is less than 0.5, it follows (see inequality (13.25)) that, when it is sufficiently deep in the money, the option should be exercised immediately before the second ex-dividend date.

We now use Black's approximation to value the option. The present value of the first dividend is

$$0.5e^{-0.1667\times0.09} = 0.4926$$

so that the value of the option, on the assumption that it expires just before the final ex-dividend date, can be calculated using the Black–Scholes formula with $S_0 = 40 - 0.4926 = 39.5074$, $K = 40$, $r = 0.09$, $\sigma = 0.30$, and $T = 0.4167$. It is \$3.52. Black's approximation involves taking the greater of this and the value of the option when it can only be exercised at the end of 6 months. From Example 13.8, we know that the latter is \$3.67. Black's approximation, therefore, gives the value of the American call as \$3.67.

The value of the option given by DerivaGem using "Binomial American" with 500 time steps is \$3.72. (Note that DerivaGem requires dividends to be input in chronological order in the table; the time to a dividend is in the first column and the amount of the dividend is in the second column.) There are two reasons for differences between the Binomial Model (BM) and Black's approximation (BA). The first concerns the timing of the early exercise decision; the second concerns the way volatility is applied. The timing of the early exercise decision tends to make BM greater than BA. In BA, the assumption is that the holder has to decide today whether the option will be exercised after 5 months or after 6 months; BM allows the decision on early exercise at the 5-month point to depend on the stock price at that time. The way in which volatility is applied tends to make BA greater than BM. In BA, when we assume exercise takes place after 5 months, the volatility is applied to the stock price less the present value of the first dividend; when we assume exercise takes place after 6 months, the volatility is applied to the stock price less the present value of both dividends.

SUMMARY

We started this chapter by examining the properties of the process for stock prices introduced in Chapter 12. The process implies that the price of a stock at some future time, given its price today, is lognormal. It also implies that the continuously compounded return from the stock in a period of time is normally distributed. Our uncertainty about future stock prices increases as we look further ahead. The standard deviation of the logarithm of the stock price is proportional to the square root of how far ahead we are looking.

To estimate the volatility σ of a stock price empirically, the stock price is observed at fixed intervals of time (e.g., every day, every week, or every month). For each time period, the natural logarithm of the ratio of the stock price at the end of the time period to the stock price at the beginning of the time period is calculated. The volatility is estimated as the standard deviation of these numbers divided by the square root of the length of the time period in years. Usually, days when the exchanges are closed are ignored in measuring time for the purposes of volatility calculations.

The differential equation for the price of any derivative dependent on a stock can be obtained by creating a riskless portfolio of the option and the stock. Because the derivative and the stock price both depend on the same underlying source of uncertainty, this can always be done. The portfolio that is created remains riskless for only a very short period of time. However, the return on a riskless porfolio must always be the risk-free interest rate if there are to be no arbitrage opportunities.

The expected return on the stock does not enter into the Black–Scholes differential equation. This leads to a useful result known as risk-neutral valuation. This result states that when valuing a derivative dependent on a stock price, we can assume that the world is risk neutral. This means that we can assume that the expected return from the stock is the risk-free interest rate, and then discount expected payoffs at the risk-free interest rate. The Black–Scholes equations for European call and put options can be derived by either solving their differential equation or by using risk-neutral valuation.

An implied volatility is the volatility that, when used in conjunction with the Black–Scholes option pricing formula, gives the market price of the option. Traders monitor implied volatilities. They often quote the implied volatility of an option rather than its price. They have developed procedures for using the volatilities implied by the prices of actively traded options to estimate volatilities for other options.

The Black–Scholes results can be extended to cover European call and put options on dividend-paying stocks. The procedure is to use the Black–Scholes formula with the stock price reduced by the present value of the dividends anticipated during the life of the option, and the volatility equal to the volatility of the stock price net of the present value of these dividends.

In theory, it can be optimal to exercise American call options immediately before any ex-dividend date. In practice, it is often only necessary to consider the final ex-dividend date. Fischer Black has suggested an approximation. This involves setting the American call option price equal to the greater of two European call option prices. The first European call option expires at the same time as the American call option; the second expires immediately prior to the final ex-dividend date.

FURTHER READING

On the Distribution of Stock Price Changes

Blattberg, R., and N. Gonedes, "A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices," *Journal of Business*, 47 (April 1974): 244–80.

Fama, E. F., "The Behavior of Stock Market Prices," *Journal of Business*, 38 (January 1965): 34–105.

Kon, S. J., "Models of Stock Returns—A Comparison," *Journal of Finance*, 39 (March 1984): 147–65.

Richardson, M., and T. Smith, "A Test for Multivariate Normality in Stock Returns," *Journal of Business*, 66 (1993): 295–321.

On the Black–Scholes Analysis

Black, F. "Fact and Fantasy in the Use of Options and Corporate Liabilities," *Financial Analysts Journal*, 31 (July/August 1975): 36–41, 61–72.

Black, F. "How We Came Up with the Option Pricing Formula," *Journal of Portfolio Management*, 15, 2 (1989): 4–8.

Black, F., and M. Scholes, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81 (May/June 1973): 637–59.

Merton, R. C., "Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science*, 4 (Spring 1973): 141–83.

On Risk-Neutral Valuation

Cox, J. C., and S. A. Ross, "The Valuation of Options for Alternative Stochastic Processes," *Journal of Financial Economics*, 3 (1976): 145–66.

Smith, C. W., "Option Pricing: A Review," *Journal of Financial Economics*, 3 (1976): 3–54.

On the Causes of Volatility

Fama, E. F. "The Behavior of Stock Market Prices." *Journal of Business*, 38 (January 1965): 34–105.

French, K. R. "Stock Returns and the Weekend Effect." *Journal of Financial Economics, 8 (March 1980): 55–69.*

French, K. R., and R. Roll "Stock Return Variances: The Arrival of Information and the Reaction of Traders." *Journal of Financial Economics*, 17 (September 1986): 5–26.

Roll R. "Orange Juice and Weather," *American Economic Review*, 74, 5 (December 1984): 861–80.

Questions and Problems (Answers in Solutions Manual)

13.1. What does the Black–Scholes stock option pricing model assume about the probability distribution of the stock price in one year? What does it assume about the continuously compounded rate of return on the stock during the year?

13.2. The volatility of a stock price is 30% per annum. What is the standard deviation of the percentage price change in one trading day?

13.3. Explain the principle of risk-neutral valuation.

13.4. Calculate the price of a 3-month European put option on a non-dividend-paying stock with a strike price of \$50 when the current stock price is \$50, the risk-free interest rate is 10% per annum, and the volatility is 30% per annum.

13.5. What difference does it make to your calculations in Problem 13.4 if a dividend of \$1.50 is expected in 2 months?

13.6. What is *implied volatility*? How can it be calculated?

13.7. A stock price is currently \$40. Assume that the expected return from the stock is 15% and that its volatility is 25%. What is the probability distribution for the rate of return (with continuous compounding) earned over a 2-year period?

13.8. A stock price follows geometric Brownian motion with an expected return of 16% and a volatility of 35%. The current price is \$38.

(a) What is the probability that a European call option on the stock with an exercise price of \$40 and a maturity date in 6 months will be exercised?

(b) What is the probability that a European put option on the stock with the same exercise price and maturity will be exercised?

13.9. Using the notation in this chapter, prove that a 95% confidence interval for S_T is between

$$S_0 e^{(\mu-\sigma^2/2)T-1.96\sigma\sqrt{T}} \quad \text{and} \quad S_0 e^{(\mu-\sigma^2/2)T+1.96\sigma\sqrt{T}}$$

13.10. A portfolio manager announces that the average of the returns realized in each year of the last 10 years is 20% per annum. In what respect is this statement misleading?

13.11. Assume that a non-dividend-paying stock has an expected return of μ and a volatility of σ. An innovative financial institution has just announced that it will trade a security that pays off a dollar amount equal to $\ln S_T$ at time T, where S_T denotes the value of the stock price at time T.

(a) Use risk-neutral valuation to calculate the price of the security at time t in terms of the stock price, S, at time t.

(b) Confirm that your price satisfies the differential equation (13.16).

13.12. Consider a derivative that pays off S_T^n at time T, where S_T is the stock price at that time. When the stock price follows geometric Brownian motion, it can be shown that its price at time t ($t \leqslant T$) has the form

$$h(t, T)S^n$$

where S is the stock price at time t and h is a function only of t and T.

(a) By substituting into the Black–Scholes–Merton partial differential equation, derive an ordinary differential equation satisfied by $h(t, T)$.

(b) What is the boundary condition for the differential equation for $h(t, T)$?

(c) Show that

$$h(t, T) = e^{[0.5\sigma^2 n(n-1)+r(n-1)](T-t)}$$

where r is the risk-free interest rate and σ is the stock price volatility.

13.13. What is the price of a European call option on a non-dividend-paying stock when the stock price is \$52, the strike price is \$50, the risk-free interest rate is 12% per annum, the volatility is 30% per annum, and the time to maturity is 3 months?

13.14. What is the price of a European put option on a non-dividend-paying stock when the stock price is \$69, the strike price is \$70, the risk-free interest rate is 5% per annum, the volatility is 35% per annum, and the time to maturity is 6 months?

13.15. Consider an American call option on a stock. The stock price is \$70, the time to maturity is 8 months, the risk-free rate of interest is 10% per annum, the exercise price is \$65, and the volatility is 32%. A dividend of \$1 is expected after 3 months and again after 6 months. Show that it can never be optimal to exercise the option on either of the two dividend dates. Use DerivaGem to calculate the price of the option.

13.16. A call option on a non-dividend-paying stock has a market price of \$$2\frac{1}{2}$. The stock price is \$15, the exercise price is \$13, the time to maturity is 3 months, and the risk-free interest rate is 5% per annum. What is the implied volatility?

13.17. With the notation used in this chapter:

(a) What is $N'(x)$?

(b) Show that $SN'(d_1) = Ke^{-r(T-t)}N'(d_2)$, where S is the stock price at time t and

$$d_1 = \frac{\ln(S/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}, \quad d_2 = \frac{\ln(S/K) + (r - \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}$$

(c) Calculate $\partial d_1/\partial S$ and $\partial d_2/\partial S$.

(d) Show that when

$$c = SN(d_1) - Ke^{-r(T-t)}N(d_2)$$

it follows that

$$\frac{\partial c}{\partial t} = -rKe^{-r(T-t)}N(d_2) - SN'(d_1)\frac{\sigma}{2\sqrt{T-t}}$$

where c is the price of a call option on a non-dividend-paying stock.

(e) Show that $\partial c/\partial S = N(d_1)$.
(f) Show that c satisfies the Black–Scholes differential equation.
(g) Show that c satisfies the boundary condition for a European call option, i.e., that $c = \max(S - K, 0)$ as $t \longrightarrow T$.

13.18. Show that the Black–Scholes formulas for call and put options satisfy put–call parity.

13.19. A stock price is currently \$50 and the risk-free interest rate is 5%. Use the DerivaGem software to translate the following table of European call options on the stock into a table of implied volatilities, assuming no dividends. Are the option prices consistent with the assumptions underlying Black–Scholes?

	Maturity (months)		
Strike price (\$)	*3*	*6*	*12*
45	7.0	8.3	10.5
50	3.7	5.2	7.5
55	1.6	2.9	5.1

13.20. Explain carefully why Black's approach to evaluating an American call option on a dividend-paying stock may give an approximate answer even when only one dividend is anticipated. Does the answer given by Black's approach understate or overstate the true option value? Explain your answer.

13.21. Consider an American call option on a stock. The stock price is \$50, the time to maturity is 15 months, the risk-free rate of interest is 8% per annum, the exercise price is \$55, and the volatility is 25%. Dividends of \$1.50 are expected in 4 months and 10 months. Show that it can never be optimal to exercise the option on either of the two dividend dates. Calculate the price of the option.

13.22. Show that the probability that a European call option will be exercised in a risk-neutral world is, with the notation introduced in this chapter, $N(d_2)$. What is an expression for the value of a derivative that pays off \$100 if the price of a stock at time T is greater than K?

13.23. Show that S^{-2r/σ^2} could be the price of a traded security.

13.24. A company has an issue of executive stock options outstanding. Should dilution be taken into account when the options are valued? Explain your answer.

13.25. A company's stock price is \$50 and 10 million shares are outstanding. The company is considering giving its employees 3 million at-the-money 5-year call options. Option exercises will be handled by issuing more shares. The stock price volatility is 25%, the 5-year risk-free rate is 5%, and the company does not pay dividends. Estimate the cost to the company of the employee stock option issue.

Assignment Questions

13.26. A stock price is currently \$50. Assume that the expected return from the stock is 18% and its volatility is 30%. What is the probability distribution for the stock price in 2 years? Calculate the mean and standard deviation of the distribution. Determine the 95% confidence interval.

13.27. Suppose that observations on a stock price (in dollars) at the end of each of 15 consecutive weeks are as follows:

30.2, 32.0, 31.1, 30.1, 30.2, 30.3, 30.6, 33.0, 32.9, 33.0, 33.5, 33.5, 33.7, 33.5, 33.2

Estimate the stock price volatility. What is the standard error of your estimate?

13.28. A financial institution plans to offer a security that pays off a dollar amount equal to S_T^2 at time T.

(a) Use risk-neutral valuation to calculate the price of the security at time t in terms of the stock price S at time t. (*Hint*: The expected value of S_T^2 can be calculated from the mean and variance of S_T given in Section 13.1.)

(b) Confirm that your price satisfies the differential equation (13.16).

13.29. Consider an option on a non-dividend-paying stock when the stock price is \$30, the exercise price is \$29, the risk-free interest rate is 5%, the volatility is 25% per annum, and the time to maturity is 4 months.

(a) What is the price of the option if it is a European call?

(b) What is the price of the option if it is an American call?

(c) What is the price of the option if it is a European put?

(d) Verify that put–call parity holds.

13.30. Assume that the stock in Problem 13.29 is due to go ex-dividend in $1\frac{1}{2}$ months. The expected dividend is 50 cents.

(a) What is the price of the option if it is a European call?

(b) What is the price of the option if it is a European put?

(c) If the option is an American call, are there any circumstances under which it will be exercised early?

13.31. Consider an American call option when the stock price is \$18, the exercise price is \$20, the time to maturity is 6 months, the volatility is 30% per annum, and the risk-free interest rate is 10% per annum. Two equal dividends are expected during the life of the option with ex-dividend dates at the end of 2 months and 5 months. Assume the dividends are 40 cents. Use Black's approximation and the DerivaGem software to value the option. How high can the dividends be without the American option being worth more than the corresponding European option?

APPENDIX

PROOF OF THE BLACK–SCHOLES–MERTON FORMULA

We will prove the Black–Scholes result by first proving another key result that will also be useful in future chapters.

Key Result

If V is lognormally distributed and the standard deviation of $\ln V$ is w, then

$$E[\max(V-K,\,0)] = E(V)N(d_1) - KN(d_2) \tag{13A.1}$$

where

$$d_1 = \frac{\ln[E(V)/K] + w^2/2}{w}$$

$$d_2 = \frac{\ln[E(V)/K] - w^2/2}{w}$$

and E denotes the expected value.

Proof of Key Result

Define $g(V)$ as the probability density function of V. It follows that

$$E[\max(V-K,\,0)] = \int_K^\infty (V-K)g(V)\,dV \tag{13A.2}$$

The variable $\ln V$ is normally distributed with standard deviation w. From the properties of the lognormal distribution, the mean of $\ln V$ is m, where[15]

$$m = \ln[E(V)] - w^2/2 \tag{13A.3}$$

Define a new variable

$$Q = \frac{\ln V - m}{w} \tag{13A.4}$$

This variable is normally distributed with a mean of zero and a standard deviation of 1.0. Denote the density function for Q by $h(Q)$ so that

$$h(Q) = \frac{1}{\sqrt{2\pi}}e^{-Q^2/2}$$

Using equation (13A.4) to convert the expression on the right-hand side of equation (13A.2) from an integral over V to an integral over Q, we get

$$E[\max(V-K,\,0)] = \int_{(\ln K - m)/w}^\infty (e^{Qw+m} - K)\,h(Q)\,dQ$$

or

$$E[\max(V-K,\,0)] = \int_{(\ln K - m)/w}^\infty e^{Qw+m}h(Q)dQ - K\int_{(\ln K - m)/w}^\infty h(Q)dQ \tag{13A.5}$$

[15] For a proof of this, see Technical Note 2 on the author's website.

Now

$$\begin{aligned} e^{Qw+m}h(Q) &= \frac{1}{\sqrt{2\pi}} e^{(-Q^2+2Qw+2m)/2} \\ &= \frac{1}{\sqrt{2\pi}} e^{[-(Q-w)^2+2m+w^2]/2} \\ &= \frac{e^{m+w^2/2}}{\sqrt{2\pi}} e^{[-(Q-w)^2]/2} \\ &= e^{m+w^2/2} h(Q-w) \end{aligned}$$

This means that equation (13A.5) becomes

$$E[\max(V-K,\,0)] = e^{m+w^2/2} \int_{(\ln K-m)/w}^{\infty} h(Q-w)\,dQ - K \int_{(\ln K-m)/w}^{\infty} h(Q)\,dQ \quad \textbf{(13A.6)}$$

If we define $N(x)$ as the probability that a variable with a mean of zero and a standard deviation of 1.0 is less than x, the first integral in equation (13A.6) is

$$1 - N[(\ln K - m)/w - w]$$

or

$$N[(-\ln K + m)/w + w]$$

Substituting for m from equation (13A.3) leads to

$$N\left(\frac{\ln[E(V)/K] + w^2/2}{w}\right) = N(d_1)$$

Similarly the second integral in equation (13A.6) is $N(d_2)$. Equation (13A.6), therefore, becomes

$$E[\max(V-K,\,0)] = e^{m+w^2/2} N(d_1) - KN(d_2)$$

Substituting for m from equation (13A.3) gives the key result.

The Black–Scholes–Merton Result

We now consider a call option on a non-dividend-paying stock maturing at time T. The strike price is K, the risk-free rate is r, the current stock price is S_0, and the volatility is σ. As shown in equation (13.22), the call price c is given by

$$c = e^{-rT} \hat{E}[\max(S_T - K,\,0)] \quad \textbf{(13A.7)}$$

where S_T is the stock price at time T and $\hat{E}$ denotes the expectation in a risk-neutral world. Under the stochastic process assumed by Black–Scholes, S_T is lognormal. Also, from equations (13.3) and (13.4), $\hat{E}(S_T) = S_0 e^{rT}$ and the standard deviation of $\ln S_T$ is $\sigma\sqrt{T}$.

From the key result just proved, equation (13A.7) implies

$$c = e^{-rT}[S_0 e^{rT} N(d_1) - KN(d_2)]$$

or

$$c = S_0 N(d_1) - K e^{-rT} N(d_2)$$

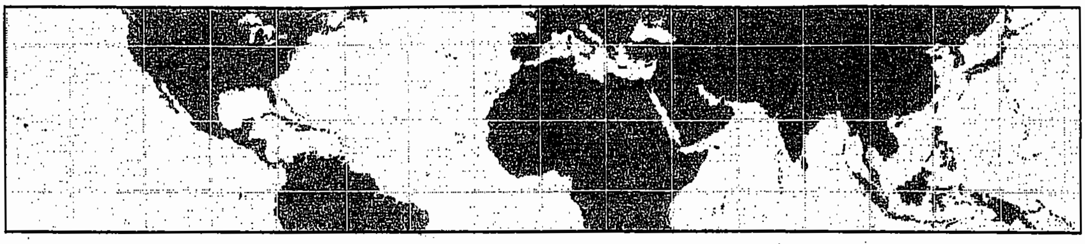
where

$$d_1 = \frac{\ln[\hat{E}(S_T)/K] + \sigma^2 T/2}{\sigma\sqrt{T}} = \frac{\ln(S_0/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}$$

and

$$d_2 = \frac{\ln[\hat{E}(S_T)/K] - \sigma^2 T/2}{\sigma\sqrt{T}} = \frac{\ln(S_0/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}}$$

This is the Black–Scholes–Merton result.



CHAPTER 14

Derivatives Markets in Developing Countries

Derivatives have become very important tools in the Western world for transferring risks from one entity to another. It is not surprising that derivatives markets are growing fast in many developing countries. This chapter focuses on what is happening in China and India. These are two countries whose economies are expected to play dominant roles in the 21st century. China's population in 2007 is estimated to be about 1.3 billion, while that of India is about 1.1 billion. (By contrast, the population of the United States is only about 0.3 billion.) The world's population is about 6.6 billion, so China and India between them account for about 36% of the world's population. India is expected to overtake China as the world's most populous nation by 2030.

There can be little doubt that China and India will have a huge impact on the development of derivatives markets throughout the world in the years to come. Other countries will also be important players. For example, Brazil, the fifth most populous country in the world, has been very successful in developing its derivatives markets. Its premier exchange, Bolsa de Mercadorias & Futuros (www.bmf.com.br) is highly regarded.

14.1 CHINA'S MARKETS

The way derivatives markets are regulated plays an important role in their growth. Too much regulation can stunt growth; not enough is liable to lead to a lack of confidence in the markets and make individuals and corporations less willing to trade.

In China, the China Securities Regulatory Commission (CSRC: www.csrc.gov.cn), the China Bank Regulatory Commission (CBRC: www.cbrc.gov.cn) and the People's Bank of China (PBC: www.pbc.gov.cn) are all involved in derivatives regulation. The CBRC issued new rules concerning derivatives in February 2004. This eased many of the restrictions that had previously been in place on the trading of derivatives. The new rules outlined the approval processes that would permit financial institutions to trade derivatives and the risk management procedures that the institutions would be required to observe. The derivatives that can be traded include forwards, futures, swaps, options, and structured products. The underlyings include foreign exchange and interest rates. The rules allowed financial institutions to enter into derivatives transactions for both hedging and profit-making purposes. This was an important change for the market's potential for growth. Under the previous regime, financial institutions were allowed to enter into derivatives transactions only for hedging purposes. As explained in Chapter 1, the main reason why derivatives markets have grown so fast in the Western world is that they can be used for three different purposes: hedging, speculation, and arbitrage.

Prior to 2004, financial institutions founds ways round the rules (usually with the help of foreign banks). However, there were some mishaps. For example, Guangdong International Trust and Investment Corporation went bankrupt in 1998. It had been a major player in derivatives markets and structured products, but the legal foundation for what it was doing was questionable. During the bankruptcy proceedings, the Chinese courts took the view that the financial institution did not have the authority to enter into its contracts and declared all the contracts null and void. The 2004 rules give much clearer guidelines than before on who can trade derivatives and what they can be used for. The rules allow foreign banks to deal directly with Chinese entities, rather than by going through a Chinese financial institution.

China's official currency is the yuan (also known as the renminbi, RMB, or CNY). The Chinese government has indicated that one of its goals is to make the yuan fully convertible. Convertibility means that the exchange rate between the yuan and other major currencies such as the euro and the US dollar will be determined by supply and demand in the same way that other exchange rates between the major currencies of the world are determined. At present the government restricts movements in the exchange rate between the yuan and the US dollar from day to day and many economists think that the yuan is currently undervalued against the US dollar, with the result that China's trade surplus is boosted.

There are a number of exchanges in China that trade futures contracts. The Shanghai Futures Exchange (SFE: www.shfe.com.cn) trades futures contracts on copper, aluminum, natural rubber, fuel oil, gold, and zinc. The Dalian Commodity Exchange (DCE: www.dce.com.cn) trades futures contracts on soybean, soybean meal, soybean oil, corn, linear low-density polyethylene, and palm olein (which is a product obtained from the fractionation of palm oil). In 2008 it became the world's second largest agricultural futures exchange (after the Chicago Board of Trade). The Zhengzhou Commodity Exchange (ZCE: www.zce.com.cn) trades futures contracts on wheat, cotton, sugar, pure terephthalic acid (a chemical used in the manufacture of clothing and plastic bottles), and rapeseed oil.

On September 8, 2006, several exchanges (including the three futures exchanges that have just been mentioned) founded the China Financial Futures Exchange (CFFEX: www.cffex.com.cn). This will launch a CSI 300 index futures, once regulatory approval is obtained. The CSI index tracks the daily price performance of the 300 stocks listed on the Shanghai Stock Exchange (SSE: www.sse.com.cn) or Shenzhen Stock Exchanges (SZSE: www.szse.cn). The index was launched in 2004 and was set equal to 1,000 on December 31, 2004. In China, as in many other countries, it has proved more difficult to obtain regulatory approval for the trading of index futures than for commodity futures. This is because they require cash settlement. Regulators are reluctant to approve contracts requiring cash settlement because they consider them closer to gambling than contracts that are settled by physical delivery of the underlying. (Arguably this is not rational because, as explained in Chapter 2, the vast majority of futures contracts are closed out before maturity and therefore effectively settled in cash.) The trading of options in China is nothing like as well developed as the trading of futures.

The contracts traded in China's over-the-counter market include forward and swap contracts on foreign exchange, interest rate swaps, and bond forwards. Convertible bonds have been popular. For example, China's second-largest telecommunications equipment maker, ZTE, announced in January 2008 that it was offering 4 billion yuan ($555 million) of convertible bonds with detachable warrants. The five-year bonds have a coupon of between 0.8 and 1.5 percent. The warrants are exercisable during a 10-day period two years after listing. Some products involving securitizations have traded. There have been asset-backed securities, mortgage-backed securities, and the occasional collateralized debt obligation.

14.2 INDIA'S MARKETS

Equity derivatives have a long history in India. Options of various kinds called teji (call options), mandi (put options), and fatak (straddles) traded in unorganized markets as early as 1900 in Mumbai. However, derivatives markets suffered a set back in 1956 when the Securities Control and Regulation Act (SCRA) banned what was considered to be undesirable speculation in securities and again in 1969 when "forward trading contracts" were banned.

The Indian currency is the rupee. In theory, this is freely convertible. In practice, the Reserve Bank of India (RBI: www.rbi.org.in) intervenes to control the exchange rate against the US dollar. (This is sometimes referred to as a dirty or managed float.) In addition, foreign nationals are forbidden from importing or exporting rupees and the extent to which Indian citizens can do this is limited. However, India is moving towards full convertibility.

Securities trading in India is regulated by the Securities and Exchange Board of India (SEBI: www.sebi.gov.in), which was set up in 1988. An important step toward the development of derivatives trading in India was the Securities Laws (Amendment) Ordinance of 1995, which lifted the prohibition on the trading of options. This led the SEBI in 1998 to set up committees to consider how derivatives should be introduced and how risks should be contained. This led to a legal framework, enacted in December 1999, within which derivatives such as options and futures are regarded as securities and can be traded. The trading of stock index futures started in June 2000, and later other products, such as stock index options, stock options, and single stock futures, were allowed.

Since 2000, derivatives trading in India has developed quite fast. Both the National Stock Exchange in Mumbai (NSE: www.nseindia.com) and the Stock Exchange, Mumbai (BSE: www.bseindia.com) trade index futures, stock futures, index options, and stock options. There are a large number of smaller exchanges trading futures and options. The exchanges use sophisticated technology and are fully electronic. The over-the-counter market is active in India, particularly in interest rate and swap products.

14.3 OTHER DEVELOPING COUNTRIES

Derivatives markets are progressing well in many other developing countries, such as Brazil, Korea, Malaysia, Mexico, Poland, and South Africa. Derivatives are sophisticated instruments and it is necessary for a country's financial markets to reach a reasonable level of sophistication before derivatives can be successful. There are a number of key conditions for the growth of derivatives. It is important for the government to set up a regulatory structure that protects investors from fraud and leaves them feeling sure that their contracts will be honored. However, the government should not impose too many restrictions on the way derivatives can be used because speculators and arbitrageurs are important for the liquidity of the market. There must be a sound financial and legal system within the country. Volatility is not bad for derivatives markets. (Indeed without volatility there would be little interest in many derivative products.) But derivatives are unlikely to thrive unless the economy of the country is reasonably stable and there is a good payments system. Stock markets, bond markets, and money markets should be reasonably well developed. (After all, stocks, bonds, and money market instruments are the underlyings for many derivatives.) Ideally the currency should be freely convertible and there should be no restrictions on the flow of the currency in and out of the country. Moreover, there should be enough swaps, bonds, and money market instruments trading for a risk-free zero-coupon yield curve to be estimated. A final very important condition is that there should be enough well educated individuals who understand the products and how they can be valued.

An intriguing idea for developing countries is the possibility of derivatives transactions between national governments. If country X exports oil to country Y and country Y exports building materials to company X, they are both subject to risks relating to the prices of their exports. It might make sense for them to enter into a swap that effectively fixes prices for several years into the future. This example can be extended so that it applies to groups of countries that trade with each other.[1]

SUMMARY

The economies of many developing countries are growing very fast. Indeed, the phrase "developing country" may no longer be an appropriate way of describing them in a decade or two. Although the trading of derivatives occasionally gets out of control (see Chapter 34), there can be no question that derivatives markets have played an important role in allowing risks to be managed in developed countries. There is no reason to suppose that the same will not in the future be true in every country of the world. Several developing countries have laid the foundations for a mature derivatives market by creating appropriate legal, financial, and regulatory frameworks. In a few decades, it is likely that the derivatives markets of countries like China, India, and Brazil will be as important, if not more important, than those of the United States and Western European countries.

---

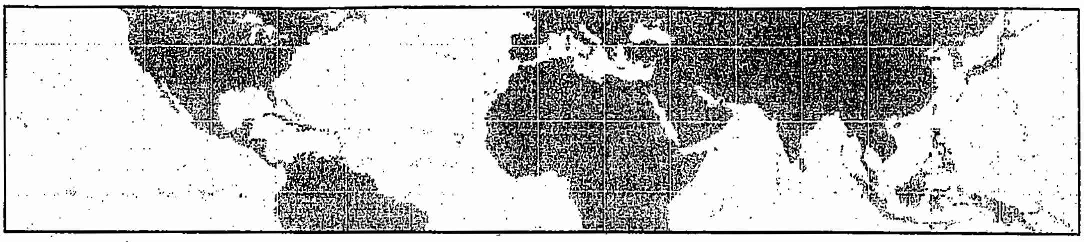
[1] This idea, which has been proposed by Robert Merton, is of course not just appropriate for developing countries.

FURTHER READING

Ahuja, N. L., "Commodity Derivatives Market in India: Development, Regulation, and Futures Prospects," *International Research Journal of Finance and Economics*, 2 (2006): 153–62.

Braga, B. S., "Derivatives Markets in Brazil: An Overview," *Journal of Derivatives*, 4 (Fall 1996): 63–78,

Ménager-Xu, M. Y., "China's Derivatives Markets," in: *China's Financial Markets: An Insider's Guide to How the Markets Work* (S. N. Neftci and M. Y. Ménager-Xu, eds.). New York: Academic Press, 2007.



CHAPTER 15

Options on Stock Indices and Currencies

Options on stock indices and currencies were introduced in Chapter 8. This chapter discusses them in more detail. It explains how they work and reviews some of the ways they can be used. In the second half of the chapter, the valuation results in Chapter 13 are extended to cover European options on a stock paying a known dividend yield. It is then argued that both stock indices and currencies are analogous to stocks paying dividend yields. This enables the results for options on a stock paying a dividend yield to be applied to these types of options as well.

15.1 OPTIONS ON STOCK INDICES

Several exchanges trade options on stock indices. Some of the indices track the movement of the market as a whole. Others are based on the performance of a particular sector (e.g., computer technology, oil and gas, transportation, or telecoms). Among the index options traded on the Chicago Board of Options Exchange are American and European options on the S&P 100 (OEX and XEO), European options on the S&P 500 (SPX), European options on the Dow Jones Industrial Average (DJX), and European options on the Nasdaq 100 (NDX). In Chapter 8, we explained that the CBOE trades LEAPS and flex options on individual stocks. It also offers these option products on indices.

One index option contract is on 100 times the index. (Note that the Dow Jones index used for index options is 0.01 times the usually quoted Dow Jones index.) Index options are settled in cash. This means that, on exercise of the option, the holder of a call option contract receives $(S - K) \times 100$ in cash and the writer of the option pays this amount in cash, where S is the value of the index at the close of trading on the day of the exercise and K is the strike price. Similarly, the holder of a put option contract receives $(K - S) \times 100$ in cash and the writer of the option pays this amount in cash.

Portfolio Insurance

Portfolio managers can use index options to limit their downside risk. Suppose that the value of an index today is S_0. Consider a manager in charge of a well-diversified portfolio whose beta is 1.0. A beta of 1.0 implies that the returns from the portfolio mirror those from the index. Assuming the dividend yield from the portfolio is the same as the dividend yield from the index, the percentage changes in the value of the portfolio can be expected to be approximately the same as the percentage changes in the value of the index. Each contract on the S&P 500 is on 100 times the index. It follows that the value of the portfolio is protected against the possibility of the index falling below K if, for each $100S_0$ dollars in the portfolio, the manager buys one put option contract with strike price K. Suppose that the manager's portfolio is worth \$500,000 and the value of the index is 1,000. The portfolio is worth 500 times the index. The manager can obtain insurance against the value of the portfolio dropping below \$450,000 in the next three months by buying five three-month put option contracts on the index with a strike price of 900.

To illustrate how the insurance works, consider the situation where the index drops to 880 in three months. The portfolio will be worth about \$440,000. The payoff from the options will be $5 \times (900 - 880) \times 100 = \$10,000$, bringing the total value of the portfolio up to the insured value of \$450,000.

When the Portfolio's Beta Is Not 1.0

If the portfolio's beta (β) is not 1.0, β put options must be purchased for each $100S_0$ dollars in the portfolio, where S_0 is the current value of the index. Suppose that the \$500,000 portfolio just considered has a beta of 2.0 instead of 1.0. We continue to assume that the S&P 500 index is 1,000. The number of put options required is

$$2.0 \times \frac{500,000}{1,000 \times 100} = 10$$

rather than 5 as before.

To calculate the appropriate strike price, the capital asset pricing model can be used. Suppose that the risk free rate is 12%, the dividend yield on both the index and the portfolio is 4%, and protection is required against the value of the portfolio dropping below \$450,000 in the next three months. Under the capital asset pricing model, the expected excess return of a portfolio over the risk-free rate is assumed to equal beta

Table 15.1 Calculation of expected value of portfolio when the index is 1,040 in three months and $\beta = 2.0$.

Value of index in three months:	1,040
Return from change in index:	40/1,000, or 4% per three months
Dividends from index:	$0.25 \times 4 = 1\%$ per three months
Total return from index:	$4 + 1 = 5\%$ per three months
Risk-free interest rate:	$0.25 \times 12 = 3\%$ per three months
Excess return from index over risk-free interest rate:	$5 - 3 = 2\%$ per three months
Expected excess return from portfolio over risk-free interest rate:	$2 \times 2 = 4\%$ per three months
Expected return from portfolio:	$3 + 4 = 7\%$ per three months
Dividends from portfolio:	$0.25 \times 4 = 1\%$ per three months
Expected increase in value of portfolio:	$7 - 1 = 6\%$ per three months
Expected value of portfolio:	$\$500,000 \times 1.06 = \$530,000$

Table 15.2 Relationship between value of index and value of portfolio for $\beta = 2.0$.

Value of index in three months	*Value of portfolio in three months ($)*
1,080	570,000
1,040	530,000
1,000	490,000
960	450,000
920	410,000
880	370,000

times the excess return of the index portfolio over the risk-free rate. The model enables the expected value of the portfolio to be calculated for different values of the index at the end of three months. Table 15.1 shows the calculations for the case where the index is 1,040. In this case the expected value of the portfolio at the end of the three months is $530,000. Similar calculations can be carried out for other values of the index at the end of the three months. The results are shown in Table 15.2. The strike price for the options that are purchased should be the index level corresponding to the protection level required on the portfolio. In this case the protection level is $450,000 and so the correct strike price for the 10 put option contracts that are purchased is 960.[1]

To illustrate how the insurance works, consider what happens if the value of the index falls to 880. As shown in Table 15.2, the value of the portfolio is then about $370,000. The put options pay off $(960 - 880) \times 10 \times 100 = \$80,000$, and this is exactly what is necessary to move the total value of the portfolio manager's position up from $370,000 to the required level of $450,000.

The examples in this section show that there are two reasons why the cost of hedging increases as the beta of a portfolio increases. More put options are required and they have a higher strike price.

15.2 CURRENCY OPTIONS

Currency options are primarily traded in the over-the-counter market. The advantage of this market is that large trades are possible, with strike prices, expiration dates, and other features tailored to meet the needs of corporate treasurers. Although European and American currency options do trade on the Philadelphia Stock Exchange in the United States, the exchange-traded market for these options is much smaller than the over-the-counter market.

An example of a European call option is a contract that gives the holder the right to buy one million euros with US dollars at an exchange rate of 1.2000 US dollars per euro. If the actual exchange rate at the maturity of the option is 1.2500, the payoff is

[1] Approximately 1% of $500,000, or $5,000, will be earned in dividends over the next three months. If we want the insured level of $450,000 to include dividends, we can choose a strike price corresponding to $445,000 rather than $450,000. This is 955.

$1{,}000{,}000 \times (1.2500 - 1.2000) = \$50{,}000$. Similarly, an example of a European put option is a contract that gives the holder the right to sell ten million Australian dollars for US dollars at an exchange rate of 0.7000 US dollars per Australian dollar. If the actual exchange rate at the maturity of the option is 0.6700, the payoff is $10{,}000{,}000 \times (0.7000 - 0.6700) = \$300{,}000$.

For a corporation wishing to hedge a foreign exchange exposure, foreign currency options are an interesting alternative to forward contracts. A company due to receive sterling at a known time in the future can hedge its risk by buying put options on sterling that mature at that time. The hedging strategy guarantees that the exchange rate applicable to the sterling will not be less than the strike price, while allowing the company to benefit from any favorable exchange-rate movements. Similarly, a company due to pay sterling at a known time in the future can hedge by buying calls on sterling that mature at that time. This hedging strategy guarantees that the cost of the sterling will not be greater than a certain amount while allowing the company to benefit from favorable exchange-rate movements. Whereas a forward contract locks in the exchange rate for a future transaction, an option provides a type of insurance. This is not free. It costs nothing to enter into a forward transaction, but options require a premium to be paid up front.

Range Forwards

A *range forward contract* is a variation on a standard forward contract for hedging foreign exchange risk. Consider a US company that knows it will receive one million pounds sterling in three months. Suppose that the three-month forward exchange rate is 1.9200 dollars per pound. The company could lock in this exchange rate for the dollars it receives by entering into a short forward contract to sell one million pounds sterling in three months. This would ensure that the amount received for the one million pounds is \$1,920,000.

An alternative is to buy a European put option with a strike price of K_1 and sell a European call option with a strike price K_2, where $K_1 < 1.9200 < K_2$. This is known as a short range forward contract. The payoff is shown in Figure 15.1(a). In both cases the options are on one million pounds. If the exchange rate in three months proves to be less than K_1, the put option is exercised and as a result the company is able to sell the one

Figure 15.1 Payoffs from (a) short and (b) long range-forward contract.

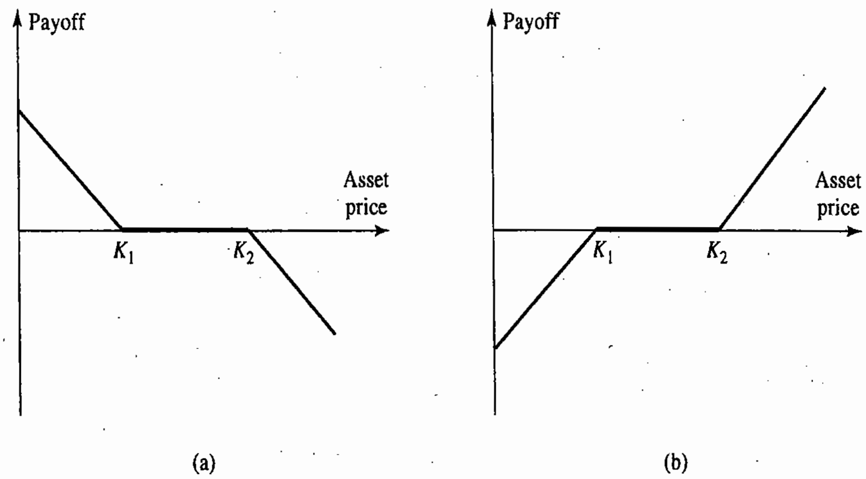
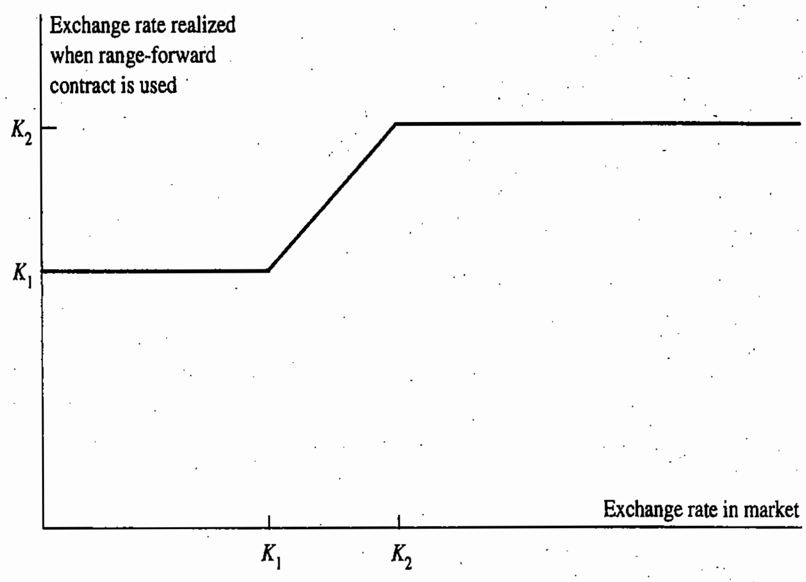


Figure 15.2 Exchange rate realized when either (a) a short range-forward contract is used to hedge a future foreign currency inflow or (b) a long range-forward contract is used to hedge a future foreign currency outflow.



million pounds at an exchange rate of K_1. If the exchange rate is between K_1 and K_2, neither option is exercised and the company gets the current exchange rate for the one million pounds. If the exchange rate is greater than K_2, the call option is exercised against the company with the result that the one million pounds is sold at an exchange rate of K_2. The exchange rate realized for the one million pounds is shown in Figure 15.2.

If the company knew it was due to pay rather than receive one million pounds in three months, it could sell a European put option with strike price K_1 and buy a European call option with strike price K_2. This is known as a long range forward contract and the payoff is shown in Figure 15.1(b). If the exchange rate in three months proves to be less than K_1, the put option is exercised against the company and as a result the company buys the one million pounds it needs at an exchange rate of K_1. If the exchange rate is between K_1 and K_2, neither option is exercised and the company buys the one million pounds at the current exchange rate. If the exchange rate is greater than K_2, the call option is exercised and the company is able to buy the one million pounds at an exchange rate of K_2. The exchange rate paid for the one million pounds is the same as that received for the one million pounds in the earlier example and is shown in Figure 15.2.

In practice, a range forward contract is set up so that the price of the put option equals the price of the call option. This means that it costs nothing to set up the range forward contract, just as it costs nothing to set up a regular forward contract. Suppose that the US and British interest rates are both 5%, so that the spot exchange rate is 1.9200 (the same as the forward exchange rate). Suppose further that the exchange rate volatility is 14%. We can use DerivaGem to show that a put with strike price 1.9000 to sell one pound has the same price as a call option with a strike price of 1.9413 to buy one pound. (Both are worth 0.04338.) Setting $K_1 = 1.9000$ and $K_2 = 1.9413$ therefore leads to a contract with zero cost in our example.

As the strike prices of the call and put options become closer in a range forward contract, the range forward contract becomes a regular forward contract. A short range forward contract becomes a short forward contract and a long range forward contract becomes a long forward contract.

15.3 OPTIONS ON STOCKS PAYING KNOWN DIVIDEND YIELDS

In this section we produce a simple rule that enables valuation results for European options on a non-dividend-paying stock to be extended so that they apply to European options on a stock paying a known dividend yield. Later we show how this enables us to value options on stock indices and currencies.

Dividends cause stock prices to reduce on the ex-dividend date by the amount of the dividend payment. The payment of a dividend yield at rate q therefore causes the growth rate in the stock price to be less than it would otherwise be by an amount q. If, with a dividend yield of q, the stock price grows from S_0 today to S_T at time T, then in the absence of dividends it would grow from S_0 today to $S_T e^{qT}$ at time T. Alternatively, in the absence of dividends it would grow from $S_0 e^{-qT}$ today to S_T at time T.

This argument shows that we get the same probability distribution for the stock price at time T in each of the following two cases:

1. The stock starts at price S_0 and provides a dividend yield at rate q.
2. The stock starts at price $S_0 e^{-qT}$ and pays no dividends.

This leads to a simple rule. When valuing a European option lasting for time T on a stock paying a known dividend yield at rate q, we reduce the current stock price from S_0 to $S_0 e^{-qT}$ and then value the option as though the stock pays no dividends.[2]

Lower Bounds for Option Prices

As a first application of this rule, consider the problem of determining bounds for the price of a European option on a stock paying a dividend yield at rate q. Substituting $S_0 e^{-qT}$ for S_0 in equation (9.1), we see that a lower bound for the European call option price, c, is given by

$$c \geqslant S_0 e^{-qT} - Ke^{-rT} \tag{15.1}$$

We can also prove this directly by considering the following two portfolios:

Portfolio A: one European call option plus an amount of cash equal to Ke^{-rT}

Portfolio B: e^{-qT} shares with dividends being reinvested in additional shares

To obtain a lower bound for a European put option, we can similarly replace S_0 by $S_0 e^{-qT}$ in equation (9.2) to get

$$p \geqslant Ke^{-rT} - S_0 e^{-qT} \tag{15.2}$$

[2] This rule is analogous to the one developed in Section 13.12 for valuing a European option on a stock paying known cash dividends. (In that case we concluded that it is correct to reduce the stock price by the present value of the dividends; in this case we discount the stock price at the dividend yield rate.)

This result can also be proved directly by considering the following portfolios:

Portfolio C: one European put option plus e^{-qT} shares with dividends on the shares being reinvested in additional shares

Portfolio D: an amount of cash equal to Ke^{-rT}

Put–Call Parity

Replacing S_0 by S_0e^{-qT} in equation (9.3) we obtain put–call parity for an option on a stock paying a dividend yield at rate q:

$$c + Ke^{-rT} = p + S_0e^{-qT} \tag{15.3}$$

This result can also be proved directly by considering the following two portfolios:

Portfolio A: one European call option plus an amount of cash equal to Ke^{-rT}

Portfolio C: one European put option plus e^{-qT} shares with dividends on the shares being reinvested in additional shares

Both portfolios are both worth $\max(S_T, K)$ at time T. They must therefore be worth the same today, and the put–call parity result in equation (15.3) follows. For American options, the put–call parity relationship is (see Problem 15.12)

$$S_0e^{-qT} - K \leqslant C - P \leqslant S_0 - Ke^{-rT}$$

Pricing Formulas

By replacing S_0 by S_0e^{-qT} in the Black–Scholes formulas, equations (13.20) and (13.21), we obtain the price, c, of a European call and the price, p, of a European put on a stock paying a dividend yield at rate q as

$$c = S_0e^{-qT}N(d_1) - Ke^{-rT}N(d_2) \tag{15.4}$$

$$p = Ke^{-rT}N(-d_2) - S_0e^{-qT}N(-d_1) \tag{15.5}$$

Since

$$\ln\frac{S_0e^{-qT}}{K} = \ln\frac{S_0}{K} - qT$$

it follows that d_1 and d_2 are given by

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

$$d_2 = \frac{\ln(S_0/K) + (r - q - \sigma^2/2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$$

These results were first derived by Merton.[3] As discussed in Chapter 13, the word *dividend* should, for the purposes of option valuation, be defined as the reduction in the stock price on the ex-dividend date arising from any dividends declared. If the dividend yield rate is known but not constant during the life of the option, equations (15.4)

[3] See R.C. Merton, "Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science*, 4 (Spring 1973): 141–83.

and (15.5) are still true, with q equal to the average annualized dividend yield during the option's life.

Differential Equation and Risk-Neutral Valuation

To prove the results in equations (15.4) and (15.5) more formally, we can either solve the differential equation that the option price must satisfy or use risk-neutral valuation.

When we include a dividend yield of q in the analysis in Section 13.6, the differential equation (13.16) becomes[4]

$$\frac{\partial f}{\partial t}+(r-q)S\frac{\partial f}{\partial S}+\tfrac{1}{2}\sigma^2S^2\frac{\partial^2 f}{\partial S^2}=rf \qquad \textbf{(15.6)}$$

Like equation (13.16), this does not involve any variable affected by risk preferences. Therefore the risk-neutral valuation procedure described in Section 13.7 can be used.

In a risk-neutral world, the total return from the stock must be r. The dividends provide a return of q. The expected growth rate in the stock price must therefore be $r-q$. It follows that the risk-neutral process for the stock price is

$$dS=(r-q)S\,dt+\sigma S\,dz \qquad \textbf{(15.7)}$$

To value a derivative dependent on a stock that provides a dividend yield equal to q, we set the expected growth rate of the stock equal to $r-q$ and discount the expected payoff at rate r. When the expected growth rate in the stock price is $r-q$, the expected stock price at time T is $S_0e^{(r-q)T}$. A similar analysis to that in the appendix to Chapter 13 gives the expected payoff for a call option in a risk-neutral world as

$$e^{(r-q)T}S_0N(d_1)-KN(d_2)$$

where d_1 and d_2 are defined as above. Discounting at rate r for time T leads to equation (15.4).

15.4 VALUATION OF EUROPEAN STOCK INDEX OPTIONS

In valuing index futures in Chapter 5, we assumed that the index could be treated as an asset paying a known yield. In valuing index options, we make similar assumptions. This means that equations (15.1) and (15.2) provide a lower bound for European index options; equation (15.3) is the put–call parity result for European index options; equations (15.4) and (15.5) can be used to value European options on an index; and the binomial tree approach can be used for American options. In all cases, S_0 is equal to the value of the index, σ is equal to the volatility of the index, and q is equal to the average annualized dividend yield on the index during the life of the option.

Example 15.1

Consider a European call option on the S&P 500 that is two months from maturity. The current value of the index is 930, the exercise price is 900, the risk-free interest rate is 8% per annum, and the volatility of the index is 20% per annum. Dividend

[4] See Technical Note 6 on the author's website for a proof of this.

yields of 0.2% and 0.3% are expected in the first month and the second month, respectively. In this case $S_0 = 930$, $K = 900$, $r = 0.08$, $\sigma = 0.2$, and $T = 2/12$. The total dividend yield during the option's life is 0.2% + 0.3% = 0.5%. This is 3% per annum. Hence, $q = 0.03$ and

$$d_1 = \frac{\ln(930/900) + (0.08 - 0.03 + 0.2^2/2) \times 2/12}{0.2\sqrt{2/12}} = 0.5444$$

$$d_2 = \frac{\ln(930/900) + (0.08 - 0.03 - 0.2^2/2) \times 2/12}{0.2\sqrt{2/12}} = 0.4628$$

$$N(d_1) = 0.7069, \qquad N(d_2) = 0.6782$$

so that the call price, c, is given by equation (15.4) as

$$c = 930 \times 0.7069e^{-0.03\times 2/12} - 900 \times 0.6782e^{-0.08\times 2/12} = 51.83$$

One contract would cost \$5,183.

The calculation of q should include only dividends whose ex-dividend date occurs during the life of the option. In the United States ex-dividend dates tend to occur during the first week of February, May, August, and November. At any given time the correct value of q is therefore likely to depend on the life of the option. This is even more true for some foreign indices. In Japan, for example, all companies tend to use the same ex-dividend dates.

If the absolute amount of the dividend that will be paid on the stocks underlying the index (rather than the dividend yield) is assumed to be known, the basic Black–Scholes formula can be used with the initial stock price being reduced by the present value of the dividends. This is the approach recommended in Chapter 13 for a stock paying known dividends. However, it may be difficult to implement for a broadly based stock index because it requires a knowledge of the dividends expected on every stock underlying the index.

It is sometimes argued that the return from a portfolio of stocks is certain to beat the return from a bond portfolio in the long run when both have the same initial value. If this were so, a long-dated put option on the stock portfolio where the strike price equaled the future value of the bond portfolio would not cost very much. In fact, as indicated by Business Snapshot 15.1, it is quite expensive.

Forward Prices

Define F_0 as the forward price of the index for a contract with maturity T. As shown by equation (5.3), $F_0 = S_0e^{(r-q)T}$. This means that the equations for the European call price c and the European put price p in equations (15.4) and (15.5) can be written

$$c = F_0e^{-rT}N(d_1) - Ke^{-rT}N(d_2) \tag{15.8}$$

$$p = Ke^{-rT}N(-d_2) - F_0e^{-rT}N(-d_1) \tag{15.9}$$

where

$$d_1 = \frac{\ln(F/K) + \sigma^2T/2}{\sigma\sqrt{T}} \quad \text{and} \quad d_2 = \frac{\ln(F/K) - \sigma^2T/2}{\sigma\sqrt{T}}$$

Business Snapshot 15.1 Can We Guarantee that Stocks Will Beat Bonds in the Long Run?

It is often said that if you are a long-term investor you should buy stocks rather than bonds. Consider a US fund manager who is trying to persuade investors to buy, as a long-term investment, an equity fund that is expected to mirror the S&P 500. The manager might be tempted to offer purchasers of the fund a guarantee that their return will be at least as good as the return on risk-free bonds over the next 10 years. Historically stocks have outperformed bonds in the United States over almost any 10-year period. It appears that the fund manager would not be giving much away.

In fact, this type of guarantee is surprisingly expensive. Suppose that an equity index is 1,000 today, the dividend yield on the index is 1% per annum, the volatility of the index is 15% per annum, and the 10-year risk-free rate is 5% per annum. To outperform bonds, the stocks underlying the index must earn more than 5% per annum. The dividend yield will provide 1% per annum. The capital gains on the stocks must therefore provide 4% per annum. This means that we require the index level to be at least $1{,}000e^{0.04\times 10} = 1{,}492$ in 10 years.

A guarantee that the return on \$1,000 invested in the index will be greater than the return on \$1,000 invested in bonds over the next 10 years is therefore equivalent to the right to sell the index for 1,492 in 10 years. This is a European put option on the index and can be valued from equation (15.5) with $S_0 = 1{,}000$, $K = 1{,}492$, $r = 5\%$, $\sigma = 15\%$, $T = 10$, and $q = 1\%$. The value of the put option is 169.7. This shows that the guarantee contemplated by the fund manager is worth about 17% of the fund—hardly something that should be given away!

The put–call parity relationship in equation (15.3) can be written

$$c + Ke^{-rT} = p + F_0e^{-rT}$$

or

$$F_0 = K + (c - p)e^{rT} \tag{15.10}$$

If, as is not uncommon in the exchange-traded markets, pairs of puts and calls with the same strike price are traded actively for a particular maturity date, this equation can be used to estimate the forward price of the index for that maturity date. Once the forward prices of the index for a number of different maturity dates have been obtained, the term structure of forward rates can be estimated, and other options can be valued using equations (15.8) and (15.9). The advantage of this approach is that the dividend yield on the index does not have to be estimated explicitly.

Implied Dividend Yields

If estimates of the dividend yield are required (e.g because an American option is being valued), calls and puts with the same strike price and time to maturity can be used. From equation (15.3),

$$q = -\frac{1}{T}\ln\frac{c - p + Ke^{-rT}}{S_0}$$

For a particular strike price and time to maturity, the estimates of q calculated from this equation are liable to be unreliable. But when the results from many matched pairs of calls and puts are combined, a clear picture of the dividend yield being assumed by the market emerges.

15.5 VALUATION OF EUROPEAN CURRENCY OPTIONS

To value currency options, we define S_0 as the spot exchange rate. To be precise, S_0 is the value of one unit of the foreign currency in US dollars. As explained in Section 5.10, a foreign currency is analogous to a stock paying a known dividend yield. The owner of foreign currency receives a yield equal to the risk-free interest rate, r_f, in the foreign currency. Equations (15.1) and (15.2), with q replaced by r_f, provide bounds for the European call price, c, and the European put price, p:

$$c \geqslant S_0 e^{-r_f T} - K e^{-rT}$$

$$p \geqslant K e^{-rT} - S_0 e^{-r_f T}$$

Equation (15.3), with q replaced by r_f, provides the put–call parity result for currency options:

$$c + K e^{-rT} = p + S_0 e^{-r_f T}$$

Finally, equations (15.4) and (15.5) provide the pricing formulas for currency options when q is replaced by r_f:

$$c = S_0 e^{-r_f T} N(d_1) - K e^{-rT} N(d_2) \qquad \textbf{(15.11)}$$

$$p = K e^{-rT} N(-d_2) - S_0 e^{-r_f T} N(-d_1) \qquad \textbf{(15.12)}$$

where

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

$$d_2 = \frac{\ln(S_0/K) + (r - r_f - \sigma^2/2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$$

Both the domestic interest rate, r, and the foreign interest rate, r_f, are the rates for a maturity T. Put and call options on a currency are symmetrical in that a put option to sell currency A for currency B at strike price K is the same as a call option to buy B with currency A at strike price $1/K$ (see Problem 15.8).

Example 15.2

Consider a four-month European call option on the British pound. Suppose that the current exchange rate is 1.6000, the exercise price is 1.6000, the risk-free interest rate in the United States is 8% per annum, the risk-free interest rate in Britain is 11% per annum, and the option price is 4.3 cents. In this case, $S_0 = 1.6$, $K = 1.6$, $r = 0.08$, $r_f = 0.11$, $T = 0.3333$, and $c = 0.043$. The implied volatility can be calculated by trial and error. A volatility of 20% gives an option price of 0.0639; a volatility of 10% gives an option price of 0.0285; and so on. The implied volatility is 14.1%.

Using Forward Exchange Rates

Because banks and other financial institutions trade forward contracts on foreign exchange rates actively, foreign exchange rates are often used for valuing options.

From equation (5.9), the forward rate, F_0, for a maturity T is given by

$$F_0 = S_0 e^{(r-r_f)T}$$

This relationship allows equations (15.11) and (15.12) to be simplified to

$$c = e^{-rT}[F_0 N(d_1) - KN(d_2)] \tag{15.13}$$

$$p = e^{-rT}[KN(-d_2) - F_0 N(-d_1)] \tag{15.14}$$

where

$$d_1 = \frac{\ln(F_0/K) + \sigma^2 T/2}{\sigma\sqrt{T}}$$

$$d_2 = \frac{\ln(F_0/K) - \sigma^2 T/2}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$$

These equations are the same as equations (15.8) and (15.9). As we shall see in Chapter 16, a European option on any asset can be valued in terms of forward or futures contracts on the asset using equations (15.13) and (15.14). Note that the maturity of the forward or futures contract must be the same as the maturity of the European option.

15.6 AMERICAN OPTIONS

As described in Chapter 11, binomial trees can be used to value American options on indices and currencies. As in the case of American options on a non-dividend-paying stock, the parameter determining the size of up movements, u, is set equal to $e^{\sigma\sqrt{\Delta t}}$, where σ is the volatility and Δt is the length of time steps. The parameter determining the size of down movements, d, is set equal to $1/u$, or $e^{-\sigma\sqrt{\Delta t}}$. For a non-dividend-paying stock, the probability of an up movement is

$$p = \frac{a-d}{u-d}$$

where $a = e^{r\Delta t}$. For options on indices and currencies, the formula for p is the same, but a is defined differently. In the case of options on an index,

$$a = e^{(r-q)\Delta t}$$

where q is the dividend yield on the index. In the case of options on a currency,

$$a = e^{(r-r_f)\Delta t}$$

where r_f is the foreign risk-free rate. Example 11.1 in Section 11.9 shows how a two-step tree can be constructed to value an option on an index. Example 11.2 shows how a three-step tree can be constructed to value an option on a currency. Further examples of the use of binomial trees to value options on indices and currencies are given in Chapter 19.

In some circumstances it is optimal to exercise American currency options prior to maturity. Thus, American currency options are worth more than their European counterparts. In general, call options on high-interest currencies and put options on low-interest currencies are the most likely to be exercised prior to maturity. The reason is that a high-interest currency is expected to depreciate and a low-interest currency is expected to appreciate.

SUMMARY

The index options that trade on exchanges are settled in cash. On exercise of an index call option, the holder receives 100 times the amount by which the index exceeds the strike price. Similarly, on exercise of an index put option contract, the holder receives 100 times the amount by which the strike price exceeds the index. Index options can be used for portfolio insurance. If the value of the portfolio mirrors the index, it is appropriate to buy one put option contract for each $100S_0$ dollars in the portfolio, where S_0 is the value of the index. If the portfolio does not mirror the index, β put option contracts should be purchased for each $100S_0$ dollars in the portfolio, where β is the beta of the portfolio calculated using the capital asset pricing model. The strike price of the put options purchased should reflect the level of insurance required.

Most currency options are traded in the over-the-counter market. They can be used by corporate treasurers to hedge a foreign exchange exposure. For example, a US corporate treasurer who knows that the company will be receiving sterling at a certain time in the future can hedge by buying put options that mature at that time. Similarly, a US corporate treasurer who knows that the company will be paying sterling at a certain time in the future can hedge by buying call options that mature at that time. Currency options can also be used to create a range forward contract. This is a zero-cost contract that provides downside protection while giving up some of the upside.

The Black–Scholes formula for valuing European options on a non-dividend-paying stock can be extended to cover European options on a stock paying a known dividend yield. The extension can be used to value European options on stock indices and currencies because:

1. A stock index is analogous to a stock paying a dividend yield. The dividend yield is the dividend yield on the stocks that make up the index.
2. A foreign currency is analogous to a stock paying a dividend yield. The foreign risk-free interest rate plays the role of the dividend yield.

Binomial trees can be used to value American options on stock indices and currencies.

FURTHER READING

Amin, K., and R.A. Jarrow. "Pricing Foreign Currency Options under Stochastic Interest Rates," *Journal of International Money and Finance*, 10 (1991): 310–29.

Biger, N., and J.C. Hull. "The Valuation of Currency Options," *Financial Management*, 12 (Spring 1983): 24–28.

Bodie, Z. "On the Risk of Stocks in the Long Run," *Financial Analysts Journal*, 51, 3 (1995): 18–22.

Garman, M.B., and S.W. Kohlhagen. "Foreign Currency Option Values," *Journal of International Money and Finance*, 2 (December 1983): 231–37.

Giddy, I.H., and G. Dufey. "Uses and Abuses of Currency Options," *Journal of Applied Corporate Finance*, 8, 3 (1995): 49–57.

Grabbe, J.O. "The Pricing of Call and Put Options on Foreign Exchange," *Journal of International Money and Finance*, 2 (December 1983): 239–53.

Jorion, P. "Predicting Volatility in the Foreign Exchange Market," *Journal of Finance* 50, 2(1995): 507–28.

Merton, R.C. "Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science*, 4 (Spring 1973): 141–83.

Questions and Problems (Answers in Solutions Manual)

15.1. A portfolio is currently worth \$10 million and has a beta of 1.0. An index is currently standing at 800. Explain how a put option on the index with a strike price of 700 can be used to provide portfolio insurance.

15.2. "Once we know how to value options on a stock paying a dividend yield, we know how to value options on stock indices and currencies." Explain this statement.

15.3. A stock index is currently 300, the dividend yield on the index is 3% per annum, and the risk-free interest rate is 8% per annum. What is a lower bound for the price of a six-month European call option on the index when the strike price is 290?

15.4. A currency is currently worth \$0.80 and has a volatility of 12%. The domestic and foreign risk-free interest rates are 6% and 8%, respectively. Use a two-step binomial tree to value (a) a European four-month call option with a strike price of 0.79 and (b) an American four-month call option with the same strike price.

15.5. Explain how corporations can use range forward contracts to hedge their foreign exchange risk.

15.6. Calculate the value of a three-month at-the-money European call option on a stock index when the index is at 250, the risk-free interest rate is 10% per annum, the volatility of the index is 18% per annum, and the dividend yield on the index is 3% per annum.

15.7. Calculate the value of an eight-month European put option on a currency with a strike price of 0.50. The current exchange rate is 0.52, the volatility of the exchange rate is 12%, the domestic risk-free interest rate is 4% per annum, and the foreign risk-free interest rate is 8% per annum.

15.8. Show that the formula in equation (15.12) for a put option to sell one unit of currency A for currency B at strike price K gives the same value as equation (15.11) for a call option to buy K units of currency B for currency A at strike price $1/K$.

15.9. A foreign currency is currently worth \$1.50. The domestic and foreign risk-free interest rates are 5% and 9%, respectively. Calculate a lower bound for the value of a six-month call option on the currency with a strike price of \$1.40 if it is (a) European and (b) American.

15.10. Consider a stock index currently standing at 250. The dividend yield on the index is 4% per annum, and the risk-free rate is 6% per annum. A three-month European call option on the index with a strike price of 245 is currently worth \$10. What is the value of a three-month put option on the index with a strike price of 245?

15.11. An index currently stands at 696 and has a volatility of 30% per annum. The risk-free rate of interest is 7% per annum and the index provides a dividend yield of 4% per annum. Calculate the value of a three-month European put with an exercise price of 700.

15.12. Show that, if C is the price of an American call with exercise price K and maturity T on a stock paying a dividend yield of q, and P is the price of an American put on the same stock with the same strike price and exercise date, then

$$S_0e^{-qT} - K < C - P < S_0 - Ke^{-rT},$$

where S_0 is the stock price, r is the risk-free rate, and $r > 0$. (*Hint*: To obtain the first half of the inequality, consider possible values of:

Portfolio A: a European call option plus an amount K invested at the risk-free rate
Portfolio B: an American put option plus e^{-qT} of stock with dividends being reinvested in the stock

To obtain the second half of the inequality, consider possible values of:

Portfolio C: an American call option plus an amount Ke^{-rT} invested at the risk-free rate
Portfolio D: a European put option plus one stock with dividends being reinvested in the stock.)

15.13. Show that a European call option on a currency has the same price as the corresponding European put option on the currency when the forward price equals the strike price.

15.14. Would you expect the volatility of a stock index to be greater or less than the volatility of a typical stock? Explain your answer.

15.15. Does the cost of portfolio insurance increase or decrease as the beta of a portfolio increases? Explain your answer.

15.16. Suppose that a portfolio is worth \$60 million and the S&P 500 is at 1,200. If the value of the portfolio mirrors the value of the index, what options should be purchased to provide protection against the value of the portfolio falling below \$54 million in one year's time?

15.17. Consider again the situation in Problem 15.16. Suppose that the portfolio has a beta of 2.0, the risk-free interest rate is 5% per annum, and the dividend yield on both the portfolio and the index is 3% per annum. What options should be purchased to provide protection against the value of the portfolio falling below \$54 million in one year's time?

15.18. An index currently stands at 1,500. European call and put options with a strike price of 1,400 and time to maturity of six months have market prices of 154.00 and 34.25, respectively. The six-month risk-free rate is 5%. What is the implied dividend yield?

15.19. A total return index tracks the return, including dividends, on a certain portfolio. Explain how you would value (a) forward contracts and (b) European options on the index.

15.20. What is the put–call parity relationship for European currency options?

15.21. Can an option on the yen–euro exchange rate be created from two options, one on the dollar–euro exchange rate, and the other on the dollar–yen exchange rate? Explain your answer.

15.22. Prove the results in equations (15.1), (15.2), and (15.3) using the portfolios indicated.

Assignment Questions

15.23. The Dow Jones Industrial Average on January 12, 2007, was 12,556 and the price of the March 126 call was \$2.25. Use the DerivaGem software to calculate the implied volatility of this option. Assume the risk-free rate was 5.3% and the dividend yield was 3%. The option expires on March 20, 2007. Estimate the price of a March 126 put. What is the volatility implied by the price you estimate for this option? (Note that options are on the Dow Jones index divided by 100.)

15.24. A stock index currently stands at 300 and has a volatility of 20%. The risk-free interest rate is 8% and the dividend yield on the index is 3%. Use a three-step binomial tree to value a six-month put option on the index with a strike price of 300 if it is (a) European and (b) American?

15.25. Suppose that the spot price of the Canadian dollar is US \$0.85 and that the Canadian dollar/US dollar exchange rate has a volatility of 4% per annum. The risk-free rates of interest in Canada and the United States are 4% and 5% per annum, respectively. Calculate the value of a European call option to buy one Canadian dollar for US \$0.85 in nine months. Use put–call parity to calculate the price of a European put option to sell one Canadian dollar for US \$0.85 in nine months. What is the price of a call option to buy US \$0.85 with one Canadian dollar in nine months?

15.26. A mutual fund announces that the salaries of its fund managers will depend on the performance of the fund. If the fund loses money, the salaries will be zero. If the fund makes a profit, the salaries will be proportional to the profit. Describe the salary of a fund manager as an option. How is a fund manager motivated to behave with this type of remuneration package?

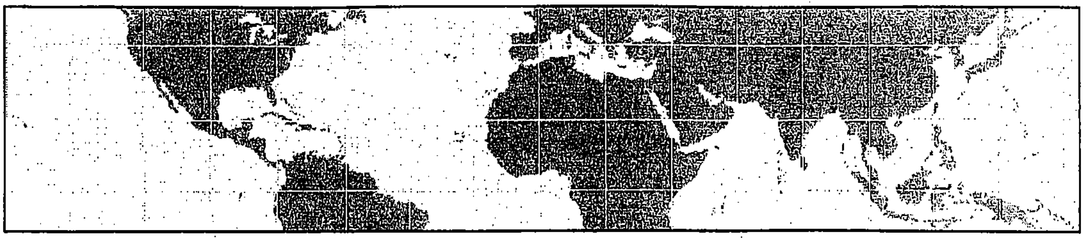
15.27. Assume that the price of currency A expressed in terms of the price of currency B follows the process

$$dS = (r_B - r_A)S\,dt + \sigma S\,dz$$

where r_A is the risk-free interest rate in currency A and r_B is the risk-free interest rate in currency B. What is the process followed by the price of currency B expressed in terms of currency A?

15.28. The three-month forward USD/euro exchange rate is 1.3000. The exchange rate volatility is 15%. A US company will have to pay 1 million euros in three months. The euro and USD risk-free rates are 5% and 4%, respectively. The company decides to use a range forward contract with the lower strike price equal to 1.2500.

(a) What should the higher strike price be to create a zero-cost contract?
(b) What position in calls and puts should the company take?
(c) Does your answer depend on the euro risk-free rate? Explain.
(d) Does your answer depend on the USD risk-free rate? Explain.



CHAPTER 16

Futures Options

The options we have considered so far provide the holder with the right to buy or sell a certain asset by a certain date. They are sometimes termed *options on spot* or *spot options* because, when the options are exercised, the sale or purchase of the asset at the agreed-on price takes place immediately. In this chapter we move on to consider *options on futures*, also known as *futures options*. In these contracts, exercise of the option gives the holder a position in a futures contract.

The Commodity Futures Trading Commission authorized the trading of options on futures on an experimental basis in 1982. Permanent trading was approved in 1987, and since then the popularity of the contract with investors has grown very fast.

In this chapter we consider how futures options work and the differences between these options and spot options. We examine how futures options can be priced using either binomial trees or formulas similar to those produced by Black, Scholes, and Merton for stock options. We also explore the relative pricing of futures options and spot options.

16.1 NATURE OF FUTURES OPTIONS

A futures option is the right, but not the obligation, to enter into a futures contract at a certain futures price by a certain date. Specifically, a call futures option is the right to enter into a long futures contract at a certain price; a put futures option is the right to enter into a short futures contract at a certain price. Futures options are generally American; that is, they can be exercised any time during the life of the contract.

If a call futures option is exercised, the holder acquires a long position in the underlying futures contract plus a cash amount equal to the most recent settlement futures price minus the strike price. If a put futures option is exercised, the holder acquires a short position in the underlying futures contract plus a cash amount equal to the strike price minus the most recent settlement futures price. As the following examples show, the effective payoff from a call futures option is the futures price at the time of exercise less the strike price; the effective payoff from a put futures option is the strike price less the futures price at the time of exercise.

Example 16.1

Suppose it is August 15 and an investor has one September futures call option contract on copper with a strike price of 240 cents per pound. One futures contract is on 25,000 pounds of copper. Suppose that the futures price of copper for delivery in September is currently 251 cents, and at the close of trading on August 14 (the last settlement) it was 250 cents. If the option is exercised, the investor receives a cash amount of

$$25{,}000 \times (250 - 240)\,\text{cents} = \$2{,}500$$

plus a long position in a futures contract to buy 25,000 pounds of copper in September. If desired, the position in the futures contract can be closed out immediately. This would leave the investor with the $2,500 cash payoff plus an amount

$$25{,}000 \times (251 - 250)\,\text{cents} = \$250$$

reflecting the change in the futures price since the last settlement. The total payoff from exercising the option on August 15 is $2,750, which equals 25,000($F - K$), where F is the futures price at the time of exercise and K is the strike price.

Example 16.2

An investor has one December futures put option on corn with a strike price of 400 cents per bushel. One futures contract is on 5,000 bushels of corn. Suppose that the current futures price of corn for delivery in December is 380, and the most recent settlement price is 379 cents. If the option is exercised, the investor receives a cash amount of

$$5{,}000 \times (400 - 379)\,\text{cents} = \$1{,}050$$

plus a short position in a futures contract to sell 5,000 bushels of corn in December. If desired, the position in the futures contract can be closed out. This would leave the investor with the $1,050 cash payoff minus an amount

$$5{,}000 \times (380 - 379)\,\text{cents} = \$50$$

reflecting the change in the futures price since the last settlement. The net payoff from exercise is $1,000, which equals 5,000($K - F$), where F is the futures price at the time of exercise and K is the strike price.

Expiration Months

Futures options are referred to by the delivery month of the underlying futures contract —not by the expiration month of the option. As mentioned earlier, most futures options are American. The expiration date of a futures option contract is usually on, or a few days before, the earliest delivery date of the underlying futures contract. (For example, the CBOT Treasury bond futures option expires on the latest Friday that precedes by at least five business days the end of the month before the futures delivery month.) An exception is the CME mid-curve Eurodollar contract where the futures contract expires either one or two years after the options contract.

Popular contracts trading in the United States are those on corn, soybeans, cotton, sugar-world, crude oil, natural gas, gold, Treasury bonds, Treasury notes, five-year Treasury notes, 30-day federal funds, Eurodollars, one-year and two-year mid-curve Eurodollars, Euribor, Eurobunds, and the S&P 500.

Options on Interest Rate Futures

The most actively traded interest rate options offered by exchanges in the United States are those on Treasury bond futures, Treasury note futures, and Eurodollar futures.

A Treasury bond futures option, which is traded on the Chicago Board of Trade, is an option to enter a Treasury bond futures contract. As mentioned in Chapter 6, one Treasury bond futures contract is for the delivery of \$100,000 of Treasury bonds. The price of a Treasury bond futures option is quoted as a percentage of the face value of the underlying Treasury bonds to the nearest sixty-fourth of 1%.

An option on Eurodollar futures, which is traded on the Chicago Mercantile Exchange, is an option to enter into a Eurodollar futures contract. As explained in Chapter 6, when the Eurodollar futures quote changes by 1 basis point, or 0.01%, there is a gain or loss on a Eurodollar futures contract of \$25. Similarly, in the pricing of options on Eurodollar futures, 1 basis point represents \$25.

Interest rate futures option contracts work in the same way as the other futures options contracts discussed in this chapter. For example, in addition to the cash payoff, the holder of a call option obtains a long position in the futures contract when the option is exercised and the option writer obtains a corresponding short position. The total payoff from the call, including the value of the futures position, is $\max(F - K, 0)$, where F is the futures price at the time of exercise and K is the strike price.

Interest rate futures prices increase when bond prices increase (i.e., when interest rates fall). They decrease when bond prices decrease (i.e., when interest rates rise). An investor who thinks that short-term interest rates will rise can speculate by buying put options on Eurodollar futures, whereas an investor who thinks the rates will fall can speculate by buying call options on Eurodollar futures. An investor who thinks that long-term interest rates will rise can speculate by buying put options on Treasury note futures or Treasury bond futures, whereas an investor who thinks the rates will fall can speculate by buying call options on these instruments.

Example 16.3

It is February and the futures price for the June Eurodollar contract is 93.82 (corresponding to a 3-month Eurodollar interest rate of 6.18% per annum). The price of a call option on the contract with a strike price of 94.00 is quoted as 0.1, or 10 basis points. This option could be attractive to an investor who feels that interest rates are likely to come down. Suppose that short-term interest rates do drop by about 100 basis points and the investor exercises the call when the Eurodollar futures price is 94.78 (corresponding to a 3-month Eurodollar interest rate of 5.22% per annum). The payoff is $25 \times (94.78 - 94.00) \times 100 = \$1{,}950$. The cost of the contract is $10 \times 25 = \$250$. The investor's profit is therefore \$1,700.

Example 16.4

It is August and the futures price for the December Treasury bond contract traded on the CBOT is 96-09 (or $96\frac{9}{32} = 96.28125$). The yield on long-term government bonds is about 6.4% per annum. An investor who feels that this yield will fall by December might choose to buy December calls with a strike price of 98. Assume that the price of these calls is 1-04 (or $1\frac{4}{64} = 1.0625\%$ of the principal). If long-term rates fall to 6% per annum and the Treasury bond futures price rises to 100-00, the investor will make a net profit per \$100 of bond futures of

$$100.00 - 98.00 - 1.0625 = 0.9375$$

Since one option contract is for the purchase or sale of instruments with a face value of \$100,000, the investor would make a profit of \$937.50 per option contract bought.

16.2 REASONS FOR THE POPULARITY OF FUTURES OPTIONS

It is natural to ask why people choose to trade options on futures rather than options on the underlying asset. The main reason appears to be that a futures contract is, in many circumstances, more liquid and easier to trade than the underlying asset. Furthermore, a futures price is known immediately from trading on the futures exchange, whereas the spot price of the underlying asset may not be so readily available.

Consider Treasury bonds. The market for Treasury bond futures is much more active than the market for any particular Treasury bond. Also, a Treasury bond futures price is known immediately from trading on the Chicago Board of Trade. By contrast, the current market price of a bond can be obtained only by contacting one or more dealers. It is not surprising that investors would rather take delivery of a Treasury bond futures contract than Treasury bonds.

Futures on commodities are also often easier to trade than the commodities themselves. For example, it is much easier and more convenient to make or take delivery of a live-cattle futures contract than it is to make or take delivery of the cattle themselves.

An important point about a futures option is that exercising it does not usually lead to delivery of the underlying asset, as in most circumstances the underlying futures contract is closed out prior to delivery. Futures options are therefore normally eventually settled in cash. This is appealing to many investors, particularly those with limited capital who may find it difficult to come up with the funds to buy the underlying asset when an option is exercised. Another advantage sometimes cited for futures options is that futures and futures options are traded in pits side by side in the same exchange. This facilitates hedging, arbitrage, and speculation. It also tends to make the markets more efficient. A final point is that futures options tend to entail lower transactions costs than spot options in many situations.

16.3 EUROPEAN SPOT AND FUTURES OPTIONS

The payoff from a European call option with strike price K on the spot price of an asset is

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

where S_T is the spot price at the option's maturity. The payoff from a European call option with the same strike price on the futures price of the asset is

$$\max(F_T - K, 0)$$

where F_T is the futures price at the option's maturity. If the futures contract matures at the same time as the option, then $F_T = S_T$ and the two options are equivalent. Similarly, a European futures put option is worth the same as its spot put option counterpart when the futures contract matures at the same time as the option.

Most of the futures options that trade are American-style. However, as we shall see, it is useful to study European futures options because the results that are obtained can be used to value the corresponding European spot options.

16.4 PUT–CALL PARITY

In Chapter 9 we derived a put–call parity relationship for European stock options. We now consider a similar argument to derive a put–call parity relationship for European futures options. Consider European call and put futures options, both with strike price K and time to expiration T. We can form two portfolios:

Portfolio A: a European call futures option plus an amount of cash equal to Ke^{-rT}
Portfolio B: a European put futures option plus a long futures contract plus an amount of cash equal to F_0e^{-rT}, where F_0 is the futures price

In portfolio A, the cash can be invested at the risk-free rate, r, and grows to K at time T. Let F_T be the futures price at maturity of the option. If $F_T > K$, the call option in portfolio A is exercised and portfolio A is worth F_T. If $F_T \leqslant K$, the call is not exercised and portfolio A is worth K. The value of portfolio A at time T is therefore

$$\max(F_T, K)$$

In portfolio B, the cash can be invested at the risk-free rate to grow to F_0 at time T. The put option provides a payoff of $\max(K - F_T, 0)$. The futures contract provides a payoff of $F_T - F_0$.[1] The value of portfolio B at time T is therefore

$$F_0 + (F_T - F_0) + \max(K - F_T, 0) = \max(F_T, K)$$

Because the two portfolios have the same value at time T and European options cannot be exercised early, it follows that they are worth the same today. The value of portfolio A today is

$$c + Ke^{-rT}$$

where c is the price of the call futures option. The marking-to-market process ensures that the futures contract in portfolio B is worth zero today. Portfolio B is therefore worth

$$p + F_0e^{-rT}$$

where p is the price of the put futures option. Hence

$$c + Ke^{-rT} = p + F_0e^{-rT} \tag{16.1}$$

The difference between this put–call parity relationship and the one for a non-dividend-paying stock in equation (9.3) is that the stock price, S_0, is replaced by the discounted futures price, F_0e^{-rT}.

[1] This analysis assumes that a futures contract is like a forward contract and settled at the end of its life rather than on a day-to-day basis.

As shown in Section 16.3, when the underlying futures contract matures at the same time as the option, European futures and spot options are the same. Equation (16.1) therefore gives a relationship between the price of a call option on the spot price, the price of a put option on the spot price, and the futures price when both options mature at the same time as the futures contract.

Example 16.5

Suppose that the price of a European call option on spot silver for delivery in six months is \$0.56 per ounce when the exercise price is \$8.50. Assume that the silver futures price for delivery in six months is currently \$8.00, and the risk-free interest rate for an investment that matures in six months is 10% per annum. From a rearrangement of equation (16.1), the price of a European put option on spot silver with the same maturity and exercise date as the call option is

$$0.56 + 8.50e^{-0.1\times 6/12} - 8.00e^{-0.1\times 6/12} = 1.04$$

We can use equation (16.1) for spot options because the futures price that is considered has the same maturity as the option price.

For American futures options, the put–call relationship is (see Problem 16.19)

$$F_0e^{-rT} - K < C - P < F_0 - Ke^{-rT} \tag{16.2}$$

16.5 BOUNDS FOR FUTURES OPTIONS

The put–call parity relationship in equation (16.1) provides bounds for European call and put options. Because the price of a put, p, cannot be negative, it follows from equation (16.1) that

$$c + Ke^{-rT} \geqslant F_0e^{-rT}$$

or

$$c \geqslant (F_0 - K)e^{-rT} \tag{16.3}$$

Similarly, because the price of a call option cannot be negative, it follows from equation (16.1) that

$$Ke^{-rT} \leqslant F_0e^{-rT} + p$$

or

$$p \geqslant (K - F_0)e^{-rT} \tag{16.4}$$

These bounds are similar to the ones derived for European stock options in Chapter 9. The prices of European call and put options are very close to their lower bounds when the options are deep in the money. To see why this is so, we return to the put–call parity relationship in equation (16.1). When a call option is deep in the money, the corresponding put option is deep out of the money. This means that p is very close to zero. The difference between c and its lower bound equals p, so that the price of the call option must be very close to its lower bound. A similar argument applies to put options.

Because American futures options can be exercised at any time, we must have

$$C \geqslant F_0 - K$$

and

$$P \geqslant K - F_0$$

Thus, assuming interest rates are positive, the lower bound for an American option price is always higher than the lower bound for a European option price. This is because there is always some chance that an American futures option will be exercised early.

16.6 VALUATION OF FUTURES OPTIONS USING BINOMIAL TREES

This section examines, more formally than in Chapter 11, how binomial trees can be used to price futures options. A key difference between futures options and stock options is that there are no up-front costs when a futures contract is entered into.

Suppose that the current futures price is 30 and that it will move either up to 33 or down to 28 over the next month. We consider a one-month call option on the futures with a strike price of 29 and ignore daily settlement. The situation is as indicated in Figure 16.1. If the futures price proves to be 33, the payoff from the option is 4 and the value of the futures contract is 3. If the futures price proves to be 28, the payoff from the option is zero and the value of the futures contract is -2.[2]

To set up a riskless hedge, we consider a portfolio consisting of a short position in one options contract and a long position in Δ futures contracts. If the futures price moves up to 33, the value of the portfolio is $3\Delta - 4$; if it moves down to 28, the value of the portfolio is -2Δ. The portfolio is riskless when these are the same, that is, when

$$3\Delta - 4 = -2\Delta$$

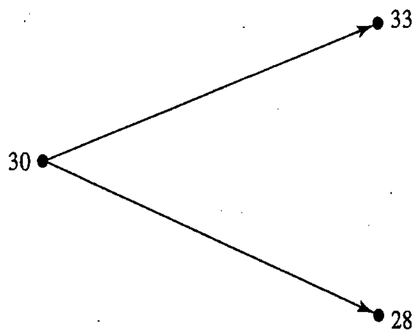
or $\Delta = 0.8$.

For this value of Δ, we know the portfolio will be worth $3 \times 0.8 - 4 = -1.6$ in one month. Assume a risk-free interest rate of 6%. The value of the portfolio today must be

$$-1.6e^{-0.06\times 1/12} = -1.592$$

The portfolio consists of one short option and Δ futures contracts. Because the value of the futures contract today is zero, the value of the option today must be 1.592.

Figure 16.1 Futures price movements in numerical example.



[2] There is an approximation here in that the gain or loss on the futures contract is not realized at time T. It is realized day by day between time 0 and time T. However, as the length of the time step in a binomial tree becomes shorter, the approximation becomes better, and in the limit, as the time step tends to zero, there is no loss of accuracy.

A Generalization

We can generalize this analysis by considering a futures price that starts at F_0 and is anticipated to rise to F_0u or move down to F_0d over the time period T. We consider an option maturing at time T and suppose that its payoff is f_u if the futures price moves up and f_d if it moves down. The situation is summarized in Figure 16.2.

The riskless portfolio in this case consists of a short position in one option combined with a long position in Δ futures contracts, where

$$\Delta = \frac{f_u - f_d}{F_0u - F_0d}$$

The value of the portfolio at time T is then always

$$(F_0u - F_0)\Delta - f_u$$

Denoting the risk-free interest rate by r, we obtain the value of the portfolio today as

$$[(F_0u - F_0)\Delta - f_u]e^{-rT}$$

Another expression for the present value of the portfolio is $-f$, where f is the value of the option today. It follows that

$$-f = [(F_0u - F_0)\Delta - f_u]e^{-rT}$$

Substituting for Δ and simplifying reduces this equation to

$$f = e^{-rT}[pf_u + (1-p)f_d] \tag{16.5}$$

where

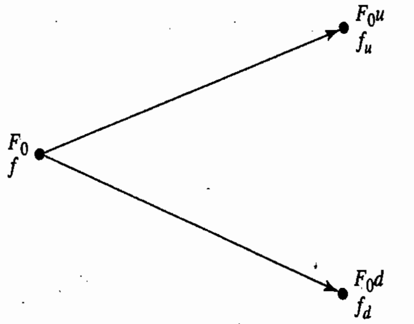
$$p = \frac{1-d}{u-d} \tag{16.6}$$

This agrees with the result in Section 11.9. Equation (16.6) gives the risk-neutral probability of an up movement.

In the numerical example considered previously (see Figure 16.1), $u = 1.1$, $d = 0.9333$, $r = 0.06$, $T = 1/12$, $f_u = 4$, and $f_d = 0$. From equation (16.6),

$$p = \frac{1 - 0.9333}{1.1 - 0.9333} = 0.4$$

Figure 16.2 Futures price and option price in the general situation.



and, from equation (16.5),

$$f = e^{-0.06\times 1/12}[0.4 \times 4 + 0.6 \times 0] = 1.592$$

This result agrees with the answer obtained for this example earlier.

Multistep Trees

Multistep binomial trees are used to value American-style futures options in much the same way that they are used to value options on stocks. This is explained in Section 11.9. The parameter u defining up movements in the futures price is $e^{\sigma\sqrt{\Delta t}}$, where σ is the volatility of the futures price and Δt is the length of one time step. The probability of an up movement in the future price is that in equation (16.6):

$$p = \frac{1-d}{u-d}$$

Example 11.13 illustrates the use of multistep binomial trees for valuing a futures option. Example 19.3 in Chapter 19 provides a further illustration.

16.7 DRIFT OF A FUTURES PRICE IN A RISK-NEUTRAL WORLD

There is a general result that allows us to use the analysis in Section 15.3 for futures options. This result is that in a risk-neutral world a futures price behaves in the same way as a stock paying a dividend yield at the domestic risk-free interest rate r.

One clue that this might be so is given by noting that the equation for p in a binomial tree for a futures price is the same as that for a stock paying a dividend yield equal to q when $q = r$. Another clue is that the put–call parity relationship for futures options prices is the same as that for options on a stock paying a dividend yield at rate q when the stock price is replaced by the futures price and $q = r$.

To prove the result formally, we calculate the drift of a futures price in a risk-neutral world. We define F_t as the futures price at time t. If we enter into a long futures contract today, its value is zero. At time Δt (the first time it is marked to market) it provides a payoff of $F_{\Delta t} - F_0$. If r is the very-short-term (Δt-period) interest rate at time 0, risk-neutral valuation gives the value of the contract at time 0 as

$$e^{-r\Delta t}\hat{E}[F_{\Delta t} - F_0]$$

where $\hat{E}$ denotes expectations in a risk-neutral world. We must therefore have

$$e^{-r\Delta t}\hat{E}(F_{\Delta t} - F_0) = 0$$

showing that

$$\hat{E}(F_{\Delta t}) = F_0$$

Similarly, $\hat{E}(F_{2\Delta t}) = F_{\Delta t}$, $\hat{E}(F_{3\Delta t}) = F_{2\Delta t}$, and so on. Putting many results like this together, we see that

$$\hat{E}(F_T) = F_0$$

for any time T.

The drift of the futures price in a risk-neutral world is therefore zero. From equation (15.7), the futures price behaves like a stock providing a dividend yield q equal to r. This result is a very general one. It is true for all futures prices and does not depend on any assumptions about interest rates, volatilities, etc.[3]

The usual assumption made for the process followed by a futures price F in the risk-neutral world is

$$dF = \sigma F\,dz \tag{16.7}$$

where σ is a constant.

Differential Equation

For another way of seeing that a futures price behaves like a stock paying a dividend yield at rate q, we can derive the differential equation satisfied by a derivative dependent on a futures price in the same way as we derived the differential equation for a derivative dependent on a non-dividend-paying stock in Section 13.6. This is[4]

$$\frac{\partial f}{\partial t} + \tfrac{1}{2}\frac{\partial^2 f}{\partial F^2}\sigma^2 F^2 = rf \tag{16.8}$$

It has the same form as equation (15.6) with q set equal to r. This confirms that, for the purpose of valuing derivatives, a futures price can be treated in the same way as a stock providing a dividend yield at rate r.

16.8 BLACK'S MODEL FOR VALUING FUTURES OPTIONS

European futures options can be valued by extending the results we have produced. Fischer Black was the first to show this in a paper published in 1976.[5] Assuming that the futures price follows the (lognormal) process in equation (16.7), the European call price c and the European put price p for a futures option are given by equations (15.4) and (15.5) with S_0 replaced by F_0 and $q = r$:

$$c = e^{-rT}[F_0 N(d_1) - KN(d_2)] \tag{16.9}$$

$$p = e^{-rT}[KN(-d_2) - F_0 N(-d_1)] \tag{16.10}$$

where

$$d_1 = \frac{\ln(F_0/K) + \sigma^2 T/2}{\sigma\sqrt{T}}$$

$$d_2 = \frac{\ln(F_0/K) - \sigma^2 T/2}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$$

and σ is the volatility of the futures price. When the cost of carry and the convenience

---

[3] As we will discover in Chapter 27, a more precise statement of the result is: "A futures price has zero drift in the traditional risk-neutral world where the numeraire is the money market account." A zero-drift stochastic process is known as a martingale. A forward price is a martingale in a different risk-neutral world. This is one where the numeraire is a zero-coupon bond maturing at time T.

[4] See Technical Note 7 on the author's website for a proof of this.

[5] See F. Black, "The Pricing of Commodity Contracts," *Journal of Financial Economics*, 3 (March 1976): 167–79.

yield are functions only of time, it can be shown that the volatility of the futures price is the same as the volatility of the underlying asset. Note that Black's model does not require the option contract and the futures contract to mature at the same time.

Example 16.6

Consider a European put futures option on crude oil. The time to the option's maturity is 4 months, the current futures price is \$20, the exercise price is \$20, the risk-free interest rate is 9% per annum, and the volatility of the futures price is 25% per annum. In this case, $F_0 = 20$, $K = 20$, $r = 0.09$, $T = 4/12$, $\sigma = 0.25$, and $\ln(F_0/K) = 0$, so that

$$d_1 = \frac{\sigma\sqrt{T}}{2} = 0.07216$$

$$d_2 = -\frac{\sigma\sqrt{T}}{2} = -0.07216$$

$$N(-d_1) = 0.4712, \quad N(-d_2) = 0.5288$$

and the put price p is given by

$$p = e^{-0.09\times 4/12}(20 \times 0.5288 - 20 \times 0.4712) = 1.12$$

or \$1.12.

Using Black's Model Instead of Black–Scholes

The results in Section 16.3 show that futures options and spot options are equivalent when the option contract matures at the same time as the futures contract. Equations (16.9) and (16.10) therefore provide a way of calculating the value of European options on the spot price of a asset.

Example 16.7

Consider a six-month European call option on the spot price of gold, that is, an option to buy one ounce of gold in six months. The strike price is \$600, the six-month futures price of gold is \$620, the risk-free rate of interest is 5% per annum, and the volatility of the futures price is 20%. The option is the same as a six-month European option on the six-month futures price. The value of the option is therefore given by equation (16.9) as

$$e^{-0.05\times 0.5}[620N(d_1) - 600N(d_2)]$$

where

$$d_1 = \frac{\ln(620/600) + 0.2^2 \times 0.5/2}{0.2 \times \sqrt{0.5}} = 0.3026$$

$$d_2 = \frac{\ln(620/600) - 0.2^2 \times 0.5/2}{0.2 \times \sqrt{0.5}} = 0.1611$$

It is \$44.19.

Traders like to use Black's model rather than Black–Scholes to value European options on a wide range of underlying assets. The variable F_0 in equations (16.9) and (16.10) is set equal to either the futures or the forward price of the underlying asset for a contract maturing at the same time as the option. Equations (15.13) and (15.14) show

Black's model being used to value European options on the spot value of a currency. They avoid the need to estimate the foreign risk-free interest rate explicitly. Equations (15.8) and (15.9) show Black's model being used to value European options on the spot value of an index. They avoid the need to estimate the dividend yield explicitly.

As explained in Section 15.4, Black's model can be used to imply a term structure of forward rates from actively traded index options. The forward rates can then be used to price other options on the index. The same approach can be used for other underlying assets.

16.9 AMERICAN FUTURES OPTIONS vs. AMERICAN SPOT OPTIONS

Traded futures options are in practice usually American. Assuming that the risk-free rate of interest, r, is positive, there is always some chance that it will be optimal to exercise an American futures option early. American futures options are therefore worth more than their European counterparts.

It is not generally true that an American futures option is worth the same as the corresponding American spot option when the futures and options contracts have the same maturity.[6] Suppose, for example, that there is a normal market with futures prices consistently higher than spot prices prior to maturity. This is the case with most stock indices, gold, silver, low-interest currencies, and some commodities. An American call futures option must be worth more than the corresponding American spot call option. The reason is that in some situations the futures option will be exercised early, in which case it will provide a greater profit to the holder. Similarly, an American put futures option must be worth less than the corresponding American spot put option. If there is an inverted market with futures prices consistently lower than spot prices, as is the case with high-interest currencies and some commodities, the reverse must be true. American call futures options are worth less than the corresponding American spot call option, whereas American put futures options are worth more than the corresponding American spot put option.

The differences just described between American futures options and American spot options hold true when the futures contract expires later than the options contract as well as when the two expire at the same time. In fact, the later the futures contract expires the greater the differences tend to be.

16.10 FUTURES-STYLE OPTIONS

Some exchanges trade what are termed *futures-style options*. These are futures contracts on the payoff from an option. Normally a trader who buys (sells) an option, whether on the spot price of an asset or on the futures price of an asset, pays (receives) cash up front. By contrast, traders who buy or sell a futures-style option post margin in the same way that they do on a regular futures contract (see Chapter 2). The contract is settled daily as with any other futures contract and the final settlement price is the payoff from the option. Just as a futures contract is a bet on what the future price of an

[6] The spot option "corresponding" to a futures option is defined here as one with the same strike price and the same expiration date.

asset will be, a futures-style option is a bet on what the payoff from an option will be.[7] If interest rates are constant, the futures price in a futures-style option is the same as the forward price in a forward contract on the option payoff. This shows that the futures price for a futures-style option is the price that would be paid for the option if payment were made in arrears. It is therefore the value of a regular option compounded forward at the risk-free rate.

