics*. Wiley.

GUIDOLIN, M., AND A. TIMMERMANN (2003): "Option prices under Bayesian learning: implied volatility dynamics and predictive densities," *Journal of Economic Dynamics and Control*, 27, 717–769.

HESTON, S., AND S. NANDI (2000): "A Closed-form GARCH Option Valuation Model," *Review of Financial Studies*, 13, 585–625.

JOHNSON, N. L., AND S. KOTZ (1977): *Urn Models and Their Applications*. Wiley.

LEROY, S., AND J. WERNER (2001): *Principles of Financial Economics*. Cambridge University Press.

LUCAS, R. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46, 1429–1445.

Lund Economic Studies

1. Guy Arvidsson — Bidrag till teorin för verkningarna av räntevariationer, 1962

2. Björn Thalberg — A Trade Cycle Analysis. Extensions of the Goodwin Model, 1966

3. Bengt Höglund — Modell och observationer. En studie av empirisk anknytning och aggregation för en linjär produktionsmodell, 1968

4. Alf Carling — Industrins struktur och konkurrensförhållanden, 1968

5. Tony Hagström — Kreditmarknadens struktur och funktionssätt, 1968

6. Göran Skogh — Straffrätt och samhällsekonomi, 1973

7. Ulf Jakobsson och Göran Norman — Inkomstbeskattningen i den ekonomiska politiken. En kvantitativ analys av systemet för personlig inkomstbeskattning 1952-71, 1974

8. Eskil Wadensjö — Immigration och samhällsekonomi. Immigrationens ekonomiska orsaker och effekter, 1973

9. Rögnvaldur Hannesson — Economics of Fisheries. Some Problems of Efficiency, 1974

10. Charles Stuart — Search and the Organization of Marketplaces, 1975

11. S Enone Metuge — An Input-Output Study of the Structure and Resource Use in the Cameroon Economy, 1976

12. Bengt Jönsson — Cost-Benefit Analysis in Public Health and Medical Care, 1976

13. Agneta Kruse och Ann-Charlotte Ståhlberg — Effekter av ATP - en samhällsekonomisk studie, 1977

14. Krister Hjalte — Sjörestaureringens ekonomi, 1977

15. Lars-Gunnar Svensson — Social Justice and Fair Distributions, 1977

16. Curt Wells — Optimal Fiscal and Monetary Policy - Experiments with an Econometric Model of Sweden, 1978

17. Karl Lidgren — Dryckesförpackningar och miljöpolitik - En studie av styrmedel, 1978

18. Mats Lundahl — Peasants and Poverty. A Study of Haiti, London, 1979

19. Inga Persson-Tanimura — Studier kring arbetsmarknad och information, 1980

20. Bengt Turner — Hyressättning på bostadsmarknaden - Från hyresreglering till bruksvärdesprövning, Stockholm 1979

21. Ingemar Hansson — Market Adjustment and Investment Determination. A Theoretical Analysis of the Firm and the Industry, Stockholm 1981

22. Daniel Boda Ndlela — Dualism in the Rhodesian Colonial Economy, 1981

23. Tom Alberts | Agrarian Reform and Rural Poverty: A Case Study of Peru, 1981

24. Björn Lindgren | Costs of Illness in Sweden 1964-75, 1981

25. Göte Hansson | Social Clauses and International Trade. An Economic Analysis of Labour Standards in Trade Policy, 1981

26. Noman Kanafani | Oil and Development. A Case Study of Iraq, 1982

27. Jan Ekberg | Inkomsteffekter av invandring, 1983

28. Stefan Hedlund | Crisis in Soviet Agriculture?, 1983

29. Ann-Marie Pålsson | Hushållen och kreditpolitiken. En studie av kreditrestriktioners effekt på hushållens konsumtion, sparande och konsumtionsmönster, 1983

30. Lennart Petersson | Svensk utrikeshandel, 1871-1980. En studie i den intraindustriella handelns framväxt, 1984

31. Bengt Assarsson | Inflation and Relative Prices in an Open Economy, 1984

32. Claudio Vedovato | Politics, Foreign Trade and Economic Development in the Dominican Republic, 1985

33. Knut Ödegaard | Cash Crop versus Food Crop Production in Tanzania: An Assessment of the Major Post-Colonial Trends, 1985

