5

The Quest for the Option Formula

This chapter describes the history of option pricing starting with Louis Bachelier's seminal work in 1900 and culminating some 70 years later in the ultimate solution by Black, Scholes and Merton. We will see that Bachelier's work was neglected for almost 50 years until the renaissance of interest in option pricing in the 1950s. During the next 20 years, there were several notable attempts to solve the problem and we will see that some of them came very close to the solution. We explain some of the technical terms more fully in the Appendix to this chapter.

LOUIS BACHELIER'S CONTRIBUTIONS[1]

To set the stage for our discussion of Bachelier's contributions it is useful to summarise some ideas from the last chapter. We have noted that, to value a derivative, you need assumptions about how the underlying asset price moves in the future as well as a method for converting the future value back to the current time. The replicating portfolio, which is maintained by dynamic hedging, has the same payoff as the derivative at maturity. The no-arbitrage argument shows that the current price of the derivative is equal to the current value of this portfolio. We can use this argument to form a portfolio of the underlying asset and the derivative that replicates the riskless asset. A portfolio of a long position in a stock and a short position in the option can therefore be constructed so that it is risk-free. Note that this portfolio must earn the risk-free rate to preclude arbitrage.

We also saw in our discussion of the discrete time model that, to construct the replicating portfolio, we need to know the distribution of future stock prices. So an important ingredient of an option formula is an assumption about how stock prices move over time.

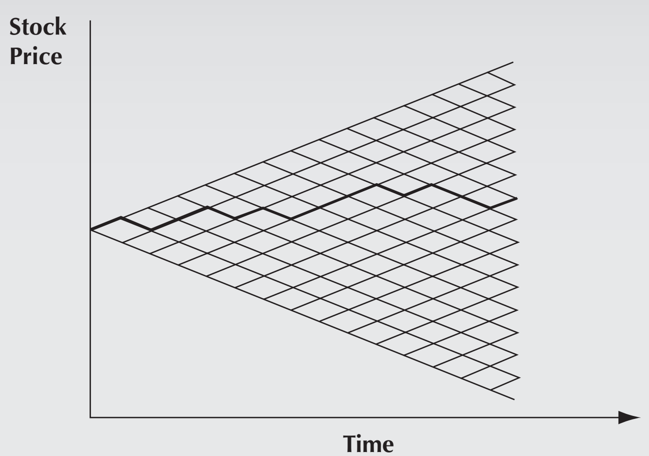
Bachelier made a number of important contributions to the modelling of stock prices and the mathematics of uncertainty in his brilliant thesis. Specifically, he modelled stock price movements in discrete time as a random walk. (We discussed random walks in Chapter 4.) We saw that we can generate a random walk by tossing a coin at each step. The coin then determines the direction of the stock price movement. If the coin comes up heads the stock price rises and if the coin comes up tails the price falls.

We saw in Chapter 4 that the stock price path will trace out a random walk if the price moves are generated by a sequence of coin tosses. The bold line in Figure 5.1 shows the stock price path for one particular sequence. When we make the number of tosses very large and the individual stock price jumps very small, this path becomes extremely erratic. As the time steps become smaller, the path traced out eventually corresponds to a particular type of movement called Brownian motion. Brownian motion is named after the English botanist Robert Brown, who studied the movement of pollen grains suspended in water. Bachelier showed that the random walk could be used to generate Brownian motion. This was five years before Einstein used Brownian motion to study the movements of dust particles suspended in water.

Bachelier's model of stock price movements essentially assumed that stock prices follow the so-called *normal distribution*.[2] This is the well-known bell curve shown in Figure 5.2. The model gives realistic movements of stock prices over a very short time period but it is not a realistic model of stock prices over a long time because a stock price that follows a normal distribution could become negative. In reality, a stock can end up worth nothing, but its price can never become negative. This is a consequence of the limited liability provision. Hence the lowest possible value of a common stock is zero.

We can adjust the Bachelier model to overcome this drawback by using *the rate of return* on the stock, rather than the actual stock prices used by Bachelier. If the stock price at the start of the year is US$100 and it is

Figure 5.1 A sample stock price path



US$120 at the end of the year then the rate of return is 20%. If the stock price at the start of the year is US$100 and it is US$80 at the end of the year, then the rate of return is –20%.

If the log of 1 plus the return on the stock follows a normal distribution we say the stock price follows a *lognormal model*. We provide an example of a lognormal distribution in Figure 5.3. The height of the graph indicates how likely a future event is. We see that in this case the most likely return is around 15%. The advantage of the lognormal model over the normal model in this context is that the lognormal model does not give rise to negative stock prices. Paul Samuelson, who also made important contributions to the development of the option formula, used the lognormal model to represent the stock price distribution.

Bachelier tested his theoretical model using actual options traded on the Paris Stock Exchange. As these options had short maturities, ranging from one day to a maximum of 45 days, Bachelier's use of the normal distribution gave reasonable results. In his thesis, Bachelier developed a very simple formula to price options where the asset price was equal to the strike price. In Bachelier's terminology such options were known as *simple options*. Bachelier stated that "The value of a simple option must be proportional to the square root of time."[3] This formula is still valid as an approximation for the prices of many short-term options.[4]