Black's model in equations (16.9) and (16.10) gives the price of a regular European option on an asset in terms of the futures (or forward) price F_0 for a contract maturing at the same time as the option. The futures price in a call futures-style option is therefore

$$F_0N(d_1) - KN(d_2)$$

and the futures price in a put futures-style option is

$$KN(-d_2) - F_0N(-d_1)$$

where d_1 and d_2 are as defined in equations (16.9) and (16.10). These formulas are correct for a futures-style option on a futures contract and a futures-style option on the spot value of an asset. In the first case, F_0 is the current futures price for the contract underlying the option; in the second case, it is the current futures price for a futures contract on the underlying asset maturing at the same time as the option.

The put–call parity relationship for a futures-style options is

$$p + F_0 = c + K$$

An American futures-style option can be exercised early, in which case there is an immediate final settlement at the option's intrinsic value. As it turns out, it is never optimal to exercise an American futures-style options on a futures contract early because the futures price of the option is always greater than the intrinsic value. This type of American futures-style option can therefore be treated as though it were the corresponding European futures-style option.

SUMMARY

Futures options require delivery of the underlying futures contract on exercise. When a call is exercised, the holder acquires a long futures position plus a cash amount equal to the excess of the futures price over the strike price. Similarly, when a put is exercised the holder acquires a short position plus a cash amount equal to the excess of the strike price over the futures price. The futures contract that is delivered usually expires slightly later than the option.

A futures price behaves in the same way as a stock that provides a dividend yield equal to the risk-free rate, r. This means that the results produced in Chapter 15 for options on a stock paying a dividend yield apply to futures options if we replace the stock price by the futures price and set the dividend yield equal to the risk-free interest

[7] For a more detailed discussion of futures-style options, see D. Lieu, "Option Pricing with Futures-Style Margining," *Journal of Futures Markets*, 10, 4 (1990), 327–38. For pricing when interest rates are stochastic, see R.-R. Chen and L. Scott, "Pricing Interest Rate Futures Options with Futures-Style Margining." *Journal of Futures Markets*, 13, 1 (1993) 15–22).

rate. Pricing formulas for European futures options were first produced by Fischer Black in 1976. They assume that the futures price is lognormally distributed at the option's expiration.

If the expiration dates for the option and futures contracts are the same, a European futures option is worth exactly the same as the corresponding European spot option. This result is often used to value European spot options. The result is not true for American options. If the futures market is normal, an American call futures is worth more than the corresponding American spot call option, while an American put futures is worth less than the corresponding American spot put option. If the futures market is inverted, the reverse is true.

FURTHER READING

Black, F. "The Pricing of Commodity Contracts," *Journal of Financial Economics*, 3 (1976): 167–79.

Hilliard, J. E., and J. Reis. "Valuation of Commodity Futures and Options under Stochastic Convenience Yields, Interest Rates, and Jump Diffusions in the Spot," *Journal of Financial and Quantitative Analysis*, 33, 1 (March 1998): 61–86.

Miltersen, K. R., and E. S. Schwartz. "Pricing of Options on Commodity Futures with Stochastic Term Structures of Convenience Yields and Interest Rates," *Journal of Financial and Quantitative Analysis*, 33, 1 (March 1998): 33–59.

Questions and Problems (Answers in Solutions Manual)

16.1. Explain the difference between a call option on yen and a call option on yen futures.

16.2. Why are options on bond futures more actively traded than options on bonds?

16.3. "A futures price is like a stock paying a dividend yield." What is the dividend yield?

16.4. A futures price is currently 50. At the end of six months it will be either 56 or 46. The risk-free interest rate is 6% per annum. What is the value of a six-month European call option on the futures with a strike price of 50?

16.5. How does the put–call parity formula for a futures option differ from put–call parity for an option on a non-dividend-paying stock?

16.6. Consider an American futures call option where the futures contract and the option contract expire at the same time. Under what circumstances is the futures option worth more than the corresponding American option on the underlying asset?

16.7. Calculate the value of a five-month European put futures option when the futures price is \$19, the strike price is \$20, the risk-free interest rate is 12% per annum, and the volatility of the futures price is 20% per annum.

16.8. Suppose you buy a put option contract on October gold futures with a strike price of \$700 per ounce. Each contract is for the delivery of 100 ounces. What happens if you exercise when the October futures price is \$680?

16.9. Suppose you sell a call option contract on April live cattle futures with a strike price of 90 cents per pound. Each contract is for the delivery of 40,000 pounds. What happens if the contract is exercised when the futures price is 95 cents?

16.10. Consider a two-month call futures option with a strike price of 40 when the risk-free interest rate is 10% per annum. The current futures price is 47. What is a lower bound for the value of the futures option if it is (a) European and (b) American?

16.11. Consider a four-month put futures option with a strike price of 50 when the risk-free interest rate is 10% per annum. The current futures price is 47. What is a lower bound for the value of the futures option if it is (a) European and (b) American?

16.12. A futures price is currently 60 and its volatility is 30%. The risk-free interest rate is 8% per annum. Use a two-step binomial tree to calculate the value of a six-month European call option on the futures with a strike price of 60? If the call were American, would it ever be worth exercising it early?

16.13. In Problem 16.12, what does the binomial tree give for the value of a six-month European put option on futures with a strike price of 60? If the put were American, would it ever be worth exercising it early? Verify that the call prices calculated in Problem 16.12 and the put prices calculated here satisfy put–call parity relationships.

16.14. A futures price is currently 25, its volatility is 30% per annum, and the risk-free interest rate is 10% per annum. What is the value of a nine-month European call on the futures with a strike price of 26?

16.15. A futures price is currently 70, its volatility is 20% per annum, and the risk-free interest rate is 6% per annum. What is the value of a five-month European put on the futures with a strike price of 65?

16.16. Suppose that a one-year futures price is currently 35. A one-year European call option and a one-year European put option on the futures with a strike price of 34 are both priced at 2 in the market. The risk-free interest rate is 10% per annum. Identify an arbitrage opportunity.

16.17. "The price of an at-the-money European call futures option always equals the price of a similar at-the-money European put futures option." Explain why this statement is true.

16.18. Suppose that a futures price is currently 30. The risk-free interest rate is 5% per annum. A three-month American call futures option with a strike price of 28 is worth 4. Calculate bounds for the price of a three-month American put futures option with a strike price of 28.

16.19. Show that, if C is the price of an American call option on a futures contract when the strike price is K and the maturity is T, and P is the price of an American put on the same futures contract with the same strike price and exercise date, then

$$F_0e^{-rT} - K < C - P < F_0 - Ke^{-rT}$$

where F_0 is the futures price and r is the risk-free rate. Assume that $r > 0$ and that there is no difference between forward and futures contracts. (*Hint*: Use an analogous approach to that indicated for Problem 15.12.)

16.20. Calculate the price of a three-month European call option on the spot value of silver. The three-month futures price is \$12, the strike price is \$13, the risk-free rate is 4% and the volatility of the price of silver is 25%.

16.21. A corporation knows that in three months it will have \$5 million to invest for 90 days at LIBOR minus 50 basis points and wishes to ensure that the rate obtained will be at least 6.5%. What position in exchange-traded options should it take?

Assignment Questions

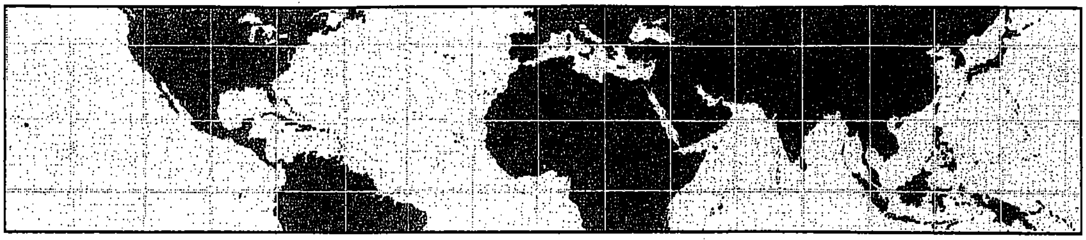
16.22. A futures price is currently 40. It is known that at the end of three months the price will be either 35 or 45. What is the value of a three-month European call option on the futures with a strike price of 42 if the risk-free interest rate is 7% per annum?

16.23. It is February 4. July call options on corn futures with strike prices of 260, 270, 280, 290, and 300 cost 26.75, 21.25, 17.25, 14.00, and 11.375, respectively. July put options with these strike prices cost 8.50, 13.50, 19.00, 25.625, and 32.625, respectively. The options mature on June 19, the current July corn futures price is 278.25, and the risk-free interest rate is 1.1%. Calculate implied volatilities for the options using DerivaGem. Comment on the results you get.

16.24. Calculate the implied volatility of soybean futures prices from the following information concerning a European put on soybean futures:

Current futures price	525
Exercise price	525
Risk-free rate	6% per annum
Time to maturity	5 months
Put price	20

16.25. Calculate the price of a six-month European put option on the spot value of the S&P 500. The six-month forward price of the index is 1,400, the strike price is 1,450, the risk-free rate is 5%, and the volatility of the index is 15%.



CHAPTER 17 The Greek Letters

A financial institution that sells an option to a client in the over-the-counter markets is faced with the problem of managing its risk. If the option happens to be the same as one that is traded on an exchange, the financial institution can neutralize its exposure by buying on the exchange the same option as it has sold. But when the option has been tailored to the needs of a client and does not correspond to the standardized products traded by exchanges, hedging the exposure is far more difficult.

In this chapter we discuss some of the alternative approaches to this problem. We cover what are commonly referred to as the "Greek letters", or simply the "Greeks". Each Greek letter measures a different dimension to the risk in an option position and the aim of a trader is to manage the Greeks so that all risks are acceptable. The analysis presented in this chapter is applicable to market makers in options on an exchange as well as to traders working in the over-the-counter market for financial institutions.

Toward the end of the chapter, we will consider the creation of options synthetically. This turns out to be very closely related to the hedging of options. Creating an option position synthetically is essentially the same task as hedging the opposite option position. For example, creating a long call option synthetically is the same as hedging a short position in the call option.

17.1 ILLUSTRATION

In the next few sections we use as an example the position of a financial institution that has sold for \$300,000 a European call option on 100,000 shares of a non-dividend-paying stock. We assume that the stock price is \$49, the strike price is \$50, the risk-free interest rate is 5% per annum, the stock price volatility is 20% per annum, the time to maturity is 20 weeks (0.3846 years), and the expected return from the stock is 13% per annum.[1] With our usual notation, this means that

$$S_0 = 49, \quad K = 50, \quad r = 0.05, \quad \sigma = 0.20, \quad T = 0.3846, \quad \mu = 0.13$$

The Black–Scholes price of the option is about \$240,000. The financial institution has

[1] As shown in Chapters 11 and 13, the expected return is irrelevant to the pricing of an option. It is given here because it can have some bearing on the effectiveness of a hedging scheme.

therefore sold the option for $60,000 more than its theoretical value, but it is faced with the problem of hedging the risks.[2]

17.2 NAKED AND COVERED POSITIONS

One strategy open to the financial institution is to do nothing. This is sometimes referred to as a *naked position*. It is a strategy that works well if the stock price is below $50 at the end of the 20 weeks. The option then costs the financial institution nothing and it makes a profit of $300,000. A naked position works less well if the call is exercised because the financial institution then has to buy 100,000 shares at the market price prevailing in 20 weeks to cover the call. The cost to the financial institution is 100,000 times the amount by which the stock price exceeds the strike price. For example, if after 20 weeks the stock price is $60, the option costs the financial institution $1,000,000. This is considerably greater than the $300,000 charged for the option.

As an alternative to a naked position, the financial institution can adopt a *covered position*. This involves buying 100,000 shares as soon as the option has been sold. If the option is exercised, this strategy works well, but in other circumstances it could lead to a significant loss. For example, if the stock price drops to $40, the financial institution loses $900,000 on its stock position. This is considerably greater than the $300,000 charged for the option.[3]

Neither a naked position nor a covered position provides a good hedge. If the assumptions underlying the Black–Scholes formula hold, the cost to the financial institution should always be $240,000 on average for both approaches.[4] But on any one occasion the cost is liable to range from zero to over $1,000,000. A good hedge would ensure that the cost is always close to $240,000.

17.3 A STOP-LOSS STRATEGY

One interesting hedging procedure that is sometimes proposed involves a *stop-loss strategy*. To illustrate the basic idea, consider an institution that has written a call option with strike price K to buy one unit of a stock. The hedging procedure involves buying one unit of the stock as soon as its price rises above K and selling it as soon as its price falls below K. The objective is to hold a naked position whenever the stock price is less than K and a covered position whenever the stock price is greater than K. The procedure is designed to ensure that at time T the institution owns the stock if the option closes in the money and does not own it if the option closes out of the money. The strategy appears to produce payoffs that are the same as the payoffs on the option. In the situation illustrated in Figure 17.1, it involves buying the stock at time t_1, selling it at time t_2, buying it at time t_3, selling it at time t_4, buying it at time t_5, and delivering it at time T.

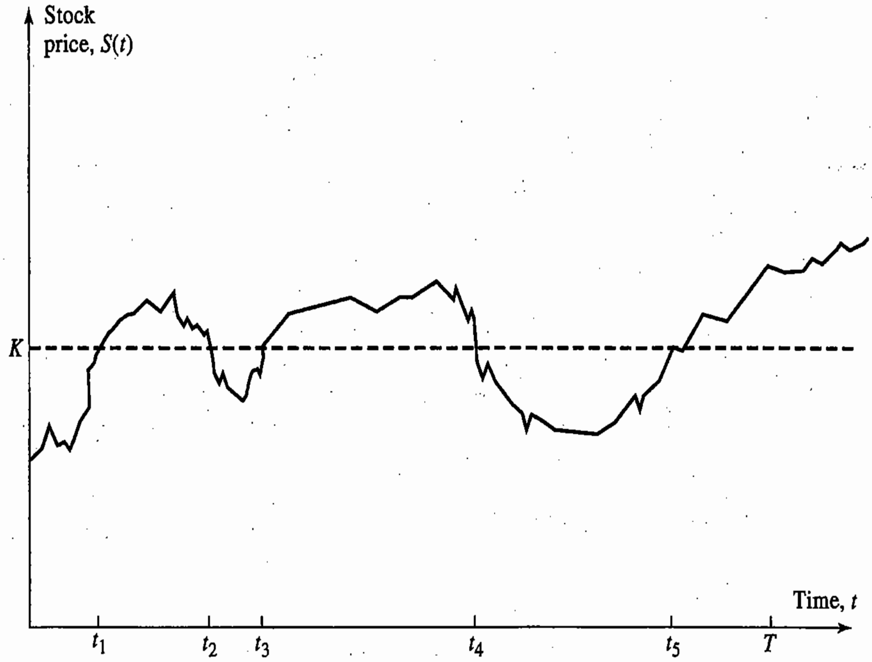
---

[2] A call option on a non-dividend-paying stock is a convenient example with which to develop our ideas. The points that will be made apply to other types of options and to other derivatives.

[3] Put–call parity shows that the exposure from writing a covered call is the same as the exposure from writing a naked put.

[4] More precisely, the present value of the expected cost is $240,000 for both approaches assuming that appropriate risk-adjusted discount rates are used.

Figure 17.1 A stop-loss strategy.



As usual, we denote the initial stock price by S_0. The cost of setting up the hedge initially is S_0 if $S_0 > K$ and zero otherwise. It seems as though the total cost, Q, of writing and hedging the option is the option's intrinsic value:

$$Q = \max(S_0 - K, 0) \tag{17.1}$$

This is because all purchases and sales subsequent to time 0 are made at price K. If this were in fact correct, the hedging scheme would work perfectly in the absence of transactions costs. Furthermore, the cost of hedging the option would always be less than its Black–Scholes price. Thus, an investor could earn riskless profits by writing options and hedging them.

There are two basic reasons why equation (17.1) is incorrect. The first is that the cash flows to the hedger occur at different times and must be discounted. The second is that purchases and sales cannot be made at exactly the same price K. This second point is critical. If we assume a risk-neutral world with zero interest rates, we can justify ignoring the time value of money. But we cannot legitimately assume that both purchases and sales are made at the same price. If markets are efficient, the hedger cannot know whether, when the stock price equals K, it will continue above or below K.

As a practical matter, purchases must be made at a price $K + \epsilon$ and sales must be made at a price $K - \epsilon$, for some small positive number ϵ. Thus, every purchase and subsequent sale involves a cost (apart from transaction costs) of 2ϵ. A natural response on the part of the hedger is to monitor price movements more closely, so that ϵ is reduced. Assuming that stock prices change continuously, ϵ can be made arbitrarily small by monitoring the stock prices closely. But as ϵ is made smaller, trades tend to occur more frequently. Thus, the lower cost per trade is offset by the

Table 17.1 Performance of stop-loss strategy. The performance measure is the ratio of the standard deviation of the cost of writing the option and hedging it to the theoretical price of the option.

Δt (weeks)	5	4	2	1	0.5	0.25
Hedge performance	1.02	0.93	0.82	0.77	0.76	0.76

increased frequency of trading. As $\epsilon \to 0$, the expected number of trades tends to infinity.[5]

A stop-loss strategy, although superficially attractive, does not work particularly well as a hedging scheme. Consider its use for an out-of-the-money option. If the stock price never reaches the strike price K, the hedging scheme costs nothing. If the path of the stock price crosses the strike price level many times, the scheme is quite expensive. Monte Carlo simulation can be used to assess the overall performance of stop-loss hedging. This involves randomly sampling paths for the stock price and observing the results of using the scheme. Table 17.1 shows the results for the option considered earlier. It assumes that the stock price is observed at the end of time intervals of length Δt.[6] The hedge performance measure is the ratio of the standard deviation of the cost of hedging the option to the Black–Scholes option price. Each result is based on 1,000 sample paths for the stock price and has a standard error of about 2%. A perfect hedge would have a hedge performance measure of zero. In this case it appears to be impossible to produce a value for the hedge performance measure below 0.70 regardless of how small Δt is made.

17.4 DELTA HEDGING

Most traders use more sophisticated hedging schemes than those mentioned so far. These involve calculating measures such as delta, gamma, and vega. In this section we consider the role played by delta.

The *delta* (Δ) of an option was introduced in Chapter 11. It is defined as the rate of change of the option price with respect to the price of the underlying asset. It is the slope of the curve that relates the option price to the underlying asset price. Suppose that the delta of a call option on a stock is 0.6. This means that when the stock price changes by a small amount, the option price changes by about 60% of that amount. Figure 17.2 shows the relationship between a call price and the underlying stock price. When the stock price corresponds to point A, the option price corresponds to point B, and Δ is the slope of the line indicated. In general,

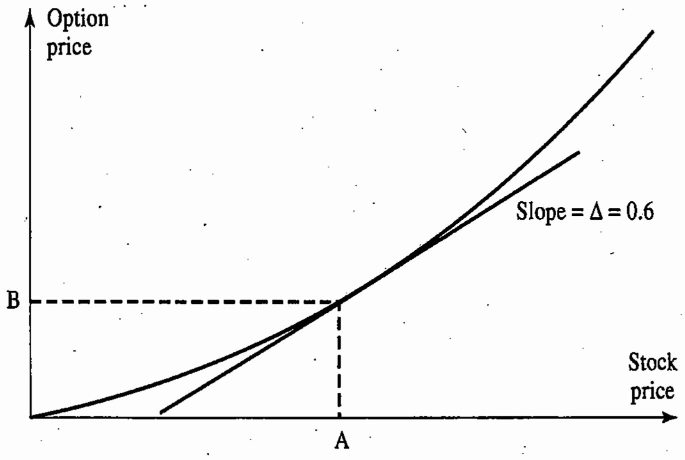
$$\Delta = \frac{\partial c}{\partial S}$$

where c is the price of the call option and S is the stock price.

[5] As mentioned in Section 12.2, the expected number of times a Wiener process equals any particular value in a given time interval is infinite.

[6] The precise hedging rule used was as follows. If the stock price moves from below K to above K in a time interval of length Δt, it is bought at the end of the interval. If it moves from above K to below K in the time interval, it is sold at the end of the interval; otherwise, no action is taken.

Figure 17.2 Calculation of delta.



Suppose that, in Figure 17.2, the stock price is \$100 and the option price is \$10. Imagine an investor who has sold 20 call option contracts—that is, options to buy 2,000 shares. The investor's position could be hedged by buying $0.6 \times 2{,}000 = 1{,}200$ shares. The gain (loss) on the option position would then tend to be offset by the loss (gain) on the stock position. For example, if the stock price goes up by \$1 (producing a gain of \$1,200 on the shares purchased), the option price will tend to go up by $0.6 \times \$1 = \$0.60$ (producing a loss of \$1,200 on the options written); if the stock price goes down by \$1 (producing a loss of \$1,200 on the shares purchased), the option price will tend to go down by \$0.60 (producing a gain of \$1,200 on the options written).

In this example, the delta of the investor's option position is

$$0.6 \times (-2{,}000) = -1{,}200$$

In other words, the investor loses $1{,}200\Delta S$ on the short option position when the stock price increases by ΔS. The delta of the stock is 1.0, so that the long position in 1,200 shares has a delta of +1,200. The delta of the investor's overall position is, therefore, zero. The delta of the stock position offsets the delta of the option position. A position with a delta of zero is referred to as being *delta neutral.*

It is important to realize that, because delta changes, the investor's position remains delta hedged (or delta neutral) for only a relatively short period of time. The hedge has to be adjusted periodically. This is known as *rebalancing*. In our example, at the end of 3 days the stock price might increase to \$110. As indicated by Figure 17.2, an increase in the stock price leads to an increase in delta. Suppose that delta rises from 0.60 to 0.65. An extra $0.05 \times 2{,}000 = 100$ shares would then have to be purchased to maintain the hedge.

The delta-hedging procedure just described is an example of *dynamic hedging*. It can be contrasted with *static hedging*, where the hedge is set up initially and never adjusted. Static hedging is sometimes also referred to as *hedge-and-forget*. Delta is closely related to the Black–Scholes–Merton analysis. As explained in Chapter 13, Black, Scholes, and Merton showed that it is possible to set up a riskless portfolio consisting of a position in an option on a stock and a position in the stock. Expressed in terms of Δ, the Black–Scholes portfolio is

-1: option
$+\Delta$: shares of the stock

Using our new terminology, we can say that Black and Scholes valued options by setting up a delta-neutral position and arguing that the return on the position should be the risk-free interest rate.

Delta of European Stock Options

For a European call option on a non-dividend-paying stock, it can be shown (see Problem 13.17) that

$$\Delta(\text{call}) = N(d_1)$$

where d_1 is defined as in equation (13.20). The formula gives the delta of a long position in one call option. The delta of a short position in one call option is $-N(d_1)$. Using delta hedging for a short position in a European call option involves maintaining a long position of $N(d_1)$ for each option sold. Similarly, using delta hedging for a long position in a European call option involves maintaining a short position of $N(d_1)$ shares for each option purchased.

For a European put option on a non-dividend-paying stock, delta is given by

$$\Delta(\text{put}) = N(d_1) - 1$$

Delta is negative, which means that a long position in a put option should be hedged with a long position in the underlying stock, and a short position in a put option should be hedged with a short position in the underlying stock. Figure 17.3 shows the variation of the delta of a call option and a put option with the stock price. Figure 17.4 shows the variation of delta with the time to maturity for in-the-money, at-the-money, and out-of-the-money call options.

Figure 17.3 Variation of delta with stock price for (a) a call option and (b) a put option on a non-dividend-paying stock.

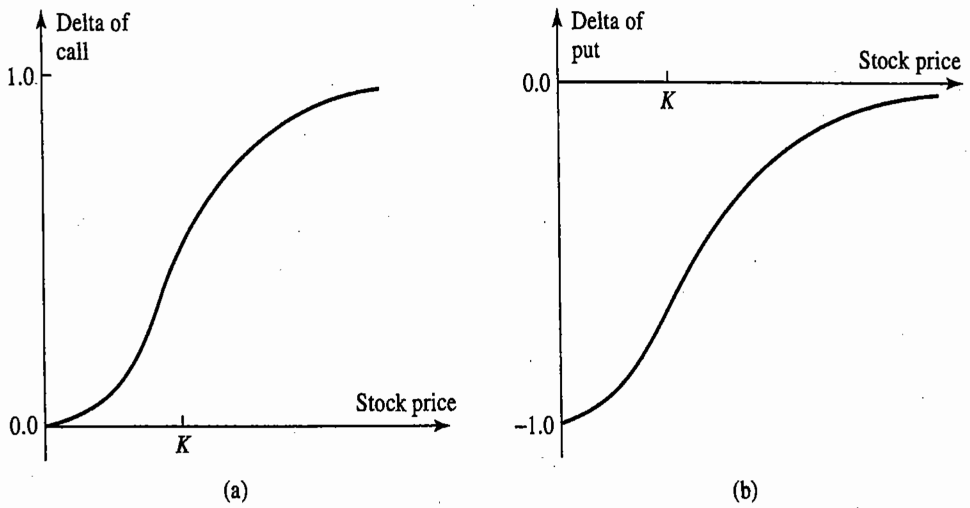
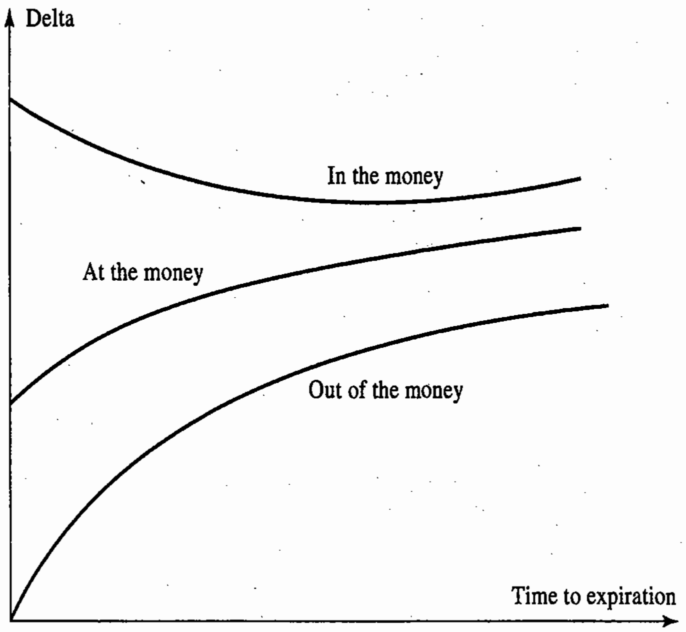


Figure 17.4 Typical patterns for variation of delta with time to maturity for a call option.



Example 17.1

Consider again the call option on a non-dividend-paying stock in Section 17.1 where the stock price is \$49, the strike price is \$50, the risk-free rate is 5%, the time to maturity is 20 weeks (= 0.3846 years), and the volatility is 20%. In this case,

$$d_1 = \frac{\ln(49/50) + (0.05 + 0.2^2/2) \times 0.3846}{0.2 \times \sqrt{0.3836}} = 0.0542$$

Delta is $N(d_1)$, or 0.522. When the stock price changes by ΔS, the option price changes by $0.522\Delta S$.

Dynamic Aspects of Delta Hedging

Tables 17.2 and 17.3 provide two examples of the operation of delta hedging for the example in Section 17.1. The hedge is assumed to be adjusted or rebalanced weekly. The initial value of delta for the option being sold is calculated in Example 17.1 as 0.522. This means that the delta of the short option position is initially −52,200. As soon as the option is written, \$2,557,800 must be borrowed to buy 52,200 shares at a price of \$49. The rate of interest is 5%. An interest cost of approximately \$2,500 is therefore incurred in the first week.

In Table 17.2 the stock price falls by the end of the first week to \$48.12. The delta of the option declines to 0.458, so that the new delta of the option position is −45,800. This means that 6,400 of the shares initially purchased are sold to maintain the hedge. The strategy realizes \$308,000 in cash, and the cumulative borrowings at the end of Week 1 are reduced to \$2,252,300. During the second week, the stock price reduces to \$47.37, delta declines again, and so on. Toward the end of the life of the option, it

Table 17.2 Simulation of delta hedging. Option closes in the money and cost of hedging is $263,300.

Week	*Stock price*	*Delta*	*Shares purchased*	*Cost of shares purchased ($000)*	*Cumulative cost including interest ($000)*	*Interest cost ($000)*
0	49.00	0.522	52,200	2,557.8	2,557.8	2.5
1	48.12	0.458	(6,400)	(308.0)	2,252.3	2.2
2	47.37	0.400	(5,800)	(274.7)	1,979.8	1.9
3	50.25	0.596	19,600	984.9	2,966.6	2.9
4	51.75	0.693	9,700	502.0	3,471.5	3.3
5	53.12	0.774	8,100	430.3	3,905.1	3.8
6	53.00	0.771	(300)	(15.9)	3,893.0	3.7
7	51.87	0.706	(6,500)	(337.2)	3,559.5	3.4
8	51.38	0.674	(3,200)	(164.4)	3,398.5	3.3
9	53.00	0.787	11,300	598.9	4,000.7	3.8
10	49.88	0.550	(23,700)	(1,182.2)	2,822.3	2.7
11	48.50	0.413	(13,700)	(664.4)	2,160.6	2.1
12	49.88	0.542	12,900	643.5	2,806.2	2.7
13	50.37	0.591	4,900	246.8	3,055.7	2.9
14	52.13	0.768	17,700	922.7	3,981.3	3.8
15	51.88	0.759	(900)	(46.7)	3,938.4	3.8
16	52.87	0.865	10,600	560.4	4,502.6	4.3
17	54.87	0.978	11,300	620.0	5,126.9	4.9
18	54.62	0.990	1,200	65.5	5,197.3	5.0
19	55.87	1.000	1,000	55.9	5,258.2	5.1
20	57.25	1.000	0	0.0	5,263.3	

becomes apparent that the option will be exercised and the delta of the option approaches 1.0. By Week 20, therefore, the hedger has a fully covered position. The hedger receives $5 million for the stock held, so that the total cost of writing the option and hedging it is $263,300.

Table 17.3 illustrates an alternative sequence of events such that the option closes out of the money. As it becomes clear that the option will not be exercised, delta approaches zero. By Week 20 the hedger has a naked position and has incurred costs totaling $256,600.

In Tables 17.2 and 17.3, the costs of hedging the option, when discounted to the beginning of the period, are close to but not exactly the same as the Black–Scholes price of $240,000. If the hedging scheme worked perfectly, the cost of hedging would, after discounting, be exactly equal to the Black–Scholes price for every simulated stock price path. The reason for the variation in the cost of hedging is that the hedge is rebalanced only once a week. As rebalancing takes place more frequently, the variation in the cost of hedging is reduced. Of course, the examples in Tables 17.2 and 17.3 are idealized in that they assume that the volatility is constant and there are no transaction costs.

Table 17.3 Simulation of delta hedging. Option closes out of the money and cost of hedging is $256,600.

Week	*Stock price*	*Delta*	*Shares purchased*	*Cost of shares purchased ($000)*	*Cumulative cost including interest ($000)*	*Interest cost ($000)*
0	49.00	0.522	52,200	2,557.8	2,557.8	2.5
1	49.75	0.568	4,600	228.9	2,789.2	2.7
2	52.00	0.705	13,700	712.4	3,504.3	3.4
3	50.00	0.579	(12,600)	(630.0)	2,877.7	2.8
4	48.38	0.459	(12,000)	(580.6)	2,299.9	2.2
5	48.25	0.443	(1,600)	(77.2)	2,224.9	2.1
6	48.75	0.475	3,200	156.0	2,383.0	2.3
7	49.63	0.540	6,500	322.6	2,707.9	2.6
8	48.25	0.420	(12,000)	(579.0)	2,131.5	2.1
9	48.25	0.410	(1,000)	(48.2)	2,085.4	2.0
10	51.12	0.658	24,800	1,267.8	3,355.2	3.2
11	51.50	0.692	3,400	175.1	3,533.5	3.4
12	49.88	0.542	(15,000)	(748.2)	2,788.7	2.7
13	49.88	0.538	(400)	(20.0)	2,771.4	2.7
14	48.75	0.400	(13,800)	(672.7)	2,101.4	2.0
15	47.50	0.236	(16,400)	(779.0)	1,324.4	1.3
16	48.00	0.261	2,500	120.0	1,445.7	1.4
17	46.25	0.062	(19,900)	(920.4)	526.7	0.5
18	48.13	0.183	12,100	582.4	1,109.6	1.1
19	46.63	0.007	(17,600)	(820.7)	290.0	0.3
20	48.12	0.000	(700)	(33.7)	256.6	

Table 17.4 shows statistics on the performance of delta hedging obtained from 1,000 random stock price paths in our example. As in Table 17.1, the performance measure is the ratio of the standard deviation of the cost of hedging the option to the Black–Scholes price of the option. It is clear that delta hedging is a great improvement over a stop-loss strategy. Unlike a stop-loss strategy, the performance of a delta-hedging strategy gets steadily better as the hedge is monitored more frequently.

Table 17.4 Performance of delta hedging. The performance measure is the ratio of the standard deviation of the cost of writing the option and hedging it to the theoretical price of the option.

Time between hedge rebalancing (weeks):	5	4	2	1	0.5	0.25
Performance measure:	0.43	0.39	0.26	0.19	0.14	0.09

Delta hedging aims to keep the value of the financial institution's position as close to unchanged as possible. Initially, the value of the written option is \$240,000. In the situation depicted in Table 17.2, the value of the option can be calculated as \$414,500 in Week 9. Thus, the financial institution has lost \$174,500 on its short option position. Its cash position, as measured by the cumulative cost, is \$1,442,900 worse in Week 9 than in Week 0. The value of the shares held has increased from \$2,557,800 to \$4,171,100. The net effect of all this is that the value of the financial institution's position has changed by only \$4,100 between Week 0 and Week 9.

Where the Cost Comes From

The delta-hedging procedure in Tables 17.2 and 17.3 creates the equivalent of a long position in the option. This neutralizes the short position the financial institution created by writing the option. As the tables illustrate, delta hedging a short position generally involves selling stock just after the price has gone down and buying stock just after the price has gone up. It might be termed a buy-high, sell-low trading strategy! The cost of \$240,000 comes from the average difference between the price paid for the stock and the price realized for it.

Delta of a Portfolio

The delta of a portfolio of options or other derivatives dependent on a single asset whose price is S is

$$\frac{\partial \Pi}{\partial S}$$

where Π is the value of the portfolio.

The delta of the portfolio can be calculated from the deltas of the individual options in the portfolio. If a portfolio consists of a quantity w_i of option i ($1 \leqslant i \leqslant n$), the delta of the portfolio is given by

$$\Delta = \sum_{i=1}^{n} w_i \, \Delta_i$$

where Δ_i is the delta of the ith option. The formula can be used to calculate the position in the underlying asset necessary to make the delta of the portfolio zero. When this position has been taken, the portfolio is referred to as being *delta neutral.*

Suppose a financial institution has the following three positions in options on a stock:

1. A long position in 100,000 call options with strike price \$55 and an expiration date in 3 months. The delta of each option is 0.533.
2. A short position in 200,000 call options with strike price \$56 and an expiration date in 5 months. The delta of each option is 0.468.
3. A short position in 50,000 put options with strike price \$56 and an expiration date in 2 months. The delta of each option is −0.508.

The delta of the whole portfolio is

$$100{,}000 \times 0.533 - 200{,}000 \times 0.468 - 50{,}000 \times (-0.508) = -14{,}900$$

This means that the portfolio can be made delta neutral by buying 14,900 shares.

Transactions Costs

Derivatives dealers usually rebalance their positions once a day to maintain delta neutrality. When the dealer has a small number of options on a particular asset, this is liable to be prohibitively expensive because of the transactions costs incurred on trades. For a large portfolio of options, it is more feasible. Only one trade in the underlying asset is necessary to zero out delta for the whole portfolio. The hedging transactions costs are absorbed by the profits on many different trades.

17.5 THETA

The *theta* (Θ) of a portfolio of options is the rate of change of the value of the portfolio with respect to the passage of time with all else remaining the same. Theta is sometimes referred to as the *time decay* of the portfolio. For a European call option on a non-dividend-paying stock, it can be shown from the Black–Scholes formula that

$$\Theta(\text{call}) = -\frac{S_0 N'(d_1)\sigma}{2\sqrt{T}} - rKe^{-rT}N(d_2)$$

where d_1 and d_2 are defined as in equation (13.20) and

$$N'(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2} \tag{17.2}$$

is the probability density function for a standard normal distribution.

For a European put option on the stock (see Problem 13.17),

$$\Theta(\text{put}) = -\frac{S_0 N'(d_1)\sigma}{2\sqrt{T}} + rKe^{-rT}N(-d_2)$$

Because $N(-d_2) = 1 - N(d_2)$, the theta of a put exceeds the theta of the corresponding call by rKe^{-rT}.

In these formulas, time is measured in years. Usually, when theta is quoted, time is measured in days, so that theta is the change in the portfolio value when 1 day passes with all else remaining the same. We can measure theta either "per calendar day" or "per trading day". To obtain the theta per calendar day, the formula for theta must be divided by 365; to obtain theta per trading day, it must be divided by 252. (DerivaGem measures theta per calendar day.)

Example 17.2

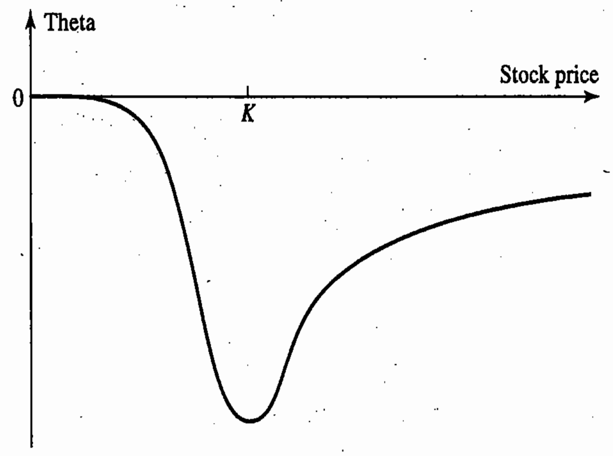
As in Example 17.1, consider a call option on a non-dividend-paying stock where the stock price is \$49, the strike price is \$50, the risk-free rate is 5%, the time to maturity is 20 weeks (= 0.3846 years), and the volatility is 20%. In this case, $S_0 = 49$, $K = 50$, $r = 0.05$, $\sigma = 0.2$, and $T = 0.3846$.

The option's theta is

$$-\frac{S_0 N'(d_1)\sigma}{2\sqrt{T}} - rKe^{-rT}N(d_2) = -4.31$$

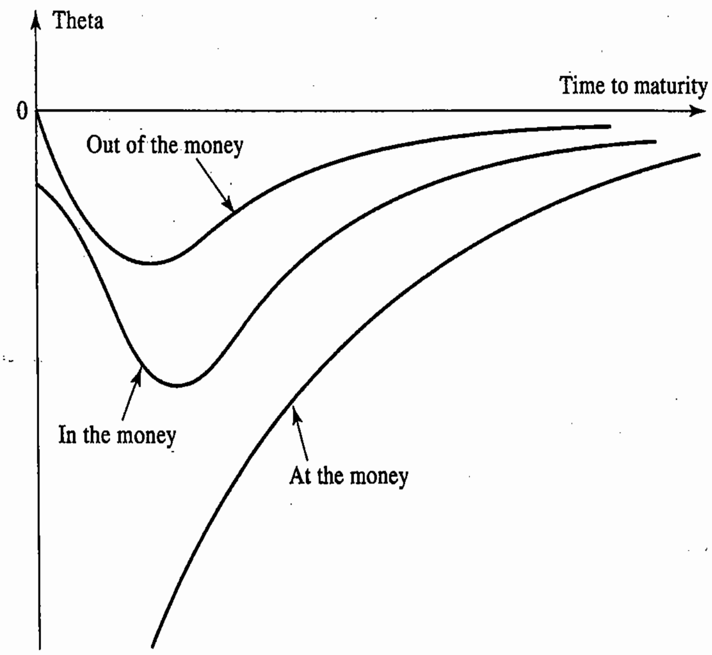
The theta is $-4.31/365 = -0.0118$ per calendar day, or $-4.31/252 = -0.0171$ per trading day.

Figure 17.5 Variation of theta of a European call option with stock price.



Theta is usually negative for an option.[7] This is because, as time passes with all else remaining the same, the option tends to become less valuable. The variation of Θ with stock price for a call option on a stock is shown in Figure 17.5. When the stock price is very low, theta is close to zero. For an at-the-money call option, theta is large and negative. As the stock price becomes larger, theta tends to $-rKe^{-rT}$. Figure 17.6 shows typical patterns for the variation of Θ with the time to maturity for in-the-money, at-the-money, and out-of-the-money call options.

Figure 17.6 Typical patterns for variation of theta of a European call option with time to maturity.



[7] An exception to this could be an in-the-money European put option on a non-dividend-paying stock or an in-the-money European call option on a currency with a very high interest rate.

Theta is not the same type of hedge parameter as delta. There is uncertainty about the future stock price, but there is no uncertainty about the passage of time. It makes sense to hedge against changes in the price of the underlying asset, but it does not make any sense to hedge against the passage of time. In spite of this, many traders regard theta as a useful descriptive statistic for a portfolio. This is because, as we shall see later, in a delta-neutral portfolio theta is a proxy for gamma.

17.6 GAMMA

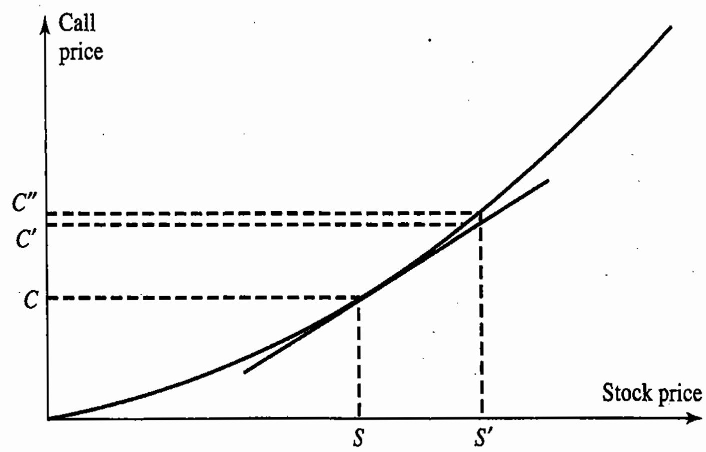
The *gamma* (Γ) of a portfolio of options on an underlying asset is the rate of change of the portfolio's delta with respect to the price of the underlying asset. It is the second partial derivative of the portfolio with respect to asset price:

$$\Gamma = \frac{\partial^2 \Pi}{\partial S^2}$$

If gamma is small, delta changes slowly, and adjustments to keep a portfolio delta neutral need to be made only relatively infrequently. However, if the absolute value of gamma is large, delta is highly sensitive to the price of the underlying asset. It is then quite risky to leave a delta-neutral portfolio unchanged for any length of time. Figure 17.7 illustrates this point. When the stock price moves from S to S', delta hedging assumes that the option price moves from C to C', when in fact it moves from C to C''. The difference between C' and C'' leads to a hedging error. The size of the error depends on the curvature of the relationship between the option price and the stock price. Gamma measures this curvature.[8]

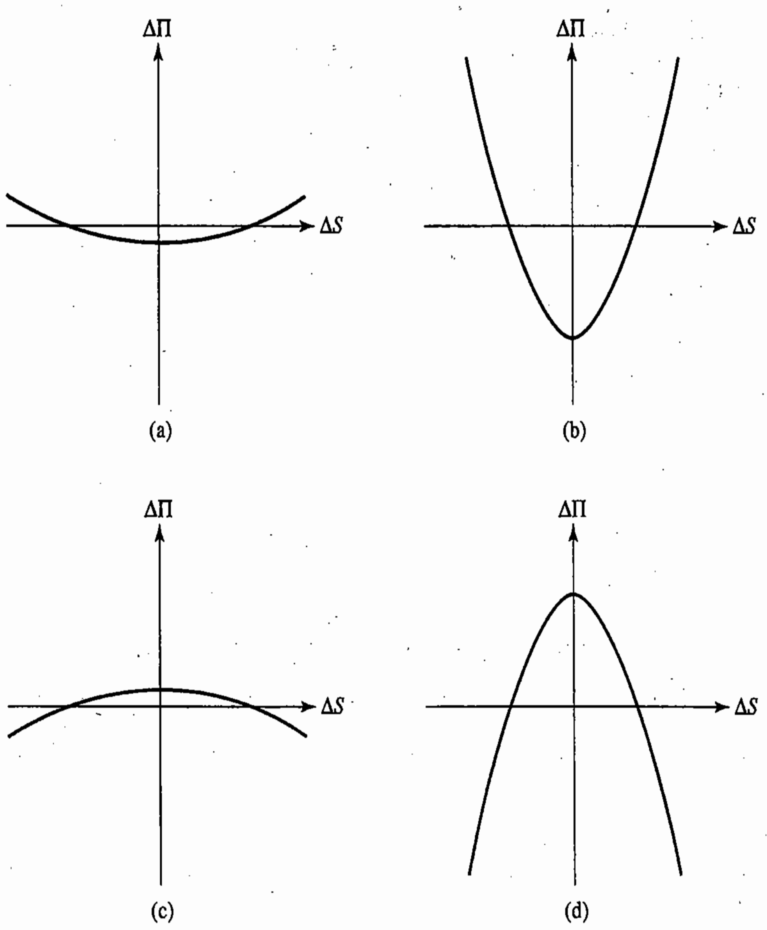
Suppose that ΔS is the price change of an underlying asset during a small interval of time, Δt, and $\Delta \Pi$ is the corresponding price change in the portfolio. The appendix at

Figure 17.7 Hedging error introduced by nonlinearity.



[8] Indeed, the gamma of an option is sometimes referred to as its *curvature* by practitioners.

Figure 17.8 Relationship between $\Delta\Pi$ and ΔS in time Δt for a delta-neutral portfolio with (a) slightly positive gamma, (b) large positive gamma, (c) slightly negative gamma, and (d) large negative gamma.



the end of this chapter shows that, if terms of order higher than Δt are ignored,

$$\Delta\Pi = \Theta\,\Delta t + \tfrac{1}{2}\Gamma\,\Delta S^2 \tag{17.3}$$

for a delta-neutral portfolio, where Θ is the theta of the portfolio. Figure 17.8 shows the nature of this relationship between $\Delta\Pi$ and ΔS. When gamma is positive, theta tends to be negative. The portfolio declines in value if there is no change in S, but increases in value if there is a large positive or negative change in S. When gamma is negative, theta tends to be positive and the reverse is true; the portfolio increases in value if there is no change in S but decreases in value if there is a large positive or negative change in S. As the absolute value of gamma increases, the sensitivity of the value of the portfolio to S increases.

Example 17.3

Suppose that the gamma of a delta-neutral portfolio of options on an asset is $-10{,}000$. Equation (17.3) shows that, if a change of $+2$ or -2 in the price of the asset occurs over a short period of time, there is an unexpected decrease in the value of the portfolio of approximately $0.5 \times 10{,}000 \times 2^2 = \$20{,}000$.

Making a Portfolio Gamma Neutral

A position in the underlying asset has zero gamma and cannot be used to change the gamma of a portfolio. What is required is a position in an instrument such as an option that is not linearly dependent on the underlying asset.

Suppose that a delta-neutral portfolio has a gamma equal to Γ, and a traded option has a gamma equal to Γ_T. If the number of traded options added to the portfolio is w_T, the gamma of the portfolio is

$$w_T \Gamma_T + \Gamma$$

Hence, the position in the traded option necessary to make the portfolio gamma neutral is $-\Gamma/\Gamma_T$. Including the traded option is likely to change the delta of the portfolio, so the position in the underlying asset then has to be changed to maintain delta neutrality. Note that the portfolio is gamma neutral only for a short period of time. As time passes, gamma neutrality can be maintained only if the position in the traded option is adjusted so that it is always equal to $-\Gamma/\Gamma_T$.

Making a portfolio gamma neutral as well as delta-neutral can be regarded as a correction for the hedging error illustrated in Figure 17.7. Delta neutrality provides protection against relatively small stock price moves between rebalancing. Gamma neutrality provides protection against larger movements in this stock price between hedge rebalancing. Suppose that a portfolio is delta neutral and has a gamma of −3,000. The delta and gamma of a particular traded call option are 0.62 and 1.50, respectively. The portfolio can be made gamma neutral by including in the portfolio a long position of

$$\frac{3{,}000}{1.5} = 2{,}000$$

in the call option. However, the delta of the portfolio will then change from zero to $2{,}000 \times 0.62 = 1{,}240$. Therefore 1,240 units of the underlying asset must be sold from the portfolio to keep it delta neutral.

Calculation of Gamma

For a European call or put option on a non-dividend-paying stock, the gamma is given by

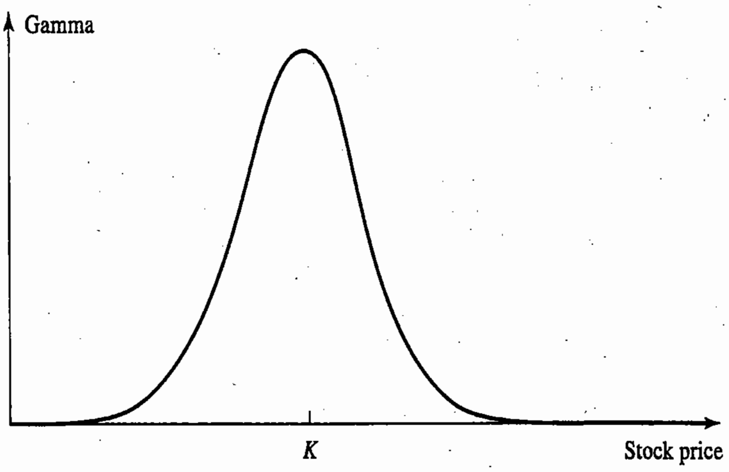
$$\Gamma = \frac{N'(d_1)}{S_0 \sigma \sqrt{T}}$$

where d_1 is defined as in equation (13.20) and $N'(x)$ is as given by equation (17.2). The gamma of a long position is always positive and varies with S_0 in the way indicated in Figure 17.9. The variation of gamma with time to maturity for out-of-the-money, at-the-money, and in-the-money options is shown in Figure 17.10. For an at-the-money option, gamma increases as the time to maturity decreases. Short-life at-the-money options have very high gammas, which means that the value of the option holder's position is highly sensitive to jumps in the stock price.

Example 17.4

As in Example 17.1, consider a call option on a non-dividend-paying stock where the stock price is \$49, the strike price is \$50, the risk-free rate is 5%, the time to

Figure 17.9 Variation of gamma with stock price for an option.



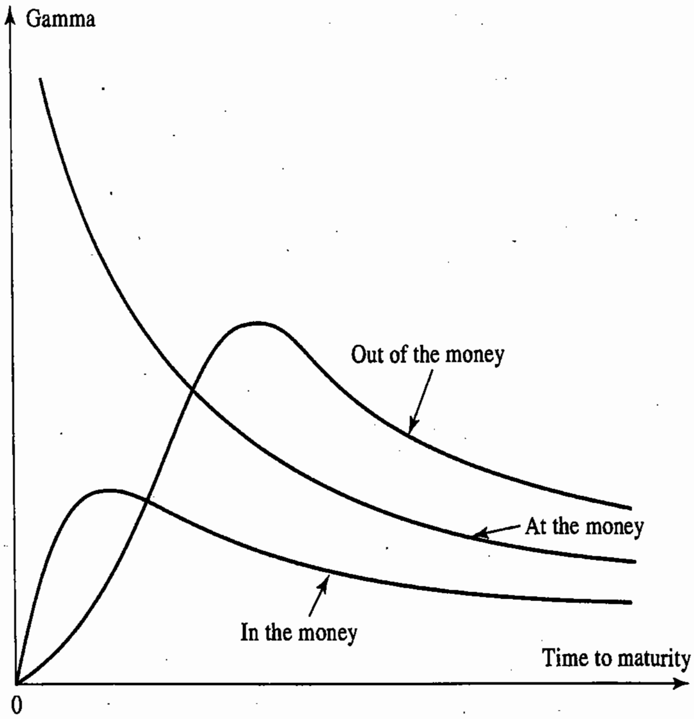
maturity is 20 weeks (= 0.3846 years), and the volatility is 20%. In this case, $S_0 = 49$, $K = 50$, $r = 0.05$, $\sigma = 0.2$, and $T = 0.3846$.

The option's gamma is

$$\frac{N'(d_1)}{S_0\sigma\sqrt{T}} = 0.066$$

When the stock price changes by ΔS, the delta of the option changes by $0.066\Delta S$.

Figure 17.10 Variation of gamma with time to maturity for a stock option.



17.7 RELATIONSHIP BETWEEN DELTA, THETA, AND GAMMA

The price of a single derivative dependent on a non-dividend-paying stock must satisfy the differential equation (13.16). It follows that the value of Π of a portfolio of such derivatives also satisfies the differential equation

$$\frac{\partial \Pi}{\partial t} + rS\frac{\partial \Pi}{\partial S} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 \Pi}{\partial S^2} = r\Pi$$

Since

$$\Theta = \frac{\partial \Pi}{\partial t}, \qquad \Delta = \frac{\partial \Pi}{\partial S}, \qquad \Gamma = \frac{\partial^2 \Pi}{\partial S^2}$$

it follows that

$$\Theta + rS\Delta + \tfrac{1}{2}\sigma^2 S^2 \Gamma = r\Pi \tag{17.4}$$

Similar results can be produced for other underlying assets (see Problem 17.19).

For a delta-neutral portfolio, $\Delta = 0$ and

$$\Theta + \tfrac{1}{2}\sigma^2 S^2 \Gamma = r\Pi$$

This shows that, when Θ is large and positive, gamma of a portfolio tends to be large and negative, and vice versa. This is consistent with the way in which Figure 17.8 has been drawn and explains why theta can to some extent be regarded as a proxy for gamma in a delta-neutral portfolio.

17.8 VEGA

Up to now we have implicitly assumed that the volatility of the asset underlying a derivative is constant. In practice, volatilities change over time. This means that the value of a derivative is liable to change because of movements in volatility as well as because of changes in the asset price and the passage of time.

The *vega* of a portfolio of derivatives, $\mathcal{V}$, is the rate of change of the value of the portfolio with respect to the volatility of the underlying asset.[9]

$$\mathcal{V} = \frac{\partial \Pi}{\partial \sigma}$$

If the absolute value of vega is high, the portfolio's value is very sensitive to small changes in volatility. If the absolute value of vega is low, volatility changes have relatively little impact on the value of the portfolio.

A position in the underlying asset has zero vega. However, the vega of a portfolio can be changed by adding a position in a traded option. If $\mathcal{V}$ is the vega of the portfolio and $\mathcal{V}_T$ is the vega of a traded option, a position of $-\mathcal{V}/\mathcal{V}_T$ in the traded option makes the portfolio instantaneously vega neutral. Unfortunately, a portfolio that is gamma neutral will not in general be vega neutral, and vice versa. If a hedger requires a portfolio to be both gamma and vega neutral, at least two traded derivatives dependent on the underlying asset must usually be used.

[9] Vega is the name given to one of the "Greek letters" in option pricing, but it is not one of the letters in the Greek alphabet.

Example 17.5

Consider a portfolio that is delta neutral, with a gamma of −5,000 and a vega of −8,000. The options shown in the table below can be traded. The portfolio can be made vega neutral by including a long position in 4,000 of Option 1. This would increase delta to 2,400 and require that 2,400 units of the asset be sold to maintain delta neutrality. The gamma of the portfolio would change from −5,000 to −3,000.

	Delta	*Gamma*	*Vega*
Portfolio	0	−5000	−8000
Option 1	0.6	0.5	2.0
Option 2	0.5	0.8	1.2

To make the portfolio gamma and vega neutral, both Option 1 and Option 2 can be used. If w_1 and w_2 are the quantities of Option 1 and Option 2 that are added to the portfolio, we require that

$$-5{,}000 + 0.5w_1 + 0.8w_2 = 0$$

and

$$-8{,}000 + 2.0w_1 + 1.2w_2 = 0$$

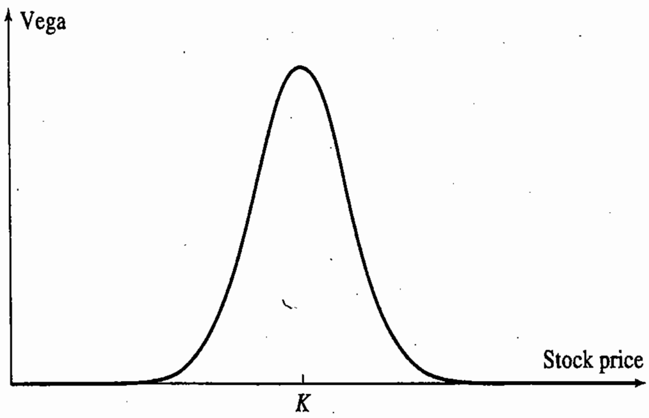
The solution to these equations is $w_1 = 400$, $w_2 = 6{,}000$. The portfolio can therefore be made gamma and vega neutral by including 400 of Option 1 and 6,000 of Option 2. The delta of the portfolio, after the addition of the positions in the two traded options, is $400 \times 0.6 + 6{,}000 \times 0.5 = 3{,}240$. Hence, 3,240 units of the asset would have to be sold to maintain delta neutrality.

For a European call or put option on a non-dividend-paying stock, vega is given by

$$\mathcal{V} = S_0\sqrt{T}\,N'(d_1)$$

where d_1 is defined as in equation (13.20). The formula for $N'(x)$ is given in equation (17.2). The vega of a long position in a European or American option is always positive. The general way in which vega varies with S_0 is shown in Figure 17.11.

Figure 17.11 Variation of vega with stock price for an option.



Example 17.6

As in Example 17.1, consider a call option on a non-dividend-paying stock where the stock price is \$49, the strike price is \$50, the risk-free rate is 5%, the time to maturity is 20 weeks (= 0.3846 years), and the volatility is 20%. In this case, $S_0 = 49$, $K = 50$, $r = 0.05$, $\sigma = 0.2$, and $T = 0.3846$.

The option's vega is

$$S_0\sqrt{T}N'(d_1) = 12.1$$

Thus a 1% (0.01) increase in the volatility from (20% to 21%) increases the value of the option by approximately $0.01 \times 12.1 = 0.121$.

Calculating vega from the Black–Scholes model and its extensions may seem strange because one of the assumptions underlying Black–Scholes is that volatility is constant. It would be theoretically more correct to calculate vega from a model in which volatility is assumed to be stochastic. However, it turns out that the vega calculated from a stochastic volatility model is very similar to the Black–Scholes vega, so the practice of calculating vega from a model in which volatility is constant works reasonably well.[10]

Gamma neutrality protects against large changes in the price of the underlying asset between hedge rebalancing. Vega neutrality protects for a variable σ. As might be expected, whether it is best to use an available traded option for vega or gamma hedging depends on the time between hedge rebalancing and the volatility of the volatility.[11]

When volatilities change, the implied volatilities of short-dated options tend to change by more than the implied volatilities of long-dated options. The vega of a portfolio is therefore often calculated by changing the volatilities of long-dated options by less than that of short-dated options. One way of doing this is discussed in Section 21.6.

17.9 RHO

The *rho* of a portfolio of options is the rate of change of the value of the portfolio with respect to the interest rate:

$$\frac{\partial \Pi}{\partial r}$$

It measures the sensitivity of the value of a portfolio to a change in the interest rate when all else remains the same. For a European call option on a non-dividend-paying stock,

$$\text{rho (call)} = KTe^{-rT}N(d_2)$$

where d_2 is defined as in equation (13.20). For a European put option,

$$\text{rho (put)} = -KTe^{-rT}N(-d_2)$$

[10] See J.C. Hull and A. White, "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Finance* 42 (June 1987): 281–300; J.C. Hull and A. White, "An Analysis of the Bias in Option Pricing Caused by a Stochastic Volatility," *Advances in Futures and Options Research* 3 (1988): 27–61.

[11] For a discussion of this issue, see J.C. Hull and A. White, "Hedging the Risks from Writing Foreign Currency Options," *Journal of International Money and Finance* 6 (June 1987): 131–52.

Example 17.7

As in Example 17.1, consider a call option on a non-dividend-paying stock where the stock price is \$49, the strike price is \$50, the risk-free rate is 5%, the time to maturity is 20 weeks (= 0.3846 years), and the volatility is 20%. In this case, $S_0 = 49$, $K = 50$, $r = 0.05$, $\sigma = 0.2$, and $T = 0.3846$.

The option's rho is

$$KTe^{-rT}N(d_2) = 8.91$$

This means that a 1% (0.01) increase in the risk-free rate (from 5% to 6%) increases the value of the option by approximately $0.01 \times 8.91 = 0.0891$.

17.10 THE REALITIES OF HEDGING

In an ideal world, traders working for financial institutions would be able to rebalance their portfolios very frequently in order to maintain all Greeks equal to zero. In practice, this is not possible. When managing a large portfolio dependent on a single underlying asset, traders usually make delta zero, or close to zero, at least once a day by trading the underlying asset. Unfortunately, a zero gamma and a zero vega are less easy to achieve because it is difficult to find options or other nonlinear derivatives that can be traded in the volume required at competitive prices. Business Snapshot 17.1 provides a discussion of how dynamic hedging is organized at financial institutions.

There are big economies of scale in trading derivatives. Maintaining delta neutrality for a small number of options on an asset by trading daily is usually not economically feasible. The trading costs per option being hedged is high.[12] But when a derivatives dealer maintains delta neutrality for a large portfolio of options on an asset, the trading costs per option hedged are likely to be much more reasonable.

17.11 SCENARIO ANALYSIS

In addition to monitoring risks such as delta, gamma, and vega, option traders often also carry out a scenario analysis. The analysis involves calculating the gain or loss on their portfolio over a specified period under a variety of different scenarios. The time period chosen is likely to depend on the liquidity of the instruments. The scenarios can be either chosen by management or generated by a model.

Consider a bank with a portfolio of options on a foreign currency. There are two main variables on which the value of the portfolio depends. These are the exchange rate and the exchange-rate volatility. Suppose that the exchange rate is currently 1.0000 and its volatility is 10% per annum. The bank could calculate a table such as Table 17.5 showing the profit or loss experienced during a 2-week period under different scenarios. This table considers seven different exchange rates and three different volatilities. Because a one-standard-deviation move in the exchange rate during a 2-week period is about 0.02, the exchange rate moves considered are approximately one, two, and three standard deviations.

---

[12] The trading costs arise from the fact that each day the hedger buys some of the underlying asset at the offer price or sells some of the underlying asset at the bid price.

Business Snapshot 17.1 Dynamic Hedging in Practice

In a typical arrangement at a financial institution, the responsibility for a portfolio of derivatives dependent on a particular underlying asset is assigned to one trader or to a group of traders working together. For example, one trader at Goldman Sachs might be assigned responsibility for all derivatives dependent on the value of the Australian dollar. A computer system calculates the value of the portfolio and Greek letters for the portfolio. Limits are defined for each Greek letter and special permission is required if a trader wants to exceed a limit at the end of a trading day.

The delta limit is often expressed as the equivalent maximum position in the underlying asset. For example, the delta limit of Goldman Sachs on Microsoft might be $10 million. If the Microsoft stock price is $50 this means that the absolute value of delta as we have calculated it can be no more that 200,000. The vega limit is usually expressed as a maximum dollar exposure per 1% change in the volatility.

As a matter of course, options traders make themselves delta neutral—or close to delta neutral—at the end of each day. Gamma and vega are monitored, but are not usually managed on a daily basis. Financial institutions often find that their business with clients involves writing options and that as a result they accumulate negative gamma and vega. They are then always looking out for opportunities to manage their gamma and vega risks by buying options at competitive prices.

There is one aspect of an options portfolio that mitigates problems of managing gamma and vega somewhat. Options are often close to the money when they are first sold, so that they have relatively high gammas and vegas. But after some time has elapsed, the underlying asset price has often changed enough for them to become deep out of the money or deep in the money. Their gammas and vegas are then very small and of little consequence. The nightmare scenario for an options trader is where written options remain very close to the money as the maturity date is approached.

In Table 17.5, the greatest loss is in the lower right corner of the table. The loss corresponds to the volatility increasing to 12% and the exchange rate moving up to 1.06. Usually the greatest loss in a table such as 17.5 occurs at one of the corners, but this is not always so. Consider, for example, the situation where a bank's portfolio consists of a short position in a butterfly spread (see Section 10.2). The greatest loss will be experienced if the exchange rate stays where it is.

Table 17.5 Profit or loss realized in 2 weeks under different scenarios ($ million).

Volatility	*Exchange rate*						
	0.94	*0.96*	*0.98*	*1.00*	*1.02*	*1.04*	*1.06*
8%	+102	+55	+25	+6	−10	−34	−80
10%	+80	+40	+17	+2	−14	−38	−85
12%	+60	+25	+9	−2	−18	−42	−90

17.12 EXTENSION OF FORMULAS

The formulas produced so far for delta, theta, gamma, vega, and rho have been for an option on a non-dividend-paying stock. Table 17.6 shows how they change when the stock pays a continuous dividend yield at rate q. The expressions for d_1 and d_2 are as for equations (15.4) and (15.5). By setting q equal to the dividend yield on an index, we obtain the Greek letters for European options on indices. By setting q equal to the foreign risk-free rate, we obtain the Greek letters for European options on a currency. By setting $q = r$, we obtain Greek letters for European options on a futures contract. An exception lies in the calculation of rho for European options on a futures contract. The rho for a call futures option is $-cT$ and the rho for a European put futures option is $-pT$.

In the case of currency options, there are two rhos corresponding to the two interest rates. The rho corresponding to the domestic interest rate is given by the formula in Table 17.6 (with d_2 as in equation (15.11)). The rho corresponding to the foreign interest rate for a European call on a currency is

$$\text{rho} = -Te^{-r_f T} S_0 N(d_1)$$

For a European put, it is

$$\text{rho} = Te^{-r_f T} S_0 N(-d_1)$$

with d_1 as in equation (15.11).

Delta of Forward Contracts

The concept of delta can be applied to financial instruments other than options. Consider a forward contract on a non-dividend-paying stock. Equation (5.5) shows that the value of a forward contract is $S_0 - Ke^{-rT}$, where K is the delivery price and T is the forward contract's time to maturity. When the price of the stock changes by ΔS, with all else remaining the same, the value of a forward contract on the stock also changes by ΔS. The delta of a long forward contract on one share of the stock is therefore always 1.0. This

Table 17.6 Greek letters for options on an asset that provides a yield at rate q

Greek letter	*Call option*	*Put option*
Delta	$e^{-qT}N(d_1)$	$e^{-qT}[N(d_1) - 1]$
Gamma	$\dfrac{N'(d_1)e^{-qT}}{S_0\sigma\sqrt{T}}$	$\dfrac{N'(d_1)e^{-qT}}{S_0\sigma\sqrt{T}}$
Theta	$-S_0N'(d_1)\sigma e^{-qT}/(2\sqrt{T}) + qS_0N(d_1)e^{-qT} - rKe^{-rT}N(d_2)$	$-S_0N'(d_1)\sigma e^{-qT}/(2\sqrt{T}) - qS_0N(-d_1)e^{-qT} + rKe^{-rT}N(-d_2)$
Vega	$S_0\sqrt{T}N'(d_1)e^{-qT}$	$S_0\sqrt{T}N'(d_1)e^{-qT}$
Rho	$KTe^{-rT}N(d_2)$	$-KTe^{-rT}N(-d_2)$

means that a long forward contract on one share can be hedged by shorting one share; a short forward contract on one share can be hedged by purchasing one share.[13]

For an asset providing a dividend yield at rate q, equation (5.7) shows that the forward contract's delta is e^{-qT}. For the delta of a forward contract on a stock index, q is set equal to the dividend yield on the index in this expression. For the delta of a forward contract, it is set equal to the foreign risk-free rate, r_f.

Delta of a Futures Contract

From equation (5.1), the futures price for a contract on a non-dividend-paying stock is S_0e^{rT}, where T is the time to maturity of the futures contract. This shows that when the price of the stock changes by ΔS, with all else remaining the same, the futures price changes by $\Delta S\, e^{rT}$. Since futures contracts are marked to market daily, the holder of a long futures position makes an almost immediate gain of this amount. The delta of a futures contract is therefore e^{rT}. For a futures position on an asset providing a dividend yield at rate q, equation (5.3) shows similarly that delta is $e^{(r-q)T}$.

It is interesting that marking to market makes the deltas of futures and forward contracts slightly different. This is true even when interest rates are constant and the forward price equals the futures price. (A related point is made in Business Snapshot 5.2.)

Sometimes a futures contract is used to achieve a delta-neutral position. Define:

T: Maturity of futures contract

H_A: Required position in asset for delta hedging

H_F: Alternative required position in futures contracts for delta hedging

If the underlying asset is a non-dividend-paying stock, the analysis we have just given shows that

$$H_F = e^{-rT} H_A \tag{17.5}$$

When the underlying asset pays a dividend yield q,

$$H_F = e^{-(r-q)T} H_A \tag{17.6}$$

For a stock index, we set q equal to the dividend yield on the index; for a currency, we set it equal to the foreign risk-free rate, r_f, so that

$$H_F = e^{-(r-r_f)T} H_A \tag{17.7}$$

Example 17.8

Suppose that a portfolio of currency options held by a US bank can be made delta neutral with a short position of 458,000 pounds sterling. Risk-free rates are 4% in the US and 7% in the UK. From equation (17.7), hedging using 9-month currency futures requires a short futures position

$$e^{-(0.04-0.07)\times 9/12} 458,000$$

or £468,442. Since each futures contract is for the purchase or sale of £62,500, seven contracts would be shorted. (Seven is the nearest whole number to 468,442/62,500.)

[13] These are hedge-and-forget schemes. Since delta is always 1.0, no changes need to be made to the position in the stock during the life of the contract.

17.13 PORTFOLIO INSURANCE

A portfolio manager is often interested in acquiring a put option on his or her portfolio. This provides protection against market declines while preserving the potential for a gain if the market does well. One approach (discussed in Section 15.1) is to buy put options on a market index such as the S&P 500. An alternative is to create the options synthetically.

Creating an option synthetically involves maintaining a position in the underlying asset (or futures on the underlying asset) so that the delta of the position is equal to the delta of the required option. The position necessary to create an option synthetically is the reverse of that necessary to hedge it. This is because the procedure for hedging an option involves the creation of an equal and opposite option synthetically.

There are two reasons why it may be more attractive for the portfolio manager to create the required put option synthetically than to buy it in the market. First, options markets do not always have the liquidity to absorb the trades required by managers of large funds. Second, fund managers often require strike prices and exercise dates that are different from those available in exchange-traded options markets.

The synthetic option can be created from trading the portfolio or from trading in index futures contracts. We first examine the creation of a put option by trading the portfolio. From Table 17.6 the delta of a European put on the portfolio is

$$\Delta = e^{-qT}[N(d_1) - 1] \tag{17.8}$$

where, with our usual notation,

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

S_0 is the value of the portfolio, K is the strike price, r is the risk-free rate, q is the dividend yield on the portfolio, σ is the volatility of the portfolio, and T is the life of the option. The volatility of the portfolio can usually be assumed to be its beta times the volatility of a well-diversified market index.

To create the put option synthetically, the fund manager should ensure that at any given time a proportion

$$e^{-qT}[1 - N(d_1)]$$

of the stocks in the original portfolio has been sold and the proceeds invested in riskless assets. As the value of the original portfolio declines, the delta of the put given by equation (17.8) becomes more negative and the proportion of the original portfolio sold must be increased. As the value of the original portfolio increases, the delta of the put becomes less negative and the proportion of the original portfolio sold must be decreased (i.e., some of the original portfolio must be repurchased).

Using this strategy to create portfolio insurance means that at any given time funds are divided between the stock portfolio on which insurance is required and riskless assets. As the value of the stock portfolio increases, riskless assets are sold and the position in the stock portfolio is increased. As the value of the stock portfolio declines, the position in the stock portfolio is decreased and riskless assets are purchased. The cost of the insurance arises from the fact that the portfolio manager is always selling after a decline in the market and buying after a rise in the market.

Example 17.9

A portfolio is worth \$90 million. To protect against market downturns the managers of the portfolio require a 6-month European put option on the portfolio with a strike price of \$87 million. The risk-free rate is 9% per annum, the dividend yield is 3% per annum, and the volatility of the portfolio is 25% per annum. The S&P 500 index stands at 900. As the portfolio is considered to mimic the S&P 500 fairly closely, one alternative, discussed in Section 15.1, is to buy 1,000 put option contracts on the S&P 500 with a strike price of 870. Another alternative is to create the required option synthetically. In this case, $S_0 = 90$ million, $K = 87$ million, $r = 0.09$, $q = 0.03$, $\sigma = 0.25$, and $T = 0.5$, so that

$$d_1 = \frac{\ln(90/87) + (0.09 - 0.03 + 0.25^2/2)0.5}{0.25\sqrt{0.5}} = 0.4499$$

and the delta of the required option is

$$e^{-qT}[N(d_1) - 1] = -0.3215$$

This shows that 32.15% of the portfolio should be sold initially to match the delta of the required option. The amount of the portfolio sold must be monitored frequently. For example, if the value of the portfolio reduces to \$88 million after 1 day, the delta of the required option changes to 0.3679 and a further 4.64% of the original portfolio should be sold. If the value of the portfolio increases to \$92 million, the delta of the required option changes to −0.2787 and 4.28% of the original portfolio should be repurchased.

Use of Index Futures

Using index futures to create options synthetically can be preferable to using the underlying stocks because the transaction costs associated with trades in index futures are generally lower than those associated with the corresponding trades in the underlying stocks. The dollar amount of the futures contracts shorted as a proportion of the value of the portfolio should from equations (17.6) and (17.8) be

$$e^{-qT}e^{-(r-q)T^*}[1 - N(d_1)] = e^{q(T^*-T)}e^{-rT^*}[1 - N(d_1)]$$

where T^* is the maturity of the futures contract. If the portfolio is worth A_1 times the index and each index futures contract is on A_2 times the index, the number of futures contracts shorted at any given time should be

$$e^{q(T^*-T)}e^{-rT^*}[1 - N(d_1)]A_1/A_2$$

Example 17.10

Suppose that in the previous example futures contracts on the S&P 500 maturing in 9 months are used to create the option synthetically. In this case initially $T = 0.5$, $T^* = 0.75$, $A_1 = 100{,}000$, $A_2 = 250$, and $d_1 = 0.4499$, so that the number of futures contracts shorted should be

$$e^{q(T^*-T)}e^{-rT^*}[1 - N(d_1)]A_1/A_2 = 122.96$$

or 123, rounding to the nearest whole number. As time passes and the index changes, the position in futures contracts must be adjusted.

Business Snapshot 17.2 Was Portfolio Insurance to Blame for the Crash of 1987?

On Monday, October 19, 1987, the Dow Jones Industrial Average dropped by more than 20%. Many people feel that portfolio insurance played a major role in this crash. In October 1987 between \$60 billion and \$90 billion of equity assets were subject to portfolio insurance trading rules where put options were created synthetically in the way discussed in Section 17.13. During the period Wednesday, October 14, 1987, to Friday, October 16, 1987, the market declined by about 10%, with much of this decline taking place on Friday afternoon. The portfolio trading rules should have generated at least \$12 billion of equity or index futures sales as a result of this decline. In fact, portfolio insurers had time to sell only \$4 billion and they approached the following week with huge amounts of selling already dictated by their models. It is estimated that on Monday, October 19, sell programs by three portfolio insurers accounted for almost 10% of the sales on the New York Stock Exchange, and that portfolio insurance sales amounted to 21.3% of all sales in index futures markets. It is likely that the decline in equity prices was exacerbated by investors other than portfolio insurers selling heavily because they anticipated the actions of portfolio insurers.

Because the market declined so fast and the stock exchange systems were overloaded, many portfolio insurers were unable to execute the trades generated by their models and failed to obtain the protection they required. Needless to say, the popularity of portfolio insurance schemes has declined significantly since 1987. One of the morals of this story is that it is dangerous to follow a particular trading strategy—even a hedging strategy—when many other market participants are doing the same thing.

This analysis assumes that the portfolio mirrors the index. When this is not the case, it is necessary to (a) calculate the portfolio's beta, (b) find the position in options on the index that gives the required protection, and (c) choose a position in index futures to create the options synthetically. As discussed in Section 15.1, the strike price for the options should be the expected level of the market index when the portfolio reaches its insured value. The number of options required is beta times the number that would be required if the portfolio had a beta of 1.0.

17.14 STOCK MARKET VOLATILITY

We discussed in Chapter 13 the issue of whether volatility is caused solely by the arrival of new information or whether trading itself generates volatility. Portfolio insurance strategies such as those just described have the potential to increase volatility. When the market declines, they cause portfolio managers either to sell stock or to sell index futures contracts. Either action may accentuate the decline (see Business Snapshot 17.2). The sale of stock is liable to drive down the market index further in a direct way. The sale of index futures contracts is liable to drive down futures prices. This creates selling pressure on stocks via the mechanism of index arbitrage (see Chapter 5), so that the market index is liable to be driven down in this case as well. Similarly, when the market rises, the portfolio insurance strategies cause portfolio managers either to buy stock or to buy futures contracts. This may accentuate the rise.

In addition to formal portfolio trading strategies, we can speculate that many investors consciously or subconsciously follow portfolio insurance rules of their own. For example, an investor may be inclined to enter the market when it is rising but will sell when it is falling to limit the downside risk.

Whether portfolio insurance trading strategies (formal or informal) affect volatility depends on how easily the market can absorb the trades that are generated by portfolio insurance. If portfolio insurance trades are a very small fraction of all trades, there is likely to be no effect. As portfolio insurance becomes more popular, it is liable to have a destabilizing effect on the market.

SUMMARY

Financial institutions offer a variety of option products to their clients. Often the options do not correspond to the standardized products traded by exchanges. The financial institutions are then faced with the problem of hedging their exposure. Naked and covered positions leave them subject to an unacceptable level of risk. One course of action that is sometimes proposed is a stop-loss strategy. This involves holding a naked position when an option is out of the money and converting it to a covered position as soon as the option moves into the money. Although superficially attractive, the strategy does not provide a good hedge.

The delta (Δ) of an option is the rate of change of its price with respect to the price of the underlying asset. Delta hedging involves creating a position with zero delta (sometimes referred to as a delta-neutral position). Because the delta of the underlying asset is 1.0, one way of hedging is to take a position of $-\Delta$ in the underlying asset for each long option being hedged. The delta of an option changes over time. This means that the position in the underlying asset has to be frequently adjusted.

Once an option position has been made delta neutral, the next stage is often to look at its gamma (Γ). The gamma of an option is the rate of change of its delta with respect to the price of the underlying asset. It is a measure of the curvature of the relationship between the option price and the asset price. The impact of this curvature on the performance of delta hedging can be reduced by making an option position gamma neutral. If Γ is the gamma of the position being hedged, this reduction is usually achieved by taking a position in a traded option that has a gamma of $-\Gamma$.

Delta and gamma hedging are both based on the assumption that the volatility of the underlying asset is constant. In practice, volatilities do change over time. The vega of an option or an option portfolio measures the rate of change of its value with respect to volatility. A trader who wishes to hedge an option position against volatility changes can make the position vega neutral. As with the procedure for creating gamma neutrality, this usually involves taking an offsetting position in a traded option. If the trader wishes to achieve both gamma and vega neutrality, two traded options are usually required.

Two other measures of the risk of an option position are theta and rho. Theta measures the rate of change of the value of the position with respect to the passage of time, with all else remaining constant. Rho measures the rate of change of the value of the position with respect to the interest rate, with all else remaining constant.

In practice, option traders usually rebalance their portfolios at least once a day to maintain delta neutrality. It is usually not feasible to maintain gamma and vega neutrality on a regular basis. Typically a trader monitors these measures. If they get too large, either corrective action is taken or trading is curtailed.

Portfolio managers are sometimes interested in creating put options synthetically for the purposes of insuring an equity portfolio. They can do so either by trading the portfolio or by trading index futures on the portfolio. Trading the portfolio involves splitting the portfolio between equities and risk-free securities. As the market declines, more is invested in risk-free securities. As the market increases, more is invested in equities. Trading index futures involves keeping the equity portfolio intact and selling index futures. As the market declines, more index futures are sold; as it rises, fewer are sold. This type of portfolio insurance works well in normal market conditions. On Monday, October 19, 1987, when the Dow Jones Industrial Average dropped very sharply, it worked badly. Portfolio insurers were unable to sell either stocks or index futures fast enough to protect their positions.

FURTHER READING

Taleb, N. N., *Dynamic Hedging: Managing Vanilla and Exotic Options*. New York: Wiley, 1996.

Questions and Problems (Answers in Solutions Manual)

17.1. Explain how a stop-loss trading rule can be implemented for the writer of an out-of-the-money call option. Why does it provide a relatively poor hedge?

17.2. What does it mean to assert that the delta of a call option is 0.7? How can a short position in 1,000 options be made delta neutral when the delta of each option is 0.7?

17.3. Calculate the delta of an at-the-money six-month European call option on a non-dividend-paying stock when the risk-free interest rate is 10% per annum and the stock price volatility is 25% per annum.

17.4. What does it mean to assert that the theta of an option position is −0.1 when time is measured in years? If a trader feels that neither a stock price nor its implied volatility will change, what type of option position is appropriate?

17.5. What is meant by the gamma of an option position? What are the risks in the situation where the gamma of a position is highly negative and the delta is zero?

17.6. "The procedure for creating an option position synthetically is the reverse of the procedure for hedging the option position." Explain this statement.

17.7. Why did portfolio insurance not work well on October 19, 1987?

17.8. The Black–Scholes price of an out-of-the-money call option with an exercise price of $40 is $4. A trader who has written the option plans to use a stop-loss strategy. The trader's plan is to buy at $40.10 and to sell at $39.90. Estimate the expected number of times the stock will be bought or sold.

17.9. Suppose that a stock price is currently $20 and that a call option with an exercise price of $25 is created synthetically using a continually changing position in the stock. Consider the following two scenarios: (a) Stock price increases steadily from $20 to $35 during the life of the option; (b) Stock price oscillates wildly, ending up at $35. Which scenario would make the synthetically created option more expensive? Explain your answer.

17.10. What is the delta of a short position in 1,000 European call options on silver futures? The options mature in 8 months, and the futures contract underlying the option matures in 9 months. The current 9-month futures price is \$8 per ounce, the exercise price of the options is \$8, the risk-free interest rate is 12% per annum, and the volatility of silver is 18% per annum.

17.11. In Problem 17.10, what initial position in 9-month silver futures is necessary for delta hedging? If silver itself is used, what is the initial position? If 1-year silver futures are used, what is the initial position? Assume no storage costs for silver.

17.12. A company uses delta hedging to hedge a portfolio of long positions in put and call options on a currency. Which of the following would give the most favorable result?
(a) A virtually constant spot rate
(b) Wild movements in the spot rate
Explain your answer.

17.13. Repeat Problem 17.12 for a financial institution with a portfolio of short positions in put and call options on a currency.

17.14. A financial institution has just sold 1,000 7-month European call options on the Japanese yen. Suppose that the spot exchange rate is 0.80 cent per yen, the exercise price is 0.81 cent per yen, the risk-free interest rate in the United States is 8% per annum, the risk-free interest rate in Japan is 5% per annum, and the volatility of the yen is 15% per annum. Calculate the delta, gamma, vega, theta, and rho of the financial institution's position. Interpret each number.

17.15. Under what circumstances is it possible to make a European option on a stock index both gamma neutral and vega neutral by adding a position in one other European option?

17.16. A fund manager has a well-diversified portfolio that mirrors the performance of the S&P 500 and is worth \$360 million. The value of the S&P 500 is 1,200, and the portfolio manager would like to buy insurance against a reduction of more than 5% in the value of the portfolio over the next 6 months. The risk-free interest rate is 6% per annum. The dividend yield on both the portfolio and the S&P 500 is 3%, and the volatility of the index is 30% per annum.
(a) If the fund manager buys traded European put options, how much would the insurance cost?
(b) Explain carefully alternative strategies open to the fund manager involving traded European call options, and show that they lead to the same result.
(c) If the fund manager decides to provide insurance by keeping part of the portfolio in risk-free securities, what should the initial position be?
(d) If the fund manager decides to provide insurance by using 9-month index futures, what should the initial position be?

17.17. Repeat Problem 17.16 on the assumption that the portfolio has a beta of 1.5. Assume that the dividend yield on the portfolio is 4% per annum.

17.18. Show by substituting for the various terms in equation (17.4) that the equation is true for:
(a) A single European call option on a non-dividend-paying stock
(b) A single European put option on a non-dividend-paying stock
(c) Any portfolio of European put and call options on a non-dividend-paying stock

17.19. What is the equation corresponding to equation (17.4) for (a) a portfolio of derivatives on a currency and (b) a portfolio of derivatives on a futures price?

17.20. Suppose that \$70 billion of equity assets are the subject of portfolio insurance schemes. Assume that the schemes are designed to provide insurance against the value of the assets declining by more than 5% within 1 year. Making whatever estimates you find necessary, use the DerivaGem software to calculate the value of the stock or futures contracts that the administrators of the portfolio insurance schemes will attempt to sell if the market falls by 23% in a single day.

17.21. Does a forward contract on a stock index have the same delta as the corresponding futures contract? Explain your answer.

17.22. A bank's position in options on the dollar/euro exchange rate has a delta of 30,000 and a gamma of −80,000. Explain how these numbers can be interpreted. The exchange rate (dollars per euro) is 0.90. What position would you take to make the position delta neutral? After a short period of time, the exchange rate moves to 0.93. Estimate the new delta. What additional trade is necessary to keep the position delta neutral? Assuming the bank did set up a delta-neutral position originally, has it gained or lost money from the exchange-rate movement?

17.23. Use the put–call parity relationship to derive, for a non-dividend-paying stock, the relationship between:
(a) The delta of a European call and the delta of a European put
(b) The gamma of a European call and the gamma of a European put
(c) The vega of a European call and the vega of a European put
(d) The theta of a European call and the theta of a European put

Assignment Questions

17.24. Consider a 1-year European call option on a stock when the stock price is \$30, the strike price is \$30, the risk-free rate is 5%, and the volatility is 25% per annum. Use the DerivaGem software to calculate the price, delta, gamma, vega, theta, and rho of the option. Verify that delta is correct by changing the stock price to \$30.1 and recomputing the option price. Verify that gamma is correct by recomputing the delta for the situation where the stock price is \$30.1. Carry out similar calculations to verify that vega, theta, and rho are correct. Use the DerivaGem Applications Builder functions to plot the option price, delta, gamma, vega, theta, and rho against the stock price for the stock option.

17.25. A financial institution has the following portfolio of over-the-counter options on sterling:

Type	*Position*	*Delta of option*	*Gamma of option*	*Vega of option*
Call	−1,000	0.50	2.2	1.8
Call	−500	0.80	0.6	0.2
Put	−2,000	−0.40	1.3	0.7
Call	−500	0.70	1.8	1.4

A traded option is available with a delta of 0.6, a gamma of 1.5, and a vega of 0.8.
(a) What position in the traded option and in sterling would make the portfolio both gamma neutral and delta neutral?
(b) What position in the traded option and in sterling would make the portfolio both vega neutral and delta neutral?

17.26. Consider again the situation in Problem 17.25. Suppose that a second traded option with a delta of 0.1, a gamma of 0.5, and a vega of 0.6 is available. How could the portfolio be made delta, gamma, and vega neutral?

17.27. A deposit instrument offered by a bank guarantees that investors will receive a return during a 6-month period that is the greater of (a) zero and (b) 40% of the return provided by a market index. An investor is planning to put $100,000 in the instrument. Describe the payoff as an option on the index. Assuming that the risk-free rate of interest is 8% per annum, the dividend yield on the index is 3% per annum, and the volatility of the index is 25% per annum, is the product a good deal for the investor?

17.28. The formula for the price c of a European call futures option in terms of the futures price F_0 is given in Chapter 16 as

$$c = e^{-rT}[F_0 N(d_1) - KN(d_2)]$$

where

$$d_1 = \frac{\ln(F_0/K) + \sigma^2 T/2}{\sigma\sqrt{T}} \quad \text{and} \quad d_2 = d_1 - \sigma\sqrt{T}$$

and K, r, T, and σ are the strike price, interest rate, time to maturity, and volatility, respectively.

(a) Prove that $F_0 N'(d_1) = KN'(d_2)$.

(b) Prove that the delta of the call price with respect to the futures price is $e^{-rT}N(d_1)$.

(c) Prove that the vega of the call price is $F_0\sqrt{T}N'(d_1)e^{-rT}$.

(d) Prove the formula for the rho of a call futures option given in Section 17.12.

The delta, gamma, theta, and vega of a call futures option are the same as those for a call option on a stock paying dividends at rate q, with q replaced by r and S_0 replaced by F_0. Explain why the same is not true of the rho of a call futures option.

17.29. Use DerivaGem to check that equation (17.4) is satisfied for the option considered in Section 17.1. (*Note*: DerivaGem produces a value of theta "per calendar day". The theta in equation (17.4) is "per year".)

17.30. Use the DerivaGem Application Builder functions to reproduce Table 17.2. (In Table 17.2 the stock position is rounded to the nearest 100 shares.) Calculate the gamma and theta of the position each week. Calculate the change in the value of the portfolio each week and check whether equation (17.3) is approximately satisfied. (*Note*: DerivaGem produces a value of theta "per calendar day". The theta in equation (17.3) is "per year".)

APPENDIX

TAYLOR SERIES EXPANSIONS AND HEDGE PARAMETERS

A Taylor series expansion of the change in the portfolio value in a short period of time shows the role played by different Greek letters. If the volatility of the underlying asset is assumed to be constant, the value Π of the portfolio is a function of the asset price S, and time t. The Taylor series expansion gives

$$\Delta\Pi = \frac{\partial\Pi}{\partial S}\Delta S + \frac{\partial\Pi}{\partial t}\Delta t + \frac{1}{2}\frac{\partial^2\Pi}{\partial S^2}\Delta S^2 + \frac{1}{2}\frac{\partial^2\Pi}{\partial t^2}\Delta t^2 + \frac{\partial^2\Pi}{\partial S\,\partial t}\Delta S\,\Delta t + \cdots \tag{17A.1}$$

where $\Delta\Pi$ and ΔS are the change in Π and S in a small time interval Δt. Delta hedging eliminates the first term on the right-hand side. The second term is nonstochastic. The third term (which is of order Δt) can be made zero by ensuring that the portfolio is gamma neutral as well as delta neutral. Other terms are of order higher than Δt.

For a delta-neutral portfolio, the first term on the right-hand side of equation (17A.1) is zero, so that

$$\Delta\Pi = \Theta\,\Delta t + \tfrac{1}{2}\Gamma\,\Delta S^2$$

when terms of order higher than Δt are ignored. This is equation (17.3).

When the volatility of the underlying asset is uncertain, Π is a function of σ, S, and t. Equation (17A.1) then becomes

$$\Delta\Pi = \frac{\partial\Pi}{\partial S}\Delta S + \frac{\partial\Pi}{\partial\sigma}\Delta\sigma + \frac{\partial\Pi}{\partial t}\Delta t + \frac{1}{2}\frac{\partial^2\Pi}{\partial S^2}\Delta S^2 + \frac{1}{2}\frac{\partial^2\Pi}{\partial\sigma^2}\Delta\sigma^2 + \cdots$$

where $\Delta\sigma$ is the change in σ in time Δt. In this case, delta hedging eliminates the first term on the right-hand side. The second term is eliminated by making the portfolio vega neutral. The third term is nonstochastic. The fourth term is eliminated by making the portfolio gamma neutral. Traders sometimes define other Greek letters to correspond to later terms in the expansion.

CHAPTER 18

Volatility Smiles

How close are the market prices of options to those predicted by Black–Scholes? Do traders really use Black–Scholes when determining a price for an option? Are the probability distributions of asset prices really lognormal? This chapter answers these questions. It explains that traders do use the Black–Scholes model—but not in exactly the way that Black and Scholes originally intended. This is because they allow the volatility used to price an option to depend on its strike price and time to maturity.

A plot of the implied volatility of an option as a function of its strike price is known as a *volatility smile*. This chapter describes the volatility smiles that traders use in equity and foreign currency markets. It explains the relationship between a volatility smile and the risk-neutral probability distribution being assumed for the future asset price. It also discusses how option traders allow volatility to be a function of option maturity and how they use volatility surfaces as pricing tools.

18.1 WHY THE VOLATILITY SMILE IS THE SAME FOR CALLS AND PUTS

This section shows that the implied volatilities of European call and put options should be equal when they have the same strike price and time to maturity.

As explained in Chapter 9, put–call parity provides a relationship between the prices of European call and put options when they have the same strike price and time to maturity. With a dividend yield on the underlying asset of q, the relationship is

$$p + S_0e^{-qT} = c + Ke^{-rT} \tag{18.1}$$

As usual, c and p are the European call and put price. They have the same strike price, K, and time to maturity, T. The variable S_0 is the price of the underlying asset today, r is the risk-free interest rate for maturity T, and q is the yield on the asset.

A key feature of the put–call parity relationship is that it is based on a relatively simple no-arbitrage argument. It does not require any assumption about the probability distribution of the asset price in the future. It is true both when the asset price distribution is lognormal and when it is not lognormal.

Suppose that, for a particular value of the volatility, p_{BS} and c_{BS} are the values of European put and call options calculated using the Black–Scholes model. Suppose further that p_{mkt} and c_{mkt} are the market values of these options. Because put–call parity holds for the Black–Scholes model, we must have

$$p_{\text{BS}} + S_0e^{-qT} = c_{\text{BS}} + Ke^{-rT}$$

In the absence of arbitrage opportunities, put–call parity also holds for the market prices, so that

$$p_{\text{mkt}} + S_0e^{-qT} = c_{\text{mkt}} + Ke^{-rT}$$

Subtracting these two equations, we get

$$p_{\text{BS}} - p_{\text{mkt}} = c_{\text{BS}} - c_{\text{mkt}} \tag{18.2}$$

This shows that the dollar pricing error when the Black–Scholes model is used to price a European put option should be exactly the same as the dollar pricing error when it is used to price a European call option with the same strike price and time to maturity.

Suppose that the implied volatility of the put option is 22%. This means that $p_{\text{BS}} = p_{\text{mkt}}$ when a volatility of 22% is used in the Black–Scholes model. From equation (18.2), it follows that $c_{\text{BS}} = c_{\text{mkt}}$ when this volatility is used. The implied volatility of the call is, therefore, also 22%. This argument shows that the implied volatility of a European call option is always the same as the implied volatility of a European put option when the two have the same strike price and maturity date. To put this another way, for a given strike price and maturity, the correct volatility to use in conjunction with the Black–Scholes model to price a European call should always be the same as that used to price a European put. This means that the volatility smile (i.e., the relationship between implied volatility and strike price for a particular maturity) is the same for calls and puts. It also means that the volatility term structure (i.e., the relationship between implied volatility and maturity for a particular strike) is the same for calls and puts.

Example 18.1

The value of the Australian dollar is \$0.60. The risk-free interest rate is 5% per annum in the United States and 10% per annum in Australia. The market price of a European call option on the Australian dollar with a maturity of 1 year and a strike price of \$0.59 is 0.0236. DerivaGem shows that the implied volatility of the call is 14.5%. For there to be no arbitrage, the put–call parity relationship in equation (18.1) must apply with q equal to the foreign risk-free rate. The price p of a European put option with a strike price of \$0.59 and maturity of 1 year therefore satisfies

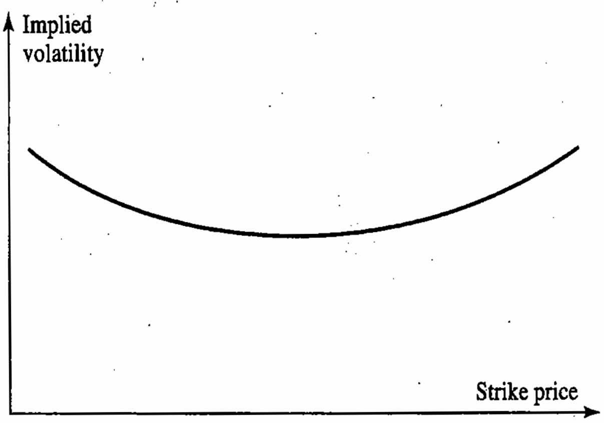
$$p + 0.60e^{-0.10\times 1} = 0.0236 + 0.59e^{-0.05\times 1}$$

so that $p = 0.0419$. DerivaGem shows that, when the put has this price, its implied volatility is also 14.5%. This is what we expect from the analysis just given.

18.2 FOREIGN CURRENCY OPTIONS

The volatility smile used by traders to price foreign currency options has the general form shown in Figure 18.1. The implied volatility is relatively low for at-the-money options. It becomes progressively higher as an option moves either into the money or out of the money.