34. Vassilios Vlachos | Temporära lönesubventioner. En studie av ett arbetsmarknadspolitiskt medel, 1985

35. Stig Tegle | Part-Time Employment. An Economic Analysis of Weekly Working Hours in Sweden 1963-1982, 1985

36. Peter Stenkula | Tre studier över resursanvändningen i högskolan, 1985

37. Carl Hampus Lyttkens | Swedish Work Environment Policy. An Economic Analysis, 1985

38. Per-Olof Bjuggren | A Transaction Cost Approach to Vertical Integration: The Case of Swedish Pulp and Paper Industry, 1985

39. Jan Petersson | Erik Lindahl och Stockholmsskolans dynamiska metod, 1987

40. Yves Bourdet | International Integration, Market Structure and Prices. A Case Study of the West-European Passenger Car Industry, 1987

41. Krister Andersson and Erik Norrman | Capital Taxation and Neutrality. A study of tax wedges with special reference to Sweden, 1987

42. Tohmas Karlsson | A Macroeconomic Disequilibrium Model. An Econometric Study of the Swedish Business Sector 1970-84, 1987

43. Rosemary Vargas-Lundius | Peasants in Distress. Poverty and Unemployment in the Dominican Republic, 1989

44. Lena Ekelund Axelson | Structural Changes in the Swedish Marketing of Vegetables, 1991

45.	Elias Kazarian	Finance and Economic Development: Islamic Banking in Egypt, 1991
46.	Anders Danielson	Public Sector Expansion and Economic Development. The Sources and Consequences of Development Finance in Jamaica 1962-84, 1991
47.	Johan Torstensson	Factor Endowments, Product Differentiation, and International Trade, 1992
48.	Tarmo Haavisto	Money and Economic Activity in Finland, 1866-1985, 1992
49.	Ulf Grönkvist	Economic Methodology. Patterns of Reasoning and the Structure of Theories, 1992
50.	Evelyne Hangali Maje	Monetization, Financial Development and the Demand for Money, 1992
51.	Michael Bergman	Essays on Economic Fluctuations, 1992
52.	Flora Mndeme Musonda	Development Strategy and Manufactured Exports in Tanzania, 1992
53.	Håkan J. Holm	Complexity in Economic Theory. An Automata Theoretical Approach, 1993
54.	Klas Fregert	Wage Contracts, Policy Regimes and Business Cycles. A Contractual History of Sweden 1908-90, 1994
55.	Per Frennberg	Essays on Stock Price Behaviour in Sweden, 1994
56.	Lisbeth Hellvin	Trade and Specialization in Asia, 1994
57.	Sören Höjgård	Long-term Unemployment in a Full Employment Economy, 1994
58.	Karolina Ekholm	Multinational Production and Trade in Technological Knowledge, 1995
59.	Fredrik Andersson	Essays in the Economics of Asymmetric Information, 1995
60.	Rikard Althin	Essays on the Measurement of Producer Performance, 1995
61.	Lars Nordén	Empirical Studies of the Market Microstructure on the Swedish Stock Exchange, 1996
62.	Kristian Bolin	An Economic Analysis of Marriage and Divorce, 1996
63.	Fredrik Sjöholm	R&D, International Spillovers and Productivity Growth, 1997
64.	Hossein Asgharian	Essays on Capital Structure, 1997
65.	Hans Falck	Aid and Economic Performance - The Case of Tanzania, 1997
66.	Bengt Liljas	The Demand for Health and the Contingent Valuation Method, 1997
67.	Lars Pålsson Syll	Utility Theory and Structural Analysis, 1997
68.	Richard Henricsson	Time Varying Parameters in Exchange Rate Models, 1997