It is clear from his thesis that Bachelier used his practical experience in

Figure 5.2 Normal distribution for stock price

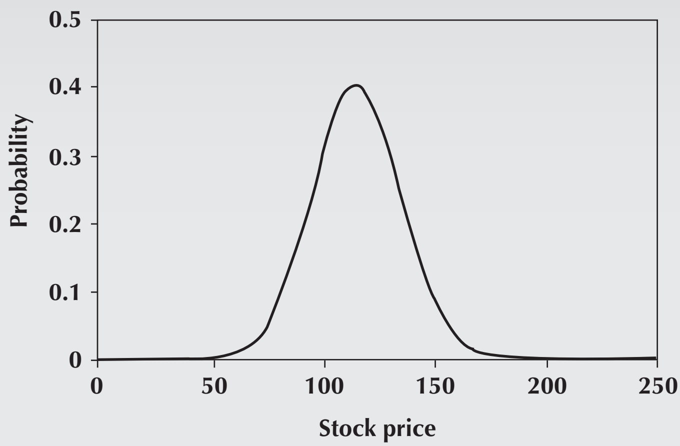
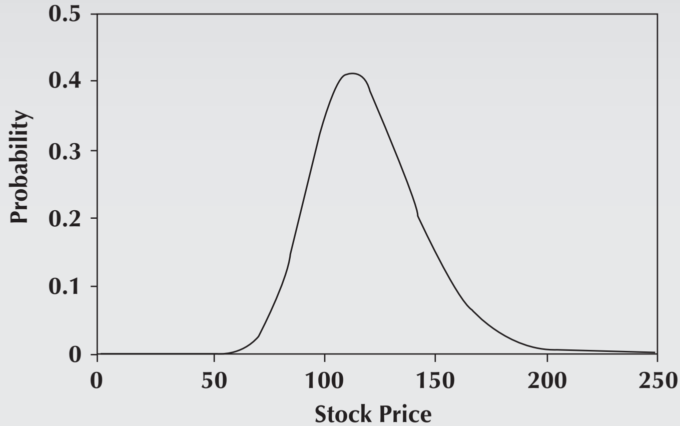


Figure 5.3 Lognormal distribution of stock price



formulating his ideas. For example, he contemplated possible strategies that would generate an arbitrage profit:

> Several years ago I noticed that it was possible, while admitting the above fact, to imagine operations in which one of the traders would make a profit regardless of eventual prices.

After explaining how this strategy could be implemented he noted that there would be no free lunch. "We will see that such spreads are never found in practice."

Bachelier's important scientific contributions to option pricing and probability were not recognised during his lifetime. Indeed, until recently little was known about the father of the modern option pricing theory. Paul Samuelson immediately recognised the importance of the work when he stumbled upon Bachelier's thesis.[5]

> In the early 1950s I was able to locate by chance this unknown book, rotting in the library of the University of Paris, and when I opened it up it was as if a whole new world was laid out before me. In fact as I was reading it, I arranged to get a translation in English, because I really wanted every precious pearl to be understood.

Bachelier's work became more widely known with the publication of Paul Cootner's 1964 book, which contained an English translation of the entire thesis.

ARBITARGE, WARRANTS AND BEN GRAHAM

The publication of Bachelier's thesis was followed by a 50-year lull in the development of a scientific model to price options. However, during this time important changes were taking place in the world of practical finance. Financial markets became more important in economic life and nowhere more so than in the United States, where the stock market had suffered the Great Depression and was poised to take off.

In 1915, a 21 year-old analyst with the Wall Street firm of Newberger, Henderson & Loeb spotted an arbitrage opportunity.[6] The Guggenheim Exploration Company, which held large positions in the shares of several copper-mining companies, was planning to dissolve and distribute its holding to its shareholders. These mining shares were traded on the New York Stock Exchange. The young analyst noted that the price of the copper shares to be received in the package was higher than the price of Guggenheim's stock. He recommended that Newberger buy Guggenheim stock and sell short the copper shares in the package. This strategy was adopted and proved highly profitable for the firm and several of its clients. Of course, as the market became more efficient such arbitrage opportunities became harder to find. The young analyst was Benjamin Graham who was later to co-author a classic book on security analysis.[7] He founded the school of value investing and became the teacher and mentor of Warren Buffet.

Graham became a partner in the firm in 1920 and ran what we would now call the firm's proprietary trading desk.[8] Graham believed that convertible bonds were good value in relative terms. A convertible bond is a package consisting of debt and an embedded call option to buy a firm's common stock. Graham would buy convertible bonds and sell short call options to hedge the price risk. If this portfolio is suitably constructed the impact of the stock price movements on the convertible bond and the short call option will offset each other. Hence this strategy can eliminate the price risk stemming from the stock price movements. Graham also used short positions in the stock as well as put options to hedge the risk. The idea of combining different securities in a portfolio to hedge out the risk of the underlying asset, lies at the very heart of modern option pricing.

The call option embedded in a convertible bond is a type of warrant. A warrant is basically an option issued by a company on its own stock. In some ways they are like call options. The main differences stem from the fact that warrants are issued by the company itself and tend to be long-term contracts. In the 1920s, warrants became popular in the US but they fell into disrepute in the 1930s when they became associated with market manipulation. They were shunned by the New York Stock Exchange until 1965 when AT&T issued warrants.

We will see that many of the researchers who contributed to the development of option pricing in the 1960s picked up their interest in this topic from their own market experience. Paul Samuelson, who was destined to play a key role in this development traces his interest in warrants to an investment newsletter to which he subscribed in about 1950.[9] Sheen Kassouf was already investing in the market when he went back to Columbia in 1962 to study for a PhD about valuing warrants. Ed Thorp was also preparing to invest in warrants before he teamed up with Kassouf in 1965.

PAUL SAMUELSON

Paul Samuelson, made a number of fundamental contributions to the pricing of warrants. He was already very interested in these securities when he discovered Bachelier's thesis. It will be recalled that Bachelier had assumed that asset-price movements follow a normal distribution and that Samuelson assumed that the returns on the underlying stock followed a *lognormal distribution*. He was able to derive a formula for the warrant price. This formula contained a number of variables. Two of these variables were the expected return on the stock and the expected return on the warrant (We discuss the expected return in the Appendix to this Chapter). These were unknown variables and it was difficult to estimate them. If Samuelson had found a way to obtaining values for these two variables, he would have solved the option-pricing problem.