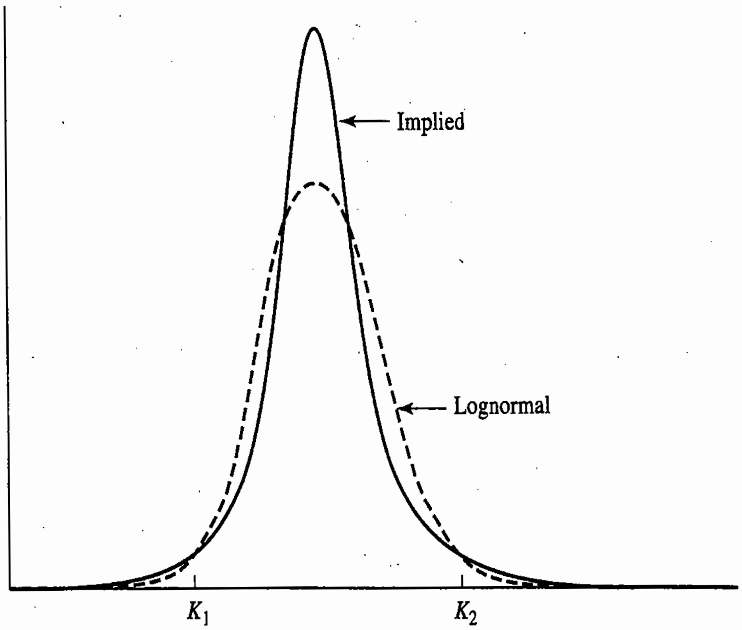
Figure 18.1 Volatility smile for foreign currency options.



In the appendix at the end of this chapter we show how to determine the risk-neutral probability distribution for an asset price at a future time from the volatility smile given by options maturing at that time. We refer to this as the *implied distribution*. The volatility smile in Figure 18.1 corresponds to the implied distribution shown by the solid line in Figure 18.2. A lognormal distribution with the same mean and standard deviation as the implied distribution is shown by the dashed line in Figure 18.2. It can be seen that the implied distribution has heavier tails than the lognormal distribution.[1]

To see that Figures 18.1 and 18.2 are consistent with each other, consider first a deep-

Figure 18.2 Implied and lognormal distribution for foreign currency options.



[1] This is known as *kurtosis*. Note that, in addition to having a heavier tail, the implied distribution is more "peaked". Both small and large movements in the exchange rate are more likely than with the lognormal distribution. Intermediate movements are less likely.

out-of-the-money call option with a high strike price of K_2. This option pays off only if the exchange rate proves to be above K_2. Figure 18.2 shows that the probability of this is higher for the implied probability distribution than for the lognormal distribution. We therefore expect the implied distribution to give a relatively high price for the option. A relatively high price leads to a relatively high implied volatility—and this is exactly what we observe in Figure 18.1 for the option. The two figures are therefore consistent with each other for high strike prices. Consider next a deep-out-of-the-money put option with a low strike price of K_1. This option pays off only if the exchange rate proves to be below K_1. Figure 18.2 shows that the probability of this is also higher for implied probability distribution than for the lognormal distribution. We therefore expect the implied distribution to give a relatively high price, and a relatively high implied volatility, for this option as well. Again, this is exactly what we observe in Figure 18.1.

Empirical Results

We have just shown that the volatility smile used by traders for foreign currency options implies that they consider that the lognormal distribution understates the probability of extreme movements in exchange rates. To test whether they are right, Table 18.1 examines the daily movements in 12 different exchange rates over a 10-year period.[2] The first step in the production of the table is to calculate the standard deviation of daily percentage change in each exchange rate. The next stage is to note how often the actual percentage change exceeded 1 standard deviation, 2 standard deviations, and so on. The final stage is to calculate how often this would have happened if the percentage changes had been normally distributed. (The lognormal model implies that percentage changes are almost exactly normally distributed over a one-day time period.)

Daily changes exceed 3 standard deviations on 1.34% of days. The lognormal model predicts that this should happen on only 0.27% of days. Daily changes exceed 4, 5, and 6 standard deviations on 0.29%, 0.08%, and 0.03% of days, respectively. The lognormal model predicts that we should hardly ever observe this happening. The table therefore provides evidence to support the existence of heavy tails (Figure 18.2) and the volatility smile used by traders (Figure 18.1). Business Snapshot 18.1 shows how you could have made money if you had done the analysis in Table 18.1 ahead of the rest of the market.

Table 18.1 Percentage of days when daily exchange rate moves are greater than one, two,..., six standard deviations (SD = standard deviation of daily change).

	Real world	*Lognormal model*
>1 SD	25.04	31.73
>2 SD	5.27	4.55
>3 SD	1.34	0.27
>4 SD	0.29	0.01
>5 SD	0.08	0.00
>6 SD	0.03	0.00

[2] This table is taken from J. C. Hull and A. White, "Value at Risk When Daily Changes in Market Variables Are Not Normally Distributed." *Journal of Derivatives*, 5, No. 3 (Spring 1998): 9–19.

Business Snapshot 18.1 Making Money from Foreign Currency Options

Suppose that most market participants think that exchange rates are lognormally distributed. They will be comfortable using the same volatility to value all options on a particular exchange rate. You have just done the analysis in Table 18.1 and know that the lognormal assumption is not a good one for exchange rates. What should you do?

The answer is that you should buy deep-out-the-money call and put options on a variety of different currencies and wait. These options will be relatively inexpensive and more of them will close in the money than the lognormal model predicts. The present value of your payoffs will on average be much greater than the cost of the options.

In the mid-1980s a few traders knew about the heavy tails of foreign exchange probability distributions. Everyone else thought that the lognormal assumption of Black–Scholes was reasonable. The few traders who were well informed followed the strategy we have described—and made lots of money. By the late 1980s everyone realized that foreign currency options should be priced with a volatility smile and the trading opportunity disappeared.

Reasons for the Smile in Foreign Currency Options

Why are exchange rates not lognormally distributed? Two of the conditions for an asset price to have a lognormal distribution are:

1. The volatility of the asset is constant.
2. The price of the asset changes smoothly with no jumps.

In practice, neither of these conditions is satisfied for an exchange rate. The volatility of an exchange rate is far from constant, and exchange rates frequently exhibit jumps.[3] It turns out that the effect of both a nonconstant volatility and jumps is that extreme outcomes become more likely.

The impact of jumps and nonconstant volatility depends on the option maturity. As the maturity of the option is increased, the percentage impact of a nonconstant volatility on prices becomes more pronounced, but the percentage impact on implied volatility usually becomes less pronounced. The percentage impact of jumps on both prices and the implied volatility becomes less pronounced as the maturity of the option is increased.[4] The result of all this is that the volatility smile becomes less pronounced as option maturity increases.

18.3 EQUITY OPTIONS

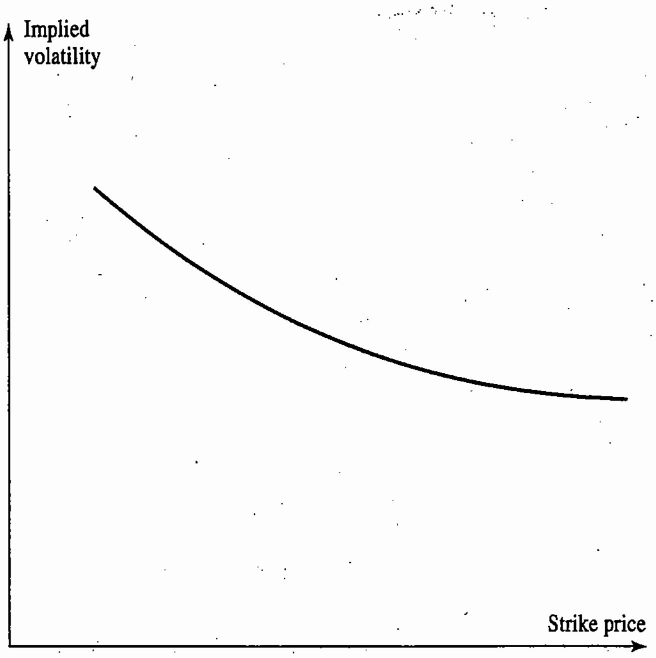
The volatility smile for equity options has been studied by Rubinstein (1985, 1994) and Jackwerth and Rubinstein (1996). Prior to 1987 there was no marked volatility smile. Since 1987 the volatility smile used by traders to price equity options (both on individual

---

[3] Sometimes the jumps are in response to the actions of central banks.

[4] When we look at sufficiently long-dated options, jumps tend to get "averaged out" so that the exchange rate distribution when there are jumps is almost indistinguishable from the one obtained when the exchange rate changes smoothly.

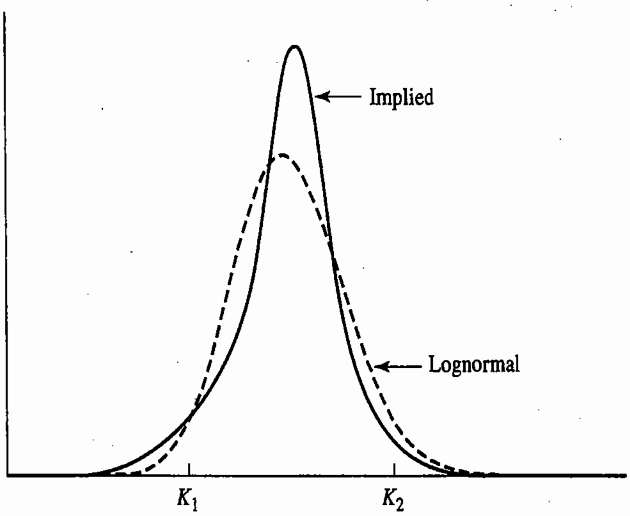
Figure 18.3 Volatility smile for equities.



stocks and on stock indices) has had the general form shown in Figure 18.3. This is sometimes referred to as a *volatility skew*. The volatility decreases as the strike price increases. The volatility used to price a low-strike-price option (i.e., a deep-out-of-the-money put or a deep-in-the-money call) is significantly higher than that used to price a high-strike-price option (i.e., a deep-in-the-money put or a deep-out-of-the-money call).

The volatility smile for equity options corresponds to the implied probability distribution given by the solid line in Figure 18.4. A lognormal distribution with the same mean and standard deviation as the implied distribution is shown by the dotted line. It can be seen that the implied distribution has a heavier left tail and a less heavy right tail than the lognormal distribution.

Figure 18.4 Implied distribution and lognormal distribution for equity options.



Business Snapshot 18.2 Crashophobia

It is interesting that the pattern in Figure 18.3 for equities has existed only since the stock market crash of October 1987. Prior to October 1987, implied volatilities were much less dependent on strike price. This has led Mark Rubinstein to suggest that one reason for the equity volatility smile may be "crashophobia". Traders are concerned about the possibility of another crash similar to October 1987, and they price options accordingly.

There is some empirical support for this explanation. Declines in the S&P 500 tend to be accompanied by a steepening of the volatility skew. When the S&P increases, the skew tends to become less steep.

To see that Figures 18.3 and 18.4 are consistent with each other, we proceed as for Figures 18.1 and 18.2 and consider options that are deep out of the money. From Figure 18.4 a deep-out-of-the-money call with a strike price of K_2 has a lower price when the implied distribution is used than when the lognormal distribution is used. This is because the option pays off only if the stock price proves to be above K_2, and the probability of this is lower for the implied probability distribution than for the lognormal distribution. Therefore, we expect the implied distribution to give a relatively low price for the option. A relatively low price leads to a relatively low implied volatility—and this is exactly what we observe in Figure 18.3 for the option. Consider next a deep-out-of-the-money put option with a strike price of K_1. This option pays off only if the stock price proves to be below K_1. Figure 18.4 shows that the probability of this is higher for the implied probability distribution than for the lognormal distribution. We therefore expect the implied distribution to give a relatively high price, and a relatively high implied volatility, for this option. Again, this is exactly what we observe in Figure 18.3.

The Reason for the Smile in Equity Options

One possible explanation for the smile in equity options concerns leverage. As a company's equity declines in value, the company's leverage increases. This means that the equity becomes more risky and its volatility increases. As a company's equity increases in value, leverage decreases. The equity then becomes less risky and its volatility decreases. This argument shows that we can expect the volatility of equity to be a decreasing function of price and is consistent with Figures 18.3 and 18.4. Another explanation is "crashophobia" (see Business Snapshot 18.2).

18.4 ALTERNATIVE WAYS OF CHARACTERIZING THE VOLATILITY SMILE

So far we have defined the volatility smile as the relationship between implied volatility and strike price. The relationship depends on the current price of the asset. For example, the lowest point of the volatility smile in Figure 18.1 is usually close to the current exchange rate. If the exchange rate increases, the volatility smile tends to move to the right; if the exchange rate decreases, the volatility smile tends to move to the left. Similarly, in Figure 18.3, when the equity price increases, the volatility skew tends to move to the right, and when the equity price decreases, it tends to move to the left.[5] For this reason the volatility smile is often calculated as the relationship between the implied volatility and K/S_0 rather than as the relationship between the implied volatility and K. The smile is then much more stable.

A refinement of this is to calculate the volatility smile as the relationship between the implied volatility and K/F_0, where F_0 is the forward price of the asset for a contract maturing at the same time as the options that are considered. Traders also often define an "at-the-money" option as an option where $K = F_0$, not as an option where $K = S_0$. The argument for this is that F_0, not S_0, is the expected stock price on the option's maturity date in a risk-neutral world.[6]

Yet another approach to defining the volatility smile is as the relationship between the implied volatility and the delta of the option (where delta is defined as in Chapter 17). This approach sometimes makes it possible to apply volatility smiles to options other than European and American calls and puts. When the approach is used, an at-the-money option is then defined as a call option with a delta of 0.5 or a put option with a delta of −0.5. These are referred to as "50-delta options".

18.5 THE VOLATILITY TERM STRUCTURE AND VOLATILITY SURFACES

In addition to a volatility smile, traders use a volatility term structure when pricing options. This means that the volatility used to price an at-the-money option depends on the maturity of the option. Volatility tends to be an increasing function of maturity when short-dated volatilities are historically low. This is because there is then an expectation that volatilities will increase. Similarly, volatility tends to be a decreasing function of maturity when short-dated volatilities are historically high. This is because there is then an expectation that volatilities will decrease.

Volatility surfaces combine volatility smiles with the volatility term structure to tabulate the volatilities appropriate for pricing an option with any strike price and any maturity. An example of a volatility surface that might be used for foreign currency options is given in Table 18.2. In this case, we assume that the smile is measured as the relationship between volatility and K/S_0.

One dimension of Table 18.2 is K/S_0; the other is time to maturity. The main body of the table shows implied volatilities calculated from the Black–Scholes model. At any given time, some of the entries in the table are likely to correspond to options for which reliable market data are available. The implied volatilities for these options are calculated directly from their market prices and entered into the table. The rest of the table is typically determined using interpolation.

When a new option has to be valued, financial engineers look up the appropriate volatility in the table. For example, when valuing a 9-month option with a ratio of strike price to asset price of 1.05, a financial engineer would interpolate between 13.4 and 14.0 in Table 18.2 to obtain a volatility of 13.7%. This is the volatility that would be used in the Black–Scholes formula or a binomial tree. When valuing a 1.5-year

---

[5] Research by Derman suggests that this adjustment is sometimes "sticky" in the case of exchange-traded options. See E. Derman, "Regimes of Volatility," *Risk*, April 1999: 55–59.

[6] As explained in Chapter 27 , whether the futures or forward price of the asset is the expected price in a risk-neutral world depends on exactly how the risk-neutral world is defined.

Table 18.2 Volatility surface.

	K/S_0				
	0.90	*0.95*	*1.00*	*1.05*	*1.10*
1 month	14.2	13.0	12.0	13.1	14.5
3 month	14.0	13.0	12.0	13.1	14.2
6 month	14.1	13.3	12.5	13.4	14.3
1 year	14.7	14.0	13.5	14.0	14.8
2 year	15.0	14.4	14.0	14.5	15.1
5 year	14.8	14.6	14.4	14.7	15.0

option with a K/S_0 ratio of 0.925, a two-dimensional interpolation would be used to give an implied volatility of 14.525%.

The shape of the volatility smile depends on the option maturity. As illustrated in Table 18.2, the smile tends to become less pronounced as the option maturity increases. Define T as the time to maturity and F_0 as the forward price of the asset for a contract maturing at the same time as the option. Some financial engineers choose to define the volatility smile as the relationship between implied volatility and

$$\frac{1}{\sqrt{T}}\ln\left(\frac{K}{F_0}\right)$$

rather than as the relationship between the implied volatility and K. The smile is then usually much less dependent on the time to maturity.[7]

The Role of the Model

How important is the pricing model if traders are prepared to use a different volatility for every option? It can be argued that the Black–Scholes model is no more than a sophisticated interpolation tool used by traders for ensuring that an option is priced consistently with the market prices of other actively traded options. If traders stopped using Black–Scholes and switched to another plausible model, then the volatility surface and the shape of the smile would change, but arguably the dollar prices quoted in the market would not change appreciably.

18.6 GREEK LETTERS

The volatility smile complicates the calculation of Greek letters. Assume that the relationship between the implied volatility and K/S for an option with a certain time to maturity remains the same.[8] As the price of the underlying asset changes, the implied

[7] For a discussion of this approach, see S. Natenberg *Option Pricing and Volatility: Advanced Trading Strategies and Techniques*, 2nd edn. McGraw-Hill, 1994; R. Tompkins *Options Analysis: A State of the Art Guide to Options Pricing*, Burr Ridge, IL: Irwin, 1994.

[8] It is interesting that this natural model is internally consistent only when the the volatility smile is flat for all maturities. See, for example, T. Daglish, J. Hull, and W. Suo, "Volatility Surfaces: Theory, Rules of Thumb, and Empirical Evidence," *Quantitative Finance*, 7, 5 (October 2007). 507–24.

volatility of the option changes to reflect the option's "moneyness" (i.e., the extent to which it is in or out of the money). The formulas for Greek letters given in Chapter 17 are no longer correct. For example, delta of a call option is given by

$$\frac{\partial c_{\mathrm{BS}}}{\partial S} + \frac{\partial c_{\mathrm{BS}}}{\partial \sigma_{\mathrm{imp}}} \frac{\partial \sigma_{\mathrm{imp}}}{\partial S}$$

where c_{BS} is the Black-Scholes price of the option expressed as a function of the asset price S and the implied volatility σ_{imp}. Consider the impact of this formula on the delta of an equity call option. Volatility is a decreasing function of K/S. This means that the implied volatility increases as the asset price increases, so that

$$\frac{\partial \sigma_{\mathrm{imp}}}{\partial S} > 0$$

As a result, delta is higher than that given by the Black–Scholes assumptions.

In practice, banks try to ensure that their exposure to the most commonly observed changes in the volatility surface is reasonably small. One technique for identifying these changes is principal components analysis, which we discuss in Chapter 20.

18.7 WHEN A SINGLE LARGE JUMP IS ANTICIPATED

Let us now consider an example of how an unusual volatility smile might arise in equity markets. Suppose that a stock price is currently \$50 and an important news announcement due in a few days is expected either to increase the stock price by \$8 or to reduce it by \$8. (This announcement could concern the outcome of a takeover attempt or the verdict in an important lawsuit.) The probability distribution of the stock price in, say, 1 month might then consist of a mixture of two lognormal distributions, the first corresponding to favorable news, the second to unfavorable news. The situation is illustrated in Figure 18.5. The solid line shows the mixture-of-

Figure 18.5 Effect of a single large jump. The solid line is the true distribution; the dashed line is the lognormal distribution.

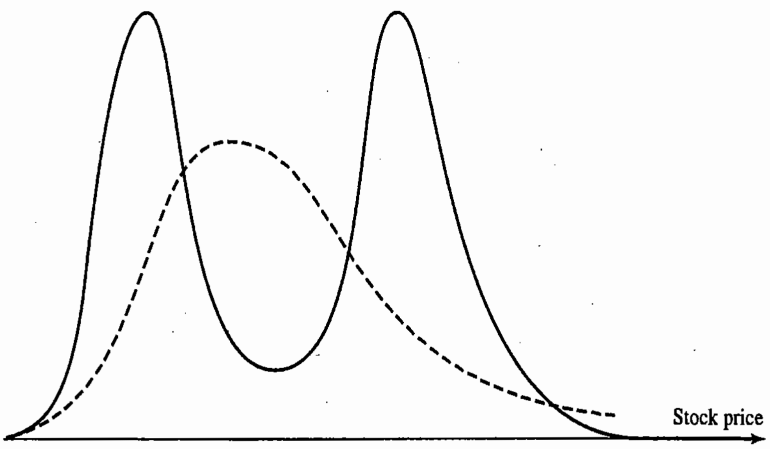
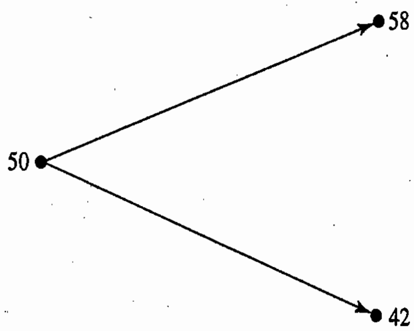


Figure 18.6 Change in stock price in 1 month.



lognormals distribution for the stock price in 1 month; the dashed line shows a lognormal distribution with the same mean and standard deviation as this distribution.

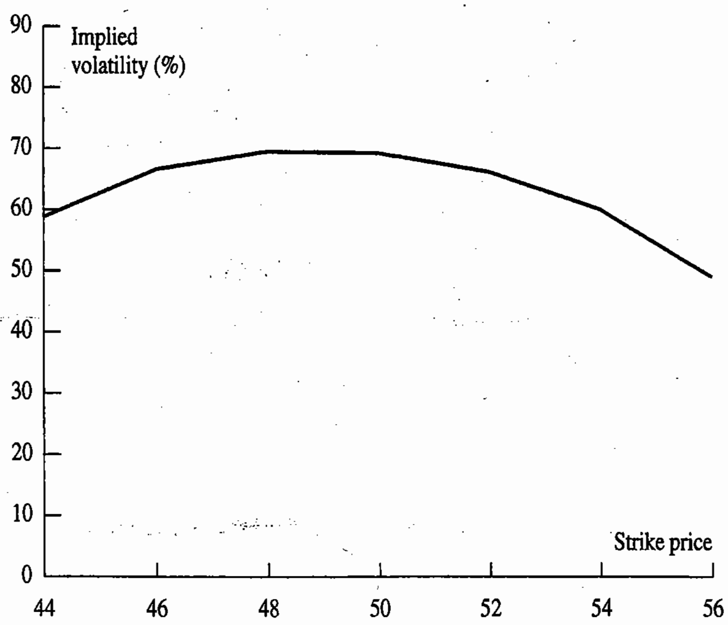
The true probability distribution is bimodal (certainly not lognormal). One easy way to investigate the general effect of a bimodal stock price distribution is to consider the extreme case where the distribution is binomial. This is what we will now do.

Suppose that the stock price is currently \$50 and that it is known that in 1 month it will be either \$42 or \$58. Suppose further that the risk-free rate is 12% per annum. The situation is illustrated in Figure 18.6. Options can be valued using the binomial model from Chapter 11. In this case $u = 1.16$, $d = 0.84$, $a = 1.0101$, and $p = 0.5314$. The results from valuing a range of different options are shown in Table 18.3. The first column shows alternative strike prices; the second column shows prices of 1-month European call options; the third column shows the prices of one-month European put option prices; the fourth column shows implied volatilities. (As shown in Section 18.1, the implied volatility of a European put option is the same as that of a European call option when they have the same strike price and maturity.) Figure 18.7 displays the volatility smile from Table 18.3. It is actually a "frown" (the opposite of that observed for currencies) with volatilities declining as we move out of or into the money. The volatility implied from an option with a strike price of 50 will overprice an option with a strike price of 44 or 56.

Table 18.3 Implied volatilities in situation where true distribution is binomial.

Strike price (\$)	*Call price* (\$)	*Put price* (\$)	*Implied volatility* (%)
42	8.42	0.00	0.0
44	7.37	0.93	58.8
46	6.31	1.86	66.6
48	5.26	2.78	69.5
50	4.21	3.71	69.2
52	3.16	4.64	66.1
54	2.10	5.57	60.0
56	1.05	6.50	49.0
58	0.00	7.42	0.0

Figure 18.7 Volatility smile for situation in Table 18.3.



SUMMARY

The Black–Scholes model and its extensions assume that the probability distribution of the underlying asset at any given future time is lognormal. This assumption is not the one made by traders. They assume the probability distribution of an equity price has a heavier left tail and a less heavy right tail than the lognormal distribution. They also assume that the probability distribution of an exchange rate has a heavier right tail and a heavier left tail than the lognormal distribution.

Traders use volatility smiles to allow for nonlognormality. The volatility smile defines the relationship between the implied volatility of an option and its strike price. For equity options, the volatility smile tends to be downward sloping. This means that out-of-the-money puts and in-the-money calls tend to have high implied volatilities whereas out-of-the-money calls and in-the-money puts tend to have low implied volatilities. For foreign currency options, the volatility smile is U-shaped. Both out-of-the-money and in-the-money options have higher implied volatilities than at-the-money options.

Often traders also use a volatility term structure. The implied volatility of an option then depends on its life. When volatility smiles and volatility term structures are combined, they produce a volatility surface. This defines implied volatility as a function of both the strike price and the time to maturity.

FURTHER READING

Bakshi, G., C. Cao, and Z. Chen. "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance*, 52, No. 5 (December 1997): 2004–49.

Bates, D.S. "Post-'87 Crash Fears in the S&P Futures Market," *Journal of Econometrics*, 94 (January/February 2000): 181–238.

Derman, E. "Regimes of Volatility," *Risk*, April 1999: 55–59.

Ederington, L. H., and W. Guan. "Why Are Those Options Smiling," *Journal of Derivatives*, 10, 2 (2002): 9–34.

Jackwerth, J. C., and M. Rubinstein. "Recovering Probability Distributions from Option Prices," *Journal of Finance*, 51 (December 1996): 1611–31.

Lauterbach, B., and P. Schultz. "Pricing Warrants: An Empirical Study of the Black–Scholes Model and Its Alternatives," *Journal of Finance*, 4, No. 4 (September 1990): 1181–1210.

Melick, W. R., and C. P. Thomas. "Recovering an Asset's Implied Probability Density Function from Option Prices: An Application to Crude Oil during the Gulf Crisis," *Journal of Financial and Quantitative Analysis*, 32, 1 (March 1997): 91–115.

Rubinstein, M. "Nonparametric Tests of Alternative Option Pricing Models Using All Reported Trades and Quotes on the 30 Most Active CBOE Option Classes from August 23, 1976, through August 31, 1978," *Journal of Finance*, 40 (June 1985): 455–80.

Rubinstein, M. "Implied Binomial Trees," *Journal of Finance*, 49, 3 (July 1994): 771–818.

Xu, X., and S. J. Taylor. "The Term Structure of Volatility Implied by Foreign Exchange Options," *Journal of Financial and Quantitative Analysis*, 29 (1994): 57–74.

Questions and Problems (Answers in Solutions Manual)

18.1. What volatility smile is likely to be observed when:
   (a) Both tails of the stock price distribution are less heavy than those of the lognormal distribution?
   (b) The right tail is heavier, and the left tail is less heavy, than that of a lognormal distribution?

18.2. What volatility smile is observed for equities?

18.3. What volatility smile is likely to be caused by jumps in the underlying asset price? Is the pattern likely to be more pronounced for a 2-year option than for a 3-month option?

18.4. A European call and put option have the same strike price and time to maturity. The call has an implied volatility of 30% and the put has an implied volatility of 25%. What trades would you do?

18.5. Explain carefully why a distribution with a heavier left tail and less heavy right tail than the lognormal distribution gives rise to a downward sloping volatility smile.

18.6. The market price of a European call is \$3.00 and its price given by Black–Scholes model with a volatility of 30% is \$3.50. The price given by this Black–Scholes model for a European put option with the same strike price and time to maturity is \$1.00. What should the market price of the put option be? Explain the reasons for your answer.

18.7. Explain what is meant by "crashophobia".

18.8. A stock price is currently \$20. Tomorrow, news is expected to be announced that will either increase the price by \$5 or decrease the price by \$5. What are the problems in using Black–Scholes to value 1-month options on the stock?

18.9. What volatility smile is likely to be observed for 6-month options when the volatility is uncertain and positively correlated to the stock price?

18.10. What problems do you think would be encountered in testing a stock option pricing model empirically?

18.11. Suppose that a central bank's policy is to allow an exchange rate to fluctuate between 0.97 and 1.03. What pattern of implied volatilities for options on the exchange rate would you expect to see?

18.12. Option traders sometimes refer to deep-out-of-the-money options as being options on volatility. Why do you think they do this?

18.13. A European call option on a certain stock has a strike price of \$30, a time to maturity of 1 year, and an implied volatility of 30%. A European put option on the same stock has a strike price of \$30, a time to maturity of 1 year, and an implied volatility of 33%. What is the arbitrage opportunity open to a trader? Does the arbitrage work only when the lognormal assumption underlying Black–Scholes holds? Explain carefully the reasons for your answer.

18.14. Suppose that the result of a major lawsuit affecting a company is due to be announced tomorrow. The company's stock price is currently \$60. If the ruling is favorable to the company, the stock price is expected to jump to \$75. If it is unfavorable, the stock is expected to jump to \$50. What is the risk-neutral probability of a favorable ruling? Assume that the volatility of the company's stock will be 25% for 6 months after the ruling if the ruling is favorable and 40% if it is unfavorable. Use DerivaGem to calculate the relationship between implied volatility and strike price for 6-month European options on the company today. The company does not pay dividends. Assume that the 6-month risk-free rate is 6%. Consider call options with strike prices of \$30, \$40, \$50, \$60, \$70, and \$80.

18.15. An exchange rate is currently 0.8000. The volatility of the exchange rate is quoted as 12% and interest rates in the two countries are the same. Using the lognormal assumption, estimate the probability that the exchange rate in 3 months will be (a) less than 0.7000, (b) between 0.7000 and 0.7500, (c) between 0.7500 and 0.8000, (d) between 0.8000 and 0.8500, (e) between 0.8500 and 0.9000, and (f) greater than 0.9000. Based on the volatility smile usually observed in the market for exchange rates, which of these estimates would you expect to be too low and which would you expect to be too high?

18.16. A stock price is \$40. A 6-month European call option on the stock with a strike price of \$30 has an implied volatility of 35%. A 6-month European call option on the stock with a strike price of \$50 has an implied volatility of 28%. The 6-month risk-free rate is 5% and no dividends are expected. Explain why the two implied volatilities are different. Use DerivaGem to calculate the prices of the two options. Use put–call parity to calculate the prices of 6-month European put options with strike prices of \$30 and \$50. Use DerivaGem to calculate the implied volatilities of these two put options.

18.17. "The Black–Scholes model is used by traders as an interpolation tool." Discuss this view.

18.18. Using Table 18.2, calculate the implied volatility a trader would use for an 8-month option with $K/S_0 = 1.04$.

Assignment Questions

18.19. A company's stock is selling for \$4. The company has no outstanding debt. Analysts consider the liquidation value of the company to be at least \$300,000 and there are 100,000 shares outstanding. What volatility smile would you expect to see?

18.20. A company is currently awaiting the outcome of a major lawsuit. This is expected to be known within 1 month. The stock price is currently \$20. If the outcome is positive, the stock price is expected to be \$24 at the end of 1 month. If the outcome is negative, it is expected to be \$18 at this time. The 1-month risk-free interest rate is 8% per annum.
(a) What is the risk-neutral probability of a positive outcome?
(b) What are the values of 1-month call options with strike prices of \$19, \$20, \$21, \$22, and \$23?
(c) Use DerivaGem to calculate a volatility smile for 1-month call options.
(d) Verify that the same volatility smile is obtained for 1-month put options.

18.21. A futures price is currently \$40. The risk-free interest rate is 5%. Some news is expected tomorrow that will cause the volatility over the next 3 months to be either 10% or 30%. There is a 60% chance of the first outcome and a 40% chance of the second outcome. Use DerivaGem to calculate a volatility smile for 3-month options.

18.22. Data for a number of foreign currencies are provided on the author's website:

`http://www.rotman.utoronto.ca/~hull`

Choose a currency and use the data to produce a table similar to Table 18.1.

18.23. Data for a number of stock indices are provided on the author's website:

`http://www.rotman.utoronto.ca/~hull`

Choose an index and test whether a three-standard-deviation down movement happens more often than a three-standard-deviation up movement.

18.24. Consider a European call and a European put with the same strike price and time to maturity. Show that they change in value by the same amount when the volatility increases from a level σ_1 to a new level σ_2 within a short period of time. (*Hint*: Use put–call parity.)

18.25. An exchange rate is currently 1.0 and the implied volatilities of 6-month European options with strike prices 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3 are 13%, 12%, 11%, 10%, 11%, 12%, 13%. The domestic and foreign risk-free rates are both 2.5%. Calculate the implied probability distribution using an approach similar to that used for Example 18A.1 in the appendix to this chapter. Compare it with the implied distribution where all the implied volatilities are 11.5%.

18.26. Using Table 18.2, calculate the implied volatility a trader would use for an 11-month option with $K/S_0 = 0.98$.

APPENDIX

DETERMINING IMPLIED RISK-NEUTRAL DISTRIBUTIONS FROM VOLATILITY SMILES

The price of a European call option on an asset with strike price K and maturity T is given by

$$c = e^{-rT} \int_{S_T=K}^{\infty} (S_T - K)\, g(S_T)\, dS_T$$

where r is the interest rate (assumed constant), S_T is the asset price at time T, and g is the risk-neutral probability density function of S_T. Differentiating once with respect to K gives

$$\frac{\partial c}{\partial K} = -e^{-rT} \int_{S_T=K}^{\infty} g(S_T)\, dS_T$$

Differentiating again with respect to K gives

$$\frac{\partial^2 c}{\partial K^2} = e^{-rT}\, g(K)$$

This shows that the probability density function g is given by

$$g(K) = e^{rT}\, \frac{\partial^2 c}{\partial K^2}$$

This result, which is from Breeden and Litzenberger (1978), allows risk-neutral probability distributions to be estimated from volatility smiles.[9] Suppose that c_1, c_2, and c_3 are the prices of T-year European call options with strike prices of $K - \delta$, K, and $K + \delta$, respectively. Assuming δ is small, an estimate of $g(K)$ is

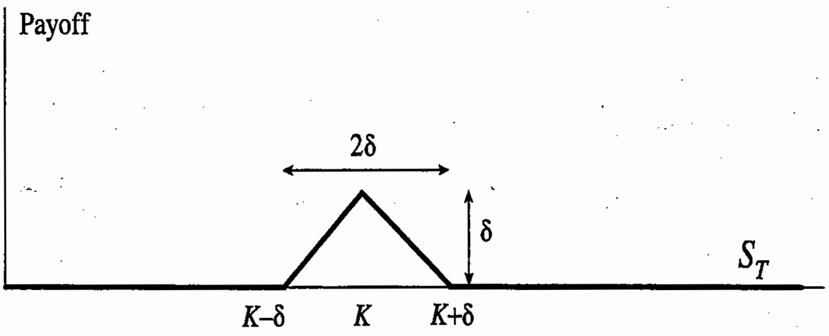
$$e^{rT}\, \frac{c_1 + c_3 - 2c_2}{\delta^2}$$

For another way of understanding this formula, suppose you set up a butterfly spread with strike prices $K - \delta$, K, and $K + \delta$, and maturity T. This means that you buy a call with strike price $K - \delta$, buy a call with strike price $K + \delta$, and sell two calls with strike price K. The value of your position is $c_1 + c_3 - 2c_2$. The value of the position can also be calculated by integrating the payoff over the risk-neutral probability distribution, $g(S_T)$, and discounting at the risk-free rate. The payoff is shown in Figure 18A.1. Since δ is small, we can assume that $g(S_T) = g(K)$ in the whole of the range $K - \delta < S_T < K + \delta$, where the payoff is nonzero. The area under the "spike" in Figure 18A.1 is $0.5 \times 2\delta \times \delta = \delta^2$. The value of the payoff (when δ is small) is therefore $e^{-rT} g(K)\delta^2$. It follows that

$$e^{-rT} g(K)\delta^2 = c_1 + c_3 - 2c_2$$

[9] See D. T. Breeden and R. H. Litzenberger, "Prices of State-Contingent Claims Implicit in Option Prices," *Joumal of Business*, 51 (1978), 621–51.

Figure 18A.1 Payoff from butterfly spread.



which leads directly to

$$g(K) = e^{rT}\frac{c_1 + c_3 - 2c_2}{\delta^2} \tag{18A.1}$$

Example 18A.1

Suppose that the price of a non-dividend-paying stock is \$10, the risk-free interest rate is 3%, and the implied volatilities of three-month European options with strike prices of \$6, \$7, \$8, \$9, \$10, \$11, \$12, \$13, \$14 are 30%, 29%, 28%, 27%, 26%, 25%, 24%, 23%, 22%, respectively. One way of applying the above results is as follows. Assume that $g(S_T)$ is constant between $S_T = 6$ and $S_T = 7$, constant between $S_T = 7$ and $S_T = 8$, and so on. Define:

$$\begin{aligned}
g(S_T) &= g_1 \quad \text{for} \quad 6 \leqslant S_T < 7\\
g(S_T) &= g_2 \quad \text{for} \quad 7 \leqslant S_T < 8\\
g(S_T) &= g_3 \quad \text{for} \quad 8 \leqslant S_T < 9\\
g(S_T) &= g_4 \quad \text{for} \quad 9 \leqslant S_T < 10\\
g(S_T) &= g_5 \quad \text{for} \quad 10 \leqslant S_T < 11\\
g(S_T) &= g_6 \quad \text{for} \quad 11 \leqslant S_T < 12\\
g(S_T) &= g_7 \quad \text{for} \quad 12 \leqslant S_T < 13\\
g(S_T) &= g_8 \quad \text{for} \quad 13 \leqslant S_T < 14
\end{aligned}$$

The value of g_1 can be calculated by interpolating to get the implied volatility for a one-year option with a strike price of \$6.5 as 29.5%. This means that options with strike prices of \$6, \$6.5, and \$7 have implied volatilities of 30%, 29.5%, and 29%, respectively. From DerivaGem their prices are \$4.045, \$3.549, and \$3.055, respectively. Using equation (18A.1), with $K = 6.5$ and $\delta = 0.5$, gives

$$g_1 = \frac{e^{0.03\times0.25}(4.045 + 3.055 - 2 \times 3.549)}{0.5^2} = 0.0057$$

Similar calculations show that

$$g_2 = 0.0444, \quad g_3 = 0.1545, \quad g_4 = 0.2781$$

$$g_5 = 0.2813, \quad g_6 = 0.1659, \quad g_7 = 0.0573, \quad g_8 = 0.0113$$

Figure 18.A.2 Implied probability distribution for Example 18A.1.

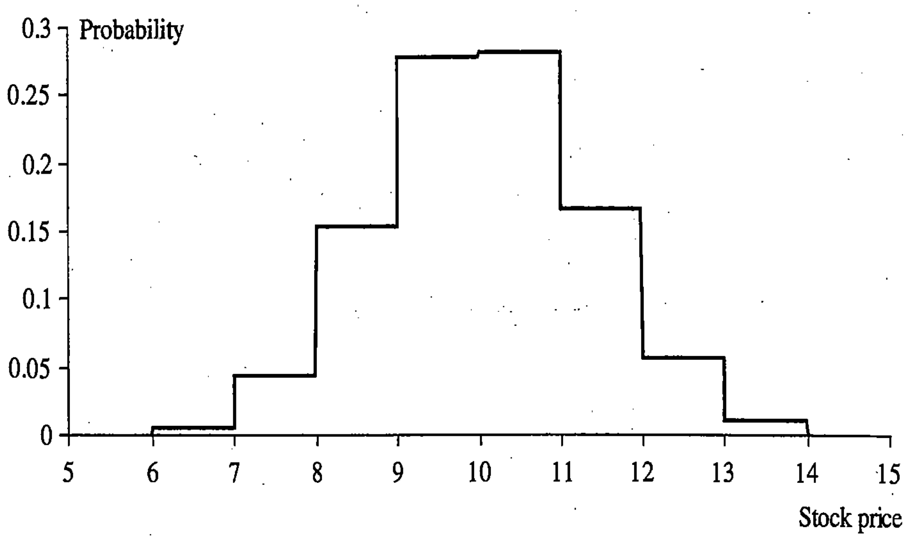
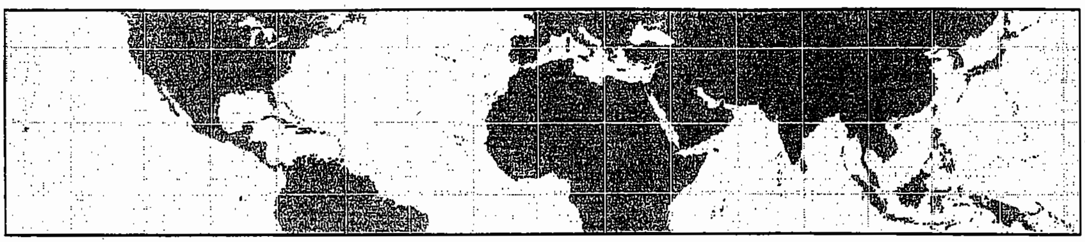


Figure 18A.2 displays the implied distribution. (Note that the area under the probability distribution is 0.9985. The probability that $S_T < 6$ or $S_T > 14$ is therefore 0.0015.) Although not obvious from Figure 18A.2, the implied distribution does have a heavier left tail and less heavy right tail than a lognormal distribution. For the lognormal distribution based on a single volatility of 26%, the probability of a stock price between \$6 and \$7 is 0.0022 (compared with 0.0057 in Figure 18A.2) and the probability of a stock price between \$13 and \$14 is 0.0141 (compared with 0.0113 in Figure 18A.2).



CHAPTER 19

Basic Numerical Procedures

This chapter discusses three numerical procedures for valuing derivatives when exact formulas are not available. The first involves representing the asset price movements in the form of a tree and was introduced in Chapter 11. The second involves Monte Carlo simulation, which we encountered briefly in Chapter 12 when stochastic processes were explained. The third involves finite difference methods.

Monte Carlo simulation is usually used for derivatives where the payoff is dependent on the history of the underlying variable or where there are several underlying variables. Trees and finite difference methods are usually used for American options and other derivatives where the holder has decisions to make prior to maturity. In addition to valuing a derivative, all the procedures can be used to calculate Greek letters such as delta, gamma, and vega.

The basic procedures discussed in this chapter can be used to handle most of the derivatives valuation problems encountered in practice. However, sometimes they have to be adapted to cope with particular situations, as will be explained in Chapter 26.

19.1 BINOMIAL TREES

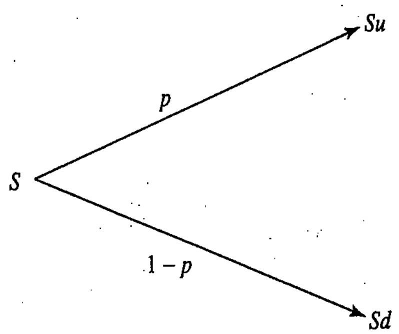
Binomial trees were introduced in Chapter 11. They can be used to value either European or American options. The Black–Scholes formulas and their extensions that were presented in Chapters 13, 15, and 16 provide analytic valuations for European options.[1] There are no analytic valuations for American options. Binomial trees are therefore most useful for valuing these types of options.[2]

As explained in Chapter 11, the binomial tree valuation approach involves dividing the life of the option into a large number of small time intervals of length Δt. It assumes that in each time interval the price of the underlying asset moves from its initial value of S to one of two new values, Su and Sd. The approach is illustrated in Figure 19.1. In

[1] The Black–Scholes formulas are based on the same set of assumptions as binomial trees. As one might expect, in the limit as the number of time steps is increased, the price given by a binomial tree for a European option converges to the Black–Scholes price.

[2] Some analytic approximations for valuing American options have been suggested. The most well-known one is the quadratic approximation approach. See Technical Note 8 on the author's website for a description of this approach.

Figure 19.1 Asset price movements in time Δt under the binomial model.



general, $u > 1$ and $d < 1$. The movement from S to Su, therefore, is an "up" movement and the movement from S to Sd is a "down" movement. The probability of an up movement will be denoted by p. The probability of a down movement is $1 - p$.

Risk-Neutral Valuation

The risk-neutral valuation principle, explained in Chapters 11 and 13, states that an option (or other derivative) can be valued on the assumption that the world is risk neutral. This means that for valuation purposes we can use the following procedure:

1. Assume that the expected return from all traded assets is the risk-free interest rate.
2. Value payoffs from the derivative by calculating their expected values and discounting at the risk-free interest rate.

This principle underlies the way trees are used.

Determination of p, u, and d

The parameters p, u, and d must give correct values for the mean and variance of asset price changes during a time interval of length Δt. Because we are working in a risk-neutral world, the expected return from the asset is the risk-free interest rate, r. Suppose that the asset provides a yield of q. The expected return in the form of capital gains must be $r - q$. This means that the expected value of the asset price at the end of a time interval of length Δt must be $Se^{(r-q)\Delta t}$, where S is the asset price at the beginning of the time interval. To match the mean return with the tree, we therefore need

$$Se^{(r-q)\Delta t} = pSu + (1 - p)Sd$$

or

$$e^{(r-q)\Delta t} = pu + (1 - p)d \tag{19.1}$$

As explained in Section 13.4, the variance of the percentage change, R, in the asset price in a small time interval of length Δt is $\sigma^2 \Delta t$, where σ is the asset price volatility. This is also the variance of $1 + R$. (Adding a constant to a variable makes no difference to its variance.) The variance of a variable Q is defined as $E(Q^2) - [E(Q)]^2$. There is a probability p that $1 + R$ is u and a probability $1 - p$ that it is d. It follows that

$$pu^2 + (1 - p)d^2 - e^{2(r-q)\Delta t} = \sigma^2 \Delta t$$

From equation (19.1), $e^{(r-q)\Delta t}(u+d) = pu^2 + (1-p)d^2 + ud$, so that

$$e^{(r-q)\Delta t}(u+d) - ud - e^{2(r-q)\Delta t} = \sigma^2 \Delta t \tag{19.2}$$

Equations (19.1) and (19.2) impose two conditions on p, u, and d. A third condition used by Cox, Ross, and Rubinstein (1979) is[3]

$$u = \frac{1}{d} \tag{19.3}$$

A solution to equations (19.1) to (19.3), when terms of higher order than Δt are ignored, is[4]

$$p = \frac{a-d}{u-d} \tag{19.4}$$

$$u = e^{\sigma\sqrt{\Delta t}} \tag{19.5}$$

$$d = e^{-\sigma\sqrt{\Delta t}} \tag{19.6}$$

where

$$a = e^{(r-q)\Delta t} \tag{19.7}$$

The variable a is sometimes referred to as the *growth factor*. Equations (19.4) to (19.7) are the same as those in Section 11.9.

Tree of Asset Prices

Figure 19.2 illustrates the complete tree of asset prices that is considered when the binomial model is used. At time zero, the asset price, S_0, is known. At time Δt, there are two possible asset prices, S_0u and S_0d; at time $2\Delta t$, there are three possible asset prices, S_0u^2, S_0, and S_0d^2; and so on. In general, at time $i\,\Delta t$, we consider $i+1$ asset prices. These are

$$S_0u^jd^{i-j}, \quad j = 0, 1, \dots, i$$

Note that the relationship $u = 1/d$ is used in computing the asset price at each node of the tree in Figure 19.2. For example, $S_0u^2d = S_0u$. Note also that the tree recombines in the sense that an up movement followed by a down movement leads to the same asset price as a down movement followed by an up movement.

Working Backward through the Tree

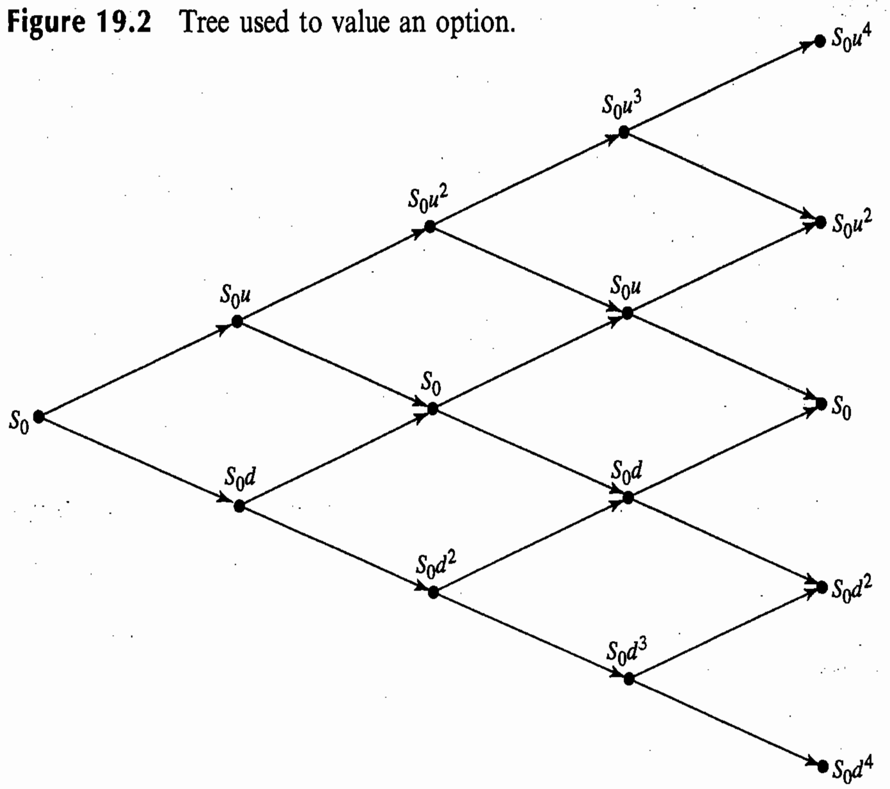
Options are evaluated by starting at the end of the tree (time T) and working backward. The value of the option is known at time T. For example, a put option is worth $\max(K - S_T, 0)$ and a call option is worth $\max(S_T - K, 0)$, where S_T is the asset price at time T and K is the strike price. Because a risk-neutral world is being assumed, the

---

[3] See J.C. Cox, S.A. Ross, and M. Rubinstein, "Option Pricing: A Simplified Approach," *Journal of Financial Economics*, 7 (October 1979), 229–63.

[4] To see this, we note that equations (19.4) and (19.7) satisfy the conditions in equations (19.1) and (19.3) exactly. The exponential function e^x can be expanded as $1 + x + x^2/2 + \cdots$. When terms of higher order than Δt are ignored, equation (19.5) implies that $u = 1 + \sigma\sqrt{\Delta t} + \frac{1}{2}\sigma^2\Delta t$ and equation (19.6) implies that $d = 1 - \sigma\sqrt{\Delta t} + \frac{1}{2}\sigma^2\Delta t$. Also, $e^{(r-q)\Delta t} = 1 + (r-q)\Delta t$ and $e^{2(r-q)\Delta t} = 1 + 2(r-q)\Delta t$. By substitution, we see that equation (19.2) is satisfied when terms of higher order than Δt are ignored.

Figure 19.2 Tree used to value an option.



value at each node at time $T - \Delta t$ can be calculated as the expected value at time T discounted at rate r for a time period Δt. Similarly, the value at each node at time $T - 2\Delta t$ can be calculated as the expected value at time $T - \Delta t$ discounted for a time period Δt at rate r, and so on. If the option is American, it is necessary to check at each node to see whether early exercise is preferable to holding the option for a further time period Δt. Eventually, by working back through all the nodes, we are able to obtain the value of the option at time zero.

Example 19.1

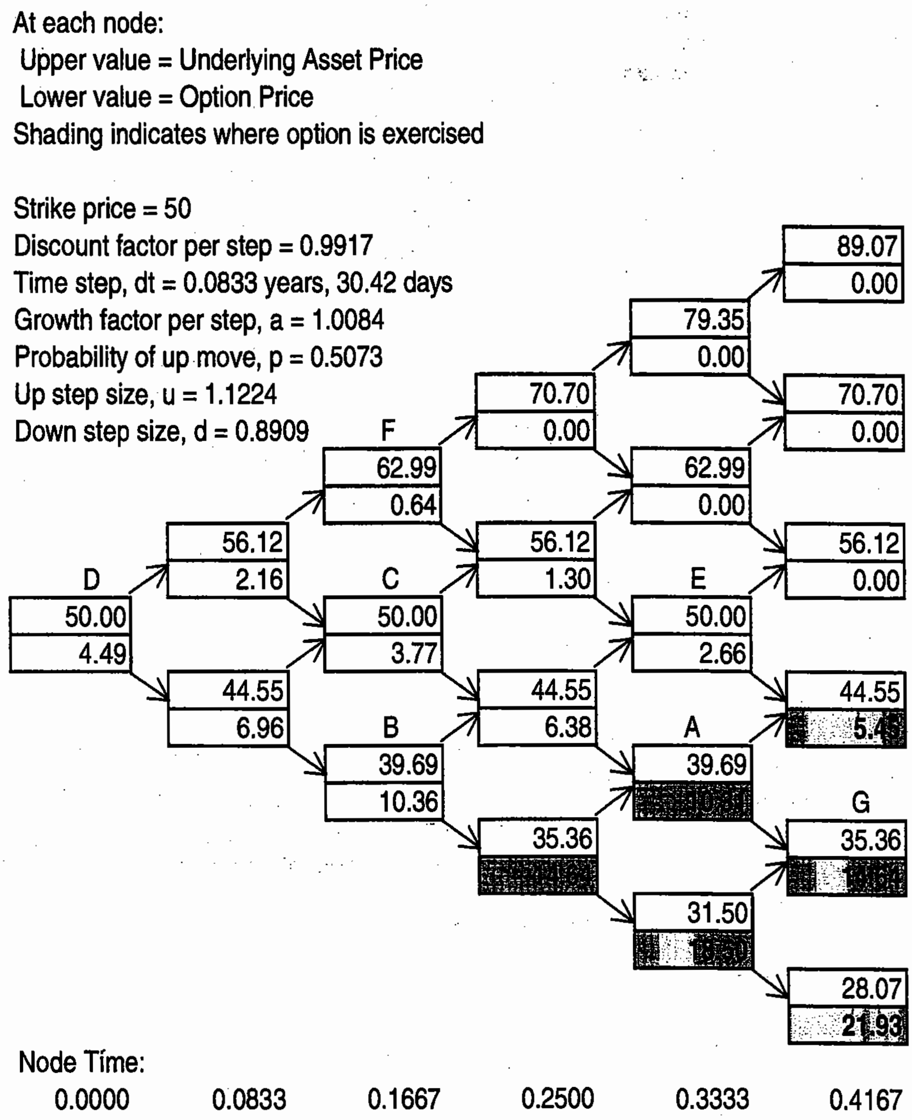
Consider a 5-month American put option on a non-dividend-paying stock when the stock price is \$50, the strike price is \$50, the risk-free interest rate is 10% per annum, and the volatility is 40% per annum. With our usual notation, this means that $S_0 = 50$, $K = 50$, $r = 0.10$, $\sigma = 0.40$, $T = 0.4167$, and $q = 0$. Suppose that we divide the life of the option into five intervals of length 1 month (= 0.0833 year) for the purposes of constructing a binomial tree. Then $\Delta t = 0.0833$ and using equations (19.4) to (19.7) gives

$$u = e^{\sigma\sqrt{\Delta t}} = 1.1224, \qquad d = e^{-\sigma\sqrt{\Delta t}} = 0.8909, \qquad a = e^{r\Delta t} = 1.0084$$

$$p = \frac{a-d}{u-d} = 0.5073, \qquad 1 - p = 0.4927$$

Figure 19.3 shows the binomial tree produced by DerivaGem. At each node there are two numbers. The top one shows the stock price at the node; the lower one shows the value of the option at the node. The probability of an up movement is always 0.5073; the probability of a down movement is always 0.4927.

Figure 19.3 Binomial tree from DerivaGem for American put on non-dividend-paying stock (Example 19.1).



The stock price at the jth node ($j = 0, 1, \dots, i$) at time $i\,\Delta t$ ($i = 0, 1, \dots, 5$) is calculated as $S_0 u^j d^{i-j}$. For example, the stock price at node A ($i = 4$, $j = 1$) (i.e., the second node up at the end of the fourth time step) is $50 \times 1.1224 \times 0.8909^3 =$ \$39.69. The option prices at the final nodes are calculated as $\max(K - S_T, 0)$. For example, the option price at node G is $50.00 - 35.36 = 14.64$. The option prices at the penultimate nodes are calculated from the option prices at the final nodes. First, we assume no exercise of the option at the nodes. This means that the option price is calculated as the present value of the expected option price one time step later. For example, at node E, the option price is calculated as

$$(0.5073 \times 0 + 0.4927 \times 5.45)e^{-0.10 \times 0.0833} = 2.66$$

whereas at node A it is calculated as

$$(0.5073 \times 5.45 + 0.4927 \times 14.64)e^{-0.10 \times 0.0833} = 9.90$$

We then check to see if early exercise is preferable to waiting. At node E, early exercise would give a value for the option of zero because both the stock price and strike price are \$50. Clearly it is best to wait. The correct value for the option at node E, therefore, is \$2.66. At node A, it is a different story. If the option is exercised, it is worth \$50.00 – \$39.69, or \$10.31. This is more than \$9.90. If node A is reached, then the option should be exercised and the correct value for the option at node A is \$10.31.

Option prices at earlier nodes are calculated in a similar way. Note that it is not always best to exercise an option early when it is in the money. Consider node B. If the option is exercised, it is worth \$50.00 – \$39.69, or \$10.31. However, if it is not exercised, it is worth

$$(0.5073 \times 6.38 + 0.4927 \times 14.64)e^{-0.10\times 0.0833} = 10.36$$

The option should, therefore, not be exercised at this node, and the correct option value at the node is \$10.36.

Working back through the tree, the value of the option at the initial node is \$4.49. This is our numerical estimate for the option's current value. In practice, a smaller value of Δt, and many more nodes, would be used. DerivaGem shows that with 30, 50, 100, and 500 time steps we get values for the option of 4.263, 4.272, 4.278, and 4.283.

Expressing the Approach Algebraically

Suppose that the life of an American put option on a non-dividend-paying stock is divided into N subintervals of length Δt. We will refer to the jth node at time $i\,\Delta t$ as the (i, j) node, where $0 \leqslant i \leqslant N$ and $0 \leqslant j \leqslant i$). Define $f_{i,j}$ as the value of the option at the (i, j) node. The stock price at the (i, j) node is $S_0 u^j d^{i-j}$. Since the value of an American put at its expiration date is $\max(K - S_T, 0)$, we know that

$$f_{N,j} = \max(K - S_0 u^j d^{N-j}, 0), \quad j = 0, 1, \ldots, N$$

There is a probability p of moving from the (i, j) node at time $i\,\Delta t$ to the $(i+1, j+1)$ node at time $(i+1)\,\Delta t$, and a probability $1-p$ of moving from the (i, j) node at time $i\,\Delta t$ to the $(i+1, j)$ node at time $(i+1)\,\Delta t$. Assuming no early exercise, risk-neutral valuation gives

$$f_{i,j} = e^{-r\Delta t}[p f_{i+1,j+1} + (1-p) f_{i+1,j}]$$

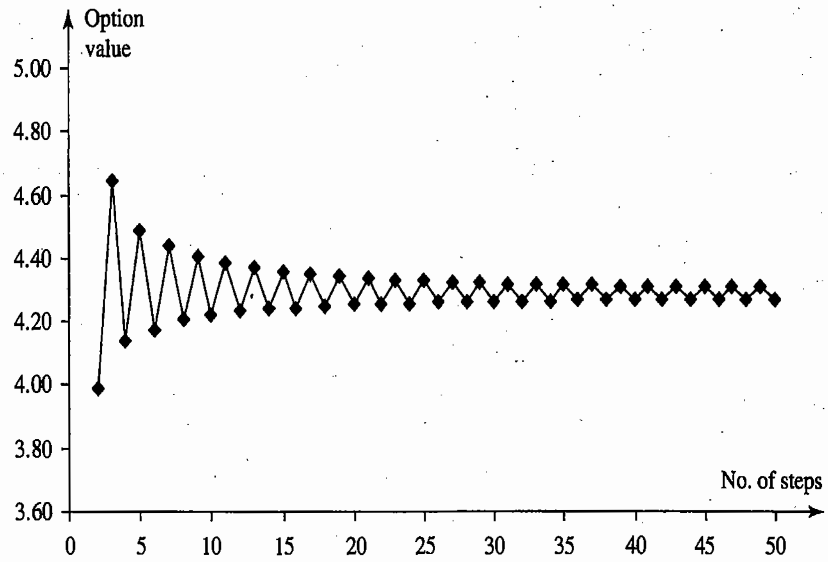
for $0 \leqslant i \leqslant N-1$ and $0 \leqslant j \leqslant i$. When early exercise is taken into account, this value for $f_{i,j}$ must be compared with the option's intrinsic value, so that

$$f_{i,j} = \max\{K - S_0 u^j d^{i-j},\ e^{-r\Delta t}[p f_{i+1,j+1} + (1-p) f_{i+1,j}]\}$$

Note that, because the calculations start at time T and work backward, the value at time $i\,\Delta t$ captures not only the effect of early exercise possibilities at time $i\,\Delta t$, but also the effect of early exercise at subsequent times.

In the limit as Δt tends to zero, an exact value for the American put is obtained. In practice, $N = 30$ usually gives reasonable results. Figure 19.4 shows the convergence of the option price in the example we have been considering. This figure was calculated

Figure 19.4 Convergence of the price of the option in Example 19.1 calculated from the DerivaGem Application Builder functions.



using the Application Builder functions provided with the DerivaGem software (see Sample Application A).

Estimating Delta and Other Greek Letters

It will be recalled that the delta (Δ) of an option is the rate of change of its price with respect to the underlying stock price. It can be calculated as

$$\frac{\Delta f}{\Delta S}$$

where ΔS is a small change in the stock price and Δf is the corresponding small change in the option price. At time Δt, we have an estimate $f_{1,1}$ for the option price when the stock price is $S_0 u$ and an estimate $f_{1,0}$ for the option price when the stock price is $S_0 d$. In other words, when $\Delta S = S_0 u - S_0 d$, $\Delta f = f_{1,1} - f_{1,0}$. Therefore an estimate of delta at time Δt is

$$\Delta = \frac{f_{1,1} - f_{1,0}}{S_0 u - S_0 d} \tag{19.8}$$

To determine gamma (Γ), note that we have two estimates of Δ at time $2\Delta t$. When $S = (S_0 u^2 + S_0)/2$ (halfway between the second and third node), delta is $(f_{2,2} - f_{2,1})/(S_0 u^2 - S_0)$; when $S = (S_0 + S_0 d^2)/2$ (halfway between the first and second node), delta is $(f_{2,1} - f_{2,0})/(S_0 - S_0 d^2)$. The difference between the two values of S is h, where

$$h = 0.5(S_0 u^2 - S_0 d^2)$$

Gamma is the change in delta divided by h:

$$\Gamma = \frac{[(f_{2,2} - f_{2,1})/(S_0 u^2 - S_0)] - [(f_{2,1} - f_{2,0})/(S_0 - S_0 d^2)]}{h} \tag{19.9}$$

These procedures provide estimates of delta at time Δt and of gamma at time $2\Delta t$. In practice, they are usually used as estimates of delta and gamma at time zero as well.[5]

A further hedge parameter that can be obtained directly from the tree is theta (Θ). This is the rate of change of the option price with time when all else is kept constant. If the tree starts at time zero, an estimate of theta is

$$\Theta = \frac{f_{2,1} - f_{0,0}}{2\Delta t} \tag{19.10}$$

Vega can be calculated by making a small change, $\Delta\sigma$, in the volatility and constructing a new tree to obtain a new value of the option. (The time step Δt should be kept the same.) The estimate of vega is

$$\mathcal{V} = \frac{f^* - f}{\Delta\sigma}$$

where f and f^* are the estimates of the option price from the original and the new tree, respectively. Rho can be calculated similarly.

Example 19.2

Consider again Example 19.1. From Figure 19.3, $f_{1,0} = 6.96$ and $f_{1,1} = 2.16$. Equation (19.8) gives an estimate for delta of

$$\frac{2.16 - 6.96}{56.12 - 44.55} = -0.41$$

From equation (19.9), an estimate of the gamma of the option can be obtained from the values at nodes B, C, and F as

$$\frac{[(0.64 - 3.77)/(62.99 - 50.00)] - [(3.77 - 10.36)/(50.00 - 39.69)]}{11.65} = 0.03$$

From equation (19.10), an estimate of the theta of the option can be obtained from the values at nodes D and C as

$$\frac{3.77 - 4.49}{0.1667} = -4.3 \quad \text{per year}$$

or -0.012 per calendar day. These are only rough estimates. They become progressively better as the number of time steps on the tree is increased. Using 50 time steps, DerivaGem provides estimates of -0.415, 0.034, and -0.0117 for delta, gamma, and theta, respectively. By making small changes to parameters and recomputing values, vega and rho are estimated as 0.123 and -0.072, respectively.

19.2 USING THE BINOMIAL TREE FOR OPTIONS ON INDICES, CURRENCIES, AND FUTURES CONTRACTS

As explained in Chapters 11, 15 and 16, stock indices, currencies, and futures contracts can, for the purposes of option valuation, be considered as assets providing known

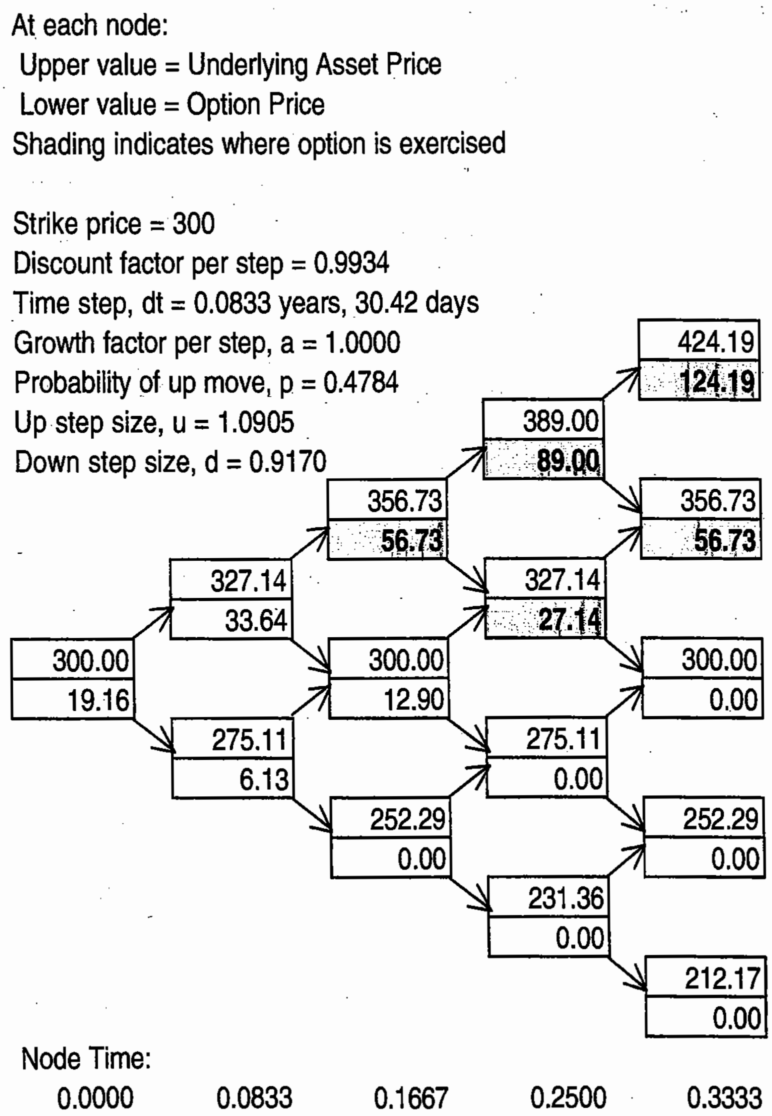
[5] If slightly more accuracy is required for delta and gamma, we can start the binomial tree at time $-2\Delta t$ and assume that the stock price is S_0 at this time. This leads to the option price being calculated for three different stock prices at time zero.

yields. For a stock index, the relevant yield is the dividend yield on the stock portfolio underlying the index; in the case of a currency, it is the foreign risk-free interest rate; in the case of a futures contract, it is the domestic risk-free interest rate. The binomial tree approach can therefore be used to value options on stock indices, currencies, and futures contracts provided that q in equation (19.7) is interpreted appropriately.

Example 19.3

Consider a 4-month American call option on index futures where the current futures price is 300, the exercise price is 300, the risk-free interest rate is 8% per annum, and the volatility of the index is 30% per annum. The life of the option is divided into four 1-month periods for the purposes of constructing the tree. In this case, $F_0 = 300$, $K = 300$, $r = 0.08$, $\sigma = 0.3$, $T = 0.3333$, and $\Delta t = 0.0833$. Because a futures contract is analogous to a stock paying dividends at a rate r, q should be set equal to r in equation (19.7). This gives $a = 1$. The other parameters

Figure 19.5 Binomial tree produced by DerivaGem for American call option on an index futures contract (Example 19.3).



necessary to construct the tree are

$$u = e^{\sigma\sqrt{\Delta t}} = 1.0905, \qquad d = \frac{1}{u} = 0.9170$$

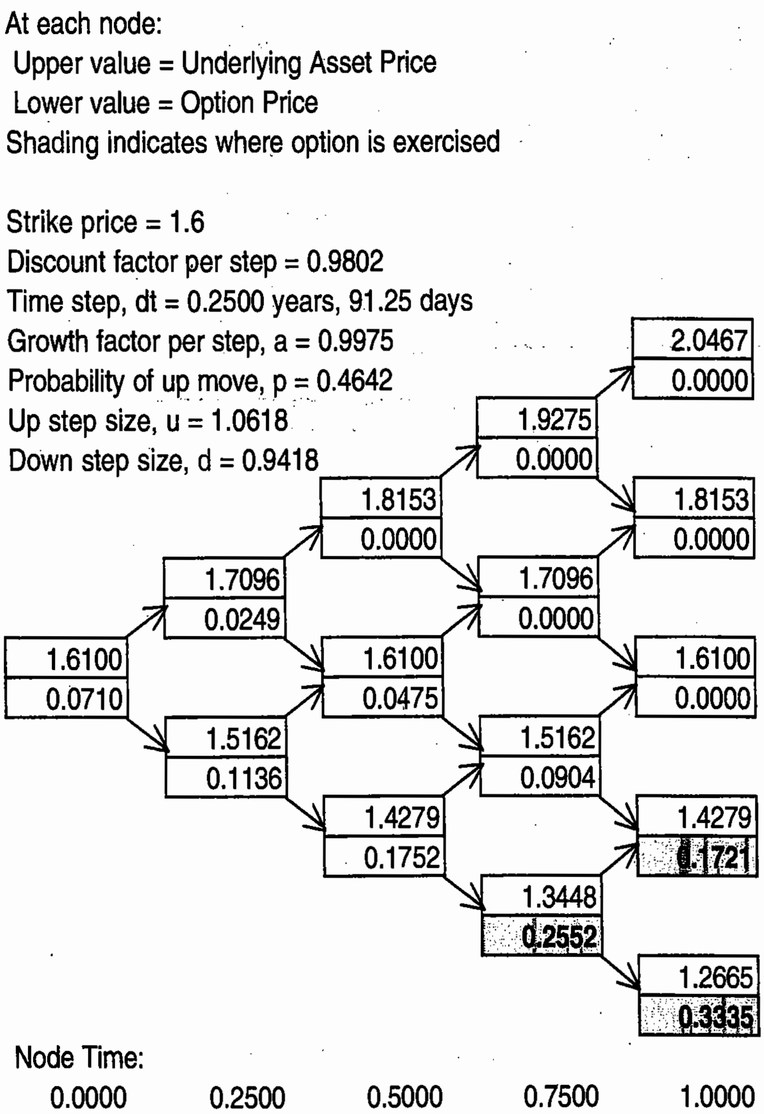
$$p = \frac{a-d}{u-d} = 0.4784, \qquad 1 - p = 0.5216$$

The tree, as produced by DerivaGem, is shown in Figure 19.5. (The upper number is the futures price; the lower number is the option price.) The estimated value of the option is 19.16. More accuracy is obtained using more steps. With 50 time steps, DerivaGem gives a value of 20.18; with 100 time steps it gives 20.22.

Example 19.4

Consider a 1-year American put option on the British pound. The current exchange rate is 1.6100, the strike price is 1.6000, the US risk-free interest rate is 8%

Figure 19.6 Binomial tree produced by DerivaGem for American put option on a currency (Example 19.4).



per annum, the sterling risk-free interest rate is 9% per annum, and the volatility of the sterling exchange rate is 12% per annum. In this case, $S_0 = 1.61$, $K = 1.60$, $r = 0.08$, $r_f = 0.09$, $\sigma = 0.12$, and $T = 1.0$. The life of the option is divided into four 3-month periods for the purposes of constructing the tree, so that $\Delta t = 0.25$. In this case, $q = r_f$ and equation (19.7) gives

$$a = e^{(0.08-0.09)\times 0.25} = 0.9975$$

The other parameters necessary to construct the tree are

$$u = e^{\sigma\sqrt{\Delta t}} = 1.0618, \qquad d = \frac{1}{u} = 0.9418$$

$$p = \frac{a-d}{u-d} = 0.4642, \qquad 1 - p = 0.5358$$

The tree, as produced by DerivaGem, is shown in Figure 19.6. (The upper number is the exchange rate; the lower number is the option price.) The estimated value of the option is \$0.0710. (Using 50 time steps, DerivaGem gives the value of the option as 0.0738; with 100 time steps it also gives 0.0738.)

19.3 BINOMIAL MODEL FOR A DIVIDEND-PAYING STOCK

We now move on to the more tricky issue of how the binomial model can be used for a dividend-paying stock. As in Chapter 13, the word *dividend* will, for the purposes of our discussion, be used to refer to the reduction in the stock price on the ex-dividend date as a result of the dividend.

Known Dividend Yield

If it is assumed that there is a single dividend, and the dividend yield (i.e., the dividend as a percentage of the stock price) is known, the tree takes the form shown in Figure 19.7 and can be analyzed in similar manner to that just described. If the time $i\,\Delta t$ is prior to the stock going ex-dividend, the nodes on the tree correspond to stock prices

$$S_0 u^j d^{i-j}, \quad j = 0, 1, \dots, i$$

where u and d are defined as in equations (19.5) and (19.6). If the time $i\,\Delta t$ is after the stock goes ex-dividend, the nodes correspond to stock prices

$$S_0(1-\delta)u^j d^{i-j}, \quad j = 0, 1, \dots, i$$

where δ is the dividend yield. Several known dividend yields during the life of an option can be dealt with similarly. If δ_i is the total dividend yield associated with all ex-dividend dates between time zero and time $i\,\Delta t$, the nodes at time $i\,\Delta t$ correspond to stock prices

$$S_0(1-\delta_i)u^j d^{i-j}$$

Known Dollar Dividend

In some situations, the most realistic assumption is that the dollar amount of the dividend rather than the dividend yield is known in advance. If the volatility of the stock, σ, is assumed constant, the tree then takes the form shown in Figure 19.8. It does not recombine, which means that the number of nodes that have to be evaluated, particularly if there are several dividends, is liable to become very large. Suppose that there is only one dividend, that the ex-dividend date, τ, is between $k\,\Delta t$ and $(k+1)\,\Delta t$, and that the dollar amount of the dividend is D. When $i \leqslant k$, the nodes on the tree at time $i\,\Delta t$ correspond to stock prices

$$S_0 u^j d^{i-j}, \quad j = 0, 1, 2, \dots, i$$

as before. When $i = k+1$, the nodes on the tree correspond to stock prices

$$S_0 u^j d^{i-j} - D, \quad j = 0, 1, 2, \dots, i$$

When $i = k+2$, the nodes on the tree correspond to stock prices

$$(S_0 u^j d^{i-1-j} - D)u \quad \text{and} \quad (S_0 u^j d^{i-1-j} - D)d$$

for $j = 0, 1, 2, \dots, i-1$, so that there are $2i$ rather than $i+1$ nodes. When $i = k+m$, there are $m(k+2)$ rather than $k+m+1$ nodes.

The problem can be simplified by assuming, as in the analysis of European options in Section 13.12, that the stock price has two components: a part that is uncertain and a part that is the present value of all future dividends during the life of the option. Suppose, as before, that there is only one ex-dividend date, τ, during the life of the

Figure 19.7 Tree when stock pays a known dividend yield at one particular time.

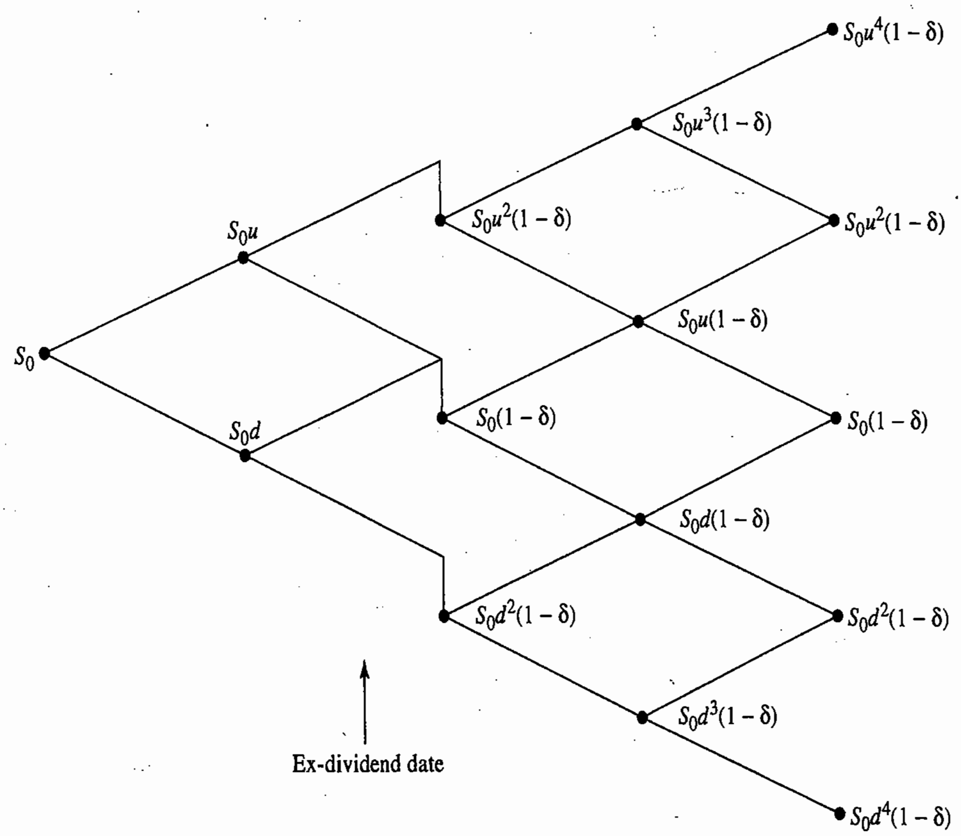
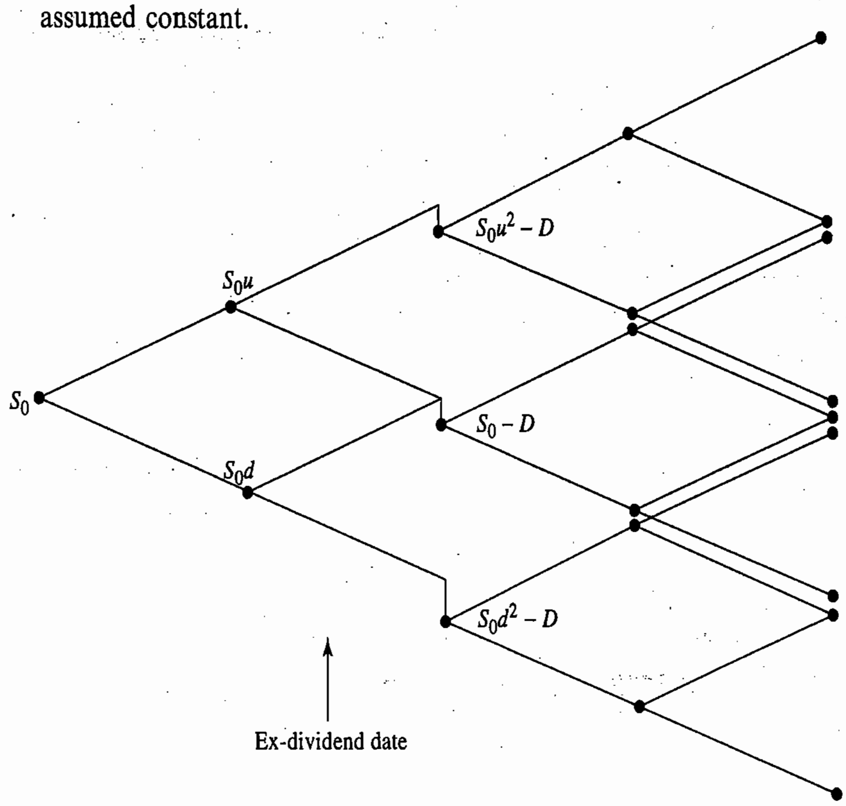


Figure 19.8 Tree when dollar amount of dividend is assumed known and volatility is assumed constant.



option and that $k\,\Delta t \leqslant \tau \leqslant (k+1)\,\Delta t$. The value of the uncertain component, S^*, at time $i\,\Delta t$ is given by

$$S^* = S \quad \text{when } i\,\Delta t > \tau$$

and

$$S^* = S - De^{-r(\tau - i\Delta t)} \quad \text{when } i\,\Delta t \leqslant \tau$$

where D is the dividend. Define σ^* as the volatility of S^* and assume that σ^* is constant.[6] The parameters p, u, and d can be calculated from equations (19.4), (19.5), (19.6), and (19.7) with σ replaced by σ^* and a tree can be constructed in the usual way to model S^*. By adding to the stock price at each node, the present value of future dividends (if any), the tree can be converted into another tree that models S. Suppose that S_0^* is the value of S^* at time zero. At time $i\,\Delta t$, the nodes on this tree correspond to the stock prices

$$S_0^* u^j d^{i-j} + De^{-r(\tau - i\,\Delta t)}, \quad j = 0, 1, \ldots, i$$

when $i\,\Delta t < \tau$ and

$$S_0^* u^j d^{i-j}, \quad j = 0, 1, \ldots, i$$

when $i\,\Delta t > \tau$. This approach, which has the advantage of being consistent with the approach for European options in Section 13.12, succeeds in achieving a situation where

[6] As mentioned in Section 13.12, σ^* is in theory slightly greater than σ, the volatility of S. In practice, the use of implied volatilities avoids the need for analysts to distinguish between σ and σ^*.

the tree recombines so that there are $i+1$ nodes at time $i\,\Delta t$. It can be generalized in a straightforward way to deal with the situation where there are several dividends.

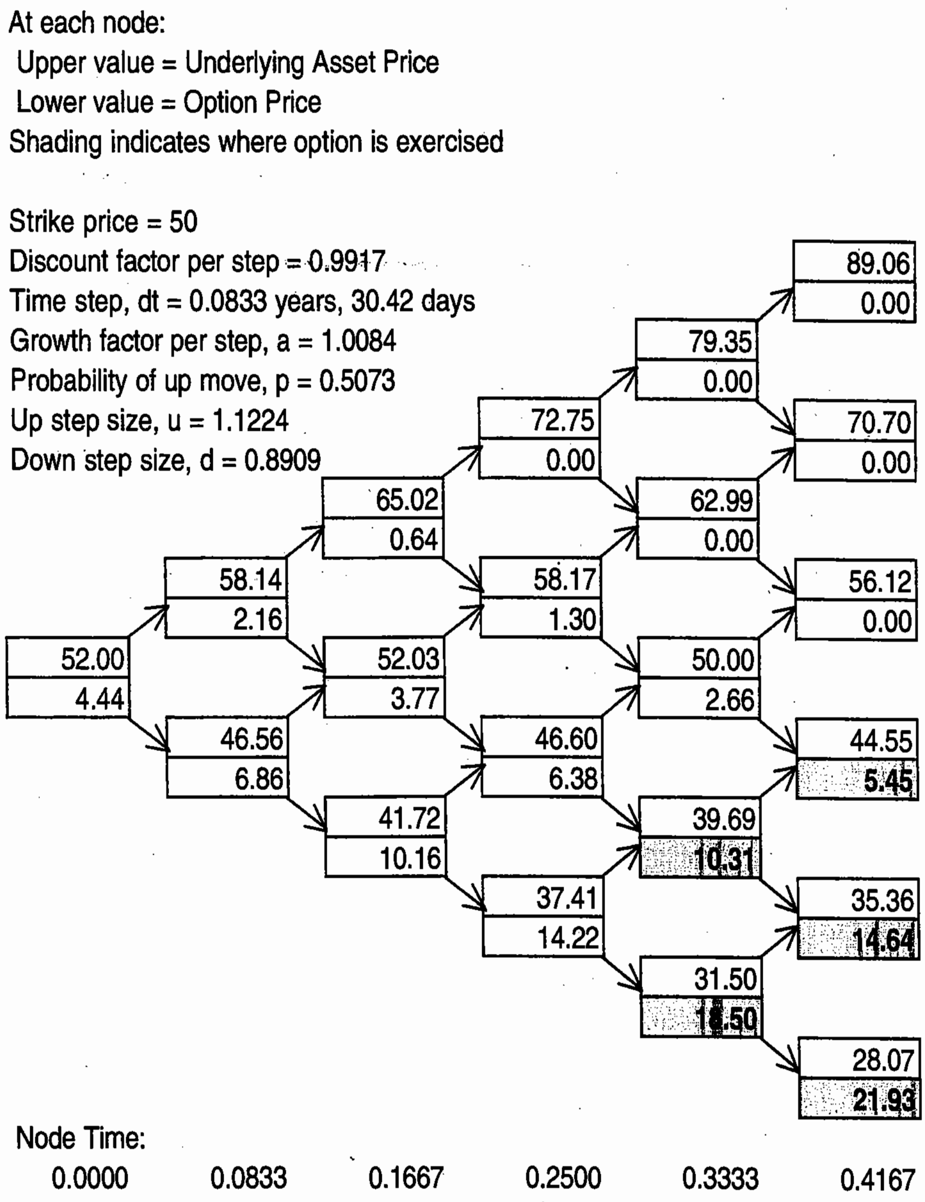
Example 19.5

Consider a 5-month American put option on a stock that is expected to pay a single dividend of \$2.06 during the life of the option. The initial stock price is \$52, the strike price is \$50, the risk-free interest rate is 10% per annum, the volatility is 40% per annum, and the ex-dividend date is in $3\frac{1}{2}$ months.

We first construct a tree to model S^*, the stock price less the present value of future dividends during the life of the option. At time zero, the present value of the dividend is

$$2.06e^{-0.2917\times 0.1} = 2.00$$

Figure 19.9 Tree produced by DerivaGem for Example 19.5.



The initial value of S^* is therefore 50.00. If we assume that the 40% per annum volatility refers to S^*, then Figure 19.3 provides a binomial tree for S^*. (This is because S^* has the same initial value and volatility as the stock price that Figure 19.3 was based upon.) Adding the present value of the dividend at each node leads to Figure 19.9, which is a binomial model for S. The probabilities at each node are, as in Figure 19.3, 0.5073 for an up movement and 0.4927 for a down movement. Working back through the tree in the usual way gives the option price as \$4.44. (Using 50 time steps, DerivaGem gives a value for the option of 4.202; using 100 steps it gives 4.212.)

When the option lasts a long time (say, 3 or more years) it is usually more appropriate to assume a known dividend yield rather than a known cash dividend because the latter cannot reasonably be assumed to be the same for all the stock prices that might be encountered in the future.[7] Often for convenience the dividend yield is assumed to be paid continuously. Valuing an option on a dividend paying stock is then similar to valuing an option on a stock index.

Control Variate Technique

A technique known as the *control variate technique* can improve the accuracy of the pricing of an American option.[8] This involves using the same tree to calculate both the value of the American option, f_A, and the value of the corresponding European option, f_E. We also calculate the Black–Scholes price of the European option, f_{BS}. The error given by the tree in the pricing of the European option is assumed equal to that given by the tree in the pricing of the American option. This gives the estimate of the price of the American option as

$$f_A + f_{BS} - f_E$$

To illustrate this approach, Figure 19.10 values the option in Figure 19.3 on the assumption that it is European. The price obtained is \$4.32. From the Black–Scholes formula, the true European price of the option is \$4.08. The estimate of the American price in Figure 19.3 is \$4.49. The control variate estimate of the American price, therefore, is

$$4.49 + 4.08 - 4.32 = 4.25$$

A good estimate of the American price, calculated using 100 steps, is 4.278. The control variate approach does, therefore, produce a considerable improvement over the basic tree estimate of 4.49 in this case.

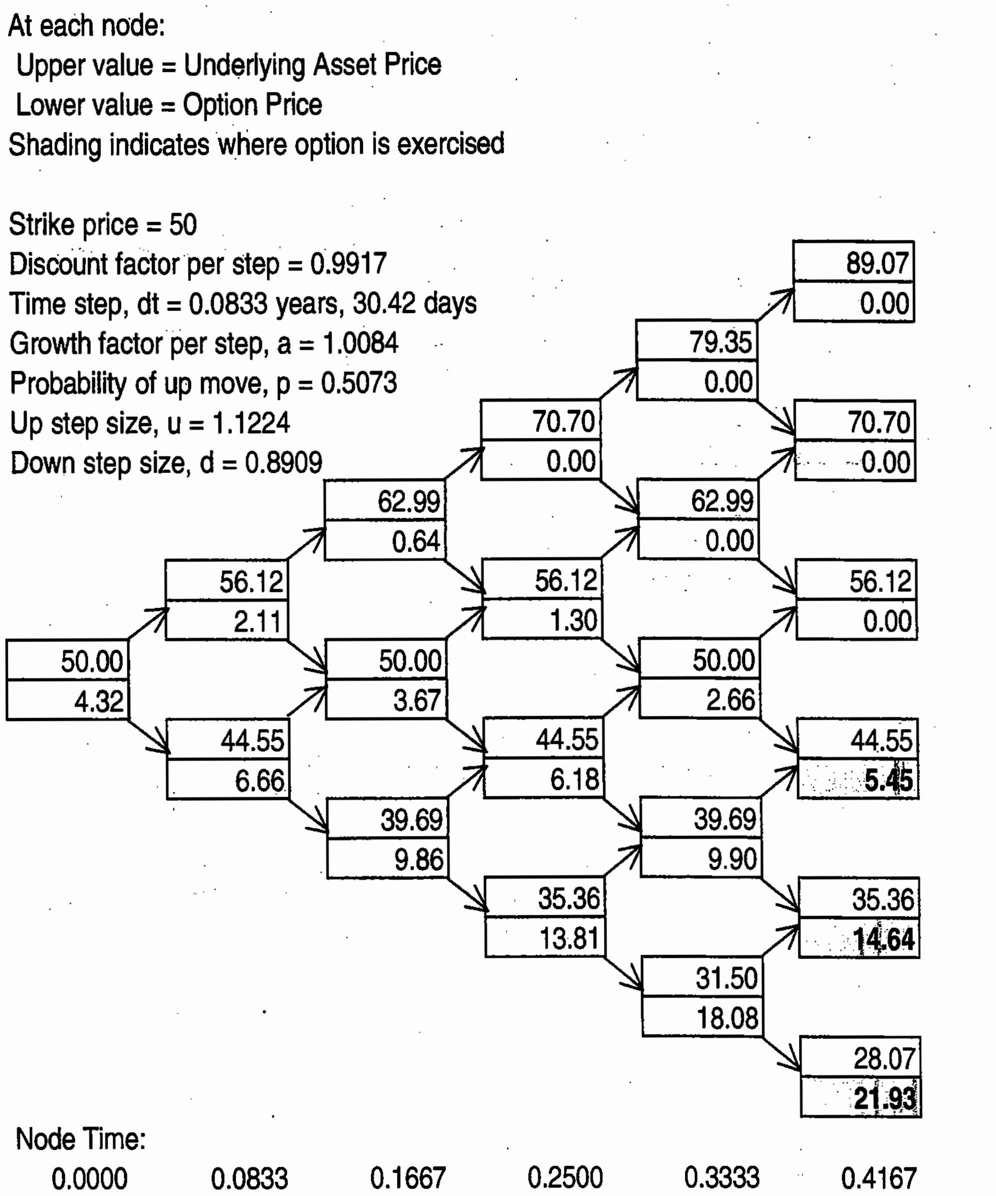
The control variate technique in effect involves using the tree to calculate the difference between the European and the American price rather than the American price itself. We give a further application of the control variate technique when we discuss Monte Carlo simulation later in the chapter.

---

[7] Another problem is that, for long-dated options, S^* is significantly less than S_0 and volatility estimates can be very high.

[8] See J. Hull and A. White, "The Use of the Control Variate Technique in Option Pricing," *Journal of Financial and Quantitative Analysis*, 23 (September 1988): 237–51.

Figure 19.10 Tree, as produced by DerivaGem, for European version of option in Figure 19.3. At each node, the upper number is the stock price, and the lower number is the option price.



19.4 ALTERNATIVE PROCEDURES FOR CONSTRUCTING TREES

The Cox, Ross, and Rubinstein approach is not the only way of building a binomial tree. Instead of imposing the assumption $u = 1/d$ on equations (19.1) and (19.2), we can set $p = 0.5$. A solution to the equations when terms of higher order than Δt are ignored is then

$$u = e^{(r-q-\sigma^2/2)\Delta t + \sigma\sqrt{\Delta t}}$$

$$d = e^{(r-q-\sigma^2/2)\Delta t - \sigma\sqrt{\Delta t}}$$

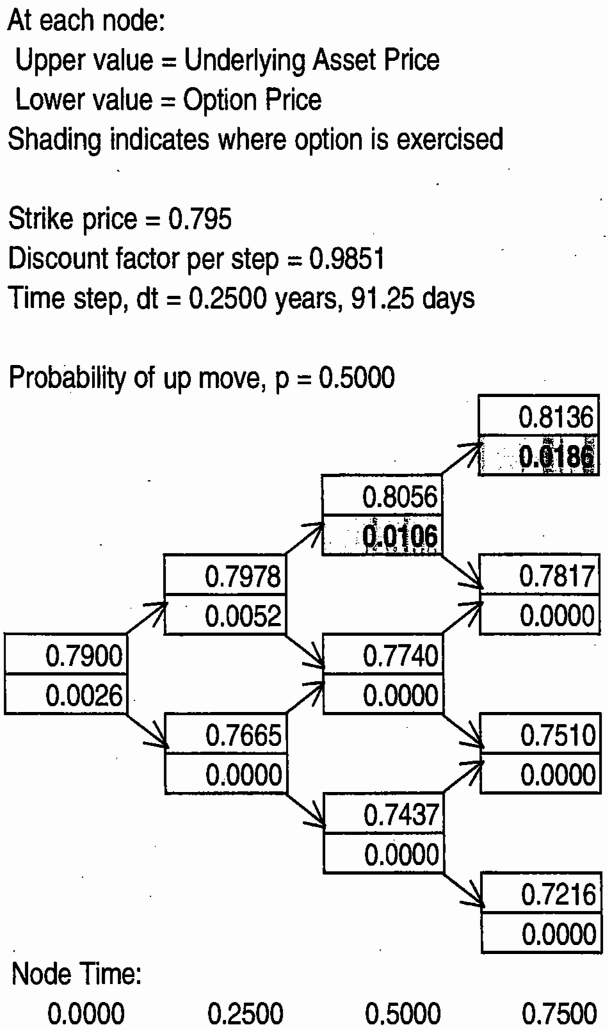
This allows trees with $p = 0.5$ to be built for options on stocks, indices, foreign exchange, and futures.

This alternative tree-building procedure has the advantage over the Cox, Ross, and Rubinstein approach that the probabilities are always 0.5 regardless of the value of σ or the number of time steps.[9] Its disadvantage is that it is not as straightforward to calculate delta, gamma, and rho from the tree because the tree is no longer centered at the initial stock price.

Example 19.6

Consider a 9-month American call option on the Canadian dollar. The current exchange rate is 0.7900, the strike price is 0.7950, the US risk-free interest rate is 6% per annum, the Canadian risk-free interest rate is 10% per annum, and the

Figure 19.11 Binomial tree for American call option on the Canadian dollar. At each node, upper number is spot exchange rate and lower number is option price. All probabilities are 0.5.