69.	Peter Hördahl	Financial Volatility and Time-Varying Risk Premia, 1997
70.	Lars Nilsson	Essays on North-South Trade, 1997
71.	Fredrik Berggren	Essays on the Demand for Alcohol in Sweden - Review and Applied Demand Studies, 1998
72.	Henrik Braconier	Essays on R&D, Technology and Growth, 1998
73.	Jerker Lundbäck	Essays on Trade, Growth and Exchange Rates, 1998
74.	Dan Anderberg	Essays on Pensions and Information, 1998
75.	P. Göran T. Hägg	An Institutional Analysis of Insurance Regulation – The Case of Sweden, 1998
76.	Hans-Peter Bermin	Essays on Lookback and Barrier Options - A Malliavin Calculus Approach, 1998
77.	Kristian Nilsson	Essays on Exchange Rates, Exports and Growth in Developing Countries, 1998
78.	Peter Jochumzen	Essays on Econometric Theory, 1998
79.	Lars Behrenz	Essays on the Employment Service and Employers' Recruitment Behaviour, 1998
80.	Paul Nystedt	Economic Aspects of Ageing, 1998
81.	Rasha M. Torstensson	Empirical Studies in Trade, Integration and Growth, 1999
82.	Mattias Ganslandt	Games and Markets – Essays on Communication, Coordination and Multi-Market Competition, 1999
83.	Carl-Johan Belfrage	Essays on Interest Groups and Trade Policy, 1999
84.	Dan-Olof Rooth	Refugee Immigrants in Sweden - Educational Investments and Labour Market Integration, 1999
85.	Karin Olofsdotter	Market Structure and Integration: Essays on Trade, Specialisation and Foreign Direct Investment, 1999
86.	Katarina Steen Carlsson	Equality of Access in Health Care, 1999
87.	Peter Martinsson	Stated preference methods and empirical analyses of equity in health, 2000
88.	Klas Bergenheim	Essays on Pharmaceutical R&D, 2000
89.	Hanna Norberg	Empirical Essays on Regional Specialization and Trade in Sweden, 2000
90.	Åsa Hansson	Limits of Tax Policy, 2000
91.	Hans Byström	Essays on Financial Markets, 2000
92.	Henrik Amilon	Essays on Financial Models, 2000
93.	Mattias Lundbäck	Asymmetric Information and The Production of Health, 2000

94. Jesper Hansson	Macroeconometric Studies of Private Consumption, Government Debt and Real Exchange Rates, 2001
95. Jonas Månsson	Essays on: Application of Cross Sectional Efficiency Analysis, 2001
96. Mattias Persson	Portfolio Selection and the Analysis of Risk and Time Diversification, 2001
97. Pontus Hansson	Economic Growth and Fiscal Policy, 2002
98. Joakim Gullstrand	Splitting and Measuring Intra-Industry Trade, 2002
99. Birger Nilsson	International Asset Pricing, Diversification and Links between National Stock Markets, 2002
100. Andreas Graflund	Financial Applications of Markov Chain Monte Carlo Methods, 2002
101. Therése Hindman Persson	Economic Analyses of Drinking Water and Sanitation in Developing Countries, 2002
102. Göran Hjelm	Macroeconomic Studies on Fiscal Policy and Real Exchange Rates, 2002
103. Klas Rikner	Sickness Insurance: Design and Behavior, 2002
104. Thomas Ericson	Essays on the Acquisition of Skills in Teams, 2002
105. Thomas Elger	Empirical Studies on the Demand for Monetary Services in the UK, 2002
106. Helena Johansson	International Competition, Productivity and Regional Spillovers, 2003
107. Fredrik Gallo	Explorations in the New Economic Geography, 2003
108. Susanna Thede	Essays on Endogenous Trade Policies, 2003
109. Fredrik CA Andersson	Interest Groups and Government Policy, A Political Economy Analysis, 2003
110. Petter Lundborg	Risky Health Behaviour among Adolescents, 2003
111. Martin W Johansson	Essays on Empirical Macroeconomics, 2003
112. Joakim Ekstrand	Currency Markets - Equilibrium and Expectations, 2003
113. Ingemar Bengtsson	Central bank power: a matter of coordination rather than money supply, 2003
114. Lars Pira	Staples, Institutions and Growth: Competitiveness of Guatemalan Exports 1524-1945, 2003
115. Andreas Bergh	Distributive Justice and the Welfare State, 2003
116. Staffan Waldo	Efficiency in Education - A Multilevel Analysis, 2003
117. Mikael Stenkula	Essays on Network Effects and Money, 2004