The key to determining the values of the two unknowns was the no-arbitrage concept. Samuelson had the key in his hands because he had already used the no-arbitrage idea.[10]

SPRENKLE, BONESS AND COOTNER

In 1958, Case Sprenkle, an economics graduate student at Yale was searching for a thesis topic. He attended a seminar given by Paul Samuelson on the subject of option pricing and this gave Sprenkle the idea for his own dissertation.

Harry Markowitz had developed his model of portfolio selection in 1952. It provided a precise method for investors to select an optimal portfolio of stocks. In order to implement Markowitz's model the investor needed estimates of the expected returns on the different stocks and a measure of the risk of each stock. This measure of risk is known as the *variance* and we describe it in the Appendix.

Sprenkle's idea was to use the options and warrants market to infer investors' expectations about the returns and variance of common stocks. He developed a pricing formula that contained these parameters and used statistical techniques to back out the market's estimates. The extraction of expectations from derivatives prices was quite perceptive and this approach emerged as a powerful tool once these markets were more fully developed.

Case Sprenkle's formula for the price of a warrant also contained two unknown parameters.[11] One of these was the expected return on the stock and the other was a discount factor related to the risk of the stock.

Meanwhile, James Boness was also working on deriving a formula for stock options in his thesis at the University of Chicago under the chairmanship of Lawrence Fisher. Boness, too, assumed that the distribution of stock prices was lognormal. His solution to the expected rate of return question was to assume that investors discounted the expected proceeds from the option at the expected rate of return on the stock. His final formula is tantalisingly close to the Black–Scholes–Merton (BSM) formula, except that it contains the expected return on the stock where the BSM formula contains the riskless rate.

Boness also performed another useful service by translating Bachelier's thesis into English. This translation of Bachelier's thesis formed the centrepiece of an influential collection of research papers that was edited by Paul Cootner. Cootner's book, published in 1964, was entitled *The Random Character of Stock Market Prices*. It brought the major papers on stock price movements together in one volume as well as the key papers on option pricing and warrant pricing. This volume became essential reading for every serious scholar in the field. Cootner wrote superb introductions to each of the four sections of the book. His first paragraph conveys the intended scope of the work:

> Wherever there are valuable commodities to be traded, there are incentives to develop markets to organize that trade more efficiently. In modern complex societies the securities markets are usually among the best organized and virtually always the largest in terms of sales. The prices of such securities are typically very sensitive, responsive to all events both real and imagined that cast light on the murky future. The subject of this book is the attempts by skilled statisticians and economic theorists to probe into this process of price formation.

THORP AND KASSOUF

In 1965, two young professors met at the University of California's newly established campus at Irvine. Sheen Kassouf was an economist and Edward Thorp was a mathematician. They soon discovered their common interest in warrant pricing. Kassouf analysed market data to detemine the key variables that influence warrant prices. Based on this analysis he developed an empirical formula that explained warrant price in terms of these variables. Kassouf collaborated with Thorp to write a book called *Beat the Market*. It discussed the hedging of warrants using the underlying stock and developed a formula for the ratio of shares of stock to options needed to create a hedged position. This important idea was used by Black and Scholes in their celebrated 1973 paper.[12]

Thorp and Kassouf knew that the conventional approach of projecting the terminal payoff under the warrant and discounting back the positive part involved two troublesome parameters: the expected rate of return on the stock and the discount rate.

Thorp and the option formula

While Ed Thorp was thinking about these issues he was also trading warrants, which he believed were overvalued. His strategy was to buy the stock and sell the warrants short. As time passed the price of the stock changed and so too did the price of the warrant. Thorp noted that in these circumstances the portfolio could be adjusted by changing the investments in the two assets. He explored how this dynamic adjustment could be done in an actual market and noted the relationship between the stock price and the warrant price as circumstances changed. In a paper published in 1969, but written in 1968, it is clear that he understood dynamic hedging.[13]

By 1967, Thorp was aware of Cootner's book and the various warrant models that were based on taking the expected value of the payoff. Thorp had previously concluded that if he assumed a lognormal distribution for the asset, this produced a plausible formula for the warrant price. However his formula still contained the two bothersome parameters: the expected return on the stock, which Thorp called *m*, and the discount rate needed to convert the payoff at expiration back to current time, which he called *d*. As he experimented with the warrant formula, Thorp noticed that a simple way to eliminate the two parameters was to set both the expected return on the stock and the discount rate equal to the riskless rate. The resulting formula is, of course, the same as the Black–Scholes formula.

Thorp goes on to note that, not only does he not have a proof of the option formula but he does not even know if it is the right formula. At this point however, it provided the practical tool he needed. He describes his experiences using the formula:

> I can't prove the formula but I decide to go ahead and use it to invest, because there is in 1967–68 an abundance of vastly overpriced (in the sense of *Beat the Market*) OTC options. I use the formula to sell short the most extremely overpriced. I have limited capital and margin requirements are unfavorable so I short the options (typically at two to three times fair value) "naked," i.e. without hedging with the underlying stock. As it happens, small company stocks are up 84% in 1967 and 36% in 1968 (Ibbotson), so naked shorts of options are a disaster. Amazingly, I end up breaking even overall, on about $100,000 worth of about 20 different options sold short at various times from late '67 through '68. The formula has proven itself in action.

Was Ed Thorp the first person to discover and use the Black–Scholes formula? We find the evidence persuasive.[14] Thorp had both the background experience in hedging warrants and the mathematical ability to make such a discovery. He also had a strong incentive. When asked why he did not go public with this key result he replied that he was planning to set up a hedge fund and that this result would provide a competitive edge.[15] Thorp's work does not diminish in any way the contribution of Black, Merton and Scholes. They were the first to prove the result and they were the first to publish it. As Thorp himself notes:[16]

> BS was a watershed – it was only after seeing their proof that I was certain that this was the formula – and they justifiably get all the credit. They did two things that are required: They proved the formula (I didn't) and they published it (I didn't).