[9] When time steps are so large that $\sigma < |(r - q)\sqrt{\Delta t}|$, the Cox, Ross, and Rubinstein tree gives negative probabilities. The alternative procedure described here does not have that drawback.

volatility of the exchange rate is 4% per annum. In this case, $S_0 = 0.79$, $K = 0.795$, $r = 0.06$, $r_f = 0.10$, $\sigma = 0.04$, and $T = 0.75$. We divide the life of the option into 3-month periods for the purposes of constructing the tree, so that $\Delta t = 0.25$. We set the probabilities on each branch to 0.5 and

$$u = e^{(0.06-0.10-0.0016/2)0.25+0.04\sqrt{0.25}} = 1.0098$$

$$d = e^{(0.06-0.10-0.0016/2)0.25-0.04\sqrt{0.25}} = 0.9703$$

The tree for the exchange rate is shown in Figure 19.11. The tree gives the value of the option as \$0.0026.

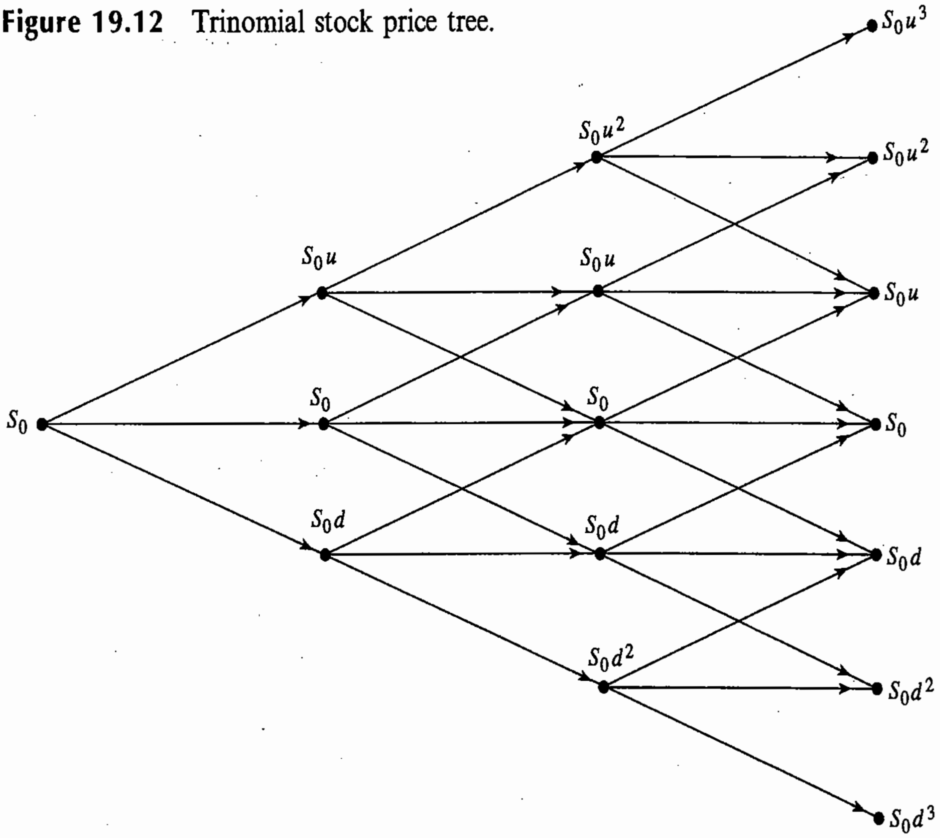
Trinomial Trees

Trinomial trees can be used as an alternative to binomial trees. The general form of the tree is as shown in Figure 19.12. Suppose that p_u, p_m, and p_d are the probabilities of up, middle, and down movements at each node and Δt is the length of the time step. For an asset paying dividends at a rate q, parameter values that match the mean and standard deviation of price changes when terms of higher order than Δt are ignored are

$$u = e^{\sigma\sqrt{3\Delta t}}, \qquad d = \frac{1}{u}$$

$$p_d = -\sqrt{\frac{\Delta t}{12\sigma^2}}\left(r - q - \frac{\sigma^2}{2}\right) + \frac{1}{6}, \qquad p_m = \frac{2}{3}, \qquad p_u = \sqrt{\frac{\Delta t}{12\sigma^2}}\left(r - q - \frac{\sigma^2}{2}\right) + \frac{1}{6}$$

Figure 19.12 Trinomial stock price tree.



Calculations for a trinomial tree are analogous to those for a binomial tree. We work from the end of the tree to the beginning. At each node we calculate the value of exercising and the value of continuing. The value of continuing is

$$e^{-r\Delta t}(p_u f_u + p_m f_m + p_d f_d)$$

where f_u, f_m, and f_d are the values of the option at the subsequent up, middle, and down nodes, respectively. The trinomial tree approach proves to be equivalent to the explicit finite difference method, which will be described in Section 19.8.

Figlewski and Gao have proposed an enhancement of the trinomial tree method, which they call the *adaptive mesh model*. In this, a high-resolution (small-Δt) tree is grafted onto a low-resolution (large-Δt) tree.[10] When valuing a regular American option, high resolution is most useful for the parts of the tree close to the strike price at the end of the life of the option.

19.5 TIME-DEPENDENT PARAMETERS

Up to now we have assumed that r, q, r_f, and σ are constants. In practice, they are usually assumed to be time dependent. The values of these variables between times t and $t + \Delta t$ are assumed to be equal to their forward values.[11]

To make r and q (or r_f) a function of time in a Cox–Ross–Rubinstein binomial tree, we set

$$a = e^{[f(t)-g(t)]\Delta t} \tag{19.11}$$

for nodes at time t, where $f(t)$ is the forward interest rate between times t and $t + \Delta t$ and $g(t)$ is the forward value of q between these times. This does not change the geometry of the tree because u and d do not depend on a. The probabilities on the branches emanating from nodes at time t are:[12]

$$p = \frac{e^{[f(t)-g(t)]\Delta t} - d}{u - d} \tag{19.12}$$

$$1 - p = \frac{u - e^{[f(t)-g(t)]\Delta t}}{u - d}$$

The rest of the way that we use the tree is the same as before, except that when discounting between times t and $t + \Delta t$ we use $f(t)$.

Making σ a function of time in a binomial tree is more challenging. One approach is to make the lengths of time steps inversely proportional to the variance rate. The values of u and d are then always the same and the tree recombines. Suppose that $\sigma(t)$ is the volatility for a maturity t so that $\sigma(t)^2 t$ is the cumulative variance by time t. Define $V = \sigma(T)^2 T$, where T is the life of the tree, and let t_i be the end of the ith time step. If there is a total

[10] See S. Figlewski and B. Gao, "The Adaptive Mesh Model: A New Approach to Efficient Option Pricing," *Journal of Financial Economics*, 53 (1999): 313–51.

[11] The forward dividend yield and forward variance rate are calculated in the same way as the forward interest rate. (The variance rate is the square of the volatility.)

[12] For a sufficiently large number of time steps, these probabilities are always positive.

Business Snapshot 19.1 Calculating Pi with Monte Carlo Simulation

Suppose the sides of the square in Figure 19.13 are one unit in length. Imagine that you fire darts randomly at the square and calculate the percentage that lie in the circle. What should you find? The square has an area of 1.0 and the circle has a radius of 0.5 The area of the circle is π times the radius squared or $\pi/4$. It follows that the proportion of darts that lie in the circle should be $\pi/4$. We can estimate π by multiplying the proportion that lie in the circle by 4.

We can use an Excel spreadsheet to simulate the dart throwing as illustrated in Table 19.1. We define both cell A1 and cell B1 as =RAND(). A1 and B1 are random numbers between 0 and 1 and define how far to the right and how high up the dart lands in the square in Figure 19.13. We then define cell C1 as

=IF((A1−0.5)^2+(B1−0.5)^2<0.5^2,4,0)

This has the effect of setting C1 equal to 4 if the dart lies in the circle and 0 otherwise.

Define the next 99 rows of the spreadsheet similarly to the first one. (This is a "select and drag" operation in Excel.) Define C102 as =AVERAGE(C1:C100) and C103 as =STDEV(C1:C100). C102 (which is 3.04 in Table 19.1) is an estimate of π calculated from 100 random trials. C103 is the standard deviation of our results and as we will see in Example 19.7 can be used to assess the accuracy of the estimate. Increasing the number of trials improves accuracy—but convergence to the correct value of 3.14159 is slow.

of N time steps, we choose t_i to satisfy $\sigma(t_i)^2 t_i = iV/N$. The variance between times t_{i-1} and t_i is then V/N for all i.

With a trinomial tree, a generalized tree-building procedure can be used to match time-dependent interest rates and volatilities (see Technical Note 9 on the author's website).

19.6 MONTE CARLO SIMULATION

We now explain Monte Carlo simulation, a quite different approach for valuing derivatives from binomial trees. Business Snapshot 19.1 illustrates the random sampling idea underlying Monte Carlo simulation by showing how a simple Excel program can be constructed to estimate π.

Figure 19.13 Calculation of π by throwing darts.

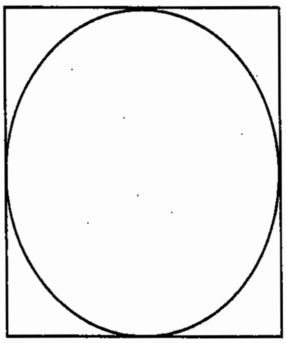


Table 19.1 Sample spreadsheet calculations in Business Snapshot 19.1.

	A	*B*	*C*
1	0.207	0.690	4
2	0.271	0.520	4
3	0.007	0.221	0
⋮	⋮	⋮	⋮
100	0.198	0.403	4
101			
102		Mean:	3.04
103		SD:	1.69

When used to value an option, Monte Carlo simulation uses the risk-neutral valuation result. We sample paths to obtain the expected payoff in a risk-neutral world and then discount this payoff at the risk-free rate. Consider a derivative dependent on a single market variable S that provides a payoff at time T. Assuming that interest rates are constant, we can value the derivative as follows:

1. Sample a random path for S in a risk-neutral world.
2. Calculate the payoff from the derivative.
3. Repeat steps 1 and 2 to get many sample values of the payoff from the derivative in a risk-neutral world.
4. Calculate the mean of the sample payoffs to get an estimate of the expected payoff in a risk-neutral world.
5. Discount this expected payoff at the risk-free rate to get an estimate of the value of the derivative.

Suppose that the process followed by the underlying market variable in a risk-neutral world is

$$dS = \hat{\mu} S\,dt + \sigma S\,dz \tag{19.13}$$

where dz is a Wiener process, $\hat{\mu}$ is the expected return in a risk-neutral world, and σ is the volatility.[13] To simulate the path followed by S, we can divide the life of the derivative into N short intervals of length Δt and approximate equation (19.13) as

$$S(t+\Delta t) - S(t) = \hat{\mu} S(t)\,\Delta t + \sigma S(t)\epsilon\sqrt{\Delta t} \tag{19.14}$$

where $S(t)$ denotes the value of S at time t, ϵ is a random sample from a normal distribution with mean zero and standard deviation of 1.0. This enables the value of S at time Δt to be calculated from the initial value of S, the value at time $2\,\Delta t$ to be calculated from the value at time Δt, and so on. An illustration of the procedure is in Section 12.3. One simulation trial involves constructing a complete path for S using N random samples from a normal distribution.

[13] If S is the price of a non-dividend-paying stock then $\hat{\mu} = r$, if it is an exchange rate then $\hat{\mu} = r - r_f$, and so on. Note that the volatility is the same in a risk-neutral world as in the real world, as illustrated in Section 11.7.

In practice, it is usually more accurate to simulate $\ln S$ rather than S. From Itô's lemma the process followed by $\ln S$ is

$$d\ln S = \left(\hat{\mu} - \frac{\sigma^2}{2}\right)dt + \sigma\,dz \tag{19.15}$$

so that

$$\ln S(t+\Delta t) - \ln S(t) = \left(\hat{\mu} - \frac{\sigma^2}{2}\right)\Delta t + \sigma\epsilon\sqrt{\Delta t}$$

or equivalently

$$S(t+\Delta t) = S(t)\exp\left[\left(\hat{\mu} - \frac{\sigma^2}{2}\right)\Delta t + \sigma\epsilon\sqrt{\Delta t}\right] \tag{19.16}$$

This equation is used to construct a path for S.

Working with $\ln S$ rather than S gives more accuracy. Also, if $\hat{\mu}$ and σ are constant, then

$$\ln S(T) - \ln S(0) = \left(\hat{\mu} - \frac{\sigma^2}{2}\right)T + \sigma\epsilon\sqrt{T}$$

is true for all T.[14] It follows that

$$S(T) = S(0)\exp\left[\left(\hat{\mu} - \frac{\sigma^2}{2}\right)T + \sigma\epsilon\sqrt{T}\right] \tag{19.17}$$

This equation can be used to value derivatives that provide a nonstandard payoff at time T. As shown in Business Snapshot 19.2, it can also be used to check the Black–Scholes formulas.

The key advantage of Monte Carlo simulation is that it can be used when the payoff depends on the path followed by the underlying variable S as well as when it depends only on the final value of S. (For example, it can be used when payoffs depend on the average value of S.) Payoffs can occur at several times during the life of the derivative rather than all at the end. Any stochastic process for S can be accommodated. As will be shown shortly, the procedure can also be extended to accommodate situations where the payoff from the derivative depends on several underlying market variables. The drawbacks of Monte Carlo simulation are that it is computationally very time consuming and cannot easily handle situations where there are early exercise opportunities.[15]

Derivatives Dependent on More than One Market Variable

Consider the situation where the payoff from a derivative depends on n variables θ_i $(1 \leqslant i \leqslant n)$. Define s_i as the volatility of θ_i, $\hat{m}_i$ as the expected growth rate of θ_i in a risk-neutral world, and ρ_{ik} as the instantaneous correlation between θ_i and θ_k.[16] As in the single-variable case, the life of the derivative must be divided into N subintervals of

[14] By contrast, equation (19.14) is exactly true only in the limit as Δt tends to zero.

[15] As discussed in Chapter 26, a number of researchers have suggested ways Monte Carlo simulation can be extended to value American options.

[16] Note that s_i, $\hat{m}_i$, and ρ_{ik} are not necessarily constant; they may depend on the θ_i.

Business Snapshot 19.2 Checking Black–Scholes

The Black–Scholes formula for a European call option can be checked by using a binomial tree with a very large number of time steps. An alternative way of checking it is to use Monte Carlo simulation. Table 19.2 shows a spreadsheet that can be constructed. The cells C2, D2, E2, F2, and G2 contain S_0, K, r, σ, and T, respectively. Cells D4, E4, and F4 calculate d_1, d_2, and the Black–Scholes price, respectively. (The Black–Scholes price is 4.817 in the sample spreadsheet.)

NORMSINV is the inverse cumulative function for the standard normal distribution. It follows that NORMSINV(RAND())gives a random sample from a standard normal distribution. We set cell A1 as

```
=$C$2*EXP(($E$2-$F$2*$F$2/2)*$G$2+$F$2*NORMSINV(RAND())*SQRT($G$2))
```

This corresponds to equation (19.17) and is a random sample from the set of all stock prices at time T. We set cell B1 as

```
=EXP(-$E$2*$G$2)*MAX(A1-$D$2,0)
```

This is the present value of the payoff from a call option. We define the next 999 rows of the spreadsheet similarly to the first one. (This is a "select and drag" operation in Excel.) Define B1002 as AVERAGE(B1:B1000) and B1003 as STDEV(B1:B1000). B1002 (which is 4.98 in the sample spreadsheet) is an estimate of the value of the option. This should be not too far from the Black–Scholes price. As we shall see in Example 19.8, B1003 can be used to assess the accuracy of the estimate.

length Δt. The discrete version of the process for θ_i is then

$$\theta_i(t+\Delta t)-\theta_i(t)=\hat{m}_i\theta_i(t)\,\Delta t+s_i\theta_i(t)\epsilon_i\sqrt{\Delta t} \tag{19.18}$$

where ϵ_i is a random sample from a standard normal distribution. The coefficient of correlation between ϵ_i and ϵ_k is ρ_{ik} ($1 \leqslant i; k \leqslant n$). One simulation trial involves obtaining N samples of the ϵ_i ($1 \leqslant i \leqslant n$) from a multivariate standardized normal distribution. These are substituted into equation (19.18) to produce simulated paths for each θ_i, thereby enabling a sample value for the derivative to be calculated.

Table 19.2 Monte Carlo simulation to check Black–Scholes

	A	*B*	*C*	*D*	*E*	*F*	*G*
1	45.95	0	S_0	K	r	σ	T
2	54.49	4.38	50	50	0.05	0.3	0.5
3	50.09	0.09		d_1	d_2	BS price	
4	47.46	0		0.2239	0.0118	4.817	
5	44.93	0					
⋮	⋮	⋮					
1000	68.27	17.82					
1001							
1002	Mean:	4.98					
1003	SD:	7.68					

Generating the Random Samples from Normal Distributions

An approximate sample from a univariate standardized normal distribution can be obtained from the formula

$$\epsilon = \sum_{i=1}^{12} R_i - 6 \tag{19.19}$$

where the R_i ($1 \leqslant i \leqslant 12$) are independent random numbers between 0 and 1, and ϵ is the required sample from $\phi(0, 1)$. This approximation is satisfactory for most purposes. An alternative approach in Excel is to use =NORMSINV(RAND()) as in Business Snapshot 19.2.

When two correlated samples ϵ_1 and ϵ_2 from standard normal distributions are required, an appropriate procedure is as follows. Independent samples x_1 and x_2 from a univariate standardized normal distribution are obtained as just described. The required samples ϵ_1 and ϵ_2 are then calculated as follows:

$$\epsilon_1 = x_1$$
$$\epsilon_2 = \rho x_1 + x_2\sqrt{1-\rho^2}$$

where ρ is the coefficient of correlation.

More generally, consider the situation where we require n correlated samples from normal distributions with the correlation between sample i and sample j being ρ_{ij}. We first sample n independent variables x_i ($1 \leqslant i \leqslant n$), from univariate standardized normal distributions. The required samples, ϵ_i ($1 \leqslant i \leqslant n$), are then defined as follows:

$$\epsilon_1 = \alpha_{11}x_1$$
$$\epsilon_2 = \alpha_{21}x_1 + \alpha_{22}x_2$$
$$\epsilon_3 = \alpha_{31}x_1 + \alpha_{32}x_2 + \alpha_{33}x_3$$

and so on. We choose the coefficients α_{ij} so that the correlations and variances are correct. This can be done step by step as follows. Set $\alpha_{11} = 1$; choose α_{21} so that $\alpha_{21}\alpha_{11} = \rho_{21}$; choose α_{22} so that $\alpha_{21}^2 + \alpha_{22}^2 = 1$; choose α_{31} so that $\alpha_{31}\alpha_{11} = \rho_{31}$; choose α_{32} so that $\alpha_{31}\alpha_{21} + \alpha_{32}\alpha_{22} = \rho_{32}$; choose α_{33} so that $\alpha_{31}^2 + \alpha_{32}^2 + \alpha_{33}^2 = 1$; and so on.[17] This procedure is known as the *Cholesky decomposition.*

Number of Trials

The accuracy of the result given by Monte Carlo simulation depends on the number of trials. It is usual to calculate the standard deviation as well as the mean of the discounted payoffs given by the simulation trials. Denote the mean by μ and the standard deviation by ω. The variable μ is the simulation's estimate of the value of the derivative. The standard error of the estimate is

$$\frac{\omega}{\sqrt{M}}$$

where M is the number of trials. A 95% confidence interval for the price f of the

[17] If the equations for the α's do not have real solutions, the assumed correlation structure is internally inconsistent This will be discussed further in Section 21.7.

derivative is therefore given by

$$\mu - \frac{1.96\omega}{\sqrt{M}} < f < \mu + \frac{1.96\omega}{\sqrt{M}}$$

This shows that uncertainty about the value of the derivative is inversely proportional to the square root of the number of trials. To double the accuracy of a simulation, we must quadruple the number of trials; to increase the accuracy by a factor of 10, the number of trials must increase by a factor of 100; and so on.

Example 19.7

In Table 19.1, π is calculated as the average of 100 numbers. The standard deviation of the numbers is 1.69. In this case, $\omega = 1.69$ and $M = 100$, so that the standard error of the estimate is $1.69/\sqrt{100} = 0.169$. The spreadsheet therefore gives a 95% confidence interval for π as $(3.04 - 1.96 \times 0.169)$ to $(3.04 + 1.96 \times 0.169)$ or 2.71 to 3.37.

Example 19.8

In Table 19.2, the value of the option is calculated as the average of 1000 numbers. The standard deviation of the numbers is 7.68. In this case, $\omega = 7.68$ and $M = 1000$. The standard error of the estimate is $7.68/\sqrt{1000} = 0.24$. The spreadsheet therefore gives a 95% confidence interval for the option value as $(4.98 - 1.96 \times 0.24)$ to $(4.98 + 1.96 \times 0.24)$, or 4.51 to 5.45.

Applications

Monte Carlo simulation tends to be numerically more efficient than other procedures when there are three or more stochastic variables. This is because the time taken to carry out a Monte Carlo simulation increases approximately linearly with the number of variables, whereas the time taken for most other procedures increases exponentially with the number of variables. One advantage of Monte Carlo simulation is that it can provide a standard error for the estimates that it makes. Another is that it is an approach that can accommodate complex payoffs and complex stochastic processes. Also, it can be used when the payoff depends on some function of the whole path followed by a variable, not just its terminal value.

Calculating the Greek Letters

The Greek letters discussed in Chapter 17 can be calculated using Monte Carlo simulation. Suppose that we are interested in the partial derivative of f with respect to x, where f is the value of the derivative and x is the value of an underlying variable or a parameter. First, Monte Carlo simulation is used in the usual way to calculate an estimate $\hat{f}$ for the value of the derivative. A small increase Δx is then made in the value of x, and a new value for the derivative, $\hat{f}^*$, is calculated in the same way as $\hat{f}$. An estimate for the hedge parameter is given by

$$\frac{\hat{f}^* - \hat{f}}{\Delta x}$$

In order to minimize the standard error of the estimate, the number of time intervals, N, the random number streams, and the number of trials, M, should be the same for calculating both $\hat{f}$ and $\hat{f}^*$.

Sampling through a Tree

Instead of implementing Monte Carlo simulation by randomly sampling from the stochastic process for an underlying variable, we can use an N-step binomial tree and sample from the 2^N paths that are possible. Suppose we have a binomial tree where the probability of an "up" movement is 0.6. The procedure for sampling a random path through the tree is as follows. At each node, we sample a random number between 0 and 1. If the number is less than 0.4, we take the down branch. If it is greater than 0.4, we take the up branch. Once we have a complete path from the initial node to the end of the tree, we can calculate a payoff. This completes the first trial. A similar procedure is used to complete more trials. The mean of the payoffs is discounted at the risk-free rate to get an estimate of the value of the derivative.[18]

Example 19.9

Suppose that the tree in Figure 19.3 is used to value an option that pays off $\max(S_{ave} - 50, 0)$, where S_{ave} is the average stock price during the 5 months (with the first and last stock price being included in the average). This is known as an Asian option. When ten simulation trials are used one possible result is shown in Table 19.3.

Table 19.3 Monte Carlo simulation to value Asian option from the tree in Figure 19.3. Payoff is amount by which average stock price exceeds \$50. U = up movement; D = down movement.

Trial	*Path*	*Average stock price*	*Option payoff*
1	UUUUD	64.98	14.98
2	UUUDD	59.82	9.82
3	DDDUU	42.31	0.00
4	UUUUU	68.04	18.04
5	UUDDU	55.22	5.22
6	UDUUD	55.22	5.22
7	DDUDD	42.31	0.00
8	UUDDU	55.22	5.22
9	UUUDU	62.25	12.25
10	DDUUD	45.56	0.00
Average			7.08

The value of the option is calculated as the average payoff discounted at the risk-free rate. In this case, the average payoff is \$7.08 and the risk-free rate is 10% and so the calculated value is $7.08e^{-0.1\times5/12} = 6.79$. (This illustrates the methodology. In practice, we would have to use more time steps on the tree and many more simulation trials to get an accurate answer.)

[18] See D. Mintz, "Less is More," *Risk*, July 1997: 42–45, for a discussion of how sampling through a tree can be made efficient.

19.7 VARIANCE REDUCTION PROCEDURES

If the simulation is carried out as described so far, a very large number of trials is usually necessary to estimate f with reasonable accuracy. This is very expensive in terms of computation time. In this section, we examine a number of variance reduction procedures that can lead to dramatic savings in computation time.

Antithetic Variable Technique

In the antithetic variable technique, a simulation trial involves calculating two values of the derivative. The first value f_1 is calculated in the usual way; the second value f_2 is calculated by changing the sign of all the random samples from standard normal distributions. (If ϵ is a sample used to calculate f_1, then $-\epsilon$ is the corresponding sample used to calculate f_2.) The sample value of the derivative calculated from a simulation trial is the average of f_1 and f_2. This works well because when one value is above the true value, the other tends to be below, and vice versa.

Denote $\bar{f}$ as the average of f_1 and f_2:

$$\bar{f} = \frac{f_1 + f_2}{2}$$

The final estimate of the value of the derivative is the average of the $\bar{f}$'s. If $\bar{\omega}$ is the standard deviation of the $\bar{f}$'s, and M is the number of simulation trials (i.e., the number of pairs of values calculated), then the standard error of the estimate is

$$\bar{\omega}/\sqrt{M}$$

This is usually much less than the standard error calculated using $2M$ random trials.

Control Variate Technique

We have already given one example of the control variate technique in connection with the use of trees to value American options (see Section 19.3). The control variate technique is applicable when there are two similar derivatives, A and B. Derivative A is the one being valued; derivative B is similar to derivative A and has an analytic solution available. Two simulations using the same random number streams and the same Δt are carried out in parallel. The first is used to obtain an estimate f_A^* of the value of A; the second is used to obtain an estimate f_B^*, of the value of B. A better estimate f_A of the value of A is then obtained using the formula

$$f_A = f_A^* - f_B^* + f_B \tag{19.20}$$

where f_B is the known true value of B calculated analytically. Hull and White provide an example of the use of the control variate technique when evaluating the effect of stochastic volatility on the price of a European call option.[19] In this case, A is the option assuming stochastic volatility and B is the option assuming constant volatility.

[19] See J. Hull and A. White, "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Finance*, 42 (June 1987): 281–300.

Importance Sampling

Importance sampling is best explained with an example. Suppose that we wish to calculate the price of a deep-out-of-the-money European call option with strike price K and maturity T. If we sample values for the underlying asset price at time T in the usual way, most of the paths will lead to zero payoff. This is a waste of computation time because the zero-payoff paths contribute very little to the determination of the value of the option. We therefore try to choose only important paths, that is, paths where the stock price is above K at maturity.

Suppose F is the unconditional probability distribution function for the stock price at time T and q, the probability of the stock price being greater than K at maturity, is known analytically. Then $G = F/q$ is the probability distribution of the stock price conditional on the stock price being greater than K. To implement importance sampling, we sample from G rather than F. The estimate of the value of the option is the average discounted payoff multiplied by q.

Stratified Sampling

Sampling representative values rather than random values from a probability distribution usually gives more accuracy. Stratified sampling is a way of doing this. Suppose we wish to take 1000 samples from a probability distribution. We would divide the distribution into 1000 equally likely intervals and choose a representative value (typically the mean or median) for each interval.

In the case of a standard normal distribution when there are n intervals, we can calculate the representative value for the ith interval as

$$N^{-1}\left(\frac{i-0.5}{n}\right)$$

where N^{-1} is the inverse cumulative normal distribution. For example, when $n = 4$ the representative values corresponding to the four intervals are $N^{-1}(0.125)$, $N^{-1}(0.375)$, $N^{-1}(0.625)$, $N^{-1}(0.875)$. The function N^{-1} can be calculated using the NORMSINV function in Excel.

Moment Matching

Moment matching involves adjusting the samples taken from a standardized normal distribution so that the first, second, and possibly higher moments are matched. Suppose that we sample from a normal distribution with mean 0 and standard deviation 1 to calculate the change in the value of a particular variable over a particular time period. Suppose that the samples are ϵ_i $(1 \leqslant i \leqslant n)$. To match the first two moments, we calculate the mean of the samples, m, and the standard deviation of the samples, s. We then define adjusted samples ϵ_i^* $(1 \leqslant i \leqslant n)$ as

$$\epsilon_i^* = \frac{\epsilon_i - m}{s}$$

These adjusted samples have the correct mean of 0 and the correct standard deviation of 1.0. We use the adjusted samples for all calculations.

Moment matching saves computation time, but can lead to memory problems because every number sampled must be stored until the end of the simulation. Moment matching is sometimes termed *quadratic resampling*. It is often used in conjunction with the antithetic variable technique. Because the latter automatically matches all odd moments, the goal of moment matching then becomes that of matching the second moment and, possibly, the fourth moment.

Using Quasi-Random Sequences

A quasi-random sequence (also called a *low-discrepancy* sequence) is a sequence of representative samples from a probability distribution.[20] Descriptions of the use of quasi-random sequences appear in Brotherton-Ratcliffe, and Press *et al.*[21] Quasi-random sequences can have the desirable property that they lead to the standard error of an estimate being proportional to $1/M$ rather than $1/\sqrt{M}$, where M is the sample size.

Quasi-random sampling is similar to stratified sampling. The objective is to sample representative values for the underlying variables. In stratified sampling, it is assumed that we know in advance how many samples will be taken. A quasi-random sampling scheme is more flexible. The samples are taken in such a way that we are always "filling in" gaps between existing samples. At each stage of the simulation, the sampled points are roughly evenly spaced throughout the probability space.

Figure 19.14 shows points generated in two dimensions using a procedure suggested by Sobol'.[22] It can be seen that successive points do tend to fill in the gaps left by previous points.

19.8 FINITE DIFFERENCE METHODS

Finite difference methods value a derivative by solving the differential equation that the derivative satisfies. The differential equation is converted into a set of difference equations, and the difference equations are solved iteratively.

To illustrate the approach, we consider how it might be used to value an American put option on a stock paying a dividend yield of q. The differential equation that the option must satisfy is, from equation (15.6),

$$\frac{\partial f}{\partial t} + (r-q)S\frac{\partial f}{\partial S} + \tfrac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = rf \tag{19.21}$$

Suppose that the life of the option is T. We divide this into N equally spaced intervals of length $\Delta t = T/N$. A total of $N+1$ times are therefore considered

$$0,\ \Delta t,\ 2\,\Delta t,\ \ldots,\ T$$

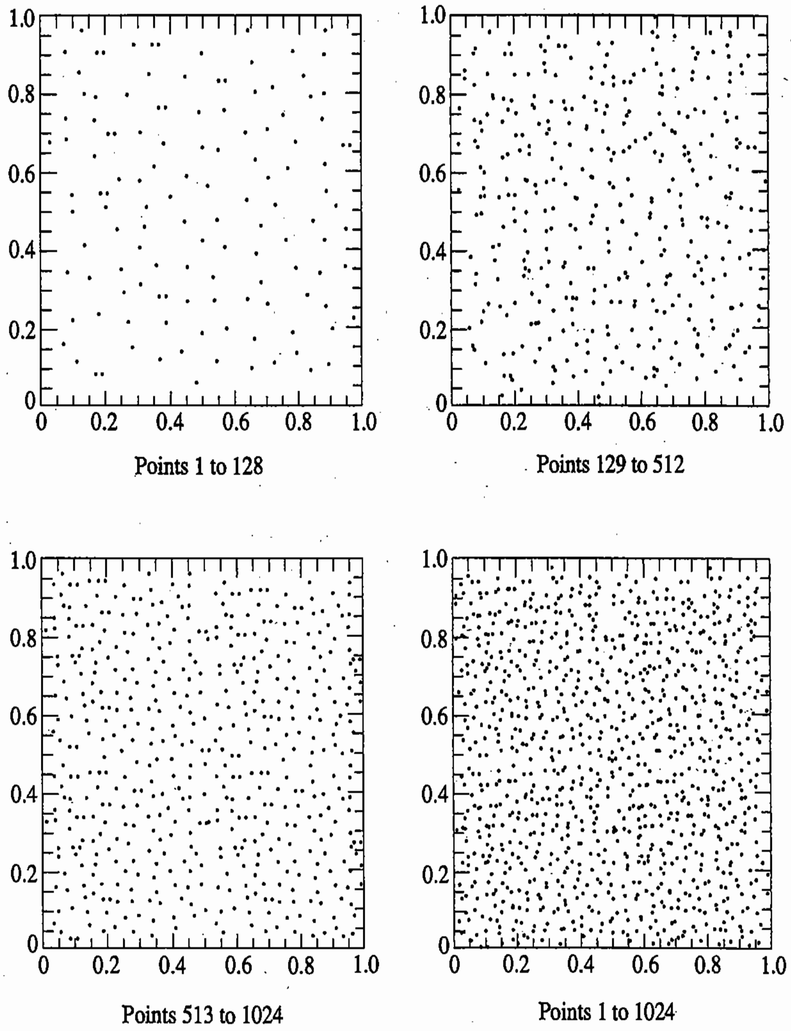
---

[20] The term *quasi-random* is a misnomer. A quasi-random sequence is totally deterministic.

[21] See R. Brotherton-Ratcliffe, "Monte Carlo Motoring," *Risk*, December 1994: 53–58; W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd edn. Cambridge University Press, 1992.

[22] See I.M. Sobol', *USSR Computational Mathematics and Mathematical Physics*, 7, 4 (1967): 86–112. A description of Sobol's procedure is in W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd edn. Cambridge University Press, 1992.

Figure 19.14 First 1024 points of a Sobol' sequence.



Suppose that S_{max} is a stock price sufficiently high that, when it is reached, the put has virtually no value. We define $\Delta S = S_{max}/M$ and consider a total of $M+1$ equally spaced stock prices:

$$0, \ \Delta S, \ 2\Delta S, \ \ldots, \ S_{max}$$

The level S_{max} is chosen so that one of these is the current stock price.

The time points and stock price points define a grid consisting of a total of $(M+1)(N+1)$ points, as shown in Figure 19.15. The (i, j) point on the grid is the point that corresponds to time $i\,\Delta t$ and stock price $j\,\Delta S$. We will use the variable $f_{i,j}$ to denote the value of the option at the (i, j) point.

Implicit Finite Difference Method

For an interior point (i, j) on the grid, $\partial f/\partial S$ can be approximated as

$$\frac{\partial f}{\partial S} = \frac{f_{i,j+1} - f_{i,j}}{\Delta S} \tag{19.22}$$

or as

$$\frac{\partial f}{\partial S} = \frac{f_{i,j} - f_{i,j-1}}{\Delta S} \tag{19.23}$$

Equation (19.22) is known as the *forward difference approximation*; equation (19.23) is known as the *backward difference approximation.* We use a more symmetrical approximation by averaging the two:

$$\frac{\partial f}{\partial S} = \frac{f_{i,j+1} - f_{i,j-1}}{2\,\Delta S} \tag{19.24}$$

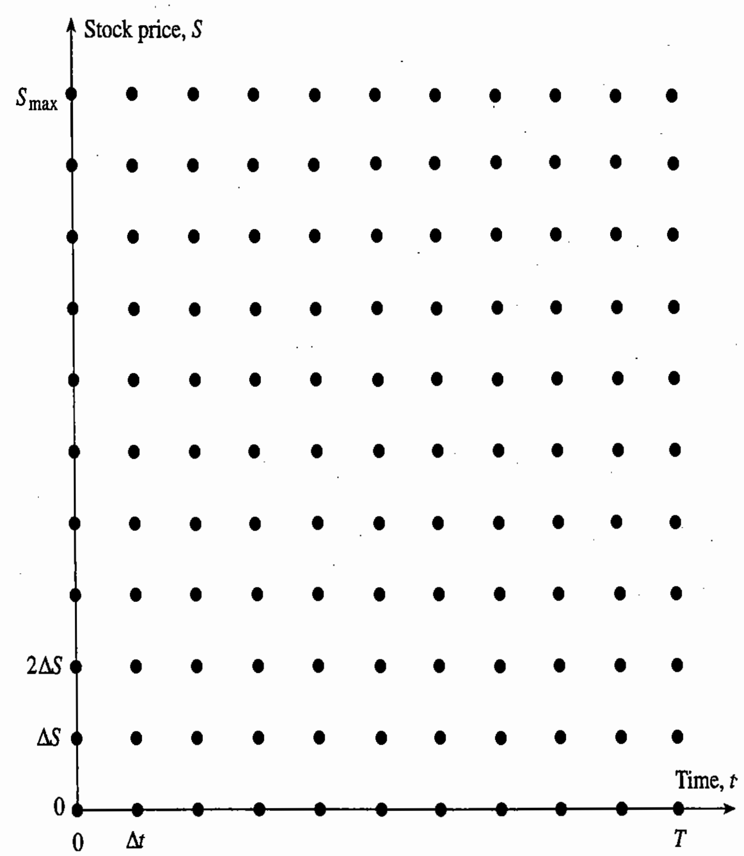
For $\partial f/\partial t$, we will use a forward difference approximation so that the value at time $i\,\Delta t$ is related to the value at time $(i+1)\,\Delta t$:

$$\frac{\partial f}{\partial t} = \frac{f_{i+1,j} - f_{i,j}}{\Delta t} \tag{19.25}$$

The backward difference approximation for $\partial f/\partial S$ at the (i, j) point is given by equation (19.23). The backward difference at the $(i, j+1)$ point is

$$\frac{f_{i,j+1} - f_{i,j}}{\Delta S}$$

Figure 19.15 Grid for finite difference approach.



Hence a finite difference approximation for $\partial^2 f/\partial S^2$ at the (i, j) point is

$$\frac{\partial^2 f}{\partial S^2} = \left(\frac{f_{i,j+1} - f_{i,j}}{\Delta S} - \frac{f_{i,j} - f_{i,j-1}}{\Delta S}\right) \Big/ \Delta S$$

or

$$\frac{\partial^2 f}{\partial S^2} = \frac{f_{i,j+1} + f_{i,j-1} - 2f_{i,j}}{\Delta S^2} \tag{19.26}$$

Substituting equations (19.24), (19.25), and (19.26) into the differential equation (19.21) and noting that $S = j\,\Delta S$ gives

$$\frac{f_{i+1,j} - f_{i,j}}{\Delta t} + (r-q)j\,\Delta S\frac{f_{i,j+1} - f_{i,j-1}}{2\,\Delta S} + \tfrac{1}{2}\sigma^2 j^2 \Delta S^2 \frac{f_{i,j+1} + f_{i,j-1} - 2f_{i,j}}{\Delta S^2} = r f_{i,j}$$

for $j = 1, 2, \dots, M-1$ and $i = 0, 1 \dots, N-1$. Rearranging terms, we obtain

$$a_j f_{i,j-1} + b_j f_{i,j} + c_j f_{i,j+1} = f_{i+1,j} \tag{19.27}$$

where

$$a_j = \tfrac{1}{2}(r-q)j\,\Delta t - \tfrac{1}{2}\sigma^2 j^2 \Delta t$$

$$b_j = 1 + \sigma^2 j^2 \Delta t + r\,\Delta t$$

$$c_j = -\tfrac{1}{2}(r-q)j\,\Delta t - \tfrac{1}{2}\sigma^2 j^2 \Delta t$$

The value of the put at time T is $\max(K - S_T, 0)$, where S_T is the stock price at time T. Hence,

$$f_{N,j} = \max(K - j\,\Delta S, 0), \quad j = 0, 1, \dots, M \tag{19.28}$$

The value of the put option when the stock price is zero is K. Hence,

$$f_{i,0} = K, \quad i = 0, 1, \dots, N \tag{19.29}$$

We assume that the put option is worth zero when $S = S_{\max}$, so that

$$f_{i,M} = 0, \quad i = 0, 1, \dots, N \tag{19.30}$$

Equations (19.28), (19.29), and (19.30) define the value of the put option along the three edges of the grid in Figure 19.15, where $S = 0$, $S = S_{\max}$, and $t = T$. It remains to use equation (19.27) to arrive at the value of f at all other points. First the points corresponding to time $T - \Delta t$ are tackled. Equation (19.27) with $i = N - 1$ gives

$$a_j\, f_{N-1,j-1} + b_j\, f_{N-1,j} + c_j\, f_{N-1,j+1} = f_{N,j} \tag{19.31}$$

for $j = 1, 2, \dots, M-1$. The right-hand sides of these equations are known from equation (19.28). Furthermore, from equations (19.29) and (19.30),

$$f_{N-1,0} = K \tag{19.32}$$

$$f_{N-1,M} = 0 \tag{19.33}$$

Equations (19.31) are therefore $M-1$ simultaneous equations that can be solved for the $M-1$ unknowns: $f_{N-1,1}, f_{N-1,2}, \ldots, f_{N-1,M-1}$.[23] After this has been done, each value of $f_{N-1,j}$ is compared with $K - j\,\Delta S$. If $f_{N-1,j} < K - j\,\Delta S$, early exercise at time $T - \Delta t$ is optimal and $f_{N-1,j}$ is set equal to $K - j\,\Delta S$. The nodes corresponding to time $T - 2\,\Delta t$ are handled in a similar way, and so on. Eventually, $f_{0,1}, f_{0,2}, f_{0,3}, \ldots, f_{0,M-1}$ are obtained. One of these is the option price of interest.

The control variate technique can be used in conjunction with finite difference methods. The same grid is used to value an option similar to the one under consideration but for which an analytic valuation is available. Equation (19.20) is then used.

Example 19.10

Table 19.4 shows the result of using the implicit finite difference method as just described for pricing the American put option in Example 19.1. Values of 20, 10, and 5 were chosen for M, N, and ΔS, respectively. Thus, the option price is evaluated at \$5 stock price intervals between \$0 and \$100 and at half-month time intervals throughout the life of the option. The option price given by the grid is \$4.07. The same grid gives the price of the corresponding European option as \$3.91. The true European price given by the Black–Scholes formula is \$4.08. The control variate estimate of the American price is therefore

$$4.07 + 4.08 - 3.91 = \$4.24$$

Explicit Finite Difference Method

The implicit finite difference method has the advantage of being very robust. It always converges to the solution of the differential equation as ΔS and Δt approach zero.[24] One of the disadvantages of the implicit finite difference method is that $M-1$ simultaneous equations have to be solved in order to calculate the $f_{i,j}$ from the $f_{i+1,j}$. The method can be simplified if the values of $\partial f/\partial S$ and $\partial^2 f/\partial S^2$ at point (i, j) on the grid are assumed to be the same as at point $(i+1, j)$. Equations (19.24) and (19.26) then become

$$\frac{\partial f}{\partial S} = \frac{f_{i+1,j+1} - f_{i+1,j-1}}{2\,\Delta S}$$

$$\frac{\partial^2 f}{\partial S^2} = \frac{f_{i+1,j+1} + f_{i+1,j-1} - 2f_{i+1,j}}{\Delta S^2}.$$

The difference equation is

$$\frac{f_{i+1,j} - f_{i,j}}{\Delta t} + (r-q)j\,\Delta S\,\frac{f_{i+1,j+1} - f_{i+1,j-1}}{2\Delta S} + \tfrac{1}{2}\sigma^2 j^2 \Delta S^2\,\frac{f_{i+1,j+1} + f_{i+1,j-1} - 2f_{i+1,j}}{\Delta S^2} = r f_{i,j}$$

[23] This does not involve inverting a matrix. The $j=1$ equation in (19.31) can be used to express $f_{N-1,2}$ in terms of $f_{N-1,1}$; the $j=2$ equation, when combined with the $j=1$ equation, can be used to express $f_{N-1,3}$ in terms of $f_{N-1,1}$; and so on. The $j = M-2$ equation, together with earlier equations, enables $f_{N-1,M-1}$ to be expressed in terms of $f_{N-1,1}$. The final $j = M-1$ equation can then be solved for $f_{N-1,1}$, which can then be used to determine the other $f_{N-1,j}$.

[24] A general rule in finite difference methods is that ΔS should be kept proportional to $\sqrt{\Delta t}$ as they approach zero.

Table 19.4 Grid to value American option in Example 19.1 using implicit finite difference methods.

Stock price (dollars)	*Time to maturity (months)*										
	5	*4.5*	*4*	*3.5*	*3*	*2.5*	*2*	*1.5*	*1*	*0.5*	*0*
100	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
95	0.02	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
90	0.05	0.04	0.03	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.00
85	0.09	0.07	0.05	0.03	0.02	0.01	0.01	0.00	0.00	0.00	0.00
80	0.16	0.12	0.09	0.07	0.04	0.03	0.02	0.01	0.00	0.00	0.00
75	0.27	0.22	0.17	0.13	0.09	0.06	0.03	0.02	0.01	0.00	0.00
70	0.47	0.39	0.32	0.25	0.18	0.13	0.08	0.04	0.02	0.00	0.00
65	0.82	0.71	0.60	0.49	0.38	0.28	0.19	0.11	0.05	0.02	0.00
60	1.42	1.27	1.11	0.95	0.78	0.62	0.45	0.30	0.16	0.05	0.00
55	2.43	2.24	2.05	1.83	1.61	1.36	1.09	0.81	0.51	0.22	0.00
50	4.07	3.88	3.67	3.45	3.19	2.91	2.57	2.17	1.66	0.99	0.00
45	6.58	6.44	6.29	6.13	5.96	5.77	5.57	5.36	5.17	5.02	5.00
40	10.15	10.10	10.05	10.01	10.00	10.00	10.00	10.00	10.00	10.00	10.00
35	15.00	15.00	15.00	15.00	15.00	15.00	15.00	15.00	15.00	15.00	15.00
30	20.00	20.00	20.00	20.00	20.00	20.00	20.00	20.00	20.00	20.00	20.00
25	25.00	25.00	25.00	25.00	25.00	25.00	25.00	25.00	25.00	25.00	25.00
20	30.00	30.00	30.00	30.00	30.00	30.00	30.00	30.00	30.00	30.00	30.00
15	35.00	35.00	35.00	35.00	35.00	35.00	35.00	35.00	35.00	35.00	35.00
10	40.00	40.00	40.00	40.00	40.00	40.00	40.00	40.00	40.00	40.00	40.00
5	45.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00
0	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00

or

$$f_{i,j} = a_j^* f_{i+1,j-1} + b_j^* f_{i+1,j} + c_j^* f_{i+1,j+1} \tag{19.34}$$

where

$$a_j^* = \frac{1}{1+r\,\Delta t}\left(-\tfrac{1}{2}(r-q)j\,\Delta t + \tfrac{1}{2}\sigma^2 j^2\,\Delta t\right)$$

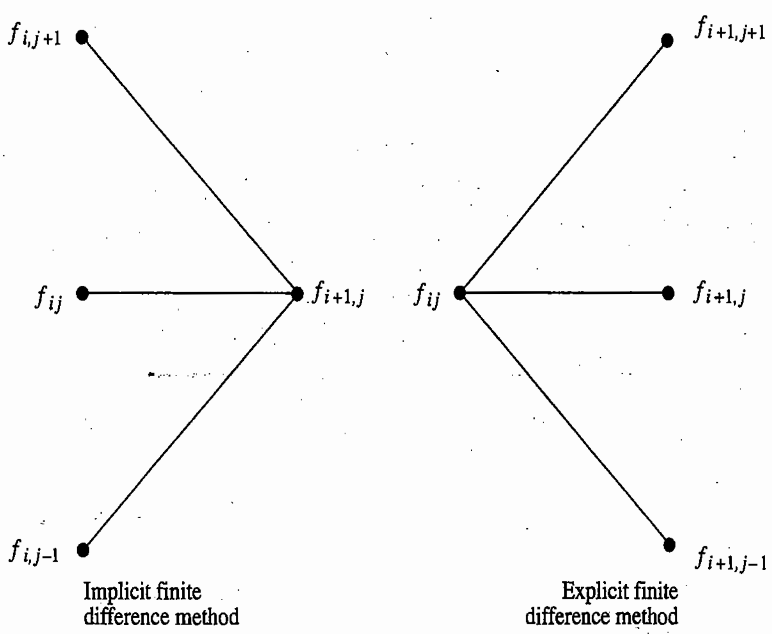
$$b_j^* = \frac{1}{1+r\,\Delta t}(1-\sigma^2 j^2\,\Delta t)$$

$$c_j^* = \frac{1}{1+r\,\Delta t}\left(\tfrac{1}{2}(r-q)j\,\Delta t + \tfrac{1}{2}\sigma^2 j^2\,\Delta t\right)$$

This creates what is known as the *explicit finite difference method*.[25] Figure 19.16 shows the difference between the implicit and explicit methods. The implicit method leads to equation (19.27), which gives a relationship between three different values of the option at time $i\,\Delta t$ (i.e., $f_{i,j-1}$, $f_{i,j}$, and $f_{i,j+1}$) and one value of the option at time $(i+1)\,\Delta t$

[25] We also obtain the explicit finite difference method if we use the backward difference approximation instead of the forward difference approximation for $\partial f/\partial t$.

Figure 19.16 Difference between implicit and explicit finite difference methods.



(i.e., $f_{i+1,j}$). The explicit method leads to equation (19.34), which gives a relationship between one value of the option at time $i\,\Delta t$ (i.e., $f_{i,j}$) and three different values of the option at time $(i+1)\,\Delta t$ (i.e., $f_{i+1,j-1}$, $f_{i+1,j}$, $f_{i+1,j+1}$).

Example 19.11

Table 19.5 shows the result of using the explicit version of the finite difference method for pricing the American put option described in Example 19.1. As in Example 19.10, values of 20, 10, and 5 were chosen for M, N, and ΔS, respectively. The option price given by the grid is \$4.26.[26]

Change of Variable

It is computationally more efficient to use finite difference methods with $\ln S$ rather than S as the underlying variable. Define $Z = \ln S$. Equation (19.21) becomes

$$\frac{\partial f}{\partial t} + \left(r - q - \frac{\sigma^2}{2}\right)\frac{\partial f}{\partial Z} + \tfrac{1}{2}\sigma^2\frac{\partial^2 f}{\partial Z^2} = rf$$

The grid then evaluates the derivative for equally spaced values of Z rather than for equally spaced values of S. The difference equation for the implicit method becomes

$$\frac{f_{i+1,j} - f_{i,j}}{\Delta t} + (r - q - \sigma^2/2)\frac{f_{i,j+1} - f_{i,j-1}}{2\Delta Z} + \tfrac{1}{2}\sigma^2\frac{f_{i,j+1} + f_{i,j-1} - 2f_{i,j}}{\Delta Z^2} = rf_{i,j}$$

or

$$\alpha_j f_{i,j-1} + \beta_j f_{i,j} + \gamma_j f_{i,j+1} = f_{i+1,j} \qquad \textbf{(19.35)}$$

[26] The negative numbers and other inconsistencies in the top left-hand part of the grid will be explained later.

Table 19.5 Grid to value American option in Example 19.1 using explicit finite difference method.

Stock price (dollars)	*Time to maturity (months)*										
	5	*4.5*	*4*	*3.5*	*3*	*2.5*	*2*	*1.5*	*1*	*0.5*	*0*
100	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
95	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
90	−0.11	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
85	0.28	−0.05	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
80	−0.13	0.20	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00
75	0.46	0.06	0.20	0.04	0.06	0.00	0.00	0.00	0.00	0.00	0.00
70	0.32	0.46	0.23	0.25	0.10	0.09	0.00	0.00	0.00	0.00	0.00
65	0.91	0.68	0.63	0.44	0.37	0.21	0.14	0.00	0.00	0.00	0.00
60	1.48	1.37	1.17	1.02	0.81	0.65	0.42	0.27	0.00	0.00	0.00
55	2.59	2.39	2.21	1.99	1.77	1.50	1.24	0.90	0.59	0.00	0.00
50	4.26	4.08	3.89	3.68	3.44	3.18	2.87	2.53	2.07	1.56	0.00
45	6.76	6.61	6.47	6.31	6.15	5.96	5.75	5.50	5.24	5.00	5.00
40	10.28	10.20	10.13	10.06	10.01	10.00	10.00	10.00	10.00	10.00	10.00
35	15.00	15.00	15.00	15.00	15.00	15.00	15.00	15.00	15.00	15.00	15.00
30	20.00	20.00	20.00	20.00	20.00	20.00	20.00	20.00	20.00	20.00	20.00
25	25.00	25.00	25.00	25.00	25.00	25.00	25.00	25.00	25.00	25.00	25.00
20	30.00	30.00	30.00	30.00	30.00	30.00	30.00	30.00	30.00	30.00	30.00
15	35.00	35.00	35.00	35.00	35.00	35.00	35.00	35.00	35.00	35.00	35.00
10	40.00	40.00	40.00	40.00	40.00	40.00	40.00	40.00	40.00	40.00	40.00
5	45.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00	45.00
0	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00	50.00

where

$$\alpha_j = \frac{\Delta t}{2\Delta Z}(r - q - \sigma^2/2) - \frac{\Delta t}{2\Delta Z^2}\sigma^2$$

$$\beta_j = 1 + \frac{\Delta t}{\Delta Z^2}\sigma^2 + r\,\Delta t$$

$$\gamma_j = -\frac{\Delta t}{2\Delta Z}(r - q - \sigma^2/2) - \frac{\Delta t}{2\Delta Z^2}\sigma^2$$

The difference equation for the explicit method becomes

$$\frac{f_{i+1,j} - f_{i,j}}{\Delta t} + (r - q - \sigma^2/2)\,\frac{f_{i+1,j+1} - f_{i+1,j-1}}{2\Delta Z} + \tfrac{1}{2}\sigma^2\,\frac{f_{i+1,j+1} + f_{i+1,j-1} - 2f_{i+1,j}}{\Delta Z^2} = rf_{i,j}$$

or

$$\alpha_j^*\, f_{i+1,j-1} + \beta_j^*\, f_{i+1,j} + \gamma_j^*\, f_{i+1,j+1} = f_{i,j} \tag{19.36}$$

where

$$\alpha_j^* = \frac{1}{1+r\,\Delta t}\left[-\frac{\Delta t}{2\Delta Z}(r-q-\sigma^2/2)+\frac{\Delta t}{2\Delta Z^2}\sigma^2\right] \tag{19.37}$$

$$\beta_j^* = \frac{1}{1+r\,\Delta t}\left(1-\frac{\Delta t}{\Delta Z^2}\sigma^2\right) \tag{19.38}$$

$$\gamma_j^* = \frac{1}{1+r\,\Delta t}\left[\frac{\Delta t}{2\Delta Z}(r-q-\sigma^2/2)+\frac{\Delta t}{2\Delta Z^2}\sigma^2\right] \tag{19.39}$$

The change of variable approach has the property that α_j, β_j, and γ_j as well as α_j^*, β_j^*, and γ_j^* are independent of j. In most cases, a good choice for ΔZ is $\sigma\sqrt{3\Delta t}$.

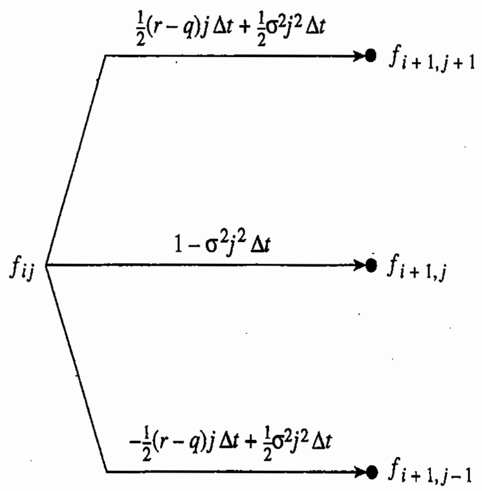
Relation to Trinomial Tree Approaches

The explicit finite difference method is equivalent to the trinomial tree approach.[27] In the expressions for a_j^*, b_j^*, and c_j^* in equation (19.34), we can interpret terms as follows:

$-\frac{1}{2}(r-q)j\,\Delta t+\frac{1}{2}\sigma^2 j^2\Delta t$:	Probability of stock price decreasing from $j\,\Delta S$ to $(j-1)\Delta S$ in time Δt.
$1-\sigma^2 j^2\Delta t$:	Probability of stock price remaining unchanged at $j\,\Delta S$ in time Δt.
$\frac{1}{2}(r-q)j\,\Delta t+\frac{1}{2}\sigma^2 j^2\Delta t$:	Probability of stock price increasing from $j\,\Delta S$ to $(j+1)\Delta S$ in time Δt.

This interpretation is illustrated in Figure 19.17. The three probabilities sum to unity. They give the expected increase in the stock price in time Δt as $(r-q)j\,\Delta S\,\Delta t = (r-q)S\,\Delta t$. This is the expected increase in a risk-neutral world. For small values

Figure 19.17 Interpretation of explicit finite difference method as a trinomial tree.



[27] It can also be shown that the implicit finite difference method is equivalent to a multinomial tree approach where there are $M+1$ branches emanating from each node.

of Δt, they also give the variance of the change in the stock price in time Δt as $\sigma^2 j^2 \Delta S^2 \Delta t = \sigma^2 S^2 \Delta t$. This corresponds to the stochastic process followed by S. The value of f at time $i\,\Delta t$ is calculated as the expected value of f at time $(i+1)\,\Delta t$ in a risk-neutral world discounted at the risk-free rate.

For the explicit version of the finite difference method to work well, the three "probabilities"

$$-\tfrac{1}{2}(r-q)j\,\Delta t + \tfrac{1}{2}\sigma^2 j^2 \Delta t,$$

$$1-\sigma^2 j^2 \Delta t$$

$$\tfrac{1}{2}(r-q)j\,\Delta t + \tfrac{1}{2}\sigma^2 j^2 \Delta t$$

should all be positive. In Example 19.11, $1-\sigma^2 j^2 \Delta t$ is negative when $j \geqslant 13$ (i.e., when $S \geqslant 65$). This explains the negative option prices and other inconsistencies in the top left-hand part of Table 19.5. This example illustrates the main problem associated with the explicit finite difference method. Because the probabilities in the associated tree may be negative, it does not necessarily produce results that converge to the solution of the differential equation.[28]

When the change-of-variable approach is used (see equations (19.36) to (19.39)), the probability that $Z = \ln S$ will decrease by ΔZ, stay the same, and increase by ΔZ are

$$-\frac{\Delta t}{2\Delta Z}(r-q-\sigma^2/2)+\frac{\Delta t}{2\Delta Z^2}\sigma^2$$

$$1-\frac{\Delta t}{\Delta z^2}\sigma^2$$

$$\frac{\Delta t}{2\Delta Z}(r-q-\sigma^2/2)+\frac{\Delta t}{2\Delta Z^2}\sigma^2$$

respectively. These movements in Z correspond to the stock price changing from S to $Se^{-\Delta Z}$, S, and $Se^{\Delta Z}$, respectively. If we set $\Delta Z = \sigma\sqrt{3\Delta t}$, then the tree and the probabilities are identical to those for the trinomial tree approach discussed in Section 19.4.

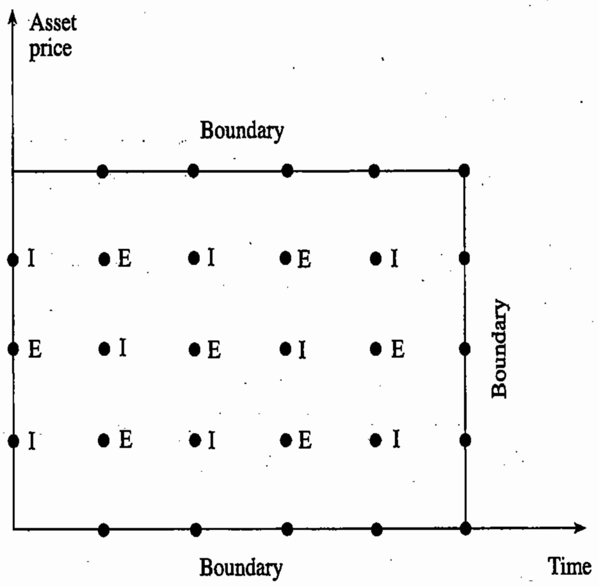
Other Finite Difference Methods

Many of the other finite difference methods that have been proposed have some of the features of the explicit finite difference method and some features of the implicit finite difference method.

In what is known as the *hopscotch method*, we alternate between the explicit and implicit calculations as we move from node to node. This is illustrated in Figure 19.18. At each time, we first do all the calculations at the "explicit nodes" (E) in the usual way. The "implicit nodes" (I) can then be handled without solving a set of simultaneous equations because the values at the adjacent nodes have already been calculated.

[28] J. Hull and A. White, "Valuing Derivative Securities Using the Explicit Finite Difference Method," *Journal of Financial and Quantitative Analysis*, 25 (March 1990): 87–100, show how this problem can be overcome. In the situation considered here it is sufficient to construct the grid in $\ln S$ rather than S to ensure convergence.

Figure 19.18 The hopscotch method. I indicates node at which implicit calculations are done; E indicates node at which explicit calculations are done.



The *Crank–Nicolson* scheme is an average of the explicit and implicit methods. For the implicit method, equation (19.27) gives

$$f_{i,j} = a_j f_{i-1,j-1} + b_j f_{i-1,j} + c_j f_{i-1,j+1}$$

For the explicit method, equation (19.34) gives

$$f_{i-1,j} = a_j^* f_{i,j-1} + b_j^* f_{i,j} + c_j^* f_{i,j+1}$$

The Crank–Nicolson method averages these two equations to obtain

$$f_{i,j} + f_{i-1,j} = a_j f_{i-1,j-1} + b_j f_{i-1,j} + c_j f_{i-1,j+1} + a_j^* f_{i,j-1} + b_j^* f_{i,j} + c_j^* f_{i,j+1}$$

Putting

$$g_{i,j} = f_{i,j} - a_j^* f_{i,j-1} - b_j^* f_{i,j} - c_j^* f_{i,j+1}$$

gives

$$g_{i,j} = a_j f_{i-1,j-1} + b_j f_{i-1,j} + c_j f_{i-1,j+1} - f_{i-1,j}$$

This shows that implementing the Crank–Nicolson method is similar to implementing the implicit finite difference method. The advantage of the Crank–Nicolson method is that it has faster convergence than either the explicit or implicit method.

Applications of Finite Difference Methods

Finite difference methods can be used for the same types of derivative pricing problems as tree approaches. They can handle American-style as well as European-style derivatives but cannot easily be used in situations where the payoff from a derivative depends on the past history of the underlying variable. Finite difference methods can, at the expense of a considerable increase in computer time, be used when there are several state variables. The grid in Figure 19.15 then becomes multidimensional.

The method for calculating Greek letters is similar to that used for trees. Delta, gamma, and theta can be calculated directly from the $f_{i,j}$ values on the grid. For vega, it is necessary to make a small change to volatility and recalculate the value of the derivative using the same grid.

SUMMARY

We have presented three different numerical procedures for valuing derivatives when no analytic solution is available. These involve the use of trees, Monte Carlo simulation, and finite difference methods.

Binomial trees assume that, in each short interval of time Δt, a stock price either moves up by a multiplicative amount u or down by a multiplicative amount d. The sizes of u and d and their associated probabilities are chosen so that the change in the stock price has the correct mean and standard deviation in a risk-neutral world. Derivative prices are calculated by starting at the end of the tree and working backwards. For an American option, the value at a node is the greater of (a) the value if it is exercised immediately and (b) the discounted expected value if it is held for a further period of time Δt.

Monte Carlo simulation involves using random numbers to sample many different paths that the variables underlying the derivative could follow in a risk-neutral world. For each path, the payoff is calculated and discounted at the risk-free interest rate. The arithmetic average of the discounted payoffs is the estimated value of the derivative.

Finite difference methods solve the underlying differential equation by converting it to a difference equation. They are similar to tree approaches in that the computations work back from the end of the life of the derivative to the beginning. The explicit method is functionally the same as using a trinomial tree. The implicit finite difference method is more complicated but has the advantage that the user does not have to take any special precautions to ensure convergence.

In practice, the method that is chosen is likely to depend on the characteristics of the derivative being evaluated and the accuracy required. Monte Carlo simulation works forward from the beginning to the end of the life of a derivative. It can be used for European-style derivatives and can cope with a great deal of complexity as far as the payoffs are concerned. It becomes relatively more efficient as the number of underlying variables increases. Tree approaches and finite difference methods work from the end of the life of a security to the beginning and can accommodate American-style as well as European-style derivatives. However, they are difficult to apply when the payoffs depend on the past history of the state variables as well as on their current values. Also, they are liable to become computationally very time consuming when three or more variables are involved.

FURTHER READING

General

Clewlow, L., and C. Strickland, *Implementing Derivatives Models*. Chichester: Wiley, 1998.

Press, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd edn. Cambridge University Press, 1992.

On Tree Approaches

Cox, J. C, S. A. Ross, and M. Rubinstein. "Option Pricing: A Simplified Approach," *Journal of Financial Economics*, 7 (October 1979): 229–64.

Figlewski, S., and B. Gao. "The Adaptive Mesh Model: A New Approach to Efficient Option Pricing," *Journal of Financial Economics*, 53 (1999): 313–51.

Hull, J. C., and A. White, "The Use of the Control Variate Technique in Option Pricing," *Journal of Financial and Quantitative Analysis*, 23 (September 1988): 237–51.

Rendleman, R., and B. Bartter, "Two State Option Pricing," *Journal of Finance*, 34 (1979): 1092–1110.

On Monte Carlo Simulation

Boyle, P.P., "Options: A Monte Carlo Approach," *Journal of Financial Economics*, 4 (1977): 323–38.

Boyle, P.P., M. Broadie, and P. Glasserman. "Monte Carlo Methods for Security Pricing," *Journal of Economic Dynamics and Control*, 21 (1997): 1267–1322.

Broadie, M., P. Glasserman, and G. Jain. "Enhanced Monte Carlo Estimates for American Option Prices," *Journal of Derivatives*, 5 (Fall 1997): 25–44.

On Finite Difference Methods

Hull, J. C., and A. White, "Valuing Derivative Securities Using the Explicit Finite Difference Method," *Journal of Financial and Quantitative Analysis*, 25 (March 1990): 87–100.

Wilmott, P., *Derivatives: The Theory and Practice of Financial Engineering*. Chichester: Wiley, 1998.

Questions and Problems (Answers in Solutions Manual)

19.1. Which of the following can be estimated for an American option by constructing a single binomial tree: delta, gamma, vega, theta, rho?

19.2. Calculate the price of a 3-month American put option on a non-dividend-paying stock when the stock price is \$60, the strike price is \$60, the risk-free interest rate is 10% per annum, and the volatility is 45% per annum. Use a binomial tree with a time interval of 1 month.

19.3. Explain how the control variate technique is implemented when a tree is used to value American options.

19.4. Calculate the price of a 9-month American call option on corn futures when the current futures price is 198 cents, the strike price is 200 cents, the risk-free interest rate is 8% per annum, and the volatility is 30% per annum. Use a binomial tree with a time interval of 3 months.

19.5. Consider an option that pays off the amount by which the final stock price exceeds the average stock price achieved during the life of the option. Can this be valued using the binomial tree approach? Explain your answer.

19.6. "For a dividend-paying stock, the tree for the stock price does not recombine; but the tree for the stock price less the present value of future dividends does recombine." Explain this statement.

19.7. Show that the probabilities in a Cox, Ross, and Rubinstein binomial tree are negative when the condition in footnote 9 holds.

19.8. Use stratified sampling with 100 trials to improve the estimate of π in Business Snapshot 19.1 and Table 19.1.

19.9. Explain why the Monte Carlo simulation approach cannot easily be used for American-style derivatives.

19.10. A 9-month American put option on a non-dividend-paying stock has a strike price of \$49. The stock price is \$50, the risk-free rate is 5% per annum, and the volatility is 30% per annum. Use a three-step binomial tree to calculate the option price.

19.11. Use a three-time-step tree to value a 9-month American call option on wheat futures. The current futures price is 400 cents, the strike price is 420 cents, the risk-free rate is 6%, and the volatility is 35% per annum. Estimate the delta of the option from your tree.

19.12. A 3-month American call option on a stock has a strike price of \$20. The stock price is \$20, the risk-free rate is 3% per annum, and the volatility is 25% per annum. A dividend of \$2 is expected in 1.5 months. Use a three-step binomial tree to calculate the option price.

19.13. A 1-year American put option on a non-dividend-paying stock has an exercise price of \$18. The current stock price is \$20, the risk-free interest rate is 15% per annum, and the volatility of the stock price is 40% per annum. Use the DerivaGem software with four 3-month time steps to estimate the value of the option. Display the tree and verify that the option prices at the final and penultimate nodes are correct. Use DerivaGem to value the European version of the option. Use the control variate technique to improve your estimate of the price of the American option.

19.14. A 2-month American put option on a stock index has an exercise price of 480. The current level of the index is 484, the risk-free interest rate is 10% per annum, the dividend yield on the index is 3% per annum, and the volatility of the index is 25% per annum. Divide the life of the option into four half-month periods and use the tree approach to estimate the value of the option.

19.15. How can the control variate approach improve the estimate of the delta of an American option when the tree approach is used?

19.16. Suppose that Monte Carlo simulation is being used to evaluate a European call option on a non-dividend-paying stock when the volatility is stochastic. How could the control variate and antithetic variable technique be used to improve numerical efficiency? Explain why it is necessary to calculate six values of the option in each simulation trial when both the control variate and the antithetic variable technique are used.

19.17. Explain how equations (19.27) to (19.30) change when the implicit finite difference method is being used to evaluate an American call option on a currency.

19.18. An American put option on a non-dividend-paying stock has 4 months to maturity. The exercise price is \$21, the stock price is \$20, the risk-free rate of interest is 10% per annum, and the volatility is 30% per annum. Use the explicit version of the finite difference approach to value the option. Use stock price intervals of \$4 and time intervals of 1 month.

19.19. The spot price of copper is \$0.60 per pound. Suppose that the futures prices (dollars per pound) are as follows:

3 months	0.59
6 months	0.57
9 months	0.54
12 months	0.50

The volatility of the price of copper is 40% per annum and the risk-free rate is 6% per annum. Use a binomial tree to value an American call option on copper with an exercise price of $0.60 and a time to maturity of 1 year. Divide the life of the option into four 3-month periods for the purposes of constructing the tree. (*Hint*: As explained in Section 16.7, the futures price of a variable is its expected future price in a risk-neutral world.)

19.20. Use the binomial tree in Problem 19.19 to value a security that pays off x^2 in 1 year where x is the price of copper.

19.21. When do the boundary conditions for $S = 0$ and $S \to \infty$ affect the estimates of derivative prices in the explicit finite difference method?

19.22. How would you use the antithetic variable method to improve the estimate of the European option in Business Snapshot 19.2 and Table 19.2?

19.23. A company has issued a 3-year convertible bond that has a face value of $25 and can be exchanged for two of the company's shares at any time. The company can call the issue, forcing conversion, when the share price is greater than or equal to $18. Assuming that the company will force conversion at the earliest opportunity, what are the boundary conditions for the price of the convertible? Describe how you would use finite difference methods to value the convertible assuming constant interest rates. Assume there is no risk of the company defaulting.

19.24. Provide formulas that can be used for obtaining three random samples from standard normal distributions when the correlation between sample i and sample j is $\rho_{i,j}$.

Assignment Questions

19.25. An American put option to sell a Swiss franc for dollars has a strike price of $0.80 and a time to maturity of 1 year. The Swiss franc's volatility is 10%, the dollar interest rate is 6%, the Swiss franc interest rate is 3%, and the current exchange rate is 0.81. Use a three-step binomial tree to value the option. Estimate the delta of the option from your tree.

19.26. A 1-year American call option on silver futures has an exercise price of $9.00. The current futures price is $8.50, the risk-free rate of interest is 12% per annum, and the volatility of the futures price is 25% per annum. Use the DerivaGem software with four 3-month time steps to estimate the value of the option. Display the tree and verify that the option prices at the final and penultimate nodes are correct. Use DerivaGem to value the European version of the option. Use the control variate technique to improve your estimate of the price of the American option.

19.27. A 6-month American call option on a stock is expected to pay dividends of $1 per share at the end of the second month and the fifth month. The current stock price is $30, the exercise price is $34, the risk-free interest rate is 10% per annum, and the volatility of the part of the stock price that will not be used to pay the dividends is 30% per annum. Use the DerivaGem software with the life of the option divided into six time steps to estimate the value of the option. Compare your answer with that given by Black's approximation (see Section 13.12).

19.28. The current value of the British pound is $1.60 and the volatility of the pound/dollar exchange rate is 15% per annum. An American call option has an exercise price of $1.62 and a time to maturity of 1 year. The risk-free rates of interest in the United States and the United Kingdom are 6% per annum and 9% per annum, respectively. Use the explicit finite difference method to value the option. Consider exchange rates at intervals of 0.20 between 0.80 and 2.40 and time intervals of 3 months.

19.29. Answer the following questions concerned with the alternative procedures for constructing trees in Section 19.4:

(a) Show that the binomial model in Section 19.4 is exactly consistent with the mean and variance of the change in the logarithm of the stock price in time Δt.

(b) Show that the trinomial model in Section 19.4 is consistent with the mean and variance of the change in the logarithm of the stock price in time Δt when terms of order $(\Delta t)^2$ and higher are ignored.

(c) Construct an alternative to the trinomial model in Section 19.4 so that the probabilities are 1/6, 2/3, and 1/6 on the upper, middle, and lower branches emanating from each node. Assume that the branching is from S to Su, Sm, or Sd with $m^2 = ud$. Match the mean and variance of the change in the logarithm of the stock price exactly.

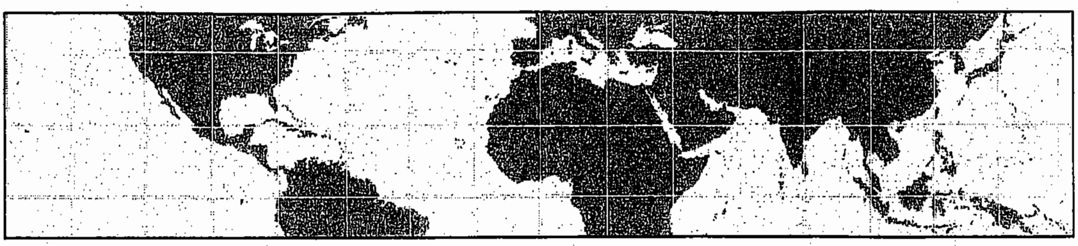
19.30. The DerivaGem Application Builder functions enable you to investigate how the prices of options calculated from a binomial tree converge to the correct value as the number of time steps increases. (See Figure 19.4 and Sample Application A in DerivaGem.) Consider a put option on a stock index where the index level is 900, the strike price is 900, the risk-free rate is 5%, the dividend yield is 2%, and the time to maturity is 2 years.

(a) Produce results similar to Sample Application A on convergence for the situation where the option is European and the volatility of the index is 20%.

(b) Produce results similar to Sample Application A on convergence for the situation where the option is American and the volatility of the index is 20%.

(c) Produce a chart showing the pricing of the American option when the volatility is 20% as a function of the number of time steps when the control variate technique is used.

(d) Suppose that the price of the American option in the market is 85.0. Produce a chart showing the implied volatility estimate as a function of the number of time steps.



CHAPTER 20

Value at Risk

Chapter 17 examined measures such as delta, gamma, and vega for describing different aspects of the risk in a portfolio of derivatives. A financial institution usually calculates each of these measures each day for every market variable to which it is exposed. Often there are hundreds, or even thousands, of these market variables. A delta–gamma–vega analysis, therefore, leads to a very large number of different risk measures being produced each day. These risk measures provide valuable information for the financial institution's traders. However, they do not provide a way of measuring the total risk to which the financial institution is exposed.

Value at Risk (VaR) is an attempt to provide a single number summarizing the total risk in a portfolio of financial assets. It has become widely used by corporate treasurers and fund managers as well as by financial institutions. Bank regulators also use VaR in determining the capital a bank is required to keep for the risks it is bearing.

This chapter explains the VaR measure and describes the two main approaches for calculating it. These are known as the *historical simulation* approach and the *model-building* approach.

20.1 THE VaR MEASURE

When using the value-at-risk measure, an analyst is interested in making a statement of the following form:

I am X percent certain there will not be a loss of more than V dollars in the next N days.

The variable V is the VaR of the portfolio. It is a function of two parameters: the time horizon (N days) and the confidence level (X%). It is the loss level over N days that has a probability of only $(100 - X)$% of being exceeded. Bank regulators require banks to calculate VaR for market risk with $N = 10$ and $X = 99$ (see the discussion in Business Snapshot 20.1).

When N days is the time horizon and X% is the confidence level, VaR is the loss corresponding to the $(100 - X)$th percentile of the distribution of the change in the value of the portfolio over the next N days. (In constructing the probability distribution of the change in value, gains are positive and losses are negative.) For example, when $N = 5$ and $X = 97$, VaR is the third percentile of the distribution of changes in the value

Business Snapshot 20.1 How Bank Regulators Use VaR

The Basel Committee on Bank Supervision is a committee of the world's bank regulators that meets regularly in Basel, Switzerland. In 1988 it published what has become known as *The 1988 BIS Accord*, or simply *The Accord*. This is an agreement between the regulators on how the capital a bank is required to hold for credit risk should be calculated. Several years later the Basel Committee published *The 1996 Amendment*, which was implemented in 1998 and required banks to hold capital for market risk as well as credit risk. *The Amendment* distinguishes between a bank's trading book and its banking book. The banking book consists primarily of loans and is not usually revalued on a regular basis for managerial and accounting purposes. The trading book consists of the myriad of different instruments that are traded by the bank (stocks, bonds, swaps, forward contracts, options, etc.) and is normally revalued daily.

The 1996 BIS Amendment calculates capital for the trading book using the VaR measure with $N = 10$ and $X = 99$. This means that it focuses on the revaluation loss over a 10-day period that is expected to be exceeded only 1% of the time. The capital it requires the bank to hold is k times this VaR measure (with an adjustment for what are termed specific risks). The multiplier k is chosen on a bank-by-bank basis by the regulators and must be at least 3.0. For a bank with excellent well-tested VaR estimation procedures, it is likely that k will be set equal to the minimum value of 3.0. For other banks it may be higher.

of the portfolio over the next 5 days. VaR is illustrated for the situation where the change in the value of the portfolio is approximately normally distributed in Figure 20.1.

VaR is an attractive measure because it is easy to understand. In essence, it asks the simple question "How bad can things get?" This is the question all senior managers want answered. They are very comfortable with the idea of compressing all the Greek letters for all the market variables underlying a portfolio into a single number.

If we accept that it is useful to have a single number to describe the risk of a portfolio, an interesting question is whether VaR is the best alternative. Some researchers have

Figure 20.1 Calculation of VaR from the probability distribution of the change in the portfolio value; confidence level is X%. Gains in portfolio value are positive; losses are negative.

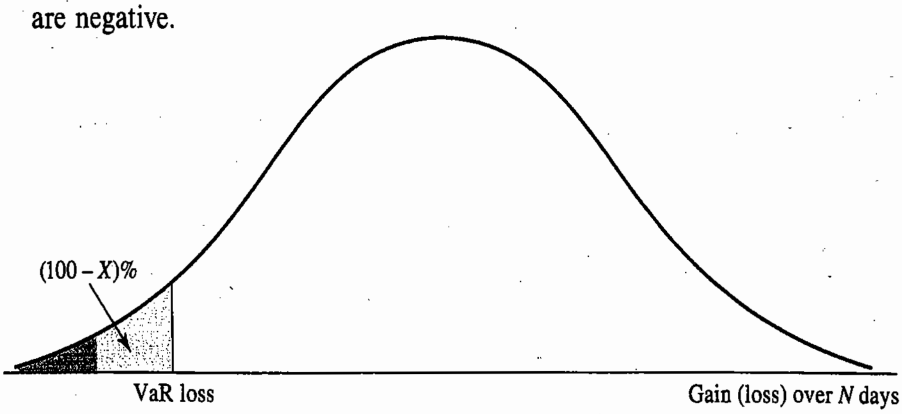
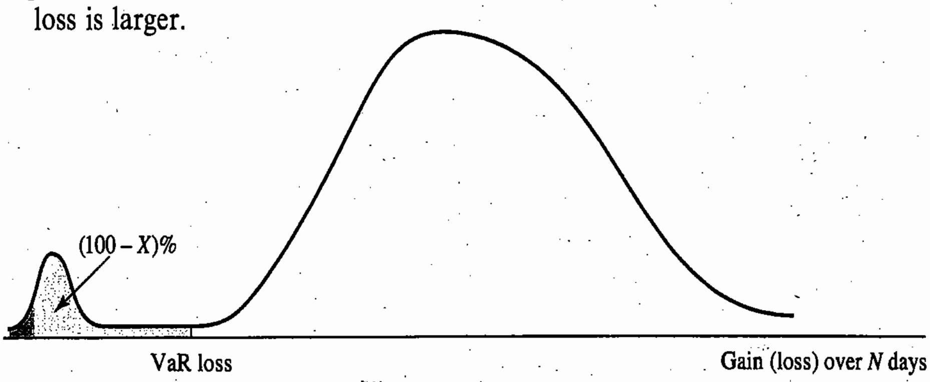


Figure 20.2 Alternative situation to Figure 20.1. VaR is the same, but the potential loss is larger.



argued that VaR may tempt traders to choose a portfolio with a return distribution similar to that in Figure 20.2. The portfolios in Figures 20.1 and 20.2 have the same VaR, but the portfolio in Figure 20.2 is much riskier because potential losses are much larger.

A measure that deals with the problem we have just mentioned is *expected shortfall*.[1] Whereas VaR asks the question "How bad can things get?", expected shortfall asks "If things do get bad, how much can the company expect to lose?" Expected shortfall is the expected loss during an N-day period conditional that an outcome in the $(100 - X)\%$ left tail of the distribution occurs. For example, with $X = 99$ and $N = 10$, the expected shortfall is the average amount the company loses over a 10-day period when the loss is in the 1% tail of the distribution.

In spite of its weaknesses, VaR (not expected shortfall) is the most popular measure of risk among both regulators and risk managers. We will therefore devote most of the rest of this chapter to how it can be measured.

The Time Horizon

VaR has two parameters: the time horizon N, measured in days, and the confidence level X. In practice, analysts almost invariably set $N = 1$ in the first instance. This is because there is not enough data to estimate directly the behavior of market variables over periods of time longer than 1 day. The usual assumption is

$$N\text{-day VaR} = 1\text{-day VaR} \times \sqrt{N}$$

This formula is exactly true when the changes in the value of the portfolio on successive days have independent identical normal distributions with mean zero. In other cases it is an approximation.

Business Snapshot 20.1 explains that regulators require a bank's capital for market risk to be at least three times the 10-day 99% VaR. Given the way a 10-day VaR is calculated, this minimum capital level is $3 \times \sqrt{10} = 9.49$ times the 1-day 99% VaR.

[1] This measure, which is also known as *C-VaR* or *tail loss*, was suggested by P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent Measures of Risk," *Mathematical Finance*, 9 (1999): 203–28. These authors define certain properties that a good risk measure should have and show that the standard VaR measure does not have all of them.

20.2 HISTORICAL SIMULATION

Historical simulation is one popular way of estimating VaR. It involves using past data in a very direct way as a guide to what might happen in the future. Suppose that VaR is to be calculated for a portfolio using a 1-day time horizon, a 99% confidence level, and 501 days of data. The first step is to identify the market variables affecting the portfolio. These will typically be exchange rates, equity prices, interest rates, and so on. Data is then collected on the movements in these market variables over the most recent 501 days. This provides 500 alternative scenarios for what can happen between today and tomorrow. Scenario 1 is where the percentage changes in the values of all variables are the same as they were between Day 0 and Day 1, scenario 2 is where they are the same as they were between Day 1 and Day 2, and so on. For each scenario, the dollar change in the value of the portfolio between today and tomorrow is calculated. This defines a probability distribution for daily changes in the value of the portfolio. The fifth-worst daily change is the first percentile of the distribution. The estimate of VaR is the loss at this first percentile point. Assuming that the last 501 days are a good guide to what could happen during the next day, the company is 99% certain that it will not take a loss greater than the VaR estimate.

The historical simulation methodology is illustrated in Tables 20.1 and 20.2. Table 20.1 shows observations on market variables over the last 501 days. The observations are taken at some particular point in time during the day (usually the close of trading). We denote the first day for which data is available as Day 0, the second as Day 1, and so on. Today is Day 500; tomorrow is Day 501.

Table 20.2 shows the values of the market variables tomorrow if their percentage changes between today and tomorrow are the same as they were between Day $i-1$ and Day i for $1 \leqslant i \leqslant 500$. The first row in Table 20.2 shows the values of market variables tomorrow assuming their percentage changes between today and tomorrow are the same as they were between Day 0 and Day 1, the second row shows the values of market variables tomorrow assuming their percentage changes between Day 1 and Day 2 occur, and so on. The 500 rows in Table 20.2 are the 500 scenarios considered.

Define v_i as the value of a market variable on Day i and suppose that today is Day m.

Table 20.1 Data for VaR historical simulation calculation.

Day	*Market variable 1*	*Market variable 2*	...	*Market variable n*
0	20.33	0.1132	...	65.37
1	20.78	0.1159	...	64.91
2	21.44	0.1162	...	65.02
3	20.97	0.1184	...	64.90
⋮	⋮	⋮	⋮	⋮
498	25.72	0.1312	...	62.22
499	25.75	0.1323	...	61.99
500	25.85	0.1343	...	62.10

Table 20.2 Scenarios generated for tomorrow (Day 501) using data in Table 20.1.

Scenario number	*Market variable 1*	*Market variable 2*	...	*Market variable n*	*Portfolio value ($ millions)*	*Change in value ($ millions)*
1	26.42	0.1375	...	61.66	23.71	0.21
2	26.67	0.1346	...	62.21	23.12	−0.38
3	25.28	0.1368	...	61.99	22.94	−0.56
⋮	⋮	⋮	⋮	⋮	⋮	⋮
499	25.88	0.1354	...	61.87	23.63	0.13
500	25.95	0.1363	...	62.21	22.87	−0.63

The ith scenario assumes that the value of the market variable tomorrow will be

$$v_m \frac{v_i}{v_{i-1}}$$

In our example, $m = 500$. For the first variable, the value today, v_{500}, is 25.85. Also $v_0 = 20.33$ and $v_1 = 20.78$. It follows that the value of the first market variable in the first scenario is

$$25.85 \times \frac{20.78}{20.33} = 26.42$$

The penultimate column of Table 20.2 shows the value of the portfolio tomorrow for each of the 500 scenarios. We suppose the value of the portfolio today is $23.50 million. This leads to the numbers in the final column for the change in the value between today and tomorrow for all the different scenarios. For Scenario 1 the change in value in our example is +$210,000, for Scenario 2 it is −$380,000, and so on.

We are interested in the 1-percentile point of the distribution of changes in the portfolio value. Because there are a total of 500 scenarios in Table 20.2 we can estimate this as the fifth-worst number in the final column of the table. Alternatively, we can use the techniques of what is known as *extreme value theory* to smooth the numbers in the left tail of the distribution in an attempt to obtain a more accurate estimate of the 1% point of the distribution.[2] As mentioned in the previous section, the N-day VaR for a 99% confidence level is calculated as $\sqrt{N}$ times the 1-day VaR.

Each day the VaR estimate in our example would be updated using the most recent 501 days of data. Consider, for example, what happens on Day 501. New values for all the market variables become available and are used to calculate a new value for our portfolio.[3] The procedure we have outlined is employed to calculate a new VaR using data on the market variables from Day 1 to Day 501. (This gives 501 observations on the percentage changes in market variables; the Day-0 values of the market variables are no longer used.) Similarly, on Day 502, data from Day 2 to Day 502 are used to determine VaR, and so on.

[2] See P. Embrechts, C. Kluppelberg, and T. Mikosch. *Modeling Extremal Events for Insurance and Finance.* New York: Springer, 1997; A. J. McNeil, "Extreme Value Theory for Risk Managers," in *Internal Modeling and CAD II.* London, Risk Books, 1999, and available from www.math.ethz.ch/~mcneil.

[3] Note that the portfolio's composition may have changed between Day 500 and Day 501.

20.3 MODEL-BUILDING APPROACH

The main alternative to historical simulation is the model-building approach. Before getting into the details of the approach, it is appropriate to mention one issue concerned with the units for measuring volatility.

Daily Volatilities

In option pricing time is usually measured in years, and the volatility of an asset is usually quoted as a "volatility per year". When using the model-building approach to calculate VaR, time is usually measured in days and the volatility of an asset is usually quoted as a "volatility per day".

What is the relationship between the volatility per year used in option pricing and the volatility per day used in VaR calculations? Let us define σ_{year} as the volatility per year of a certain asset and σ_{day} as the equivalent volatility per day of the asset. Assuming 252 trading days in a year, equation (13.2) gives the standard deviation of the continuously compounded return on the asset in 1 year as either σ_{year} or $\sigma_{\text{day}}\sqrt{252}$. It follows that

$$\sigma_{\text{year}} = \sigma_{\text{day}}\sqrt{252}$$

or

$$\sigma_{\text{day}} = \frac{\sigma_{\text{year}}}{\sqrt{252}}$$

so that daily volatility is about 6% of annual volatility.

As pointed out in Section 13.4, σ_{day} is approximately equal to the standard deviation of the percentage change in the asset price in one day. For the purposes of calculating VaR we assume exact equality. We define the daily volatility of an asset price (or any other variable) as equal to the standard deviation of the percentage change in one day.

Our discussion in the next few sections assumes that estimates of daily volatilities and correlations are available. Chapter 21 discusses how the estimates can be produced.

Single-Asset Case

Consider how VaR is calculated using the model-building approach in a very simple situation where the portfolio consists of a position in a single stock: \$10 million in shares of Microsoft. We suppose that $N = 10$ and $X = 99$, so that we are interested in the loss level over 10 days that we are 99% confident will not be exceeded. Initially, we consider a 1-day time horizon.

Assume that the volatility of Microsoft is 2% per day (corresponding to about 32% per year). Because the size of the position is \$10 million, the standard deviation of daily changes in the value of the position is 2% of \$10 million, or \$200,000.

It is customary in the model-building approach to assume that the expected change in a market variable over the time period considered is zero. This is not strictly true, but it is a reasonable assumption. The expected change in the price of a market variable over a short time period is generally small when compared with the standard deviation of the change. Suppose, for example, that Microsoft has an expected return of 20% per annum. Over a 1-day period, the expected return is 0.20/252, or about 0.08%, whereas the standard deviation of the return is 2%. Over a 10-day period, the expected return is

0.08×10, or about 0.8%, whereas the standard deviation of the return is $2\sqrt{10}$, or about 6.3%.

So far, we have established that the change in the value of the portfolio of Microsoft shares over a 1-day period has a standard deviation of \$200,000 and (at least approximately) a mean of zero. We assume that the change is normally distributed.[4] From the tables at the end of this book, $N(-2.33) = 0.01$. This means that there is a 1% probability that a normally distributed variable will decrease in value by more than 2.33 standard deviations. Equivalently, it means that we are 99% certain that a normally distributed variable will not decrease in value by more than 2.33 standard deviations. The 1-day 99% VaR for our portfolio consisting of a \$10 million position in Microsoft is therefore

$$2.33 \times 200{,}000 = \$466{,}000$$

As discussed earlier, the N-day VaR is calculated as $\sqrt{N}$ times the 1-day VaR. The 10-day 99% VaR for Microsoft is therefore

$$466{,}000 \times \sqrt{10} = \$1{,}473{,}621$$

Consider next a portfolio consisting of a \$5 million position in AT&T, and suppose the daily volatility of AT&T is 1% (approximately 16% per year). A similar calculation to that for Microsoft shows that the standard deviation of the change in the value of the portfolio in 1 day is

$$5{,}000{,}000 \times 0.01 = 50{,}000$$

Assuming the change is normally distributed, the 1-day 99% VaR is

$$50{,}000 \times 2.33 = \$116{,}500$$

and the 10-day 99% VaR is

$$116{,}500 \times \sqrt{10} = \$368{,}405$$

Two-Asset Case

Now consider a portfolio consisting of both \$10 million of Microsoft shares and \$5 million of AT&T shares. We suppose that the returns on the two shares have a bivariate normal distribution with a correlation of 0.3. A standard result in statistics tells us that, if two variables X and Y have standard deviations equal to σ_X and σ_Y with the coefficient of correlation between them equal to ρ, the standard deviation of $X + Y$ is given by

$$\sigma_{X+Y} = \sqrt{\sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y}$$

To apply this result, we set X equal to the change in the value of the position in Microsoft over a 1-day period and Y equal to the change in the value of the position in AT&T over a 1-day period, so that

$$\sigma_X = 200{,}000 \quad \text{and} \quad \sigma_Y = 50{,}000$$

[4] To be consistent with the option pricing assumption in Chapter 13, we could assume that the price of Microsoft is lognormal tomorrow. Because 1 day is such a short period of time, this is almost indistinguishable from the assumption we do make—that the change in the stock price between today and tomorrow is normal.

The standard deviation of the change in the value of the portfolio consisting of both stocks over a 1-day period is therefore

$$\sqrt{200{,}000^2 + 50{,}000^2 + 2 \times 0.3 \times 200{,}000 \times 50{,}000} = 220{,}227$$

The mean change is assumed to be zero and the change is normally distributed. So the 1-day 99% VaR is therefore

$$220{,}227 \times 2.33 = \$513{,}129$$

The 10-day 99% VaR is $\sqrt{10}$ times this, or \$1,622,657.

The Benefits of Diversification

In the example we have just considered:

1. The 10-day 99% VaR for the portfolio of Microsoft shares is \$1,473,621.
2. The 10-day 99% VaR for the portfolio of AT&T shares is \$368,405.
3. The 10-day 99% VaR for the portfolio of both Microsoft and AT&T shares is \$1,622,657.

The amount

$$(1{,}473{,}621 + 368{,}405) - 1{,}622{,}657 = \$219{,}369$$

represents the benefits of diversification. If Microsoft and AT&T were perfectly correlated, the VaR for the portfolio of both Microsoft and AT&T would equal the VaR for the Microsoft portfolio plus the VaR for the AT&T portfolio. Less than perfect correlation leads to some of the risk being "diversified away".[5]

20.4 LINEAR MODEL

The examples we have just considered are simple illustrations of the use of the linear model for calculating VaR. Suppose that a portfolio worth P consists of n assets with an amount α_i being invested in asset i ($1 \leqslant i \leqslant n$). We define Δx_i as the return on asset i in 1 day. It follows that the dollar change in the value of the investment in asset i in 1 day is $\alpha_i\,\Delta x_i$ and

$$\Delta P = \sum_{i=1}^{n} \alpha_i\,\Delta x_i \qquad \textbf{(20.1)}$$

where ΔP is the dollar change in the value of the whole portfolio in 1 day.

In the example considered in the previous section, \$10 million was invested in the first asset (Microsoft) and \$5 million was invested in the second asset (AT&T), so that (in millions of dollars) $\alpha_1 = 10$, $\alpha_2 = 5$, and

$$\Delta P = 10\Delta x_1 + 5\Delta x_2$$

If we assume that the Δx_i in equation (20.1) are multivariate normal, ΔP is normally

---

[5] Harry Markowitz was one of the first researchers to study the benefits of diversification to a portfolio manager. He was awarded a Nobel prize for this research in 1990. See H. Markowitz, "Portfolio Selection," *Journal of Finance*, 7, 1 (March 1952): 77–91.

distributed. To calculate VaR, we therefore need to calculate only the mean and standard deviation of ΔP. We assume, as discussed in the previous section, that the expected value of each Δx_i is zero. This implies that the mean of ΔP is zero.

To calculate the standard deviation of ΔP, we define σ_i as the daily volatility of the ith asset and ρ_{ij} as the coefficient of correlation between returns on asset i and asset j. This means that σ_i is the standard deviation of Δx_i, and ρ_{ij} is the coefficient of correlation between Δx_i and Δx_j. The variance of ΔP, which we will denote by σ_P^2, is given by

$$\sigma_P^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij}\alpha_i\alpha_j\sigma_i\sigma_j$$

This equation can also be written

$$\sigma_P^2 = \sum_{i=1}^{n} \alpha_i^2\sigma_i^2 + 2\sum_{i=1}^{n} \sum_{j<i} \rho_{ij}\alpha_i\alpha_j\sigma_i\sigma_j \tag{20.2}$$

The standard deviation of the change over N days is $\sigma_P\sqrt{N}$, and the 99% VaR for an N-day time horizon is $2.33\sigma_P\sqrt{N}$.

In the example considered in the previous section, $\sigma_1 = 0.02$, $\sigma_2 = 0.01$, and $\rho_{12} = 0.3$. As already noted, $\alpha_1 = 10$ and $\alpha_2 = 5$, so that

$$\sigma_P^2 = 10^2 \times 0.02^2 + 5^2 \times 0.01^2 + 2 \times 10 \times 5 \times 0.3 \times 0.02 \times 0.01 = 0.0485$$

and $\sigma_P = 0.220$. This is the standard deviation of the change in the portfolio value per day (in millions of dollars). The 10-day 99% VaR is $2.33 \times 0.220 \times \sqrt{10} =$ \$1.623 million. This agrees with the calculation in the previous section.