MEANWHILE BACK IN BOSTON

The story now moves back to MIT. By the late 1960s, Robert Merton was working with Paul Samuelson as his research assistant and graduate student. In 1969 they published a paper on warrant pricing that took a somewhat different approach.[17] They went back to the basic economic idea that in equilibrium the price adjusts so that supply is equal to demand. Samuelson and Merton were able to use this approach to obtain a relationship between the values of the warrant at successive time steps. Their approach, with some additional assumptions, provides another method of reaching the Black–Scholes formula. Samuelson and Merton came close to discovering the option formula. Some of the concepts they used have a very contemporary flavour and are now part of the toolkit of modern derivative pricing.[18] The equilibrium approach requires us to make more assumptions than the no-arbitrage approach but it is more versatile in that it can be applied to a broader class of problems.

FISCHER BLACK

The final steps in solving the option puzzle were made in papers by Fischer Black, Myron Scholes and Robert Merton.[19] The fascinating story of how they arrived at the formula has been told by Peter Bernstein in his book *Capital Ideas*.[20] We will sketch the details of their contributions and achievements.

In 1965, Fischer Black joined the consulting company Arthur D Little in Boston, where he met Jack Treynor who stimulated his interest in finance. Treynor was also a creative individual and made a contribution to the disovery of the so-called *capital asset pricing model* (CAPM). This model is derived from the basic idea that in any market the price is determined by the balancing of supply and demand. It tells us how the expected return on any common stock is related to the expected return on a portfolio that contains all the stocks in the market: the so-called market portfolio. We describe it more fully in the Appendix. Black became intrigued with the concept of equilibrium.

Black's fascination with this essentially economic concept was unusual in view of his background. Black had majored in physics as an undergraduate and his PhD was in applied mathematics at Harvard. He had never taken a formal course in economics or finance in his life.

Black was a remarkable individual. Jack Treynor has summarised Black's contributions as follows: "Fischer's research was about developing clever models – insightful, elegant models that changed the way we look at the world."[21]

Emanuel Derman, who was a colleague of Fischer Black's at Goldman Sachs & Co, has given an insightful account of Black's approach:[22]

> To me, Fischer's approach to modeling seemed to consist of unafraid hard thinking, intuition and no great reliance on advanced mathematics. This was inspiring. He attacked puzzles in a direct way, with whatever skills he had at his command, and often it worked.

Black became interested in the problem of pricing warrants after he teamed up with Treynor. He explored the relationship between the expected rate of return on the warrant and the expected rate of return on underlying stock. Over each short time period, Black assumed that these returns would conform to the CAPM, which was originally developed for common stocks.

Black was able to use the CAPM to derive an equation for the option's price. (We provide further details of Black's derivation in the Appendix to this Chapter.) The equation involved a relationship between the option price and its rate of change with respect to time as well as the asset price. Such equations are known as *differential equations* and have been used for a long time in physics and mathematics but until that point, had not been used much in finance. Black's final equation for the option price did not contain some of the variables he had started with. This meant that the eventual formula would not depend on these variables. The only risk term remaining was the total risk of the stock as measured by its volatility. Black was fascinated to note that the option price equation did not include the stock's expected return nor indeed any other asset's expected return.

BLACK AND SCHOLES

At this point Black had made significant progress. The solution to the differential equation would be the option price. He tried to produce a solution but noted that he was not familiar with the standard solution methods.[23] Black put the problem aside but started working on it again in 1969 with Myron Scholes, who was also interested in the warrant problem. Scholes had obtained a PhD in finance from the University of Chicago, where his mentor had been Merton Miller. In 1968, Scholes had joined MIT as an assistant professor of finance. Together Black and Scholes would solve the problem and produce the most acclaimed formula in finance.

Black has described the thinking that guided them toward the solution. His equation indicated that the option formula depended on the stock's volatility and not its expected return. The implication was that the formula could be derived using any expected return. So, they pretended that the expected return on the stock was equal to the riskless rate.[24] They also assumed, as had the other researchers in the 1960s, that the stock's returns were lognormal. This meant they could compute the expected value of the option at maturity. However this was not the option's current price – only its expected terminal value.

Black and Scholes then had an important insight. They could treat the option's expected return in the same way as they had dealt with the stock's expected return. They could assume, for valuation purposes, that it too had an expected return equal to the riskless rate. So, they could convert the expected final value of the option to its current value by discounting it at the riskless rate and when they did so they discovered the option formula. They confirmed that the formula satisfied the differential equation that Black had derived earlier. The task started by Bachelier was now complete.

HARNESSING THE POWER OF ITÔ CALCULUS

While Black and Scholes were working on their formula, they had several discussions with Robert Merton who was also working on option valuation. One of Merton's important contributions to finance was to introduce rigorous mathematical tools to deal properly with the modelling of uncertainty in continuous time. This framework was known as *stochastic calculus* and was developed by mathematicians. The most important contribution to this development was made by a Japanese mathematician, Kiyoshi Itô who gave a precise mathematical framework for modelling the evolution of uncertainty over time.[25]

Itô's work provided a rigorous mathematical foundation for the ideas of Bachelier and provided Merton with the perfect instrument for the analysis of stock price movements in continuous time. Merton also used this approach to model how individuals select investments over time. He extended the static one-period models to the much more sophisticated continuous time models.