Handling Interest Rates

It is out of the question in the model-building approach to define a separate market variable for every single bond price or interest rate to which a company is exposed. Some simplifications are necessary when the model-building approach is used. One possibility is to assume that only parallel shifts in the yield curve occur. It is then necessary to define only one market variable: the size of the parallel shift. The changes in the value of a bond portfolio can then be calculated using the duration relationship

$$\Delta P = -DP\Delta y$$

where P is the value of the portfolio, ΔP is the its change in P in one day, D is the modified duration of the portfolio, and Δy is the parallel shift in 1 day.

This approach does not usually give enough accuracy. The procedure usually followed is to choose as market variables the prices of zero-coupon bonds with standard maturities: 1 month, 3 months, 6 months, 1 year, 2 years, 5 years, 7 years, 10 years, and 30 years. For the purposes of calculating VaR, the cash flows from instruments in the portfolio are mapped into cash flows occurring on the standard maturity dates. Consider a \$1 million position in a Treasury bond lasting 1.2 years that pays a coupon of 6% semiannually. Coupons are paid in 0.2, 0.7, and 1.2 years, and the principal is paid in 1.2 years. This bond is, therefore, in the first instance regarded as a \$30,000 position in 0.2-year zero-coupon bond plus a \$30,000 position in a 0.7-year zero-coupon bond plus a \$1.03 million position in a 1.2-year zero-coupon bond. The position in the 0.2-year bond is then replaced by an equivalent position in 1-month and 3-month zero-coupon bonds; the position in the 0.7-year bond is replaced by an equivalent position in 6-month and 1-year zero-coupon bonds; and the position in the 1.2-year bond is replaced by an equivalent position in 1-year and 2-year zero-coupon bonds. The result is that the position in the 1.2-year coupon-bearing bond is for VaR purposes regarded as a position in zero-coupon bonds having maturities of 1 month, 3 months, 6 months, 1 year, and 2 years.

This procedure is known as *cash-flow mapping*. One way of doing it is explained in the appendix at the end of this chapter. Note that cash-flow mapping is not necessary when the historical simulation approach is used. This is because the complete term structure of interest rates can be calculated for each of the scenarios considered.

Applications of the Linear Model

The simplest application of the linear model is to a portfolio with no derivatives consisting of positions in stocks, bonds, foreign exchange, and commodities. In this case, the change in the value of the portfolio is linearly dependent on the percentage changes in the prices of the assets comprising the portfolio. Note that, for the purposes of VaR calculations, all asset prices are measured in the domestic currency. The market variables considered by a large bank in the United States are therefore likely to include the value of the Nikkei 225 index measured in dollars, the price of a 10-year sterling zero-coupon bond measured in dollars, and so on.

An example of a derivative that can be handled by the linear model is a forward contract to buy a foreign currency. Suppose the contract matures at time T. It can be regarded as the exchange of a foreign zero-coupon bond maturing at time T for a domestic zero-coupon bond maturing at time T. For the purposes of calculating VaR, the forward contract is therefore treated as a long position in the foreign bond combined with a short position in the domestic bond. Each bond can be handled using a cash-flow mapping procedure.

Consider next an interest rate swap. As explained in Chapter 7, this can be regarded as the exchange of a floating-rate bond for a fixed-rate bond. The fixed-rate bond is a regular coupon-bearing bond. The floating-rate bond is worth par just after the next payment date. It can be regarded as a zero-coupon bond with a maturity date equal to the next payment date. The interest rate swap therefore reduces to a portfolio of long and short positions in bonds and can be handled using a cash-flow mapping procedure.

The Linear Model and Options

We now consider how we might try to use the linear model when there are options. Consider first a portfolio consisting of options on a single stock whose current price is S. Suppose that the delta of the position (calculated in the way described in Chapter 17) is δ.[6] Since δ is the rate of change of the value of the portfolio with S, it is approximately true that

$$\delta = \frac{\Delta P}{\Delta S}$$

[6] Normally we denote the delta and gamma of a portfolio by Δ and Γ. In this section and the next, we use the lower case Greek letters δ and γ to avoid overworking Δ.

or

$$\Delta P = \delta \, \Delta S \tag{20.3}$$

where ΔS is the dollar change in the stock price in 1 day and ΔP is, as usual, the dollar change in the portfolio in 1 day. Define Δx as the percentage change in the stock price in 1 day, so that

$$\Delta x = \frac{\Delta S}{S}$$

It follows that an approximate relationship between ΔP and Δx is

$$\Delta P = S\delta \, \Delta x$$

When we have a position in several underlying market variables that includes options, we can derive an approximate linear relationship between ΔP and the Δx_i similarly. This relationship is

$$\Delta P = \sum_{i=1}^{n} S_i \delta_i \, \Delta x_i \tag{20.4}$$

where S_i is the value of the ith market variable and δ_i is the delta of the portfolio with respect to the ith market variable. This corresponds to equation (20.1):

$$\Delta P = \sum_{i=1}^{n} \alpha_i \, \Delta x_i \tag{20.5}$$

with $\alpha_i = S_i \delta_i$. Equation (20.2) can therefore be used to calculate the standard deviation of ΔP.

Example 20.1

A portfolio consists of options on Microsoft and AT&T. The options on Microsoft have a delta of 1,000, and the options on AT&T have a delta of 20,000. The Microsoft share price is \$120, and the AT&T share price is \$30. From equation (20.4), it is approximately true that

$$\Delta P = 120 \times 1{,}000 \times \Delta x_1 + 30 \times 20{,}000 \times \Delta x_2$$

or

$$\Delta P = 120{,}000 \Delta x_1 + 600{,}000 \Delta x_2$$

where Δx_1 and Δx_2 are the returns from Microsoft and AT&T in 1 day and ΔP is the resultant change in the value of the portfolio. (The portfolio is assumed to be equivalent to an investment of \$120,000 in Microsoft and \$600,000 in AT&T.) Assuming that the daily volatility of Microsoft is 2% and the daily volatility of AT&T is 1% and the correlation between the daily changes is 0.3, the standard deviation of ΔP (in thousands of dollars) is

$$\sqrt{(120 \times 0.02)^2 + (600 \times 0.01)^2 + 2 \times 120 \times 0.02 \times 600 \times 0.01 \times 0.3} = 7.099$$

Since $N(-1.65) = 0.05$, the 5-day 95% VaR is $1.65 \times \sqrt{5} \times 7{,}099 = \$26{,}193$.

20.5 QUADRATIC MODEL

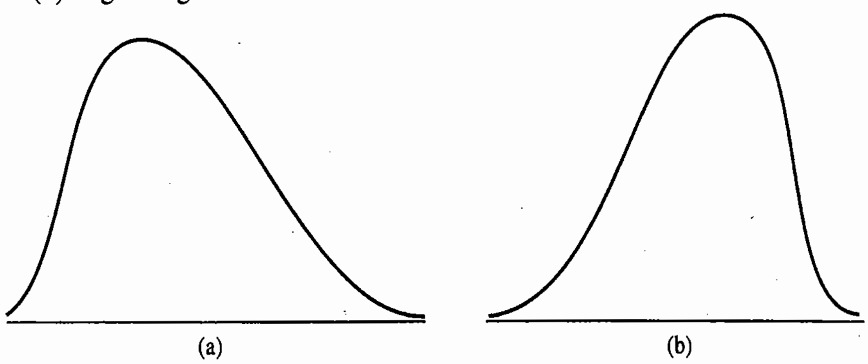
When a portfolio includes options, the linear model is an approximation. It does not take account of the gamma of the portfolio. As discussed in Chapter 17, delta is defined as the rate of change of the portfolio value with respect to an underlying market variable and gamma is defined as the rate of change of the delta with respect to the market variable. Gamma measures the curvature of the relationship between the portfolio value and an underlying market variable.

Figure 20.3 shows the impact of a nonzero gamma on the probability distribution of the value of the portfolio. When gamma is positive, the probability distribution tends to be positively skewed; when gamma is negative, it tends to be negatively skewed. Figures 20.4 and 20.5 illustrate the reason for this result. Figure 20.4 shows the relationship between the value of a long call option and the price of the underlying asset. A long call is an example of an option position with positive gamma. The figure shows that, when the probability distribution for the price of the underlying asset at the end of 1 day is normal, the probability distribution for the option price is positively skewed.[7] Figure 20.5 shows the relationship between the value of a short call position and the price of the underlying asset. A short call position has a negative gamma. In this case, we see that a normal distribution for the price of the underlying asset at the end of 1 day gets mapped into a negatively skewed distribution for the value of the option position.

The VaR for a portfolio is critically dependent on the left tail of the probability distribution of the portfolio value. For example, when the confidence level used is 99%, the VaR is the value in the left tail below which there is only 1% of the distribution. As indicated in Figures 20.3(a) and 20.4, a positive gamma portfolio tends to have a less heavy left tail than the normal distribution. If the distribution of ΔP is normal, the calculated VaR tends to be too high. Similarly, as indicated in Figures 20.3(b) and 20.5, a negative gamma portfolio tends to have a heavier left tail than the normal distribution. If the distribution of ΔP is normal, the calculated VaR tends to be too low.

For a more accurate estimate of VaR than that given by the linear model, both delta and gamma measures can be used to relate ΔP to the Δx_i. Consider a portfolio

Figure 20.3 Probability distribution for value of portfolio: (a) positive gamma; (b) negative gamma.



[7] As mentioned in footnote 4, we can use the normal distribution as an approximation to the lognormal distribution in VaR calculations.

Figure 20.4 Translation of normal probability distribution for asset into probability distribution for value of a long call on asset.

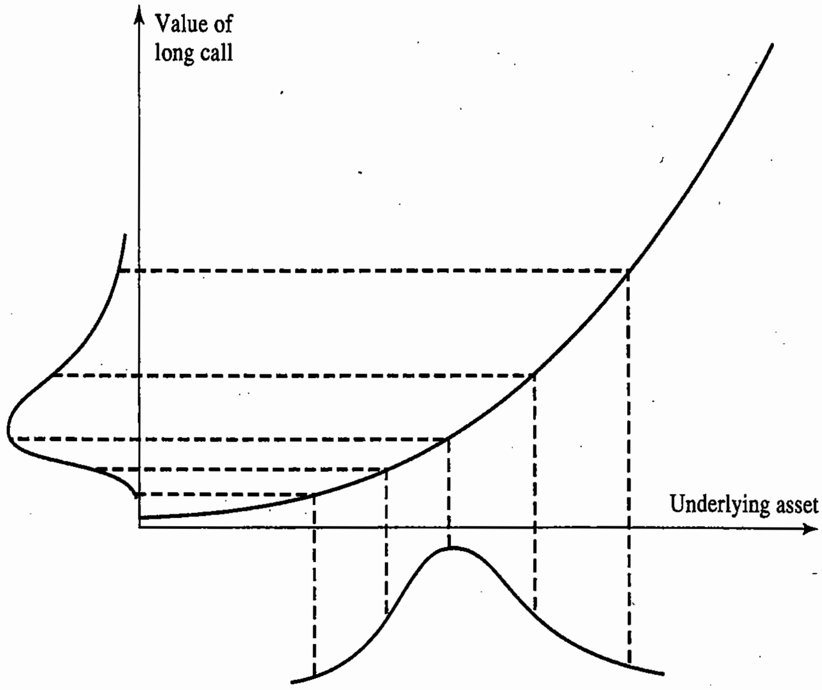
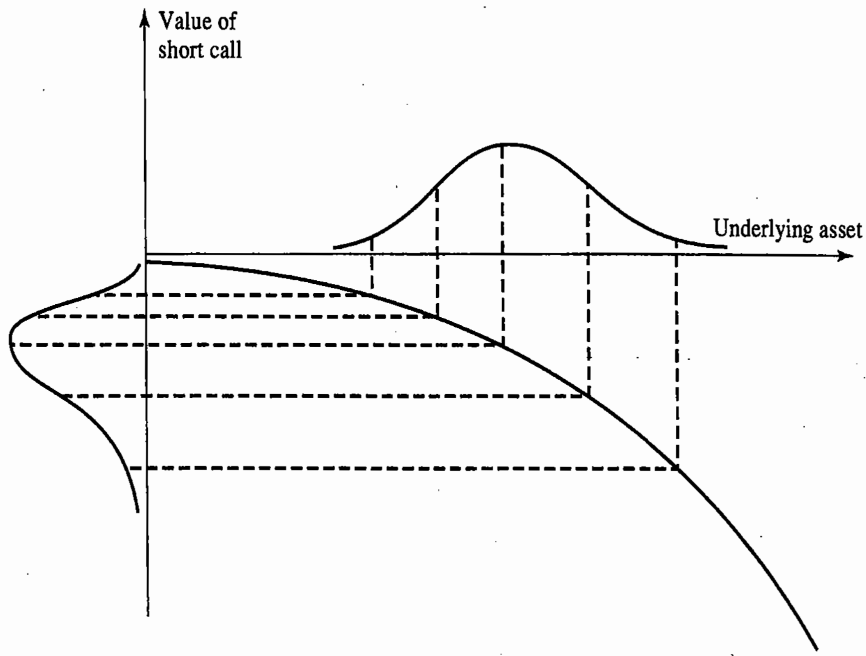


Figure 20.5 Translation of normal probability distribution for asset into probability distribution for value of a short call on asset.



dependent on a single asset whose price is S. Suppose δ and γ are the delta and gamma of the portfolio. From the appendix to Chapter 17, the equation

$$\Delta P = \delta\,\Delta S + \tfrac{1}{2}\gamma(\Delta S)^2$$

is an improvement over the approximation in equation (20.3).[8] Setting

$$\Delta x = \frac{\Delta S}{S}$$

reduces this to

$$\Delta P = S\delta\,\Delta x + \tfrac{1}{2}S^2\gamma(\Delta x)^2 \tag{20.6}$$

More generally for a portfolio with n underlying market variables, with each instrument in the portfolio being dependent on only one of the market variables, equation (20.6) becomes

$$\Delta P = \sum_{i=1}^{n} S_i\delta_i\,\Delta x_i + \sum_{i=1}^{n} \tfrac{1}{2}S_i^2\gamma_i\,(\Delta x_i)^2$$

where S_i is the value of the ith market variable, and δ_i and γ_i are the delta and gamma of the portfolio with respect to the ith market variable. When individual instruments in the portfolio may be dependent on more than one market variable, this equation takes the more general form

$$\Delta P = \sum_{i=1}^{n} S_i\delta_i\,\Delta x_i + \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{2}S_iS_j\gamma_{ij}\,\Delta x_i\,\Delta x_j \tag{20.7}$$

where γ_{ij} is a "cross gamma" defined as

$$\gamma_{ij} = \frac{\partial^2 P}{\partial S_i\,\partial S_j}$$

Equation (20.7) is not as easy to work with as equation (20.5), but it can be used to calculate moments for ΔP. A result in statistics known as the Cornish–Fisher expansion can be used to estimate percentiles of the probability distribution from the moments.[9]

20.6 MONTE CARLO SIMULATION

As an alternative to the procedure described so far, the model-building approach can be implemented using Monte Carlo simulation to generate the probability distribution for

[8] The Taylor series expansion in the appendix to Chapter 17 suggests the approximation

$$\Delta P = \Theta\,\Delta t + \delta\,\Delta S + \tfrac{1}{2}\gamma(\Delta S)^2$$

when terms of higher order than Δt are ignored. In practice, the $\Theta\,\Delta t$ term is so small that it is usually ignored.

[9] See Technical Note 10 on the author's website for details of the calculation of moments and the use of Cornish–Fisher expansions. When there is a single underlying variable, $E(\Delta P) = 0.5S^2\gamma\sigma^2$, $E(\Delta P^2) = S^2\delta^2\sigma^2 + 0.75S^4\gamma^2\sigma^4$, and $E(\Delta P^3) = 4.5S^4\delta^2\gamma\sigma^4 + 1.875S^6\gamma^3\sigma^6$, where S is the value of the variable and σ is its daily volatility. Sample Application E in the DerivaGem Application Builder implements the Cornish–Fisher expansion method for this case.

ΔP. Suppose we wish to calculate a 1-day VaR for a portfolio. The procedure is as follows:

1. Value the portfolio today in the usual way using the current values of market variables.
2. Sample once from the multivariate normal probability distribution of the Δx_i.[10]
3. Use the values of the Δx_i that are sampled to determine the value of each market variable at the end of one day.
4. Revalue the portfolio at the end of the day in the usual way.
5. Subtract the value calculated in Step 1 from the value in Step 4 to determine a sample ΔP.
6. Repeat Steps 2 to 5 many times to build up a probability distribution for ΔP.

The VaR is calculated as the appropriate percentile of the probability distribution of ΔP. Suppose, for example, that we calculate 5,000 different sample values of ΔP in the way just described. The 1-day 99% VaR is the value of ΔP for the 50th worst outcome; the 1-day VaR 95% is the value of ΔP for the 250th worst outcome; and so on.[11] The N-day VaR is usually assumed to be the 1-day VaR multiplied by $\sqrt{N}$.[12]

The drawback of Monte Carlo simulation is that it tends to be slow because a company's complete portfolio (which might consist of hundreds of thousands of different instruments) has to be revalued many times.[13] One way of speeding things up is to assume that equation (20.7) describes the relationship between ΔP and the Δx_i. We can then jump straight from Step 2 to Step 5 in the Monte Carlo simulation and avoid the need for a complete revaluation of the portfolio. This is sometimes referred to as the *partial simulation approach*.

20.7 COMPARISON OF APPROACHES

We have discussed two methods for estimating VaR: the historical simulation approach and the model-building approach. The advantages of the model-building approach are that results can be produced very quickly and it can easily be used in conjunction with volatility updating schemes such as those we will describe in the next chapter. The main disadvantage of the model-building approach is that it assumes that the market variables have a multivariate normal distribution. In practice, daily changes in market variables often have distributions that are quite different from normal (see, e.g., Table 18.1).

The historical simulation approach has the advantage that historical data determine the joint probability distribution of the market variables. It also avoids the need for cash-flow mapping (see Problem 20.2). The main disadvantages of historical simulation

---

[10] One way of doing so is given in Chapter 19.

[11] As in the case of historical simulation, extreme value theory can be used to "smooth the tails" so that better estimates of extreme percentiles are obtained.

[12] This is only approximately true when the portfolio includes options, but it is the assumption that is made in practice for most VaR calculation methods.

[13] An approach for limiting the number of portfolio revaluations is proposed in F. Jamshidian and Y. Zhu "Scenario simulation model: theory and methodology," *Finance and Stochastics*, 1 (1997), 43–67.

are that it is computationally slow and does not easily allow volatility updating schemes to be used.[14]

One disadvantage of the model-building approach is that it tends to give poor results for low-delta portfolios (see Problem 20.21).

20.8 STRESS TESTING AND BACK TESTING

In addition to calculating VaR, many companies carry out what is known as *stress testing*. This involves estimating how a company's portfolio would have performed under some of the most extreme market moves seen in the last 10 to 20 years.

For example, to test the impact of an extreme movement in US equity prices, a company might set the percentage changes in all market variables equal to those on October 19, 1987 (when the S&P 500 moved by 22.3 standard deviations). If this is considered to be too extreme, the company might choose January 8, 1988 (when the S&P 500 moved by 6.8 standard deviations). To test the effect of extreme movements in UK interest rates, the company might set the percentage changes in all market variables equal to those on April 10, 1992 (when 10-year bond yields moved by 7.7 standard deviations).

The scenarios used in stress testing are also sometimes generated by senior management. One technique sometimes used is to ask senior management to meet periodically and "brainstorm" to develop extreme scenarios that might occur given the current economic environment and global uncertainties.

Stress testing can be considered as a way of taking into account extreme events that do occur from time to time but are virtually impossible according to the probability distributions assumed for market variables. A 5-standard-deviation daily move in a market variable is one such extreme event. Under the assumption of a normal distribution, it happens about once every 7,000 years, but, in practice, it is not uncommon to see a 5-standard-deviation daily move once or twice every 10 years.

Whatever the method used for calculating VaR, an important reality check is *back testing*. It involves testing how well the VaR estimates would have performed in the past. Suppose that we are calculating a 1-day 99% VaR. Back testing would involve looking at how often the loss in a day exceeded the 1-day 99% VaR that would have been calculated for that day. If this happened on about 1% of the days, we can feel reasonably comfortable with the methodology for calculating VaR. If it happened on, say, 7% of days, the methodology is suspect.

20.9 PRINCIPAL COMPONENTS ANALYSIS

One approach to handling the risk arising from groups of highly correlated market variables is principal components analysis. This takes historical data on movements in the market variables and attempts to define a set of components or factors that explain the movements.

[14] For a way of adapting the historical simulation approach to incorporate volatility updating, see J. Hull and A. White. "Incorporating volatility updating into the historical simulation method for value-at-risk," *Journal of Risk* 1, No. 1 (1998): 5–19.

Table 20.3 Factor loadings for US Treasury data.

	PC1	*PC2*	*PC3*	*PC4*	*PC5*	*PC6*	*PC7*	*PC8*	*PC9*	*PC10*
3m	0.21	−0.57	0.50	0.47	−0.39	−0.02	0.01	0.00	0.01	0.00
6m	0.26	−0.49	0.23	−0.37	0.70	0.01	−0.04	−0.02	−0.01	0.00
12m	0.32	−0.32	−0.37	−0.58	−0.52	−0.23	−0.04	−0.05	0.00	0.01
2y	0.35	−0.10	−0.38	0.17	0.04	0.59	0.56	0.12	−0.12	−0.05
3y	0.36	0.02	−0.30	0.27	0.07	0.24	−0.79	0.00	−0.09	−0.00
4y	0.36	0.14	−0.12	0.25	0.16	−0.63	0.15	0.55	−0.14	−0.08
5y	0.36	0.17	−0.04	0.14	0.08	−0.10	0.09	−0.26	0.71	0.48
7y	0.34	0.27	0.15	0.01	0.00	−0.12	0.13	−0.54	0.00	−0.68
10y	0.31	0.30	0.28	−0.10	−0.06	0.01	0.03	−0.23	−0.63	0.52
30y	0.25	0.33	0.46	−0.34	−0.18	0.33	−0.09	0.52	0.26	−0.13

The approach is best illustrated with an example. The market variables we will consider are 10 US Treasury rates with maturities between 3 months and 30 years. Tables 20.3 and 20.4 shows results produced by Frye for these market variables using 1,543 daily observations between 1989 and 1995.[15] The first column in Table 20.3 shows the maturities of the rates that were considered. The remaining 10 columns in the table show the 10 factors (or principal components) describing the rate moves. The first factor, shown in the column labeled PC1, corresponds to a roughly parallel shift in the yield curve. When there is one unit of that factor, the 3-month rate increases by 0.21 basis points, the 6-month rate increases by 0.26 basis points, and so on. The second factor is shown in the column labeled PC2. It corresponds to a "twist" or "steepening" of the yield curve. Rates between 3 months and 2 years move in one direction; rates between 3 years and 30 years move in the other direction. The third factor corresponds to a "bowing" of the yield curve. Rates at the short end and long end of the yield curve move in one direction; rates in the middle move in the other direction. The interest rate move for a particular factor is known as *factor loading*. In the example, the first factor's loading for the three-month rate is 0.21.[16]

Because there are 10 rates and 10 factors, the interest rate changes observed on any given day can always be expressed as a linear sum of the factors by solving a set of 10 simultaneous equations. The quantity of a particular factor in the interest rate changes on a particular day is known as the *factor score* for that day.

The importance of a factor is measured by the standard deviation of its factor score. The standard deviations of the factor scores in our example are shown in Table 20.4 and

Table 20.4 Standard deviation of factor scores (basis points).

PC1	*PC2*	*PC3*	*PC4*	*PC5*	*PC6*	*PC7*	*PC8*	*PC9*	*PC10*
17.49	6.05	3.10	2.17	1.97	1.69	1.27	1.24	0.80	0.79

[15] See J. Frye, "Principals of Risk: Finding VAR through Factor-Based Interest Rate Scenarios," in *VAR: Understanding and Applying Value at Risk*, pp. 275–88. London: Risk Publications, 1997.

[16] The factor loadings have the property that the sum of their squares for each factor is 1.0.

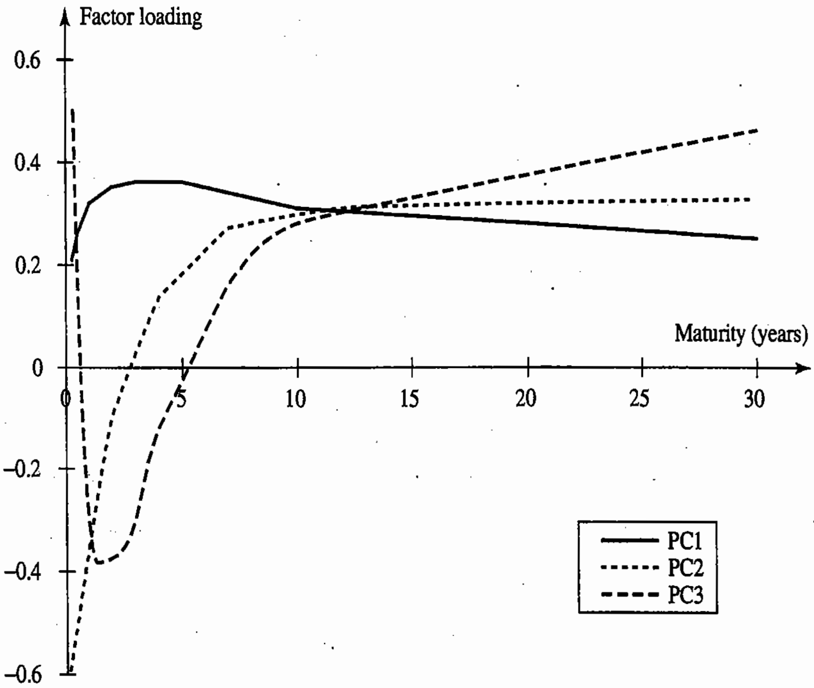
the factors are listed in order of their importance. The numbers in Table 20.4 are measured in basis points. A quantity of the first factor equal to one standard deviation, therefore, corresponds to the 3-month rate moving by $0.21 \times 17.49 = 3.67$ basis points, the 6-month rate moving by $0.26 \times 17.49 = 4.55$ basis points, and so on.

The technical details of how the factors are determined are not covered here. It is sufficient for us to note that the factors are chosen so that the factor scores are uncorrelated. For instance, in our example, the first factor score (amount of parallel shift) is uncorrelated with the second factor score (amount of twist) across the 1,543 days. The variances of the factor scores (i.e., the squares of the standard deviations) have the property that they add up to the total variance of the data. From Table 20.4, the total variance of the original data (i.e., sum of the variance of the observations on the 3-month rate, the variance of the observations on the 6-month rate, and so on) is

$$17.49^2 + 6.05^2 + 3.10^2 + \cdots + 0.79^2 = 367.9$$

From this it can be seen that the first factor accounts for $17.49^2/367.9 = 83.1\%$ of the variance in the original data; the first two factors account for $(17.49^2 + 6.05^2)/367.9 = 93.1\%$ of the variance in the data; the third factor accounts for a further 2.8% of the variance. This shows most of the risk in interest rate moves is accounted for by the first two or three factors. It suggests that we can relate the risks in a portfolio of interest rate dependent instruments to movements in these factors instead of considering all 10 interest rates. The three most important factors from Table 20.3 are plotted in Figure 20.6.[17]

Figure 20.6 The three most important factors driving yield curve movements.



[17] Similar results to those described here, in respect of the nature of the factors and the amount of the total risk they account for, are obtained when a principal components analysis is used to explain the movements in almost any yield curve in any country.

Using Principal Components Analysis to Calculate VaR

To illustrate how a principal components analysis can be used to calculate VaR, consider a portfolio with the exposures to interest rate moves shown in Table 20.5. A 1-basis-point change in the 1-year rate causes the portfolio value to increase by \$10 million, a 1-basis-point change in the 2-year rate causes it to increase by \$4 million, and so on. Suppose the first two factors are used to model rate moves. (As mentioned above, this captures 93.1% of the uncertainty in rate moves.) Using the data in Table 20.3, the exposure to the first factor (measured in millions of dollars per factor score basis point) is

$$10 \times 0.32 + 4 \times 0.35 - 8 \times 0.36 - 7 \times 0.36 + 2 \times 0.36 = -0.08$$

and the exposure to the second factor is

$$10 \times (-0.32) + 4 \times (-0.10) - 8 \times 0.02 - 7 \times 0.14 + 2 \times 0.17 = -4.40$$

Suppose that f_1 and f_2 are the factor scores (measured in basis points). The change in the portfolio value is, to a good approximation, given by

$$\Delta P = -0.08 f_1 - 4.40 f_2$$

The factor scores are uncorrelated and have the standard deviations given in Table 20.4. The standard deviation of ΔP is therefore

$$\sqrt{0.08^2 \times 17.49^2 + 4.40^2 \times 6.05^2} = 26.66$$

Hence, the 1-day 99% VaR is $26.66 \times 2.33 = 62.12$. Note that the data in Table 20.5 are such that there is very little exposure to the first factor and significant exposure to the second factor. Using only one factor would significantly understate VaR (see Problem 20.13). The duration-based method for handling interest rates, mentioned in Section 20.4, would also significantly understate VaR as it considers only parallel shifts in the yield curve.

A principal components analysis can in theory be used for market variables other than interest rates. Suppose that a financial institution has exposures to a number of different stock indices. A principal components analysis can be used to identify factors describing movements in the indices and the most important of these can be used to replace the market indices in a VaR analysis. How effective a principal components analysis is for a group of market variables depends on how closely correlated they are.

As explained earlier in the chapter, VaR is usually calculated by relating the actual changes in a portfolio to percentage changes in market variables (the Δx_i). For a VaR calculation, it may therefore be most appropriate to carry out a principal components analysis on percentage changes in market variables rather than actual changes.

Table 20.5 Change in portfolio value for a 1-basis-point rate move (\$ millions).

1-year rate	*2-year rate*	*3-year rate*	*4-year rate*	*5-year rate*
+10	+4	−8	−7	+2

SUMMARY

A value at risk (VaR) calculation is aimed at making a statement of the form: "We are X percent certain that we will not lose more than V dollars in the next N days." The variable V is the VaR, $X\%$ is the confidence level, and N days is the time horizon.

One approach to calculating VaR is historical simulation. This involves creating a database consisting of the daily movements in all market variables over a period of time. The first simulation trial assumes that the percentage changes in each market variable are the same as those on the first day covered by the database; the second simulation trial assumes that the percentage changes are the same as those on the second day; and so on. The change in the portfolio value, ΔP, is calculated for each simulation trial, and the VaR is calculated as the appropriate percentile of the probability distribution of ΔP.

An alternative is the model-building approach. This is relatively straightforward if two assumptions can be made:

1. The change in the value of the portfolio (ΔP) is linearly dependent on percentage changes in market variables.
2. The percentage changes in market variables are multivariate normally distributed.

The probability distribution of ΔP is then normal, and there are analytic formulas for relating the standard deviation of ΔP to the volatilities and correlations of the underlying market variables. The VaR can be calculated from well-known properties of the normal distribution.

When a portfolio includes options, ΔP is not linearly related to the percentage changes in market variables. From knowledge of the gamma of the portfolio, we can derive an approximate quadratic relationship between ΔP and percentage changes in market variables. Monte Carlo simulation can then be used to estimate VaR.

In the next chapter we discuss how volatilities and correlations can be estimated and monitored.

FURTHER READING

Artzner P., F. Delbaen, J.-M. Eber, and D. Heath. "Coherent Measures of Risk," *Mathematical Finance*, 9 (1999): 203–28.

Basak, S., and A. Shapiro. "Value-at-Risk-Based Risk Management: Optimal Policies and Asset Prices," *Review of Financial Studies*, 14, 2 (2001): 371–405.

Beder, T. "VaR: Seductive but Dangerous," *Financial Analysts Journal*, 51, 5 (1995): 12–24.

Boudoukh, J., M. Richardson, and R. Whitelaw. "The Best of Both Worlds," *Risk*, May 1998: 64–67.

Dowd, K. *Beyond Value at Risk: The New Science of Risk Management*. New York: Wiley, 1998.

Duffie, D., and J. Pan. "An Overview of Value at Risk," *Journal of Derivatives*, 4, 3 (Spring 1997): 7–49.

Embrechts, P., C. Kluppelberg, and T. Mikosch. *Modeling Extremal Events for Insurance and Finance*. New York: Springer, 1997.

Frye, J. "Principals of Risk: Finding VAR through Factor-Based Interest Rate Scenarios" in *VAR: Understanding and Applying Value at Risk*, pp. 275–88. London: Risk Publications, 1997.

Hendricks, D. "Evaluation of Value-at-Risk Models Using Historical Data,"*Economic Policy Review*, Federal Reserve Bank of New York, 2 (April 1996): 39–69.

Hopper, G. "Value at Risk: A New Methodology for Measuring Portfolio Risk," *Business Review*, Federal Reserve Bank of Philadelphia, July/August 1996: 19–29.

Hua P., and P. Wilmott, "Crash Courses," *Risk*, June 1997: 64–67.

Hull, J. C., and A. White. "Value at Risk When Daily Changes in Market Variables Are Not Normally Distributed," *Journal of Derivatives*, 5 (Spring 1998): 9–19.

Hull, J. C., and A. White. "Incorporating Volatility Updating into the Historical Simulation Method for Value at Risk," *Journal of Risk*, 1, 1 (1998): 5–19.

Jackson, P., D. J. Maude, and W. Perraudin. "Bank Capital and Value at Risk." *Journal of Derivatives*, 4, 3 (Spring 1997): 73–90.

Jamshidian, F., and Y. Zhu. "Scenario Simulation Model: Theory and Methodology," *Finance and Stochastics*, 1 (1997): 43–67.

Jorion, P. *Value at Risk*, 3rd edn. McGraw-Hill, 2007.

Longin, F. M. "Beyond the VaR," *Journal of Derivatives*, 8, 4 (Summer 2001): 36–48.

Marshall, C., and M. Siegel. "Value at Risk: Implementing a Risk Measurement Standard," *Journal of Derivatives* 4, 3 (Spring 1997): 91–111.

McNeil, A. J. "Extreme Value Theory for Risk Managers," in *Internal Modeling and CAD II*, London: Risk Books, 1999. See also: www.math.ethz.ch/~mcneil.

Neftci, S. N. "Value at Risk Calculations, Extreme Events and Tail Estimation," *Journal of Derivatives*, 7, 3 (Spring 2000): 23–38.

Rich, D. "Second Generation VaR and Risk-Adjusted Return on Capital," *Journal of Derivatives*, 10, 4 (Summer 2003): 51–61.

Questions and Problems (Answers in Solutions Manual)

20.1. Consider a position consisting of a \$100,000 investment in asset A and a \$100,000 investment in asset B. Assume that the daily volatilities of both assets are 1% and that the coefficient of correlation between their returns is 0.3. What is the 5-day 99% VaR for the portfolio?

20.2. Describe three ways of handling instruments that are dependent on interest rates when the model-building approach is used to calculate VaR. How would you handle these instruments when historical simulation is used to calculate VaR?

20.3. A financial institution owns a portfolio of options on the US dollar–sterling exchange rate. The delta of the portfolio is 56.0. The current exchange rate is 1.5000. Derive an approximate linear relationship between the change in the portfolio value and the percentage change in the exchange rate. If the daily volatility of the exchange rate is 0.7%, estimate the 10-day 99% VaR.

20.4. Suppose you know that the gamma of the portfolio in the previous question is 16.2. How does this change your estimate of the relationship between the change in the portfolio value and the percentage change in the exchange rate?

20.5. Suppose that the daily change in the value of a portfolio is, to a good approximation, linearly dependent on two factors, calculated from a principal components analysis. The delta of a portfolio with respect to the first factor is 6 and the delta with respect to the second factor is −4. The standard deviations of the factor are 20 and 8, respectively. What is the 5-day 90% VaR?

20.6. Suppose that a company has a portfolio consisting of positions in stocks, bonds, foreign exchange, and commodities. Assume that there are no derivatives. Explain the assumptions underlying (a) the linear model and (b) the historical simulation model for calculating VaR.

20.7. Explain how an interest rate swap is mapped into a portfolio of zero-coupon bonds with standard maturities for the purposes of a VaR calculation.

20.8. Explain the difference between value at risk and expected shortfall.

20.9. Explain why the linear model can provide only approximate estimates of VaR for a portfolio containing options.

20.10. Verify that the 0.3-year zero-coupon bond in the cash-flow mapping example in the appendix to this chapter is mapped into a \$37,397 position in a 3-month bond and a \$11,793 position in a 6-month bond.

20.11. Suppose that the 5-year rate is 6%, the 7-year rate is 7% (both expressed with annual compounding), the daily volatility of a 5-year zero-coupon bond is 0.5%, and the daily volatility of a 7-year zero-coupon bond is 0.58%. The correlation between daily returns on the two bonds is 0.6. Map a cash flow of \$1,000 received at time 6.5 years into a position in a 5-year bond and a position in a 7-year bond using the approach in the appendix. What cash flows in 5 and 7 years are equivalent to the 6.5-year cash flow?

20.12. Some time ago a company entered into a forward contract to buy £1 million for \$1.5 million. The contract now has 6 months to maturity. The daily volatility of a 6-month zero-coupon sterling bond (when its price is translated to dollars) is 0.06% and the daily volatility of a 6-month zero-coupon dollar bond is 0.05%. The correlation between returns from the two bonds is 0.8. The current exchange rate is 1.53. Calculate the standard deviation of the change in the dollar value of the forward contract in 1 day. What is the 10-day 99% VaR? Assume that the 6-month interest rate in both sterling and dollars is 5% per annum with continuous compounding.

20.13. The text calculates a VaR estimate for the example in Table 20.5 assuming two factors. How does the estimate change if you assume (a) one factor and (b) three factors.

20.14. A bank has a portfolio of options on an asset. The delta of the options is –30 and the gamma is –5. Explain how these numbers can be interpreted. The asset price is 20 and its volatility is 1% per day. Adapt Sample Application E in the DerivaGem Application Builder software to calculate VaR.

20.15. Suppose that in Problem 20.14 the vega of the portfolio is –2 per 1% change in the annual volatility. Derive a model relating the change in the portfolio value in 1 day to delta, gamma, and vega. Explain without doing detailed calculations how you would use the model to calculate a VaR estimate.

Assignment Questions

20.16. A company has a position in bonds worth \$6 million. The modified duration of the portfolio is 5.2 years. Assume that only parallel shifts in the yield curve can take place and that the standard deviation of the daily yield change (when yield is measured in percent) is 0.09. Use the duration model to estimate the 20-day 90% VaR for the portfolio. Explain carefully the weaknesses of this approach to calculating VaR. Explain two alternatives that give more accuracy.

20.17. Consider a position consisting of a \$300,000 investment in gold and a \$500,000 investment in silver. Suppose that the daily volatilities of these two assets are 1.8% and 1.2%, respectively, and that the coefficient of correlation between their returns is 0.6. What is the 10-day 97.5% VaR for the portfolio? By how much does diversification reduce the VaR?

20.18. Consider a portfolio of options on a single asset. Suppose that the delta of the portfolio is 12, the value of the asset is \$10, and the daily volatility of the asset is 2%. Estimate the 1-day 95% VaR for the portfolio from the delta. Suppose next that the gamma of the portfolio is −2.6. Derive a quadratic relationship between the change in the portfolio value and the percentage change in the underlying asset price in one day. How would you use this in a Monte Carlo simulation?

20.19. A company has a long position in a 2-year bond and a 3-year bond, as well as a short position in a 5-year bond. Each bond has a principal of \$100 and pays a 5% coupon annually. Calculate the company's exposure to the 1-year, 2-year, 3-year, 4-year, and 5-year rates. Use the data in Tables 20.3 and 20.4 to calculate a 20-day 95% VaR on the assumption that rate changes are explained by (a) one factor, (b) two factors, and (c) three factors. Assume that the zero-coupon yield curve is flat at 5%.

20.20. A bank has written a call option on one stock and a put option on another stock. For the first option the stock price is 50, the strike price is 51, the volatility is 28% per annum, and the time to maturity is 9 months. For the second option the stock price is 20, the strike price is 19, the volatility is 25% per annum, and the time to maturity is 1 year. Neither stock pays a dividend, the risk-free rate is 6% per annum, and the correlation between stock price returns is 0.4. Calculate a 10-day 99% VaR:
(a) Using only deltas
(b) Using the partial simulation approach
(c) Using the full simulation approach

20.21. A common complaint of risk managers is that the model-building approach (either linear or quadratic) does not work well when delta is close to zero. Test what happens when delta is close to zero by using Sample Application E in the DerivaGem Application Builder software. (You can do this by experimenting with different option positions and adjusting the position in the underlying to give a delta of zero.) Explain the results you get.

APPENDIX

CASH-FLOW MAPPING

In this appendix we explain one procedure for mapping cash flows to standard maturity dates. We will illustrate the procedure by considering a simple example of a portfolio consisting of a long position in a single Treasury bond with a principal of $1 million maturing in 0.8 years. We suppose that the bond provides a coupon of 10% per annum payable semiannually. This means that the bond provides coupon payments of $50,000 in 0.3 years and 0.8 years. It also provides a principal payment of $1 million in 0.8 years. The Treasury bond can therefore be regarded as a position in a 0.3-year zero-coupon bond with a principal of $50,000 and a position in a 0.8-year zero-coupon bond with a principal of $1,050,000.

The position in the 0.3-year zero-coupon bond is mapped into an equivalent position in 3-month and 6-month zero-coupon bonds. The position in the 0.8-year zero-coupon bond is mapped into an equivalent position in 6-month and 1-year zero-coupon bonds. The result is that the position in the 0.8-year coupon-bearing bond is, for VaR purposes, regarded as a position in zero-coupon bonds having maturities of 3 months, 6 months, and 1 year.

The Mapping Procedure

Consider the $1,050,000 that will be received in 0.8 years. We suppose that zero rates, daily bond price volatilities, and correlations between bond returns are as shown in Table 20A.1.

The first stage is to interpolate between the 6-month rate of 6.0% and the 1-year rate of 7.0% to obtain a 0.8-year rate of 6.6%. (Annual compounding is assumed for all rates.) The present value of the $1,050,000 cash flow to be received in 0.8 years is

$$\frac{1{,}050{,}000}{1.066^{0.8}} = 997{,}662$$

We also interpolate between the 0.1% volatility for the 6-month bond and the 0.2% volatility for the 1-year bond to get a 0.16% volatility for the 0.8-year bond.

Table 20A.1 Data to illustrate mapping procedure.

Maturity:	*3-month*	*6-month*	*1-year*
Zero rate (% with annual compounding):	5.50	6.00	7.00
Bond price volatility (% per day):	0.06	0.10	0.20

Correlation between daily returns	*3-month bond*	*6-month bond*	*1-year bond*
3-month bond	1.0	0.9	0.6
6-month bond	0.9	1.0	0.7
1-year bond	0.6	0.7	1.0

Table 20A.2 The cash-flow mapping result.

	\$50,000 received in 0.3 years	*\$1,050,000 received in 0.8 years*	*Total*
Position in 3-month bond (\$):	37,397		37,397
Position in 6-month bond (\$):	11,793	319,589	331,382
Position in 1-year bond (\$):		678,074	678,074

Suppose we allocate α of the present value to the 6-month bond and $1-\alpha$ of the present value to the 1-year bond. Using equation (20.2) and matching variances, we obtain

$$0.0016^2 = 0.001^2\alpha^2 + 0.002^2(1-\alpha)^2 + 2 \times 0.7 \times 0.001 \times 0.002\alpha(1-\alpha)$$

This is a quadratic equation that can be solved in the usual way to give $\alpha = 0.320337$. This means that 32.0337% of the value should be allocated to a 6-month zero-coupon bond and 67.9663% of the value should be allocated to a 1-year zero coupon bond. The 0.8-year bond worth \$997,662 is therefore replaced by a 6-month bond worth

$$997{,}662 \times 0.320337 = \$319{,}589$$

and a 1-year bond worth

$$997{,}662 \times 0.679663 = \$678{,}074$$

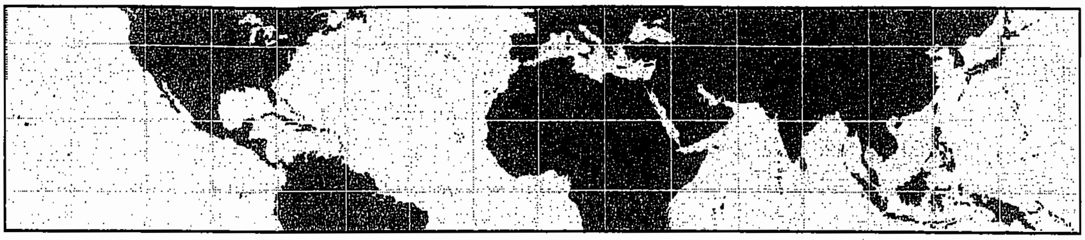
This cash-flow mapping scheme has the advantage that it preserves both the value and the variance of the cash flow. Also, it can be shown that the weights assigned to the two adjacent zero-coupon bonds are always positive.

For the \$50,000 cash flow received at time 0.3 years, we can carry out similar calculations (see Problem 20.10). It turns out that the present value of the cash flow is \$49,189. It can be mapped into a position worth \$37,397 in a 3-month bond and a position worth \$11,793 in a 6-month bond.

The results of the calculations are summarized in Table 20A.2. The 0.8-year coupon-bearing bond is mapped into a position worth \$37,397 in a 3-month bond, a position worth \$331,382 in a 6-month bond, and a position worth \$678,074 in a 1-year bond. Using the volatilities and correlations in Table 20A.1, equation (20.2) gives the variance of the change in the price of the 0.8-year bond with $n = 3$, $\alpha_1 = 37{,}397$, $\alpha_2 = 331{,}382$, $\alpha_3 = 678{,}074$, $\sigma_1 = 0.0006$, $\sigma_2 = 0.001$, $\sigma_3 = 0.002$, and $\rho_{12} = 0.9$, $\rho_{13} = 0.6$, $\rho_{23} = 0.7$. This variance is 2,628,518. The standard deviation of the change in the price of the bond is therefore $\sqrt{2{,}628{,}518} = 1{,}621.3$. Because we are assuming that the bond is the only instrument in the portfolio, the 10-day 99% VaR is

$$1621.3 \times \sqrt{10} \times 2.33 = 11{,}946$$

or about \$11,950.



CHAPTER 21

Estimating Volatilities and Correlations

In this chapter we explain how historical data can be used to produce estimates of the current and future levels of volatilities and correlations. The chapter is relevant both to the calculation of value at risk using the model-building approach and to the valuation of derivatives. When calculating value at risk, we are most interested in the current levels of volatilities and correlations because we are assessing possible changes in the value of a portfolio over a very short period of time. When valuing derivatives, forecasts of volatilities and correlations over the whole life of the derivative are usually required.

The chapter considers models with imposing names such as exponentially weighted moving average (EWMA), autoregressive conditional heteroscedasticity (ARCH), and generalized autoregressive conditional heteroscedasticity (GARCH). The distinctive feature of the models is that they recognize that volatilities and correlations are not constant. During some periods, a particular volatility or correlation may be relatively low, whereas during other periods it may be relatively high. The models attempt to keep track of the variations in the volatility or correlation through time.

21.1 ESTIMATING VOLATILITY

Define σ_n as the volatility of a market variable on day n, as estimated at the end of day $n-1$. The square of the volatility, σ_n^2, on day n is the *variance rate*. We described the standard approach to estimating σ_n from historical data in Section 13.4. Suppose that the value of the market variable at the end of day i is S_i. The variable u_i is defined as the continuously compounded return during day i (between the end of day $i-1$ and the end of day i):

$$u_i = \ln\frac{S_i}{S_{i-1}}$$

An unbiased estimate of the variance rate per day, σ_n^2, using the most recent m observations on the u_i is

$$\sigma_n^2 = \frac{1}{m-1}\sum_{i=1}^{m}(u_{n-i} - \bar{u})^2 \tag{21.1}$$

where $\bar{u}$ is the mean of the u_is:

$$\bar{u} = \frac{1}{m}\sum_{i=1}^{m} u_{n-i}$$

For the purposes of monitoring daily volatility, the formula in equation (21.1) is usually changed in a number of ways:

1. u_i is defined as the percentage change in the market variable between the end of day $i-1$ and the end of day i, so that:[1]

$$u_i = \frac{S_i - S_{i-1}}{S_{i-1}} \tag{21.2}$$

2. $\bar{u}$ is assumed to be zero.[2]
3. $m-1$ is replaced by m.[3]

These three changes make very little difference to the estimates that are calculated, but they allow us to simplify the formula for the variance rate to

$$\sigma_n^2 = \frac{1}{m}\sum_{i=1}^{m} u_{n-i}^2 \tag{21.3}$$

where u_i is given by equation (21.2).[4]

Weighting Schemes

Equation (21.3) gives equal weight to $u_{n-1}^2, u_{n-2}^2, \ldots, u_{n-m}^2$. Our objective is to estimate the current level of volatility, σ_n. It therefore makes sense to give more weight to recent data. A model that does this is

$$\sigma_n^2 = \sum_{i=1}^{m} \alpha_i u_{n-i}^2 \tag{21.4}$$

The variable α_i is the amount of weight given to the observation i days ago. The α's are positive. If we choose them so that $\alpha_i < \alpha_j$ when $i > j$, less weight is given to older observations. The weights must sum to unity, so that

$$\sum_{i=1}^{m} \alpha_i = 1$$

[1] This is consistent with the point made in Section 20.3 about the way that volatility is defined for the purposes of VaR calculations.

[2] As explained in Section 20.3, this assumption usually has very little effect on estimates of the variance because the expected change in a variable in one day is very small when compared with the standard deviation of changes.

[3] Replacing $m-1$ by m moves us from an unbiased estimate of the variance to a maximum likelihood estimate. Maximum likelihood estimates are discussed later in the chapter.

[4] Note that the u's in this chapter play the same role as the Δx's in Chapter 20. Both are daily percentage changes in market variables. In the case of the u's, the subscripts count observations made on different days on the same market variable. In the case of the Δx's, they count observations made on the same day on different market variables. The use of subscripts for σ is similarly different between the two chapters. In this chapter, the subscripts refer to days; in Chapter 20 they referred to market variables.

An extension of the idea in equation (21.4) is to assume that there is a long-run average variance rate and that this should be given some weight. This leads to the model that takes the form

$$\sigma_n^2 = \gamma V_L + \sum_{i=1}^{m} \alpha_i u_{n-i}^2 \tag{21.5}$$

where V_L is the long-run variance rate and γ is the weight assigned to V_L. Since the weights must sum to unity, it follows that

$$\gamma + \sum_{i=1}^{m} \alpha_i = 1$$

This is known as an ARCH(m) model. It was first suggested by Engle.[5] The estimate of the variance is based on a long-run average variance and m observations. The older an observation, the less weight it is given. Defining $\omega = \gamma V_L$, the model in equation (21.5) can be written

$$\sigma_n^2 = \omega + \sum_{i=1}^{m} \alpha_i u_{n-i}^2 \tag{21.6}$$

In the next two sections we discuss two important approaches to monitoring volatility using the ideas in equations (21.4) and (21.5).

21.2 THE EXPONENTIALLY WEIGHTED MOVING AVERAGE MODEL

The exponentially weighted moving average (EWMA) model is a particular case of the model in equation (21.4) where the weights α_i decrease exponentially as we move back through time. Specifically, $\alpha_{i+1} = \lambda\alpha_i$, where λ is a constant between 0 and 1.

It turns out that this weighting scheme leads to a particularly simple formula for updating volatility estimates. The formula is

$$\sigma_n^2 = \lambda\sigma_{n-1}^2 + (1-\lambda)u_{n-1}^2 \tag{21.7}$$

The estimate, σ_n, of the volatility of a variable for day n (made at the end of day $n-1$) is calculated from σ_{n-1} (the estimate that was made at the end of day $n-2$ of the volatility for day $n-1$) and u_{n-1} (the most recent daily percentage change in the variable).

To understand why equation (21.7) corresponds to weights that decrease exponentially, we substitute for σ_{n-1}^2 to get

$$\sigma_n^2 = \lambda[\lambda\sigma_{n-2}^2 + (1-\lambda)u_{n-2}^2] + (1-\lambda)u_{n-1}^2$$

or

$$\sigma_n^2 = (1-\lambda)(u_{n-1}^2 + \lambda u_{n-2}^2) + \lambda^2\sigma_{n-2}^2$$

Substituting in a similar way for σ_{n-2}^2 gives

$$\sigma_n^2 = (1-\lambda)(u_{n-1}^2 + \lambda u_{n-2}^2 + \lambda^2 u_{n-3}^2) + \lambda^3\sigma_{n-3}^2$$

---

[5] See R. Engle "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation," *Econometrica*, 50 (1982): 987–1008.

Continuing in this way gives

$$\sigma_n^2 = (1-\lambda)\sum_{i=1}^{m} \lambda^{i-1} u_{n-i}^2 + \lambda^m \sigma_{n-m}^2$$

For large m, the term $\lambda^m \sigma_{n-m}^2$ is sufficiently small to be ignored, so that equation (21.7) is the same as equation (21.4) with $\alpha_i = (1-\lambda)\lambda^{i-1}$. The weights for the u_i decline at rate λ as we move back through time. Each weight is λ times the previous weight.

Example 21.1

Suppose that λ is 0.90, the volatility estimated for a market variable for day $n-1$ is 1% per day, and during day $n-1$ the market variable increased by 2%. This means that $\sigma_{n-1}^2 = 0.01^2 = 0.0001$ and $u_{n-1}^2 = 0.02^2 = 0.0004$. Equation (21.7) gives

$$\sigma_n^2 = 0.9 \times 0.0001 + 0.1 \times 0.0004 = 0.00013$$

The estimate of the volatility, σ_n, for day n is therefore $\sqrt{0.00013}$, or 1.14%, per day. Note that the expected value of u_{n-1}^2 is σ_{n-1}^2, or 0.0001. In this example, the realized value of u_{n-1}^2 is greater than the expected value, and as a result our volatility estimate increases. If the realized value of u_{n-1}^2 had been less than its expected value, our estimate of the volatility would have decreased.

The EWMA approach has the attractive feature that relatively little data need to be stored. At any given time, only the current estimate of the variance rate and the most recent observation on the value of the market variable need be remembered. When a new observation on the market variable is obtained, a new daily percentage change is calculated and equation (21.7) is used to update the estimate of the variance rate. The old estimate of the variance rate and the old value of the market variable can then be discarded.

The EWMA approach is designed to track changes in the volatility. Suppose there is a big move in the market variable on day $n-1$, so that u_{n-1}^2 is large. From equation (21.7) this causes the estimate of the current volatility to move upward. The value of λ governs how responsive the estimate of the daily volatility is to the most recent daily percentage change. A low value of λ leads to a great deal of weight being given to the u_{n-1}^2 when σ_n is calculated. In this case, the estimates produced for the volatility on successive days are themselves highly volatile. A high value of λ (i.e., a value close to 1.0) produces estimates of the daily volatility that respond relatively slowly to new information provided by the daily percentage change.

The RiskMetrics database, which was originally created by J. P. Morgan and made publicly available in 1994, uses the EWMA model with $\lambda = 0.94$ for updating daily volatility estimates in its RiskMetrics database. The company found that, across a range of different market variables, this value of λ gives forecasts of the variance rate that come closest to the realized variance rate.[6] The realized variance rate on a particular day was calculated as an equally weighted average of the u_i^2 on the subsequent 25 days (see Problem 21.17).

[6] See J. P. Morgan, *RiskMetrics Monitor*, Fourth Quarter, 1995. We will explain an alternative (maximum likelihood) approach to estimating parameters later in the chapter.

21.3 THE GARCH(1,1) MODEL

We now move on to discuss what is known as the GARCH(1,1) model, proposed by Bollerslev in 1986.[7] The difference between the GARCH(1,1) model and the EWMA model is analogous to the difference between equation (21.4) and equation (21.5). In GARCH(1,1), σ_n^2 is calculated from a long-run average variance rate, V_L, as well as from σ_{n-1} and u_{n-1}. The equation for GARCH(1,1) is

$$\sigma_n^2 = \gamma V_L + \alpha u_{n-1}^2 + \beta\sigma_{n-1}^2 \tag{21.8}$$

where γ is the weight assigned to V_L, α is the weight assigned to u_{n-1}^2, and β is the weight assigned to σ_{n-1}^2. Since the weights must sum to unity, it follows that

$$\gamma + \alpha + \beta = 1$$

The EWMA model is a particular case of GARCH(1,1) where $\gamma = 0$, $\alpha = 1 - \lambda$, and $\beta = \lambda$.

The "(1,1)" in GARCH(1,1) indicates that σ_n^2 is based on the most recent observation of u^2 and the most recent estimate of the variance rate. The more general GARCH(p,q) model calculates σ_n^2 from the most recent p observations on u^2 and the most recent q estimates of the variance rate.[8] GARCH(1,1) is by far the most popular of the GARCH models.

Setting $\omega = \gamma V_L$, the GARCH(1,1) model can also be written

$$\sigma_n^2 = \omega + \alpha u_{n-1}^2 + \beta\sigma_{n-1}^2 \tag{21.9}$$

This is the form of the model that is usually used for the purposes of estimating the parameters. Once ω, α, and β have been estimated, we can calculate γ as $1 - \alpha - \beta$. The long-term variance V_L can then be calculated as ω/γ. For a stable GARCH(1,1) process we require $\alpha + \beta < 1$. Otherwise the weight applied to the long-term variance is negative.

Example 21.2

Suppose that a GARCH(1,1) model is estimated from daily data as

$$\sigma_n^2 = 0.000002 + 0.13u_{n-1}^2 + 0.86\sigma_{n-1}^2$$

This corresponds to $\alpha = 0.13$, $\beta = 0.86$, and $\omega = 0.000002$. Because $\gamma = 1 - \alpha - \beta$, it follows that $\gamma = 0.01$. Because $\omega = \gamma V_L$, it follows that $V_L = 0.0002$. In other words, the long-run average variance per day implied by the model is 0.0002. This corresponds to a volatility of $\sqrt{0.0002} = 0.014$, or 1.4%, per day.

[7] See T. Bollerslev, "Generalized Autoregressive Conditional Heteroscedasticity," *Journal of Econometrics*, 31 (1986): 307–27.

[8] Other GARCH models have been proposed that incorporate asymmetric news. These models are designed so that σ_n depends on the sign of u_{n-1}. Arguably, the models are more appropriate for equities than GARCH(1,1). As mentioned in Chapter 18, the volatility of an equity's price tends to be inversely related to the price so that a negative u_{n-1} should have a bigger effect on σ_n than the same positive u_{n-1}. For a discussion of models for handling asymmetric news, see D. Nelson, "Conditional Heteroscedasticity and Asset Returns: A New Approach," *Econometrica*, 59 (1990): 347–70; R. F. Engle and V. Ng, "Measuring and Testing the Impact of News on Volatility," *Journal of Finance*, 48 (1993): 1749–78.

Suppose that the estimate of the volatility on day $n-1$ is 1.6% per day, so that $\sigma_{n-1}^2 = 0.016^2 = 0.000256$, and that on day $n-1$ the market variable decreased by 1%, so that $u_{n-1}^2 = 0.01^2 = 0.0001$. Then

$$\sigma_n^2 = 0.000002 + 0.13 \times 0.0001 + 0.86 \times 0.000256 = 0.00023516$$

The new estimate of the volatility is therefore $\sqrt{0.00023516} = 0.0153$, or 1.53%, per day.

The Weights

Substituting for σ_{n-1}^2 in equation (21.9) gives

$$\sigma_n^2 = \omega + \alpha u_{n-1}^2 + \beta(\omega + \alpha u_{n-2}^2 + \beta\sigma_{n-2}^2)$$

or

$$\sigma_n^2 = \omega + \beta\omega + \alpha u_{n-1}^2 + \alpha\beta u_{n-2}^2 + \beta^2\sigma_{n-2}^2$$

Substituting for σ_{n-2}^2 gives

$$\sigma_n^2 = \omega + \beta\omega + \beta^2\omega + \alpha u_{n-1}^2 + \alpha\beta u_{n-2}^2 + \alpha\beta^2 u_{n-3}^2 + \beta^3\sigma_{n-3}^2$$

Continuing in this way, we see that the weight applied to u_{n-i}^2 is $\alpha\beta^{i-1}$. The weights decline exponentially at rate β. The parameter β can be interpreted as a "decay rate". It is similar to λ in the EWMA model. It defines the relative importance of the observations on the u's in determining the current variance rate. For example, if $\beta = 0.9$, then u_{n-2}^2 is only 90% as important as u_{n-1}^2; u_{n-3}^2 is 81% as important as u_{n-1}; and so on. The GARCH(1,1) model is similar to the EWMA model except that, in addition to assigning weights that decline exponentially to past u^2, it also assigns some weight to the long-run average volatility.

Mean Reversion

The GARCH (1,1) model recognizes that over time the variance tends to get pulled back to a long-run average level of V_L. The amount of weight assigned to V_L is $\gamma = 1 - \alpha - \beta$. The GARCH(1,1) is equivalent to a model where the variance V follows the stochastic process

$$dV = a(V_L - V)\,dt + \xi V\,dz$$

where time is measured in days, $a = 1 - \alpha - \beta$, and $\xi = \alpha\sqrt{2}$ (see Problem 21.14). This is a mean-reverting model. The variance has a drift that pulls it back to V_L at rate a. When $V > V_L$, the variance has a negative drift; when $V < V_L$, it has a positive drift. Superimposed on the drift is a volatility ξ. Chapter 26 discusses this type of model further.

21.4 CHOOSING BETWEEN THE MODELS

In practice, variance rates do tend to be mean reverting. The GARCH(1,1) model incorporates mean reversion, whereas the EWMA model does not. GARCH (1,1) is therefore theoretically more appealing than the EWMA model.

In the next section, we will discuss how best-fit parameters ω, α, and β in GARCH(1,1) can be estimated. When the parameter ω is zero, the GARCH(1,1) reduces to EWMA. In circumstances where the best-fit value of ω turns out to be negative, the GARCH(1,1) model is not stable and it makes sense to switch to the EWMA model.

21.5 MAXIMUM LIKELIHOOD METHODS

It is now appropriate to discuss how the parameters in the models we have been considering are estimated from historical data. The approach used is known as the *maximum likelihood method.* It involves choosing values for the parameters that maximize the chance (or likelihood) of the data occurring.

To illustrate the method, we start with a very simple example. Suppose that we sample 10 stocks at random on a certain day and find that the price of one of them declined on that day and the prices of the other nine either remained the same or increased. What is the best estimate of the probability of a price decline? The natural answer is 0.1. Let us see if this is what the maximum likelihood method gives.

Suppose that the probability of a price decline is p. The probability that one particular stock declines in price and the other nine do not is $p(1-p)^9$. Using the maximum likelihood approach, the best estimate of p is the one that maximizes $p(1-p)^9$. Differentiating this expression with respect to p and setting the result equal to zero, we find that $p = 0.1$ maximizes the expression. This shows that the maximum likelihood estimate of p is 0.1, as expected.

Estimating a Constant Variance

Our next example of maximum likelihood methods considers the problem of estimating the variance of a variable X from m observations on X when the underlying distribution is normal with zero mean. Assume that the observations are $u_1, u_2, \dots, u_m$. Denote the variance by v. The likelihood of u_i being observed is defined as the probability density function for X when $X = u_i$. This is

$$\frac{1}{\sqrt{2\pi v}}\exp\left(\frac{-u_i^2}{2v}\right)$$

The likelihood of m observations occurring in the order in which they are observed is

$$\prod_{i=1}^{m}\left[\frac{1}{\sqrt{2\pi v}}\exp\left(\frac{-u_i^2}{2v}\right)\right] \tag{21.10}$$

Using the maximum likelihood method, the best estimate of v is the value that maximizes this expression.

Maximizing an expression is equivalent to maximizing the logarithm of the expression. Taking logarithms of the expression in equation (21.10) and ignoring constant multiplicative factors, it can be seen that we wish to maximize

$$\sum_{i=1}^{m}\left[-\ln(v) - \frac{u_i^2}{v}\right] \tag{21.11}$$

or

$$-m\ln(v) - \sum_{i=1}^{m} \frac{u_i^2}{v}$$

Differentiating this expression with respect to v and setting the resulting equation to zero, we see that the maximum likelihood estimator of v is[9]

$$\frac{1}{m}\sum_{i=1}^{m} u_i^2$$

Estimating GARCH (1,1) Parameters

We now consider how the maximum likelihood method can be used to estimate the parameters when GARCH (1,1) or some other volatility updating scheme is used. Define $v_i = \sigma_i^2$ as the variance estimated for day i. Assume that the probability distribution of u_i conditional on the variance is normal. A similar analysis to the one just given shows the best parameters are the ones that maximize

$$\prod_{i=1}^{m}\left[\frac{1}{\sqrt{2\pi v_i}}\exp\left(\frac{-u_i^2}{2v_i}\right)\right]$$

Taking logarithms, we see that this is equivalent to maximizing

$$\sum_{i=1}^{m}\left[-\ln(v_i) - \frac{u_i^2}{v_i}\right] \tag{21.12}$$

This is the same as the expression in equation (21.11), except that v is replaced by v_i. It is necessary to search iteratively to find the parameters in the model that maximize the expression in equation (21.12).

The spreadsheet in Table 21.1 indicates how the calculations could be organized for the GARCH(1,1) model. The table analyzes data on the Japanese yen exchange rate between January 6, 1988, and August 15, 1997.[10] The numbers in the table are based on trial estimates of the three GARCH(1,1) parameters: ω, α, and β. The first column in the table records the date. The second column counts the days. The third column shows the exchange rate, S_i, at the end of day i. The fourth column shows the proportional change in the exchange rate between the end of day $i-1$ and the end of day i. This is $u_i = (S_i - S_{i-1})/S_{i-1}$. The fifth column shows the estimate of the variance rate, $v_i = \sigma_i^2$, for day i made at the end of day $i-1$. On day 3, we start things off by setting the variance equal to u_2^2. On subsequent days, equation (21.9) is used. The sixth column tabulates the likelihood measure, $-\ln(v_i) - u_i^2/v_i$. The values in the fifth and sixth columns are based on the current trial estimates of ω, α, and β. We are interested in choosing ω, α, and β to maximize the sum of the numbers in the sixth column. This involves an iterative search procedure.[11]

---

[9] This confirms the point made in footnote 3.

[10] The data can be downloaded from www.rotman.utoronto.ca/~hull/data

[11] As discussed later, a general purpose algorithm such as Solver in Microsoft's Excel can be used. Alternatively, a special purpose algorithm, such as Levenberg–Marquardt, can be used. See, e.g., W. H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling. *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, 1988.

Table 21.1 Estimation of parameters in GARCH(1,1) model.

Date	*Day i*	S_i	u_i	$v_i = \sigma_i^2$	$-\ln(v_i) - u_i^2/v_i$
06-Jan-88	1	0.007728			
07-Jan-88	2	0.007779	0.006599		
08-Jan-88	3	0.007746	−0.004242	0.00004355	9.6283
11-Jan-88	4	0.007816	0.009037	0.00004198	8.1329
12-Jan-88	5	0.007837	0.002687	0.00004455	9.8568
13-Jan-88	6	0.007924	0.011101	0.00004220	7.1529
⋮	⋮	⋮	⋮	⋮	⋮
13-Aug-97	2421	0.008643	0.003374	0.00007626	9.3321
14-Aug-97	2422	0.008493	−0.017309	0.00007092	5.3294
15-Aug-97	2423	0.008495	0.000144	0.00008417	9.3824
					22,063.5763

Trial estimates of GARCH parameters
$\omega = 0.00000176 \quad \alpha = 0.0626 \quad \beta = 0.8976$

In our example, the optimal values of the parameters turn out to be

$$\omega = 0.00000176, \quad \alpha = 0.0626, \quad \beta = 0.8976$$

and the maximum value of the function in equation (21.12) is 22,063.5763. The numbers shown in Table 21.1 were calculated on the final iteration of the search for the optimal ω, α, and β.

The long-term variance rate, V_L, in our example is

$$\frac{\omega}{1-\alpha-\beta} = \frac{0.00000176}{0.0398} = 0.00004422$$

The long-term volatility is $\sqrt{0.00004422}$, or 0.665%, per day.

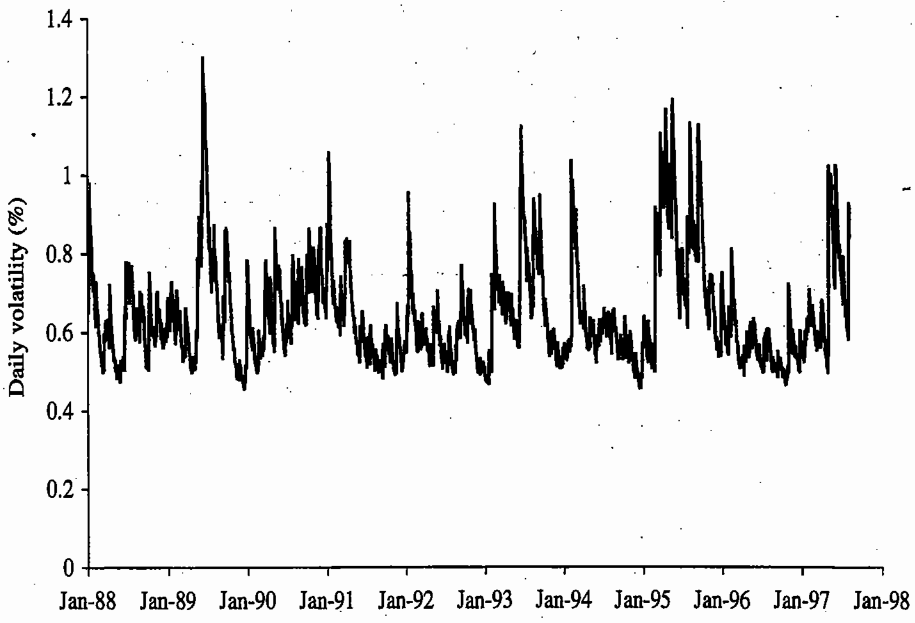
Figure 21.1 shows the way the GARCH (1,1) volatility for the Japanese yen changed over the 10-year period covered by the data. Most of the time, the volatility was between 0.4% and 0.8% per day, but volatilities over 1% were experienced during some periods.

An alternative and more robust approach to estimating parameters in GARCH(1,1) is known as *variance targeting*.[12] This involves setting the long-run average variance rate, V_L, equal to the sample variance calculated from the data (or to some other value that is believed to be reasonable). The value of ω then equals $V_L(1-\alpha-\beta)$ and only two parameters have to be estimated. For the data in Table 21.1, the sample variance is 0.00004341, which gives a daily volatility of 0.659%. Setting V_L equal to the sample variance, the values of α and β that maximize the objective function in equation (21.12) are 0.0607 and 0.8990, respectively. The value of the objective function is 22,063.5274, only marginally below the value of 22,063.5763 obtained using the earlier procedure.

When the EWMA model is used, the estimation procedure is relatively simple. We set $\omega = 0$, $\alpha = 1-\lambda$, and $\beta = \lambda$, and only one parameter has to be estimated. In the data in Table 21.1, the value of λ that maximizes the objective function in equation (21.12) is 0.9686 and the value of the objective function is 21,995.8377.

[12] See R. Engle and J. Mezrich, "GARCH for Groups," *Risk*, August 1996: 36–40.

Figure 21.1 Daily volatility of the yen/USD exchange rate, 1988–1997.



Both GARCH (1,1) and the EWMA method can be implemented by using the Solver routine in Excel to search for the values of the parameters that maximize the likelihood function. The routine works well provided that the spreadsheet is structured so that the parameters being searched for have roughly equal values. For example, in GARCH (1,1) we could let cells A1, A2, and A3 contain $\omega \times 10^5$, α, and 0.1β. We could then set B1 = A1/100000, B2 = A2, and B3 = 10 * A3. We would use B1, B2, and B3 to calculate the likelihood function. We would ask Solver to calculate the values of A1, A2, and A3 that maximize the likelihood function.

How Good Is the Model?

The assumption underlying a GARCH model is that volatility changes with the passage of time. During some periods volatility is relatively high; during other periods it is relatively low. To put this another way, when u_i^2 is high, there is a tendency for u_{i+1}^2, $u_{i+2}^2, \ldots$ to be high; when u_i^2 is low, there is a tendency for u_{i+1}^2, $u_{i+2}^2, \ldots$ to be low. We can test how true this is by examining the autocorrelation structure of the u_i^2.

Let us assume the u_i^2 do exhibit autocorrelation. If a GARCH model is working well, it should remove the autocorrelation. We can test whether it has done so by considering the autocorrelation structure for the variables u_i^2/σ_i^2. If these show very little autocorrelation, our model for σ_i has succeeded in explaining autocorrelations in the u_i^2.

Table 21.2 shows results for the yen/dollar exchange rate data referred to above. The first column shows the lags considered when the autocorrelation is calculated. The second shows autocorrelations for u_i^2; the third shows autocorrelations for u_i^2/σ_i^2.[13] The table shows that the autocorrelations are positive for u_i^2 for all lags between 1 and 15. In the case of u_i^2/σ_i^2, some of the autocorrelations are positive and some are negative. They are all much smaller in magnitude than the autocorrelations for u_i^2.

[13] For a series x_i, the autocorrelation with a lag of k is the coefficient of correlation between x_i and x_{i+k}.

Table 21.2 Autocorrelations before and after the use of a GARCH model.

Time lag	*Autocorrelation for u_i^2*	*Autocorrelation for u_i^2/σ_i^2*
1	0.072	0.004
2	0.041	−0.005
3	0.057	0.008
4	0.107	0.003
5	0.075	0.016
6	0.066	0.008
7	0.019	−0.033
8	0.085	0.012
9	0.054	0.010
10	0.030	−0.023
11	0.038	−0.004
12	0.038	−0.021
13	0.057	−0.001
14	0.040	0.002
15	0.007	−0.028

The GARCH model appears to have done a good job in explaining the data. For a more scientific test, we can use what is known as the Ljung–Box statistic.[14] If a certain series has m observations the Ljung–Box statistic is

$$m\sum_{k=1}^{K} w_k \eta_k^2$$

where η_k is the autocorrelation for a lag of k, K is the number of lags considered, and

$$w_k = \frac{m+2}{m-k}$$

For $K = 15$, zero autocorrelation can be rejected with 95% confidence when the Ljung–Box statistic is greater than 25.

From Table 21.2, the Ljung–Box Statistic for the u_i^2 series is about 123. This is strong evidence of autocorrelation. For the u_i^2/σ_i^2 series, the Ljung–Box statistic is 8.2, suggesting that the autocorrelation has been largely removed by the GARCH model.

21.6 USING GARCH(1,1) TO FORECAST FUTURE VOLATILITY

The variance rate estimated at the end of day $n-1$ for day n, when GARCH(1,1) is used, is

$$\sigma_n^2 = (1-\alpha-\beta)V_L + \alpha u_{n-1}^2 + \beta\sigma_{n-1}^2$$

[14] See G. M. Ljung and G. E. P. Box, "On a Measure of Lack of Fit in Time Series Models," *Biometrica*, 65 (1978): 297–303.

so that

$$\sigma_n^2 - V_L = \alpha(u_{n-1}^2 - V_L) + \beta(\sigma_{n-1}^2 - V_L)$$

On day $n + t$ in the future,

$$\sigma_{n+t}^2 - V_L = \alpha(u_{n+t-1}^2 - V_L) + \beta(\sigma_{n+t-1}^2 - V_L)$$

The expected value of u_{n+t-1}^2 is σ_{n+t-1}^2. Hence,

$$E[\sigma_{n+t}^2 - V_L] = (\alpha + \beta)E[\sigma_{n+t-1}^2 - V_L]$$

where E denotes expected value. Using this equation repeatedly yields

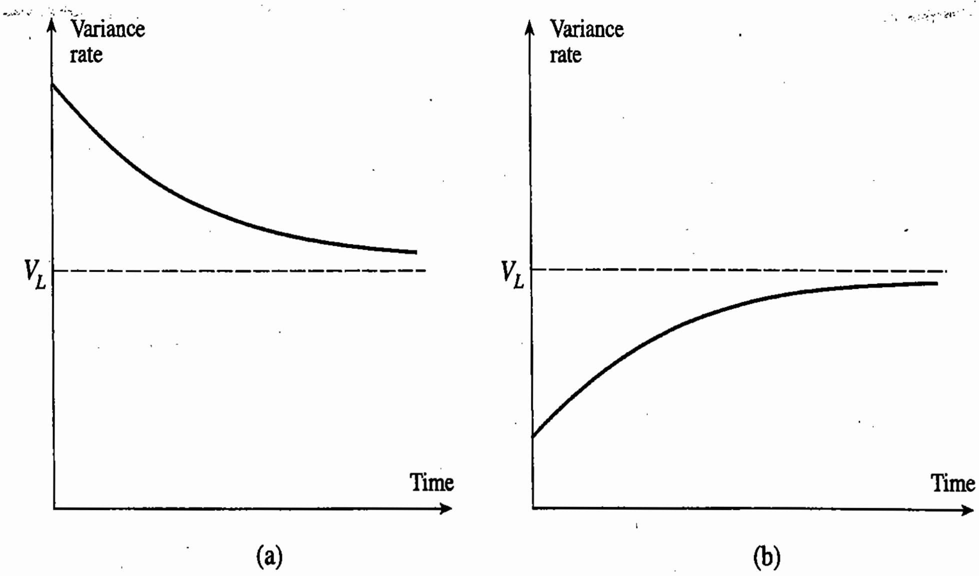
$$E[\sigma_{n+t}^2 - V_L] = (\alpha + \beta)^t(\sigma_n^2 - V_L)$$

or

$$E[\sigma_{n+t}^2] = V_L + (\alpha + \beta)^t(\sigma_n^2 - V_L) \tag{21.13}$$

This equation forecasts the volatility on day $n + t$ using the information available at the end of day $n - 1$. In the EWMA model, $\alpha + \beta = 1$ and equation (21.13) shows that the expected future variance rate equals the current variance rate. When $\alpha + \beta < 1$, the final term in the equation becomes progressively smaller as t increases. Figure 21.2 shows the expected path followed by the variance rate for situations where the current variance rate is different from V_L. As mentioned earlier, the variance rate exhibits mean reversion with a reversion level of V_L and a reversion rate of $1 - \alpha - \beta$. Our forecast of the future variance rate tends towards V_L as we look further and further ahead. This analysis emphasizes the point that we must have $\alpha + \beta < 1$ for a stable GARCH(1,1) process.

Figure 21.2 Expected path for the variance rate when (a) current variance rate is above long-term variance rate and (b) current variance rate is below long-term variance rate.



When $\alpha + \beta > 1$, the weight given to the long-term average variance is negative and the process is "mean fleeing" rather than "mean reverting".

In the yen-dollar exchange rate example considered earlier $\alpha + \beta = 0.9602$ and $V_L = 0.00004422$. Suppose that the estimate of the current variance rate per day is 0.00006. (This corresponds to a volatility of 0.77% per day.) In 10 days the expected variance rate is

$$0.00004422 + 0.9602^{10}(0.00006 - 0.00004422) = 0.00005473$$

The expected volatility per day is 0.74%, still well above the long-term volatility of 0.665% per day. However, the expected variance rate in 100 days is

$$0.00004422 + 0.9602^{100}(0.00006 - 0.00004422) = 0.00004449$$

and the expected volatility per day is 0.667%, very close to the long-term volatility.

Volatility Term Structures

Suppose it is day n. Define:

$$V(t) = E(\sigma_{n+t}^2)$$

and

$$a = \ln\frac{1}{\alpha + \beta}$$

so that equation (21.13) becomes

$$V(t) = V_L + e^{-at}[V(0) - V_L]$$

Here, $V(t)$ is an estimate of the instantaneous variance rate in t days. The average variance rate per day between today and time T is given by

$$\frac{1}{T}\int_0^T V(t)\,dt = V_L + \frac{1 - e^{-aT}}{aT}[V(0) - V_L]$$

The larger T is, the closer this is to V_L. Define $\sigma(T)$ as the volatility per annum that should be used to price a T-day option under GARCH(1,1). Assuming 252 days per year, $\sigma(T)^2$ is 252 times the average variance rate per day, so that

$$\sigma(T)^2 = 252\left(V_L + \frac{1 - e^{-aT}}{aT}[V(0) - V_L]\right) \qquad \textbf{(21.14)}$$

As discussed in Chapter 18, the market prices of different options on the same asset are often used to calculate a *volatility term structure*. This is the relationship between the implied volatilities of the options and their maturities. Equation (21.14) can be used to estimate a volatility term structure based on the GARCH(1,1) model. The estimated volatility term structure is not usually the same as the actual volatility term structure. However, as we will show, it is often used to predict the way that the actual volatility term structure will respond to volatility changes.

When the current volatility is above the long-term volatility, the GARCH(1,1) model estimates a downward-sloping volatility term structure. When the current volatility is below the long-term volatility, it estimates an upward-sloping volatility term structure. In the case of the yen/dollar exchange rate, $a = \ln(1/0.9602) = 0.0406$

Table 21.3 Yen/dollar volatility term structure predicted from GARCH(1,1).

Option life (days)	10	30	50	100	500
Option volatility (% per annum)	12.00	11.59	11.33	11.00	10.65

and $V_L = 0.00004422$. Suppose that the current variance rate per day, $V(0)$, is estimated as 0.00006 per day. It follows from equation (21.14) that

$$\sigma(T)^2 = 252\left(0.00004422 + \frac{1 - e^{-0.0406T}}{0.0406T}(0.00006 - 0.000044220)\right)$$

where T is measured in days. Table 21.3 shows the volatility per year for different values of T.

Impact of Volatility Changes

Equation (21.14) can be written

$$\sigma(T)^2 = 252\left[V_L + \frac{1 - e^{-aT}}{aT}\left(\frac{\sigma(0)^2}{252} - V_L\right)\right]$$

When $\sigma(0)$ changes by $\Delta\sigma(0)$, $\sigma(T)$ changes by

$$\frac{1 - e^{-aT}}{aT}\frac{\sigma(0)}{\sigma(T)}\Delta\sigma(0) \tag{21.15}$$

Table 21.4 shows the effect of a volatility change on options of varying maturities for the yen/dollar exchange rate example. We assume as before that $V(0) = 0.00006$, so that $\sigma(0) = 12.30\%$. The table considers a 100-basis-point change in the instantaneous volatility from 12.30% per year to 13.30% per year. This means that $\Delta\sigma(0) = 0.01$, or 1%.

Many financial institutions use analyses such as this when determining the exposure of their books to volatility changes. Rather than consider an across-the-board increase of 1% in implied volatilities when calculating vega, they relate the size of the volatility increase that is considered to the maturity of the option. Based on Table 21.4, a 0.84% volatility increase would be considered for a 10-day option, a 0.61% increase for a 30-day option, a 0.46% increase for a 50-day option, and so on.

Table 21.4 Impact of 1% change in the instantaneous volatility predicted from GARCH(1,1).

Option life (days)	10	30	50	100	500
Increase in volatility (%)	0.84	0.61	0.46	0.27	0.06

21.7 CORRELATIONS

The discussion so far has centered on the estimation and forecasting of volatility. As explained in Chapter 20, correlations also play a key role in the calculation of VaR. In this section, we show how correlation estimates can be updated in a similar way to volatility estimates.

The correlation between two variables X and Y can be defined as

$$\frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y}$$

where σ_X and σ_Y are the standard deviations of X and Y and $\text{cov}(X, Y)$ is the covariance between X and Y. The covariance between X and Y is defined as

$$E[(X - \mu_X)(Y - \mu_Y)]$$

where μ_X and μ_Y are the means of X and Y, and E denotes the expected value. Although it is easier to develop intuition about the meaning of a correlation than it is for a covariance, it is covariances that are the fundamental variables of our analysis.[15]

Define x_i and y_i as the percentage changes in X and Y between the end of day $i-1$ and the end of day i:

$$x_i = \frac{X_i - X_{i-1}}{X_{i-1}}, \qquad y_i = \frac{Y_i - Y_{i-1}}{Y_{i-1}}$$

where X_i and Y_i are the values of X and Y at the end of day i. We also define the following:

$\sigma_{x,n}$: Daily volatility of variable X, estimated for day n

$\sigma_{y,n}$: Daily volatility of variable Y, estimated for day n

cov_n: Estimate of covariance between daily changes in X and Y, calculated on day n.

The estimate of the correlation between X and Y on day n is

$$\frac{\text{cov}_n}{\sigma_{x,n}\,\sigma_{y,n}}$$

Using an equal-weighting scheme and assuming that the means of x_i and y_i are zero, equation (21.3) shows that the variance rates of X and Y can be estimated from the most recent m observations as

$$\sigma_{x,n}^2 = \frac{1}{m}\sum_{i=1}^{m} x_{n-i}^2, \qquad \sigma_{y,n}^2 = \frac{1}{m}\sum_{i=1}^{m} y_{n-i}^2$$

A similar estimate for the covariance between X and Y is

$$\text{cov}_n = \frac{1}{m}\sum_{i=1}^{m} x_{n-i}\, y_{n-i} \tag{21.16}$$

[15] An analogy here is that variance rates were the fundamental variables for the EWMA and GARCH schemes in first part of this chapter, even though volatilities are easier to understand.

One alternative for updating covariances is an EWMA model similar to equation (21.7). The formula for updating the covariance estimate is then

$$\text{cov}_n = \lambda\,\text{cov}_{n-1} + (1-\lambda)x_{n-1}\,y_{n-1}$$

A similar analysis to that presented for the EWMA volatility model shows that the weights given to observations on the $x_i\,y_i$ decline as we move back through time. The lower the value of λ, the greater the weight that is given to recent observations.

Example 21.3

Suppose that $\lambda = 0.95$ and that the estimate of the correlation between two variables X and Y on day $n-1$ is 0.6. Suppose further that the estimate of the volatilities for the X and Y on day $n-1$ are 1% and 2%, respectively. From the relationship between correlation and covariance, the estimate of the covariance between the X and Y on day $n-1$ is

$$0.6 \times 0.01 \times 0.02 = 0.00012$$

Suppose that the percentage changes in X and Y on day $n-1$ are 0.5% and 2.5%, respectively. The variance and covariance for day n would be updated as follows:

$$\sigma_{x,n}^2 = 0.95 \times 0.01^2 + 0.05 \times 0.005^2 = 0.00009625$$

$$\sigma_{y,n}^2 = 0.95 \times 0.02^2 + 0.05 \times 0.025^2 = 0.00041125$$

$$\text{cov}_n = 0.95 \times 0.00012 + 0.05 \times 0.005 \times 0.025 = 0.00012025$$

The new volatility of X is $\sqrt{0.00009625} = 0.981\%$ and the new volatility of Y is $\sqrt{0.00041125} = 2.028\%$. The new coefficient of correlation between X and Y is

$$\frac{0.00012025}{0.00981 \times 0.02028} = 0.6044$$

GARCH models can also be used for updating covariance estimates and forecasting the future level of covariances. For example, the GARCH(1,1) model for updating a covariance is

$$\text{cov}_n = \omega + \alpha x_{n-1}\,y_{n-1} + \beta\,\text{cov}_{n-1}$$

and the long-term average covariance is $\omega/(1-\alpha-\beta)$. Formulas similar to those in equations (21.13) and (21.14) can be developed for forecasting future covariances and calculating the average covariance during the life of an option.[16]

Consistency Condition for Covariances

Once all the variances and covariances have been calculated, a variance–covariance matrix can be constructed. When $i \neq j$, the (i, j) element of this matrix shows the covariance between variable i and variable j. When $i = j$, it shows the variance of variable i.

---

[16] The ideas in this chapter can be extended to multivariate GARCH models, where an entire variance–covariance matrix is updated in a consistent way. For a discussion of alternative approaches, see R. Engle and J. Mezrich, "GARCH for Groups," *Risk*, August 1996: 36–40.

Not all variance–covariance matrices are internally consistent. The condition for an $N \times N$ variance–covariance matrix Ω to be internally consistent is

$$w^{\mathsf{T}}\Omega w \geqslant 0 \tag{21.17}$$

for all $N \times 1$ vectors w, where w^{T} is the transpose of w. A matrix that satisfies this property is known as *positive-semidefinite.*

To understand why the condition in equation (21.17) must hold, suppose that w^{T} is $[w_1, w_2, \dots, w_n]$. The expression $w^{\mathsf{T}}\Omega w$ is the variance of $w_1x_1 + w_2x_2 + \cdots + w_nx_n$, where x_i is the value of variable i. As such, it cannot be negative.

To ensure that a positive-semidefinite matrix is produced, variances and covariances should be calculated consistently. For example, if variances are calculated by giving equal weight to the last m data items, the same should be done for covariances. If variances are updated using an EWMA model with $\lambda = 0.94$, the same should be done for covariances.

An example of a variance–covariance matrix that is not internally consistent is

$$\begin{bmatrix} 1 & 0 & 0.9 \\ 0 & 1 & 0.9 \\ 0.9 & 0.9 & 1 \end{bmatrix}$$

The variance of each variable is 1.0, and so the covariances are also coefficients of correlation. The first variable is highly correlated with the third variable and the second variable is highly correlated with the third variable. However, there is no correlation at all between the first and second variables. This seems strange. When w is set equal to $(1, 1, -1)$, the condition in equation (21.17) is not satisfied, proving that the matrix is not positive-semidefinite.[17]

SUMMARY

Most popular option pricing models, such as Black–Scholes, assume that the volatility of the underlying asset is constant. This assumption is far from perfect. In practice, the volatility of an asset, like the asset's price, is a stochastic variable. Unlike the asset price, it is not directly observable. This chapter has discussed procedures for attempting to keep track of the current level of volatility.

We define u_i as the percentage change in a market variable between the end of day $i-1$ and the end of day i. The variance rate of the market variable (that is, the square of its volatility) is calculated as a weighted average of the u_i^2. The key feature of the procedures that have been discussed here is that they do not give equal weight to the observations on the u_i^2. The more recent an observation, the greater the weight assigned to it. In the EWMA and the GARCH(1,1) models, the weights assigned to observations decrease exponentially as the observations become older. The GARCH(1,1) model differs from the EWMA model in that some weight is also assigned to the long-run average variance rate. Both the EWMA and GARCH(1,1) models have structures that enable forecasts of the future level of variance rate to be produced relatively easily.

---

[17] It can be shown that the condition for a 3×3 matrix of correlations to be internally consistent is

$$\rho_{12}^2 + \rho_{13}^2 + \rho_{23}^2 - 2\rho_{12}\rho_{13}\rho_{23} \leqslant 1$$

where ρ_{ij} is the coefficient of correlation between variables i and j.

Maximum likelihood methods are usually used to estimate parameters from historical data in GARCH(1,1) and similar models. These methods involve using an iterative procedure to determine the parameter values that maximize the chance or likelihood that the historical data will occur. Once its parameters have been determined, a model can be judged by how well it removes autocorrelation from the u_i^2.

For every model that is developed to track variances, there is a corresponding model that can be developed to track covariances. The procedures described here can therefore be used to update the complete variance–covariance matrix used in value at risk calculations.

FURTHER READING

Bollerslev, T. "Generalized Autoregressive Conditional Heteroscedasticity," *Journal of Econometrics*, 31 (1986): 307–27.

Cumby, R., S. Figlewski, and J. Hasbrook. "Forecasting Volatilities and Correlations with EGARCH Models," *Journal of Derivatives*, 1, 2 (Winter 1993): 51–63.

Engle, R. F. "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of UK Inflation," *Econometrica* 50 (1982): 987–1008.

Engle R. F., and J. Mezrich. "Grappling with GARCH," *Risk*, September 1995: 112–117.

Engle, R. F., and J. Mezrich, "GARCH for Groups," *Risk*, August 1996: 36–40.

Engle, R. F., and V. Ng, "Measuring and Testing the Impact of News on Volatility," *Journal of Finance*, 48 (1993): 1749–78.

Nelson, D. "Conditional Heteroscedasticity and Asset Returns: A New Approach," *Econometrica*, 59 (1990): 347–70.

Noh, J., R. F. Engle, and A. Kane. "Forecasting Volatility and Option Prices of the S&P 500 Index," *Journal of Derivatives*, 2 (1994): 17–30.

Questions and Problems (Answers in Solutions Manual)

21.1. Explain the exponentially weighted moving average (EWMA) model for estimating volatility from historical data.

21.2. What is the difference between the exponentially weighted moving average model and the GARCH(1,1) model for updating volatilities?

21.3. The most recent estimate of the daily volatility of an asset is 1.5% and the price of the asset at the close of trading yesterday was \$30.00. The parameter λ in the EWMA model is 0.94. Suppose that the price of the asset at the close of trading today is \$30.50. How will this cause the volatility to be updated by the EWMA model?

21.4. A company uses an EWMA model for forecasting volatility. It decides to change the parameter λ from 0.95 to 0.85. Explain the likely impact on the forecasts.

21.5. The volatility of a certain market variable is 30% per annum. Calculate a 99% confidence interval for the size of the percentage daily change in the variable.

21.6. A company uses the GARCH(1,1) model for updating volatility. The three parameters are ω, α, and β. Describe the impact of making a small increase in each of the parameters while keeping the others fixed.

21.7. The most recent estimate of the daily volatility of the US dollar/sterling exchange rate is 0.6% and the exchange rate at 4 p.m. yesterday was 1.5000. The parameter λ in the EWMA model is 0.9. Suppose that the exchange rate at 4 p.m. today proves to be 1.4950. How would the estimate of the daily volatility be updated?

21.8. Assume that S&P 500 at close of trading yesterday was 1,040 and the daily volatility of the index was estimated as 1% per day at that time. The parameters in a GARCH(1, 1) model are $\omega = 0.000002$, $\alpha = 0.06$, and $\beta = 0.92$. If the level of the index at close of trading today is 1,060, what is the new volatility estimate?

21.9. Suppose that the daily volatilities of asset A and asset B, calculated at the close of trading yesterday, are 1.6% and 2.5%, respectively. The prices of the assets at close of trading yesterday were \$20 and \$40 and the estimate of the coefficient of correlation between the returns on the two assets was 0.25. The parameter λ used in the EWMA model is 0.95.
(a) Calculate the current estimate of the covariance between the assets.
(b) On the assumption that the prices of the assets at close of trading today are \$20.5 and \$40.5, update the correlation estimate.

21.10. The parameters of a GARCH(1, 1) model are estimated as $\omega = 0.000004$, $\alpha = 0.05$, and $\beta = 0.92$. What is the long-run average volatility and what is the equation describing the way that the variance rate reverts to its long-run average? If the current volatility is 20% per year, what is the expected volatility in 20 days?

21.11. Suppose that the current daily volatilities of asset X and asset Y are 1.0% and 1.2%, respectively. The prices of the assets at close of trading yesterday were \$30 and \$50 and the estimate of the coefficient of correlation between the returns on the two assets made at this time was 0.50. Correlations and volatilities are updated using a GARCH(1, 1) model. The estimates of the model's parameters are $\alpha = 0.04$ and $\beta = 0.94$. For the correlation $\omega = 0.000001$, and for the volatilities $\omega = 0.000003$. If the prices of the two assets at close of trading today are \$31 and \$51, how is the correlation estimate updated?

21.12. Suppose that the daily volatility of the FTSE 100 stock index (measured in pounds sterling) is 1.8% and the daily volatility of the dollar/sterling exchange rate is 0.9%. Suppose further that the correlation between the FTSE 100 and the dollar/sterling exchange rate is 0.4. What is the volatility of the FTSE 100 when it is translated to US dollars? Assume that the dollar/sterling exchange rate is expressed as the number of US dollars per pound sterling. (*Hint*: When $Z = XY$, the percentage daily change in Z is approximately equal to the percentage daily change in X plus the percentage daily change in Y.)

21.13. Suppose that in Problem 21.12 the correlation between the S&P 500 Index (measured in dollars) and the FTSE 100 Index (measured in sterling) is 0.7, the correlation between the S&P 500 Index (measured in dollars) and the dollar/sterling exchange rate is 0.3, and the daily volatility of the S&P 500 index is 1.6%. What is the correlation between the S&P 500 index (measured in dollars) and the FTSE 100 index when it is translated to dollars? (*Hint*: For three variables X, Y, and Z, the covariance between $X + Y$ and Z equals the covariance between X and Z plus the covariance between Y and Z.)

21.14. Show that the GARCH (1,1) model $\sigma_n^2 = \omega + \alpha u_{n-1}^2 + \beta\sigma_{n-1}^2$ in equation (21.9) is equivalent to the stochastic volatility model $dV = a(V_L - V)\,dt + \xi V\,dz$, where time is measured in days, V is the square of the volatility of the asset price, and

$$a = 1 - \alpha - \beta, \qquad V_L = \frac{\omega}{1 - \alpha - \beta}, \qquad \xi = \alpha\sqrt{2}$$

What is the stochastic volatility model when time is measured in years? (*Hint*: The variable u_{n-1} is the return on the asset price in time Δt. It can be assumed to be normally distributed with mean zero and standard deviation σ_{n-1}. It follows that the mean of u_{n-1}^2 and u_{n-1}^4 are σ_{n-1}^2 and $3\sigma_{n-1}^4$, respectively.)