Merton showed how the Black–Scholes model could be derived without the use of the CAPM. Merton's approach corresponds to setting up a portfolio of the stock and the option and dynamically adjusting this portfolio over time. Thanks to his use of the Itô calculus, Merton was able to do this in continuous time. By adjusting the portfolio at every instant, all the random fluctuations can be hedged away. From the no-arbitrage principle, this portfolio must earn the riskless interest rate. However, it would seem that this approach has a Catch 22 feature: to work out the correct amount of the option to hold in the portfolio we need to know how the option

PANEL 1

ITÔ USES MUSIC TO DESCRIBE HIS WORK

In precisely built mathematical structures, mathematicians find the same sort of beauty others find in enchanting pieces of music, or in magnificent architecture. There is, however, one great difference between the beauty of mathematical structures and that of great art. Music by Mozart, for instance, impresses greatly even those who do not know musical theory; the cathedral in Cologne overwhelms spectators even if they know nothing about Christianity. The beauty in mathematical structures, however, cannot be appreciated without understanding of a group of numerical formulae that express laws of logic. Only mathematicians can read “musical scores” containing many numerical formulae, and play that “music” in their hearts. Accordingly, I once believed that without numerical formulae, I could never communicate the sweet melody played in my heart. Stochastic differential equations, called “Itô Formula”, are currently in wide use for describing phenomena of random fluctuations over time. When I first set forth stochastic differential equations, however, my paper did not attract attention. It was over ten years after my paper that other mathematicians began reading my “musical scores” and playing my “music” with their “instruments”. By developing my “original musical scores” into more elaborate “music”, these researchers have contributed greatly to developing “Itô Formula”. In recent years, I find that my “music” is played in various fields, in addition to mathematics. Never did I expect that my “music” would be found in such various fields, its echo benefiting the practical world, as well as adding abstract beauty to the field of mathematics. On this opportunity of the Kyoto Prize lectures, I would like to express my sincerest gratitude and render homage to my senior researchers, who repeatedly encouraged me, hearing subtle sounds in my “Unfinished Symphony”.

(Extract from Lecture by Professor K. Itô (1998) on occasion of being awarded the Kyoto Prize, the most prestiguous scientific award in Japan.)

changes as the stock price changes. But it is the option price that we have to find in terms of the things that affect it.

There is a way out. When setting up the hedge it is enough to assume that the option depends on the current stock price and time to maturity. Both of these variables will change as time passes. It turns out we can derive an equation for the option price. The equation involves the option price directly and also includes other terms that depend on the option price. For example, one of these terms shows how the option price changes as the stock price changes. If the stock price moves by a dollar, this term shows how much the option price moves in response. The details of the derivation are beyond the scope of this book but when the dust settles we have an equation for the option price, which is exactly the same differential equation that Black had derived and that Black and Scholes had solved. The hedging argument, developed by Merton, led to the same equation and to the same formula for the option price. When Black and Scholes published their paper in 1973, they first derived the formula using Merton's approach.[26] Their original approach, based on the CAPM, was also included.

Merton also published a remarkable paper in 1973, which included a number of important extensions of the Black–Scholes model. Merton constructed a rigorous and general theory of option pricing based on the foundation of no arbitrage and the Itô calculus.[27] He showed just how far the no-arbitrage assumption can go, in deriving relations among different securities. Merton derived the Black–Scholes model under more general conditions than Black and Scholes originally specified. An option's value depends on the dividends payable on the asset over the option's future lifetime. Merton showed how to handle this in the valuation. He also predicted when a rational investor should exercise an American option and showed how to value American options (see Chapter 2).

In 1973, therefore, with the publication of these two seminal papers, the classic option valuation problem – a problem that had baffled some of the greatest minds in the finance profession – was solved. It had taken a long time and the efforts of many bright minds. The biggest obstacle was how to handle the expected rates of growth and the rate to be used to discount. The solution was simple: blindingly simple, but also deceptively simple. The option could be valued as if all the assets earned the riskless rate. The expected return on the stock and the expected return on the option did not appear at all.

Since the publication of these two papers, there has been an explosive growth in derivatives and this growth is related to the intellectual advances in the field, inspired by the BSM solution. New pricing and hedging technologies fuel this growth and, at the same time, the quest for the solution of practical problems has inspired new ideas. As Merton himself noted:[28]

> While reaffirming old insights, the continuous time model also provides new ones. Perhaps no better example is the seminal contribution of Black and Scholes that, virtually on the day it was published, brought the field to closure on the subject of option and corporate liability pricing. As the Black–Scholes work was closing gates on fundamental research in these areas it was simultaneously opening new gates: in applied and empirical study and setting the foundation for a new branch of finance called contingent claims analysis.

The BSM formula is so important that it has been included as part of the Appendix to this chapter.

APPENDIX

Statistical concepts: expected value and variance

This Appendix contains four sections dealing in more detail with some technical topics that have been mentioned in this chapter. The first section deals with basic statistical concepts. The second describes the capital asset pricing model. The third summarises Black's derivation of the equation for the option price and the last section gives the actual formula itself.

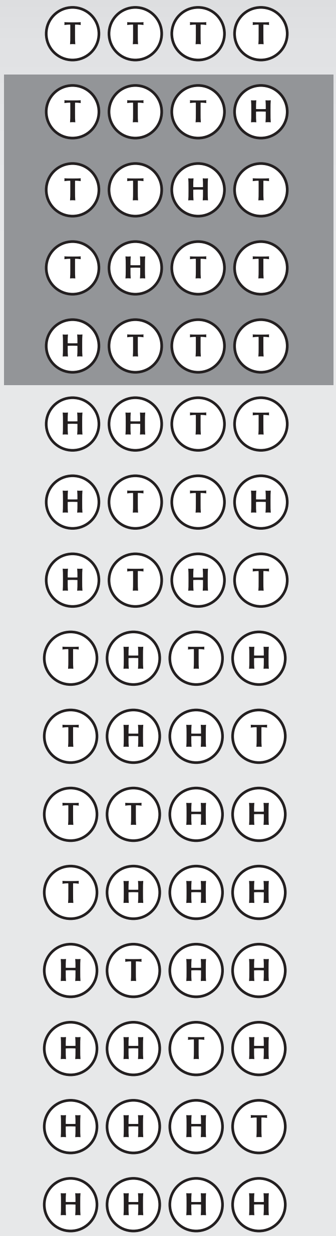
We will explain the concepts of expected value and variance using a simple example. The expected value corresponds to the familiar notion of average and the variance is a measure of the uncertainty associated with a set of uncertain outcomes. The set of uncertain outcomes is said to form a distribution, eg, the prices of a certain stock one year from now or the number of heads in four tosses of a coin. In the coin example, the number of heads could be zero, one, two, three or four and we will now use this example as an illustration.