Assignment Questions

21.15. Suppose that the price of gold at close of trading yesterday was \$600 and its volatility was estimated as 1.3% per day. The price at the close of trading today is \$596. Update the volatility estimate using
(a) The EWMA model with $\lambda = 0.94$
(b) The GARCH(1,1) model with $\omega = 0.000002$, $\alpha = 0.04$, and $\beta = 0.94$

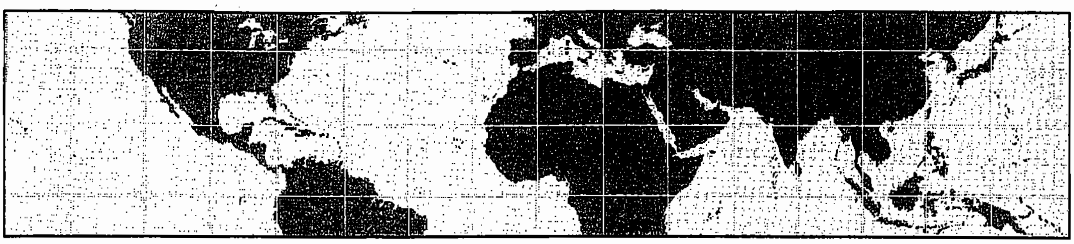
21.16. Suppose that in Problem 21.15 the price of silver at the close of trading yesterday was \$16, its volatility was estimated as 1.5% per day, and its correlation with gold was estimated as 0.8. The price of silver at the close of trading today is unchanged at \$16. Update the volatility of silver and the correlation between silver and gold using the two models in Problem 21.15. In practice, is the ω parameter likely to be the same for gold and silver?

21.17. An Excel spreadsheet containing over 900 days of daily data on a number of different exchange rates and stock indices can be downloaded from the author's website:

`http://www.rotman.utoronto.ca/~hull`

Choose one exchange rate and one stock index. Estimate the value of λ in the EWMA model that minimizes the value of $\sum_i (v_i - \beta_i)^2$, where v_i is the variance forecast made at the end of day $i-1$ and β_i is the variance calculated from data between day i and day $i+25$. Use the Solver tool in Excel. Set the variance forecast at the end of the first day equal to the square of the return on that day to start the EWMA calculations.

21.18. Suppose that the parameters in a GARCH (1,1) model are $\alpha = 0.03$, $\beta = 0.95$, and $\omega = 0.000002$.
(a) What is the long-run average volatility?
(b) If the current volatility is 1.5% per day, what is your estimate of the volatility in 20, 40, and 60 days?
(c) What volatility should be used to price 20-, 40-, and 60-day options?
(d) Suppose that there is an event that increases the current volatility by 0.5% to 2% per day. Estimate the effect on the volatility in 20, 40, and 60 days.
(e) Estimate by how much the event increases the volatilities used to price 20-, 40-, and 60-day options?



CHAPTER 22

Credit Risk

The value-at-risk measure we covered in Chapter 20 and the Greek letters we studied in Chapter 17 are aimed at quantifying market risk. In this chapter we consider another important risk for financial institutions: credit risk. Most financial institutions devote considerable resources to the measurement and management of credit risk. Regulators have for many years required banks to keep capital to reflect the credit risks they are bearing. This capital is in addition to the capital, described in Business Snapshot 20.1, that is required for market risk.

Credit risk arises from the possibility that borrowers and counterparties in derivatives transactions may default. This chapter discusses a number of different approaches to estimating the probability that a company will default and explains the key difference between risk-neutral and real-world probabilities of default. It examines the nature of the credit risk in over-the-counter derivatives transactions and discusses the clauses derivatives dealers write into their contracts to reduce credit risk. It also covers default correlation, Gaussian copula models, and the estimation of credit value at risk.

Chapter 23 will discuss credit derivatives and show how ideas introduced in this chapter can be used to value these instruments.

22.1 CREDIT RATINGS

Rating agencies, such as Moody's, S&P, and Fitch, are in the business of providing ratings describing the creditworthiness of corporate bonds. The best rating assigned by Moody's is Aaa. Bonds with this rating are considered to have almost no chance of defaulting. The next best rating is Aa. Following that comes A, Baa, Ba, B, Caa, Ca, and C. Only bonds with ratings of Baa or above are considered to be *investment grade*. The S&P ratings corresponding to Moody's Aaa, Aa, A, Baa, Ba, B, Caa, Ca, and C are AAA, AA, A, BBB, BB, B, CCC, CC, and C, respectively. Fitch's rating categories are similar to those of S&P. To create finer rating measures, Moody's divides its Aa rating category into Aa1, Aa2, and Aa3, its A category into A1, A2, and A3, and so on. Similarly, S&P divides its AA rating category into AA+, AA, and AA−, its A rating category into A+, A, and A−, and so on. Moody's Aaa category and S&P's AAA category are not subdivided, nor usually are the two lowest rating categories.

22.2 HISTORICAL DEFAULT PROBABILITIES

Table 22.1 is typical of the data produced by rating agencies. It shows the default experience during a 20-year period of bonds that had a particular rating at the beginning of the period. For example, a bond with a credit rating of Baa has a 0.181% chance of defaulting by the end of the first year, a 0.506% chance of defaulting by the end of the second year, and so on. The probability of a bond defaulting during a particular year can be calculated from the table. For example, the probability that a bond initially rated Baa will default during the second year is $0.506 - 0.181 = 0.325\%$.

Table 22.1 shows that, for investment grade bonds, the probability of default in a year tends to be an increasing function of time (e.g., the probabilities of an A-rated bond defaulting during years 0–5, 5–10, 10–15, and 15–20 are 0.472%, 0.815%, 1.077%, and 1.874%, respectively). This is because the bond issuer is initially considered to be creditworthy, and the more time that elapses, the greater the possibility that its financial health will decline. For bonds with a poor credit rating, the probability of default is often a decreasing function of time (e.g., the probabilities that a B-rated bond will default during years 0–5, 5–10, 10–15, and 15–20 are 26.794%, 16.549%, 8.832%, and 2.246%, respectively). The reason here is that, for a bond with a poor credit rating, the next year or two may be critical. The longer the issuer survives, the greater the chance that its financial health improves.

Default Intensities

From Table 22.1 we can calculate the probability of a bond rated Caa or below defaulting during the third year as $39.717 - 30.494 = 9.223\%$. We will refer to this as the *unconditional default probability*. It is the probability of default during the third year as seen at time 0. The probability that the bond will survive until the end of year 2 is $100 - 30.494 = 69.506\%$. The probability that it will default during the third year conditional on no earlier default is therefore $0.09223/0.69506$, or 13.27%. Conditional default probabilities are referred to as *default intensities* or *hazard rates*.

The 13.27% we have just calculated is for a 1-year time period. Suppose instead that we consider a short time period of length Δt. The default intensity $\lambda(t)$ at time t is then defined so that $\lambda(t)\,\Delta t$ is the probability of default between time t and $t + \Delta t$ conditional on no earlier default. If $V(t)$ is the cumulative probability of the company surviving to time t (i.e., no default by time t), the conditional probability of default between time t

Table 22.1 Average cumulative default rates (%), 1970–2006. *Source*: Moody's.

Term (years):	*1*	*2*	*3*	*4*	*5*	*7*	*10*	*15*	*20*
Aaa	0.000	0.000	0.000	0.026	0.099	0.251	0.521	0.992	1.191
Aa	0.008	0.019	0.042	0.106	0.177	0.343	0.522	1.111	1.929
A	0.021	0.095	0.220	0.344	0.472	0.759	1.287	2.364	4.238
Baa	0.181	0.506	0.930	1.434	1.938	2.959	4.637	8.244	11.362
Ba	1.205	3.219	5.568	7.958	10.215	14.005	19.118	28.380	35.093
B	5.236	11.296	17.043	22.054	26.794	34.771	43.343	52.175	54.421
Caa–C	19.476	30.494	39.717	46.904	52.622	59.938	69.178	70.870	70.870

and $t + \Delta t$ is $[V(t) - V(t + \Delta t)]/V(t)$. Since this equals $\lambda(t)\,\Delta t$, it follows that

$$V(t + \Delta t) - V(t) = -\lambda(t)V(t)\,\Delta t$$

Taking limits

$$\frac{dV(t)}{dt} = -\lambda(t)V(t)$$

from which

$$V(t) = e^{-\int_0^t \lambda(\tau)d\tau}$$

Defining $Q(t)$ as the probability of default by time t, so that $Q(t) = 1 - V(t)$, gives

$$Q(t) = 1 - e^{-\int_0^t \lambda(\tau)d\tau}$$

or

$$Q(t) = 1 - e^{-\bar{\lambda}(t)t} \tag{22.1}$$

where $\bar{\lambda}(t)$ is the average default intensity (hazard rate) between time 0 and time t.

22.3 RECOVERY RATES

When a company goes bankrupt, those that are owed money by the company file claims against the assets of the company.[1] Sometimes there is a reorganization in which these creditors agree to a partial payment of their claims. In other cases the assets are sold by the liquidator and the proceeds are used to meet the claims as far as possible. Some claims typically have priority over other claims and are met more fully.

The recovery rate for a bond is normally defined as the bond's market value immediately after a default, as a percent of its face value. Table 22.2 provides historical data on average recovery rates for different categories of bonds in the United States. It shows that senior secured debt holders had an average recovery rate of 54.44 cents per dollar of face value while junior subordinated debt holders had an average recovery rate of only 24.47 cents per dollar of face value.

Recovery rates are significantly negatively correlated with default rates. Moody's looked at average recovery rates and average default rates each year between 1982

Table 22.2 Recovery rates on corporate bonds as a percentage of face value, 1982–2003. *Source*: Moody's.

Class	*Average recovery rate (%)*
Senior secured	54.44
Senior unsecured	38.39
Senior subordinated	32.85
Subordinated	31.61
Junior subordinated	24.47

[1] In the United States, the claim made by a bondholder is the bond's face value plus accrued interest.

and 2006. It found that the following relationship provides a good fit to the data:[2]

$$\text{Recovery rate} = 59.1 - 8.356 \times \text{Default rate}$$

where the recovery rate is the average recovery rate on senior unsecured bonds in a year measured as a percentage and the default rate is the corporate default rate in the year measured as a percentage.

This relationship means that a bad year for the default rate is usually doubly bad because it is accompanied by a low recovery rate. For example, when the average default rate in a year is only 0.1%, the expected recovery rate is relatively high at 58.3%. When the default rate is relatively high at 3%, the expected recovery rate is only 34.0%.

22.4 ESTIMATING DEFAULT PROBABILITIES FROM BOND PRICES

The probability of default for a company can be estimated from the prices of bonds it has issued. The usual assumption is that the only reason a corporate bond sells for less than a similar risk-free bond is the possibility of default.[3]

Consider first an approximate calculation. Suppose that a bond yields 200 basis points more than a similar risk-free bond and that the expected recovery rate in the event of a default is 40%. The holder of a corporate bond must be expecting to lose 200 basis points (or 2% per year) from defaults. Given the recovery rate of 40%, this leads to an estimate of the probability of a default per year conditional on no earlier default of 0.02/(1 − 0.4), or 3.33%. In general,

$$\bar{\lambda} = \frac{s}{1-R} \tag{22.2}$$

where $\bar{\lambda}$ is the average default intensity (hazard rate) per year, s is the spread of the corporate bond yield over the risk-free rate, and R is the expected recovery rate.

A More Exact Calculation

For a more exact calculation, suppose that the corporate bond we have been considering lasts for 5 years, provides a coupon 6% per annum (paid semiannually) and that the yield on the corporate bond is 7% per annum (with continuous compounding). The yield on a similar risk-free bond is 5% (with continuous compounding). The yields imply that the price of the corporate bond is 95.34 and the price of the risk-free bond is 104.09. The expected loss from default over the 5-year life of the bond is therefore 104.09 − 95.34, or \$8.75. Suppose that the probability of default per year (assumed in this simple example to be the same each year) is Q. Table 22.3 calculates the expected loss from default in terms of Q on the assumption that defaults can happen at times 0.5,

[2] See D. T. Hamilton, P. Varma, S. Ou, and R. Cantor, "Default and Recovery Rates of Corporate Bond Issuers," Moody's Investor's Services, January 2004. The R^2 of the regression is 0.6. The correlation is also identified and discussed in E. I. Altman, B. Brady, A. Resti, and A. Sironi, "The Link between Default and Recovery Rates: Implications for Credit Risk Models and Procyclicality," Working Paper, New York University, 2003.

[3] This assumption is not perfect. In practice the price of a corporate bond is affected by its liquidity. The lower the liquidity, the lower the price.

Table 22.3 Calculation of loss from default on a bond in terms of the default probabilities per year, Q. Notional principal = \$100.

Time (years)	*Default probability*	*Recovery amount (\$)*	*Risk-free value (\$)*	*Loss given default (\$)*	*Discount factor*	*PV of expected loss (\$)*
0.5	Q	40	106.73	66.73	0.9753	$65.08Q$
1.5	Q	40	105.97	65.97	0.9277	$61.20Q$
2.5	Q	40	105.17	65.17	0.8825	$57.52Q$
3.5	Q	40	104.34	64.34	0.8395	$54.01Q$
4.5	Q	40	103.46	63.46	0.7985	$50.67Q$
Total						$288.48Q$

1.5, 2.5, 3.5, and 4.5 years (immediately before coupon payment dates). Risk-free rates for all maturities are assumed to be 5% (with continuous compounding).

To illustrate the calculations, consider the 3.5-year row in Table 22.3. The expected value of the corporate bond at time 3.5 years (calculated using forward interest rates and assuming no possibility of default) is

$$3 + 3e^{-0.05\times0.5} + 3e^{-0.05\times1.0} + 103e^{-0.05\times1.5} = 104.34$$

Given the definition of recovery rates in the previous section, the amount recovered if there is a default is 40, so that the loss given default is 104.34 − 40, or \$64.34. The present value of this loss is 54.01. The expected loss is therefore $54.01Q$.

The total expected loss is $288.48Q$. Setting this equal to 8.75, we obtain a value for Q of 8.75/288.48, or 3.03%. The calculations we have given assume that the default probability is the same in each year and that defaults take place at just one time during the year. We can extend the calculations to assume that defaults can take place more frequently. Also, instead of assuming a constant unconditional probability of default we can assume a constant default intensity (hazard rate) or assume a particular pattern for the variation of default probabilities with time. With several bonds we can estimate several parameters describing the term structure of default probabilities. Suppose, for example, we have bonds maturing in 3, 5, 7, and 10 years. We could use the first bond to estimate a default probability per year for the first 3 years, the second bond to estimate default probability per year for years 4 and 5, the third bond to estimate a default probability for years 6 and 7, and the fourth bond to estimate a default probability for years 8, 9, and 10 (see Problems 22.15 and 22.29). This approach is analogous to the bootstrap procedure in Section 4.5 for calculating a zero-coupon yield curve.

The Risk-Free Rate

A key issue when bond prices are used to estimate default probabilities is the meaning of the terms "risk-free rate" and "risk-free bond". In equation (22.2), the spread s is the excess of the corporate bond yield over the yield on a similar risk-free bond. In Table 22.3, the risk-free value of the bond must be calculated using the risk-free discount rate. The benchmark risk-free rate that is usually used in quoting corporate bond yields is the yield on similar Treasury bonds. (For example, a bond trader might quote the yield on a particular corporate bond as being a spread of 250 basis points over Treasuries.)

As discussed in Section 4.1, traders usually use LIBOR/swap rates as proxies for risk-free rates when valuing derivatives. Traders also often use LIBOR/swap rates as risk-free rates when calculating default probabilities. For example, when they determine default probabilities from bond prices, the spread s in equation (22.2) is the spread of the bond yield over the LIBOR/swap rate. Also, the risk-free discount rates used in the calculations in Table 22.3 are LIBOR/swap zero rates.

Credit default swaps (which will be discussed in the next chapter) can be used to imply the risk-free rate assumed by traders. The implied rate appears to be approximately equal to the LIBOR/swap rate minus 10 basis points on average.[4] This estimate is plausible. As explained in Section 7.5, the credit risk in a swap rate is the credit risk from making a series of 6-month loans to AA-rated counterparties and 10 basis points is a reasonable default risk premium for a AA-rated 6-month instrument.

Asset Swaps

In practice, traders often use asset swap spreads as a way of extracting default probabilities from bond prices. This is because asset swap spreads provide a direct estimate of the spread of bond yields over the LIBOR/swap curve.

To explain how asset swaps work, consider the situation where an asset swap spread for a particular bond is quoted as 150 basis points. There are three possible situations:

1. The bond sells for its par value of 100. The swap then involves one side (company A) paying the coupon on the bond and the other side (company B) paying LIBOR plus 150 basis points. Note that it is the promised coupons that are exchanged. The exchanges take place regardless of whether the bond defaults.
2. The bond sells below its par value, say, for 95. The swap is then structured so that, in addition to the coupons, company A pays \$5 per \$100 of notional principal at the outset. Company B pays LIBOR plus 150 basis points.
3. The underlying bond sells above par, say, for 108. The swap is then structured so that, in addition to LIBOR plus 150 basis points, company B makes a payment of \$8 per \$100 of principal at the outset. Company A pays the coupons.

The effect of all this is that the present value of the asset swap spread is the amount by which the price of the corporate bond is exceeded by the price of a similar risk-free bond where the risk-free rate is assumed to be given by the LIBOR/swap curve (see Problem 22.24).

Consider again the example in Table 22.3 where the LIBOR/swap zero curve is flat at 5%. Suppose that instead of knowing the bond's price we know that the asset swap spread is 150 basis points. This means that the amount by which the value of the risk-free bond exceeds the value of the corporate bond is the present value of 150 basis points per year for 5 years. Assuming semiannual payments, this is \$6.55 per \$100 of principal. The total loss in Table 22.3 would in this case be set equal to \$6.55. This means that the default probability per year, Q, would be 6.55/288.48, or 2.27%.

[4] See J. Hull, M. Predescu, and A. White, "The Relationship between Credit Default Swap Spreads, Bond Yields, and Credit Rating Announcements," *Journal of Banking and Finance*, 28 (November 2004): 2789-2811.

22.5 COMPARISON OF DEFAULT PROBABILITY ESTIMATES

The default probabilities estimated from historical data are much less than those derived from bond prices. Table 22.4 illustrates this.[5] It shows, for companies that start with a particular rating, the average annual default intensity over 7 years calculated from (a) historical data and (b) bond prices.

The calculation of default intensities from historical data are based on equation (22.1) with $t = 7$ and Table 22.1. From equation (22.1),

$$\bar{\lambda}(7) = -\tfrac{1}{7}\ln[1 - Q(7)]$$

where $\bar{\lambda}(t)$ is the average default intensity (or hazard rate) by time t and $Q(t)$ is the cumulative probability of default by time t. The values of $Q(7)$ are taken directly from Table 22.1. Consider, for example, an A-rated company. The value of $Q(7)$ is 0.00759. The average 7-year default intensity is therefore

$$\bar{\lambda}(7) = -\tfrac{1}{7}\ln(0.99241) = 0.0011$$

or 0.11%.

The calculations of average default intensities from bond prices are based on equation (22.2) and bond yields published by Merrill Lynch. The results shown are averages between December 1996 and October 2007. The recovery rate is assumed to be 40% and, for the reasons discussed in the previous section, the risk-free interest rate is assumed to be the 7-year swap rate minus 10 basis points. For example, for A-rated bonds, the average Merrill Lynch yield was 5.993%. The average swap rate was 5.398%, so that the average risk-free rate was 5.289%. This gives the average 7-year default probability as

$$\frac{0.05993 - 0.05298}{1 - 0.4} = 0.0116$$

or 1.16%.

Table 22.4 shows that the ratio of the default probability backed out from bond prices to the default probability calculated from historical data is very high for investment grade companies and tends to decline as a company's credit rating declines. The difference between the two default probabilities tends to increase as the credit rating declines.

Table 22.4 Seven-year average default intensities (% per annum).

Rating	*Historical default intensity*	*Default intensity from bonds*	*Ratio*	*Difference*
Aaa	0.04	0.60	16.7	0.56
Aa	0.05	0.74	14.6	0.68
A	0.11	1.16	10.5	1.04
Baa	0.43	2.13	5.0	1.71
Ba	2.16	4.67	2.2	2.54
B	6.10	7.97	1.3	1.98
Caa and lower	13.07	18.16	1.4	5.50

[5] The results in Tables 22.4 and 22.5 are updates of the results in J. Hull, M. Predescu, and A. White, "Bond Prices, Default Probabilities, and Risk Premiums," *Journal of Credit Risk*, 1, 2 (Spring 2005): 53–60.

Table 22.5 Expected excess return on bonds (basis points).

Rating	*Bond yield spread over Treasuries*	*Spread of risk-free rate over Treasuries*	*Spread for historical defaults*	*Excess return*
Aaa	78	42	2	34
Aa	87	42	3	42
A	112	42	7	63
Baa	170	42	26	102
Ba	323	42	129	151
B	521	42	366	112
Caa	1132	42	784	305

Table 22.5 provides another way of looking at these results. It shows the excess return over the risk-free rate (still assumed to be the 7-year swap rate minus 10 basis points) earned by investors in bonds with different credit rating. Consider again an A-rated bond. The average spread over Treasuries is 112 basis points. Of this, 42 basis points are accounted for by the average spread between 7-year Treasuries and our proxy for the risk-free rate. A spread of 7 basis points is necessary to cover expected defaults. (This equals the real-world probability of default from Table 22.4 multiplied by 0.6 to allow for recoveries.) This leaves an excess return (after expected defaults have been taken into account) of 63 basis points.

Tables 22.4 and 22.5 show that a large percentage difference between default probability estimates translates into a small (but significant) excess return on the bond. For Aaa-rated bonds the ratio of the two default probabilities is 16.8, but the expected excess return is only 38 basis points. The excess return tends to increase as credit quality declines.[6]

The excess return in Table 22.5 does not remain constant through time. Credit spreads, and therefore excess returns, were high in 2001, 2002, and the first half of 2003. After that they were fairly low until the credit crunch in July 2007 when they started to increase rapidly. (See Business Snapshot 23.3 for a discussion of the year 2007 credit crunch.)

Real-World vs. Risk-Neutral Probabilities

The default probabilities implied from bond yields are risk-neutral probabilities of default. To explain why this is so, consider the calculations of default probabilities in Table 22.3. The calculations assume that expected default losses can be discounted at the risk-free rate. The risk-neutral valuation principle shows that this is a valid procedure providing the expected losses are calculated in a risk-neutral world. This means that the default probability Q in Table 22.3 must be a risk-neutral probability.

By contrast, the default probabilities implied from historical data are real-world default probabilities (sometimes also called *physical probabilities*). The expected excess return in Table 22.5 arises directly from the difference between real-world and risk-

[6] The results for B-rated bonds in Tables 22.4 and 22.5 run counter to the overall pattern.

neutral default probabilities. If there were no expected excess return, then the real-world and risk-neutral default probabilities would be the same, and vice versa.

Why do we see such big differences between real-world and risk-neutral default probabilities? As we have just argued, this is the same as asking why corporate bond traders earn more than the risk-free rate on average.

One reason for the results is that corporate bonds are relatively illiquid and the returns on bonds are higher than they would otherwise be to compensate for this. But this is a small part of what is going on. It explains perhaps 25 basis points of the excess return in Table 22.5. Another possible reason for the results is that the subjective default probabilities of bond traders are much higher than the those given in Table 22.1. Bond traders may be allowing for depression scenarios much worse than anything seen during the 1970 to 2003 period. However, it is difficult to see how this can explain a large part of the excess return that is observed.[7]

By far the most important reason for the results in Tables 22.4 and 22.5 is that bonds do not default independently of each other. There are periods of time when default rates are very low and periods of time when they are very high. Evidence for this can be obtained by looking at the default rates in different years. Moody's statistics show that between 1970 and 2006 the default rate per year ranged from a low of 0.09% in 1979 to a high of 3.81% in 2001. The year-to-year variation in default rates gives rise to systematic risk (i.e., risk that cannot be diversified away) and bond traders earn an excess expected return for bearing the risk. The variation in default rates from year to year may be because of overall economic conditions and it may be because a default by one company has a ripple effect resulting in defaults by other companies. (The latter is referred to by researchers as *credit contagion.*)

In addition to the systematic risk we have just talked about there is nonsystematic (or idiosyncratic) risk associated with each bond. If we were talking about stocks, we would argue that investors can diversify the nonsystematic risk by choosing a portfolio of, say, 30 stocks. They should not therefore demand a risk premium for bearing nonsystematic risk. For bonds, the arguments are not so clear-cut. Bond returns are highly skewed with limited upside. (For example, on an individual bond, there might be a 99.75% chance of a 7% return in a year, and a 0.25% chance of a −60% return in the year, the first outcome corresponding to no default and the second to default.) This type of risk is difficult to "diversify away".[8] It would require tens of thousands of different bonds. In practice, many bond portfolios are far from fully diversified. As a result, bond traders may earn an extra return for bearing nonsystematic risk as well as for bearing the systematic risk mentioned in the previous paragraph.

At this stage it is natural to ask whether we should use real-world or risk-neutral default probabilities in the analysis of credit risk. The answer depends on the purpose of the analysis. When valuing credit derivatives or estimating the impact of default risk on the pricing of instruments, risk-neutral default probabilities should be used. This is because the analysis calculates the present value of expected future cash flows and

[7] In addition to producing Table 22.1, which is based on the 1970 to 2006 period, Moody's produces a similar table based on the 1920 to 2006 period. When this table is used, historical default intensities for investment grade bonds in Table 22.4 rise somewhat. The Aaa default intensity increases from 4 to 6 basis points; the Aa increases from 5 to 22 basis points; the A increases from 11 to 29 basis points; the Baa increases from 43 to 73 basis points. However, the noninvestment grade historical default intensities decline.

[8] See J. D. Amato and E. M. Remolona, "The credit spread puzzle," *BIS Quarterly Review*, 5 (December 2003): 51–63.

almost invariably (implicitly or explicitly) involves using risk-neutral valuation. When carrying out scenario analyses to calculate potential future losses from defaults, real-world default probabilities should be used.

22.6 USING EQUITY PRICES TO ESTIMATE DEFAULT PROBABILITIES

When we use a table such as Table 22.1 to estimate a company's real-world probability of default, we are relying on the company's credit rating. Unfortunately, credit ratings are revised relatively infrequently. This has led some analysts to argue that equity prices can provide more up-to-date information for estimating default probabilities.

In 1974, Merton proposed a model where a company's equity is an option on the assets of the company.[9] Suppose, for simplicity, that a firm has one zero-coupon bond outstanding and that the bond matures at time T. Define:

V_0: Value of company's assets today
V_T: Value of company's assets at time T
E_0: Value of company's equity today
E_T: Value of company's equity at time T
D: Debt repayment due at time T
σ_V: Volatility of assets (assumed constant)
σ_E: Instantaneous volatility of equity.

If $V_T < D$, it is (at least in theory) rational for the company to default on the debt at time T. The value of the equity is then zero. If $V_T > D$, the company should make the debt repayment at time T and the value of the equity at this time is $V_T - D$. Merton's model, therefore, gives the value of the firm's equity at time T as

$$E_T = \max(V_T - D, 0)$$

This shows that the equity is a call option on the value of the assets with a strike price equal to the repayment required on the debt. The Black–Scholes formula gives the value of the equity today as

$$E_0 = V_0 N(d_1) - De^{-rT} N(d_2) \tag{22.3}$$

where

$$d_1 = \frac{\ln V_0/D + (r + \sigma_V^2/2)T}{\sigma_V \sqrt{T}} \quad \text{and} \quad d_2 = d_1 - \sigma_V \sqrt{T}$$

The value of the debt today is $V_0 - E_0$.

The risk-neutral probability that the company will default on the debt is $N(-d_2)$. To calculate this, we require V_0 and σ_V. Neither of these are directly observable. However, if the company is publicly traded, we can observe E_0. This means that equation (22.3) provides one condition that must be satisfied by V_0 and σ_V. We can also estimate σ_E from historical data or options. From Itô's lemma,

$$\sigma_E E_0 = \frac{\partial E}{\partial V} \sigma_V V_0$$

[9] See R. Merton "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, 29 (1974): 449–70.

or

$$\sigma_E E_0 = N(d_1)\sigma_V V_0 \tag{22.4}$$

This provides another equation that must be satisfied by V_0 and σ_V. Equations (22.3) and (22.4) provide a pair of simultaneous equations that can be solved for V_0 and σ_V.[10]

Example 22.1

The value of a company's equity is \$3 million and the volatility of the equity is 80%. The debt that will have to be paid in 1 year is \$10 million. The risk-free rate is 5% per annum. In this case $E_0 = 3$, $\sigma_E = 0.80$, $r = 0.05$, $T = 1$, and $D = 10$. Solving equations (22.3) and (22.4) yields $V_0 = 12.40$ and $\sigma_V = 0.2123$. The parameter d_2 is 1.1408, so that the probability of default is $N(-d_2) = 0.127$, or 12.7%. The market value of the debt is $V_0 - E_0$, or 9.40. The present value of the promised payment on the debt is $10e^{-0.05\times 1} = 9.51$. The expected loss on the debt is therefore $(9.51 - 9.40)/9.51$, or about 1.2% of its no-default value. Comparing this with the probability of default gives the expected recovery in the event of a default as $(12.7 - 1.2)/12.7$, or about 91% of the debt's no-default value.

The basic Merton model we have just presented has been extended in a number of ways. For example, one version of the model assumes that a default occurs whenever the value of the assets falls below a barrier level.

How well do the default probabilities produced by Merton's model and its extensions correspond to actual default experience? The answer is that Merton's model and its extensions produce a good ranking of default probabilities (risk-neutral or real-world). This means that a monotonic transformation can be used to convert the probability of default output from Merton's model into a good estimate of either the real-world or risk-neutral default probability.[11]

22.7 CREDIT RISK IN DERIVATIVES TRANSACTIONS

The credit exposure on a derivatives transaction is more complicated than that on a loan. This is because the claim that will be made in the event of a default is more uncertain. Consider a financial institution that has one derivatives contract outstanding with a counterparty. Three possible situations can be distinguished:

1. Contract is always a liability to the financial institution
2. Contract is always an asset to the financial institution
3. Contract can become either an asset or a liability to the financial institution

An example of a derivatives contract in the first category is a short option position; an example in the second category is a long option position; an example in the third category is a forward contract.

[10] To solve two nonlinear equations of the form $F(x, y) = 0$ and $G(x, y) = 0$, the Solver routine in Excel can be asked to find the values of x and y that minimize $[F(x, y)]^2 + [G(x, y)]^2$.

[11] Moody's KMV provides a service that transforms a default probability produced by Merton's model into a real-world default probability (which it refers to as an EDF, short for expected default frequency). CreditGrades use Merton's model to estimate credit spreads, which are closely linked to risk-neutral default probabilities.

Derivatives in the first category have no credit risk to the financial institution. If the counterparty goes bankrupt, there will be no loss. The derivative is one of the counterparty's assets. It is likely to be retained, closed out, or sold to a third party. The result is no loss (or gain) to the financial institution.

Derivatives in the second category always have credit risk to the financial institution. If the counterparty goes bankrupt, a loss is likely to be experienced. The derivative is one of the counterparty's liabilities. The financial institution has to make a claim against the assets of the counterparty and may receive some percentage of the value of the derivative. (Typically, a claim arising from a derivatives transaction is unsecured and junior.)

Derivatives in the third category may or may not have credit risk. If the counterparty defaults when the value of the derivative is positive to the financial institution, a claim will be made against the assets of the counterparty and a loss is likely to be experienced. If the counterparty defaults when the value is negative to the financial institution, no loss is made because the derivative is retained, closed out, or sold to a third party.[12]

Adjusting Derivatives' Valuations for Counterparty Default Risk

How should a financial institution (or end-user of derivatives) adjust the value of a derivative to allow for counterparty credit risk? Consider a derivative that lasts until time T and has a value of f_0 today assuming no defaults. Let us suppose that defaults can take place at times $t_1, t_2, \ldots, t_n$, where $t_n = T$, and that the value of the derivative to the financial institution (assuming no defaults) at time t_i is f_i. Define the risk-neutral probability of default at time t_i as q_i and the expected recovery rate as R.[13]

The exposure at time t_i is the financial institution's potential loss. This is $\max(f_i, 0)$. Assume that the expected recovery in the event of a default is R times the exposure. Assume also that the recovery rate and the probability of default are independent of the value of the derivative. The risk-neutral expected loss from default at time t_i is

$$q_i(1 - R)\hat{E}[\max(f_i, 0)]$$

where $\hat{E}$ denotes expected value in a risk-neutral world. Taking present values leads to the cost of defaults being

$$\sum_{i=1}^{n} u_i v_i \tag{22.5}$$

where u_i equals $q_i(1 - R)$ and v_i is the value today of an instrument that pays off the exposure on the derivative under consideration at time t_i.

Consider again the three categories of derivatives mentioned earlier. The first category (where the derivative is always a liability to the financial institution) is easy to deal with. The value of f_i is always negative and so the total expected loss from defaults given by equation (22.5) is always zero. The financial institution needs to make no adjustments for the cost of defaults. (Of course, the counterparty may want to take account of the possibility of the financial institution defaulting in its own pricing.)

[12] Note that a company usually defaults because of a deterioration in its overall financial health, not because of the value of any one transaction.

[13] The probability of default could be calculated from bond prices in the way described in Section 22.4.

For the second category (where the derivative is always an asset to the financial institution), f_i is always positive. This means than the expression $\max(f_i, 0)$ always equals f_i. Suppose that the only payoff from the derivative is at time T, the end of its life. In this case, f_0 must be the present value of f_i, so that $v_i = f_0$ for all i. The expression in equation (22.5) for the present value of the cost of defaults becomes

$$f_0 \sum_{i=1}^{n} q_i(1-R)$$

If f_0^* is the actual value of the derivative (after allowing for possible defaults), it follows that

$$f_0^* = f_0 - f_0 \sum_{i=1}^{n} q_i(1-R) = f_0\left[1 - \sum_{i=1}^{n} q_i(1-R)\right] \tag{22.6}$$

One particular instrument that falls into the second category we are considering is an unsecured zero-coupon bond that promises \$1 at time T and is issued by the counterparty in the derivatives transaction. Define B_0 as the value of the bond assuming no possibility of default and B_0^* as the actual value of the bond. If we make the simplifying assumption that the recovery on the bond as a percent of its no-default value is the same as that on the derivative, then

$$B_0^* = B_0\left[1 - \sum_{i=1}^{n} q_i(1-R)\right] \tag{22.7}$$

From equations (22.6) and (22.7),

$$\frac{f_0^*}{f_0} = \frac{B_0^*}{B_0} \tag{22.8}$$

If y is the yield on a risk-free zero-coupon bond maturing at time T and y^* is the yield on a zero-coupon bond issued by the counterparty that matures at time T, then $B_0 = e^{-yT}$ and $B_0^* = e^{-y^*T}$, so that equation (22.8) gives

$$f_0^* = f_0 e^{-(y^*-y)T} \tag{22.9}$$

This shows that any derivative promising a payoff at time T can be valued by increasing the discount rate that is applied to the expected payoff in a risk-neutral world from the risk-free rate y to the risky rate y^*.

Example 22.2

Consider a 2-year over-the-counter option sold by company X with a value, assuming no possibility of default, of \$3. Suppose that 2-year zero-coupon bonds issued by the company X have a yield that is 1.5% greater than a similar risk-free zero-coupon bond. The value of the option is

$$3e^{-0.015\times 2} = 2.91$$

or \$2.91.

For the third category of derivatives, the sign of f_i is uncertain. The variable v_i is a call option on f_i with a strike price of zero. One way of calculating v_i is to simulate the underlying market variables over the life of the derivative. Sometimes approximate analytic calculations are possible (see, e.g., Problems 22.17 and 22.18).

The analyses we have presented assume that the probability of default is independent of the value of the derivative. This is likely to be a reasonable approximation in circumstances when the derivative is a small part of the portfolio of the counterparty or when the counterparty is using the derivative for hedging purposes. When a counterparty wants to enter into a large derivatives transaction for speculative purposes a financial institution should be wary. When the transaction has a large negative value for the counterparty (and a large positive value for the financial institution), the chance of counterparty declaring bankruptcy may be much higher than when the situation is the other way round.

Traders working for a financial institution use the term *right-way risk* to describe the situation where a counterparty is most likely to default when the financial institution has zero, or very little, exposure. They use the term *wrong-way risk* to describe the situation where the counterparty is most likely to default when the financial institution has a big exposure.

22.8 CREDIT RISK MITIGATION

In many instances the analysis we just have presented overstates the credit risk in a derivatives transaction. This is because there are a number of clauses that derivatives dealers include in their contracts to mitigate credit risk.

Netting

A clause that has become standard in over-the-counter derivatives contracts is known as *netting*. This states that, if a company defaults on one contract it has with a counterparty, it must default on all outstanding contracts with the counterparty.

Netting has been successfully tested in the courts in most jurisdictions. It can substantially reduce credit risk for a financial institution. Consider, for example, a financial institution that has three contracts outstanding with a particular counterparty. The contracts are worth +\$10 million, +\$30 million, and −\$25 million to the financial institution. Suppose the counterparty runs into financial difficulties and defaults on its outstanding obligations. To the counterparty the three contracts have values of −\$10 million, −\$30 million, and +\$25 million, respectively. Without netting, the counterparty would default on the first two contracts and retain the third for a loss to the financial institution of \$40 million. With netting, it is compelled to default on all three contracts for a loss to the financial institution of \$15 million.[14]

Suppose a financial institution has a portfolio of N derivatives contracts with a particular counterparty. Suppose that the no-default value of the ith contract is V_i and the amount recovered in the event of default is the recovery rate times this no default value. Without netting, the financial institution loses

$$(1 - R)\sum_{i=1}^{N}\max(V_i, 0)$$

[14] Note that if the third contract were worth −\$45 million to the financial institution instead of −\$25 million, the counterparty would choose not to default and there would be no loss to the financial institution.

where R is the recovery rate. With netting, it loses

$$(1 - R)\max\left(\sum_{i=1}^{N} V_i, 0\right)$$

Without netting, its loss is the payoff from a portfolio of call options on the contracts where each option has a strike price of zero. With netting, it is the payoff from a single option on the portfolio of contracts with a strike price of zero. The value of an option on a portfolio is never greater than, and is often considerably less than, the value of the corresponding portfolio of options.

The analysis presented in the previous section can be extended so that equation (22.5) gives the present value of the expected loss from all contracts with a counterparty when netting agreements are in place. This is achieved by redefining v_i in the equation as the present value of a derivative that pays off the exposure at time t_i on the portfolio of all contracts with a counterparty.

A challenging task for a financial institution when considering whether it should enter into a new derivatives contract with a counterparty is to calculate the incremental effect on expected credit losses. This can be done by using equation (22.5) in the way just described to calculate expected default costs with and without the contract. It is interesting to note that, because of netting, the incremental effect of a new contract on expected default losses can be negative. This happens when the value of the new contract is negatively correlated with the value of existing contracts.

Collateralization

Another clause frequently used to mitigate credit risks is known as *collateralization*. Suppose that a company and a financial institution have entered into a number of derivatives contracts. A typical collateralization agreement specifies that the contracts be valued periodically. If the total value of the contracts to the financial institution is above a specified threshold level, the agreement requires the cumulative collateral posted by the company to equal the difference between the value of the contracts to the financial institution and the threshold level. If, after the collateral has been posted, the value of the contracts moves in favor of the company so that the difference between value of the contract to the financial institution and the threshold level is less than the total margin already posted, the company can reclaim margin. In the event of a default by the company, the financial institution can seize the collateral. If the company does not post collateral as required, the financial institution can close out the contracts.

Suppose, for example, that the threshold level for the company is \$10 million and the contracts are marked to market daily for the purposes of collateralization. If on a particular day the value of the contracts to the financial institution rises from \$9 million to \$10.5 million, it can ask for \$0.5 million of collateral. If the next day the value of the contracts rises further to \$11.4 million it can ask for a further \$0.9 million of collateral. If the value of the contracts falls to \$10.9 million on the following day, the company can ask for \$0.5 million of the collateral to be returned. Note that the threshold (\$10 million in this case) can be regarded as a line of credit that the financial institution is prepared to grant to the company.

The margin must be deposited by the company with the financial institution in cash or in the form of acceptable securities such as bonds. The securities are subject to a discount known as a *haircut* applied to their market value for the purposes of margin calculations. Interest is normally paid on cash.

If the collateralization agreement is a two-way agreement a threshold will also be specified for the financial institution. The company can then ask the financial institution to post collateral when the value of the outstanding contracts to the company exceeds the threshold.

Collateralization agreements provide a great deal of protection against the possibility of default (just as the margin accounts discussed in Chapter 2 provide protection for people who trade futures on an exchange). However, the threshold amount is not subject to protection. Furthermore, even when the threshold is zero, the protection is not total. This is because, when a company gets into financial difficulties, it is likely to stop responding to requests to post collateral. By the time the counterparty exercises its right to close out contracts, their value may have moved further in its favor.

Downgrade Triggers

Another credit risk mitigation technique used by a financial institution is known as a *downgrade trigger*. This is a clause stating that if the credit rating of the counterparty falls below a certain level, say Baa, the financial institution has the option to close out a derivatives contract at its market value.

Downgrade triggers do not provide protection from a big jump in a company's credit rating (for example, from A to default). Also, downgrade triggers work well only if relatively little use is made of them. If a company has many downgrade triggers outstanding with its counterparties, they are liable to provide little protection to the counterparties (see Business Snapshot 22.1).

22.9 DEFAULT CORRELATION

The term *default correlation* is used to describe the tendency for two companies to default at about the same time. There are a number of reasons why default correlation exists. Companies in the same industry or the same geographic region tend to be affected similarly by external events and as a result may experience financial difficulties at the same time. Economic conditions generally cause average default rates to be higher in some years than in other years. A default by one company may cause a default by another—the credit contagion effect. Default correlation means that credit risk cannot be completely diversified away and is the major reason why risk-neutral default probabilities are greater than real-world default probabilities (see Section 22.5).

Default correlation is important in the determination of probability distributions for default losses from a portfolio of exposures to different counterparties. Two types of default correlation models that have been suggested by researchers are referred to as *reduced form models* and *structural models*.

Reduced form models assume that the default intensities for different companies follow stochastic processes and are correlated with macroeconomic variables. When the default intensity for company A is high there is a tendency for the default intensity for company B to be high. This induces a default correlation between the two companies.

Reduced form models are mathematically attractive and reflect the tendency for

Business Snapshot 22.1 Downgrade Triggers and Enron's Bankruptcy

In December 2001, Enron, one of the largest companies in the United States, went bankrupt. Right up to the last few days, it had an investment grade credit rating. The Moody's rating immediately prior to default was Baa3 and the S&P rating was BBB−. The default was, however, anticipated to some extent by the stock market because Enron's stock price fell sharply in the period leading up to the bankruptcy. The probability of default estimated by models such as the one described in Section 22.6 increased sharply during this period.

Enron had entered into a huge number of derivatives contracts with downgrade triggers. The downgrade triggers stated that, if its credit rating fell below investment grade (i.e., below Baa3/BBB−), its counterparties would have the option of closing out the contracts. Suppose that Enron had been downgraded to below investment grade in, say, October 2001. The contracts that counterparties would choose to close out would be those with negative values to Enron (and positive values to the counterparties). So, Enron would have been required to make huge cash payments to its counterparties. It would not have been able to do this and immediate bankruptcy would have resulted.

This example illustrates that downgrade triggers provide protection only when relatively little use is made of them. When a company enters into a huge number of contracts with downgrade triggers, they may actually cause a company to go bankrupt prematurely. In Enron's case, we could argue that it was going to go bankrupt anyway and accelerating the event by two months would not have done any harm. In fact, Enron did have a chance of survival in October 2001. Attempts were being made to work out a deal with another energy company, Dynergy, and so forcing bankruptcy in October 2001 was not in the interests of either creditors or shareholders.

The credit rating companies found themselves in a difficult position. If they downgraded Enron to recognize its deteriorating financial position, they were signing its death warrant. If they did not do so, there was a chance of Enron surviving.

economic cycles to generate default correlations. Their main disadvantage is that the range of default correlations that can be achieved is limited. Even when there is a perfect correlation between the default intensities of the two companies, the probability that they will both default during the same short period of time is usually very low. This is liable to be a problem in some circumstances. For example, when two companies operate in the same industry and the same country or when the financial health of one company is for some reason heavily dependent on the financial health of another company, a relatively high default correlation may be warranted. One approach to solving this problem is by extending the model so that the default intensity exhibits large jumps.

Structural models are based on a model similar to Merton's model (see Section 22.6). A company defaults if the value of its assets is below a certain level. Default correlation between companies A and B is introduced into the model by assuming that the stochastic process followed by the assets of company A is correlated with the stochastic process followed by the assets of company B. Structural models have the advantage over reduced form models that the correlation can be made as high as desired. Their main disadvantage is that they are liable to be computationally quite slow.

The Gaussian Copula Model for Time to Default

A model that has become a popular practical tool is the Gaussian copula model for the time to default. It can be characterized as a simplified structural model. It assumes that all companies will default eventually and attempts to quantify the correlation between the probability distributions of the times to default for two different companies.

The model can be used in conjunction with either real-world or risk-neutral default probabilities. The left tail of the real-world probability distribution for the time to default of a company can be estimated from data produced by rating agencies such as that in Table 22.1. The left tail of the risk-neutral probability distribution of the time to default can be estimated from bond prices using the approach in Section 22.4.

Define t_1 as the time to default of company 1 and t_2 as the time to default of company 2. If the probability distributions of t_1 and t_2 were normal, we could assume that the joint probability distribution of t_1 and t_2 is bivariate normal. As it happens, the probability distribution of a company's time to default is not even approximately normal. This is where a Gaussian copula model comes in. We transform t_1 and t_2 into new variables x_1 and x_2 using

$$x_1 = N^{-1}[Q_1(t_1)], \qquad x_2 = N^{-1}[Q_2(t_2)]$$

where Q_1 and Q_2 are the cumulative probability distributions for t_1 and t_2, respectively, and N^{-1} is the inverse of the cumulative normal distribution ($u = N^{-1}(v)$ when $v = N(u)$). These are "percentile-to-percentile" transformations. The 5-percentile point in the probability distribution for t_1 is transformed to $x_1 = -1.645$, which is the 5-percentile point in the standard normal distribution; the 10-percentile point in the probability distribution for t_1 is transformed to $x_1 = -1.282$, which is the 10-percentile point in the standard normal distribution, and so on. The t_2-to-x_2 transformation is similar.

By construction, x_1 and x_2 have normal distributions with mean zero and unit standard deviation. The model assumes that the joint distribution of x_1 and x_2 is bivariate normal. This assumption is referred to as using a *Gaussian copula*. The assumption is convenient because it means that the joint probability distribution of t_1 and t_2 is fully defined by the cumulative default probability distributions Q_1 and Q_2 for t_1 and t_2, together with a single correlation parameter.

The attraction of the Gaussian copula model is that it can be extended to many companies. Suppose that we are considering n companies and that t_i is the time to default of the ith company. We transform each t_i into a new variable, x_i, that has a standard normal distribution. The transformation is the percentile-to-percentile transformation

$$x_i = N^{-1}[Q_i(t_i)]$$

where Q_i is the cumulative probability distribution for t_i. It is then assumed that the x_i are multivariate normal. The default correlation between t_i and t_j is measured as the correlation between x_i and x_j. This is referred to as the *copula correlation*.[15]

The Gaussian copula is a useful way of representing the correlation structure between variables that are not normally distributed. It allows the correlation structure

[15] As an approximation, the copula correlation between t_i and t_j is often assumed to be the correlation between the equity returns for companies i and j.

of the variables to be estimated separately from their marginal (unconditional) distributions. Although the variables themselves are not multivariate normal, the approach assumes that after a transformation is applied to each variable they are multivariate normal.

Example 22.3

Suppose that we wish to simulate defaults during the next 5 years in 10 companies. The copula default correlations between each pair of companies is 0.2. For each company the cumulative probability of a default during the next 1, 2, 3, 4, 5 years is 1%, 3%, 6%, 10%, 15%, respectively. When a Gaussian copula is used we sample from a multivariate normal distribution to obtain the x_i ($1 \leqslant i \leqslant 10$) with the pairwise correlation between the x_i being 0.2. We then convert the x_i to t_i, a time to default. When the sample from the normal distribution is less than $N^{-1}(0.01) = -2.33$, a default takes place within the first year; when the sample is between -2.33 and $N^{-1}(0.03) = -1.88$, a default takes place during the second year; when the sample is between -1.88 and $N^{-1}(0.06) = -1.55$, a default takes place during the third year; when the sample is between -1.55 and $N^{-1}(0.10) = -1.28$, a default takes place during the fourth year; when the sample is between -1.28 and $N^{-1}(0.15) = -1.04$, a default takes place during the fifth year. When the sample is greater than -1.04, there is no default during the 5 years.

A Factor-Based Correlation Structure

To avoid defining a different correlation between x_i and x_j for each pair of companies i and j in the Gaussian copula model, a one-factor model is often used. The assumption is that

$$x_i = a_i F + \sqrt{1 - a_i^2}\, Z_i \tag{22.10}$$

In this equation, F is a common factor affecting defaults for all companies and Z_i is a factor affecting only company i. The variable F and the variables Z_i have independent standard normal distributions. The a_i are constant parameters between -1 and $+1$. The correlation between x_i and x_j is $a_i a_j$.[16]

Suppose that the probability that company i will default by a particular time T is $Q_i(T)$. Under the Gaussian copula model, a default happens by time T when $N(x_i) < Q_i(T)$ or $x_i < N^{-1}[Q_i(T)]$. From equation (22.10), this condition is

$$a_i F + \sqrt{1 - a_i^2}\, Z_i < N^{-1}[Q_i(T)]$$

or

$$Z_i < \frac{N^{-1}[Q_i(T)] - a_i F}{\sqrt{1 - a_i^2}}$$

Conditional on the value of the factor F, the probability of default is therefore

$$Q_i(T \mid F) = N\left(\frac{N^{-1}[Q_i(T)] - a_i F}{\sqrt{1 - a_i^2}}\right) \tag{22.11}$$

[16] The parameter a_i is sometimes approximated as the correlation of company i's equity returns with a well-diversified market index.

A particular case of the one-factor Gaussian copula model is where the probability distributions of default are the same for all i and the correlation between x_i and x_j is the same for all i and j. Suppose that $Q_i(T) = Q(T)$ for all i and that the common correlation is ρ, so that $a_i = \sqrt{\rho}$ for all i. Equation (22.11) becomes

$$Q(T \mid F) = N\left(\frac{N^{-1}[Q(T)] - \sqrt{\rho}\,F}{\sqrt{1-\rho}}\right) \tag{22.12}$$

Binomial Correlation Measure

An alternative correlation measure that has been used by rating agencies is the *binomial correlation measure*. For two companies A and B, this is the coefficient of correlation between:

1. A variable that equals 1 if company A defaults between times 0 and T, and 0 otherwise; and
2. A variable that equals 1 if company B defaults between times 0 and T, and 0 otherwise.

The measure is

$$\beta_{AB}(T) = \frac{P_{AB}(T) - Q_A(T)Q_B(T)}{\sqrt{[Q_A(T) - Q_A(T)^2][Q_B(T) - Q_B(T)^2]}} \tag{22.13}$$

where $P_{AB}(T)$ is the joint probability of A and B defaulting between time 0 and time T, $Q_A(T)$ is the cumulative probability that company A will default by time T, and $Q_B(T)$ is the cumulative probability that company B will default by time T. Typically $\beta_{AB}(T)$ depends on T, the length of the time period considered. Usually it increases as T increases.

From the definition of a Gaussian copula model, $P_{AB}(T) = M[x_A(T), x_B(T); \rho_{AB}]$, where $x_A(T) = N^{-1}(Q_A(T))$ and $x_B(T) = N^{-1}(Q_B(T))$ are the transformed times to default for companies A and B, and ρ_{AB} is the Gaussian copula correlation for the times to default for A and B. The function $M(a, b; \rho)$ is the probability that, in a bivariate normal distribution where the correlation between the variables is ρ, the first variable is less than a and the second variable is less than b.[17] It follows that

$$\beta_{AB}(T) = \frac{M[x_A(T),\ x_B(T);\ \rho_{AB}] - Q_A(T)Q_B(T)}{\sqrt{[Q_A(T) - Q_A(T)^2][Q_B(T) - Q_B(T)^2]}} \tag{22.14}$$

This shows that, if $Q_A(T)$ and $Q_B(T)$ are known, $\beta_{AB}(T)$ can be calculated from ρ_{AB} and vice versa. Usually ρ_{AB} is markedly greater than $\beta_{AB}(T)$. This illustrates the important point that the magnitude of a correlation measure depends on the way it is defined.

Example 22.4

Suppose that the probability of company A defaulting in a 1-year period is 1% and the probability of company B defaulting in a 1-year period is also 1%. In this case, $x_A(1) = x_B(1) = N^{-1}(0.01) = -2.326$. If ρ_{AB} is 0.20, $M(x_A(1), x_B(1), \rho_{AB}) = 0.000337$ and equation (22.14) shows that $\beta_{AB}(T) = 0.024$ when $T = 1$.

[17] See Technical Note 5 on the author's website for the calculation of $M(a, b; \rho)$.

22.10 CREDIT VaR

Credit value at risk can be defined analogously to the way value at risk is defined for market risks (see Chapter 20). For example, a credit VaR with a confidence level of 99.9% and a 1-year time horizon is the credit loss that we are 99.9% confident will not be exceeded over 1 year.

Consider a bank with a very large portfolio of similar loans. As an approximation, assume that the probability of default is the same for each loan and the correlation between each pair of loans is the same. When the Gaussian copula model for time to default is used, the right-hand side of equation (22.12) is to a good approximation equal to the percentage of defaults by time T as a function of F. The factor F has a standard normal distribution. We are X% certain that its value will be greater than $N^{-1}(1-X) = -N^{-1}(X)$. We are therefore X% certain that the percentage of losses over T years on a large portfolio will be less than $V(X, T)$, where

$$V(X,T) = N\left(\frac{N^{-1}[Q(T)] + \sqrt{\rho}\,N^{-1}(X)}{\sqrt{1-\rho}}\right) \qquad (22.15)$$

This result was first produced by Vasicek.[18] As in equation (22.12), $Q(T)$ is the probability of default by time T and ρ is the copula correlation between any pair of loans.

A rough estimate of the credit VaR when an X% confidence level is used and the time horizon is T is therefore $L(1-R)V(X,T)$, where L is the size of the loan portfolio and R is the recovery rate. The contribution of a particular loan of size L_i to the credit VaR is $L_i(1-R)V(X,T)$. This model underlies the formulas that regulators use for credit risk capital (see Business Snapshot 22.2).

Example 22.4

Suppose that a bank has a total of \$100 million of retail exposures. The 1-year probability of default averages 2% and the recovery rate averages 60%. The copula correlation parameter is estimated as 0.1. In this case,

$$V(0.999, 1) = N\left(\frac{N^{-1}(0.02) + \sqrt{0.1}\,N^{-1}(0.999)}{\sqrt{1-0.1}}\right) = 0.128$$

showing that the 99.9% worst case default rate is 12.8%. The 1-year 99.9% credit VaR is therefore $100 \times 0.128 \times (1-0.6)$ or \$5.13 million.

CreditMetrics

Many banks have developed other procedures for calculating credit VaR for internal use. One popular approach is known as CreditMetrics. This involves estimating a probability distribution of credit losses by carrying out a Monte Carlo simulation of the credit rating changes of all counterparties. Suppose we are interested in determining the probability distribution of losses over a 1-year period. On each simulation trial, we sample to determine the credit rating changes and defaults of all counterparties during the year. We then revalue our outstanding contracts to determine the total of credit losses for the

[18] See O. Vasicek, "Probability of Loss on a Loan Portfolio," Working Paper, KMV, 1987. Vasicek's results were published in *Risk* magazine in December 2002 under the title "Loan Portfolio Value".

Business Snapshot 22.2 Basel II

The Basel Committee on Bank Supervision is planning an overhaul of its procedures for calculating the capital banks are required to keep for the risks they are bearing. This is known as Basel II. No changes are planned to the way market risk capital is calculated (see Business Snapshot 20.1). A new capital requirement for operational risk is planned and significant changes have been proposed for the way in which capital is calculated for credit risk.

For banks eligible to use the Internal Ratings Based (IRB) approach, credit risk capital for a transaction is calculated as

$$UDR \times LGD \times EAD \times MatAd$$

Here UDR, the unexpected default rate, is the excess of the 99.9% worst case 1-year default rate over the expected 1-year default rate. It is calculated, using equation (22.15), as $V(X, T) - Q(T)$ with $X = 99.9\%$ and $T = 1$. The variable LGD is the percentage loss given default (similar to the variable we have been denoting by $1 - R$); EAD is the exposure at default; $MatAd$ is a maturity adjustment.

The rules for determining these numbers are complicated. For UDR, the 1-year probability of default, $Q(1)$, and a correlation parameter ρ are required. The 1-year probability of default is estimated by the bank and the rules for determining the correlation parameter depend on the type of exposure (retail, corporate, sovereign, etc.). For retail exposures, banks also determine LGD and EAD internally. For corporate exposures, banks using the "Advanced IRB" determine LGD and EAD internally, but for banks using the "Foundation IRB" approach there are rules prescribed for determining LGD and EAD. The maturity adjustment is an increasing function of the maturity of the instrument and equals 1.0 when the maturity of the instrument is in 1 year.

year. After a large number of simulation trials, a probability distribution for credit losses is obtained. This can be used to calculate credit VaR.

This approach is liable to be computationally quite time intensive. However, it has the advantage that credit losses are defined as those arising from credit downgrades as well as defaults. Also the impact of credit mitigation clauses such as those described in Section 22.8 can be approximately incorporated into the analysis.

Table 22.6 is typical of the historical data provided by rating agencies on credit rating changes and could be used as a basis for a CreditMetrics Monte Carlo simulation. It shows the percentage probability of a bond moving from one rating category to another during a 1-year period. For example, a bond that starts with an A credit rating has a 91.84% chance of still having an A rating at the end of 1 year. It has a 0.02% chance of defaulting during the year, a 0.10% chance of dropping to B, and so on.[19]

In sampling to determine credit losses, the credit rating changes for different counterparties should not be assumed to be independent. A Gaussian copula model is typically used to construct a joint probability distribution of rating changes similarly to the way it is used in the model in the previous section to describe the joint probability

[19] Technical Note 11 on the author's website explains how a table such as Table 22.6 can be used to calculate transition matrices for periods other than 1 year.

Table 22.6 One-year ratings transition matrix, 1970–2006, with probabilities expressed as percentages and adjustments for transitions to the WR (without rating) category. *Source*: Moody's.

Initial rating	*Rating at year-end*								
	Aaa	*Aa*	*A*	*Baa*	*Ba*	*B*	*Caa*	*Ca–C*	*Default*
Aaa	91.56	7.73	0.69	0.00	0.02	0.00	0.00	0.00	0.00
Aa	0.86	91.43	7.33	0.29	0.06	0.02	0.00	0.00	0.01
A	0.06	2.64	91.48	5.14	0.53	0.10	0.02	0.00	0.02
Baa	0.05	0.22	5.16	88.70	4.60	0.84	0.23	0.03	0.19
Ba	0.01	0.07	0.52	6.17	83.10	8.25	0.58	0.05	1.26
B	0.01	0.05	0.19	0.41	6.27	81.65	5.17	0.75	5.50
Caa	0.00	0.04	0.04	0.25	0.79	10.49	65.47	4.44	18.47
Ca–C	0.00	0.00	0.00	0.00	0.46	2.78	11.07	47.83	37.85
Default	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00

distribution of times to default. The copula correlation between the rating transitions for two companies is usually set equal to the correlation between their equity returns using a factor model similar to that in Section 22.9.

As an illustration of the CreditMetrics approach suppose that we are simulating the rating change of a Aaa and a Baa company over a 1-year period using the transition matrix in Table 22.6. Suppose that the correlation between the equities of the two companies is 0.2. On each simulation trial, we would sample two variables x_A and x_B from normal distributions so that their correlation is 0.2. The variable x_A determines the new rating of the Aaa company and variable x_B determines the new rating of the Baa company. Since $N^{-1}(0.9156) = 1.3761$, $N^{-1}(0.9156 + 0.0773) = 2.4522$, and $N^{-1}(0.9156 + 0.0773 + 0.0069) = 3.5401$, the Aaa company stays Aaa-rated if $x_A < 1.3761$, it becomes Aa-rated if $1.3761 \leqslant x_A < 2.4522$, and it becomes A-rated if $2.4522 \leqslant x_A < 3.5401$. Similarly, since $N^{-1}(0.0005) = -3.2905$, $N^{-1}(0.0005 + 0.0022) = -2.7822$, and $N^{-1}(0.0005 + 0.0022 + 0.0516) = -1.6045$, the Baa company becomes Aaa-rated if $x_B < -3.2905$, it becomes Aa-rated if $-3.2905 \leqslant x_B < -2.7822$, and it becomes A-rated if $-2.7822 \leqslant x_B < -1.6045$. The Aaa-rated company never defaults during the year. The Baa-rated company defaults when $x_B > N^{-1}(0.9981)$, that is, when $x_B > 2.8943$.

SUMMARY

The probability that a company will default during a particular period of time in the future can be estimated from historical data, bond prices, or equity prices. The default probabilities calculated from bond prices are risk-neutral probabilities, whereas those calculated from historical data are real-world probabilities. Real-world probabilities should be used for scenario analysis and the calculation of credit VaR. Risk-neutral probabilities should be used for valuing credit-sensitive instruments. Risk-neutral default probabilities are often significantly higher than real-world default probabilities.

The expected loss experienced from a counterparty default is reduced by what is known as netting. This is a clause in most contracts written by a financial institution stating that, if a counterparty defaults on one contract it has with the financial institution, it must default on all contracts it has with the financial institution. Losses are also reduced by collateralization and downgrade triggers. Collateralization requires the counterparty to post collateral and a downgrade trigger gives a financial institution the option to close out a contract if the credit rating of a counterparty falls below a specified level.

Credit VaR can be defined similarly to the way VaR is defined for market risk. One approach to calculating it is the Gaussian copula model of time to default. This is used by regulators in the calculation of capital for credit risk. Another popular approach for calculating credit VaR is CreditMetrics. This uses a Gaussian copula model for credit rating changes.

FURTHER READING

Altman, E. I., "Measuring Corporate Bond Mortality and Performance," *Journal of Finance*, 44 (1989): 902–22.

Duffie, D., and K. Singleton, "Modeling Term Structures of Defaultable Bonds," *Review of Financial Studies*, 12 (1999): 687–720.

Finger, C.C, "A Comparison of Stochastic Default Rate Models," *RiskMetrics Journal*, 1 (November 2000): 49–73.

Hull, J., M. Predescu, and A. White, "Relationship between Credit Default Swap Spreads, Bond Yields, and Credit Rating Announcements," *Journal of Banking and Finance*, 28 (November 2004): 2789–2811.

Kealhofer, S., "Quantifying Default Risk I: Default Prediction," *Financial Analysts Journal*, 59, 1 (2003a): 30–44.

Kealhofer, S., "Quantifying Default Risk II: Debt Valuation," *Financial Analysts Journal*, 59, 3 (2003b): 78–92.

Li, D. X. "On Default Correlation: A Copula Approach," *Journal of Fixed Income*, March 2000: 43–54.

Litterman, R., and T. Iben, "Corporate Bond Valuation and the Term Structure of Credit Spreads," *Journal of Portfolio Management*, Spring 1991: 52–64.

Merton, R. C., "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, 29 (1974): 449–70.

Rodriguez, R. J., "Default Risk, Yield Spreads, and Time to Maturity," *Journal of Financial and Quantitative Analysis*, 23 (1988): 111–17.

Vasicek, O., "Loan Portfolio Value," *Risk* (December 2002), 160–62.

Questions and Problems (Answers in the Solutions Manual)

22.1. The spread between the yield on a 3-year corporate bond and the yield on a similar risk-free bond is 50 basis points. The recovery rate is 30%. Estimate the average default intensity per year over the 3-year period.

22.2. Suppose that in Problem 22.1 the spread between the yield on a 5-year bond issued by the same company and the yield on a similar risk-free bond is 60 basis points. Assume the same recovery rate of 30%. Estimate the average default intensity per year over the 5-year period. What do your results indicate about the average default intensity in years 4 and 5?

22.3. Should researchers use real-world or risk-neutral default probabilities for (a) calculating credit value at risk and (b) adjusting the price of a derivative for defaults?

22.4. How are recovery rates usually defined?

22.5. Explain the difference between an unconditional default probability density and a default intensity.

22.6. Verify (a) that the numbers in the second column of Table 22.4 are consistent with the numbers in Table 22.1 and (b) that the numbers in the fourth column of Table 22.5 are consistent with the numbers in Table 22.4 and a recovery rate of 40%.

22.7. Describe how netting works. A bank already has one transaction with a counterparty on its books. Explain why a new transaction by a bank with a counterparty can have the effect of increasing or reducing the bank's credit exposure to the counterparty.

22.8. Suppose that the measure $\beta_{AB}(T)$ in equation (22.9) is the same in the real world and the risk-neutral world. Is the same true of the Gaussian copula measure, ρ_{AB}?

22.9. What is meant by a "haircut" in a collateralization agreement. A company offers to post its own equity as collateral. How would you respond?

22.10. Explain the difference between the Gaussian copula model for the time to default and CreditMetrics as far as the following are concerned: (a) the definition of a credit loss and (b) the way in which default correlation is modeled.

22.11. Suppose that the probability of company A defaulting during a 2-year period is 0.2 and the probability of company B defaulting during this period is 0.15. If the Gaussian copula measure of default correlation is 0.3, what is the binomial correlation measure?

22.12. Suppose that the LIBOR/swap curve is flat at 6% with continuous compounding and a 5-year bond with a coupon of 5% (paid semiannually) sells for 90.00. How would an asset swap on the bond be structured? What is the asset swap spread that would be calculated in this situation?

22.13. Show that the value of a coupon-bearing corporate bond is the sum of the values of its constituent zero-coupon bonds when the amount claimed in the event of default is the no-default value of the bond, but that this is not so when the claim amount is the face value of the bond plus accrued interest.

22.14. A 4-year corporate bond provides a coupon of 4% per year payable semiannually and has a yield of 5% expressed with continuous compounding. The risk-free yield curve is flat at 3% with continuous compounding. Assume that defaults can take place at the end of each year (immediately before a coupon or principal payment) and that the recovery rate is 30%. Estimate the risk-neutral default probability on the assumption that it is the same each year.

22.15. A company has issued 3- and 5-year bonds with a coupon of 4% per annum payable annually. The yields on the bonds (expressed with continuous compounding) are 4.5% and 4.75%, respectively. Risk-free rates are 3.5% with continuous compounding for all maturities. The recovery rate is 40%. Defaults can take place halfway through each year. The risk-neutral default rates per year are Q_1 for years 1 to 3 and Q_2 for years 4 and 5. Estimate Q_1 and Q_2.

22.16. Suppose that a financial institution has entered into a swap dependent on the sterling interest rate with counterparty X and an exactly offsetting swap with counterparty Y. Which of the following statements are true and which are false?

(a) The total present value of the cost of defaults is the sum of the present value of the cost of defaults on the contract with X plus the present value of the cost of defaults on the contract with Y.
(b) The expected exposure in 1 year on both contracts is the sum of the expected exposure on the contract with X and the expected exposure on the contract with Y.
(c) The 95% upper confidence limit for the exposure in 1 year on both contracts is the sum of the 95% upper confidence limit for the exposure in 1 year on the contract with X and the 95% upper confidence limit for the exposure in 1 year on the contract with Y.

Explain your answers.

22.17. A company enters into a 1-year forward contract to sell $100 for AUD150. The contract is initially at the money. In other words, the forward exchange rate is 1.50. The 1-year dollar risk-free rate of interest is 5% per annum. The 1-year dollar rate of interest at which the counterparty can borrow is 6% per annum. The exchange rate volatility is 12% per annum. Estimate the present value of the cost of defaults on the contract. Assume that defaults are recognized only at the end of the life of the contract.

22.18. Suppose that in Problem 22.17, the 6-month forward rate is also 1.50 and the 6-month dollar risk-free interest rate is 5% per annum. Suppose further that the 6-month dollar rate of interest at which the counterparty can borrow is 5.5% per annum. Estimate the present value of the cost of defaults assuming that defaults can occur either at the 6-month point or at the 1-year point? (If a default occurs at the 6-month point, the company's potential loss is the market value of the contract.)

22.19. "A long forward contract subject to credit risk is a combination of a short position in a no-default put and a long position in a call subject to credit risk." Explain this statement.

22.20. Explain why the credit exposure on a matched pair of forward contracts resembles a straddle.

22.21. Explain why the impact of credit risk on a matched pair of interest rate swaps tends to be less than that on a matched pair of currency swaps.

22.22. "When a bank is negotiating currency swaps, it should try to ensure that it is receiving the lower interest rate currency from a company with a low credit risk." Explain why.

22.23. Does put–call parity hold when there is default risk? Explain your answer.

22.24. Suppose that in an asset swap B is the market price of the bond per dollar of principal, B^* is the default-free value of the bond per dollar of principal, and V is the present value of the asset swap spread per dollar of principal. Show that $V = B^* - B$.

22.25. Show that under Merton's model in Section 22.6 the credit spread on a T-year zero-coupon bond is $-\ln[N(d_2) + N(-d_1)/L]/T$, where $L = De^{-rT}/V_0$.

22.26. Suppose that the spread between the yield on a 3-year zero-coupon riskless bond and a 3-year zero-coupon bond issued by a corporation is 1%. By how much does Black–Scholes overstate the value of a 3-year European option sold by the corporation.

22.27. Give an example of (a) right-way risk and (b) wrong-way risk.

Assignment Questions

22.28. Suppose a 3-year corporate bond provides a coupon of 7% per year payable semi-annually and has a yield of 5% (expressed with semiannual compounding). The yields for all maturities on risk-free bonds is 4% per annum (expressed with semiannual compounding). Assume that defaults can take place every 6 months (immediately before a coupon payment) and the recovery rate is 45%. Estimate the default probabilities assuming (a) that the unconditional default probabilities are the same on each possible default date and (b) that the default probabilities conditional on no earlier default are the same on each possible default date.

22.29. A company has 1- and 2-year bonds outstanding, each providing a coupon of 8% per year payable annually. The yields on the bonds (expressed with continuous compounding) are 6.0% and 6.6%, respectively. Risk-free rates are 4.5% for all maturities. The recovery rate is 35%. Defaults can take place halfway through each year. Estimate the risk-neutral default rate each year.

22.30. Explain carefully the distinction between real-world and risk-neutral default probabilities. Which is higher? A bank enters into a credit derivative where it agrees to pay $100 at the end of 1 year if a certain company's credit rating falls from A to Baa or lower during the year. The 1-year risk-free rate is 5%. Using Table 22.6, estimate a value for the derivative. What assumptions are you making? Do they tend to overstate or understate the value of the derivative.

22.31. The value of a company's equity is $4 million and the volatility of its equity is 60%. The debt that will have to be repaid in 2 years is $15 million. The risk-free interest rate is 6% per annum. Use Merton's model to estimate the expected loss from default, the probability of default, and the recovery rate in the event of default. Explain why Merton's model gives a high recovery rate. (*Hint*: The Solver function in Excel can be used for this question, as indicated in footnote 10.)

22.32. Suppose that a bank has a total of $10 million of exposures of a certain type. The 1-year probability of default averages 1% and the recovery rate averages 40%. The copula correlation parameter is 0.2. Estimate the 99.5% 1-year credit VaR.

CHAPTER 23 Credit Derivatives

The most exciting developments in derivatives markets since the late 1990s have been in the credit derivatives area. In 2000 the total notional principal for outstanding credit derivatives contracts was about $800 billion. By June 2007 this had grown to over $42 trillion. Credit derivatives are contracts where the payoff depends on the creditworthiness of one or more companies or countries. This chapter explains how credit derivatives work and discusses some valuation issues.

Credit derivatives allow companies to trade credit risks in much the same way that they trade market risks. Banks and other financial institutions used to be in the position where they could do little once they had assumed a credit risk except wait (and hope for the best). Now they can actively manage their portfolios of credit risks, keeping some and entering into credit derivatives contracts to protect themselves from others. As indicated in Business Snapshot 23.1, banks have been the biggest buyers of credit protection and insurance companies have been the biggest sellers.

Credit derivatives can be categorized as "single name" or "multiname". The most popular single-name credit derivative is a credit default swap. The payoff from this instrument depends on what happens to one company or country. There are two sides to the contract: the buyer and seller of protection. There is a payoff from the seller of protection to the buyer of protection if the specified entity (company or country) defaults on its obligations. The most popular multiname credit derivative is a collateralized debt obligation. In this, a portfolio of debt instruments is specified and a complex structure is created where the cash flows from the portfolio are channelled to different categories of investors. Multiname credit derivatives increased in popularity relative to single-name credit derivatives up to June 2007. In December 2004 they accounted for about 20% of the credit derivatives market, but by June 2007 their share of the market had risen to over 43%.

In July 2007 investors lost confidence in the subprime mortgage market in the United States. The events leading up to this are explained in Section 23.7 and Business Snapshot 23.3. Interest in multiname structured credit products, whether they involved subprime mortgages or not, declined. Single-name credit derivatives continued to be actively traded after July 2007, but it may be some time before multiname structured credit products regain their former popularity.

This chapter starts by explaining how credit default swaps work and how they are valued. It then covers the trading of forwards and options on credit default swaps and

Business Snapshot 23.1 Who Bears the Credit Risk?

Traditionally banks have been in the business of making loans and then bearing the credit risk that the borrower will default. But this is changing. Banks have for some time been reluctant to keep loans on their balance sheets. This is because, after the capital required by regulators has been accounted for, the average return earned on loans is often less attractive than that on other assets. During the 1990s, banks created asset-backed securities (similar to the mortgage-backed securities discussed in Chapter 31) to pass loans (and their credit risk) on to investors. In the late 1990s and early 2000s, banks have made extensive use of credit derivatives to shift the credit risk in their loans to other parts of the financial system.

If banks have been net buyers of credit protection, who have been net sellers? The answer is insurance companies. Insurance companies are not regulated in the same way as banks and as a result are sometimes more willing to bear credit risks than banks.

The result of all this is that the financial institution bearing the credit risk of a loan is often different from the financial institution that did the original credit checks. Whether this proves to be good for the overall health of the financial system remains to be seen.

total return swaps. It explains credit indices, basket credit default swaps, asset-backed securities, and collateralized debt obligations. It expands on the material in Chapter 22 to show how the Gaussian copula model of default correlation can be used to value tranches of collateralized debt obligations.

23.1 CREDIT DEFAULT SWAPS

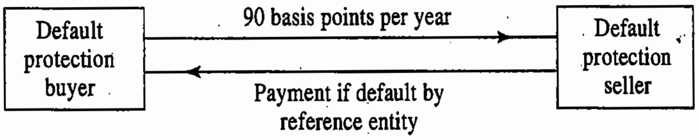
The most popular credit derivative is a *credit default swap* (CDS). This is a contract that provides insurance against the risk of a default by particular company. The company is known as the *reference entity* and a default by the company is known as a *credit event*. The buyer of the insurance obtains the right to sell bonds issued by the company for their face value when a credit event occurs and the seller of the insurance agrees to buy the bonds for their face value when a credit event occurs.[1] The total face value of the bonds that can be sold is known as the credit default swap's *notional principal*.

The buyer of the CDS makes periodic payments to the seller until the end of the life of the CDS or until a credit event occurs. These payments are typically made in arrears every quarter, every half year, or every year. The settlement in the event of a default involves either physical delivery of the bonds or a cash payment.

An example will help to illustrate how a typical deal is structured. Suppose that two parties enter into a 5-year credit default swap on March 1, 2009. Assume that the notional principal is $100 million and the buyer agrees to pay 90 basis points annually for protection against default by the reference entity.

[1] The face value (or par value) of a coupon-bearing bond is the principal amount that the issuer repays at maturity if it does not default.

Figure 23.1 Credit default swap.



The CDS is shown in Figure 23.1. If the reference entity does not default (i.e., there is no credit event), the buyer receives no payoff and pays $900,000 on March 1 of each of the years 2010, 2011, 2012, 2013, and 2014. If there is a credit event, a substantial payoff is likely. Suppose that the buyer notifies the seller of a credit event on June 1, 2012 (a quarter of the way into the fourth year). If the contract specifies physical settlement, the buyer has the right to sell bonds issued by the reference entity with a face value of $100 million for $100 million. If the contract requires cash settlement, an independent calculation agent will poll dealers to determine the mid-market value of the cheapest deliverable bond a predesignated number of days after the credit event. Suppose this bond is worth $35 per $100 of face value. The cash payoff would be $65 million.

The regular quarterly, semiannual, or annual payments from the buyer of protection to the seller of protection cease when there is a credit event. However, because these payments are made in arrears, a final accrual payment by the buyer is usually required. In our example, where there is a default on June 1, 2012, the buyer would be required to pay to the seller the amount of the annual payment accrued between March 1, 2012, and June 1, 2012 (approximately $225,000), but no further payments would be required.

The total amount paid per year, as a percent of the notional principal, to buy protection is known as the *CDS spread.* Several large banks are market makers in the credit default swap market. When quoting on a new 5-year credit default swap on a company, a market maker might bid 250 basis points and offer 260 basis points. This means that the market maker is prepared to buy protection by paying 250 basis points per year (i.e., 2.5% of the principal per year) and to sell protection for 260 basis points per year (i.e., 2.6% of the principal per year).

Many different companies and countries are reference entities for the CDS contracts that trade. Under the most popular arrangement, payments are made quarterly in arrears. Contracts with maturities of 5 years are most popular, but other maturities such as 1, 2, 3, 7, and 10 years are not uncommon. Usually contracts mature on one of the following standard dates: March 20, June 20, September 20, and December 20. The effect of this is that the actual time to maturity of a contract when it is initiated is close to, but not necessarily the same as, the number of years to maturity that is specified. Suppose you call a dealer on November 15, 2009, to buy 5-year protection on a company. The contract would probably last until December 20, 2014. Your first payment would be due on December 20, 2009, and would equal an amount covering the November 15, 2009, to December 20, 2009, period.[2] A key aspect of a CDS contract is the definition of default. In contracts on European reference entities restructuring is typically included as a credit event, whereas in contracts on North American reference entities it is not.

[2] If the time to the first standard date is less than 1 month, then the first payment is typically made on the second standard payment date; otherwise it is made on the first standard payment date.

Credit Default Swaps and Bond Yields

A CDS can be used to hedge a position in a corporate bond. Suppose that an investor buys a 5-year corporate bond yielding 7% per year for its face value and at the same time enters into a 5-year CDS to buy protection against the issuer of the bond defaulting. Suppose that the CDS spread is 200 basis points, or 2%, per annum. The effect of the CDS is to convert the corporate bond to a risk-free bond (at least approximately). If the bond issuer does not default the investor earns 5% per year when the CDS spread is netted against the corporate bond yield. If the bond does default the investor earns 5% up to the time of the default. Under the terms of the CDS, the investor is then able to exchange the bond for its face value. This face value can be invested at the risk-free rate for the remainder of the 5 years.

The n-year CDS spread should be approximately equal to the excess of the par yield on an n-year corporate bond over the par yield on an n-year risk-free bond. If it is markedly less than this, an investor can earn more than the risk-free rate by buying the corporate bond and buying protection. If it is markedly greater than this, an investor can borrow at less than the risk-free rate by shorting the corporate bond and selling CDS protection. These are not perfect arbitrages. But they are close to perfect and do give a good guide to the relationship between CDS spreads and bond yields. CDS spreads can be used to imply the risk-free rates used by market participants. As discussed in Section 22.4, the average implied risk-free rate appears to be approximately equal to the LIBOR/swap rate minus 10 basis points.

The Cheapest-to-Deliver Bond

As explained in Section 22.3, the recovery rate on a bond is defined as the value of the bond immediately after default as a percent of face value. This means that the payoff from a CDS is $L(1 - R)$, where L is the notional principal and R is the recovery rate.

Usually a CDS specifies that a number of different bonds can be delivered in the event of a default. The bonds typically have the same seniority, but they may not sell for the same percentage of face value immediately after a default.[3] This gives the holder of a CDS a cheapest-to-deliver bond option. When a default happens the buyer of protection (or the calculation agent in the event of cash settlement) will review alternative deliverable bonds and choose for delivery the one that can be purchased most cheaply.

23.2 VALUATION OF CREDIT DEFAULT SWAPS

Mid-market CDS spreads on individual reference entities (i.e., the average of the bid and offer CDS spreads quoted by brokers) can be calculated from default probability estimates. We will illustrate how this is done with a simple example.

Suppose that the probability of a reference entity defaulting during a year conditional on no earlier default is 2%. Table 23.1 shows survival probabilities and unconditional default probabilities (i.e., default probabilities as seen at time zero) for each of the

[3] There are a number of reasons for this. The claim that is made in the event of a default is typically equal to the bond's face value plus accrued interest. Bonds with high accrued interest at the time of default therefore tend to have higher prices immediately after default. Also the market may judge that in the event of a reorganization of the company some bond holders will fare better than others.

Table 23.1 Unconditional default probabilities and survival probabilities.

Time (years)	*Default probability*	*Survival probability*
1	0.0200	0.9800
2	0.0196	0.9604
3	0.0192	0.9412
4	0.0188	0.9224
5	0.0184	0.9039

5 years. The probability of a default during the first year is 0.02 and the probability the reference entity will survive until the end of the first year is 0.98. The probability of a default during the second year is $0.02 \times 0.98 = 0.0196$ and the probability of survival until the end of the second year is $0.98 \times 0.98 = 0.9604$. The probability of default during the third year is $0.02 \times 0.9604 = 0.0192$, and so on.

We will assume that defaults always happen halfway through a year and that payments on the credit default swap are made once a year, at the end of each year. We also assume that the risk-free (LIBOR) interest rate is 5% per annum with continuous compounding and the recovery rate is 40%. There are three parts to the calculation. These are shown in Tables 23.2, 23.3, and 23.4.

Table 23.2 shows the calculation of the present value of the expected payments made on the CDS assuming that payments are made at the rate of s per year and the notional principal is \$1. For example, there is a 0.9412 probability that the third payment of s is made. The expected payment is therefore $0.9412s$ and its present value is $0.9412se^{-0.05\times3} = 0.8101s$. The total present value of the expected payments is $4.0704s$.

Table 23.3 shows the calculation of the present value of the expected payoff assuming a notional principal of \$1. As mentioned earlier, we are assuming that defaults always happen halfway through a year. For example, there is a 0.0192 probability of a payoff halfway through the third year. Given that the recovery rate is 40% the expected payoff at this time is $0.0192 \times 0.6 \times 1 = 0.0115$. The present value of the expected payoff is $0.0115e^{-0.05\times2.5} = 0.0102$. The total present value of the expected payoffs is \$0.0511.

Table 23.2 Calculation of the present value of expected payments. Payment = s per annum.

Time (years)	*Probability of survival*	*Expected payment*	*Discount factor*	*PV of expected payment*
1	0.9800	$0.9800s$	0.9512	$0.9322s$
2	0.9604	$0.9604s$	0.9048	$0.8690s$
3	0.9412	$0.9412s$	0.8607	$0.8101s$
4	0.9224	$0.9224s$	0.8187	$0.7552s$
5	0.9039	$0.9039s$	0.7788	$0.7040s$
Total				$4.0704s$

Table 23.3 Calculation of the present value of expected payoff. Notional principal = $1.

Time (years)	*Probability of default*	*Recovery rate*	*Expected payoff ($)*	*Discount factor*	*PV of expected payoff ($)*
0.5	0.0200	0.4	0.0120	0.9753	0.0117
1.5	0.0196	0.4	0.0118	0.9277	0.0109
2.5	0.0192	0.4	0.0115	0.8825	0.0102
3.5	0.0188	0.4	0.0113	0.8395	0.0095
4.5	0.0184	0.4	0.0111	0.7985	0.0088
Total					0.0511

As a final step Table 23.4 considers the accrual payment made in the event of a default. For example, there is a 0.0192 probability that there will be a final accrual payment halfway through the third year. The accrual payment is $0.5s$. The expected accrual payment at this time is therefore $0.0192 \times 0.5s = 0.0096s$. Its present value is $0.0096se^{-0.05\times 2.5} = 0.0085s$. The total present value of the expected accrual payments is $0.0426s$.

From Tables 23.2 and 23.4, the present value of the expected payments is

$$4.0704s + 0.0426s = 4.1130s$$

From Table 23.3, the present value of the expected payoff is 0.0511. Equating the two gives

$$4.1130s = 0.0511$$

or $s = 0.0124$. The mid-market CDS spread for the 5-year deal we have considered should be 0.0124 times the principal or 124 basis points per year. This example is designed to illustrate the calculation methodology. In practice, we are likely to find that calculations are more extensive than those in Tables 23.2 to 23.4 because (a) payments are often made more frequently than once a year and (b) we are likely to want to assume that defaults can happen more frequently than once a year.

Table 23.4 Calculation of the present value of accrual payment.

Time (years)	*Probability of default*	*Expected accrual payment*	*Discount factor*	*PV of expected accrual payment*
0.5	0.0200	$0.0100s$	0.9753	$0.0097s$
1.5	0.0196	$0.0098s$	0.9277	$0.0091s$
2.5	0.0192	$0.0096s$	0.8825	$0.0085s$
3.5	0.0188	$0.0094s$	0.8395	$0.0079s$
4.5	0.0184	$0.0092s$	0.7985	$0.0074s$
Total				$0.0426s$

Marking to Market a CDS

At the time it is negotiated, a CDS, like most other swaps, is worth close to zero. Later it may have a positive or negative value. Suppose, for example the credit default swap in our example had been negotiated some time ago for a spread of 150 basis points, the present value of the payments by the buyer would be $4.1130 \times 0.0150 = 0.0617$ and the present value of the payoff would be 0.0511 as above. The value of swap to the seller would therefore be $0.0617 - 0.0511$, or 0.0106 times the principal. Similarly the mark-to-market value of the swap to the buyer of protection would be -0.0106 times the principal.

Estimating Default Probabilities

The default probabilities used to value a CDS should be risk-neutral default probabilities, not real-world default probabilities (see Section 22.5 for a discussion of the difference between the two). Risk-neutral default probabilities can be estimated from bond prices or asset swaps as explained in Chapter 22. An alternative is to imply them from CDS quotes. The latter approach is similar to the practice in options markets of implying volatilities from the prices of actively traded options.

Suppose we change the example in Tables 23.2, 23.3 and 23.4 so that we do not know the default probabilities. Instead we know that the mid-market CDS spread for a newly issued 5-year CDS is 100 basis points per year. We can reverse engineer our calculations to conclude that the implied default probability per year (conditional on no earlier default) is 1.61% per year.[4]

Binary Credit Default Swaps

A binary credit default swap is structured similarly to a regular credit default swap except that the payoff is a fixed dollar amount. Suppose that, in the example we considered in Tables 23.1 to 23.4, the payoff is \$1 instead of $1 - R$ dollars and the swap spread is s. Tables 23.1, 23.2 and 23.4 are the same, but Table 23.3 is replaced by

Table 23.5 Calculation of the present value of expected payoff from a binary credit default swap. Principal = \$1.

Time (years)	*Probability of default*	*Expected payoff (\$)*	*Discount factor*	*PV of expected payoff (\$)*
0.5	0.0200	0.0200	0.9753	0.0195
1.5	0.0196	0.0196	0.9277	0.0182
2.5	0.0192	0.0192	0.8825	0.0170
3.5	0.0188	0.0188	0.8395	0.0158
4.5	0.0184	0.0184	0.7985	0.0147
Total				0.0852

[4] Ideally we would like to estimate a different default probability for each year instead of a single default intensity. We could do this if we had spreads for 1-, 2-, 3-, 4-, and 5-year CDS swaps or bond prices.

Table 23.5. The CDS spread for a new binary CDS is given by $4.1130s = 0.0852$, so that the CDS spread, s, is 0.0207, or 207 basis points.

How Important is the Recovery Rate?

Whether we use CDS spreads or bond prices to estimate default probabilities we need an estimate of the recovery rate. However, provided that we use the same recovery rate for (a) estimating risk-neutral default probabilities and (b) valuing a CDS, the value of the CDS (or the estimate of the CDS spread) is not very sensitive to the recovery rate. This is because the implied probabilities of default are approximately proportional to $1/(1 - R)$ and the payoffs from a CDS are proportional to $1 - R$.

This argument does not apply to the valuation of binary CDS. Implied probabilities of default are still approximately proportional to $1/(1 - R)$. However, for a binary CDS, the payoffs from the CDS are independent of R. If we have a CDS spread for both a plain vanilla CDS and a binary CDS, we can estimate both the recovery rate and the default probability (see Problem 23.26).

The Future of the CDS Market

The credit default swap market survived the credit crunch of 2007 well. Credit default swaps have become important tools for managing credit risk. A financial institution can reduce its credit exposure to particular companies by buying protection. It can also use CDSs to diversify credit risk. For example, if a financial institution has too much credit exposure to a particular business sector, it can buy protection against defaults by companies in the sector and at the same time sell protection against default by companies in other unrelated sectors.

Some market participants think the growth of the CDS market will continue and that it will be as big as the interest rate swap market by 2010. Others are less optimistic. There is a potential asymmetric information problem in the CDS market that is not present in other over-the-counter derivatives markets (see Business Snapshot 23.2).

23.3 CREDIT INDICES

Participants in credit markets have developed indices to track credit default swap spreads. In 2004 there were agreements between different producers of indices that led to some consolidation. Two important standard portfolios used by index providers are:

1. CDX NA IG, a portfolio of 125 investment grade companies in North America
2. iTraxx Europe, a portfolio of 125 investment grade names in Europe

These portfolios are updated on March 20 and September 20 each year. Companies that are no longer investment grade are dropped from the portfolios and new investment grade companies are added.[5]

Suppose that the 5-year CDX NA IG index is quoted by a market maker as bid 65 basis points, offer 66 basis points. (This is referred to as the index spread.) Roughly

[5] On September 20, 2007, the Series 8 iTraxx Europe portfolio and the Series 9 CDX NA IG portfolio were defined. The series numbers indicate that by the end of September 2007 the iTraxx Europe portfolio had been updated seven times and the CDX NA IG portfolio had been updated eight times.

Business Snapshot 23.2 Is the CDS Market a Fair Game?

There is one important difference between credit default swaps and the other over-the-counter derivatives that we have considered in this book. The other over-the-counter derivatives depend on interest rates, exchange rates, equity indices, commodity prices, and so on. There is no reason to assume that any one market participant has better information than any other market participant about these variables.

Credit default swaps spreads depend on the probability that a particular company will default during a particular period of time. Arguably some market participants have more information to estimate this probability than others. A financial institution that works closely with a particular company by providing advice, making loans, and handling new issues of securities is likely to have more information about the creditworthiness of the company than another financial institution that has no dealings with the company. Economists refer to this as an *asymmetric information* problem.

Whether asymmetric information will curtail the expansion of the credit default swap market remains to be seen. Financial institutions emphasize that the decision to buy protection against the risk of default by a company is normally made by a risk manager and is not based on any special information that may exist elsewhere in the financial institution about the company.

speaking, this means that a trader can buy CDS protection on all 125 companies in the index for 66 basis points per company. Suppose a trader wants \$800,000 of protection on each company. The total cost is $0.0066 \times 800{,}000 \times 125$, or \$660,000 per year. The trader can similarly sell \$800,000 of protection on each of the 125 companies for a total of \$650,000 per annum. When a company defaults, the protection buyer receives the usual CDS payoff and the annual payment is reduced by $660{,}000/125 = \$5{,}280$. There is an active market in buying and selling CDS index protection for maturities of 3, 5, 7, and 10 years. The maturities for these types of contracts on the index are usually December 20 and June 20. (This means that a "5-year" contract actually lasts between $4\frac{3}{4}$ and $5\frac{1}{4}$ years.) Roughly speaking, the index is the average of the CDS spreads on the companies in the underlying portfolio.[6]

The precise way in which the contract works is a little more complicated than has just been described. For each index and each maturity a "coupon" is specified. A price is calculated from the quoted index spread using the following procedure:

1. Assume a recovery rate of 40% and four payments per year, made in arrears.
2. Imply a default intensity (hazard rate) from the quoted spread for the index. This involves calculations similar to those in Section 23.2. An iterative search is used to determine the default intensity that leads to the quoted spread.

[6] More precisely, the index is slightly lower than the average of the credit default swap spreads for the companies in the portfolio. To understand the reason for this consider a portfolio consisting of two companies, one with a spread of 1,000 basis points and the other with a spread of 10 basis points. To buy protection on the companies would cost slightly less than 505 basis points per company. This is because the 1,000 basis points is not expected to be paid for as long as the 10 basis points and should therefore carry less weight. Another complication for CDX NA IG, but not iTraxx Europe, is that the definition of default applicable to the index includes restructuring, whereas the definition for CDS contracts on the underlying companies does not.

3. Calculate a duration D for the CDS payments. This is the number that the index spread is multiplied by to get the present value of the spread payments. (In the example in Section 23.2, it is 4.1130.)
4. The price P is given by $P = 100 - 100 \times D \times (S - C)$, where S is the index spread and C is the coupon expressed in decimal form.

When a trader buys index protection the trader pays $100 - P$ per \$100 of the total remaining notional and the seller of protection receives this amount. (If $100 - P$ is negative, the buyer of protection receives money and the seller of protection pays money.) The buyer of protection then pays the coupon times the remaining notional on each payment date. (The remaining notional is the number of names in the index that have not yet defaulted multiplied by the principal per name.) The payoff when there is a default is calculated in the usual way. This arrangement facilitates trading because the regular quarterly payments made by the buyer of protection are independent of the index spread at the time the buyer enters into the contract.

Example 23.1

Suppose that the iTraxx Europe index quote is 34 basis points and the coupon is 40 basis points for a contract lasting exactly 5 years, with both quotes being expressed using a 30/360 day count. (This is the usual day count convention in CDS and CDS index markets.) The equivalent actual/actual quotes are 0.345% for the index and 0.406% for the coupon. Suppose that the yield curve is flat at 4% per year (actual/actual, continuously compounded). Assuming a recovery rate of 40% and four payments per year the implied hazard rate is 0.5717%. The duration is 4.447 years. The price is therefore

$$100 - 100 \times 4.447 \times (0.00345 - 0.00406) = 100.27$$

Consider a contract where protection is \$1 million per name. Initially, the seller of protection would pay the buyer $\$1{,}000{,}000 \times 125 \times 0.0027$. Thereafter, the buyer of protection would make quarterly payments in arrears at an annual rate of $\$1{,}000{,}000 \times 0.00406 \times n$, where n is the number of companies that have not defaulted. When a company defaults, the payoff is calculated in the usual way and there is an accrual payment from the buyer to the seller calculated at the rate of 0.406% per year on \$1 million.

23.4 CDS FORWARDS AND OPTIONS

Once the CDS market was well established, it was natural for derivatives dealers to trade forwards and options on credit default swap spreads.[7]

A forward credit default swap is the obligation to buy or sell a particular credit default swap on a particular reference entity at a particular future time T. If the reference entity defaults before time T, the forward contract ceases to exist. Thus a bank could enter into a forward contract to sell 5-year protection on a company for 280 basis points starting in 1 year. If the company defaulted before the 1-year point, the forward contract would cease to exist.

[7] The valuation of these instruments is discussed in J.C. Hull and A. White, "The Valuation of Credit Default Swap Options," *Journal of Derivatives*, 10, 5 (Spring 2003): 40–50.

A credit default swap option is an option to buy or sell a particular credit default swap on a particular reference entity at a particular future time T. For example, a trader could negotiate the right to buy 5-year protection on a company starting in 1 year for 280 basis points. This is a call option. If the 5-year CDS spread for the company in 1 year turns out to be more than 280 basis points, the option will be exercised; otherwise it will not be exercised. The cost of the option would be paid up front. Similarly an investor might negotiate the right to sell 5-year protection on a company for 280 basis points starting in 1 year. This is a put option. If the 5-year CDS spread for the company in 1 year turns out to be less than 280 basis points, the option will be exercised; otherwise it will not be exercised. Again the cost of the option would be paid up front. Like CDS forwards, CDS options are usually structured so that they will cease to exist if the reference entity defaults before option maturity.

23.5 BASKET CREDIT DEFAULT SWAPS

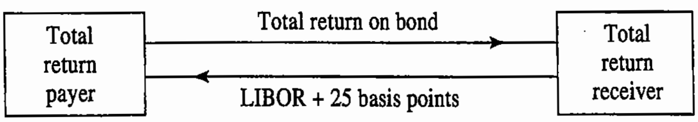
In what is referred to as a *basket credit default swap* there are a number of reference entities. An *add-up basket* CDS provides a payoff when any of the reference entities default. A *first-to-default* CDS provides a payoff only when the first default occurs. A *second-to-default* CDS provides a payoff only when the second default occurs. More generally, a *kth-to-default* CDS provides a payoff only when the kth default occurs. Payoffs are calculated in the same way as for a regular CDS. After the relevant default has occurred, there is a settlement. The swap then terminates and there are no further payments by either party.