Suppose we are interested in how likely it is that we will observe a given number of heads in the four tosses. For example, suppose we wish to know how likely it is that we will observe exactly one head in the four tosses. A natural approach might be to take a coin and perform a large number of experiments, in which each experiment consists of a sequence of four tosses. If we count the number of experiments that produced exactly one head and compare this to the total number of experiments, the ratio of these two numbers will give us an estimate of how likely it is to obtain one head. There are two difficulties with this approach: first, we might get tired of repeating such a boring task; second, it is not clear how many experiments we should conduct. Should we run the experiment 100 times or a 1,000 times? However, if we used a computer to simulate this experiment then we could run it a million times and this would give us a very good estimate of how often we could expect to find exactly one head.

The number of experiments that produce exactly one head divided by the total number of experiments provides the relative frequency of this event. In this example, the relative frequency of the event in question will approach 25% as the total number of experiments becomes large. If we are now asked to predict how likely it is that, in a single experiment of four tosses we will get exactly one head, we can assume that there will be a 25% chance of this happening. We can rephrase this as follows: there is a 25% probability of obtaining exactly one head in a series of four tosses of a fair coin.

In this case, we can work out why the probability is 25%. The first toss has two possible outcomes: either heads (H) or tails (T). The second also has two outcomes: either (H) or (T). In same way, the third has two

Figure A1 16 possible outcomes for a series of four tosses of a coin



outcomes and so does the fourth. Since each toss can have two outcomes, there are in total, 16 different possible patterns. We record the distinct outcomes in Figure A1 below, where the number of different patterns with exactly one head are shown in bold. The total possible number is 16 and the ratio of four to 16 is 25%. Hence the probability of obtaining one head in four tosses is 25%.

We apply the same logic to find the probability of obtaining zero heads, two heads, three heads and four heads. If a series of four tosses produce no heads, then all four must show tails and there is just one way that this can happen. Therefore, the probability of obtaining no heads in a series of four tosses is one divided by 16 or 6.25%. In the same way we can find that the probability of two heads is the ratio of six to 16 (37.5%), the probability of three heads is four over 16 (25%) and the probability of having all four turn up heads is one over 16 (6.25%). Note that the sum of all the probabilities is exactly one (100%).

Armed with these probabilities we can compute the expected number of heads. This is the average number we would obtain if we conducted a large number of experiments, where each experiment consists of four tosses. The expected number of heads, for our example, is given by:

$$0.0625(0) + 0.25(1) + 0.375(2) + 0.25(3) + 0.0625(4) = 2$$

Hence, the expected number of heads is two. This number is also known as the *expected value* of the distribution or the *mean* of the distribution.

Table A1 Details of steps in computing the variance of the distribution of heads

Number of heads	Probability of this number of heads	Difference between number of heads and expected value	Difference squared	Product of column 4 and probability
0	0.0625	–2	4	0.25
1	0.25	–1	1	0.25
2	0.375	0	0	0
3	0.25	1	1	0.25
4	0.0625	2	4	0.25

The variance measures the dispersion of a distribution around its expected value. To obtain the variance, we first find the difference between each outcome and its expected value; we square these differences and take the expected value of these squared differences. As an example, we now compute the variance of the distribution of the number of heads in four tosses. The main steps in computing the variance are outlined in Table A1. The first column gives the number of heads and the second gives the probability of obtaining this number of heads. The third column shows the difference between the number of heads and the expected number of heads, eg, if the number of heads is zero this difference is negative two. The fourth column shows the square of this difference. The last column is obtained by multiplying the second column by the fourth column and the final step is to add the numbers in the final column. In this case, the sum is 1, so the variance of this distribution is 1.

The variance corresponds to the expected value of the square of the distance from the mean. If the observations are all tightly bunched around the mean, then the variance will be smaller than if the observations are widely dispersed. The standard deviation is defined to be the square root of the variance. In this example, the standard deviation also happens to be 1 but this is because the variance is 1. If the variance is 9, the standard deviation would be 3. In the case of the normal or bell curve distribution, as shown in Figure 5.2 about 68% of the observations lie within one standard deviation of the mean.

In finance applications, we are often interested in the distribution of the rate of return of an asset and its associated standard deviation. In option applications, we often refer to the standard deviation of the return as the volatility of the return. The BSM formula depends, critically, on the volatility of the return on the underlying stock.

The capital asset pricing model

The *capital asset pricing model* (CAPM) tells us how the return on a stock relates to the return on the market as a whole. The market can be represented by a well-diversified portfolio of common stocks or a representative stock market index. Such a well-diversified portfolio is called the market portfolio. Suppose we pick a particular stock, eg, stock A. If the market goes up we expect that the price of stock A will also go up and likewise, if the market falls we expect the price of stock A to fall. The degree to which stock A moves will depend on how sensitive it is to movements in the market.

We need to introduce two concepts before we explain the capital asset pricing model. The first is the *beta* of a security, which represents the sensitivity of the security's return to the market return. If a stock has a high beta, its return is very sensitive to the market return and if it has a low beta its return is less sensitive to the market as a whole. The second concept is the *excess return* on a security, which is defined to be the return on the security over and above the riskless rate. Thus, the excess return on the riskless security itself is zero. If a stock earns 12% per annum and the riskless rate is 5%, the excess return on this stock is 7%. The *expected excess return* on a security is the average rate we expect to earn on a stock in excess of the riskless rate.

The CAPM states that the expected excess return on a security is equal to the beta of the security, multiplied by the expected excess return on the market. A numerical example may help at this stage. Suppose that stock A has a beta of 2, we expect the market to earn 15% and the riskless rate is 5%. In this case, the model predicts that stock A has an expected return of 25%. This is explained thus:

$$\textit{Expected return on stock A} = 0.05 + 2(0.15 - 0.05) = 0.25$$

However, if we have another stock, (B) that has a beta of 0.8, then the CAPM predicts it will have an average return of 13%. This is because:

$$\textit{Expected return on stock B} = 0.05 + 0.8(0.15 - 0.05) = 0.13$$

The CAPM expresses a very simple and powerful intuition. It demonstrates that there is a relation between return and risk and it shows that beta is the right measure of risk to use in this context.

The CAPM holds for combinations of securities in a very simple way. If an investor puts US$50,000 in Enron and US$50,000 in IBM, the beta of their portfolio is simply the average of the betas of these two stocks. By the same logic, the model holds for portfolios of securities including the market portfolio of all the securities. The beta of the market portfolio is therefore unity. We can go long and short securities to mix matters up so that we end up with a portfolio whose beta is zero. For example, if two stocks, C and D, each have the same beta, we can construct a portfolio with a zero beta by going long stock C and going short stock D. Fischer Black used the concept of a zero beta portfolio to derive the equation for the option price.

Summary of Black's approach

Black used the CAPM to formulate expressions for the expected return on the option and the expected return on the stock. The stock's expected return was equal to the riskless rate plus another term, proportional to the beta of the stock. Similarly, the option's expected return was equal to the riskless rate, as well as a risk term proportional to the beta of the option. This would seem to be a rather unpromising start as the ultimate option formula does not depend on these factors. The correct formula does *not* depend on expected returns nor does it depend on the betas. The returns on the option and the stock however, should always move in the same direction since owning the option is similar to a levered investment in the stock. Hence, they move in a synchronised fashion and their movements are also related to the market movements. There is a direct relation between the beta of the option and the beta of the stock that means we can value the option using only the beta of the stock.

The basic idea of Black's derivation lies in the construction of a portfolio that has zero beta. This portfolio consists of a short position in the call option and a long position in the right amount of the stock, to minimise the risk of the position. This was the strategy advocated by Thorp and Kassouf (1967) in *Beat the Market*. There are two criteria that could be used to select the right amount of stock to hold. One is to use the delta of the option, where delta corresponds to the change in the option induced by a dollar change in the stock. The other is to adjust the stock position so that the beta of the portfolio is zero. The result is that for short time periods, both criteria lead to the same stock position since the beta of the portfolio is zero, its *expected* return must also be equal to the riskless interest rate.

Black also knew that the key variables which affect the option's price as time passes are changes in the stock price and the passage of time itself. This insight, (together with the ideas of Professor Itô) can be used to derive a direct expression for the expected return on the option that is related to changes in the stock price and the passage of time. However, since the portfolio is a combination of the stock and the option we can also find the expected return on the portfolio if we know the expected returns on its component pieces. When we include these two pieces of information, the final expression for the expected return on the portfolio does not contain the stock's expected return. Now we can set the expected return on the portfolio equal to the riskless rate because the portfolio has a beta of zero. This gives us an equation that must be satisfied by the price of the option.

Black's analysis resulted in an equation for the option's price. The equation involved a relationship between the option price and its rate of change with respect to time as well as the asset price. His equation is an example of a *partial differential equation*. We have stressed that the final equation did not contain some of the variables he had started with because they had dropped out along the way. This meant that the eventual formula for the option price did not depend on these variables. The beta of the stock did not appear. The only risk term remaining was the total risk of the stock as measured by its volatility. Black was fascinated to note that the option price equation did not include the stock's expected return, nor indeed any other asset's expected return.

The Black–Scholes–Merton formula

The BSM formula gives the price of a standard call option European option in terms of five inputs. These inputs are:

❑ the current asset price;
❑ the option's strike price;
❑ the volatility of the asset's return;
❑ the time to option maturity; and
❑ the riskless rate.

It is conventional to use the following notation for these variables:

❑ S: current asset price;
❑ K: strike price;
❑ σ: volatility;
❑ T: time to maturity; and
❑ r: riskless rate.

It is more intuitive to present the formula in steps. First, we consider a zero coupon bond which matures on the same date as the option and has a maturity payment equal to the strike price. The current price of this bond, B is given by

$$B = Ke^{-rT}$$

The Black–Scholes formula can be written in terms of a long position in the stock and a short position in this bond as follows:

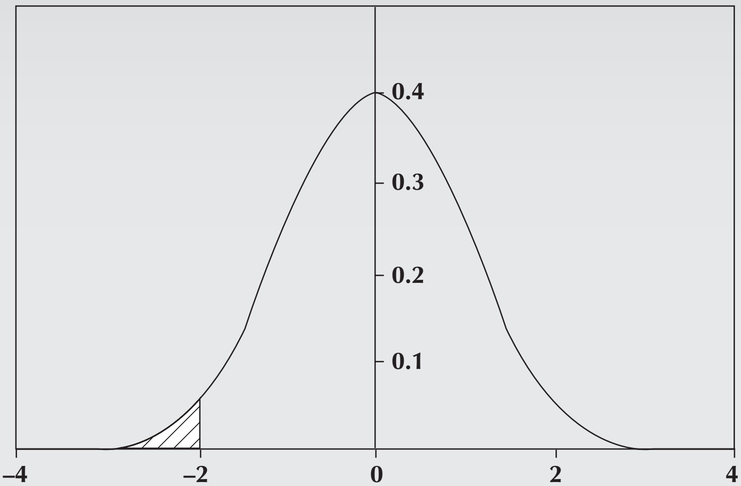
$$\Delta_1 S - \Delta_2 B$$

where Δ_1 and Δ_2 are functions of the five variables that determine the option's price. By writing the formula in this way, we see that it corresponds to a portfolio of Δ_1 units of the stock and Δ_2 units of the bond. Indeed, this shows how the current call price can be related to the replicating portfolio that we discussed in Chapter 4.