23.6 TOTAL RETURN SWAPS

A *total return swap* is a type of credit derivative. It is an agreement to exchange the total return on a bond (or any portfolio of assets) for LIBOR plus a spread. The total return includes coupons, interest, and the gain or loss on the asset over the life of the swap.

An example of a total return swap is a 5-year agreement with a notional principal of \$100 million to exchange the total return on a corporate bond for LIBOR plus 25 basis points. This is illustrated in Figure 23.2. On coupon payment dates the payer pays the coupons earned on an investment of \$100 million in the bond. The receiver pays interest at a rate of LIBOR plus 25 basis points on a principal of \$100 million. (LIBOR is set on one coupon date and paid on the next as in a plain vanilla interest rate swap.) At the end of the life of the swap there is a payment reflecting the change in value of the bond. For example, if the bond increases in value by 10% over the life of the swap, the payer

Figure 23.2 Total return swap.



is required to pay \$10 million (= 10% of \$100 million) at the end of the 5 years. Similarly, if the bond decreases in value by 15%, the receiver is required to pay \$15 million at the end of the 5 years. If there is a default on the bond, the swap is usually terminated and the receiver makes a final payment equal to the excess of \$100 million over the market value of the bond.

If the notional principal is added to both sides at the end of the life of the swap, the total return swap can be characterized as follows. The payer pays the cash flows on an investment of \$100 million in the corporate bond. The receiver pays the cash flows on a \$100 million bond paying LIBOR plus 25 basis points. If the payer owns the corporate bond, the total return swap allows it to pass the credit risk on the bond to the receiver. If it does not own the bond, the total return swap allows it to take a short position in the bond.

Total return swaps are often used as a financing tool. One scenario that could lead to the swap in Figure 23.2 is as follows. The receiver wants financing to invest \$100 million in the reference bond. It approaches the payer (which is likely to be a financial institution) and agrees to the swap. The payer then invests \$100 million in the bond. This leaves the receiver in the same position as it would have been if it had borrowed money at LIBOR plus 25 basis points to buy the bond. The payer retains ownership of the bond for the life of the swap and faces less credit risk than it would have done if it had lent money to the receiver to finance the purchase of the bond, with the bond being used as collateral for the loan. If the receiver defaults the payer does not have the legal problem of trying to realize on the collateral. Total return swaps are similar to repos (see Section 4.1) in that they are structured to minimize credit risk when securities are being financed.

The spread over LIBOR received by the payer is compensation for bearing the risk that the receiver will default. The payer will lose money if the receiver defaults at a time when the reference bond's price has declined. The spread therefore depends on the credit quality of the receiver, the credit quality of the bond issuer, and the correlation between the two.

There are a number of variations on the standard deal we have described. Sometimes, instead of there being a cash payment for the change in value of the bond, there is physical settlement where the payer exchanges the underlying asset for the notional principal at the end of the life of the swap. Sometimes the change-in-value payments are made periodically rather than all at the end.

23.7 ASSET-BACKED SECURITIES

An *asset-backed security* (ABS) is a security created from a portfolio of loans, bonds, credit card receivables, mortgages, auto loans, aircraft leases, or other financial assets. (Even assets as unusual as royalties from the future sale of a piece of music are sometimes included.) As an example of the creation of an asset-backed security consider a bank that has made a large number of auto loans. The loans would typically be classified, according to the credit quality of the borrower, as "prime", "nonprime", and "subprime". Suppose there are 10,000 nonprime loans. Rather than keeping these as assets on its balance sheet the bank might decide to sell them to a special purpose vehicle (SPV), also known as a trust or a *conduit*. The SPV issues securities that are backed by the cash flows of the loans and proceeds to sell the securities to investors. The arrangement has the effect of insulating investors from the credit risk of the bank that issued the loans. The investors' return depends solely on the cash flows from the loans. The bank earns a fee for originating and servicing the loans. However, the credit risk associated with the loans is passed on to the investors.

Many different types of ABS can be created. Often the credit risk is allocated to a number of tranches. One simple arrangement involving a $100 million portfolio of debt instruments and three tranches is shown Figure 23.3. Suppose that the life of the ABS is five years. The first tranche, known as the equity tranche, finances 5% of the total principal and is promised a return of 30%; the second tranche, known as the mezzanine tranche, finances 20% of the total principal and is promised a return of 10%; the third tranche, the senior tranche, finances 75% of the principal and is promised a return of 6%. Tranches receive their return in order of seniority using a set of rules known as a *waterfall*. The cash flows from the portfolio of assets are first used to pay the investors in Tranche 3 their promised return of 6%. As far as is possible, they are then used to provide the investors in Tranche 2 with their promised return of 10%. Finally, residual cash flows from the portfolio are used to provide the Tranche 1 investors with a return of up to 30%. Consider what happens when the portfolio of assets starts to experience default losses. The return to the Tranche 1 investors is affected first. They earn less than 30% on their original investment and are likely to fail to get some of their principal back. When defaults get sufficiently high, Tranche 2 starts to be affected; and, if defaults are really high, Tranche 3 may not get its promised return.

Typically the senior tranche is rated AAA. The mezzanine tranche might be rated BBB. The equity tranche is usually not rated and is sometimes retained by the creator of the ABS. An asset-backed security is therefore a way of taking a portfolio of risky loans with a principal of $100 million and creating from it $75 million of AAA-rated debt. The SPV or conduit buys instruments from issuers and creates securities from them in the way we have described.

Figure 23.3 Possible structure for an ABS.

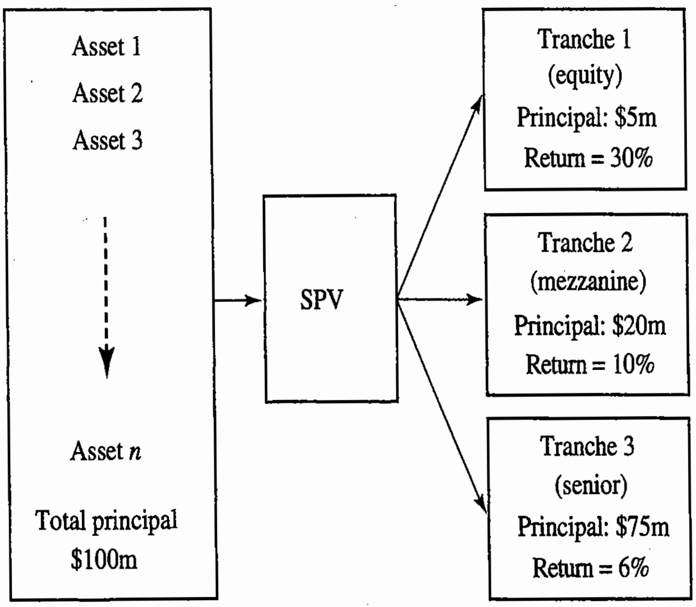
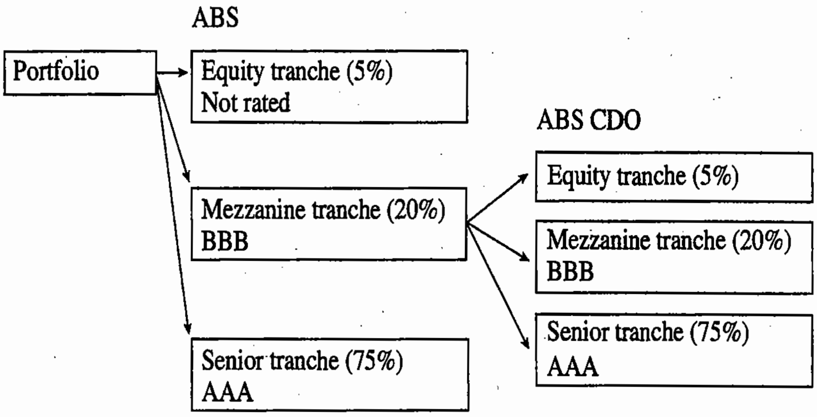


Figure 23.4 The creation of an ABS CDO.



The ABS mezzanine tranche is repackaged with other similar mezzanine tranches to form the ABS CDO

Dealers have been very creative—perhaps too creative—in their use of this type of structure. Mezzanine tranches are difficult to sell. To overcome this problem, dealers have put the mezzanine tranches from, say, 20 different asset-backed securities into a new asset-backed security. (This is known as an ABS CDO.) They then convinced rating agencies to assign a AAA rating to the most senior tranche of the new structure. A possible structure is shown in Figure 23.4. The AAA rating for the senior tranche of the ABS CDO is reasonable if the losses experienced by different mezzanine tranches are independent of each other. However, if all of the mezzanine tranches are likely to experience a high loss rate at the same time, the AAA-rated tranche is quite risky and liable to experience losses. This happened in mid-2007, as described in Business Snapshot 23.3. Investors who bought AAA-rated tranches that were created from BBB-rated mezzanine tranches that were in turn created from subprime mortgages found that their investments were steeply downgraded by rating agencies.

23.8 COLLATERALIZED DEBT OBLIGATIONS

A type of asset-backed security that has been particularly popular is a *collateralized debt obligation* (CDO). In this the assets being securitized are bonds issued by corporations or countries. The design of the instrument is similar to that in Figure 23.3 (except that there are usually more than three tranches). The creator of the CDO acquires a portfolio of bonds. These are passed on to an SPV which passes the income generated by the bonds to a series of tranches. The income from the bonds is first used to provide the promised return to the most senior tranche, then to the next most senior tranche, and so on. As is usual with asset-backed securities, the structure is designed so that the most senior tranche is rated AAA. The most junior (equity) tranche is sometimes retained by the arranger of the CDO. Assuming that the mezzanine tranche is rated BBB, the structure shown in Figure 23.3 could be used to take a $100 million portfolio consisting of, say, A-rated bonds and converting it to $75 million of AAA-rated instruments, $20 million of

Business Snapshot 23.3 The Credit Crunch of 2007

In July 2007, credit markets experienced a severe jolt. Credit spreads jumped up by about 50% and it became increasingly difficult for a wide range of private individuals and companies to borrow money. Why did this happen?

It is first necessary to look at the real-estate market in the United States. Between 2000 and 2006 house prices increased fast in most regions. As a result, mortgage lenders thought they were taking very little credit risk. It was considered that, if a borrower defaulted, the value of the house would be more than enough to cover the loan. Mortgage lenders changed their lending practices. Previously they had required borrowers to provide at least 20% of the cost of a house themselves. This requirement was relaxed. They also allowed the credit quality of borrowers to decline. The mortgage application process was lax and there were few checks of the income that mortgage applicants reported. As a result, some of the loans, which at the time were categorized as "subprime", are now termed "liar loans". The acronym "NINJA" (no income, no job, no assets) has also been coined to describe some of the loans. Another aspect of the lending was the popularity of adjustable rate mortgages or ARMs. These are arrangements where there is a low "starter interest rate" (e.g., 6%) that increases to the applicable floating rate at the end of 1 or 2 or 3 years. While the starter interest rate is in effect, the principal owing on the mortgage grows—often to more than the amount paid for the house. Many borrowers found their initial mortgage payments affordable but were unable (or unwilling) to handle the higher monthly payments that were due when the starter interest rate period ended.

Why did mortgage lenders become so casual in their approach to lending? There was a feeling that the good times would last for ever and house prices would continue to rise. The asset-backed securities market for subprime mortgages was flourishing and the mortgage originator did not usually keep the credit risk. The key question for many lenders was not "Do I want to lend money to this person and bear the credit risk for the life of the mortgage?" Instead it was "Can I sell this mortgage to a special purpose vehicle (conduit) and make a profit?"

In 2007 the bubble burst. Defaults on subprime mortgages started to increase because of the lax lending practices and because the period covered by the starter interest rates ended. Losses on subprime mortgages were forecast to be about 12%. This means that investors in the mezzanine tranches created from subprime mortgages were likely to experience losses and investors in the AAA-rated tranches created from these mezzanine tranches (see Figure 23.4) were also likely to experience losses. Securities in special purpose vehicles that had not yet been packaged and sold to investors were often financed by short-term commercial paper. It became impossible to roll over this commercial paper, creating a liquidity crisis.

The credit crunch had its origins in the US real-estate market, but was a worldwide phenomenon. Funds as far afield as Europe, Japan, and Australia owned the tranches of ABSs and ABS CDOs. They realized belatedly that they knew very little about the underlying assets (except that their tranches had originally been rated AAA) and the market for asset-backed securities dried up. There was a "flight to quality", where many investors were only willing to invest in safe assets such as Treasury bonds. As a result, credit spreads increased and new loans (whether short-term commercial paper or longer-term facilities) were very difficult to arrange.

BBB-rated instruments, and \$5 million of unrated instruments. This can add value because many investors want AAA-rated instruments and there is a limited supply of regular AAA-rated bonds. The objective of the originator of the CDO is to make money by selling the tranches to investors for more than the amount paid for the bonds. In many instances the bonds in a cash CDO can be traded by the originator of the CDO provided that certain conditions concerning the diversification of the portfolio are adhered to.

Synthetic CDOs

The structure we have just described is known as a *cash CDO*. A long position in a corporate bond has essentially the same credit risk as a short position in the corresponding credit default swap (i.e., a credit default swap where protection has been sold). This observation has led to an alternative way of creating a CDO. Instead of forming a portfolio of corporate bonds, the originator of the CDO forms a portfolio consisting of short positions in credit default swaps. The credit risks are then passed on to tranches. A CDO created in this way is known as a *synthetic CDO*.

A synthetic CDO is structured so that default losses on the credit default swap portfolio are allocated to tranches. Suppose that the total notional principal underlying the portfolio of credit default swaps is \$100 million and that there are three tranches. The situation might be as follows:

1. Tranche 1 is responsible for the first \$5 million of losses. As compensation for this, it earns 15% on the remaining Tranche 1 principal.
2. Tranche 2 is responsible for the next \$20 million of losses. As compensation for this, it earns 100 basis points on the remaining Tranche 2 principal.
3. Tranche 3 is responsible for all losses in excess of \$25 million. As compensation for this, it earns 10 basis points on the remaining Tranche 3 principal.

The principals for Tranches 1, 2, and 3 in this example are initially \$5 million, \$20 million, and \$75 million, respectively. The notional principal in a tranche is reduced by the losses that are paid for by the tranche holders. Initially Tranche 1 earns 15% on a principal of \$5 million. Suppose that after six months losses of \$1 million are experienced by the portfolio of credit default swaps. These losses are paid for by Tranche 1 and the notional principal of Tranche 1 is reduced by \$1 million, so that 15% is then earned on \$4 million rather than on \$5 million. If losses exceed \$5 million, Tranche 1 gets wiped out and Tranche 2 becomes responsible for losses. When losses reach \$7 million, the notional principal in Tranche 2 is \$18 million and Tranche 2 has by that time paid for losses totaling \$2 million.[8]

Single Tranche Trading

In Section 23.3 we discussed the portfolios of 125 companies that are used to generate CDX and iTraxx indices. The market uses these portfolios to define standard CDO tranches. The trading of these standard tranches is known as *single tranche trading*.

[8] In practice, tranche holders are required to post the initial tranche principal as collateral up front. This collateral earns LIBOR. When the tranche is responsible for the payoff on a CDS, the money is taken out of the collateral. The recovery amounts when defaults occur are usually used to retire principal on the most senior tranche. Note that the senior tranche does not lose the principal that is retired. It just fails to earn the spread (10 basis points in our example) on it.

A single tranche trade is an agreement where one side agrees to sell protection against losses on a tranche and the other side agrees to buy the protection. The tranche is not part of a synthetic CDO that someone has created but cash flows are calculated in the same way as if it were part of such a synthetic CDO.

In the case of the CDX NA IG index, the equity tranche covers losses between 0% and 3% of the principal. The second tranche, the *mezzanine tranche*, covers losses between 3% and 7%. The remaining tranches cover losses from 7% and 10%, 10% to 15%, 15% to 30%, and 30% to 100%. In the case of the iTraxx Europe index, the equity tranche covers losses between 0% and 3%. The mezzanine tranche covers losses between 3% and 6%. The remaining tranches cover losses from 6% to 9%, 9% to 12%, 12% to 22%, and 22% to 100%.

Table 23.6 shows the mid-market quotes for 5-year CDX NA IG and iTraxx Europe tranches on March 28, 2007. On that date the 5-year CDX NA IG index level was 38 basis points and the 5-year iTraxx Europe index was 24 basis points. The table shows that (ignoring bid–offer spreads) Trader A could buy 5-year protection against losses on the underlying portfolio that are in the range 7% to 10% from Trader B for 20.3 basis points. Suppose that the amount of protection bought is \$6 million. This is the initial tranche principal. Payoffs from Trader B to Trader A depend on default losses on the CDX NA IG portfolio. While the cumulative loss is less than 7% of the portfolio principal there is no payoff. As soon as the cumulative loss exceeds 7% of the portfolio principal, payoffs start. If, at the end of year 3, the cumulative loss rises from 7% to 8% of the portfolio principal, Trader B pays Trader A \$2 million and the tranche principal reduces to \$4 million. If, at the end of year 4, the cumulative loss increases from 8% to 10% of the portfolio principal, Trader B pays Trader A an additional \$4 million and the tranche principal reduces to zero. Subsequent losses then give rise to no payments. Payments from Trader A to Trader B are made quarterly in arrears at a rate of 0.203% per year with this being applied to the remaining tranche principal. Initially the payments are at the rate of $0.00203 \times 6{,}000{,}000 = \$12{,}180$ per year.

Note that the equity tranche is quoted differently from the other tranches. The market quote of 26.85% for CDX NA IG means that the protection seller receives an initial payment of 26.85% of the tranche principal plus 500 basis points per year on the remaining tranche principal. Similarly the market quote of 11.25% for iTraxx Europe means that the protection seller receives an initial payment of 11.25% of the tranche principal plus 500 basis points per year on the remaining tranche principal.

Table 23.6 Five-year CDX NA IG and iTraxx Europe tranches on March 28, 2007. Quotes are 30/360 in basis points except for 0%–3% tranche, where the quote indicates the percent of the tranche principal that must be paid up front in addition to 500 basis points per year.

CDX NA IG						
Tranche	0–3%	3–7%	7–10%	10–15%	15–30%	30–100%
Quote	26.85%	103.8	20.3	10.3	4.3	2.0
iTraxx Europe						
Tranche	0–3%	3–6%	6–9%	9–12%	12–22%	22–100%
Quote	11.25%	57.7	14.4	6.4	2.6	1.2

23.9 ROLE OF CORRELATION IN A BASKET CDS AND CDO

The cost of protection in a kth-to-default CDS or a tranche of a CDO is critically dependent on default correlation. Suppose that a basket of 100 reference entities is used to define a 5-year kth-to-default CDS and that each reference entity has a risk-neutral probability of 2% of defaulting during the 5 years. When the default correlation between the reference entities is zero the binomial distribution shows that the probability of one or more defaults during the 5 years is 86.74% and the probability of 10 or more defaults is 0.0034%. A first-to-default CDS is therefore quite valuable whereas a tenth-to-default CDS is worth almost nothing.

As the default correlation increases the probability of one or more defaults declines and the probability of 10 or more defaults increases. In the limit where the default correlation between the reference entities is perfect the probability of one or more defaults equals the probability of ten or more defaults and is 2%. This is because in this extreme situation the reference entities are essentially the same. Either they all default (with probability 2%) or none of them default (with probability 98%).

The valuation of a tranche of a synthetic CDO is similarly dependent on default correlation. If the correlation is low, the junior equity tranche is very risky and the senior tranches are very safe. As the default correlation increases, the junior tranches become less risky and the senior tranches become more risky. In the limit where the default correlation is perfect and the recovery rate is zero, the tranches are equally risky.

23.10 VALUATION OF A SYNTHETIC CDO

Suppose that the payment dates on a synthetic CDO tranche are at times $\tau_1, \tau_2, \dots, \tau_m$ and $\tau_0 = 0$. Define E_j as the expected tranche principal at time τ_j and $v(\tau)$ as the present value of \$1 received at time τ. Suppose that the spread on a particular tranche (i.e., the number of basis points paid for protection) is s per year. This spread is paid on the remaining tranche principal. The present value of the expected regular spread payments on the CDO is therefore given by sA, where

$$A = \sum_{j=1}^{m} (\tau_j - \tau_{j-1}) E_j v(\tau_j) \tag{23.1}$$

The expected loss between times τ_{j-1} and τ_j is $E_{j-1} - E_j$. Assume that the loss occurs at the midpoint of the time interval (i.e., at time $0.5\tau_{j-1} + 0.5\tau_j$). The present value of the expected payoffs on the CDO tranche is

$$C = \sum_{j=1}^{m} (E_{j-1} - E_j) v(0.5\tau_{j-1} + 0.5\tau_j) \tag{23.2}$$

The accrual payment due on the losses is given by sB, where

$$B = \sum_{j=1}^{m} 0.5(\tau_j - \tau_{j-1})(E_{j-1} - E_j) v(0.5\tau_{j-1} + 0.5\tau_j) \tag{23.3}$$

The value of the tranche to the protection buyer is $C - sA - sB$. The breakeven spread on the tranche occurs when the present value of the payments equals the present value of the payoffs or

$$C = sA + sB$$

The breakeven spread is therefore

$$s = \frac{C}{A + B} \tag{23.4}$$

Equations (23.1) to (23.3) show the key role played by the expected tranche principal in calculating the breakeven spread for a tranche. If we know the expected principal for a tranche on all payment dates and we also know the zero-coupon yield curve, the breakeven tranche spread can be calculated from equations (23.1) to (23.4).

Using the Gaussian Copula Model of Time to Default

The one-factor Gaussian copula model of time to default was introduced in Section 22.9. Suppose that there are n companies in a portfolio, t_i is the time to default of the ith company, and Q_i is the unconditional cumulative probability distribution of t_i (i.e., $Q_i(t)$ is the probability that $t_i < t$). The model assumes that

$$x_i = a_i F + \sqrt{1 - a_i^2}\, Z_i \tag{23.5}$$

where $x_i = N^{-1}[Q_i(t_i)]$ and F and Z_i have independent standard normal distributions. From equation (22.11),

$$Q_i(t \mid F) = N\left(\frac{N^{-1}[Q_i(t)] - a_i F}{\sqrt{1 - a_i^2}}\right) \tag{23.6}$$

where $Q_i(t \mid F)$ is the probability of the ith company defaulting by time t conditional on the value of the factor F. Denote the probability of exactly k defaults by time t as $P(k, t)$ and the corresponding probability conditional on F as $P(k, t \mid F)$. When we fix the value of F, the default probabilities are independent. It is this key point that makes it possible to calculate A, B, and C in equations (23.1) to (23.3) relatively easily.

In the standard market model, it is assumed that the time-to-default distribution Q_i and the parameter a_i are the same for all companies in the portfolio. This means that we can write $a_i = a$, $Q_i(t) = Q(t)$, and $Q_i(t \mid F) = Q(t \mid F)$. The standard market model is therefore

$$x_i = aF + \sqrt{1 - a^2} Z_i \tag{23.7}$$

where $x_i = N^{-1}[Q(t_i)]$ and F and Z_i are independent normal distributions. Equation (23.6) becomes

$$Q(t \mid F) = N\left(\frac{N^{-1}[Q(t)] - \sqrt{\rho} F}{\sqrt{1 - \rho}}\right) \tag{23.8}$$

where ρ, the copula correlation, equals a^2. This is equivalent to equation (22.12). In the calculation of $Q(t)$, it is usually assumed that the default intensity (hazard rate) for a company is constant and consistent with the index spread. The default intensity that is assumed can be calculated by using the CDS valuation approach in Section 23.2 and searching for the default intensity that gives the index spread. Suppose that this default intensity is λ. Then, from equation (22.1),

$$Q(t) = 1 - e^{-\lambda t} \tag{23.9}$$

From the properties of the binomial distribution, the standard market model gives

$$P(k, t \mid F) = \frac{n!}{(n-k)!\,k!} Q(t \mid F)^k [1 - Q(t \mid F)]^{n-k} \tag{23.10}$$

Suppose that the tranche under consideration covers losses on the portfolio between α_L and α_H. The parameter α_L is known as the *attachment point* and the parameter α_H is known as the *detachment point*. Define

$$n_L = \frac{\alpha_L n}{1-R} \quad \text{and} \quad n_H = \frac{\alpha_H n}{1-R}$$

where R is the recovery rate. Also, define $m(x)$ as the smallest integer greater than x. Without loss of generality, we assume that the initial tranche principal is 1. The tranche principal stays 1 while the number of defaults, k, is less than $m(n_L)$. It is zero when the number of defaults is greater than or equal to $m(n_H)$. Otherwise, the tranche principal is

$$\frac{\alpha_H - k(1-R)/n}{\alpha_H - \alpha_L}$$

Define $E_j(F)$ as the expected tranche principal at time τ_j conditional on the value of the factor F. It follows that

$$E_j(F) = \sum_{k=0}^{m(n_L)-1} P(k, \tau_j \mid F) + \sum_{k=m(n_L)}^{m(n_H)-1} P(k, \tau_j \mid F) \frac{\alpha_H - k(1-R)/n}{\alpha_H - \alpha_L} \tag{23.11}$$

Define $A(F)$, $B(F)$, and $C(F)$ as the values of A, B, and C conditional on F. Similarly to equations (23.1) to (23.3),

$$A(F) = \sum_{j=1}^{m} (\tau_j - \tau_{j-1}) E_j(F) v(\tau_j) \tag{23.12}$$

$$B(F) = \sum_{j=1}^{m} 0.5(\tau_j - \tau_{j-1})(E_{j-1}(F) - E_j(F)) v(0.5\tau_{j-1} + 0.5\tau_j) \tag{23.13}$$

$$C(F) = \sum_{j=1}^{m} (E_{j-1}(F) - E_j(F)) v(0.5\tau_{j-1} + 0.5\tau_j) \tag{23.14}$$

The variable F has a standard normal distribution. To calculate the unconditional values of A, B, and C, it is necessary to integrate $A(F)$, $B(F)$, and $C(F)$ over a standard normal distribution. Once the unconditional values have been calculated, the breakeven spread on the tranche can be calculated as $C/(A+B)$.[9]

The integration is best accomplished with a procedure known as Gaussian quadrature.

[9] In the case of the equity tranche, the quote is the upfront payment that must be made in addition to 500 basis points per year. The breakeven upfront payment is $C - 0.05(A+B)$.

It involves the following approximation:

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-F^2/2} g(F)\, dF \approx \sum_{k=1}^{k=M} w_k g(F_k) \tag{23.15}$$

As M increases, accuracy increases. The values of w_k and F_k for different values of M are given on the author's website.[10] A relatively large value of M is necessary to value senior tranches. Usually $M = 60$ gives sufficient accuracy.

Example 23.2

Consider the mezzanine tranche of iTraxx Europe when the copula correlation is 0.15 and the recovery rate is 40%. In this case, $\alpha_L = 0.03$, $\alpha_H = 0.06$, $n = 125$, $n_L = 6.25$, and $n_H = 12.5$. We suppose that the term structure of interest rates is flat at 3.5%, payments are made quarterly, and the CDS spread on the index is 50 basis points. A calculation similar to that in Section 23.2 shows that the constant hazard rate corresponding to the CDS spread is 0.83%. An extract from the remaining calculations is shown in Table 23.7. A value of $M = 60$ is used in equation (23.15). The factor values, F_k, and their weights, w_k, are shown in first segment of the table. The expected tranche principals on payment dates conditional on the factor values are calculated from equations (23.8) to (23.11) and shown in the second segment of the table. The values of A, B, and C conditional on the factor values are calculated in the last three segments of the table using equations (23.12) to (23.14). The unconditional values of A, B, and C are calculated by integrating $A(F)$, $B(F)$, and $C(F)$ over the probability distribution of F. This is done by setting $g(F)$ equal in turn to $A(F)$, $B(F)$, and $C(F)$ in equation (23.15). The result is $A = 0.1496$, $B = 4.2846$, and $C = 0.0187$. The breakeven tranche spread is $0.1496/(4.2846 + 0.0187) = 0.0348$, or 348 basis points. This is much higher than the spread of 57.7 basis points for this tranche in Table 23.6. This is largely because we assumed a spread for the index of 50 basis points and the index spread on March 28, 2007, was only 24 basis points.

Valuation of *k*th-to-Default CDS

A kth-to-default CDS (see Section 23.5) can also be valued using the standard market model by conditioning on the factor F. The conditional probability that the kth default happens between times τ_{j-1} and τ_j is the probability that there are k or more defaults by time τ_j minus the probability that there are k or more defaults by time τ_{j-1}. This can be calculated from equations (23.8) to (23.10) as

$$\sum_{q=k}^{n} P(q, \tau_j \mid F) - \sum_{q=k}^{n} P(q, \tau_{j-1} \mid F)$$

Defaults between time τ_{j-1} and τ_j can be assumed to happen at time $0.5\tau_{j-1} + 0.5\tau_j$. This allows the present value of payments and of payoffs, conditional on F, to be calculated in the same way as for regular CDS payoffs (see Section 23.2). By integrating over F, the unconditional present values of payments and payoffs can be calculated.

[10] The parameters w_k and F_k are calculated from the roots of Hermite polynomials. For more information on Gaussian quadrature, see Technical Note 21 on the author's website.

Table 23.7 Valuation of CDO in Example 23.2: principal = 1; payments are per unit of spread.

Weights and values for factors								
w_k	...	...	0.1579	0.1579	0.1342	0.0969	...	...
F_k	...	...	0.2020	−0.2020	−0.6060	−1.0104	...	...
Expected principal, $E_j(F_k)$								
Time								
$j = 1$	...	...	1.0000	1.0000	1.0000	1.0000	...	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
$j = 19$	...	...	0.9953	0.9687	0.8636	0.6134	...	...
$j = 20$	...	...	0.9936	0.9600	0.8364	0.5648	...	...
PV expected payoff, $A(F_k)$								
$j = 1$	...	...	0.0000	0.0000	0.0000	0.0000	...	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
$j = 19$	...	...	0.0011	0.0062	0.0211	0.0412	...	...
$j = 20$	...	...	0.0014	0.0074	0.0230	0.0410	...	...
Total	...	...	0.0055	0.0346	0.1423	0.3823	...	...
PV expected payment, $B(F_k)$								
$j = 1$	...	...	0.2457	0.2457	0.2457	0.2457	...	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
$j = 19$	...	...	0.2107	0.2051	0.1828	0.1299	...	...
$j = 20$	...	...	0.2085	0.2015	0.1755	0.1185	...	...
Total	...	...	4.5624	4.5345	4.4080	4.0361	...	...
PV expected accrual payment, $C(F_k)$								
$j = 1$	...	...	0.0000	0.0000	0.0000	0.0000	...	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
$j = 19$	...	...	0.0001	0.0008	0.0026	0.0051	...	...
$j = 20$	...	...	0.0002	0.0009	0.0029	0.0051	...	...
Total	...	...	0.0007	0.0043	0.0178	0.0478	...	...

Example 23.3

Consider a portfolio consisting of 10 bonds each with the default probabilities in Table 23.1 and suppose we are interested in valuing a third-to-default CDS where payments are made annually in arrears. Assume that the copula correlation is 0.3, the recovery rate is 40%, and all risk-free rates are 5%. As in Table 23.7, we consider 60 different factor values. The unconditional cumulative probability of each bond defaulting by years 1, 2, 3, 4, 5 is 0.0200, 0.0396, 0.0588, 0.0776, 0.0961, respectively. Equation (23.8) shows that, conditional on $F = -1.0104$, these default probabilities are 0.0365, 0.0754, 0.1134, 0.1498, 0.1848, respectively. From the binomial distribution, the conditional probability of three or more defaults by times 1, 2, 3, 4, 5 years is 0.0048, 0.0344, 0.0950, 0.1794, 0.2767, respectively. The conditional probability of the third default happening during years 1, 2, 3, 4, 5 is therefore 0.0048, 0.296, 0.0606, 0.0844, 0.0974, respectively. An analysis similar to that in Section 23.2 shows that the present values of payoffs, regular payments, and accrual payments conditional on $F = -1.0104$ are 0.1405, $3.8344s$, and $0.1171s$, where s is the spread. Similar calculations are carried out for the other 59 factor values and equation (23.15) is used to integrate over F. The unconditional present values of payoffs, regular payments, and accrual payments are 0.0637, $4.0543s$, and $0.0531s$. The breakeven CDS spread is therefore

$$0.0637/(4.0543 + 0.0531) = 0.0155$$

or 155 basis points.

Implied Correlation

In the standard market model, the recovery rate R is usually assumed to be 40%. This leaves only a as an unknown parameter in the model. Equivalently, the copula correlation $\rho = a^2$ is the only unknown parameter. This makes the model similar to Black–Scholes where there is only one unknown parameter, the volatility. Market participants like to imply a correlation from the market quotes for tranches in the same way that they imply a volatility from the market prices of options.

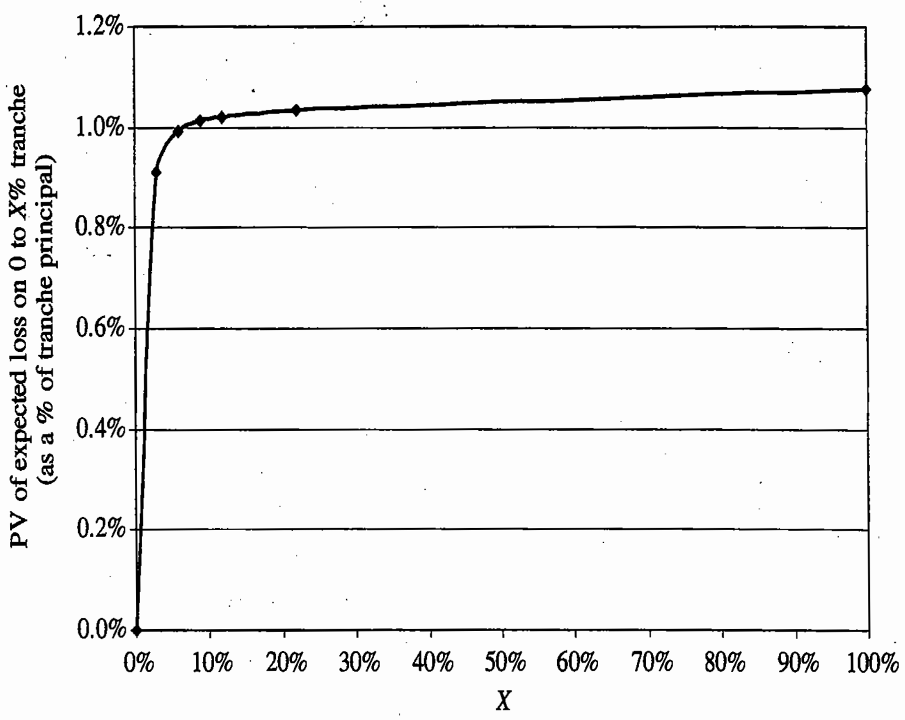
Suppose that the values of $\{\alpha_L, \alpha_H\}$ for successively more senior tranches are $\{\alpha_0, \alpha_1\}$, $\{\alpha_1, \alpha_2\}$, $\{\alpha_2, \alpha_3\}, \dots$, with $\alpha_0 = 0$. (For example, in the case of iTraxx Europe, $\alpha_0 = 0$, $\alpha_1 = 0.03$, $\alpha_2 = 0.06$, $\alpha_3 = 0.09$, $\alpha_4 = 0.12$, $\alpha_5 = 0.22$, $\alpha_6 = 1.00$.) There are two alternative implied correlations measures. One is *compound correlation*. For a tranche $\{\alpha_{q-1}, \alpha_q\}$, this is the value of the correlation, ρ, that leads to the spread calculated from the model being the same as the spread in the market. It is found using an iterative search. The other is *base correlation*. For a particular value of α_q ($q \geqslant 1$), this is the value of ρ that leads to the $\{0, \alpha_q\}$ tranche being priced consistently with the market. It is obtained using the following steps:

1. Calculate the compound correlation for each tranche.
2. Use the compound correlation to calculate the present value of the expected loss on each tranche during the life of the CDO as a percent of the initial tranche principal. This is the variable we have defined as C above. Suppose that the value of C for the α_{q-1} to α_q tranche is C_q.
3. Calculate the present value of the expected loss on the $\{0, \alpha_q\}$ tranche as a percent of the total principal of the underlying portfolio. This is $\sum_{p=1}^{q} C_p(\alpha_p - \alpha_{p-1})$.
4. The C-value for the $\{0, \alpha_q\}$ tranche is the value calculated in step 3 divided by α_q. The base correlation is the value of the correlation parameter, ρ, that is consistent with this C-value. It is found using an iterative search.

The present value of the loss as a percent of underlying portfolio that would be calculated in step 3 for the iTraxx Europe quotes in Table 23.6 are shown in Figure 23.5. The implied correlations for these quotes are shown in Table 23.8. The correlation patterns in the table are typical of those usually observed. The compound correlations exhibit a "correlation smile". As the tranche becomes more senior, the implied correlation first decreases and then increases. The base correlations exhibit a correlation skew where the implied correlation is an increasing function of the tranche detachment point.

If market prices were consistent with the one-factor Gaussian copula model, then the implied correlations (both compound and base) would be the same for all tranches.

Figure 23.5 Present value of expected loss on 0 to X% tranche as a percent of total underlying principal for iTraxx Europe on March 28, 2007.



From the pronounced smiles and skews that are observed in practice, we can infer that market prices are not consistent with this model.

Valuing Nonstandard Tranches

We do not need a model to value the standard tranches of a standard portfolio such as iTraxx Europe because the spreads for these tranches can be observed in the market. Sometimes quotes need to be produced for nonstandard tranches of a standard portfolio. Suppose that you need a quote for the 4–8% iTraxx Europe tranche. One approach is to interpolate base correlations so as to estimate the base correlation for the 0–4% tranche and the 0–8% tranche. These two base correlations allow the present value of expected loss (as a percent of the underlying portfolio principal) to be estimated for these tranches. The present value of the expected loss for the 4–8%

Table 23.8 Implied correlations for 5-year iTraxx Europe tranches on March 28, 2007.

Compound correlations					
Tranche	0–3%	3–6%	6–9%	9–12%	12–22%
Quote	18.3%	9.3%	14.3%	18.2%	24.1%
Base correlations					
Tranche	0–3%	0–6%	0–9%	0–12%	0–22%
Quote	18.3%	27.3%	34.9%	41.4%	58.1%

tranche (as a percent of the underlying principal) can be estimated as the difference between the present value of expected losses for the 0–8% and 0–4% tranches. This can be used to imply a compound correlation and a breakeven spread for the tranche.

It is now recognized that this is not the best way to proceed. A better approach is to calculate expected losses for each of the standard tranches and produce a chart such as Figure 23.5 showing the variation of expected loss for the 0–X% tranche with X. Values on this chart can be interpolated to give the expected loss for the 0–4% and the 0–8% tranches. The difference between these expected losses is a better estimate of the expected loss on the 4–8% tranche than that obtained from the base correlation approach.

It can be shown that for no arbitrage the expected losses in Figure 23.5 must increase at a decreasing rate. If base correlations are interpolated and then used to calculate expected losses, this no-arbitrage condition is often not satisfied. (The problem here is that the base correlation for the 0–X% tranche is a nonlinear function of the expected loss on the 0–X% tranche.) The direct approach of interpolating expected losses is therefore much better than the indirect approach of interpolating base correlations. What is more, it can be done so as to ensure that the no-arbitrage condition just mentioned is satisfied.

23.11 ALTERNATIVES TO THE STANDARD MARKET MODEL

This section outlines a number of alternatives to the one-factor Gaussian copula model that has become the market standard.

Heterogeneous Model

The standard market model is a homogeneous model in the sense that the time-to-default probability distributions are assumed to be the same for all companies and the copula correlations for any pair of companies are the same. The homogeneity assumption can be relaxed so that the more general model in equations (23.5) and (23.6) is used. However, this model is more complicated to implement because each company has a different probability of defaulting by any given time and $P(k, t \mid F)$ can no longer be calculated using the binomial formula in equation (23.10). It is necessary to use a numerical procedure such as that described in Andersen *et al.* (2003) and Hull and White (2004).[11]

Other Copulas

The one-factor Gaussian copula model in equations (23.7) and (23.8), or the more general model in equations (23.5) and (23.6), is a particular model of the correlation between times to default. Many other one-factor copula models have been proposed. These include the Student t copula, the Clayton copula, Archimedean copula, and Marshall–Olkin copula. We can also create new one-factor copulas by assuming that F and the Z_i in equation (23.7) have nonnormal distributions with mean 0 and standard deviation 1. Suppose that the cumulative probability distribution of Z_i is Θ and the

[11] See L. Andersen, J. Sidenius, and S. Basu, "All Your Hedges in One Basket," *Risk*, November 2003; and J. C. Hull and A. White, "Valuation of a CDO and *n*th-to-Default Swap without Monte Carlo Simulation," *Journal of Derivatives*, 12, 2 (Winter 2004), 8–23.

cumulative probability distribution of x_i (which might have to be calculated numerically from the distributions of F and Z_i) is Φ. Then $x_i = \Phi^{-1}[Q(t_i)]$ and equation (23.8) becomes

$$Q(t \mid F) = \Theta\left(\frac{\Phi^{-1}[Q(t)] - aF}{\sqrt{1-a^2}}\right)$$

Hull and White show that a good fit to the market is obtained when F and the Z_i have Student t distributions with four degrees of freedom.[12] They refer to this as the *double t copula.*

Multiple Factors

If, instead of a single factor F, there are two factors, F_1 and F_2, the model in equation (23.7) becomes

$$x_i = a_1 F_1 + a_2 F_2 + \sqrt{1 - a_1^2 - a_2^2}\, Z_i$$

and equation (23.8) becomes

$$Q(t \mid F_1, F_2) = N\left(\frac{N^{-1}[Q(t)] - a_1 F_1 - a_2 F_2}{\sqrt{1 - a_1^2 - a_2^2}}\right)$$

This model is slower to implement than the standard market model because it is necessary to integrate over two normal distributions instead of one. The model can similarly be extended to three or more factors, but the computation time increases exponentially with the number of factors.

Random Factor Loadings

Andersen and Sidenius have suggested an alternative to the model in equation (23.7) where[13]

$$x_i = a(F)F + \sqrt{1 - a(F)}\, Z_i$$

This differs from the standard market model in that the factor loading a is a function of F. In general, a increases as F decreases. This means that in states of the world where the default rate is high (i.e., states of the world where F is low) the default correlation is also high. There is empirical evidence suggesting that this is the case.[14] Andersen and Sidenius find that this model fits market quotes much better than the standard market model.

---

[12] See J.C. Hull and A. White, "Valuation of a CDO and *n*th-to-Default Swap without Monte Carlo Simulation," *Journal of Derivatives*, 12, 2 (Winter 2004), 8–23.

[13] See L. Andersen and J. Sidenius, "Extension of the Gaussian Copula Model: Random Recovery and Random Factor Loadings," *Journal of Credit Risk*, 1, 1 (Winter 2004), 29–70.

[14] See, for example, A. Sevigny and O. Renault, "Default Correlation: Empirical Evidence," Working Paper, Standard and Poors (2002); S.R. Das, L. Freed, G. Geng, and N. Kapadia, "Correlated Default Risk," *Journal of Fixed Income*, 16 (2006), 2, 7–32, J.C. Hull, M. Predescu, and A. White, "The Valuation of Correlation-Dependent Credit Derivatives Using a Structural Model," Working Paper, University of Toronto, 2005; and A. Ang and J. Chen, "Asymmetric Correlation of Equity Portfolios," *Journal of Financial Economics*, 63 (2002), 443–494.

The Implied Copula Model

Hull and White show how a copula can be implied from market quotes.[15] The simplest version of the model assumes that a certain average hazard rate applies to all companies in a portfolio over the life of a CDO. That average hazard rate has a probability distribution that can be implied from the pricing of tranches. The procedure for implying the probability distribution is to specify a number of alternative hazard rates and then search for probabilities to apply to the hazard rates so that each tranche of the CDO and the index are priced correctly. (The probabilities must sum to 1 and be nonnegative.) A smoothness condition is used to choose among the alternative solutions. The calculation of the implied copula is similar in concept to the idea, discussed in Chapter 18, of calculating an implied probability distribution for a stock price from option prices.

Dynamic Models

The models discussed so far can be characterized as static models. In essence they model the average default environment over the life of the CDO. The model constructed for a 5-year CDO is different from that constructed for a 7-year CDO, which is in turn different from that constructed for a 10-year CDO. Dynamic models are different from static models in that they attempt to model the evolution of the loss on a portfolio through time. There are three different types of dynamic models:

1. *Structural Models*: These are similar to the models described in Section 22.6 except that the stochastic processes for the asset prices of many companies are modeled simultaneously. When the asset price for a company reaches a barrier, there is a default. The processes followed by the assets are correlated. The problem with these types of models is that they have to be implemented with Monte Carlo simulation and calibration is therefore difficult.
2. *Reduced Form Models*: In these models the hazard rates of companies are modeled. In order to build in a realistic amount of correlation, it is necessary to assume that there are jumps in the hazard rates.
3. *Top Down Models*: These are models where the total loss on a portfolio is modeled directly. The models do not consider what happens to individual companies.

SUMMARY

Credit derivatives enable banks and other financial institutions to actively manage their credit risks. They can be used to transfer credit risk from one company to another and to diversify credit risk by swapping one type of exposure for another.

The most common credit derivative is a credit default swap. This is a contract where one company buys insurance against another company defaulting on its obligations. The payoff is usually the difference between the face value of a bond issued by the second company and its value immediately after a default. Credit default swaps can be analyzed by calculating the present value of the expected payments and the present value of the expected payoff in a risk-neutral world.

[15] See J. C. Hull and A. White, "Valuing Credit Derivatives Using an Implied Copula Approach," *Journal of Derivatives*, 14 (2006), 8–28.

A forward credit default swap is an obligation to enter into a particular credit default swap on a particular date. A credit default swap option is the right to enter into a particular credit default swap on a particular date. Both instruments cease to exist if the reference entity defaults before the date. A *k*th-to-default CDS is defined as a CDS that pays off when the *k*th default occurs in a portfolio of companies.

A total return swap is an instrument where the total return on a portfolio of credit-sensitive assets is exchanged for LIBOR plus a spread. Total return swaps are often used as financing vehicles. A company wanting to purchase a portfolio of assets will approach a financial institution to buy the assets on its behalf. The financial institution then enters into a total return swap with the company where it pays the return on the assets to the company and receives LIBOR plus a spread. The advantage of this type of arrangement is that the financial institution reduces its exposure to defaults by the company.

In a collateralized debt obligation a number of different securities are created from a portfolio of corporate bonds or commercial loans. There are rules for determining how credit losses are allocated. The result of the rules is that securities with both very high and very low credit ratings are created from the portfolio. A synthetic collateralized debt obligation creates a similar set of securities from credit default swaps. The standard market model for pricing both a *k*th-to-default CDS and tranches of a CDO is the one-factor Gaussian copula model for time to default. Traders use the model to imply correlations from market quotes.

FURTHER READING

Andersen, L., and J. Sidenius, "Extensions to the Gaussian Copula: Random Recovery and Random Factor Loadings," *Journal of Credit Risk*, 1, No. 1 (Winter 2004): 29–70.

Andersen, L., J. Sidenius, and S. Basu, "All Your Hedges in One Basket," *Risk*, November 2003.

Das, S., *Credit Derivatives: Trading & Management of Credit & Default Risk*, 3rd edn. New York: Wiley, 2005.

Hull, J.C., and A. White, "Valuation of a CDO and *n*th to Default Swap without Monte Carlo Simulation," *Journal of Derivatives*, 12, No. 2 (Winter 2004): 8–23.

Hull, J.C., and A. White, "Valuing Credit Derivatives Using an Implied Copula Approach," *Journal of Derivatives*, 14, 2 (Winter 2006), 8–28.

Laurent, J.-P., and J. Gregory, "Basket Default Swaps, CDOs and Factor Copulas," *Journal of Risk*, 7, 4 (2005), 8–23.

Li, D.X., "On Default Correlation: A Copula Approach," *Journal of Fixed Income*, March 2000: 43–54.

Schönbucher, P.J., *Credit Derivatives Pricing Models*. New York: Wiley, 2003.

Tavakoli, J.M., *Credit Derivatives & Synthetic Structures: A Guide to Instruments and Applications*, 2nd edn. New York: Wiley, 1998.

Questions and Problems (Answers in Solutions Manual)

23.1. Explain the difference between a regular credit default swap and a binary credit default swap.

23.2. A credit default swap requires a semiannual payment at the rate of 60 basis points per year. The principal is $300 million and the credit default swap is settled in cash. A default occurs after 4 years and 2 months, and the calculation agent estimates that the price of the cheapest deliverable bond is 40% of its face value shortly after the default. List the cash flows and their timing for the seller of the credit default swap.

23.3. Explain the two ways a credit default swap can be settled.

23.4. Explain how a cash CDO and a synthetic CDO are created.

23.5. Explain what a first-to-default credit default swap is. Does its value increase or decrease as the default correlation between the companies in the basket increases? Explain.

23.6. Explain the difference between risk-neutral and real-world default probabilities.

23.7. Explain why a total return swap can be useful as a financing tool.

23.8. Suppose that the risk-free zero curve is flat at 7% per annum with continuous compounding and that defaults can occur halfway through each year in a new 5-year credit default swap. Suppose that the recovery rate is 30% and the default probabilities each year conditional on no earlier default is 3%. Estimate the credit default swap spread. Assume payments are made annually.

23.9. What is the value of the swap in Problem 23.8 per dollar of notional principal to the protection buyer if the credit default swap spread is 150 basis points?

23.10. What is the credit default swap spread in Problem 23.8 if it is a binary CDS?

23.11. How does a 5-year *n*th-to-default credit default swap work? Consider a basket of 100 reference entities where each reference entity has a probability of defaulting in each year of 1%. As the default correlation between the reference entities increases what would you expect to happen to the value of the swap when (a) $n = 1$ and (b) $n = 25$. Explain your answer.

23.12. What is the formula relating the payoff on a CDS to the notional principal and the recovery rate?

23.13. Show that the spread for a new plain vanilla CDS should be $(1 - R)$ times the spread for a similar new binary CDS, where R is the recovery rate.

23.14. Verify that if the CDS spread for the example in Tables 23.1 to 23.4 is 100 basis points and the probability of default in a year (conditional on no earlier default) must be 1.61%. How does the probability of default change when the recovery rate is 20% instead of 40%? Verify that your answer is consistent with the implied probability of default being approximately proportional to $1/(1 - R)$, where R is the recovery rate.

23.15. A company enters into a total return swap where it receives the return on a corporate bond paying a coupon of 5% and pays LIBOR. Explain the difference between this and a regular swap where 5% is exchanged for LIBOR.

23.16. Explain how forward contracts and options on credit default swaps are structured.

23.17. "The position of a buyer of a credit default swap is similar to the position of someone who is long a risk-free bond and short a corporate bond." Explain this statement.

23.18. Why is there a potential asymmetric information problem in credit default swaps?

23.19. Does valuing a CDS using real-world default probabilities rather than risk-neutral default probabilities overstate or understate its value? Explain your answer.

23.20. What is the difference between a total return swap and an asset swap?

23.21. Suppose that in a one-factor Gaussian copula model the 5-year probability of default for each of 125 names is 3% and the pairwise copula correlation is 0.2. Calculate, for factor values of −2, −1, 0, 1, and 2: (a) the default probability conditional on the factor value and (b) the probability of more than 10 defaults conditional on the factor value.

23.22. Explain the difference between base correlation and compound correlation.

23.23. In the ABS CDO structure in Figure 23.4, suppose that there is a 12% loss on each portfolio. What is the percentage loss experienced by each of the six tranches shown.

23.24. In Example 23.2, what is the tranche spread for the 9% to 12% tranche?

Assignment Questions

23.25. Suppose that the risk-free zero curve is flat at 6% per annum with continuous compounding and that defaults can occur at times 0.25 years, 0.75 years, 1.25 years, and 1.75 years in a 2-year plain vanilla credit default swap with semiannual payments. Suppose that the recovery rate is 20% and the unconditional probabilities of default (as seen at time zero) are 1% at times 0.25 years and 0.75 years, and 1.5% at times 1.25 years and 1.75 years. What is the credit default swap spread? What would the credit default spread be if the instrument were a binary credit default swap?

23.26. Assume that the default probability for a company in a year, conditional on no earlier defaults is λ and the recovery rate is R. The risk-free interest rate is 5% per annum. Default always occurs halfway through a year. The spread for a 5-year plain vanilla CDS where payments are made annually is 120 basis points and the spread for a 5-year binary CDS where payments are made annually is 160 basis points. Estimate R and λ.

23.27. Explain how you would expect the returns offered on the various tranches in a synthetic CDO to change when the correlation between the bonds in the portfolio increases.

23.28. Suppose that:

(a) The yield on a 5-year risk-free bond is 7%.

(b) The yield on a 5-year corporate bond issued by company X is 9.5%.

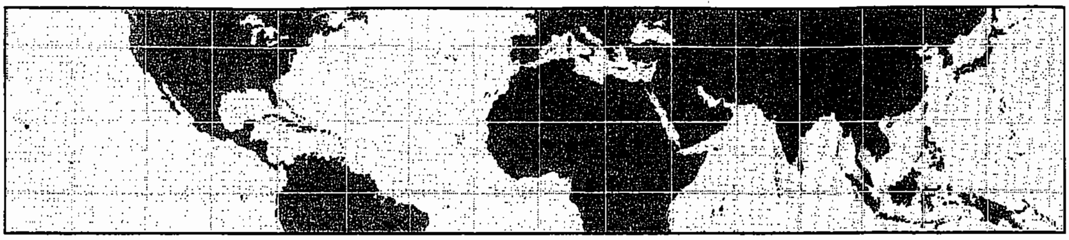
(c) A 5-year credit default swap providing insurance against company X defaulting costs 150 basis points per year.

What arbitrage opportunity is there in this situation? What arbitrage opportunity would there be if the credit default spread were 300 basis points instead of 150 basis points? Give two reasons why arbitrage opportunities such as those you identify are less than perfect.

23.29. In the ABS CDO structure in Figure 23.4, suppose that there is a 20% loss on each portfolio. What is the percentage loss experienced by each of the six tranches shown?

23.30. In Example 23.3, what is the spread for (a) a first-to-default CDS and (b) a second-to default CDS?

23.31. In Example 23.2, what is the tranche spread for the 6% to 9% tranche?



CHAPTER 24

Exotic Options

Derivatives such as European and American call and put options are what are termed *plain vanilla products*. They have standard well-defined properties and trade actively. Their prices or implied volatilities are quoted by exchanges or by brokers on a regular basis. One of the exciting aspects of the over-the-counter derivatives market is the number of nonstandard products that have been created by financial engineers. These products are termed *exotic options*, or simply *exotics*. Although they are usually a relatively small part of its portfolio, these exotics are important to a derivatives dealer because they are generally much more profitable than plain vanilla products.

Exotic products are developed for a number of reasons. Sometimes they meet a genuine hedging need in the market; sometimes there are tax, accounting, legal, or regulatory reasons why corporate treasurers, fund managers, and financial institutions find exotic products attractive; sometimes the products are designed to reflect a view on potential future movements in particular market variables; occasionally an exotic product is designed by an investment bank to appear more attractive than it is to an unwary corporate treasurer or fund manager.

In this chapter we describe different types of exotic options and discuss their valuation. We use a categorization of exotic options similar to that in an excellent series of articles written by Eric Reiner and Mark Rubinstein for *Risk* magazine in 1991 and 1992. We assume that the asset provides a yield at rate q. As discussed in Chapters 15 and 16, for an option on a stock index q should be set equal to the dividend yield on the index, for an option on a currency it should be set equal to the foreign risk-free rate, and for an option on a futures contract it should be set equal to the domestic risk-free rate. Most of the options discussed in this chapter can be valued using the DerivaGem software.

24.1 PACKAGES

A *package* is a portfolio consisting of standard European calls, standard European puts, forward contracts, cash, and the underlying asset itself. We discussed a number of different types of packages in Chapter 10: bull spreads, bear spreads, butterfly spreads, calendar spreads, straddles, strangles, and so on.

Often a package is structured by traders so that it has zero cost initially. An example is a *range forward contract.*[1] This was discussed in Section 15.2. It consists of a long call and a short put or a short call and a long put. The call strike price is greater than the put strike price and the strike prices are chosen so that the value of the call equals the value of the put.

It is worth noting that any derivative can be converted into a zero-cost product by deferring payment until maturity. Consider a European call option. If c is the cost of the option when payment is made at time zero, then $A = ce^{rT}$ is the cost when payment is made at time T, the maturity of the option. The payoff is then $\max(S_T - K, 0) - A$ or $\max(S_T - K - A, -A)$. When the strike price, K, equals the forward price, other names for a deferred payment option are break forward, Boston option, forward with optional exit, and cancelable forward.

24.2 NONSTANDARD AMERICAN OPTIONS

In a standard American option, exercise can take place at any time during the life of the option and the exercise price is always the same. The American options that are traded in the over-the-counter market sometimes have nonstandard features. For example:

1. Early exercise may be restricted to certain dates. The instrument is then known as a *Bermudan option.* (Bermuda is between Europe and America!)
2. Early exercise may be allowed during only part of the life of the option. For example, there may be an initial "lock out" period with no early exercise.
3. The strike price may change during the life of the option.

The warrants issued by corporations on their own stock often have some or all of these features. For example, in a 7-year warrant, exercise might be possible on particular dates during years 3 to 7, with the strike price being \$30 during years 3 and 4, \$32 during the next 2 years, and \$33 during the final year.

Nonstandard American options can usually be valued using a binomial tree. At each node, the test (if any) for early exercise is adjusted to reflect the terms of the option.

24.3 FORWARD START OPTIONS

Forward start options are options that will start at some time in the future. Sometimes employee stock options, which were discussed in Chapter 14, can be viewed as forward start options. This is because the company commits (implicitly or explicitly) to granting at-the-money options to employees in the future.

Consider a forward start at-the-money European call option that will start at time T_1 and mature at time T_2. Suppose that the asset price is S_0 at time zero and S_1 at time T_1. To value the option, we note from the European option pricing formulas in Chapters 13 and 15 that the value of an at-the-money call option on an asset is proportional to the asset price. The value of the forward start option at time T_1 is therefore cS_1/S_0, where c

[1] Other names used for a range forward contract are zero-cost collar, flexible forward, cylinder option, option fence, min–max, and forward band.

is the value at time zero of an at-the-money option that lasts for $T_2 - T_1$. Using risk-neutral valuation, the value of the forward start option at time zero is

$$e^{-rT_1}\hat{E}\left[c\frac{S_1}{S_0}\right]$$

where $\hat{E}$ denotes the expected value in a risk-neutral world. Since c and S_0 are known and $\hat{E}[S_1] = S_0e^{(r-q)T_1}$, the value of the forward start option is ce^{-qT_1}. For a non-dividend-paying stock, $q = 0$ and the value of the forward start option is exactly the same as the value of a regular at-the-money option with the same life as the forward start option.

24.4 COMPOUND OPTIONS

Compound options are options on options. There are four main types of compound options: a call on a call, a put on a call, a call on a put, and a put on a put. Compound options have two strike prices and two exercise dates. Consider, for example, a call on a call. On the first exercise date, T_1, the holder of the compound option is entitled to pay the first strike price, K_1, and receive a call option. The call option gives the holder the right to buy the underlying asset for the second strike price, K_2, on the second exercise date, T_2. The compound option will be exercised on the first exercise date only if the value of the option on that date is greater than the first strike price.

When the usual geometric Brownian motion assumption is made, European-style compound options can be valued analytically in terms of integrals of the bivariate normal distribution.[2] With our usual notation, the value at time zero of a European call option on a call option is

$$S_0e^{-qT_2}M(a_1, b_1; \sqrt{T_1/T_2}) - K_2e^{-rT_2}M(a_2, b_2; \sqrt{T_1/T_2}) - e^{-rT_1}K_1N(a_2)$$

where

$$a_1 = \frac{\ln(S_0/S^*) + (r - q + \sigma^2/2)T_1}{\sigma\sqrt{T_1}}, \qquad a_2 = a_1 - \sigma\sqrt{T_1}$$

$$b_1 = \frac{\ln(S_0/K_2) + (r - q + \sigma^2/2)T_2}{\sigma\sqrt{T_2}}, \qquad b_2 = b_1 - \sigma\sqrt{T_2}$$

The function $M(a, b : \rho)$ is the cumulative bivariate normal distribution function that the first variable will be less than a and the second will be less than b when the coefficient of correlation between the two is ρ.[3] The variable S^* is the asset price at time T_1 for which the option price at time T_1 equals K_1. If the actual asset price is above S^* at time T_1, the first option will be exercised; if it is not above S^*, the option expires worthless.

With similar notation, the value of a European put on a call is

$$K_2e^{-rT_2}M(-a_2, b_2; -\sqrt{T_1/T_2}) - S_0e^{-qT_2}M(-a_1, b_1; -\sqrt{T_1/T_2}) + e^{-rT_1}K_1N(-a_2)$$

[2] See R. Geske, "The Valuation of Compound Options," *Journal of Financial Economics*, 7 (1979): 63–81; M. Rubinstein, "Double Trouble," *Risk*, December 1991/January 1992: 53–56.

[3] See Technical Note 5 on the author's website for a numerical procedure for calculating M. A function for calculating M is also on the website.

The value of a European call on a put is

$$K_2e^{-rT_2}M(-a_2, -b_2; \sqrt{T_1/T_2}) - S_0e^{-qT_2}M(-a_1, -b_1; \sqrt{T_1/T_2}) - e^{-rT_1}K_1N(-a_2)$$

The value of a European put on a put is

$$S_0e^{-qT_2}M(a_1, -b_1; -\sqrt{T_1/T_2}) - K_2e^{-rT_2}M(a_2, -b_2; -\sqrt{T_1/T_2}) + e^{-rT_1}K_1N(a_2)$$

24.5 CHOOSER OPTIONS

A *chooser* option (sometimes referred to as an *as you like it* option) has the feature that, after a specified period of time, the holder can choose whether the option is a call or a put. Suppose that the time when the choice is made is T_1. The value of the chooser option at this time is

$$\max(c, p)$$

where c is the value of the call underlying the option and p is the value of the put underlying the option.

If the options underlying the chooser option are both European and have the same strike price, put–call parity can be used to provide a valuation formula. Suppose that S_1 is the asset price at time T_1, K is the strike price, T_2 is the maturity of the options, and r is the risk-free interest rate. Put–call parity implies that

$$\begin{aligned}\max(c, p) &= \max(c,\ c + Ke^{-r(T_2-T_1)} - S_1e^{-q(T_2-T_1)}) \\ &= c + e^{-q(T_2-T_1)}\max(0,\ Ke^{-(r-q)(T_2-T_1)} - S_1)\end{aligned}$$

This shows that the chooser option is a package consisting of:

1. A call option with strike price K and maturity T_2
2. $e^{-q(T_2-T_1)}$ put options with strike price $Ke^{-(r-q)(T_2-T_1)}$ and maturity T_1

As such, it can readily be valued.

More complex chooser options can be defined where the call and the put do not have the same strike price and time to maturity. They are then not packages and have features that are somewhat similar to compound options.

24.6 BARRIER OPTIONS

Barrier options are options where the payoff depends on whether the underlying asset's price reaches a certain level during a certain period of time.

A number of different types of barrier options regularly trade in the over-the-counter market. They are attractive to some market participants because they are less expensive than the corresponding regular options. These barrier options can be classified as either *knock-out options* or *knock-in options*. A knock-out option ceases to exist when the underlying asset price reaches a certain barrier; a knock-in option comes into existence only when the underlying asset price reaches a barrier.

Equations (15.4) and (15.5) show that the values at time zero of a regular call and put option are

$$c = S_0 e^{-qT} N(d_1) - K e^{-rT} N(d_2)$$
$$p = K e^{-rT} N(-d_2) - S_0 e^{-qT} N(-d_1)$$

where

$$d_1 = \frac{\ln(S_0/K) + (r - q + \sigma^2/2)T}{\sigma\sqrt{T}}$$
$$d_2 = \frac{\ln(S_0/K) + (r - q - \sigma^2/2)T}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T}$$

A *down-and-out call* is one type of knock-out option. It is a regular call option that ceases to exist if the asset price reaches a certain barrier level H. The barrier level is below the initial asset price. The corresponding knock-in option is a *down-and-in call*. This is a regular call that comes into existence only if the asset price reaches the barrier level.

If H is less than or equal to the strike price, K, the value of a down-and-in call at time zero is

$$c_{\text{di}} = S_0 e^{-qT}(H/S_0)^{2\lambda} N(y) - K e^{-rT}(H/S_0)^{2\lambda-2} N(y - \sigma\sqrt{T})$$

where

$$\lambda = \frac{r - q + \sigma^2/2}{\sigma^2}$$
$$y = \frac{\ln[H^2/(S_0 K)]}{\sigma\sqrt{T}} + \lambda\sigma\sqrt{T}$$

Because the value of a regular call equals the value of a down-and-in call plus the value of a down-and-out call, the value of a down-and-out call is given by

$$c_{\text{do}} = c - c_{\text{di}}$$

If $H \geqslant K$, then

$$c_{\text{do}} = S_0 N(x_1) e^{-qT} - K e^{-rT} N(x_1 - \sigma\sqrt{T}) - S_0 e^{-qT}(H/S_0)^{2\lambda} N(y_1) + K e^{-rT}(H/S_0)^{2\lambda-2} N(y_1 - \sigma\sqrt{T})$$

and

$$c_{\text{di}} = c - c_{\text{do}}$$

where

$$x_1 = \frac{\ln(S_0/H)}{\sigma\sqrt{T}} + \lambda\sigma\sqrt{T}$$
$$y_1 = \frac{\ln(H/S_0)}{\sigma\sqrt{T}} + \lambda\sigma\sqrt{T}$$

An *up-and-out call* is a regular call option that ceases to exist if the asset price reaches a barrier level, H, that is higher than the current asset price. An *up-and-in call* is a regular call option that comes into existence only if the barrier is reached. When H is less than or equal to K, the value of the up-and-out call, c_{uo}, is zero and the value of the up-and-in call, c_{ui}, is c. When H is greater than K,

$$c_{\text{ui}} = S_0 N(x_1) e^{-qT} - K e^{-rT} N(x_1 - \sigma\sqrt{T}) - S_0 e^{-qT}(H/S_0)^{2\lambda}[N(-y) - N(-y_1)] + K e^{-rT}(H/S_0)^{2\lambda-2}[N(-y + \sigma\sqrt{T}) - N(-y_1 + \sigma\sqrt{T})]$$

and

$$c_{uo} = c - c_{ui}$$

Put barrier options are defined similarly to call barrier options. An *up-and-out put* is a put option that ceases to exist when a barrier, H, that is greater than the current asset price is reached. An *up-and-in put* is a put that comes into existence only if the barrier is reached. When the barrier, H, is greater than or equal to the strike price, K, their prices are

$$p_{ui} = -S_0 e^{-qT}(H/S_0)^{2\lambda}N(-y) + Ke^{-rT}(H/S_0)^{2\lambda-2}N(-y+\sigma\sqrt{T})$$

and

$$p_{uo} = p - p_{ui}$$

When H is less than or equal to K,

$$p_{uo} = -S_0 N(-x_1)e^{-qT} + Ke^{-rT}N(-x_1+\sigma\sqrt{T}) + S_0 e^{-qT}(H/S_0)^{2\lambda}N(-y_1) - Ke^{-rT}(H/S_0)^{2\lambda-2}N(-y_1+\sigma\sqrt{T})$$

and

$$p_{ui} = p - p_{uo}$$

A *down-and-out put* is a put option that ceases to exist when a barrier less than the current asset price is reached. A *down-and-in put* is a put option that comes into existence only when the barrier is reached. When the barrier is greater than the strike price, $p_{do} = 0$ and $p_{di} = p$. When the barrier is less than the strike price,

$$p_{di} = -S_0 N(-x_1)e^{-qT} + Ke^{-rT}N(-x_1+\sigma\sqrt{T}) + S_0 e^{-qT}(H/S_0)^{2\lambda}[N(y) - N(y_1)] - Ke^{-rT}(H/S_0)^{2\lambda-2}[N(y-\sigma\sqrt{T}) - N(y_1-\sigma\sqrt{T})]$$

and

$$p_{do} = p - p_{di}$$

All of these valuations make the usual assumption that the probability distribution for the asset price at a future time is lognormal. An important issue for barrier options is the frequency that the asset price, S, is observed for purposes of determining whether the barrier has been reached. The analytic formulas given in this section assume that S is observed continuously and sometimes this is the case.[4] Often, the terms of a contract state that S is observed periodically; for example, once a day at 12 noon. Broadie, Glasserman, and Kou provide a way of adjusting the formulas we have just given for the situation where the price of the underlying is observed discretely.[5] The barrier level H is replaced by $He^{0.5826\sigma\sqrt{T/m}}$ for an up-and-in or up-and-out option and by $He^{-0.5826\sigma\sqrt{T/m}}$ for a down-and-in or down-and-out option, where m is the number of times the asset price is observed (so that T/m is the time interval between observations).

Barrier options often have quite different properties from regular options. For example, sometimes vega is negative. Consider an up-and-out call option when the asset price is close to the barrier level. As volatility increases, the probability that the

[4] One way to track whether a barrier has been reached from below (above) is to send a limit order to the exchange to sell (buy) the asset at the barrier price and see whether the order is filled.

[5] M. Broadie, P. Glasserman, and S. G. Kou, "A Continuity Correction for Discrete Barrier Options," *Mathematical Finance* 7, 4 (October 1997): 325–49.

barrier will be hit increases. As a result, a volatility increase can cause the price of the barrier option to decrease in these circumstances.

24.7 BINARY OPTIONS

Binary options are options with discontinuous payoffs. A simple example of a binary option is a *cash-or-nothing call.* This pays off nothing if the asset price ends up below the strike price at time T and pays a fixed amount, Q, if it ends up above the strike price. In a risk-neutral world, the probability of the asset price being above the strike price at the maturity of an option is, with our usual notation, $N(d_2)$. The value of a cash-or-nothing call is therefore $Qe^{-rT}N(d_2)$. A *cash-or-nothing put* is defined analogously to a cash-or-nothing call. It pays off Q if the asset price is below the strike price and nothing if it is above the strike price. The value of a cash-or-nothing put is $Qe^{-rT}N(-d_2)$.

Another type of binary option is an *asset-or-nothing call.* This pays off nothing if the underlying asset price ends up below the strike price and pays the asset price if it ends up above the strike price. With our usual notation, the value of an asset-or-nothing call is $S_0e^{-qT}N(d_1)$. An *asset-or-nothing put* pays off nothing if the underlying asset price ends up above the strike price and the asset price if it ends up below the strike price. The value of an asset-or-nothing put is $S_0e^{-qT}N(-d_1)$.

A regular European call option is equivalent to a long position in an asset-or-nothing call and a short position in a cash-or-nothing call where the cash payoff in the cash-or-nothing call equals the strike price. Similarly, a regular European put option is equivalent to a long position in a cash-or-nothing put and a short position in an asset-or-nothing put where the cash payoff on the cash-or-nothing put equals the strike price.

24.8 LOOKBACK OPTIONS

The payoffs from lookback options depend on the maximum or minimum asset price reached during the life of the option. The payoff from a *floating lookback call* is the amount that the final asset price exceeds the minimum asset price achieved during the life of the option. The payoff from a *floating lookback put* is the amount by which the maximum asset price achieved during the life of the option exceeds the final asset price.

Valuation formulas have been produced for floating lookbacks.[6] The value of a floating lookback call at time zero is

$$c_{\text{fl}} = S_0e^{-qT}N(a_1) - S_0e^{-qT}\frac{\sigma^2}{2(r-q)}N(-a_1) - S_{\text{min}}e^{-rT}\left[N(a_2) - \frac{\sigma^2}{2(r-q)}e^{Y_1}N(-a_3)\right]$$

[6] See B. Goldman, H. Sosin, and M. A. Gatto, "Path-Dependent Options: Buy at the Low, Sell at the High," *Journal of Finance*, 34 (December 1979): 1111–27.; M. Garman, "Recollection in Tranquility," *Risk*, March (1989): 16–19.

where

$$a_1 = \frac{\ln(S_0/S_{\min}) + (r - q + \sigma^2/2)T}{\sigma\sqrt{T}}$$

$$a_2 = a_1 - \sigma\sqrt{T},$$

$$a_3 = \frac{\ln(S_0/S_{\min}) + (-r + q + \sigma^2/2)T}{\sigma\sqrt{T}}$$

$$Y_1 = -\frac{2(r - q - \sigma^2/2)\ln(S_0/S_{\min})}{\sigma^2}$$

and $S_{\min}$ is the minimum asset price achieved to date. (If the lookback has just been originated, $S_{\min} = S_0$.) See Problem 24.23 for the $r = q$ case.

The value of a floating lookback put is

$$p_{\mathrm{fl}} = S_{\max}e^{-rT}\left[N(b_1) - \frac{\sigma^2}{2(r-q)}e^{Y_2}N(-b_3)\right] + S_0e^{-qT}\frac{\sigma^2}{2(r-q)}N(-b_2) - S_0e^{-qT}N(b_2)$$

where

$$b_1 = \frac{\ln(S_{\max}/S_0) + (-r + q + \sigma^2/2)T}{\sigma\sqrt{T}}$$

$$b_2 = b_1 - \sigma\sqrt{T}$$

$$b_3 = \frac{\ln(S_{\max}/S_0) + (r - q - \sigma^2/2)T}{\sigma\sqrt{T}}$$

$$Y_2 = \frac{2(r - q - \sigma^2/2)\ln(S_{\max}/S_0)}{\sigma^2}$$

and $S_{\max}$ is the maximum asset price achieved to date. (If the lookback has just been originated, then $S_{\max} = S_0$.)

A floating lookback call is a way that the holder can buy the underlying asset at the lowest price achieved during the life of the option. Similarly, a floating lookback put is a way that the holder can sell the underlying asset at the highest price achieved during the life of the option.

Example 24.1

Consider a newly issued floating lookback put on a non-dividend-paying stock where the stock price is 50, the stock price volatility is 40% per annum, the risk-free rate is 10% per annum, and the time to maturity is 3 months. In this case, $S_{\max} = 50$, $S_0 = 50$, $r = 0.1$, $q = 0$, $\sigma = 0.4$, and $T = 0.25$, $b_1 = -0.025$, $b_2 = -0.225$, $b_3 = 0.025$, and $Y_2 = 0$, so that the value of the lookback put is 7.79. A newly issued floating lookback call on the same stock is worth 8.04.

In a fixed lookback option, a strike price is specified. For a *fixed lookback call option*, the payoff is the same as a regular European call option except that the final asset price is replaced by the maximum asset price achieved during the life of the option. For a *fixed lookback put option*, the payoff is the same as a regular European put option except that the the final asset price is replaced by the minimum asset price achieved during the life of the option. Define $S^*_{\max} = \max(S_{\max}, K)$, where $S_{\max}$ is the maximum asset price achieved so far during the life of the option and K is the strike price. Also, define p^*_{fl}

as the value of a floating lookback put which lasts for the same period as the fixed lookback call when the actual maximum asset price so far, $S_{\max}$, is replaced by $S^*_{\max}$. A put–call parity type of argument shows that the value of the fixed lookback call option, c_{fix} is given by[7]

$$c_{\text{fix}} = p^*_{\text{fl}} + S_0 e^{-qT} - Ke^{-rT}$$

Similarly, if $S^*_{\min} = \min(S_{\min}, K)$, then the value of a fixed lookback put option, p_{fix}, is given by

$$p_{\text{fix}} = c^*_{\text{fl}} + Ke^{-rT} - S_0 e^{-qT}$$

where c^*_{fl} is the value of a floating lookback call that lasts for the same period as the fixed lookback put when the actual minimum asset price so far, $S_{\min}$, is replaced by $S^*_{\min}$. This shows that the equations given above for floating lookbacks can be modified to price fixed lookbacks.

Lookbacks are appealing to investors, but very expensive when compared with regular options. As with barrier options, the value of a lookback option is liable to be sensitive to the frequency that the asset price is observed for the purposes of computing the maximum or minimum. The formulas above assume that the asset price is observed continuously. Broadie, Glasserman, and Kou provide a way of adjusting the formulas we have just given for the situation where the asset price is observed discretely.[8]

24.9 SHOUT OPTIONS

A *shout option* is a European option where the holder can "shout" to the writer at one time during its life. At the end of the life of the option, the option holder receives either the usual payoff from a European option or the intrinsic value at the time of the shout, whichever is greater. Suppose the strike price is \$50 and the holder of a call shouts when the price of the underlying asset is \$60. If the final asset price is less than \$60, the holder receives a payoff of \$10. If it is greater than \$60, the holder receives the excess of the asset price over \$50.

A shout option has some of the same features as a lookback option, but is considerably less expensive. It can be valued by noting that if the holder shouts at a time τ when the asset price is S_τ the payoff from the option is

$$\max(0, S_T - S_\tau) + (S_\tau - K)$$

where, as usual, K is the strike price and S_T is the asset price at time T. The value at time τ if the holder shouts is therefore the present value of $S_\tau - K$ (received at time T) plus the value of a European option with strike price S_τ. The latter can be calculated using Black–Scholes formulas.

A shout option is valued by constructing a binomial or trinomial tree for the underlying asset in the usual way. Working back through the tree, the value of the

[7] The argument was proposed by H. Y. Wong and Y. K. Kwok, "Sub-replication and Replenishing Premium: Efficient Pricing of Multi-state Lookbacks," *Review of Derivatives Research*, 6 (2003), 83–106.

[8] M. Broadie, P. Glasserman, and S. G. Kou, "Connecting Discrete and Continuous Path-Dependent Options," *Finance and Stochastics*, 2 (1998): 1–28.

option if the holder shouts and the value if the holder does not shout can be calculated at each node. The option's price at the node is the greater of the two. The procedure for valuing a shout option is therefore similar to the procedure for valuing a regular American option.

24.10 ASIAN OPTIONS

Asian options are options where the payoff depends on the average price of the underlying asset during at least some part of the life of the option. The payoff from an *average price call* is $\max(0, S_{\text{ave}} - K)$ and that from an *average price put* is $\max(0, K - S_{\text{ave}})$, where S_{ave} is the average value of the underlying asset calculated over a predetermined averaging period. Average price options are less expensive than regular options and are arguably more appropriate than regular options for meeting some of the needs of corporate treasurers. Suppose that a US corporate treasurer expects to receive a cash flow of 100 million Australian dollars spread evenly over the next year from the company's Australian subsidiary. The treasurer is likely to be interested in an option that guarantees that the average exchange rate realized during the year is above some level. An average price put option can achieve this more effectively than regular put options.

Another type of Asian option is an average strike option. An *average strike call* pays off $\max(0, S_T - S_{\text{ave}})$ and an *average strike put* pays off $\max(0, S_{\text{ave}} - S_T)$. Average strike options can guarantee that the average price paid for an asset in frequent trading over a period of time is not greater than the final price. Alternatively, it can guarantee that the average price received for an asset in frequent trading over a period of time is not less than the final price.

If the underlying asset price, S, is assumed to be lognormally distributed and S_{ave} is a geometric average of the S's, analytic formulas are available for valuing European average price options.[9] This is because the geometric average of a set of lognormally distributed variables is also lognormal. Consider a newly issued option that will provide a payoff at time T based on the geometric average calculated between time zero and time T. In a risk-neutral world, it can be shown that the probability distribution of the geometric average of a asset price over a certain period is the same as that of the asset price at the end of the period if the asset's expected growth rate is set equal to $(r - q - \sigma^2/6)/2$ (rather than $r - q$) and its volatility is set equal to $\sigma/\sqrt{3}$ (rather than σ). The geometric average price option can, therefore, be treated like a regular option with the volatility set equal to $\sigma/\sqrt{3}$ and the dividend yield equal to

$$r - \tfrac{1}{2}\left(r - q - \frac{\sigma^2}{6}\right) = \tfrac{1}{2}\left(r + q + \frac{\sigma^2}{6}\right)$$

When, as is nearly always the case, Asian options are defined in terms of arithmetic averages, exact analytic pricing formulas are not available. This is because the distribution of the arithmetic average of a set of lognormal distributions does not have analytically tractable properties. However, the distribution is approximately lognormal

[9] See A. Kemna and A. Vorst, "A Pricing Method for Options Based on Average Asset Values," *Journal of Banking and Finance*, 14 (March 1990): 113–29.

and this leads to a good analytic approximation for valuing average price options. Analysts calculate the first two moments of the probability distribution of the arithmetic average in a risk-neutral world exactly and then fit a lognormal distribution to the moments.[10]

Consider a newly issued Asian option that provides a payoff at time T based on the arithmetic average between time zero and time T. The first moment, M_1, and the second moment, M_2, of the average in a risk-neutral world can be shown to be

$$M_1 = \frac{e^{(r-q)T} - 1}{(r-q)T} S_0$$

and

$$M_2 = \frac{2e^{[2(r-q)+\sigma^2]T} S_0^2}{(r-q+\sigma^2)(2r-2q+\sigma^2)T^2} + \frac{2S_0^2}{(r-q)T^2}\left[\frac{1}{2(r-q)+\sigma^2} - \frac{e^{(r-q)T}}{r-q+\sigma^2}\right]$$

when $q \neq r$ (see Problem 24.23 for the $q = r$ case).

By assuming that the average asset price is lognormal, an analyst can use Black's model. In equations (16.9) and (16.10),

$$F_0 = M_1 \tag{24.1}$$

and

$$\sigma^2 = \frac{1}{T}\ln\left(\frac{M_2}{M_1^2}\right) \tag{24.2}$$

Example 24.2

Consider a newly issued average price call option on a non-dividend-paying stock where the stock price is 50, the strike price is 50, the stock price volatility is 40% per annum, the risk-free rate is 10% per annum, and the time to maturity is 1 year. In this case, $S_0 = 50$, $K = 50$, $r = 0.1$, $q = 0$, $\sigma = 0.4$, and $T = 1$. If the average is a geometric average, the option can be valued as a regular option with the volatility equal to $0.4/\sqrt{3}$, or 23.09%, and dividend yield equal to $(0.1 + 0.4^2/6)/2$, or 6.33%. The value of the option is 5.13. If the average is an arithmetic average, we first calculate $M_1 = 52.59$ and $M_2 = 2{,}922.76$. From equations (24.1) and (24.2), $F_0 = 52.59$ and $\sigma = 23.54\%$. Equation (16.9), with $K = 50$, $T = 1$, and $r = 0.1$, gives the value of the option as 5.62.

The formulas just given for M_1 and M_2 assume that the average is calculated from continuous observations on the asset price. The appendix to this chapter shows how M_1 and M_2 can be obtained when the average is calculated from observations on the asset price at discrete points in time.

We can modify the analysis to accommodate the situation where the option is not newly issued and some prices used to determine the average have already been observed. Suppose that the averaging period is composed of a period of length t_1 over which prices have already been observed and a future period of length t_2 (the remaining life of the option). Suppose that the average asset price during the first time period is $\bar{S}$. The

[10] See S. M. Turnbull and L. M. Wakeman, "A Quick Algorithm for Pricing European Average Options," *Journal of Financial and Quantitative Analysis*, 26 (September 1991): 377–89.

payoff from an average price call is

$$\max\left(\frac{\bar{S}t_1 + S_{\text{ave}}t_2}{t_1 + t_2} - K, 0\right)$$

where S_{ave} is the average asset price during the remaining part of the averaging period. This is the same as

$$\frac{t_2}{t_1 + t_2}\max(S_{\text{ave}} - K^*, 0)$$

where

$$K^* = \frac{t_1 + t_2}{t_2}K - \frac{t_1}{t_2}\bar{S}$$

When $K^* > 0$, the option can be valued in the same way as a newly issued Asian option provided that we change the strike price from K to K^* and multiply the result by $t_2/(t_1 + t_2)$. When $K^* < 0$ the option is certain to be exercised and can be valued as a forward contract. The value is

$$\frac{t_2}{t_1 + t_2}[M_1 e^{-rt_2} - K^* e^{-rt_2}]$$

24.11 OPTIONS TO EXCHANGE ONE ASSET FOR ANOTHER

Options to exchange one asset for another (sometimes referred to as *exchange options*) arise in various contexts. An option to buy yen with Australian dollars is, from the point of view of a US investor, an option to exchange one foreign currency asset for another foreign currency asset. A stock tender offer is an option to exchange shares in one stock for shares in another stock.

Consider a European option to give up an asset worth U_T at time T and receive in return an asset worth V_T. The payoff from the option is

$$\max(V_T - U_T, 0)$$

A formula for valuing this option was first produced by Margrabe.[11] Suppose that the asset prices, U and V, both follow geometric Brownian motion with volatilities σ_U and σ_V. Suppose further that the instantaneous correlation between U and V is ρ, and the yields provided by U and V are q_U and q_V, respectively. The value of the option at time zero is

$$V_0 e^{-q_V T} N(d_1) - U_0 e^{-q_U T} N(d_2) \tag{24.3}$$

where

$$d_1 = \frac{\ln(V_0/U_0) + (q_U - q_V + \hat{\sigma}^2/2)T}{\hat{\sigma}\sqrt{T}}, \qquad d_2 = d_1 - \hat{\sigma}\sqrt{T}$$

and

$$\hat{\sigma} = \sqrt{\sigma_U^2 + \sigma_V^2 - 2\rho\sigma_U\sigma_V}$$

and U_0 and V_0 are the values of U and V at times zero.

[11] See W. Margrabe, "The Value of an Option to Exchange One Asset for Another," *Journal of Finance*, 33 (March 1978): 177–86.

This result will be proved in Chapter 27. It is interesting to note that equation (24.3) is independent of the risk-free rate r. This is because, as r increases, the growth rate of both asset prices in a risk-neutral world increases, but this is exactly offset by an increase in the discount rate. The variable $\hat{\sigma}$ is the volatility of V/U. Comparisons with equation (16.4) show that the option price is the same as the price of U_0 European call options on an asset worth V/U when the strike price is 1.0, the risk-free interest rate is q_U, and the dividend yield on the asset is q_V. Mark Rubinstein shows that the American version of this option can be characterized similarly for valuation purposes.[12] It can be regarded as U_0 American options to buy an asset worth V/U for 1.0 when the risk-free interest rate is q_U and the dividend yield on the asset is q_V. The option can therefore be valued as described in Chapter 19 using a binomial tree.

An option to obtain the better or worse of two assets can be regarded as a position in one of the assets combined with an option to exchange it for the other asset:

$$\min(U_T, V_T) = V_T - \max(V_T - U_T, 0)$$

$$\max(U_T, V_T) = U_T + \max(V_T - U_T, 0)$$

24.12 OPTIONS INVOLVING SEVERAL ASSETS

Options involving two or more risky assets are sometimes referred to as *rainbow options.* One example is the bond futures contract traded on the CBOT described in Chapter 6. The party with the short position is allowed to choose between a large number of different bonds when making delivery.

Probably the most popular option involving several assets is a *basket option.* This is an option where the payoff is dependent on the value of a portfolio (or basket) of assets. The assets are usually either individual stocks or stock indices or currencies. A European basket option can be valued with Monte Carlo simulation, by assuming that the assets follow correlated geometric Brownian motion processes. A much faster approach is to calculate the first two moments of the basket at the maturity of the option in a risk-neutral world, and then assume that value of the basket is lognormally distributed at that time. The option can then be valued using Black's model with the parameters shown in equations (24.1) and (24.2). The appendix to this chapter shows how the moments of the value of the basket at a future time can be calculated from the volatilities of, and correlations between, the assets. Correlations are typically estimated from historical data.

24.13 VOLATILITY AND VARIANCE SWAPS

A volatility swap is an agreement to exchange the realized volatility of an asset between time 0 and time T for a prespecifed fixed volatility. The realized volatility is calculated as described in Section 13.4 but with the assumption that the mean daily return is zero. Suppose that there are n daily observations on the asset price during the period between

[12] See M. Rubinstein, "One for Another," *Risk*, July/August 1991: 30–32

time 0 and time T. The realized volatility is

$$\bar{\sigma} = \sqrt{\frac{252}{n-2}\sum_{i=1}^{n-1}\left[\ln\left(\frac{S_{i+1}}{S_i}\right)\right]^2}$$

where S_i is the ith observation on the asset price.

The payoff from the volatility swap at time T to the payer of the fixed volatility is $L_{\text{vol}}(\bar{\sigma} - \sigma_K)$, where L_{vol} is the notional principal and σ_K is the fixed volatility. Whereas an option provides a complex exposure to the asset price and volatility, a volatility swap is simpler in that it has exposure only to volatility.

A variance swap is an agreement to exchange the realized variance rate $\bar{V}$ between time 0 and time T for a prespecified variance rate. The variance rate is the square of the volatility ($\bar{V} = \bar{\sigma}^2$). Variance swaps are easier to value than volatility swaps. This is because the variance rate between time 0 and time T can be replicated using a portfolio of put and call options. The payoff from a variance swap at time T to the payer of the fixed variance rate is $L_{\text{var}}(\bar{V} - V_K)$, where L_{var} is the notional principal and V_K is the fixed variance rate. Often the notional principal for a variance swap is expressed in terms of the corresponding notional principal for a volatility swap using $L_{\text{var}} = L_{\text{vol}}/(2\sigma_K)$.

Valuation of Variance Swap

Technical Note 22 on the author's website shows that, for any value S^* of the asset price, the expected value of the average variance between time 0 and time T is given by

$$\hat{E}(\bar{V}) = \frac{2}{T}\ln\frac{F_0}{S^*} - \frac{2}{T}\left[\frac{F_0}{S^*} - 1\right] + \frac{2}{T}\left[\int_{K=0}^{S^*}\frac{1}{K^2}e^{rT}p(K)\,dK + \int_{K=S^*}^{\infty}\frac{1}{K^2}e^{rT}c(K)\,dK\right] \quad \textbf{(24.4)}$$

where F_0 is the forward price of the asset for a contract maturing at time T, $c(K)$ is the price of a European call option with strike price K and time to maturity T, and $p(K)$ is the price of a European put option with strike price K and time to maturity T.

This provides a way of valuing a variance swap.[13] The value of an agreement to receive the realized variance between time 0 and time T and pay a variance rate of V_K, with both being applied to a principal of L_{var}, is

$$L_{\text{var}}[\hat{E}(\bar{V}) - V_K]e^{-rT} \quad \textbf{(24.5)}$$

Suppose that the prices of European options with strike prices K_i ($1 \leqslant i \leqslant n$) are known, where $K_1 < K_2 < \cdots < K_n$. A standard approach for implementing equation (24.4) is to set S^* equal to the first strike price below F_0 and then approximate the integrals as

$$\int_{K=0}^{S^*}\frac{1}{K^2}e^{rT}p(K)dK + \int_{K=S^*}^{\infty}\frac{1}{K^2}e^{rT}c(K)dK = \sum_{i=1}^{n}\frac{\Delta K_i}{K_i^2}e^{rT}Q(K_i) \quad \textbf{(24.6)}$$

where $\Delta K_i = 0.5(K_{i+1} - K_{i-1})$ for $2 \leqslant i \leqslant n-1$, $\Delta K_1 = K_2 - K_1$, $\Delta K_n = K_n - K_{n-1}$.

[13] See also K. Demeterfi, E. Derman, M. Kamal, and J. Zou, "More Than You Ever Wanted to Know About Volatility Swaps," *The Journal of Derivatives*, 6, 4 (Summer 1999), 9–32. For options on variance and volatility, see P. Carr and R. Lee, "Realized Volatility and Variance: Options via Swaps," *Risk*, May 2007, 76–83.

The function $Q(K_i)$ is the price of a European put option with strike price K_i if $K_i < S^*$ and the price of a European call option with strike price K_i if $K_i > S^*$. When $K_i = S^*$, the function $Q(K_i)$ is equal to the average of the prices of a European call and a European put with strike price K_i.

Example 24.3

Consider a 3-month contract to pay the realized variance rate of an index over the 3 months and receive a variance rate of 0.045 on a principal of \$100 million. The risk-free rate is 4% and the dividend yield on the index is 1%. The current level of the index is 1020. Suppose that, for strike prices of 800, 850, 900, 950, 1,000, 1,050, 1,100, 1,150, 1,200, the 3-month implied volatilities of the index are 29%, 28%, 27%, 26%, 25%, 24%, 23%, 22%, 21%, respectively. In this case, $n = 9$, $K_1 = 800$, $K_2 = 850, \ldots, K_9 = 1{,}200$, $F_0 = 1{,}020e^{(0.04-0.01)\times 0.25} = 1{,}027.68$, and $S^* = 1{,}000$. DerivaGem shows that $Q(K_1) = 2.22$, $Q(K_2) = 5.22$, $Q(K_3) = 11.05$, $Q(K_4) = 21.27$, $Q(K_5) = 51.21$, $Q(K_6) = 38.94$, $Q(K_7) = 20.69$, $Q(K_8) = 9.44$, $Q(K_9) = 3.57$. Also, $\Delta K_i = 0.5$ for all i. Hence,

$$\sum_i^n \frac{\Delta K_i}{K_i^2} e^{rT} Q(K_i) = 0.008139$$

From equations (24.4) and (24.6), it follows that

$$\hat{E}(\bar{V}) = \frac{2}{0.25}\ln\left(\frac{1027.68}{1{,}000}\right) - \frac{2}{0.25}\left(\frac{1027.68}{1{,}000} - 1\right) + \frac{2}{0.25} \times 0.008139 = 0.0621$$

From equation (24.5), the value of the variance swap (in millions of dollars) is $100 \times (0.0621 - 0.045)e^{-0.04\times 0.25} = 1.69$.

Valuation of a Volatility Swap

To value a volatility swap, we require $\hat{E}(\bar{\sigma})$, where $\bar{\sigma}$ is the average value of volatility between time 0 and time T. We can write

$$\bar{\sigma} = \sqrt{\hat{E}(\bar{V})}\sqrt{1 + \frac{\bar{V} - \hat{E}(\bar{V})}{\hat{E}(\bar{V})}}.$$

Expanding the second term on the right-hand side in a series gives

$$\bar{\sigma} = \sqrt{\hat{E}(\bar{V})}\left\{1 + \frac{\bar{V} - E(\bar{V})}{2\hat{E}(\bar{V})} - \frac{1}{8}\left[\frac{\bar{V} - \hat{E}(\bar{V})}{\hat{E}(\bar{V})}\right]^2\right\}$$

Taking expectations,

$$\hat{E}(\bar{\sigma}) = \sqrt{\hat{E}(\bar{V})}\left\{1 - \frac{1}{8}\left[\frac{\mathrm{var}(\bar{V})}{\hat{E}(\bar{V})^2}\right]\right\} \tag{24.7}$$

where $\mathrm{var}(\bar{V})$ is the variance of $\bar{V}$. The valuation of a volatility swap therefore requires an estimate of the variance of the average variance rate during the life of the contract. The value of an agreement to receive the realized volatility between time 0 and time T and pay a volatility of σ_K, with both being applied to a principal of L_{vol}, is

$$L_{\mathrm{vol}}[\hat{E}(\bar{\sigma}) - \sigma_K]e^{-RT}$$

Example 24.4

For the situation in Example 24.3, consider a volatility swap where the realized volatility is received and a volatility of 23% is paid on a principal of \$100 million. In this case $\hat{E}(\bar{V}) = 0.0621$. Suppose that the standard deviation of the average variance over 3 months has been estimated as 0.01. This means that $\mathrm{var}(\bar{V}) = 0.0001$. Equation (24.7) gives

$$\hat{E}(\bar{\sigma}) = \sqrt{0.0621}\left(1 - \tfrac{1}{8} \times \frac{0.0001}{0.0621^2}\right) = 0.2484$$

The value of the swap in (millions of dollars) is

$$100 \times (0.2484 - 0.23)e^{-0.04\times0.25} = 1.82$$

The VIX Index

In equation (24.4), the ln function can be approximated by the first two terms in a series expansion:

$$\ln\left(\frac{F_0}{S^*}\right) = \left(\frac{F_0}{S^*} - 1\right) - \frac{1}{2}\left(\frac{F_0}{S^*} - 1\right)^2$$

This means that the risk-neutral expected cumulative variance is calculated as

$$\hat{E}(\bar{V})T = -\left(\frac{F_0}{S^*} - 1\right)^2 + 2\sum_{i=1}^{n}\frac{\Delta K_i}{K_i^2}e^{RT}Q(K_i) \tag{24.8}$$

Since 2004 the VIX volatility index (see Section 13.11) has been based on equation (24.8). The procedure used on any given day is to calculate $\hat{E}(\bar{V})T$ for options that trade in the market and have maturities immediately above and below 30 days. The 30-day risk-neutral expected cumulative variance is calculated from these two numbers using interpolation. This is then multiplied by 365/30 and the index is set equal to the square root of the result. More details on the calculation can be found on:

www.cboe.com/micro/vix/vixwhite.pdf

24.14 STATIC OPTIONS REPLICATION

If the procedures described in Chapter 17 are used for hedging exotic options, some are easy to handle, but others are very difficult because of discontinuities (see Business Snapshot 24.1). For the difficult cases, a technique known as static options replication is sometimes useful.[14] This involves searching for a portfolio of actively traded options that approximately replicates the exotic option. Shorting this position provides the hedge.[15]

The basic principle underlying static options replication is as follows. If two portfolios are worth the same on a certain boundary, they are also worth the same at

[14] See E. Derman, D. Ergener, and I. Kani, "Static Options Replication," *Journal of Derivatives* 2, 4 (Summer 1995): 78–95.