The BSM formula tells us what the values of Δ_1 and Δ_2 are in terms of the five input variables that determine the option's price. To do so we need to take a little detour to introduce a function that is obtained from the normal distribution.

Recall that we discussed the normal distribution in Chapter 5. If we have a normal distribution that has an expected value of zero and a standard deviation of 1, it is called the *standard normal distribution*. Figure A2 shows the shape of this distribution. The total area that lies underneath the standard normal curve in Figure A2 is equal to 1. Hence, if we pick any number on the horizontal axis, say at a distance d from the origin and draw a vertical through this point, the area to the left of the line under the curve will be a positive number less than 1. This area is denoted by the $N(d)$ and it has an important probabilistic interpretation.

Figure A2 Standard normal distribution



Here is the interpretation. Suppose I want to determine the probability that a number picked at random from the standard normal distribution is less than *d*. This probability is given by the quantity *N*(*d*), which as we saw is the area under the normal curve to the left of the line through *d*. A few numerical examples may help. If *d* is zero then since the normal curve is symmetrical around zero and the whole area under the curve is 1, the area to the left of zero must be one half, ie, N(0) equals one half. Most of the area under the standard normal curve is concentrated in the region of two standard deviations on either side of the mean (zero in this case). This means that most of the area lines in the region are bounded by the vertical lines through –2 and +2. So, the probability that a number picked at random from the standard normal distribution is less than 2 is quite high; in fact it is 0.9772. We can confirm this from standard tables of the function *N*(*d*), which gives *N*(2) = *0.9772*. In the same way, the probability that a number picked at random from a standard normal distribution is less than –2 is very small; in exact terms it is 0.0227 since *N*(–2) = *0.0228*. The shaded area in Figure A2 is therefore 0.0228.

Both Δ_1 and Δ_2 in the BSM formula can be written in terms of this function N. In fact

$$\Delta_1 = N(d_1)$$

where d_1 is equal to

$$\frac{\log \frac{S}{B} + \frac{1}{2} \sigma^2 T}{\sigma \sqrt{T}}$$

In the same way,

$$\Delta_2 = N(d_2)$$

where d_2 is equal to

$$\frac{\log \frac{S}{B} + \frac{1}{2} \sigma^2 T}{\sigma \sqrt{T}}$$

Putting things together the BSM formula for the standard European call option is:

$$N(d_1)\, S - N(d_2)\, B = S\, N(d_1) - B\, N(d_2)$$

This formula tells us not only what factors influence the call option price. We now give a numerical example. Suppose we have a European call option based on the inputs below.

Table A2 Inputs for a European call

Variable	Symbol	Numerical value
Stock price	S	100
Strike price	K	135
Stock return volatility	σ	20% per annum
Time to option maturity	T	Five years
Riskless rate	r	6% per annum

If we substitute these numerical values into the Black–Scholes formula we obtain the following values:

Table A3 Black–Scholes values

Symbol	Numerical value for this example
d_1	0.22337
d_2	–0.22384
$N(d_1)$	0.58838
$N(d_2)$	0.41144

In this case the discounted value of the strike price, *B* is equal to 100.01. Hence the BSM option price is:

$$100(0.58838) - 100.01(0.41144) = 17.69$$

In this case, the replicating portfolio consists of a long position of 58.838% shares of stock and short position of 41.144% units of the riskless bond. The formula can be used to show what happens to the option price if one of the inputs is changed. For example, if the volatility is 19% instead of 20% and the other four inputs remain the same, then the option price drops to 16.82.

1 See Bachelier (1900).
2 If a price follows Brownian motion in continuous time, then over any discrete time interval its distribution is normal. See Feller (1968).
3 Bachelier's thesis, p. 45 as reprinted in Cootner (1965); also reprinted by Risk Books (2000) with an introduction by Andrew Lo.
4 For example, if we consider a three-month call option with representative parameters the square root formula of Bachelier gives prices that are remarkably close to those obtained by the modern Black–Scholes–Merton formula. See Boyle and Ananthananarayanan (1979) for the derivation of the square root approximation to the Black–Scholes formula.
5 As reported in the transcript of the PBS television programme "NOVA 2704: The Trillion Dollar Bet". (Broadcast February 8, 2000.)
6 Editorial, 1968, *Financial Analysts Journal* (January/February), pp. 15-6.
7 Graham and Dodd: *Security Analysis*.
8 The proprietary trading desk trades the firm's own money as distinct from its customers' money.
9 See Bernstein (1992), p. 115.
10 Samuelson (1965) had considered the idea of using the no-arbitrage principle: "Mere arbitrage can take us no further than equation (19). The rest must be experience – the recorded facts of life".
11 There is a description of Sprenkle's formula in Black and Scholes (1973).
12 Black and Scholes (1973) note that, "One of the concepts we use in developing our model is expressed by Thorp and Kassouf".

13 See Thorp (1969).
14 This evidence is based on extensive correspondence with Ed Thorp in June and July 2000 and a study of his published papers and notes.
15 Telephone interview with Ed Thorp, 20 June 2000.
16 E-mail from Ed Thorp, 26 July 2000.
17 See Samuelson and Merton (1969).
18 Samuelson and Merton (1969) derived a formula where the warrant price is expressed in terms of its discounted expected value. The expectation is taken with respect to what the authors termed the utility probability density which in modern terms is the risk neutral measure.
19 See Black and Scholes (1973) and Merton (1973).
20 See Bernstein (1992).
21 See Treynor (1996).
22 See Derman (1996).
23 See Black (1989).
24 The word "pretend" is used advisedly. The expected return on the stock will be higher than the riskless rate. Indeed the capital asset pricing model states that this must be so. The use of the riskless rate here does not mean that any investor actually believes that the expected return on the stock will be this rate. It is a useful trick but one that puzzles even the best students.
25 See Itô (1951).
26 See Black and Scholes (1973).
27 See Merton (1973).
28 See Merton (1990).