[15] Technical Note 22 on the author's website provides an example of static replication. It shows that the variance rate of an asset can be replicated by a position in the asset and out-of-the-money options on the asset. This result, which leads to equation (24.4), can be used to hedge variance swaps.

Business Snapshot 24.1 Is Delta Hedging Easier or More Difficult for Exotics?

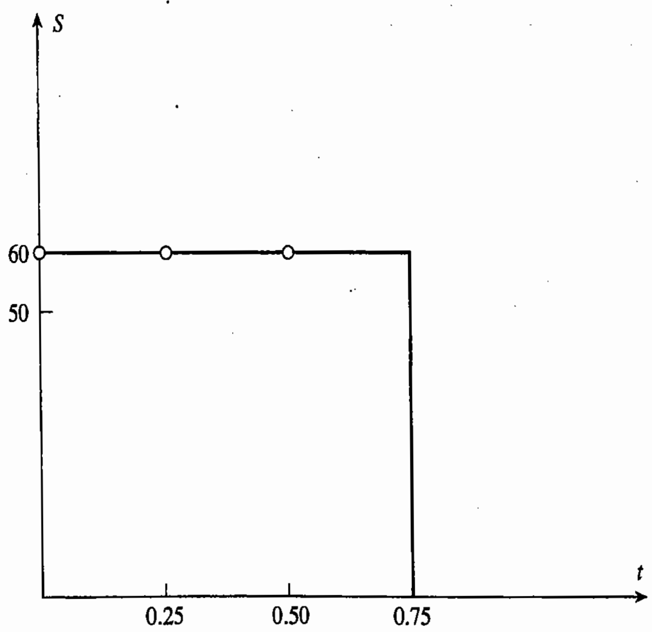
As described in Chapter 17 we can approach the hedging of exotic options by creating a delta neutral position and rebalancing frequently to maintain delta neutrality. When we do this we find some exotic options are easier to hedge than plain vanilla options and some are more difficult.

An example of an exotic option that is relatively easy to hedge is an average price option where the averaging period is the whole life of the option. As time passes, we observe more of the asset prices that will be used in calculating the final average. This means that our uncertainty about the payoff decreases with the passage of time. As a result, the option becomes progressively easier to hedge. In the final few days, the delta of the option always approaches zero because price movements during this time have very little impact on the payoff.

By contrast barrier options are relatively difficult to hedge. Consider a down-and-out call option on a currency when the exchange rate is 0.0005 above the barrier. If the barrier is hit, the option is worth nothing. If the barrier is not hit, the option may prove to be quite valuable. The delta of the option is discontinuous at the barrier making conventional hedging very difficult.

all interior points of the boundary. Consider as an example a 9-month up-and-out call option on a non-dividend-paying stock where the stock price is 50, the strike price is 50, the barrier is 60, the risk-free interest rate is 10% per annum, and the volatility is 30% per annum. Suppose that $f(S, t)$ is the value of the option at time t for a stock price of S. Any boundary in (S, t) space can be used for the purposes of producing the replicating portfolio. A convenient one to choose is shown in Figure 24.1. It is defined by $S = 60$

Figure 24.1 Boundary points used for static options replication example.



and $t = 0.75$. The values of the up-and-out option on the boundary are given by

$$f(S, 0.75) = \max(S - 50,\ 0) \quad \text{when } S < 60$$
$$f(60, t) = 0 \quad \text{when } 0 \leqslant t \leqslant 0.75$$

There are many ways that these boundary values can be approximately matched using regular options. The natural option to match the first boundary is a 9-month European call with a strike price of 50. The first component of the replicating portfolio is therefore one unit of this option. (We refer to this option as option A.)

One way of matching the $f(60, t)$ boundary is to proceed as follows:

1. Divide the life of the option into N steps of length Δt
2. Choose a European call option with a strike price of 60 and maturity at time $N\Delta t$ (= 9 months) to match the boundary at the $\{60, (N-1)\Delta t\}$ point
3. Choose a European call option with a strike price of 60 and maturity at time $(N-1)\Delta t$ to match the boundary at the $\{60, (N-2)\Delta t\}$ point

and so on. Note that the options are chosen in sequence so that they have zero value on the parts of the boundary matched by earlier options.[16] The option with a strike price of 60 that matures in 9 months has zero value on the vertical boundary that is matched by option A. The option maturing at time $i\,\Delta t$ has zero value at the point $\{60, i\,\Delta t\}$ that is matched by the option maturing at time $(i+1)\Delta t$ for $1 \leqslant i \leqslant N-1$.

Suppose that $\Delta = 0.25$. In addition to option A, the replicating portfolio consists of positions in European options with strike price 60 that mature in 9, 6, and 3 months. We will refer to these as options B, C, and D, respectively. Given our assumptions about volatility and interest rates, option B is worth 4.33 at the $\{60, 0.5\}$ point. Option A is worth 11.54 at this point. The position in option B necessary to match the boundary at the $\{60, 0.5\}$ point is therefore $-11.54/4.33 = -2.66$. Option C is worth 4.33 at the $\{60, 0.25\}$ point. The position taken in options A and B is worth -4.21 at this point. The position in option C necessary to match the boundary at the $\{60, 0.25\}$ point is therefore $4.21/4.33 = 0.97$. Similar calculations show that the position in option D necessary to match the boundary at the $\{60, 0\}$ point is 0.28.

The portfolio chosen is summarized in Table 24.1. (See also Sample Application F of the DerivaGem Application Builder.) It is worth 0.73 initially (i.e., at time zero when

Table 24.1 The portfolio of European call options used to replicate an up-and-out option.

Option	*Strike price*	*Maturity (years)*	*Position*	*Initial value*
A	50	0.75	1.00	+6.99
B	60	0.75	−2.66	−8.21
C	60	0.50	0.97	+1.78
D	60	0.25	0.28	+0.17

[16] This is not a requirement. If K points on the boundary are to be matched, we can choose K options and solve a set of K linear equations to determine required positions in the options.

the stock price is 50). This compares with 0.31 given by the analytic formula for the up-and-out call earlier in this chapter. The replicating portfolio is not exactly the same as the up-and-out option because it matches the latter at only three points on the second boundary. If we use the same procedure, but match at 18 points on the second boundary (using options that mature every half month), the value of the replicating portfolio reduces to 0.38. If 100 points are matched, the value reduces further to 0.32.

To hedge a derivative, the portfolio that replicates its boundary conditions must be shorted. The portfolio must be unwound when any part of the boundary is reached.

Static options replication has the advantage over delta hedging that it does not require frequent rebalancing. It can be used for a wide range of derivatives. The user has a great deal of flexibility in choosing the boundary that is to be matched and the options that are to be used.

SUMMARY

Exotic options are options with rules governing the payoff that are more complicated than standard options. We have discussed 12 different types of exotic options: packages, nonstandard American options, forward start options, compound options, chooser options, barrier options, binary options, lookback options, shout options, Asian options, options to exchange one asset for another, and options involving several assets. We have discussed how these can be valued using the same assumptions as those used to derive the Black–Scholes model in Chapter 13. Some can be valued analytically, but using much more complicated formulas than those for regular European calls and puts, some can be handled using analytic approximations, and some can be valued using extensions of the numerical procedures in Chapter 19. We will present more numerical procedures for valuing exotic options in Chapter 26.

Some exotic options are easier to hedge than the corresponding regular options; others are more difficult. In general, Asian options are easier to hedge because the payoff becomes progressively more certain as we approach maturity. Barrier options can be more difficult to hedge because delta is discontinuous at the barrier. One approach to hedging an exotic option, known as static options replication, is to find a portfolio of regular options whose value matches the value of the exotic option on some boundary. The exotic option is hedged by shorting this portfolio.

FURTHER READING

Carr, P., and R. Lee, "Realized Volatility and Variance: Options via Swaps," *Risk*, May 2007, 76–83.

Clewlow, L., and C. Strickland, *Exotic Options: The State of the Art*. London: Thomson Business Press, 1997.

Demeterfi, K., E. Derman, M. Kamal, and J. Zou, "More than You Ever Wanted to Know about Volatility Swaps," *Journal of Derivatives*, 6, 4 (Summer, 1999), 9–32.

Derman, E., D. Ergener, and I. Kani, "Static Options Replication," *Journal of Derivatives*, 2, 4 (Summer 1995): 78–95.

Geske, R., "The Valuation of Compound Options," *Journal of Financial Economics*, 7 (1979): 63–81.

Goldman, B., H. Sosin, and M. A. Gatto, "Path Dependent Options: Buy at the Low, Sell at the High," *Journal of Finance*, 34 (December 1979); 1111–27.

Margrabe, W., "The Value of an Option to Exchange One Asset for Another," *Journal of Finance*, 33 (March 1978): 177–86.

Milevsky, M. A., and S. E. Posner, "Asian Options: The Sum of Lognormals and the Reciprocal Gamma Distribution," *Journal of Financial and Quantitative Analysis*, 33, 3 (September 1998), 409–22.

Ritchken, P. "On Pricing Barrier Options," *Journal of Derivatives*, 3, 2 (Winter 1995): 19–28.

Ritchken P., L. Sankarasubramanian, and A. M. Vijh, "The Valuation of Path Dependent Contracts on the Average," *Management Science*, 39 (1993): 1202–13.

Rubinstein, M., and E. Reiner, "Breaking Down the Barriers," *Risk*, September (1991): 28–35.

Rubinstein, M., "Double Trouble," *Risk*, December/January (1991/1992): 53–56.

Rubinstein, M., "One for Another," *Risk*, July/August (1991): 30–32.

Rubinstein, M., "Options for the Undecided," *Risk*, April (1991): 70–73.

Rubinstein, M., "Pay Now, Choose Later," *Risk*, February (1991): 44–47.

Rubinstein, M., "Somewhere Over the Rainbow," *Risk*, November (1991): 63–66.

Rubinstein, M., "Two in One," *Risk* May (1991): 49.

Rubinstein, M., and E. Reiner, "Unscrambling the Binary Code," *Risk*, October 1991: 75–83.

Stulz, R. M., "Options on the Minimum or Maximum of Two Assets," *Journal of Financial Economics*, 10 (1982): 161–85.

Turnbull, S. M., and L. M. Wakeman, "A Quick Algorithm for Pricing European Average Options," *Journal of Financial and Quantitative Analysis*, 26 (September 1991): 377–89.

Questions and Problems (Answers in Solutions Manual)

24.1. Explain the difference between a forward start option and a chooser option.

24.2. Describe the payoff from a portfolio consisting of a lookback call and a lookback put with the same maturity.

24.3. Consider a chooser option where the holder has the right to choose between a European call and a European put at any time during a 2-year period. The maturity dates and strike prices for the calls and puts are the same regardless of when the choice is made. Is it ever optimal to make the choice before the end of the 2-year period? Explain your answer.

24.4. Suppose that c_1 and p_1 are the prices of a European average price call and a European average price put with strike price K and maturity T, c_2 and p_2 are the prices of a European average strike call and European average strike put with maturity T, and c_3 and p_3 are the prices of a regular European call and a regular European put with strike price K and maturity T. Show that $c_1 + c_2 - c_3 = p_1 + p_2 - p_3$.

24.5. The text derives a decomposition of a particular type of chooser option into a call maturing at time T_2 and a put maturing at time T_1. Derive an alternative decomposition into a call maturing at time T_1 and a put maturing at time T_2.

24.6. Section 24.6 gives two formulas for a down-and-out call. The first applies to the situation where the barrier, H, is less than or equal to the strike price, K. The second applies to the situation where $H \geqslant K$. Show that the two formulas are the same when $H = K$.

24.7. Explain why a down-and-out put is worth zero when the barrier is greater than the strike price.

24.8. Suppose that the strike price of an American call option on a non-dividend-paying stock grows at rate g. Show that if g is less than the risk-free rate, r, it is never optimal to exercise the call early.

24.9. How can the value of a forward start put option on a non-dividend-paying stock be calculated if it is agreed that the strike price will be 10% greater than the stock price at the time the option starts?

24.10. If a stock price follows geometric Brownian motion, what process does $A(t)$ follow where $A(t)$ is the arithmetic average stock price between time zero and time t?

24.11. Explain why delta hedging is easier for Asian options than for regular options.

24.12. Calculate the price of a 1-year European option to give up 100 ounces of silver in exchange for 1 ounce of gold. The current prices of gold and silver are \$380 and \$4, respectively; the risk-free interest rate is 10% per annum; the volatility of each commodity price is 20%; and the correlation between the two prices is 0.7. Ignore storage costs.

24.13. Is a European down-and-out option on an asset worth the same as a European down-and-out option on the asset's futures price for a futures contract maturing at the same time as the option?

24.14. Answer the following questions about compound options:
(a) What put–call parity relationship exists between the price of a European call on a call and a European put on a call? Show that the formulas given in the text satisfy the relationship.
(b) What put–call parity relationship exists between the price of a European call on a put and a European put on a put? Show that the formulas given in the text satisfy the relationship.

24.15. Does a floating lookback call become more valuable or less valuable as we increase the frequency with which we observe the asset price in calculating the minimum?

24.16. Does a down-and-out call become more valuable or less valuable as we increase the frequency with which we observe the asset price in determining whether the barrier has been crossed? What is the answer to the same question for a down-and-in call?

24.17. Explain why a regular European call option is the sum of a down-and-out European call and a down-and-in European call. Is the same true for American call options?

24.18. What is the value of a derivative that pays off \$100 in 6 months if the S&P 500 index is greater than 1,000 and zero otherwise? Assume that the current level of the index is 960, the risk-free rate is 8% per annum, the dividend yield on the index is 3% per annum, and the volatility of the index is 20%.

24.19. In a 3-month down-and-out call option on silver futures the strike price is \$20 per ounce and the barrier is \$18. The current futures price is \$19, the risk-free interest rate is 5%, and the volatility of silver futures is 40% per annum. Explain how the option works and calculate its value. What is the value of a regular call option on silver futures with the same terms? What is the value of a down-and-in call option on silver futures with the same terms?

24.20. A new European-style floating lookback call option on a stock index has a maturity of 9 months. The current level of the index is 400, the risk-free rate is 6% per annum, the dividend yield on the index is 4% per annum, and the volatility of the index is 20%. Use DerivaGem to value the option.

24.21. Estimate the value of a new 6-month European-style average price call option on a non-dividend-paying stock. The initial stock price is \$30, the strike price is \$30, the risk-free interest rate is 5%, and the stock price volatility is 30%.

24.22. Use DerivaGem to calculate the value of:

(a) A regular European call option on a non-dividend-paying stock where the stock price is \$50, the strike price is \$50, the risk-free rate is 5% per annum, the volatility is 30%, and the time to maturity is one year

(b) A down-and-out European call which is as in (a) with the barrier at \$45

(c) A down-and-in European call which is as in (a) with the barrier at \$45.

Show that the option in (a) is worth the sum of the values of the options in (b) and (c).

24.23. Explain adjustments that have to be made when $r = q$ for (a) the valuation formulas for lookback call options in Section 24.8 and (b) the formulas for M_1 and M_2 in Section 24.10.

24.24. Value the variance swap in Example 24.3 of Section 24.13 assuming that the implied volatilities for options with strike prices 800, 850, 900, 950, 1,000, 1,050, 1,100, 1,150, 1,200 are 20%, 20.5%, 21%, 21.5%, 22%, 22.5%, 23%, 23.5%, 24%, respectively.

Assignment Questions

24.25. What is the value in dollars of a derivative that pays off £10,000 in 1 year provided that the dollar/sterling exchange rate is greater than 1.5000 at that time? The current exchange rate is 1.4800. The dollar and sterling interest rates are 4% and 8% per annum, respectively. The volatility of the exchange rate is 12% per annum.

24.26. Consider an up-and-out barrier call option on a non-dividend-paying stock when the stock price is 50, the strike price is 50, the volatility is 30%, the risk-free rate is 5%, the time to maturity is 1 year, and the barrier at \$80. Use the software to value the option and graph the relationship between (a) the option price and the stock price, (b) the delta and the stock price, (c) the option price and the time to maturity, and (d) the option price and the volatility. Provide an intuitive explanation for the results you get. Show that the delta, gamma, theta, and vega for an up-and-out barrier call option can be either positive or negative.

24.27. Sample Application F in the DerivaGem Application Builder Software considers the static options replication example in Section 24.13. It shows the way a hedge can be constructed using four options (as in Section 24.13) and two ways a hedge can be constructed using 16 options.

(a) Explain the difference between the two ways a hedge can be constructed using 16 options. Explain intuitively why the second method works better.

(b) Improve on the four-option hedge by changing Tmat for the third and fourth options.

(c) Check how well the 16-option portfolios match the delta, gamma, and vega of the barrier option.

24.28. Consider a down-and-out call option on a foreign currency. The initial exchange rate is 0.90, the time to maturity is 2 years, the strike price is 1.00, the barrier is 0.80, the domestic risk-free interest rate is 5%, the foreign risk-free interest rate is 6%, and the volatility is 25% per annum. Use DerivaGem to develop a static option replication strategy involving five options.

24.29. Suppose that a stock index is currently 900. The dividend yield is 2%, the risk-free rate is 5%, and the volatility is 40%. Use the results in the appendix to calculate the value of a 1-year average price call where the strike price is 900 and the index level is observed at the end of each quarter for the purposes of the averaging. Compare this with the price calculated by DerivaGem for a 1-year average price option where the price is observed continuously. Provide an intuitive explanation for any differences between the prices.

24.30. Use the DerivaGem Application Builder software to compare the effectiveness of daily delta hedging for (a) the option considered in Tables 17.2 and 17.3 and (b) an average price call with the same parameters. Use Sample Application C. For the average price option you will find it necessary to change the calculation of the option price in cell C16, the payoffs in cells H15 and H16, and the deltas (cells G46 to G186 and N46 to N186). Carry out 20 Monte Carlo simulation runs for each option by repeatedly pressing F9. On each run record the cost of writing and hedging the option, the volume of trading over the whole 20 weeks and the volume of trading between weeks 11 and 20. Comment on the results.

24.31. In the DerivaGem Application Builder Software modify Sample Application D to test the effectiveness of delta and gamma hedging for a call on call compound option on a 100,000 units of a foreign currency where the exchange rate is 0.67, the domestic risk-free rate is 5%, the foreign risk-free rate is 6%, the volatility is 12%. The time to maturity of the first option is 20 weeks, and the strike price of the first option is 0.015. The second option matures 40 weeks from today and has a strike price of 0.68. Explain how you modified the cells. Comment on hedge effectiveness.

24.32. Outperformance certificates (also called "sprint certificates", "accelerator certificates", or "speeders") are offered to investors by many European banks as a way of investing in a company's stock. The initial investment equals the stock price, S_0. If the stock price goes up between time 0 and time T, the investor gains k times the increase at time T, where k is a constant greater than 1.0. However, the stock price used to calculate the gain at time T is capped at some maximum level M. If the stock price goes down, the investor's loss is equal to the decrease. The investor does not receive dividends.

(a) Show that an outperformance certificate is a package.

(b) Calculate using DerivaGem the value of a one-year outperformance certificate when the stock price is 50 euros, $k = 1.5$, $M = 70$ euros, the risk-free rate is 5%, and the stock price volatility is 25%. Dividends equal to 0.5 euros are expected in 2 months, 5 months, 8 months, and 11 months.

24.33. Carry out the analysis in Example 24.3 of Section 24.13 to value the variance swap on the assumption that the life of the swap is 1 month rather than 3 months.

APPENDIX

CALCULATION OF MOMENTS FOR VALUATION OF BASKET OPTIONS AND ASIAN OPTIONS

Consider first the problem of calculating the first two moments of the value of a basket of assets at a future time, T, in a risk-neutral world. The price of each asset in the basket is assumed to be lognormal. Define:

n: The number of assets

S_i: The value of the ith asset at time T[17]

F_i: The forward price of the ith asset for a contract maturing at time T

σ_i: The volatility of the ith asset between time zero and time T

ρ_{ij}: Correlation between returns from the ith and jth asset

P: Value of basket at time T

M_1: First moment of P in a risk-neutral world

M_2: Second moment of P in a risk-neutral world

Because $P = \sum_{i=1}^{n} S_i$, $\hat{E}(S_i) = F_i$, $M_1 = \hat{E}(P)$ and $M_2 = \hat{E}(P^2)$, where $\hat{E}$ denotes expected value in a risk-neutral world, it follows that

$$M_1 = \sum_{i=1}^{n} F_i$$

Also,

$$P^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} S_i S_j$$

From the properties of lognormal distributions,

$$\hat{E}(S_i S_j) = F_i F_j e^{\rho_{ij}\sigma_i\sigma_j T}$$

Hence

$$M_2 = \sum_{i=1}^{n} \sum_{j=1}^{n} F_i F_j e^{\rho_{ij}\sigma_i\sigma_j T}$$

Asian Options

We now move on to the related problem of calculating the first two moments of the arithmetic average price of an asset in a risk-neutral world when the average is calculated from discrete observations. Suppose that the asset price is observed at times T_i $(1 \leqslant i \leqslant m)$. Redefine variables as follows:

S_i: The value of the asset at time T_i

F_i: The forward price of the asset for a contract maturing at time T_i

σ_i: The implied volatility for an option on the asset with maturity T_i

---

[17] If the ith asset is a certain stock and there are, say, 200 shares of the stock in the basket, then (for the purposes of the first part of the appendix) the ith "asset" is defined as 200 shares of the stock and S_i is the value of 200 shares of the stock.

ρ_{ij}: Correlation between return on asset up to time T_i and the return on the asset up to time T_j

P: Value of the arithmetic average

M_1: First moment of P in a risk-neutral world

M_2: Second moment of P in a risk-neutral world

In this case,

$$M_1 = \frac{1}{m}\sum_{i=1}^{m} F_i$$

Also,

$$P^2 = \frac{1}{m^2}\sum_{i=1}^{m}\sum_{j=1}^{m} S_i S_j$$

In this case,

$$\hat{E}(S_i S_j) = F_i F_j e^{\rho_{ij}\sigma_i\sigma_j\sqrt{T_i T_j}}$$

It can be shown that, when $i < j$,

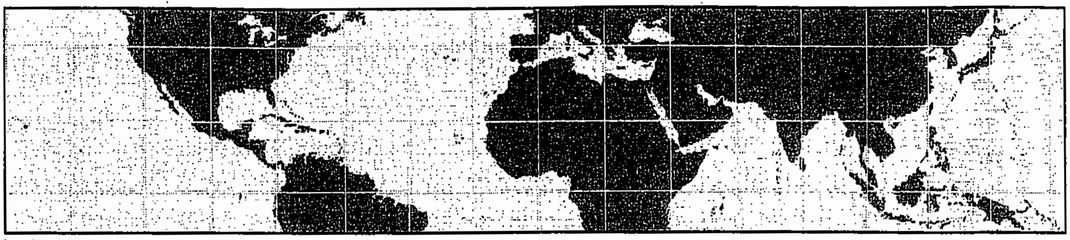
$$\rho_{ij} = \frac{\sigma_i\sqrt{T_i}}{\sigma_j\sqrt{T_j}}$$

so that

$$\hat{E}(S_i S_j) = F_i F_j e^{\sigma_i^2 T_i}$$

and

$$M_2 = \frac{1}{m^2}\left[\sum_{i=1}^{m} F_i^2 e^{\sigma_i^2 T_i} + 2\sum_{j=1}^{m}\sum_{i=1}^{j-1} F_i F_j e^{\sigma_i^2 T_i}\right]$$



CHAPTER 25

Weather, Energy, and Insurance Derivatives

The most common underlying variables in derivatives contracts are stock prices, stock indices, exchange rates, interest rates, and commodity prices. Futures, forward, option, and swap contracts on these variables have been outstandingly successful. As discussed in Chapter 23, credit derivatives have also become very popular in recent years. Chapter 24 shows that one way dealers have expanded the derivatives market is by developing nonstandard (or exotic) structures for defining payoffs. This chapter discusses another way they have expanded the market. This is by trading derivatives on new underlying variables.

The chapter examines the products that have been developed to manage weather risk, energy price risk, and insurance risks. The markets that it will talk about are in some cases in the early stages of their development. As they evolve there may well be significant changes in both the products that are offered and the ways they are used.

25.1 REVIEW OF PRICING ISSUES

Chapters 11 and 13 explained the risk-neutral valuation result. This involves pricing a derivative on the assumption that investors are risk neutral. The expected payoff is calculated in a risk-neutral world and then discounted at the risk-free interest rate. The approach gives the correct price—not just in a risk-neutral world, but in all other worlds as well.

An alternative pricing approach sometimes adopted is to use historical data to calculate the expected payoff and then discount this expected payoff at the risk free rate to obtain the price. We will refer to this as the historical data approach. Historical data give an estimate of the expected payoff in the real world. It follows that the historical data approach is correct only when the expected payoff from the derivative is the same in both the real world and the risk-neutral world.

Section 11.7 shows that when we move from the real world to the risk-neutral world, the volatilities of variables remain the same, but their expected growth rates are liable to change. For example, the expected growth rate of a stock market index decreases by perhaps 4% or 5% when we move from the real world to the risk-neutral world. The expected growth rate of a variable can be assumed to be the same in both the real world and the risk-neutral world if the variable has zero systematic risk so that percentage changes in the variable have zero correlation with stock market returns. This means that the historical data approach to valuing a derivative gives the right answer if all underlying variables have zero systematic risk. A common feature of most of the derivatives considered in this chapter is that the underlying variables can reasonably be assumed to have zero systematic risk, so that the historical data approach can be used.

25.2 WEATHER DERIVATIVES

Many companies are in the position where their performance is liable to be adversely affected by the weather.[1] It makes sense for these companies to consider hedging their weather risk in much the same way as they hedge foreign exchange or interest rate risks.

The first over-the-counter weather derivatives were introduced in 1997. To understand how they work, we explain two variables:

HDD: Heating degree days

CDD: Cooling degree days

A day's HDD is defined as

$$\text{HDD} = \max(0, 65 - A)$$

and a day's CDD is defined as

$$\text{CDD} = \max(0, A - 65)$$

where A is the average of the highest and lowest temperature during the day at a specified weather station, measured in degrees Fahrenheit. For example, if the maximum temperature during a day (midnight to midnight) is 68° Fahrenheit and the minimum temperature is 44° Fahrenheit, then $A = 56$. The daily HDD is then 9 and the daily CDD is 0.

A typical over-the-counter product is a forward or option contract providing a payoff dependent on the cumulative HDD or CDD during a month (i.e., the total of the HDDs or CDDs for every day in the month). For example, a derivatives dealer could in January 2008 sell a client a call option on the cumulative HDD during February 2009 at the Chicago O'Hare Airport weather station with a strike price of 700 and a payment rate of \$10,000 per degree day. If the actual cumulative HDD is 820, the payoff is \$1.2 million. Contracts often include a payment cap. If the payment cap in our example is \$1.5 million, the contract is the equivalent of a bull spread. The client has a long call option on cumulative HDD with a strike price of 700 and a short call option with a strike price of 850.

A day's HDD is a measure of the volume of energy required for heating during the day. A day's CDD is a measure of the volume of energy required for cooling during the day. Most weather derivative contracts are entered into by energy producers and energy consumers. But retailers, supermarket chains, food and drink manufacturers, health service companies, agricultural companies, and companies in the leisure industry are also potential users of weather derivatives. The Weather Risk Management

[1] The US Department of Energy has estimated that one-seventh of the US economy is subject to weather risk.

Association (www.wrma.org) has been formed to serve the interests of the weather risk management industry.

In September 1999 the Chicago Mercantile Exchange began trading weather futures and European options on weather futures. The contracts are on the cumulative HDD and CDD for a month observed at a weather station.[2] The contracts are settled in cash just after the end of the month once the HDD and CDD are known. One futures contract is on $100 times the cumulative HDD or CDD. The HDD and CDD are calculated by a company, Earth Satellite Corporation, using automated data-collection equipment.

The temperature at a certain location can reasonably be assumed to have zero systematic risk. It follows from Section 25.1 that weather derivatives can be priced using the historical data approach. Consider, for example, the call option on the February 2009 HDD at Chicago O'Hare airport mentioned earlier. By collecting, say, 50 years of data, a probability distribution for the HDD in February can be estimated. This in turn can be used to provide a probability distribution for the option payoff. An estimate of the value of the option is the mean of this distribution discounted at the risk-free rate. It might be desirable to adjust the probability distribution for temperature trends. For example, a linear regression might show that (perhaps because of global warming) the HDD in February is decreasing at a rate of 4 per year on average. If so, the output from the regression could be used to estimate a trend-adjusted probability distribution for the HDD in February 2009.

25.3 ENERGY DERIVATIVES

Energy companies are among the most active and sophisticated users of derivatives. Many energy products trade in both the over-the-counter market and on exchanges. In this section, the trading of crude oil, natural gas, and electricity derivatives is examined.

Crude Oil

Crude oil is one of the most important commodities in the world, with global demand amounting to about 80 million barrels daily. Ten-year fixed-price supply contracts have been commonplace in the over-the-counter market for many years. These are swaps where oil at a fixed price is exchanged for oil at a floating price.

In the 1970s the price of oil was highly volatile. The 1973 war in the Middle East led to a tripling of oil prices. The fall of the Shah of Iran in the late 1970s again increased prices. These events led oil producers and users to a realization that they needed more sophisticated tools for managing oil-price risk. In the 1980s both the over-the-counter market and the exchange-traded market developed products to meet this need.

In the over-the-counter market, virtually any derivative that is available on common stocks or stock indices is now available with oil as the underlying asset. Swaps, forward contracts, and options are popular. Contracts sometimes require settlement in cash and sometimes require settlement by physical delivery (i.e., by delivery of the oil).

[2] The CME has introduced contracts for 10 different weather stations (Atlanta, Chicago, Cincinnati, Dallas, Des Moines, Las Vegas, New York, Philadelphia, Portland, and Tucson).

Exchange-traded contracts are also popular. The New York Mercantile Exchange (NYMEX) and the International Petroleum Exchange (IPE) trade a number of oil futures and futures options contracts. Some of the futures contracts are settled in cash; others are settled by physical delivery. For example the Brent crude oil futures traded on the IPE has cash settlement based on the Brent index price; the light sweet crude oil futures traded on NYMEX requires physical delivery. In both cases the amount of oil underlying one contract is 1,000 barrels. NYMEX also trades popular contracts on two refined products: heating oil and gasoline. In both cases one contract is for the delivery of 42,000 gallons.

Natural Gas

The natural gas industry throughout the world has been going through a period of deregulation and the elimination of government monopolies. The supplier of natural gas is now not necessarily the same company as the producer of the gas. Suppliers are faced with the problem of meeting daily demand.

A typical over-the-counter contract is for the delivery of a specified amount of natural gas at a roughly uniform rate over a one-month period. Forward contracts, options, and swaps are available in the over-the-counter market. The seller of gas is usually responsible for moving the gas through pipelines to the specified location.

NYMEX trades a contract for the delivery of 10,000 million British thermal units of natural gas. The contract, if not closed out, requires physical delivery to be made during the delivery month at a roughly uniform rate to a particular hub in Louisiana. The IPE trades a similar contract in London.

Electricity

Electricity is an unusual commodity because it cannot easily be stored.[3] The maximum supply of electricity in a region at any moment is determined by the maximum capacity of all the electricity-producing plants in the region. In the United States there are 140 regions known as *control areas*. Demand and supply are first matched within a control area, and any excess power is sold to other control areas. It is this excess power that constitutes the wholesale market for electricity. The ability of one control area to sell power to another depends on the transmission capacity of the lines between the two areas. Transmission from one area to another involves a transmission cost, charged by the owner of the line, and there are generally some transmission or energy losses.

A major use of electricity is for air-conditioning systems. As a result the demand for electricity, and therefore its price, is much greater in the summer months than in the winter months. The nonstorability of electricity causes occasional very large movements in the spot price. Heat waves have been known to increase the spot price by as much as 1000% for short periods of time.

Like natural gas, electricity has been going through a period of deregulation and the elimination of government monopolies. This has been accompanied by the development of an electricity derivatives market. NYMEX now trades a futures contract on the price of electricity, and there is an active over-the-counter market in forward contracts,

[3] Electricity producers with spare capacity sometimes use it to pump water to the top of their hydroelectric plants so that it can be used to produce electricity at a later time. This is the closest they can get to storing this commodity.

options, and swaps. A typical contract (exchange traded or over the counter) allows one side to receive a specified number of megawatt-hours for a specified price at a specified location during a particular month. In a 5×8 contract, power is received for 5 days a week (Monday to Friday) during the off-peak period (11 p.m. to 7 a.m.) for the specified month. In a 5×16 contract, power is received 5 days a week during the on-peak period (7 a.m. to 11 p.m.) for the specified month. In a 7×24 contract, it is received around the clock every day during the month. Option contracts have either daily exercise or monthly exercise. In the case of daily exercise, the option holder can choose on each day of the month (by giving one day's notice) to receive the specified amount of power at the specified strike price. When there is monthly exercise, a single decision on whether to receive power for the whole month at the specified strike price is made at the beginning of the month.

An interesting contract in electricity and natural gas markets is what is known as a *swing option* or *take-and-pay option*. In this contract a minimum and maximum for the amount of power that must be purchased at a certain price by the option holder is specified for each day during a month and for the month in total. The option holder can change (or swing) the rate at which the power is purchased during the month, but usually there is a limit on the total number of changes that can be made.

Modeling Energy Prices

A plausible model for energy and other commodity prices should incorporate both mean reversion and volatility. One possible model is:

$$d \ln S = [\theta(t) - a \ln S]\, dt + \sigma\, dz \tag{25.1}$$

where S is the energy price, and a and σ are constant parameters. This is similar to the models for interest rates described in Chapter 30. The parameter σ is the volatility of S, and a measures the speed with which it reverts to a long-run average level. The $\theta(t)$ term captures seasonality and trends. Chapter 33 shows how to construct a trinomial tree for the model in equation (25.1) with $\theta(t)$ being estimated from futures prices. The parameters a and σ can be estimated from historical data or implied from derivative prices.

The parameters a and σ are different for different sources of energy. For crude oil, the reversion rate parameter a in equation (25.1) is about 0.5 and the volatility parameter σ is about 20%; for natural gas, a is about 1.0 and σ is about 40%; for electricity, a is typically between 10 and 20, while σ is 100 to 200%. The seasonality of electricity prices is also greater.[4]

How an Energy Producer Can Hedge Risks

There are two components to the risks facing an energy producer. One is the price risk; the other is the volume risk. Although prices do adjust to reflect volumes, there is a less-than-perfect relationship between the two, and energy producers have to take both into account when developing a hedging strategy. The price risk can be hedged using the energy derivative contracts discussed in this section. The volume risks can be hedged using the weather derivatives discussed in the previous section.

[4] For a fuller discussion of the spot price behavior of energy products, see D. Pilipovic, *Energy Risk*. New York: McGraw-Hill, 1997.

Define:

Y: Profit for a month

P: Average energy prices for the month

T: Relevant temperature variable (HDD or CDD) for the month

An energy producer can use historical data to obtain a best-fit linear regression relationship of the form

$$Y = a + bP + cT + \epsilon$$

where ϵ is the error term. The energy producer can then hedge risks for the month by taking a position of $-b$ in energy forwards or futures and a position of $-c$ in weather forwards or futures. The relationship can also be used to analyze the effectiveness of alternative option strategies.

25.4 INSURANCE DERIVATIVES

When derivative contracts are used for hedging purposes, they have many of the same characteristics as insurance contracts. Both types of contracts are designed to provide protection against adverse events. It is not surprising that many insurance companies have subsidiaries that trade derivatives and that many of the activities of insurance companies are becoming very similar to those of investment banks.

Traditionally the insurance industry has hedged its exposure to catastrophic (CAT) risks such as hurricanes and earthquakes using a practice known as reinsurance. Reinsurance contracts can take a number of forms. Suppose that an insurance company has an exposure of \$100 million to earthquakes in California and wants to limit this to \$30 million. One alternative is to enter into annual reinsurance contracts that cover on a pro rata basis 70% of its exposure. If California earthquake claims in a particular year total \$50 million, the costs to the company would then be only $0.3 \times \$50$, or \$15 million. Another more popular alternative, involving lower reinsurance premiums, is to buy a series of reinsurance contracts covering what are known as *excess cost layers*. The first layer might provide indemnification for losses between \$30 million and \$40 million; the next layer might cover losses between \$40 million and \$50 million; and so on. Each reinsurance contract is known as an *excess-of-loss* reinsurance contract. The reinsurer has written a bull spread on the total losses. It is long a call option with a strike price equal to the lower end of the layer and short a call option with a strike price equal to the upper end of the layer.[5]

The principal providers of CAT reinsurance have traditionally been reinsurance companies and Lloyds syndicates (which are unlimited liability syndicates of wealthy individuals). In recent years the industry has come to the conclusion that its reinsurance needs have outstripped what can be provided from these traditional sources. It has searched for new ways in which capital markets can provide reinsurance. One of the events that caused the industry to rethink its practices was Hurricane Andrew in 1992, which caused about \$15 billion of insurance costs in Florida. This exceeded the total of relevant insurance premiums received in Florida during the previous seven years. If

[5] Reinsurance is also sometimes offered in the form of a lump sum if a certain loss level is reached. The reinsurer is then writing a cash-or-nothing binary call option on the losses.

Hurricane Andrew had hit Miami, it is estimated that insured losses would have exceeded $40 billion. Hurricane Andrew and other catastrophes have led to increases in insurance/reinsurance premiums.

Exchange-traded insurance futures contracts have been developed by the CBOT, but have not been highly successful. The over-the-counter market has come up with a number of products that are alternatives to traditional reinsurance. The most popular is a CAT bond. This is a bond issued by a subsidiary of an insurance company that pays a higher-than-normal interest rate. In exchange for the extra interest the holder of the bond agrees to provide an excess-of-loss reinsurance contract. Depending on the terms of the CAT bond, the interest or principal (or both) can be used to meet claims. In the example considered above where an insurance company wants protection for California earthquake losses between $30 million and $40 million, the insurance company could issue CAT bonds with a total principal of $10 million. In the event that the insurance company's California earthquake losses exceeded $30 million, bond holders would lose some or all of their principal. As an alternative the insurance company could cover this excess cost layer by making a much bigger bond issue where only the bondholders' interest is at risk.

CAT bonds typically give a high probability of an above-normal rate of interest and a low-probability of a high loss. Why would investors be interested in such instruments? The answer is that there are no statistically significant correlations between CAT risks and market returns.[6] CAT bonds are therefore an attractive addition to an investor's portfolio. They have no systematic risk, so that their total risk can be completely diversified away in a large portfolio. If a CAT bond's expected return is greater than the risk-free interest rate (and typically it is), it has the potential to improve risk–return trade-offs.

SUMMARY

This chapter has shown that when there are risks to be managed, derivative markets have been very innovative in developing products to meet the needs of market participants.

In the weather derivatives market, two measures, HDD and CDD, have been developed to describe the temperature during a month. These are used to define the payoffs on both exchange-traded and over-the-counter derivatives. No doubt, as the weather derivatives market develops, we will see contracts on rainfall, snow, and similar variables become more commonplace.

In energy markets, oil derivatives have been important for some time and play a key role in helping oil producers and oil consumers manage their price risk. Natural gas and electricity derivatives are relatively new. They became important for risk management when these markets were deregulated and government monopolies discontinued.

Insurance derivatives are now beginning to be an alternative to traditional reinsurance as a way for insurance companies to manage the risks of a catastrophic event such as a hurricane or an earthquake. No doubt we will see other sorts of insurance (e.g., life insurance and automobile insurance) being securitized in a similar way as this market develops.

[6] See R. H. Litzenberger, D. R. Beaglehole, and C. E. Reynolds, "Assessing Catastrophe Reinsurance-Linked Securities as a New Asset Class," *Journal of Portfolio Management*, Winter (1996): 76–86.

Most weather, energy, and insurance derivatives have the property that percentage changes in the underlying variables have negligible correlations with market returns. This means that we can value derivatives by calculating expected payoffs using historical data and then discounting the expected payoffs at the risk-free rate.

FURTHER READING

On Weather Derivatives

Arditti, F., L. Cai, M. Cao, and R. McDonald, "Whether to Hedge," *Risk*, Supplement on Weather Risk (1999): 9–12.

Cao, M., and J. Wei, "Weather Derivatives Valuation and the Market Price of Weather Risk," *Journal of Futures Markets*, 24, 11 (November 2004): 1065–89.

Hunter, R., "Managing Mother Nature," *Derivatives Strategy*, February (1999).

On Energy Derivatives

Clewlow, L., and C. Strickland, *Energy Derivatives: Pricing and Risk Management*, Lacima Group, 2000.

Eydeland, A., and H. Geman, "Pricing Power Derivatives," *Risk*, October (1998): 71–73.

Joskow, P., "Electricity Sectors in Transition," *The Energy Journal*, 19 (1998): 25–52.

Kendall, R., "Crude Oil: Price Shocking," *Risk* Supplement on Commodity Risk, May (1999).

On Insurance Derivatives

Canter, M. S., J. B. Cole, and R. L. Sandor, "Insurance Derivatives: A New Asset Class for the Capital Markets and a New Hedging Tool for the Insurance Industry," *Journal of Applied Corporate Finance*, Autumn (1997): 69–83.

Froot, K. A., "The Market for Catastrophe Risk: A Clinical Examination," *Journal of Financial Economics*, 60 (2001): 529–71.

Froot, K. A., *The Financing of Catastrophe Risk*. University of Chicago Press, 1999.

Geman, H., "CAT Calls," *Risk*, September (1994): 86–89.

Hanley, M., "A Catastrophe Too Far," *Risk* Supplement on Insurance, July (1998).

Litzenberger, R. H., D. R. Beaglehole, and C. E. Reynolds, "Assessing Catastrophe Reinsurance-Linked Securities as a New Asset Class," *Journal of Portfolio Management*, Winter (1996): 76–86.

Questions and Problems (Answers in Solutions Manual)

25.1. What is meant by HDD and CDD?

25.2. How is a typical natural gas forward contract structured?

25.3. Distinguish between the historical data and the risk-neutral approach to valuing a derivative. Under what circumstance do they give the same answer.

25.4. Suppose that each day during July the minimum temperature is 68° Fahrenheit and the maximum temperature is 82° Fahrenheit. What is the payoff from a call option on the cumulative CDD during July with a strike of 250 and a payment rate of $5,000 per degree-day?

25.5. Why is the price of electricity more volatile than that of other energy sources?

25.6. Why is the historical data approach appropriate for pricing a weather derivatives contract and a CAT bond?

25.7. "HDD and CDD can be regarded as payoffs from options on temperature." Explain this statement.

25.8. Suppose that you have 50 years of temperature data at your disposal. Explain carefully the analyses you would carry out to value a forward contract on the cumulative CDD for a particular month.

25.9. Would you expect the volatility of the 1-year forward price of oil to be greater than or less than the volatility of the spot price? Explain your answer.

25.10. What are the characteristics of an energy source where the price has a very high volatility and a very high rate of mean reversion? Give an example of such an energy source.

25.11. How can an energy producer use derivatives markets to hedge risks?

25.12. Explain how a 5×8 option contract for May 2009 on electricity with daily exercise works. Explain how a 5×8 option contract for May 2009 on electricity with monthly exercise works. Which is worth more?

25.13. Explain how CAT bonds work.

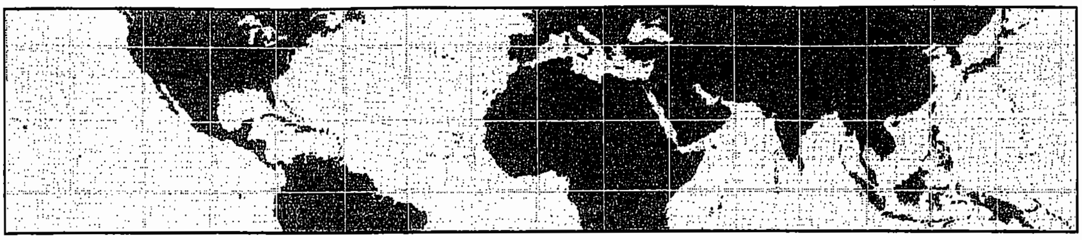
25.14. Consider two bonds that have the same coupon, time to maturity, and price. One is a B-rated corporate bond. The other is a CAT bond. An analysis based on historical data shows that the expected losses on the two bonds in each year of their life is the same. Which bond would you advise a portfolio manager to buy and why?

Assignment Question

25.15. An insurance company's losses of a particular type are to a reasonable approximation normally distributed with a mean of \$150 million and a standard deviation of \$50 million. (Assume no difference between losses in a risk-neutral world and losses in the real world.) The 1-year risk-free rate is 5%. Estimate the cost of the following:

(a) A contract that will pay in 1 year's time 60% of the insurance company's costs on a pro rata basis

(b) A contract that pays \$100 million in 1 year's time if losses exceed \$200 million



CHAPTER 26

More on Models and Numerical Procedures

Up to now the models we have used to value options have been based on the geometric Brownian motion model of asset price behavior that underlies the Black–Scholes formulas and the numerical procedures we have used have been relatively straightforward. In this chapter we introduce a number of new models and explain how the numerical procedures can be adapted to cope with particular situations.

Chapter 18 explained how traders overcome the weaknesses in the geometric Brownian motion model by using volatility surfaces. A volatility surface determines an appropriate volatility to substitute into Black–Scholes when pricing plain vanilla options. Unfortunately it says little about the volatility that should be used for exotic options when the pricing formulas of Chapter 24 are used. Suppose the volatility surface shows that the correct volatility to use when pricing a one-year plain vanilla option with a strike price of \$40 is 27%. This is liable to be totally inappropriate for pricing a barrier option (or some other exotic option) that has a strike price of \$40 and a life of one year.

The first part of this chapter discusses a number of alternatives to geometric Brownian motion that are designed to deal with the problem of pricing exotic options consistently with plain vanilla options. These alternative asset price processes fit the market prices of plain vanilla options better than geometric Brownian motion. As a result, we can have more confidence in using them to value exotic options.

The second part of the chapter extends the discussion of numerical procedures. It explains how convertible bonds and some types of path-dependent derivatives can be valued using trees. It discusses the special problems associated with valuing barrier options numerically and how these problems can be handled. Finally, it outlines alternative ways of constructing trees for two correlated variables and shows how Monte Carlo simulation can be used to value derivatives when there are early exercise opportunities.

As in earlier chapters, results are presented for derivatives dependent on an asset providing a yield at rate q. For an option on a stock index q should be set equal to the dividend yield on the index, for an option on a currency it should be set equal to the foreign risk-free rate, and for an option on a futures contract it should be set equal to the domestic risk-free rate.

26.1 ALTERNATIVES TO BLACK–SCHOLES

The Black–Scholes model assumes that an asset's price changes continously in a way that produces a lognormal distribution for the price at any future time. There are many alternative processes that can be assumed. One possibility is to retain the property that the asset price changes continuously, but assume a process other than geometric Brownian motion. Another alternative is to overlay continuous asset price changes with jumps. Yet another alternative is to assume a process where all the asset price changes that take place are jumps. We will consider examples of all three types of processes in this section. A model where stock prices change continuously is known as a *diffusion model.* A model where continuous changes are overlaid with jumps is known as a *mixed jump–diffusion model.* A model where all stock price changes are jumps is known as a *pure jump model.* These types of processes are known collectively as *Levy processes.*[1]

The Constant Elasticity of Variance Model

One alternative to Black–Scholes is the *constant elasticity of variance* (CEV) model. This is a diffusion model where the risk-neutral process for a stock price S is

$$dS = (r - q)S\,dt + \sigma S^{\alpha}\,dz$$

where r is the risk-free rate, q is the dividend yield, dz is a Wiener process, σ is a volatility parameter, and α is a positive constant.[2]

When $\alpha = 1$, the CEV model is the geometric Brownian motion model we have been using up to now. When $\alpha < 1$, the volatility increases as the stock price decreases. This creates a probability distribution similar to that observed for equities with a heavy left tail and less heavy right tail (see Figure 18.4).[3] When $\alpha > 1$, the volatility increases as the stock price increases. This creates a probability distribution with a heavy right tail and a less heavy left tail. This corresponds to a volatility smile where the implied volatility is an increasing function of the strike price. This type of volatility smile is sometimes observed for options on futures (see Assignment 16.23).

The valuation formulas for European call and put options under the CEV model are

$$c = S_0e^{-qT}[1 - \chi^2(a, b+2, c)] - Ke^{-rT}\chi^2(c, b, a)$$
$$p = Ke^{-rT}[1 - \chi^2(c, b, a)] - S_0e^{-qT}\chi^2(a, b+2, c)$$

when $0 < \alpha < 1$, and

$$c = S_0e^{-qT}[1 - \chi^2(c, -b, a)] - Ke^{-rT}\chi^2(a, 2-b, c)$$
$$p = Ke^{-rT}[1 - \chi^2(a, 2-b, c)] - S_0e^{-qT}\chi^2(c, -b, a)$$

---

[1] Roughly speaking, a Levy process is a continuous-time stochastic process with stationary independent increments.

[2] See J. C. Cox and S. A. Ross, "The Valuation of Options for Alternative Stochastic Processes," *Journal of Financial Economics*, 3 (March 1976): 145–66.

[3] The reason is as follows. As the stock price decreases, the volatility increases making even lower stock price more likely; when the stock price increases, the volatility decreases making higher stock prices less likely.

when $\alpha > 1$, with

$$a = \frac{[Ke^{-(r-q)T}]^{2(1-\alpha)}}{(1-\alpha)^2 v}, \qquad b = \frac{1}{1-\alpha}, \qquad c = \frac{S^{2(1-\alpha)}}{(1-\alpha)^2 v}$$

where

$$v = \frac{\sigma^2}{2(r-q)(\alpha-1)}[e^{2(r-q)(\alpha-1)T} - 1]$$

and $\chi^2(z, k, v)$ is the cumulative probability that a variable with a noncentral χ^2 distribution with noncentrality parameter v and k degrees of freedom is less than z. A procedure for computing $\chi^2(z, k, v)$ is provided in Technical Note 12 on the author's website.

The CEV model is particularly useful for valuing exotic equity options. The parameters of the model can be chosen to fit the prices of plain vanilla options as closely as possible by minimizing the sum of the squared differences between model prices and market prices.

Merton's Mixed Jump–Diffusion Model

Merton has suggested a model where jumps are combined with continuous changes.[4] Define:

λ: Average number of jumps per year

k: Average jump size measured as a percentage of the asset price

The percentage jump size is assumed to be drawn from a probability distribution in the model.

The probability of a jump in time Δt is $\lambda\,\Delta t$. The average growth rate in the asset price from the jumps is therefore λk. The risk-neutral process for the asset price is

$$\frac{dS}{S} = (r - q - \lambda k)\,dt + \sigma\,dz + dp$$

where dz is a Wiener process, dp is the Poisson process generating the jumps, and σ is the volatility of the geometric Brownian motion. The processes dz and dp are assumed to be independent.

An important particular case of Merton's model is where the logarithm of the size of the percentage jump is normal. Assume that the standard deviation of the normal distribution is s. Merton shows that a European option price can then be written

$$\sum_{n=0}^{\infty} \frac{e^{-\lambda' T}(\lambda' T)^n}{n!} f_n$$

where $\lambda' = \lambda(1+k)$. The variable f_n is the Black–Scholes option price when the dividend yield is q, the variance rate is

$$\sigma^2 + \frac{ns^2}{T}$$

[4] See R. C. Merton, "Option Pricing When Underlying Stock Returns Are Discontinuous," *Journal of Financial Economics*, 3 (March 1976): 125–44.

and the risk-free rate is

$$r - \lambda k + \frac{n\gamma}{T}$$

where $\gamma = \ln(1+k)$.

This model gives rise to heavier left and heavier right tails than Black–Scholes. It can be used for pricing currency options. As in the case of the CEV model, the model parameters are chosen by minimizing the sum of the squared differences between model prices and market prices.

The Variance-Gamma Model

An example of a pure jump model that is proving quite popular is the *variance-gamma model*.[5] Define a variable g as the change over time T in a variable that follows a gamma process with mean rate of 1 and variance rate of v. A gamma process is a pure jump process where small jumps occur very frequently and large jumps occur only occasionally. The probability density for g is

$$\phi(g) = \frac{g^{T/v-1} e^{-g/v}}{v^{T/v}\Gamma(T/v)}$$

where $\Gamma(\cdot)$ denotes the gamma function. This probability density can be computed in Excel using the GAMMADIST($\cdot, \cdot, \cdot, \cdot$) function. The first argument of the function is g, the second is T/v, the third is v, and the fourth is TRUE or FALSE, where TRUE returns the cumulative probability distribution function and FALSE returns the probability density function we have just given.

As usual, we define S_T as the asset price at time T, S_0 as the asset price today, r as the risk-free interest rate, and q as the dividend yield. In a risk-neutral world $\ln S_T$, under the variance-gamma model, has a probability distribution that, conditional on g, is normal. The conditional mean is

$$\ln S_0 + (r-q)T + \omega + \theta g$$

and the conditional standard deviation is

$$\sigma\sqrt{g}$$

where

$$\omega = \frac{T}{v}\ln(1 - \theta v - \sigma^2 v/2)$$

The variance-gamma model has three parameters: v, σ, and θ.[6] The parameter v is the variance rate of the gamma process, σ is the volatility, and θ is a parameter defining skewness. When $\theta = 0$, $\ln S_T$ is symmetric; when $\theta < 0$, it is negatively skewed (as for equities); and when $\theta > 0$, it is positively skewed.

Suppose that we are interested in using Excel to obtain 10,000 random samples of the change in an asset price between time 0 and time T using the variance-gamma model.

[5] See D.B. Madan, P.P. Carr, and E.C. Chang, "The Variance-Gamma Process and Option Pricing," *European Finance Review*, 2 (1998): 79–105.

[6] Note that all these parameters are liable to change when we move from the real world to the risk-neutral world. This is in contrast to pure diffusion models where the volatility remains the same.

As a preliminary, we could set cells E1, E2, E3, E4, E5, E6, and E7 equal to T, v, θ, σ, r, q, and S_0, respectively. We could also set E8 equal to ω by defining it as

= \$E\$1 * LN(1 − \$E\$3 * \$E\$2 − \$E\$4 * \$E\$4 * \$E\$2/2)/\$E\$2

We could then proceed as follows:

1. Sample values for g using the GAMMAINV function. Set the contents of cells A1, A2, ..., A10000 as

   = GAMMAINV(RAND(), \$E\$1/\$E\$2, \$E\$2)

2. For each value of g we sample a value z for a variable that is normally distributed with mean θg and standard deviation $\sigma\sqrt{g}$. This can be done by defining cell B1 as

   = A1 * \$E\$3 + SQRT(A1) * \$E\$4 * NORMSINV(RAND())

   and cells B2, B3, ..., B10000 similarly.

3. The stock price S_T is given by

$$S_T = S_0 \exp[(r - q)T + \omega + z]$$

   By defining C1 as

   = \$E\$7 * EXP((\$E\$5 − \$E\$6) * \$E\$1 + B1 + \$E\$8)

   and C2, C3, ..., C10000 similarly, random samples from the distribution of S_T are created in these cells.

Figure 26.1 shows the probability distribution that is obtained using the variance-gamma model for S_T when $S_0 = 100$, $T = 0.5$, $v = 0.5$, $\theta = 0.1$, $\sigma = 0.2$, and $r = q = 0$. For comparison it also shows the distribution given by geometric Brownian motion when the volatility, σ is 0.2 (or 20%). Although not clear in Figure 26.1, the variance-gamma distribution has heavier tails than the lognormal distribution given by geometric Brownian motion.

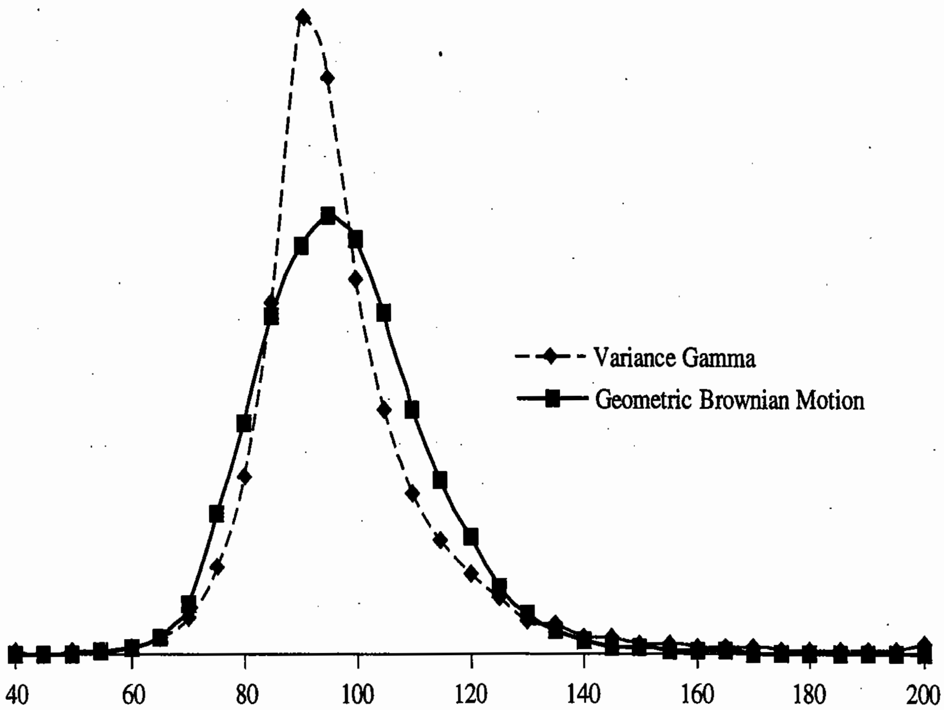
One way of characterizing the variance-gamma distribution is that g defines the rate at which information arrives during time T. If g is large, a great deal of information arrives and the sample we take from a normal distribution in step 2 above has a relatively large mean and variance. If g is small, relatively little information arrives and the sample we take has a relatively small mean and variance. The parameter T is the usual time measure, and g is sometimes referred to as measuring economic time or time adjusted for the flow of information.

Semi-analytic European option valuation formulas are provided by Madan *et al.* (1998). The variance-gamma model tends to produce a U-shaped volatility smile. The smile is not necessarily symmetrical. It is very pronounced for short maturities and "dies away" for long maturities. The model can be fitted to either equity or foreign currency plain vanilla option prices.

26.2 STOCHASTIC VOLATILITY MODELS

The Black–Scholes model assumes that volatility is constant. In practice, as discussed in Chapter 21, volatility varies through time. The variance-gamma model reflects this with

Figure 26.1 Distributions obtained with variance-gamma process and geometric Brownian motion.



its g parameter. Low values of g correspond to a low arrival rate for information and a low volatility; high values of g correspond to a high arrival rate for information and a high volatility.

An alternative to the variance-gamma model is a model where the process followed by the volatility variable is specified explicitly. Suppose first that the volatility parameter in the geometric Brownian motion is a known function of time. The risk-neutral process followed by the asset price is then

$$dS = (r - q)S\,dt + \sigma(t)S\,dz \tag{26.1}$$

The Black–Scholes formulas are then correct provided that the variance rate is set equal to the average variance rate during the life of the option (see Problem 26.6). The variance rate is the square of the volatility. Suppose that during a 1-year period the volatility of a stock will be 20% during the first 6 months and 30% during the second 6 months. The average variance rate is

$$0.5 \times 0.20^2 + 0.5 \times 0.30^2 = 0.065$$

It is correct to use Black–Scholes with a variance rate of 0.065. This corresponds to a volatility of $\sqrt{0.065} = 0.255$, or 25.5%.

Equation (26.1) assumes that the instantaneous volatility of an asset is perfectly predictable. In practice, volatility varies stochastically. This has led to the development of more complex models with two stochastic variables: the stock price and its volatility.

One model that has been used by researchers is

$$\frac{dS}{S} = (r - q)\,dt + \sqrt{V}\,dz_S \tag{26.2}$$

$$dV = a(V_L - V)\,dt + \xi V^{\alpha}\,dz_V \tag{26.3}$$

where a, V_L, ξ, and α are constants, and dz_S and dz_V are Wiener processes. The variable V in this model is the asset's variance rate. The variance rate has a drift that pulls it back to a level V_L at rate a.

Hull and White show that, when volatility is stochastic but uncorrelated with the asset price, the price of a European option is the Black–Scholes price integrated over the probability distribution of the average variance rate during the life of the option.[7] Thus a European call price is

$$\int_0^{\infty} c(\bar{V})g(\bar{V})\,d\bar{V}$$

where $\bar{V}$ is the average value of the variance rate, c is the Black–Scholes price expressed as a function of $\bar{V}$, and g is the probability density function of $\bar{V}$ in a risk-neutral world. This result can be used to show that Black–Scholes overprices options that are at the money or close to the money, and underprices options that are deep in or deep out of the money. The model is consistent with the pattern of implied volatilities observed for currency options (see Section 18.2).

The case where the asset price and volatility are correlated is more complicated. Option prices can be obtained using Monte Carlo simulation. In the particular case where $\alpha = 0.5$, Hull and White provide a series expansion and Heston provides an analytic result.[8] The pattern of implied volatilities obtained when the volatility is negatively correlated with the asset price is similar to that observed for equities (see Section 18.3).[9]

Chapter 21 discusses exponentially weighted moving average (EWMA) and GARCH(1,1) models. These are alternative approaches to characterizing a stochastic volatility model. Duan shows that it is possible to use GARCH(1,1) as the basis for an internally consistent option pricing model.[10] (See Problem 21.14 for the equivalence of GARCH(1,1) and stochastic volatility models.)

Stochastic volatility models can be fitted to the prices of plain vanilla options and then used to price exotic options.[11] For options that last less than a year, the impact of a stochastic volatility on pricing is fairly small in absolute terms (although in percentage

[7] See J.C. Hull and A. White, "The Pricing of Options on Assets with Stochastic Volatilities," *Journal of Finance*, 42 (June 1987): 281–300. This result is independent of the process followed by the variance rate.

[8] See J.C. Hull and A. White, "An Analysis of the Bias in Option Pricing Caused by a Stochastic Volatility," *Advances in Futures and Options Research*, 3 (1988): 27–61; S.L. Heston, "A Closed Form Solution for Options with Stochastic Volatility with Applications to Bonds and Currency Options," *Review of Financial Studies*, 6, 2 (1993): 327–43.

[9] The reason is given in footnote 3.

[10] See J.-C. Duan, "The GARCH Option Pricing Model," *Mathematical Finance*, vol. 5 (1995), 13–32; and J.-C. Duan, "Cracking the Smile" *RISK*, vol. 9 (December 1996), 55-59.

[11] For an example of this, see J. C. Hull and W. Suo, "A Methodology for the Assessment of Model Risk and its Application to the Implied Volatility Function Model," *Journal of Financial and Quantitative Analysis*, 37, 2 (June 2002): 297–318.

terms it can be quite large for deep-out-of-the-money options). It becomes progressively larger as the life of the option increases. The impact of a stochastic volatility on the performance of delta hedging is generally quite large. Traders recognize this and, as described in Chapter 17, monitor their exposure to volatility changes by calculating vega.

26.3 THE IVF MODEL

The parameters of the models we have discussed so far can be chosen so that they provide an approximate fit to the prices of plain vanilla options on any given day. Financial institutions sometimes want to go one stage further and use a model that provides an exact fit to the prices of these options.[12] In 1994 Derman and Kani, Dupire, and Rubinstein developed a model that is designed to do this. It has become known as the *implied volatility function* (IVF) model or the *implied tree* model.[13] It provides an exact fit to the European option prices observed on any given day, regardless of the shape of the volatility surface.

The risk-neutral process for the asset price in the model has the form

$$dS = [r(t) - q(t)]S\,dt + \sigma(S, t)S\,dz$$

where $r(t)$ is the instantaneous forward interest rate for a contract maturing at time t and $q(t)$ is the dividend yield as a function of time. The volatility $\sigma(S, t)$ is a function of both S and t and is chosen so that the model prices all European options consistently with the market. It is shown both by Dupire and by Andersen and Brotherton-Ratcliffe that $\sigma(S, t)$ can be calculated analytically:[14]

$$[\sigma(K, T)]^2 = 2\frac{\partial c_{\text{mkt}}/\partial T + q(T)c_{\text{mkt}} + K[r(T) - q(T)]\partial c_{\text{mkt}}/\partial K}{K^2(\partial^2 c_{\text{mkt}}/\partial K^2)} \tag{26.4}$$

where $c_{\text{mkt}}(K, T)$ is the market price of a European call option with strike price K and maturity T. If a sufficiently large number of European call prices are available in the market, this equation can be used to estimate the $\sigma(S, t)$ function.[15]

Andersen and Brotherton-Ratcliffe implement the model by using equation (26.4) together with the implicit finite difference method. An alternative approach, the *implied tree* methodology suggested by Derman and Kani and Rubinstein, involves constructing a tree for the asset price that is consistent with option prices in the market.

When it is used in practice the IVF model is recalibrated daily to the prices of plain vanilla options. It is a tool to price exotic options consistently with plain vanilla options. As discussed in Chapter 18 plain vanilla options define the risk-neutral

---

12 There is a practical reason for this. If the bank does not use a model with this property, there is a danger that traders working for the bank will spend their time arbitraging the bank's internal models.

13 See B. Dupire, "Pricing with a Smile," *Risk*, February (1994): 18–20; E. Derman and I. Kani, "Riding on a Smile," *Risk*, February (1994): 32–39; M. Rubinstein, "Implied Binomial Trees" *Journal of Finance*, 49, 3 (July 1994), 771–818.

14 See B. Dupire, "Pricing with a Smile," *Risk*, February (1994), 18–20; L. B. G. Andersen and R. Brotherton-Ratcliffe "The Equity Option Volatility Smile: An Implicit Finite Difference Approach," *Journal of Computation Finance* 1, No. 2 (Winter 1997/98): 5–37. Dupire considers the case where r and q are zero; Andersen and Brotherton-Ratcliffe consider the more general situation.

15 Some smoothing of the observed volatility surface is typically necessary.

probability distribution of the asset price at all future times. It follows that the IVF model gets the risk-neutral probability distribution of the asset price at all future times correct. This means that options providing payoffs at just one time (e.g., all-or-nothing and asset-or-nothing options) are priced correctly by the IVF model. However, the model does not necessarily get the joint distribution of the asset price at two or more times correct. This means that exotic options such as compound options and barrier options may be priced incorrectly.[16]

26.4 CONVERTIBLE BONDS

We now move on to discuss how the numerical procedures presented in Chapter 19 can be modified to handle particular valuation problems. We start by considering convertible bonds.

Convertible bonds are bonds issued by a company where the holder has the option to exchange the bonds for the company's stock at certain times in the future. The *conversion ratio* is the number of shares of stock obtained for one bond (this can be a function of time). The bonds are almost always callable (i.e., the issuer has the right to buy them back at certain times at a predetermined prices). The holder always has the right to convert the bond once it has been called. The call feature is therefore usually a way of forcing conversion earlier than the holder would otherwise choose. Sometimes the holder's call option is conditional on the price of the company's stock being above a certain level.

Credit risk plays an important role in the valuation of convertibles. If credit risk is ignored, poor prices are obtained because the coupons and principal payments on the bond are overvalued. Ingersoll provides a way of valuing convertibles using a model similar to Merton's (1974) model discussed in Section 22.6.[17] He assumes geometric Brownian motion for the issuer's total assets and models the company's equity, its convertible debt, and its other debt as claims contingent on the value of the assets. Credit risk is taken into account because the debt holders get repaid in full only if the value of the assets exceeds the amount owing to them.

A simpler model that is widely used in practice involves modeling the issuer's stock price. It is assumed that the stock follows geometric Brownian motion except that there is a probability $\lambda \, \Delta t$ that there will be a default in each short period of time Δt. In the event of a default the stock price falls to zero and there is a recovery on the bond. The variable λ is the risk-neutral default intensity defined in Section 22.2.

The stock price process can be represented by varying the usual binomial tree so that at each node there is:

1. A probability p_u of a percentage up movement of size u over the next time period of length Δt

---

[16] Hull and Suo test the IVF model by assuming that all derivative prices are determined by a stochastic volatility model. They found that the model works reasonably well for compound options, but sometimes gives serious errors for barrier options. See J. C. Hull and W. Suo, "A Methodology for the Assessment of Model Risk and its Application to the Implied Volatility Function Model," *Journal of Financial and Quantitative Analysis*, 37, 2 (June 2002): 297–318

[17] See J. E. Ingersoll, "A Contingent Claims Valuation of Convertible Securities," *Journal of Financial Economics*, 4, (May 1977), 289–322.

2. A probability p_d of a percentage down movement of size d over the next time period of length Δt
3. A probability $\lambda\,\Delta t$, or more accurately $1 - e^{-\lambda \Delta t}$, that there will be a default with the stock price moving to zero over the next time period of length Δt

Parameter values, chosen to match the first two moments of the stock price distribution, are:

$$p_u = \frac{a - de^{-\lambda \Delta t}}{u - d}, \qquad p_d = \frac{ue^{-\lambda \Delta t} - a}{u - d}, \qquad u = e^{\sqrt{(\sigma^2 - \lambda)\Delta t}}, \qquad d = \frac{1}{u}$$

where $a = e^{(r-q)\Delta t}$, r is the risk-free rate, and q is the dividend yield on the stock.

The life of the tree is set equal to the life of the convertible bond. The value of the convertible at the final nodes of the tree is calculated based on any conversion options that the holder has at that time. We then roll back through the tree. At nodes where the terms of the instrument allow conversion we test whether conversion is optimal. We also test whether the position of the issuer can be improved by calling the bonds. If so, we assume that the bonds are called and retest whether conversion is optimal. This is equivalent to setting the value at a node equal to

$$\max[\min(Q_1, Q_2), Q_3]$$

where Q_1 is the value given by the rollback (assuming that the bond is neither converted nor called at the node), Q_2 is the call price, and Q_3 is the value if conversion takes place.

Example 26.1

Consider a 9-month zero-coupon bond issued by company XYZ with a face value of \$100. Suppose that it can be exchanged for two shares of company XYZ's stock at any time during the 9 months. Assume also that it is callable for \$113 at any time. The initial stock price is \$50, its volatility is 30% per annum, and there are no dividends. The default intensity λ is 1% per year, and all risk-free rates for all maturities are 5%. Suppose that in the event of a default the bond is worth \$40 (i.e., the recovery rate, as it is usually defined, is 40%).

Figure 26.2 shows the stock price tree that can be used to value the convertible when there are three time steps ($\Delta t = 0.25$). The upper number at each node is the stock price; the lower number is the price of the convertible bond. The tree parameters are:

$$u = e^{\sqrt{(0.09-0.01)\times 0.25}} = 1.1519, \quad d = 1/u = 0.8681$$

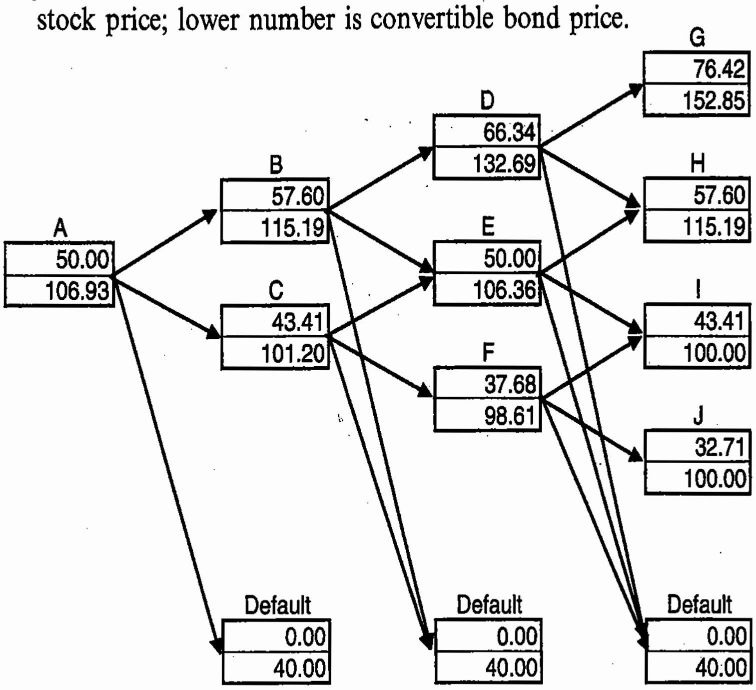
$$a = e^{0.05\times 0.25} = 1.0126, \quad p_u = 0.5167, \quad p_d = 0.4808$$

The probability of a default (i.e., of moving to the lowest nodes on the tree is $1 - e^{-0.01\times 0.25} = 0.002497$. At the three default nodes the stock price is zero and the bond price is 40.

Consider first the final nodes. At nodes G and H the bond should be converted and is worth twice the stock price. At nodes I and J the bond should not be converted and is worth 100.

Moving back through the tree enables the value to be calculated at earlier nodes. Consider, for example, node E. The value, if the bond is converted, is $2 \times 50 = \$100$. If it is not converted, then there is (a) a probability 0.5167 that it will move to node H, where the bond is worth 115.19, (b) a 0.4808 probability

Figure 26.2 Tree for valuing convertible. Upper number at each node is stock price; lower number is convertible bond price.



that it will move down to node I, where the bond is worth 100, and (c) a 0.002497 probability that it will default and be worth 40. The value of the bond if it is not converted is therefore

$$(0.5167 \times 115.19 + 0.4808 \times 100 + 0.002497 \times 40) \times e^{-0.05\times 0.25} = 106.36$$

This is more than the value of 100 that it would have if converted. We deduce that it is not worth converting the bond at node E. Finally, we note that the bond issuer would not call the bond at node E because this would be offering 113 for a bond worth 106.36.

As another example consider node B. The value of the bond if it is converted is $2 \times 57.596 = 115.19$. If it is not converted a similar calculation to that just given for node E gives its value as 118.31. The convertible bond holder will therefore choose not to convert. However, at this stage the bond issuer will call the bond for 113 and the bond holder will then decide that converting is better than being called. The value of the bond at node B is therefore 115.19. A similar argument is used to arrive at the value at node D. With no conversion the value is 132.79. However, the bond is called, forcing conversion and reducing the value at the node to 132.69.

The value of the convertible is its value at the initial node A, or 106.93.

When interest is paid on the debt, it must be taken into account. At each node, when valuing the bond on the assumption that it is not converted, the present value of any interest payable on the bond in the next time step should be included. The risk-neutral default intensity λ can be estimated from either bond prices or credit default swap spreads. In a more general implementation, λ, σ, and r are functions of time. This can be handled using a trinomial rather than a binomial tree (see Section 19.4).

One disadvantage of the model we have presented is that the probability of default is independent of the stock price. This has led some researchers to suggest an implicit finite difference method implementation of the model where the default intensity λ is a function of the stock price as well as time.[18]

26.5 PATH-DEPENDENT DERIVATIVES

A path-dependent derivative (or history-dependent derivative) is a derivative where the payoff depends on the path followed by the price of the underlying asset, not just its final value. Asian options and lookback options are examples of path-dependent derivatives. As explained in Chapter 24, the payoff from an Asian option depends on the average price of the underlying asset; the payoff from a lookback option depends on its maximum or minimum price. One approach to valuing path-dependent options when analytic results are not available is Monte Carlo simulation, as discussed in Chapter 19. A sample value of the derivative can be calculated by sampling a random path for the underlying asset in a risk-neutral world, calculating the payoff, and discounting the payoff at the risk-free interest rate. An estimate of the value of the derivative is found by obtaining many sample values of the derivative in this way and calculating their mean.

The main problem with Monte Carlo simulation is that the computation time necessary to achieve the required level of accuracy can be unacceptably high. Also, American-style path-dependent derivatives (i.e., path-dependent derivatives where one side has exercise opportunities or other decisions to make) cannot easily be handled. In this section, we show how the binomial tree methods presented in Chapter 19 can be extended to cope with some path-dependent derivatives.[19] The procedure can handle American-style path-dependent derivatives and is computationally more efficient than Monte Carlo simulation for European-style path-dependent derivatives.

For the procedure to work, two conditions must be satisfied:

1. The payoff from the derivative must depend on a single function, F, of the path followed by the underlying asset.
2. It must be possible to calculate the value of F at time $\tau + \Delta t$ from the value of F at time τ and the value of the underlying asset at time $\tau + \Delta t$.

Illustration Using Lookback Options

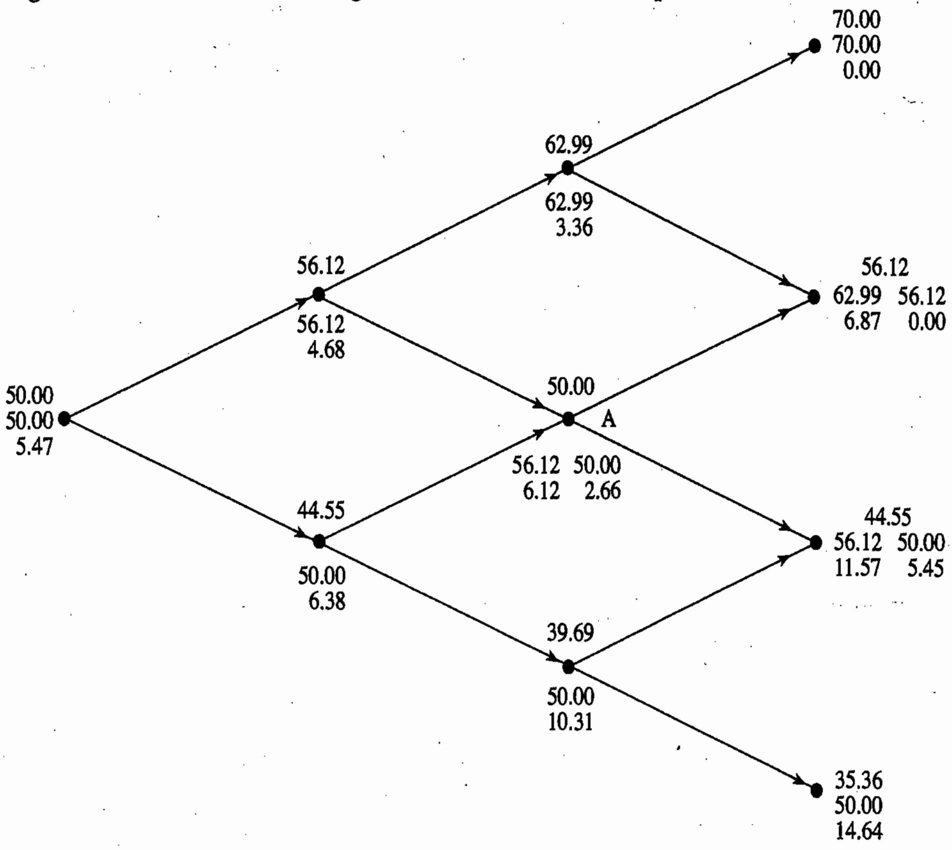
As a first illustration of the procedure, consider an American floating lookback put option on a non-dividend-paying stock.[20] If exercised at time τ, this pays off the amount by which the maximum stock price between time 0 and time τ exceeds the current stock

[18] See, e.g., L. Andersen and D. Buffum, "Calibration and Implementation of Convertible Bond Models," *Journal of Computational Finance*, 7, 1 (Winter 2003/04), 1–34. These authors suggest assuming that the default intensity is inversely proportional to S^{α}, where S is the stock price and α is a positive constant.

[19] This approach was suggested in J. Hull and A. White, "Efficient Procedures for Valuing European and American Path-Dependent Options," *Journal of Derivatives*, 1, 1 (Fall 1993): 21–31.

[20] This example is used as a first illustration of the general procedure for handling path dependence. For a more efficient approach to valuing American-style lookback options, see Technical Note 13 on the author's website.

Figure 26.3 Tree for valuing an American lookback option.



price. Suppose that the initial stock price is \$50, the stock price volatility is 40% per annum, the risk-free interest rate is 10% per annum, the total life of the option is three months, and that stock price movements are represented by a three-step binomial tree. With our usual notation this means that $S_0 = 50$, $\sigma = 0.4$, $r = 0.10$, $\Delta t = 0.08333$, $u = 1.1224$, $d = 0.8909$, $a = 1.0084$, and $p = 0.5073$.

The tree is shown in Figure 26.3. In this case, the path function F is the maximum stock price so far. The top number at each node is the stock price. The next level of numbers at each node shows the possible maximum stock prices achievable on paths leading to the node. The final level of numbers shows the values of the derivative corresponding to each of the possible maximum stock prices.

The values of the derivative at the final nodes of the tree are calculated as the maximum stock price minus the actual stock price. To illustrate the rollback procedure, suppose that we are at node A, where the stock price is \$50. The maximum stock price achieved thus far is either 56.12 or 50. Consider first the situation where it is equal to 50. If there is an up movement, the maximum stock price becomes 56.12 and the value of the derivative is zero. If there is a down movement, the maximum stock price stays at 50 and the value of the derivative is 5.45. Assuming no early exercise, the value of the derivative at A when the maximum achieved so far is 50 is, therefore,

$$(0 \times 0.5073 + 5.45 \times 0.4927)e^{-0.1 \times 0.08333} = 2.66$$

Clearly, it is not worth exercising at node A in these circumstances because the payoff from doing so is zero. A similar calculation for the situation where the maximum value at node A is 56.12 gives the value of the derivative at node A, without early exercise, to be

$$(0 \times 0.5073 + 11.57 \times 0.4927)e^{-0.1\times0.08333} = 5.65$$

In this case, early exercise gives a value of 6.12 and is the optimal strategy. Rolling back through the tree in the way we have indicated gives the value of the American lookback as \$5.47.

Generalization

The approach just described is computationally feasible when the number of alternative values of the path function, F, at each node does not grow too fast as the number of time steps is increased. The example we used, a lookback option, presents no problems because the number of alternative values for the maximum asset price at a node in a binomial tree with n time steps is never greater than n.

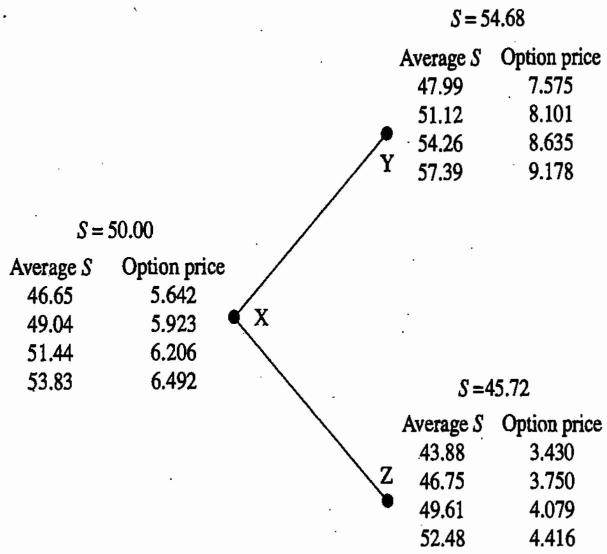
Luckily, the approach can be extended to cope with situations where there are a very large number of different possible values of the path function at each node. The basic idea is as follows. Calculations are carried out at each node for a small number of representative values of F. When the value of the derivative is required for other values of the path function, it is calculated from the known values using interpolation.

The first stage is to work forward through the tree establishing the maximum and minimum values of the path function at each node. Assuming the value of the path function at time $\tau + \Delta t$ depends only on the value of the path function at time τ and the value of the underlying variable at time $\tau + \Delta t$, the maximum and minimum values of the path function for the nodes at time $\tau + \Delta t$ can be calculated in a straightforward way from those for the nodes at time τ. The second stage is to choose representative values of the path function at each node. There are a number of approaches. A simple rule is to choose the representative values as the maximum value, the minimum value, and a number of other values that are equally spaced between the maximum and the minimum. As we roll back through the tree, we value the derivative for each of the representative values of the path function.

To illustrate the nature of the calculation, consider the problem of valuing the average price call option in Example 24.2 of Section 24.10 when the payoff depends on the arithmetic average stock price. The initial stock price is 50, the strike price is 50, the risk-free interest rate is 10%, the stock price volatility is 40%, and the time to maturity is 1 year. For 20 time steps, the binomial tree parameters are $\Delta t = 0.05$, $u = 1.0936$, $d = 0.9144$, $p = 0.5056$, and $1 - p = 0.4944$. The path function is the arithmetic average of the stock price.

Figure 26.4 shows the calculations that are carried out in one small part of the tree. Node X is the central node at time 0.2 year (at the end of the fourth time step). Nodes Y and Z are the two nodes at time 0.25 year that are reachable from node X. The stock price at node X is 50. Forward induction shows that the maximum average stock price that is achievable in reaching node X is 53.83. The minimum is 46.65. (The initial and final stock prices are included when calculating the average.) From node X, the tree branches to one of the two nodes Y and Z. At node Y, the stock price is 54.68 and the bounds for the average are 47.99 and 57.39. At node Z, the stock price is 45.72 and the bounds for the average stock price are 43.88 and 52.48.

Figure 26.4 Part of tree for valuing option on the arithmetic average.



Suppose that the representative values of the average are chosen to be four equally spaced values at each node. This means that, at node X, averages of 46.65, 49.04, 51.44, and 53.83 are considered. At node Y, the averages 47.99, 51.12, 54.26, and 57.39 are considered. At node Z, the averages 43.88, 46.75, 49.61, and 52.48 are considered. Assume that backward induction has already been used to calculate the value of the option for each of the alternative values of the average at nodes Y and Z. Values are shown in Figure 26.4 (e.g., at node Y when the average is 51.12, the value of the option is 8.101).

Consider the calculations at node X for the case where the average is 51.44. If the stock price moves up to node Y, the new average will be

$$\frac{5 \times 51.44 + 54.68}{6} = 51.98$$

The value of the derivative at node Y for this average can be found by interpolating between the values when the average is 51.12 and when it is 54.26. It is

$$\frac{(51.98 - 51.12) \times 8.635 + (54.26 - 51.98) \times 8.101}{54.26 - 51.12} = 8.247$$

Similarly, if the stock price moves down to node Z, the new average will be

$$\frac{5 \times 51.44 + 45.72}{6} = 50.49$$

and by interpolation the value of the derivative is 4.182.

The value of the derivative at node X when the average is 51.44 is, therefore,

$$(0.5056 \times 8.247 + 0.4944 \times 4.182)e^{-0.1 \times 0.05} = 6.206$$

The other values at node X are calculated similarly. Once the values at all nodes at time 0.2 year have been calculated, the nodes at time 0.15 year can be considered.

The value given by the full tree for the option at time zero is 7.17. As the number of time steps and the number of averages considered at each node is increased, the value of the option converges to the correct answer. With 60 time steps and 100 averages at each node, the value of the option is 5.58. The analytic approximation for the value of the option, as calculated in Example 24.2, is 5.62.

A key advantage of the method described here is that it can handle American options. The calculations are as we have described them except that we test for early exercise at each node for each of the alternative values of the path function at the node. (In practice, the early exercise decision is liable to depend on both the value of the path function and the value of the underlying asset.) Consider the American version of the average price call considered here. The value calculated using the 20-step tree and four averages at each node is 7.77; with 60 time steps and 100 averages, the value is 6.17.

The approach just described can be used in a wide range of different situations. The two conditions that must be satisfied were listed at the beginning of this section. Efficiency is improved somewhat if quadratic rather than linear interpolation is used at each node.

26.6 BARRIER OPTIONS

Chapter 24 presented analytic results for standard barrier options. This section considers numerical procedures that can be used for barrier options when there are no analytic results.

In principle, a barrier option can be valued using the binomial and trinomial trees discussed in Chapter 19. Consider an up-and-out option. A simple approach is to value this in the same way as a regular option except that, when a node above the barrier is encountered, the value of the option is set equal to zero.

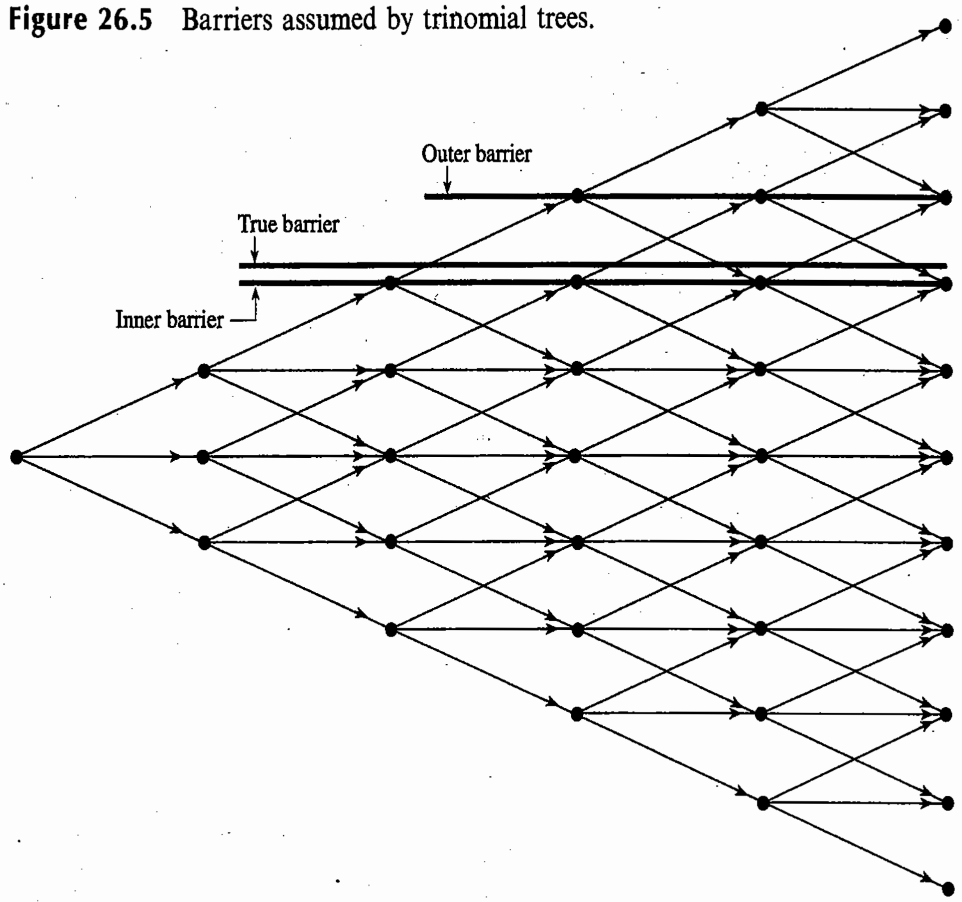
Trinomial trees work better than binomial trees, but even for them convergence is very slow when the simple approach is used. A large number of time steps are required to obtain a reasonably accurate result. The reason for this is that the barrier being assumed by the tree is different from the true barrier.[21] Define the *inner barrier* as the barrier formed by nodes just on the inside of the true barrier (i.e., closer to the center of the tree) and *the outer barrier* as the barrier formed by nodes just outside the true barrier (i.e., farther away from the center of the tree). Figure 26.5 shows the inner and outer barrier for a trinomial tree on the assumption that the true barrier is horizontal. The usual tree calculations implicitly assume that the outer barrier is the true barrier because the barrier conditions are first used at nodes on this barrier. When the time step is Δt, the vertical spacing between the nodes is of order $\sqrt{\Delta t}$. This means that errors created by the difference between the true barrier and the outer barrier also tend to be of order $\sqrt{\Delta t}$.

One approach to overcoming this problem is to:

1. Calculate the price of the derivative on the assumption that the inner barrier is the true barrier.

[21] For a discussion of this, see P.P. Boyle and S.H. Lau, "Bumping Up Against the Barrier with the Binomial Method," *Journal of Derivatives*, 1, 4 (Summer 1994): 6–14.

Figure 26.5 Barriers assumed by trinomial trees.



2. Calculate the value of the derivative on the assumption that the outer barrier is the true barrier.
3. Interpolate between the two prices.

Another approach is to ensure that nodes lie on the barrier. Suppose that the initial stock price is S_0 and that the barrier is at H. In a trinomial tree, there are three possible movements in the asset's price at each node: up by a proportional amount u; stay the same; and down by a proportional amount d, where $d = 1/u$. We can always choose u so that nodes lie on the barrier. The condition that must be satisfied by u is

$$H = S_0 u^N$$

or

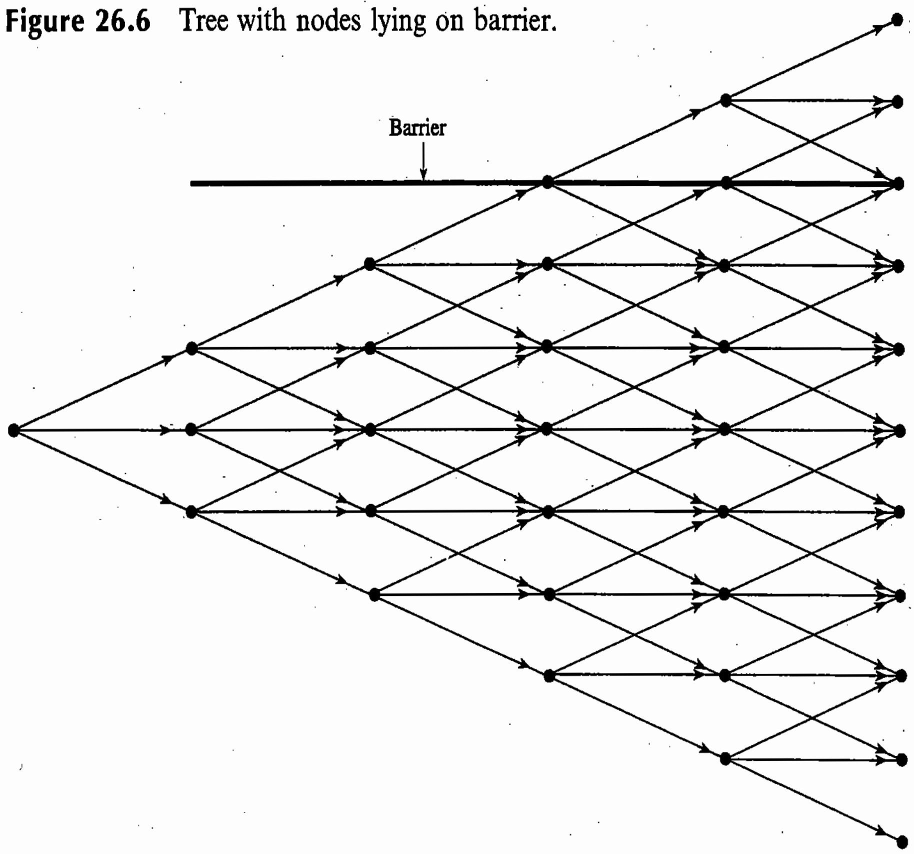
$$\ln H = \ln S_0 + N \ln u$$

for some positive or negative N.

When discussing trinomial trees in Section 19.4, the value suggested for u was $e^{\sigma\sqrt{3\Delta t}}$, so that $\ln u = \sigma\sqrt{3\Delta t}$. In the situation considered here, a good rule is to choose $\ln u$ as close as possible to this value, consistent with the condition given above. This means that

$$\ln u = \frac{\ln H - \ln S_0}{N}$$

Figure 26.6 Tree with nodes lying on barrier.



where

$$N = \text{int}\left[\frac{\ln H - \ln S_0}{\sigma\sqrt{3\Delta t}} + 0.5\right]$$

and int(x) is the integral part of x.

This leads to a tree of the form shown in Figure 26.6. The probabilities p_u, p_m, and p_d on the upper, middle, and lower branches of the tree are chosen to match the first two moments of the return, so that

$$p_d = -\frac{(r-q-\sigma^2/2)\Delta t}{2\ln u} + \frac{\sigma^2 \Delta t}{2(\ln u)^2}, \quad p_m = 1 - \frac{\sigma^2 \Delta t}{(\ln u)^2}, \quad p_u = \frac{(r-q-\sigma^2/2)\Delta t}{2\ln u} + \frac{\sigma^2 \Delta t}{2(\ln u)^2}$$

where p_u, p_m, and p_d are the probabilities on the upper, middle, and lower branches.

The Adaptive Mesh Model

The methods presented so far work reasonably well when the initial asset price is not close to the barrier. When the initial asset price is close to a barrier, the adaptive mesh model, which was introduced in Section 19.4, can be used.[22] The idea behind the model is that computational efficiency can be improved by grafting a fine tree onto a coarse

[22] See S. Figlewski and B. Gao, "The Adaptive Mesh Model: A New Approach to Efficient Option Pricing," *Journal of Financial Economics*, 53 (1999): 313–51.